



پردازش تصویر رقمی

رافائل سی. گونزالس
ریچارد ای. وودز



ترجمه:

دکتر مرتضی خادمی - مهندس داود جعفری

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِيْمِ



دانشگاه فردوسی مشهد

انتشارات دانشگاه فردوسی مشهد، شماره ۳۶۹

پردازش تصویر رقمه‌ی

تألیف:

رافائل سی. گونزالس

ریچارد ای. وودز

ترجمه:

دکتر هرتضی خادمی

عضو هیأت علمی دانشکده مهندسی دانشگاه فردوسی مشهد

داود جعفری

کارشناس ارشد مهندسی برق - مخابرات

<p>Gonzalez, Rafael C.</p> <p>پردازش تصویر رقemi / تأليف رافائل سى. گونزالس، ريجارد اي. وودز؛ ترجمه مرتضى خادمى، داود چعفرى. — مشهد: دانشگاه فردوسى مشهد، ۱۳۸۲.</p> <p>ص: مصور، جدول، نمودار. — (اتشارات دانشگاه فردوسى مشهد؛ شماره ۳۶۹، ۸۲۹، ISBN: 964-386-003-5)</p> <p>فهرست نويسى بر اساس اطلاعات فيها.</p> <p>Digital image processing.</p> <p>عنوان اصلى:</p> <p>كتابنامه: ص، [۸۰۹ - ۸۲۹]</p> <p>۱. عکس پردازی — روشهای رفعی. الف. وودز، ريجارد یوجین، ۱۹۵۴ — . م. Woods, Richard Engene Woods، Richard Engene، مرتضى، ۱۳۷۷ — ، مترجم. ج. چعفرى، داود، ۱۳۵۲ — ، مترجم. د. دانشگاه فردوسى مشهد. ه. عنوان.</p> <p>۶۲۱/۳۶۷</p> <p>TA ۱۶۳۲، ۶۹</p> <p>كتابخانه ملي اiran</p>
--



اتشارات دانشگاه فردوسى مشهد، شماره ۳۶۹

پردازش تصویر رقemi

تأليف

رافائل سى. گونزالس - ريجارد اي. وودز

مترجمان

مرتضى خادمى - داود چعفرى

ویراستار علمي

دکتر علیرضا سیدین

وزیری، ۸۳۲، ۱۳۸۵

Amir Fathi and Javad Gholami (Eds.)

بها: ۵۰۰۰۰ ریال

فهرست

۱۱	مقدمه مترجمان
۱۵	پیشگفتار
۱۹	فصل ۱ مقدمه
۱۹	۱. زمینه
۲۴	۲. نمایش تصویر رقمه
۲۵	۳. مراحل اساسی پردازش تصویر
۲۹	۴. عناصر سامانه های پردازش تصویر رقمه
۲۹	۴.۱. تصویربرداری
۳۴	۴.۲. ذخیره سازی
۳۶	۴.۳. پردازش
۳۷	۴.۴. مخابرات
۳۸	۴.۵. نمایش
۳۹	۵. سازمان کتاب
۴۰	مراجع
۴۲	فصل ۲ مبانی تصویررقمه
۴۲	۱. عناصر دریافت بصری
۴۴	۱.۱. ساختمان چشم انسان
۴۷	۱.۲. تشکیل تصویر در چشم
۴۸	۱.۳. تطابق و جداسازی روشنایی
۵۳	۲. یک مدل ساده تصویر
۵۵	۳. نمونه برداری و چندی کردن

پردازش تصویر رقمه ۴

۵۵	۲. ۳. ۱. نمونه برداری و چندی کردن یکنواخت
۶۴	۲. ۳. ۲. نمونه برداری و چندی کردن غیریکنواخت
۶۵	۴. ۲. چند رابطه پایه ای بین پیکسلها
۶۵	۴. ۲. ۱. همسایه های یک پیکسل
۶۶	۴. ۲. ۲. اتصال
۶۸	۴. ۲. ۳. بر حسب زنی اجزای متصل
۶۹	۴. ۲. ۴. رابطه ها، هم ارزی، و بسته بودن تراکنده ری
۷۲	۴. ۲. ۵. سنجه های فاصله
۷۴	۴. ۲. ۶. اعمال حسابی / منطقی
۷۹	۵. ۲. هندسه تصویربرداری
۸۰	۵. ۲. ۱. چند تبدیل پایه ای
۸۴	۵. ۲. ۲. تبدیل پرسپکتیو
۹۰	۵. ۲. ۳. مدل دوربین
۹۷	۵. ۲. ۴. تنظیم دوربین
۹۹	۵. ۲. ۵. تصویربرداری استریو
۱۰۲	۶. ۲. فیلم عکاسی
۱۰۲	۶. ۲. ۱. ساختمان و نور خوردنگی فیلم
۱۰۳	۶. ۲. ۲. مشخصات فیلم
۱۰۶	۶. ۲. ۳. تنظیم های دیافراگم و دریچه
۱۰۷	۷. ۲. نکات پایانی
۱۰۸	مراجع
۱۰۸	مسائل
۱۱۵	فصل ۳. تبدیل های تصویر
۱۱۵	۱. ۳. مقدمه ای بر تبدیل فوریه
۱۲۲	۲. ۳. تبدیل فوریه گستته
۱۲۶	۳. ۳. چند خاصیت از تبدیل فوریه دو بعدی
۱۲۸	۳. ۳. ۱. جدایزیری
۱۲۹	۳. ۳. ۲. انتقال
۱۳۰	۳. ۳. ۳. تناوب، و تقارن مزدوج
۱۳۲	۴. ۳. ۳. چرخش
۱۳۳	۳. ۳. ۵. توزیع پذیری و مقیاس بندی

۵ فهرست

۱۳۵	۶. ۳. ۳ مقدار متوسط
۱۳۵	۷. ۳. ۳ لابلسین
۱۳۵	۸. ۳. ۳ کانولوشن و همبستگی
۱۴۸	۹. ۳. ۳ نمونه برداری
۱۵۷	۴. ۳ تبدیل فوریه سریع
۱۵۸	۱۰. ۴. ۳ الگوریتم FFT
۱۶۱	۲. ۴. ۳ تعداد عملیات
۱۶۲	۳. ۴. ۳ معکوس FFT
۱۶۳	۴. ۴. ۳ پیاده سازی
۱۶۷	۵. ۳ سایر تبدیل های جدا پذیر تصویر
۱۷۰	۱. ۵. ۳ تبدیل والش
۱۷۶	۲. ۵. ۳ تبدیل هادامارد
۱۸۴	۳. ۵. ۳ تبدیل کسینوسی گسته
۱۸۶	۴. ۵. ۳ تبدیل هار آر
۱۸۸	۵. ۵. ۳ تبدیل اسلنت
۱۸۹	۶. ۳ تبدیل هتلینگ
۱۹۶	۷. ۳ نکات پایانی
۱۹۷	مراجع
۱۹۹	مسائل
۲۰۳	فصل ۴ ارتقای تصویر
۲۰۴	۱. ۴ زمینه
۲۰۴	۱. ۱. ۴ روش های میدان مکان
۲۰۶	۱. ۲. ۴ روش های میدان فرکانس
۲۰۹	۴. ۰. ۴ ارتقا با پردازش نقطه ای
۲۰۹	۱. ۰. ۲. ۴ چند تبدیل شدت ساده
۲۱۵	۲. ۰. ۲. ۴ پردازش بافت نگار
۲۲۱	۳. ۰. ۲. ۴ تفریق تصویر
۲۲۲	۴. ۰. ۲. ۴ متوسط گیری تصویر
۲۲۳	۴. ۰. ۳. ۴ فیلتر کردن مکانی
۲۲۳	۱. ۰. ۳. ۴ زمینه
۲۲۷	۲. ۰. ۳. ۴ فیلتر های آرام کننده

۶ پردازش تصویر رقمه

۲۴۰	۴. ۳. فیلتر های تیزکننده
۲۴۸	۴.۰.۴ ارتقاد در میدان فرکانس
۲۴۹	۴. ۱. فیلتر کردن پایین گذر
۲۵۷	۴. ۲. فیلتر کردن بالاگذر
۲۶۳	۴. ۳. فیلتر کردن هم شکلی
۲۶۶	۴. ۴. تولید نقاب های مکانی از مشخصات میدان فرکانس
۲۷۰	۴.۰.۴ پردازش تصویر رنگی
۲۷۱	۴. ۱. مبانی رنگ
۲۷۵	۴. ۲. مدل های رنگ
۲۸۹	۴. ۳. ۶. پردازش تصویر شبه رنگی
۲۹۸	۴. ۴. ۶. پردازش تصویر تمام رنگی
۳۰۲	۴. ۵. نکات پایانی
۳۰۲	۴. ۶. مراجع
۳۰۴	۴. ۷. مسائل

۳۰۹	۵. فصل بازیابی تصویر
۳۱۰	۵. ۱. مدل تخریب
۳۱۰	۵. ۱. ۱. چند تعریف
۳۱۱	۵. ۱. ۲. مدل تخریب برای توابع پیوسته
۳۱۳	۵. ۱. ۳. بیان گسته
۳۱۸	۵. ۲. قطربی سازی ماتریسهای چرخشی و بلوک-چرخشی
۳۱۹	۵. ۲. ۱. ماتریسهای چرخشی
۳۲۱	۵. ۲. ۲. ماتریسهای بلوک-چرخشی
۳۲۲	۵. ۲. ۳. اثرات قطربی سازی بر مدل تخریب
۳۲۷	۵. ۳. رهیافت جبری بازیابی
۳۲۷	۵. ۳. ۱. بازیابی نامقید
۳۲۸	۵. ۳. ۲. بازیابی مقید
۳۲۹	۵. ۴. فیلتر کردن معکوس
۳۲۹	۵. ۴. ۱. بیان روابط
۳۲۲	۵. ۴. ۲. حذف ماتری حاصل از حرکت خطی یکنواخت
۳۳۹	۵. ۵. فیلتر کمینه میانگین مربعات (وبن)
۳۴۲	۵. ۶. بازیابی کمینه مربعات مقید

۷ نهرست

۳۴۹	۷.۵ بازیابی تعاملی
۳۵۸	۸.۵ بازیابی در حوزه مکان
۳۵۹	۹.۵ تبدیل های هندسی
۳۶۰	۹.۵.۱ تبدیل های مکانی
۳۶۳	۹.۵.۲ درونیابی سطح خاکستری
۳۶۶	۱۰.۵ نکات پایانی
۳۶۷	مراجع
۳۶۸	مسائل
۳۷۳	فصل ۶ فشرده سازی تصویر
۳۷۵	۱.۶ مبانی
۳۷۶	۱.۶.۱ افزونگی رمزنگاری
۳۷۹	۱.۶.۲ افزونگی بین پیکسلی
۳۸۴	۱.۶.۳ افزونگی روان بصری
۳۸۷	۱.۶.۴ معیار مشابه
۳۸۹	۲.۶ مدل های فشرده سازی تصویر
۳۹۰	۲.۶.۱ رمزگذار و رمزبردار منبع
۳۹۲	۲.۶.۲ رمزگذار و رمزبردار کانال
۳۹۳	۳.۶ عناصر نظریه اطلاعات
۳۹۳	۳.۶.۱ اندازه گیری اطلاعات
۳۹۴	۳.۶.۲ کانتال اطلاعات
۴۰۰	۳.۶.۳ قضایای اساسی رمزنگاری
۴۱۲	۳.۶.۴ استفاده از نظریه اطلاعات
۴۱۵	۴.۶ فشرده سازی بدون خطای
۴۱۶	۴.۶.۱ رمزنگاری طول متغیر
۴۲۴	۴.۶.۲ رمزنگاری بیت-صفحه ای
۴۲۴	۴.۶.۳ رمزنگاری پیشگویی بی اتلاف
۴۲۸	۵.۶ فشرده سازی با اتلاف
۴۳۹	۵.۶.۱ رمزنگاری پیشگویی با اتلاف
۴۵۲	۵.۶.۲ رمزنگاری تبدیلی
۴۶۹	۶.۶ استانداردهای فشرده سازی تصویر
۴۶۹	۶.۶.۱ استانداردهای فشرده سازی تصویر دو سطحی (دودویی)

۴۷۵	۶.۲. استانداردهای فشرده سازی تصویر پیوسته
۴۸۶	۷. نکات پایانی
۴۸۷	مراجع
۴۸۹	مسائل
فصل ۷ بخش بندی تصویر	
۴۹۵	۱.۱. آشکارسازی ناپیوستگی ها
۴۹۶	۱.۱.۱. آشکارسازی نقطه
۴۹۷	۱.۱.۲. آشکارسازی خط
۴۹۷	۱.۱.۳. آشکارسازی لبه
۴۹۸	۱.۱.۴. آشکارسازی ترکیبی
۵۰۶	۲. پیوند لبه و آشکارسازی مرز
۵۱۳	۲-۱-۲-۷ پردازش محلی
۵۱۶	۲-۲-۷ پردازش سراسری با تبدیل هاف
۵۲۳	۳-۲-۷ پردازش سراسری با روشهای نظریه 'گراف
۳. آستانه گیری	
۵۲۹	۱.۳.۰.۷ مبانی
۵۳۰	۲.۳.۰.۷ نقش روشنایی
۵۳۳	۳.۳.۰.۷ آستانه گیری سراسری ساده
۵۳۴	۴.۳.۰.۷ آستانه گیری بهینه
۵۳۹	۵.۳.۰.۷ انتخاب آستانه براساس مشخصات مرز
۵۴۲	۶.۳.۰.۷ آستانه های مبتنی بر چند متغیر
۴. بخش بندی ناحیه گرا	
۵۴۵	۱.۰.۴.۰.۷ بیان پایه ای
۵۴۶	۲.۰.۴.۰.۷ رشد ناحیه با پیوستن پیکسل
۵۴۶	۳.۰.۴.۰.۷ تقسیم و ادغام ناحیه
۵۵۰	۵.۰.۰.۷ استفاده از حرکت در بخش بندی
۵۵۴	۱.۰.۵.۰.۷ روشهای مکانی
۵۶۱	۲.۰.۵.۰.۷ روشهای حوزه فرکانس
۶. نکات پایانی	
۵۶۶	مراجع
۵۶۷	مسائل
۵۶۹	

۹ نهرست

۵۷۵	فصل ۸ نمایش و توصیف
۵۷۵	۱۰.۸ روشاهای نمایش
۵۷۶	۱۱.۸ ۱. رمزهای زنجیره ای
۵۷۸	۱۲.۸ ۲. تقریبهای چند ضلعی
۵۸۰	۱۳.۸ ۳. نمایه ها
۵۸۳	۱۴.۸ ۴. قطعات مرزی
۵۸۴	۱۵.۸ ۵. اسکلت یک ناحیه
۵۸۸	۱۶.۸ توصیف گرهای مرز
۵۸۸	۱۷.۸ ۱. چند توصیفگر ساده
۵۹۰	۱۸.۸ ۲. اعداد شکلی
۵۹۳	۱۹.۸ ۳. توصیفگرهای فوریه
۵۹۷	۲۰.۸ ۴. گشتاورها
۵۹۸	۲۱.۸ توصیف گرهای ناحیه ای
۵۹۸	۲۲.۸ ۱. چند توصیفگر ساده
۵۹۹	۲۳.۸ ۲. توصیفگرهای توپولوژیکی
۶۰۲	۲۴.۸ ۳. بافت
۶۱۱	۲۵.۸ ۴. گشتاورها
۶۱۵	۲۶.۸ شکل شناسی
۶۱۶	۲۷.۸ ۱. گسترش و سایش
۶۲۲	۲۸.۸ ۲. باز کردن و بستن
۶۲۷	۲۹.۸ ۳. تبدیل HOM
۶۲۹	۳۰.۸ ۴. چند الگوریتم پایه ای شکل شناسی
۶۵۰	۳۱.۸ ۵. تعمیم به تصاویر خاکستری
۶۶۴	۳۲.۸ ۵. توصیف گرهای رابطه ای
۶۶۸	۳۳.۸ ۶. نکات پایانی
۶۶۹	۳۴.۸ مراجع
۶۷۱	۳۵.۸ مسائل
۶۷۷	فصل ۹ تشخیص و تعبیر
۶۷۸	۱۰.۹ ۱. عناصر تحلیل تصویر
۶۸۰	۱۱.۹ ۲. الگوهای دسته های الگو
۶۸۶	۱۲.۹ ۳. روش های نظریه تصمیم

۱۰ پردازش تصویر رقمنی

۶۸۷	۱.۳.۹ تطبيق
۶۹۴	۲.۳.۹ دسته بندی کننده های آماری بهینه
۷۰۵	۳.۳.۹ شبکه های عصبی
۷۳۴	۴.۹ روش های ساختاری
۷۳۴	۱.۴.۹ اعداد شکلی تطبیقی
۷۳۶	۲.۴.۹ تطبيق رشته
۷۳۷	۳.۴.۹ روش های قالبی
۷۵۸	۵.۹ تعبیر
۷۵۸	۱.۵.۹ زمینه
۷۵۹	۲.۵.۹ انواع دانش
۷۶۰	۳.۵.۹ سامانه های منطقی (حساب گزاره ای)
۷۷۲	۴.۵.۹ شبکه های معنایی
۷۷۴	۵.۵.۹ سامانه های فرآوری (خبره)
۷۸۰	۶.۹ نکات پایانی
۷۸۰	مراجع
۷۸۱	مسائل
۷۸۹	A ضمیمه تولید تصاویر نیم تن
۷۹۵	B ضمیمه تصاویر رمزشده
۸۰۹	کتاب شناسی

مقدمه مترجمان

پردازش تصویر رقمی دانش جدیدی است که سابقه آن به پس از اختراع رایانه های رقمی باز می گردد. با این حال این علم نوپا در چند دهه اخیر از هر دو جنبه نظری و عملی پیشرفت های چشمگیری داشته است. سرعت این پیشرفت به اندازه ای بوده است که هم اکنون و پس از این مدت نسبتاً کوتاه، به راحتی می توان ردد پایی پردازش تصویر رقمی را در بسیاری از علوم و صنایع مشاهده نمود. بعضی از این کاربردها آنچنان به پردازش تصویر وابسته هستند که بدون آن، اساساً قابل استفاده نمی باشند. یک نمونه از چنین کاربردهای تشخیص تومور با استفاده از تصاویر CT-Scan (تومورنگاری رایانه ای) است. اگرچه ذکر تمامی کاربردهای پردازش تصویر رقمی در حوصله این نوشتار نمی باشد، اما معرفی چند دسته مهم از کاربردهای پردازش تصویر به شرح زیر مفید به نظر می رسد:

الف) کاربردهای عکاسی مانند ارتقا، بازسازی تصاویر قدیمی، بازسازی تصاویر خراب شده با نویز و بهبود ظاهر تصاویر معمولی

ب) کاربردهای پزشکی مانند ارتقای ویژگیهای تصاویر اشعه ایکس، تولید تصاویر MRI و CT-Scan

ج) کاربردهای امنیتی مانند تشخیص حرکت (در دزدگیرها)، تشخیص اثر انگشت، تشخیص چهره و تشخیص امضاء

د) کاربردهای نظامی مانند تشخیص و رهگیری خودکار اهداف متحرک یا ثابت از هوا یا از زمین

ه) کاربردهای سنجش از راه دور مانند ارتقا و تحلیل تصاویر هوایی و ماهواره ای (برداشته شده از مناطق مختلف جغرافیایی) که در کاربردهای نقشه برداری، کشاورزی، هواشناسی و موارد دیگر مفید هستند.

و) کاربردهای صنعتی مرتبط با خودکارسازی صنایع مانند تفکیک محصولات مختلف براساس شکل یا اندازه، آشکارسازی نواقص و شکستگی های موجود در محصولات، تعیین محل اشیاء و اجرای فرآیند تولید با استفاده از روباتها و بینایی ماشینی. علاوه بر آن، بسیاری از فعالیتهای صنعتی و نیز آزمایش های شبیه سازی و اتمی که برای انسان مضر و در مواردی غیر ممکن است، رامی توان با استفاده از دانش بینایی ماشینی که مبتنی بر پردازش تصویر است، انجام داد.

ز) کاربردهای فشرده سازی تصویر مانند ذخیره سازی، ارسال تصاویر تلویزیون با کیفیت بالا و ارسال تصاویر متحرک و زنده از روی شبکه اینترنت و یا خط تلفن.

ح) موارد متفرقه دیگری نیز مانند تصویربرداری از اسناد و ارسال آنها توسط دورنگار و تشخیص خودکار نویسه در ردیف کاربردهای پردازش تصویر قرار دارند.

و سعت کاربردهای پردازش تصویر در طی سالیان اخیر، ضرورت فراگیری جامع چنین مبحثی را برای دانشجویان رشته های برق و رایانه و حتی سایر رشته های مهندسی چون مکانیک و مواد، به طور جدی مطرح نموده است. به همین دلیل هم اکنون تقریباً در اغلب دانشگاه های معترض جهان و دانشگاه های بزرگ کشور عزیزان، این مبحث به عنوان یک درس اصلی در مقطع تحصیلات تکمیلی و نیز یک درس اختیاری در مقطع کارشناسی تدریس می گردد. تحت این شرایط، لزوم انتخاب یک کتاب جامع با خصوصیات آموزشی مناسب برای تدریس این درس بر کسی پوشیده نیست. خوشبختانه کتابی که هم اکنون ترجمه آن را پیش رو دارد، در حد بسیار خوبی دارای ویژگی های یک کتاب درسی بوده و به همین دلیل به عنوان مطرح ترین و پرمصرف ترین کتاب درسی این مبحث در اغلب دانشگاه های جهان و ایران تلقی می گردد.

مؤلفین این کتاب آقایان Gonzalez و Woods از مدرسان دانشگاه Tennessee و مهندسان شرکت Perceptics هستند که این کتاب را براساس ویرایش های سال های ۱۹۷۷ و ۱۹۸۷ کتاب «پردازش تصویر رقمی» تألیف Gonzalez و Wintz (مؤلف اول در هر سه ویرایش مشترک بوده ولی مؤلف دوم و ناشر متفاوت هستند) تألیف کرده اند. مؤلفان کتاب سالها سابقه کار عملی در زمینه تجاری سازی پردازش تصویر رقمی در شرکت Perceptics و نیز سابقه تدریس طولانی در دانشگاه Tennessee داشته اند. این دو عامل باعث شده که مباحث مطرح شده علاوه بر مستدل بودن و داشتن روح علمی و دانشگاهی، حاوی مثالهای گویا و جالب و نیز تصاویر کاربردی فراوان باشند. همچنین فصول و مباحث کتاب به ترتیبی سازماندهی شده اند که از ابتدای آنها تمام مراحل و لوازم یک سامانه پردازش تصویر کامل را پوشش می دهند. وجود

خصوصیات فوق، انتخاب این کتاب را به عنوان مناسب ترین کتاب تدریس پردازش تصویر رقمنی توجیه می نماید. این کتاب حاوی ۹ فصل و ۲ ضمیمه است که خواننده محترم برای آشنایی با ساختار مطالب مطرح شده می تواند به بخش ۱ . ۵ کتاب مراجعه نماید.

گمان ما بر این است که این ترجمه می تواند به عنوان یک کتاب درسی جامع برای دانشگاهها و نیز مرجع پایه برای مطالعه مهندسین فعال در صنایع مرتبط مورد استفاده قرار گیرد. مهندسی که می خواهد فارغ از مشکلات زبان دوم و در مدت زمان کوتاه مطالب مقدماتی و پایه ای دانش پردازش تصویر را بیاموزند. همان طوری که مؤلفان نیز خاطرنشان ساخته اند، دانش لازم برای فهم مطالب این کتاب، آشنایی با تحلیل ریاضی، نظریه ماتریس، احتمالات و برنامه نویسی را يانه است.

گرایش مترجمان در این ترجمه اصالت دادن به حفظ امانت و وفاداری به متن اصلی بوده است. به موازات تلاش گردیده از کلمات متداول و جملات روان استفاده گردد. در خصوص کلمات و اصطلاحات خاص، سعی کرده ایم آنها را به معادل متداول یا نزدیکترین کلمه در زبان فارسی ترجمه نشود و البته معادل انگلیسی آنها را نیز در پاورقی ذکر نماییم. در خصوص عبارات و جملات، با این که سعی در حفظ سیاق آنها داشته ایم، اما تأکید بیشتر بر حفظ مفهوم و رسایی آنها بوده است.

در پایان صادقانه اذعان می نماییم علی رغم سعی فراوان برای حداقل شدن اشکالات، با توجه به این واقعیت که این ترجمه، اولین کتاب فارسی در زمینه پردازش تصویر رقمنی می باشد، هنوز موارد زیادی برای بهبود در آن یافت خواهد شد. از خوانندهان محترم تقاضا داریم در صورت مشاهده هرگونه اشکال، با یادآوری آن، مارا سپاسگزار خود نمایند.

پیشگفتار

موضوع «پردازش تصویر واقعی» به طور دائم در حال توسعه است. در چند سال گذشته میزان علاقه به «شکل شناسی تصویر^۱»، «شبکه های عصبی»، «پردازش تصویر تمام رنگی^۲»، «فسرده سازی داده های تصویر^۳»، «تشخیص تصویر^۴» و «سامانه های تحلیل تصویر مبتنی بر دانش^۵» افزایش زیادی یافته است. موضوعات فوق، محور اصلی فعالیت هایی هستند که این متن حاصل آنهاست. این کتاب براساس ویرایش های سالهای ۱۹۷۷ و ۱۹۸۷ کتاب پردازش تصویر رقemi^۶ تأثیف Gonzalez و Wintz و نیز حاصل دستاوردها و تجارب موفق یک دهه تلاش این نویسندها در تجاری سازی پردازش تصویر رقemi در شرکت Perceptics است. ارائه مطالب به شیوه دانشگاهی متأثر از تدریس و فعالیتهای تحقیقاتی نویسندها کتاب در دانشگاه Tennessee است.

در این کتاب توضیحات جدید و جامعی درباره شکل شناسی تصاویر دودویی و خاکستری، شبکه های عصبی، و پردازش تصویر تمام رنگی ارائه می شود. در این ویرایش فشرده سازی داده های تصویر به طور کامل با شیوه ای جدید و شامل استانداردهای فشرده سازی تصویر بازنویسی شد. یک فصل جدید نیز که به تشخیص تصویر و تعبیر مبتنی بر دانش آن می پردازد، به کتاب افزوده شد. سایر بخش های متن نیز از تغییرات اساسی در راستای گسترش چند مبحث معمولی و نیز بازنگری مثالها با استفاده از تصاویر جدید دارای تفکیک^۷ بالاتر، برخوردار بوده اند. این ویرایش، دارای ۱۵۱ بخش جدید، بیش از ۲۵۰ تصویر و

1) *Image morphology* 2) *Full-Color Image processing* 3) *Image data compression* 4) *Image recognition*

5) *Knowledge-based image analysis systems*

6) *Digital image processing*

7) *Resolution*

طرح جدید، ۹۵ مسأله و ۱۵۲ مرجع جدید است. یک راهنمای تدریس که شامل حل دقیق تمام مسائل، راهنمایی هایی برای مباحث درسی و پیشنهاد پروژه های متعدد می باشد، نیز در نزد ناشر موجود است. همانند ویرایش های گذشته، این ویرایش نیز با کمک دانشجویان و همراهی مدرسان این درس تهیه شد. بنابراین هنوز هم اهداف اصلی کتاب، معرفی مفاهیم و روش های اصلی پردازش تصویر رقمه و ایجاد بینایی که بتوان از آن به عنوان مبنای مطالعه و تحقیق بیشتر در این زمینه استفاده کرد، است. برای نیل به این اهداف، توجه خود را بر موضوعی که به نظر ما پایه ای است و کاربرد آن به مسائل خاصی محدود نمی شود، متمرکز می کنیم. علاوه بر این، پیچیدگی ریاضی کتاب، در حد فهم دانشجویان کارشناسی سال آخر دوره کارشناسی و نیز دانشجویان تحصیلات تکمیلی رشته های مهندسی و رایانه، که آمادگی مقدماتی در تحلیل ریاضی، نظریه ماتریس، احتمالات و برنامه نویسی رایانه دارند، ارائه شده است.

خلاصه مطالب جدید عبارت است از: شرح فصل یک که درباره عناصر سامانه های پردازش تصویر بود، برای همگامی با پیشرفت های جدید در سخت افزار و نرم افزار پردازش تصویر بهنگام شد. یک بخش جدید که به استخراج اجزای متصل می پردازد، به فصل ۲ افزوده شد. موضوع اجزای متصل، در پردازش تصویر دودویی برای خودکارسازی^۱، تصویربرداری اسناد، و تشخیص نویسه، بسیار مورد توجه است. فصل ۳ به منظور ارائه تبدیلات تصویر بیشتر نسبت به گذشته، شامل تبدیلات اسلت^۲ و ها آر^۳، گسترش یافت. شرح تبدیل کسینوسی گستره^۴ نیز گستردۀ تر شد. در چند سال اخیر، استفاده از این تبدیل ویژه در استانداردهای جهانی فشرده سازی تصویر کاملاً متداول شده است. در فصل ۴ نیز تا حد زیادی تغییر ساختاری بوجود آمد. چند روش ارتقا شامل پردازش های نقطه ای، نظری نگاشت های محدوده خاکستری برای تصویرهای منفی^۵، کشش تمايز^۶، فشرده سازی محدوده تغییر^۷، بخش بندی محدوده خاکستری^۸ و بخش بندی بیت- صفحه ای^۹ به آن افزوده شدند. شرحی درباره تفریق تصویر و نتایج جدیدی درباره متوسط گیری تصویر^{۱۰} نیز افزوده شدند. ساختار جدید فصل ۴ به طور جداگانه به روش های مکانی و تبدیلی می پردازد. شرحی جامع درباره مبانی فیلتر های^{۱۱} مکانی، شامل انواع مهم فیلتر های آرام ساز^{۱۲} و تیز کننده^{۱۳}، به این فصل افزوده شد. با این افزایش، بحث ارتقا در میدان فرکانس با تصاویر جدیدی که

-
- 1) Automation 2) Slant 3) Haar 4) Discrete cosine transform 5) Image negatives
 6) Contrast Stretching 7) Dynamic range compression 8) Gray-level slicing 9) Bit-plane slicing
 10) Image averaging 11) Filters 12) Smoothing 13) Sharpening

کیفیت و تفکیک بالاتری دارند، کاملتر شد. همچنین چند بخش جدید که به مدل های رنگی و پردازش تصویر تمام رنگی می پردازند، به فصل ۴ افزوده شد. در فصل ۵، فقط بعضی از مراجع تغییر کرد و به تعداد تمرین ها نیز افزوده شد؛ بدنه اصلی این فصل، «شیوه واحدی برای بازیابی تصویر با روش های جبری می باشد. فصل ۶ کاملاً بازنویسی شد. مباحث این فصل شامل مروری بر نظریه اساسی اطلاعات، شرحی از روش های فشرده سازی دودویی و پیوسته (تک رنگ و تمام رنگ)، روش های فشرده سازی بی خطاب و با اتلاف، و استانداردهای فشرده سازی تصویر می باشد. گرچه موضوع فشرده سازی تصویر، سالهاست که در کتابهای پردازش تصویر رقیع مطرح است، به تازگی، به صورت یک موضوع تجاری فعال و جذاب درآمده است. نیاز به استفاده از دستگاه دورنگار و تصویربرداری استاد، از عواملی هستند که این علاقه را تشیدید می کنند. پذیرش عملی این فن آوری بوسیله استانداردها تقویت گردید. در فصل ۷ بعضی مطالب بازنگری شد؛ «تبديل هاف^۱» که قبلاً در فصل ۳ بیان شده بود، به فصل ۷ که در آن مبحث کاربرد پیشتری دارد، انتقال یافت و مراجع و مسائل جدیدی نیز به آن اضافه شد. فصل ۸ گسترش یافت تا شرح بهتری از توصیف گرهای فوریه را شامل شود. یک بخش جدید و جامع درباره 'شکل شناسی دودویی و خاکستری نیز به آن افزوده شد. با این که بیش از یک دهه است که شکل شناسی یک موضوع تحقیقاتی فعال در اروپا می باشد، فقط چند سال است که در آمریکا مورد توجه واقع گردیده است. این مبحث با چند مثال تصویری -که برای درک بهتر مطالب افزوده شد- ارائه گردیده است. فصل ۹ جدید است. این فصل به تشخیص و تعبیر تصویر، که فرآیندهای اساسی در تحلیل تصویر هستند، می پردازد. در این فصل با شرح روش های کلاسیک «حداقل فاصله»، «همبستگی»^۲، «تشخیص بیز»^۳ و نیز شرح «پرسپترونها»^۴ و «شبکه های عصبی»، بخش مهمی از روش های تشخیص را راهه کرده ایم. اخیراً به کاربرد شبکه های عصبی در پردازش تصویر بسیار توجه می شود. در این فصل روش های تشخیص ساختاری را نیز مطرح می کنیم و آن را با شرحی درباره 'تغییر تصویر مبتنی بر دانش شامل «حساب گزاره ای»^۵، «شبکه های معنایی»^۶ و «سامانه های خبره»^۷ پایان می دهیم.

طبق تجربه ما یکی از ویژگی های اصلی درس «پردازش تصویر» که دانشجویان را به آن علاقه مند می کند، امکان پیاده سازی و آزمودن مفاهیم و الگوریتم های آن با داده های واقعی است. محیط انجام این

1) Hough transform

2) Correlation

3) Bayes recognition

4) Perceptron

5) Predicate calculus

6) Semantic network

7) Expert systems

کار، یک سامانه پردازش تصویر است که دارای رقمه ساز تصویر، رایانه همه منظوره و وسیله ای برای نمایش تصویر باشد. وقتی چنین تجهیزاتی در دسترس نباشد، ضمایم این کتاب راه جایگزینی برای بهره برداری خوانندگان فراهم می آورند. «ضمیمه الف» در مورد روش های ایجاد تصویر خاکستری روی وسایل خروجی متدال نظیر چاپگرها و نمایشگرهای دودوبی است. این روش ها در اغلب زبانهای برنامه نویسی رایانه به آسانی قابل اجرا هستند. «ضمیمه ب» حاوی مجموعه ای از تصاویر رمزگذاری شده است که برای آزمودن روش های مشروح در متن مناسب هستند. با بهره گیری از مطالب این دو ضمیمه و یک رایانه معمولی و پیاده سازی الگوریتم های پردازش تصویر و در نهایت نمایش تصویری نتایج، به روشی پایه ای برای بدست آوردن تجربه بیشتر دست یافته ایم.

ما وام دار بعضی صاحب نظران دانشگاهی، دولتی و صنعتی هستیم که به شیوه های مختلف اما به میزان چشمگیر، در تهیه مطالب کتاب سهیم بوده اند. بویژه از F.N. Green، J.M. Goooge، W.I. Shadrowan، M.T. Borelli، R.C. Kryter، M.G. Thomason، W.T. Snyder، R. Weaver، Peebles E.R. Dougherty، D.Brazakovic، M.A. Abidi، H. Alter، B. Rock، Thornton J. Perez، R. Eason، M. Goldston، C. Hayden، D. Cate، P. Saba، B. Fittes همچنین از، A. Perez، R. Eason، M. Goldston، C. Hayden، D. Cate، P. Saba، B. Fittes، M.A. O'Neal، R. Eason، A. Morris، R. Salinas، F. Contreras، Z. Bell، Herrera David Eileen Bernadette Moran و Adison-Wesley از انتشارات Dwyer به خاطر همکاری بسیار خوبشان در تمام امور مربوط به چاپ سپاس ویژه داریم. در نهایت از افراد و سازمان هایی که نامشان در شرح اشکال متعدد کتاب ذکر شد، به خاطر اجازه استفاده از آن اشکال تشکر می کنیم.

این ویرایش از کتاب «پردازش تصویر رقمه» تنها انعکاسی از پیشرفت چشمگیری است که تنها در پنج سال اخیر (سالهای ۱۹۸۷ - ۱۹۹۲) در این زمینه حاصل شده است. طبق معمول، حتی پس از پایان کار روی متنون، پیشرفت ادامه می یابد. یکی از دلایلی که ویرایشها قبلی این کتاب آن طور خوب در سراسر دنیا مورد اقبال واقع شد، تأکید آن بر مفاهیم اساسی بود، که در آنها سعی شده بود حداقل نسبت لازم در شرح دانشی که به سرعت در حال توسعه است، ایجاد گردد. ماسعی کرده ایم که آن اصل را در تهیه این ویرایش از کتاب نیز مدنظر داشته باشیم.

فصل ۱

مقدمه

یک عکس بیش از ده هزار کلمه می‌ارزد.
گمنام

۱.۱ زمینه

علاقه به روش های پردازش تصویر رقیعی از دو محدوده کاربردی اصلی نشأت می‌گیرد که آن محدوده ها عبارتند از: بهبود اطلاعات تصویری به منظور تعبیر انسانی، و پردازش داده های صحنه برای ادراف ماشینی مستقل. یکی از اولین کاربردهای فنون پردازش تصویر در دسته اول، بهبود عکس های رقیعی شده روزنامه بود که از طریق کابل زیردریایی بین لندن و نیویورک منتقل می‌شد. استفاده از سامانه کابلی انتقال عکس Bartlane در ابتدای دهه ۱۹۲۰^۱، زمان مورد نیاز برای انتقال یک عکس از عرض دریای آتلانتیک را از بیشتر از یک هفته، به کمتر از سه ساعت کاهش داد. تجهیزات مخصوصی عکس هارا برای انتقال کابلی به صورت رمز در آورده و سپس آهارا در محل دریافت بازسازی می‌کردند. شکل ۱.۱ به این روش انتقال یافت و در یک چاپگر تلگراف که الگوی نیم تن^۱ را شبیه سازی می‌کرد، بازسازی شد.

بعضی از مشکلات اولیه در مورد بهبود کیفیت دیداری این تصاویر اولیه، با انتخاب شیوه های چاپ و توزیع سطوح روشنایی مرتبط بود. روش چاپ مورد استفاده در شکل ۱.۱ تا پایان سال ۱۹۲۱ با روش دیگری که مبتنی بر تجدید چاپ عکس با استفاده از نوارهای سوراخ شده در پایانه تلگراف بود، جایگزین شد. شکل ۱.۲. یک تصویر حاصل از این روش را نشان می‌دهد. بهبود نسبت به شکل ۱.۱ هم در کیفیت «ثُنی»^۲ و هم در تفکیک، آشکار است. سامانه های اولیه Bartlane قادر به رمزگاری تصویر با ۵ سطح روشنایی متفاوت بود. این توانایی در سال ۱۹۲۹ م. به ۱۵ سطح افزایش یافت. شکل ۱.۲.

¹⁾ Halftone



شکل ۱۰.۱ عکس رقمه که در سال ۱۹۲۱ از یک نوار رمز شده به وسیله یک چاپگر تلگرافی که دارای نوکهای چاپ مخصوص بود، تهیه شد (از [۱۹۷۲] McFarlane).

نوع تصویری را که می توان با استفاده از یک وسیله ۱۵ رنگ بدست آورده، نشان می دهد. در خلال این مدت، اختراع سامانه ظهر صفحه فیلم با استفاده از پرتوهای نوری مدوله شده با نوار عکس رمز شده، به مقدار زیادی فرآیند تجدید چاپ را بهبود داد.

در طی ۳۵ سال بعد، روش های پردازش عکس های رقمه انتقال یافته، بهبود یافت. با این حال توجه به مقاومت پردازش تصویر تازمان اختراع رایانه های بزرگ رقمه و کاربرد آنها در برنامه های فضایی به تأخیر افتاد. در سال ۱۹۶۴ در Jet Propulsion Laboratory (Pasadena, California) کار روی استفاده از فنون رایانه ای برای بهبود تصاویر ارسالی کاوشگر فضایی شروع شد. در آن هنگام عکس های ماه ارسالی از ۷ Ranger به وسیله رایانه پردازش می شد تا انواع گوناگون اعوجاج تصویر، ناشی از دوربین تلویزیونی کاوشگر را تصحیح کند. این فنون به عنوان مبنای روش های بهبود یافته 'ارتقا و بازیابی تصاویر حاصل از مأموریت های Surveyor' به ماه، مجموعه مأموریت های Mariner به مریخ،



شکل ۲۰.۱ عکس رقمه که در سال ۱۹۲۲ از نواری که پس از دو بار عبور سیگنالها از روی دریای آتلانتیک سوراخ شده بود، تولید شد. چند خط آشکار است. (از [۱۹۷۲] McFarlane).

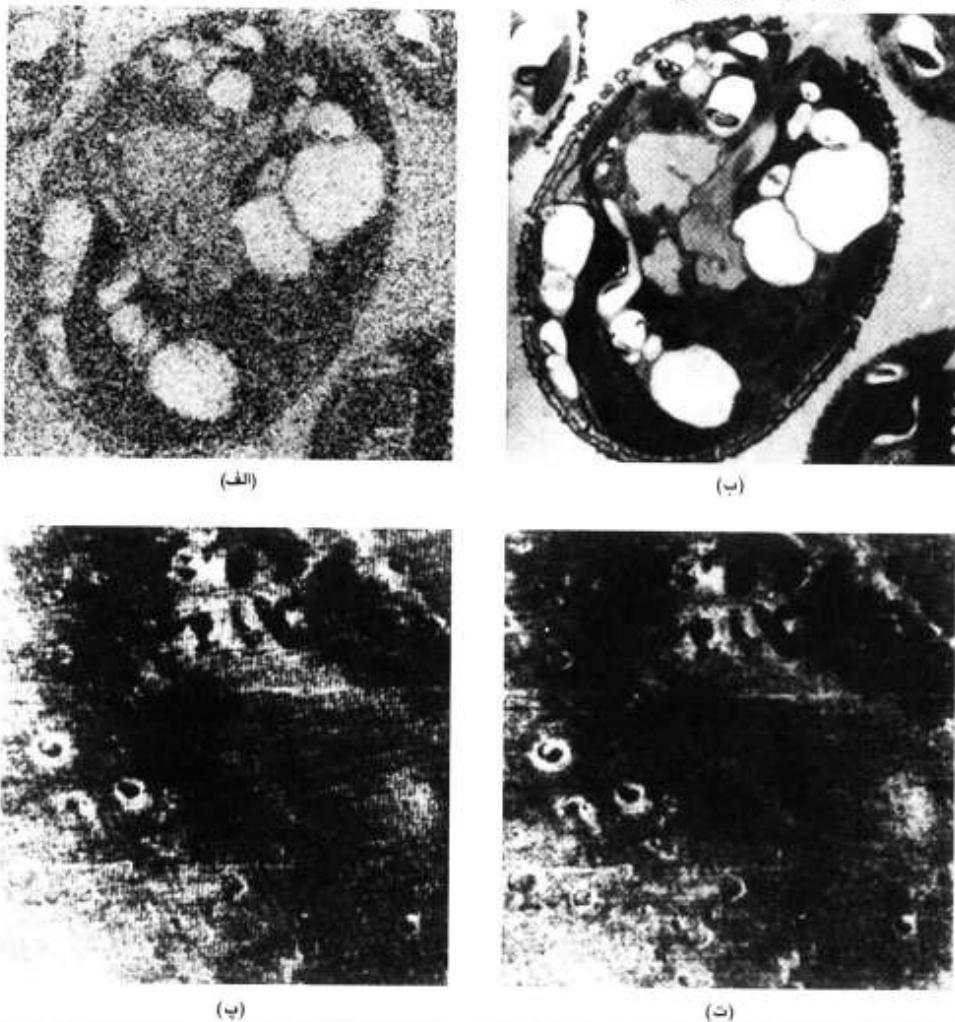


شکل ۳۰۱ عکس کابلی اصلاح نشده از *Pershing* و *Foch*، که در سال ۱۹۲۹ با یک دستگاه ۱۵ رنگ از لندن به نیویورک انتقال یافت. (از [۱۹۷۲] *McFarlane*).

پروازهای سرنشین دار Apollo به ماه و غیره به کار رفت.

از سال ۱۹۶۴ تاکنون، موضوع پردازش تصویر، رشد فراوانی کرده است. علاوه بر برنامه تحقیقات فضایی، اکنون از فنون پردازش تصویر، در موارد متعددی استفاده می شود. گرچه اغلب این مسائل با هم نامرتب هستند، اما عموماً نیازمند روش هایی هستند که قادر به ارتقای اطلاعات تصویری برای تعییر و تحلیل انسان باشد. برای نمونه در پزشکی شیوه های رایانه ای تمایز^۱ تصویر را ارتقا می دهند یا این که برای تعییر آسانتر تصاویر اشعه ایکس یا سایر تصاویر پزشکی، سطوح شدت روشنایی را بارنگ، رمز می کنند. متخصصان جغرافی نیز از این روش های روش های مشابه برای مطالعه الگوهای آلودگی هوا که با تصویربرداری هوایی و ماهواره ای بدست آمده است، استفاده می کنند. روش های ارتقا و بازیابی تصویر برای پردازش تصاویر آسیب دیده اشیای غیرقابل بازیافت یا تصاویر آزمایش هایی که تکرار آنها بسیار پرهزینه است، به کار می روند. در باستان شناسی نیز روش های پردازش تصویر برای بازیابی عکس های مات شده ای که تنها باقی مانده آثار هنری نادر هستند، مورد استفاده قرار می گیرد. در فیزیک و زمینه های مرتبط، فنون رایانه ای بارها تصاویر آزمایش های مربوط به موضوعاتی نظیر پلاسماهای پرانرژی و تصاویر ریزیبنی الکترونی را ارتقا داده اند. کاربردهای موفق دیگری از پردازش تصویر را نیز می توان در نجوم، زیست شناسی، پزشکی هسته ای، اجرای قانون، دفاع و صنعت بیان کرد.

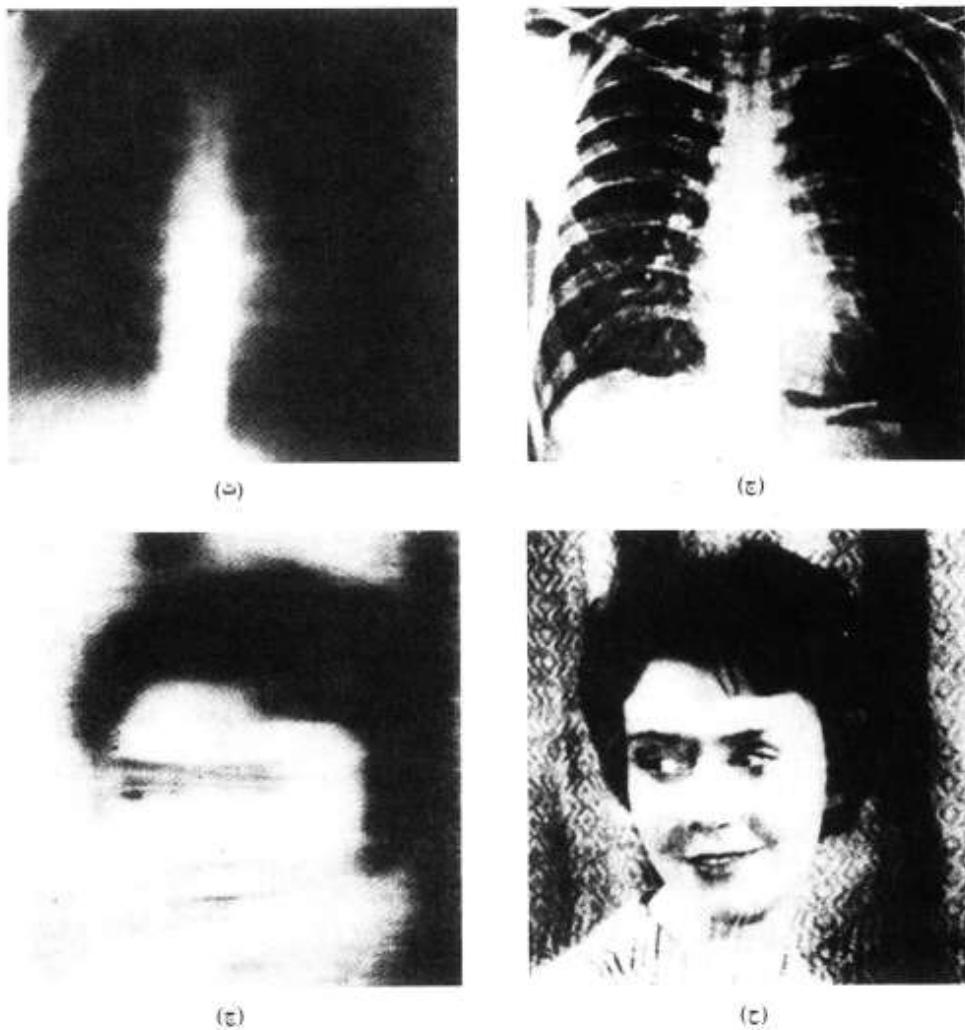
1) Contrast



شکل ۱.۴. مثالهایی از پردازش تصویر رقومی. ستون سمت چپ: تصاویر رقومی اولیه. ستون سمت راست: تصاویر پردازش شده

شکل ۱.۴. چند مثال نوعی از تابع قابل حصول با فنون پردازش تصویر رقومی را نشان می دهد.
تصاویر اولیه در سمت چپ و تصاویر پردازش شده متناظر آنها در سمت راست است. شکل ۱.۴.(الف) تصویر یک سلول است که نویز الکتریکی آن را شدید خراب کرده است. شکل ۱.۴.(ب) حاصل متوسط گیری از چند تصویر نویزدار را، که روش متدالوی برای کاهش نویز است، نشان می دهد. شکل ۱.۴.(ب) عکسی از سطح مریخ است که تداخل^۱، آن را در طی انتقال از کاوشگر فضایی به زمین خراب

1) interference



شکل ۱.۱ ادامه

کرده است. همان طور که شکل ۱.۴(ت) نشان می دهد، تداخل را، که در این حالت به صورت مجموعه ای از خطوط عمودی مشخص است، می توان تقریباً به طور کامل با پردازش رایانه ای برطرف کرد. اشکال ۱.۴(ب) و (ج) بهبود ممکن راروی تصویر اشعه ایکس به وسیله ارتقای تمایز و لبه نشان می دهد. تصویر شکل ۱.۴(ج) در خلال نور خوردنگی به وسیله حرکت یکنواخت مات شد، و تصویر شکل ۱.۴(ح) پس از اجرای الگوریتم مات زدایی به دست آمد. این نمونه ها نماینده مواردی هستند که به تفصیل در فصول ۴ و ۵ تشریح می شوند.

این مثالها نتایج پردازشی را نشان می دهد که برای بهبود تعبیر و دریافت انسانی اجرا شد. از موارد

مهم دیگر در کاربرد فنون پردازش رقمه که در ابتدای این بخش ذکر شد، حل مسائل مربوط به درک ماشینی است. در این مورد، شبیه های استخراج اطلاعات از تصویر به طوری که برای پردازش رایانه ای مناسب باشد، مورد توجه قرار می گیرد. غالباً این اطلاعات، به ویژگی های دیداری که انسانها برای تعییر محتوای تصویر استفاده می کنند، شbahت کمی دارند. مثالهایی از نوع اطلاعات مورد استفاده در ادراک ماشینی عبارتند از: گشتاورهای آماری، ضرایب تبدیل فوریه و سنجه های فاصله چند بعدی.

مسائل نوعی در ادراک ماشینی که به طور مکرر از فنون پردازش تصویر بهره می گیرند، عبارتند از: تشخیص خودکار نویسه^۱، بینایی ماشینی برای ساخت و وارسی تولیدات صنعتی، تشخیص نظامی، پردازش خودکار اثر انگشتها، نمایش تصاویر اشعه ایکس و نمونه های خونی، و پردازش ماشینی تصاویر هوایی و ماهواره ای برای پیش بینی های هواشناسی و ارزیابی محصولات کشاورزی.

۴.۱ نمایش تصویر رقمه

عبارت تصویر تک رنگ^۲ یا به طور ساده، تصویر به یک تابع شدت روشنایی دو بعدی (u, v) اشاره می کند، که u و v مختصات مکانی است و مقدار f در هر نقطه (u, v) متناسب با روشنایی (یا سطح خاکستری^۳) تصویر در آن نقطه است. شکل ۱. ۵. قرارداد محورها را که در سراسر کتاب استفاده می شود، نشان می دهد. گاهی دیدن تصویر به شکل پرسپکتیو که محور سوم آن روشنایی باشد، مفید است. اگر شکل ۱. ۵ به این طریق دیده شود، به صورت مجموعه ای از قله های تیز در نواحی دارای تغییرات متعدد روشنایی، و نواحی هموارتر یا ثابت در جاهایی که سطح روشنایی تغییرات کمی دارد یا ثابت است، دیده می شود. استفاده از قرارداد انتساب مقادیر نسبتاً بزرگتر به نواحی روشن تر، بلندی اجزای شکل را با روشنایی آن اجرا در تصویر متناسب می سازد.

تصویر رقمه، یک تابع تصویر (u, v) است که هم در مختصات مکانی و هم در روشنایی گستته شده باشد. تصویر رقمه را می توان ماتریسی دو بعدی در نظر گرفت که هر زوج از اندیس های سطری و ستونی آن یک نقطه تصویر را مشخص می کند. عناصر چنین آرایه رقمه را عناصر تصویر^۴، عناصر عکس^۵، پیکسلها^۶ یا پلها^۷ می خوانند که دو نام آخر مخفف های متداول برای «عناصر عکس» است.

1) Character

2) Monochrome image

3) Gray level

4) Image element

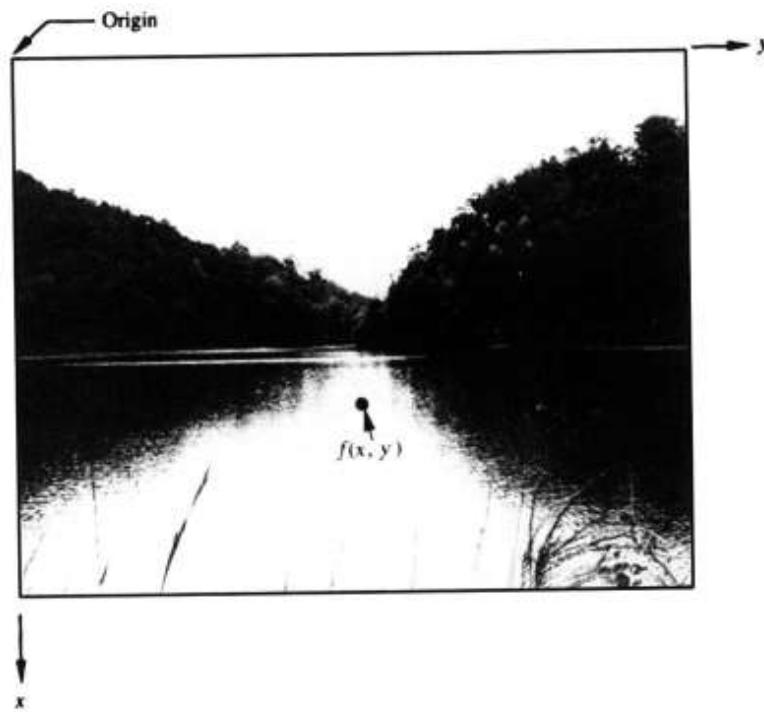
5) Picture element

6) Pixel

7) Pel

۱. مراحل اساسی پردازش تصویر

۲۵



شکل ۱. ۵. قرارداد محورها مورد استفاده در نمایش تصویر رقمه

گرچه ابعاد تصویر رقمه بسته به کاربرد متغیر است، فضول بعدی فواید زیاد انتخاب آرایه های مربعی را که ابعاد و تعداد سطوح خاکستری آنها توان های صحیح ۲ باشد، روشن می سازند. مثلاً یک اندازه 'نوعی' که از نظر کیفی با تصویر تلویزیونی قابل مقایسه است، آرایه 512×512 با 128 سطح خاکستری می باشد.

به استثنای شرح فصل ۴ درباره 'روش های رنگی ارتقای تصویر و یک شرح مختصر در فصل ۷ که به استفاده از رنگ در بخش بندی تصویر می پردازد، تمام تصاویر مورد بررسی در این کتاب تصاویر تک رنگ رقمه به شکلی که قبلآ توصیف شد، هستند. بنابراین در این کتاب به تحلیل صفحه سه بعدی و فنون نوری پردازش تصویر نمی پردازیم.

۱. مراحل اساسی پردازش تصویر

پردازش تصویر رقمه محدوده وسیعی از ساخت افزار، نرم افزار و مبانی نظری را در بر می گیرد. در

این بخش مراحل اساسی مورد نیاز برای اجرای یک پردازش بر تصویر را شرح می دهیم. موارد سخت افزاری و نرم افزاری در بخش ۱.۴ بیان می شوند.

برای روشن ساختن موضوع این بخش، از یک مثال پایه ای^۱ که در سراسر این بحث بکار رفته استفاده می کنیم. یک کاربرد که فهم آن بدون اطلاع قبلی از مفاهیم تصویربرداری نیز نسبتاً آسان است، استفاده از روش های پردازش تصویر برای خواندن آدرس روی بسته های پستی می باشد. همانطور که شکل ۱.۶ نشان می دهد، هدف کلی این فرآیند به دست آوردن خروجی ها از حوزه 'مساله'^۲ به وسیله 'پردازش تصویر' است. حوزه 'مساله' در این مثال، بسته های پستی، و هدف خواندن آدرس روی هر بسته است. بنابراین خروجی مطلوب در اینجا، رشتہ ای از نویسه های حرفی- عددی^۳ است.

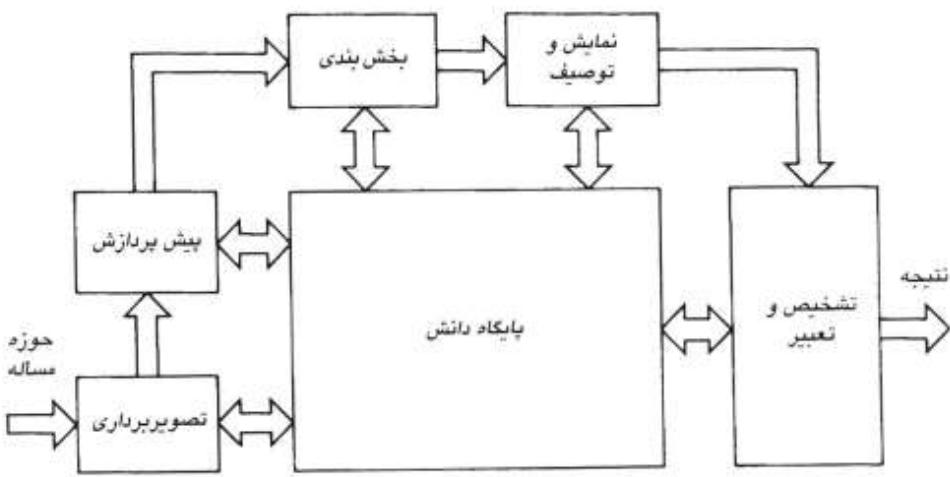
مرحله اول این فرآیند، تصویربرداری^۴ - یعنی به دست آوردن تصویر رقمه- است. انجام دادن چنین کاری نیازمند یک حسگر تصویربردار^۵ و قابلیت رقمه سازی سیگنال خروجی حسگر می باشد. همان طور که در بخش ۱.۴ با مقداری تفصیل بیان می شود، حسگر ممکن است یک دوربین تلویزیونی تک رنگ یا رنگی باشد که در هر $\frac{1}{3}$ ثانیه تصویر کاملی از حوزه 'مساله' تولید می کند. همچنین حسگر تصویربردار ممکن است یک دوربین پیماشگر خطی^۶ باشد که در هر پیماش یک خط تصویر را مرور می کند. در این حالت با حرکت شیء از مقابل پیماشگر خطی یک تصویر دو بعدی تولید می شود. اگر در ابتدا خروجی دوربین یا سایر حسگرهای تصویربردار به شکل رقمه نباشد، یک مبدل 'آنالوگ به رقمه'^۷ آن را رقمه می کند. ماهیت حسگر و تصویری که تولید می کند، به نوع کاربرد بستگی دارد. بنابراین طبق این مثال، خواندن آدرس بسته های پستی اغلب با استفاده از دوربینهای پیماشگر خطی انجام می شود.

پس از اینکه تصویر رقمه به دست آمد، مرحله بعدی پیش پردازش آن است. وظیفه اصلی پیش پردازش، بهبود تصویر به روش هایی است که امکان توفیق سایر پردازشها را نیز افزایش دهد. در این مثال، پیش پردازش، به طور معمول به روش هایی برای ارتقای تمایز، حذف نویز و جداسازی آن نواحی که زمینه شان نشان دهنده احتمال وجود اطلاعات حرفی- عددی است، می پردازد.

مرحله بعدی به بخش بندی می پردازد. در تعریف وسیع، بخش بندی^۷ فرآیندی است که تصویر

- | | | | |
|-------------------|----------------------|----------------------------|---------------------|
| 1) Theme example | 2) Problem domain | 3) Alpha-numeric character | 4) Image acquistion |
| 5) Imaging sensor | 6) Line- Scan camera | 7) Segmentation | |

۳.۱ مراحل اساسی پردازش تصویر



شكل ۱. مراحل اساسی در پردازش تصویر رقی

ورودی را به قسمت های اجزای تشکیل دهنده اش تقسیم می کند. به طور کلی بخش بندی یکی از مشکل ترین کارها در پردازش تصویر رقی است، از طرفی، یک شیوه قوی بخش بندی، تا حد زیادی فرآیند را به حل موفق مسأله مورد نظر نزدیک می کند. از طرف دیگر الگوریتم های ضعیف یا خطدار بخش بندی، تقریباً همیشه باعث خرابی اتفاقی^۱ می شوند. در تشخیص نویسه، نقش اصلی بخش بندی، استخراج نویسه ها و کلمات مجزا از متن است.

خروجی مرحله بخش بندی معمولاً، داده های پیکسلی خام است که یا مرز یک ناحیه یا تمام نقاط درون آن ناحیه را تشکیل می دهند. در هر دو حالت باید داده ها را به شکل مناسب برای پردازش رایانه ای تبدیل نمود. اولین تصمیمی که باید گرفته شود این است که آیا داده ها باید به صورت مرز یا به صورت یک ناحیه کامل نمایش داده شود. نمایش مرزی وقتی مفید است که مشخصات خارجی شکل نظیر گوشه ها یا خمیدگی ها مورد نظر باشد. نمایش ناحیه ای وقتی مفید است که خواص درونی بخش های تصویر نظیر بافت یا استخوان بندی شکل مورد توجه باشد. در هر حال در بعضی کاربردها هر دو نمایش به کار می رود. این وضعیت در کاربردهای تشخیص نویسه که اغلب نیازمند الگوریتم هایی براساس شکل مرز و نیز شکل استخوان بندی و سایر خواص درونی هستند، رخ می دهد.

1) Eventual failure

انتخاب یک روش نمایش، تنها قسمتی از راه تبدیل داده های خام به شکل مناسب برای پردازش بعدی رایانه ای است. روشی نیز باید برای توصیف داده ها تعیین شود به طوری که ویژگی های مورد علاقه برجسته شوند. توصیف^۱، که انتخاب ویژگی^۲ نیز خوانده می شود، به استخراج ویژگی هایی که مقداری از اطلاعات کمی مورد نظر را به ما می دهند یا برای تشخیص گروهی از اشیا از گروه دیگر، اساسی هستند، می پردازد. از نظر تشخیص نویسه، توصیف گرهایی نظیر حفره ها (سوراخها) و شکافها ویژگی های مهمی است که کمک می کند تا بعضی حروف را از سایر آنها تشخیص دهیم.

مرحله آخر شکل ۱.۶، شامل تشخیص و تعبیر است. تشخیص^۳ فرآیندی است که براساس اطلاعات حاصل از توصیف گرها یک برچسب را به یک شیء متناسب می کند. تعبیر^۴ شامل انتساب معنا به یک مجموعه از اشیای تشخیص داده شده است. از نظر مثال کاربردی اخیر، شناسایی یک نویسه، مثلاً به عنوان C، نیازمند مربوط نمودن توصیف گرهای آن نویسه با برچسب C است. فرآیند تعبیر می خواهد به هر مجموعه از چیزهای برچسب خورده معنای را مرتبط کند. مثلاً یک رشته پنج عددی - پنج عدد، یک خط و چهار عدد دیگر - را می توان به عنوان یک شناسه' پستی^۵ تعبیر کرد.

تاکنون مطلبی درباره 'نیاز به دانش قبلی یا درمورد تعامل بین پایگاه دانش^۶ و واحد های پردازش موجود در شکل ۱.۶ نگفته ایم. دانش درباره 'حوزه' مسأله به شکل پایگاه داده دانش در درون سامانه پردازش تصویر، ذخیره می شود. این دانش ممکن است، تنها دانستن محل نواحی دارای جزئیات مورد علاقه باشد. بنابراین جستجوی مورد نیاز برای آن اطلاعات محدود می شود. پایگاه دانش ممکن است کاملاً پیچیده باشد، نظیر فهرست به هم مرتبط تمام مقایض اصلی ممکن در یک مسأله بازرسی مواد یا یک پایگاه داده تصویری که حاوی تصاویر ماهواره ای نفکیک بالا از یک منطقه در ارتباط با کاربردهای آشکار سازی تغییر^۷ باشد. پایگاه دانش علاوه بر هدایت عمل هر واحد پردازش، بر تعامل بین واحد ها نیز نظارت می کند. شکل ۱.۶ این تفاوت را با استفاده از پیکان های دوسر بین واحد های پردازش و پایگاه دانش و پیکان های یک سر که واحد های پردازش را به هم پیوند می دهند، نشان می دهد. این نمودار نشان می دهد که ارتباط بین واحد های پردازش اغلب براساس دانش قبلی در مورد نتیجه پردازش است. مثلاً برای این که ماشین نتیجه بگیرد که رشته ای از نویسه ها یک شناسه' پستی است، باید دانشی به

1) Description	2) Feature selection	3) Recognition	4) Interpretation
5) ZIP code	6) Knowledge base	7) Change detection	

۴.۱ عناصر سامانه های پردازش تصویر رقمه

سامانه داده شود که بتواند محل این رشتہ را نسبت به سایر اجزای میدان آدرس تشخیص دهد. این پایگاه دانش نه تنها عمل هر واحد را هدایت می کند، بلکه به عملیات بازخورد^۱ بین واحد های نیز کمک می کند. برای نمونه، رشتہ ای از اعداد در محل صحیح اما دارای تنها چهار نوبت (که یکی از آنها را نمی توان تشخیص داد) ممکن است باعث شود که واحد تعبیر در مورد اتصال دونویسه شک کند. تقاضای بازخورد از طریق پایگاه دانش به طبقه بخش بندی برای «نگاه» مجدد مثالی از کاربرد دانش در پردازش تصویر است.

گرچه اکنون نمایش تصویر را مبسوط شرح نمی دهیم، اما باید بدانیم که نتایج پردازش تصویر ممکن است در خروجی هر یک از مراحل شکل ۱.۶ نظاره شوند. همچنین یادآور می شویم که تمام کاربردهای پردازش تصویر به تعامل های پیچیده شکل ۱.۶ نیاز ندارند؛ کاربردهای عملی متعددی به وسیله توابعی که در مسیر بیرونی شکل ۱.۶ فراهم شده است، انجام می شود. در بعضی موارد، حتی هیچ یک از آن واحد ها نیاز نیستند. مثلاً ارتقای تصویر برای تعبیر دیداری انسانی بذرگان از مرحله پیش پردازش فراتر می رود. به طور کلی کارهای پردازشی که شامل تشخیص و تعبیر است، به کاربردهایی از تحلیل تصویر مبسوط می شوند که در آنها هدف استخراج خودکار یا حتی نیمه خودکار اطلاعات از تصویر است. تشخیص، تنها یک مثال آن است. کاربردهای دیگری نیز در ادامه مطالب کتاب ارائه می شود.

۴.۲ عناصر سامانه های پردازش تصویر رقمه

عناصر یک سامانه همه منظوره که قادر به انجام دادن اعمال پردازش تصویر مشروح در بخش ۱.۳ باشد، در شکل ۱.۷ نشان داده می شود. از این نوع سامانه معمولاً برای اعمال (۱) تصویربرداری (۲) ذخیره سازی (۳) پردازش (۴) ارتباطات و (۵) نمایش تصویر استفاده می شود.

۱۰.۱ تصویربرداری

دو عامل برای به دست آوردن تصاویر رقمه مورد نیاز است: یک وسیله فیزیکی که حساس به نواری در طیف انرژی الکترومغناطیسی (نظیر نوارهای اشعه ایکس، ماوراءنفکش، مرئی یا مادون قرمز) باشد و یک سیگنال الکتریکی خروجی مناسب با سطح انرژی احساس شده تولید کند؛ و یک رقمه ساز^۲،

1) Feed back

2) Digitizer

وسیله ای که برای تبدیل خروجی الکتریکی حسگر به شکل رقمی است. به عنوان مثال اصول سامانه های تصویر برداری اشعه ایکس را در نظر آورید. خروجی منبع اشعه ایکس روی شیی انداخته می شود و یک رسانه حساس به اشعه ایکس در طرف دیگر شیئ قرار می گیرد. این رسانه، تصویری از بدن (نظیر استخوانها و اعضای بدن) که با درجات مختلفی اشعه ایکس را جذب می کند، ارائه می کند. رسانه مورد استفاده، ممکن است یا فیلم، یا دوربین تلویزیونی همراه با مبدل اشعه ایکس به فوتون، یا آشکارسازهای گسته که با ترکیب خروجی هایشان تصویری رقمی بازسازی می شود، باشد.

دسته اصلی دیگری از حسگرها با نور مرئی و مادون قرمز کار می کند. چند تا از وسایلی که بیش از موارد دیگر برای این منظور استفاده شده عبارت است از: ریز چگالی سنج ها^۱، تشریح کننده های تصویر^۲، دوربینهای ویدیکون^۳، و آرایه های حالت جامد حساس به نور. در اولین وسیله تصویری که باید رقمی شود، لازم است شفاف (نظیر فیلم منفی)^۴ یا عکس باشد. دوربین های ویدیکون و آرایه های حالت جامد می توانند تصاویر ضبط شده با این روش را دریافت کنند، و همچنین می توانند تصاویر طبیعی که شدت نور کافی برای تحریک آشکارساز داشته باشد را رقمی کند.

در ریز چگالی سنج ها، ورق شفاف یا فیلم روی یک سطح صاف قرار می گیرد یا این که حول یک استوانه پیچیده می شود. پیماش با تمرکز یک پرتو نوری (مانند لیزر) روی فیلم و حرکت دادن سطح زیرین یا چرخاندن استوانه نسبت به آن پرتو انجام می شود. در مورد ورقهای شفاف، پرتو از میان فیلم عبور می کند و در عکس ها از سطح تصویر منعکس می شود. در هر دو مورد، اشعه روی یک آشکارساز نوری متمرکز می شود و سطح خاکستری هر نقطه تصویر براساس شدت پرتو به وسیله آشکارساز ضبط می شود. اگر تنها مقادیر گسته شدت و محل نقاط تصویر در خروجی مجاز باشد، تصویری رقمی به دست می آید. گرچه ریز چگالی سنج ها وسایل کندی است، برای طبیعت ذاتاً پیوسته حرکت مکانیکی مورد استفاده در فرآیند رقمی سازی، قادر است به درجات بالایی از دقت مکانی دست یابد.

کار دوربینهای ویدیکون بر اساس اصل نور رسانایی (رسانش بر اثر نور)^۵ است. تصویری که روی سطح لامپ متمرکز می شود، الگویی بر رسانایی متغیر تولید می کند که با توزیع روشنایی تصویر مطابقت

1) Microdensitometer

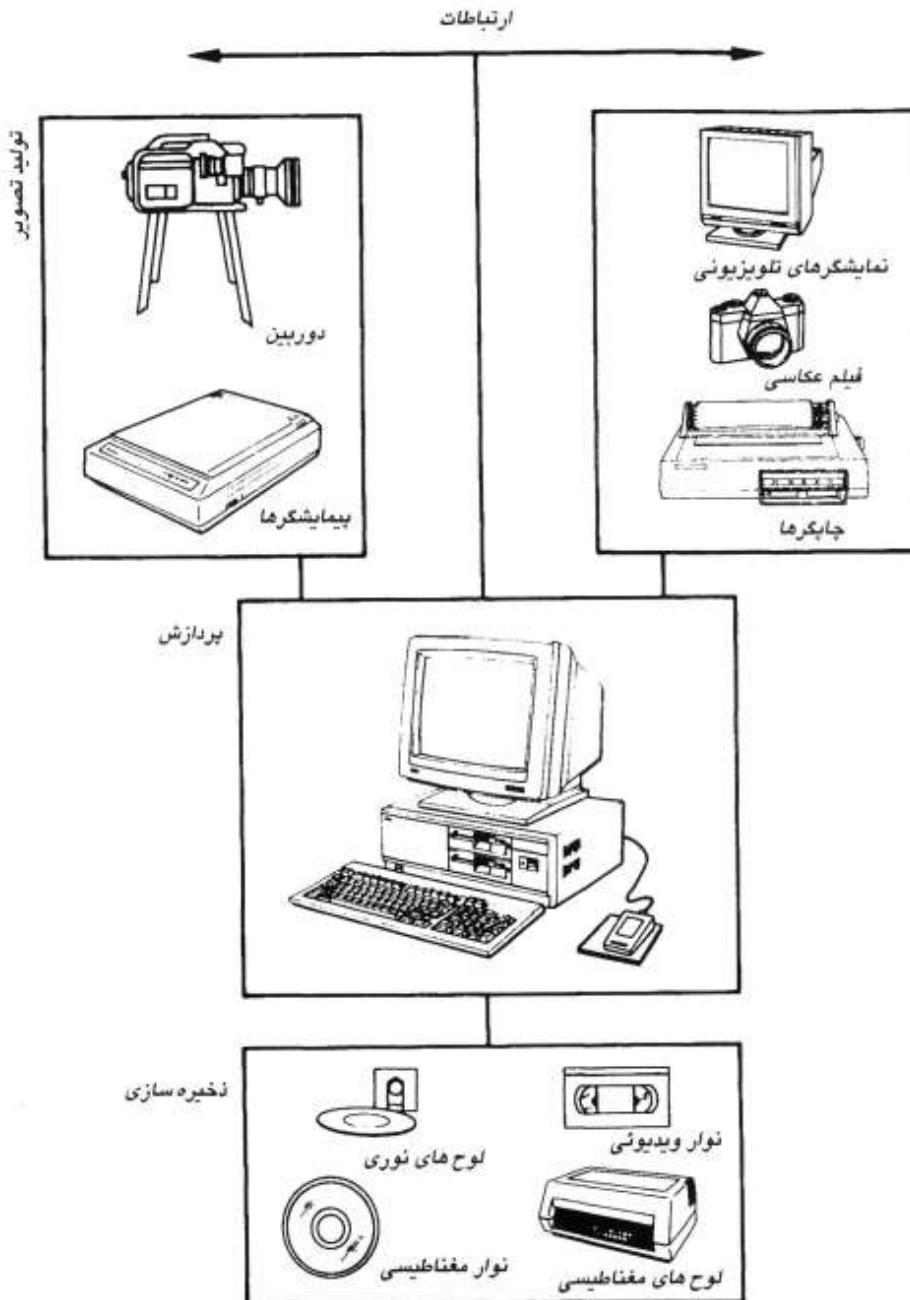
2) Image dissector

3) Vidicon camera

4) Film negative

5) Photo conductivity

۷.۱ عناصر سامانه های پردازش تصویر رقمنی



شکل ۱.۱ عناصر عملیات پایه ای در سامانه پردازش تصویر: تصویربرداری، ذخیره سازی، پردازش، ارتباطات، و نمایش. اجزای رسم شده درون هر جعبه مثال هایی از وسائل مورد استفاده در چنان سامانه ای

هستند

دارد. یک اشعه الکترونی مستقل کاملاً مرکز سطح پشت هدف نور رسانا را پیمایش می کند تا با خشی سازی بار الکتریکی هر نقطه آن، سیگنال ولتاژی مناسب باروشنای آن نقطه تصویر در جمع کننده^۱ تولید کند. با چندی سازی سیگنال حاصل و محل پرتو پیمایشگر، تصویری رقمی به دست می آید.

آرایه های حالت جامد از عناصر تصویربردار گسته سیلیکونی با نام فوتوسایت^۲ تشکیل می شود که ولتاژ خروجی آنها مناسب با شدت نور دریافتی است. آرایه های حالت جامد در یکی از این دو آرایش هندسی زیر مرتب می شوند: حسگرهای پیمایشگر خطی^۳ و حسگرهای ناحیه ای^۴. حسگر پیمایشگر خطی دارای یک ردیف فوتوسایت است که با حرکت نسبی صحنه و فوتوسایتها، تصویری دو بعدی تولید می کند. مثلاً حسگرهای پیمایشگر خطی به زیادی در پیمایشگرهای تخت تصویر به کار می روند. یک حسگر ناحیه ای مشکل از ماتریسی از فوتوسایتها است و بنا بر این می تواند به همان روش لامپ ویدیکون تصویربرداری کند. مزیت مهم حسگرهای آرایه ای حالت جامد این است که می توان آنها را با سرعت های خیلی بالا (مثلاً یک بار در هر $\frac{1}{1000}$ ثانیه) باز و بسته کرد^۵. این مزیت، آنها را برای کاربردهایی که باید تصویر بدون حرکت^۶ باشد، مطلوب می سازد.

فناوری مورد استفاده در حسگرهای تصویربردار حالت جامد از فناوری وسائل تزویج بار^۷ (سی سی دی) گرفته شده است. همان طور که شکل ۱.۸ (الف) نشان می دهد، حسگر سی سی دی پیمایشگر خطی نوعی دارای این اجزا است: یک ردیف از فوتوسایتها، دو دریچه انتقال برای انتقال زمان بندی شده محتويات ثبات های حمل به درون ثبات های حمل، و یک دریچه خروجی برای حمل زمان بندی شده محتويات ثبات های حمل به درون تقویت کننده. در خروجی تقویت کننده سیگنال ولتاژ مناسب با محتويات ردیف فوتوسایتها تولید می شود.

همان طور که شکل ۱.۸ (ب) نشان می دهد، آرایه های سی سی دی شبیه به حسگرهای پیمایشگر خطی است، جز این که در سی سی دی فوتوسایتها در قالب ماتریسی مرتب می شود و ترکیبی از دریچه و ثبات حمل ستونهای فوتوسایت را از هم جدا می کند. محتويات فوتوسایتهای با شماره فرد به ترتیب وارد ثبات های حمل عمودی و سپس وارد ثبات حمل افقی می شود. محتوا ای این ثبات نیز به تقویت کننده ای که خروجیش خطی از تصویر است، وارد می شود. تکرار این روند برای خطوط با شماره زوج، میدان

1) Collector

2) Photosite

3) Line-scan sensor

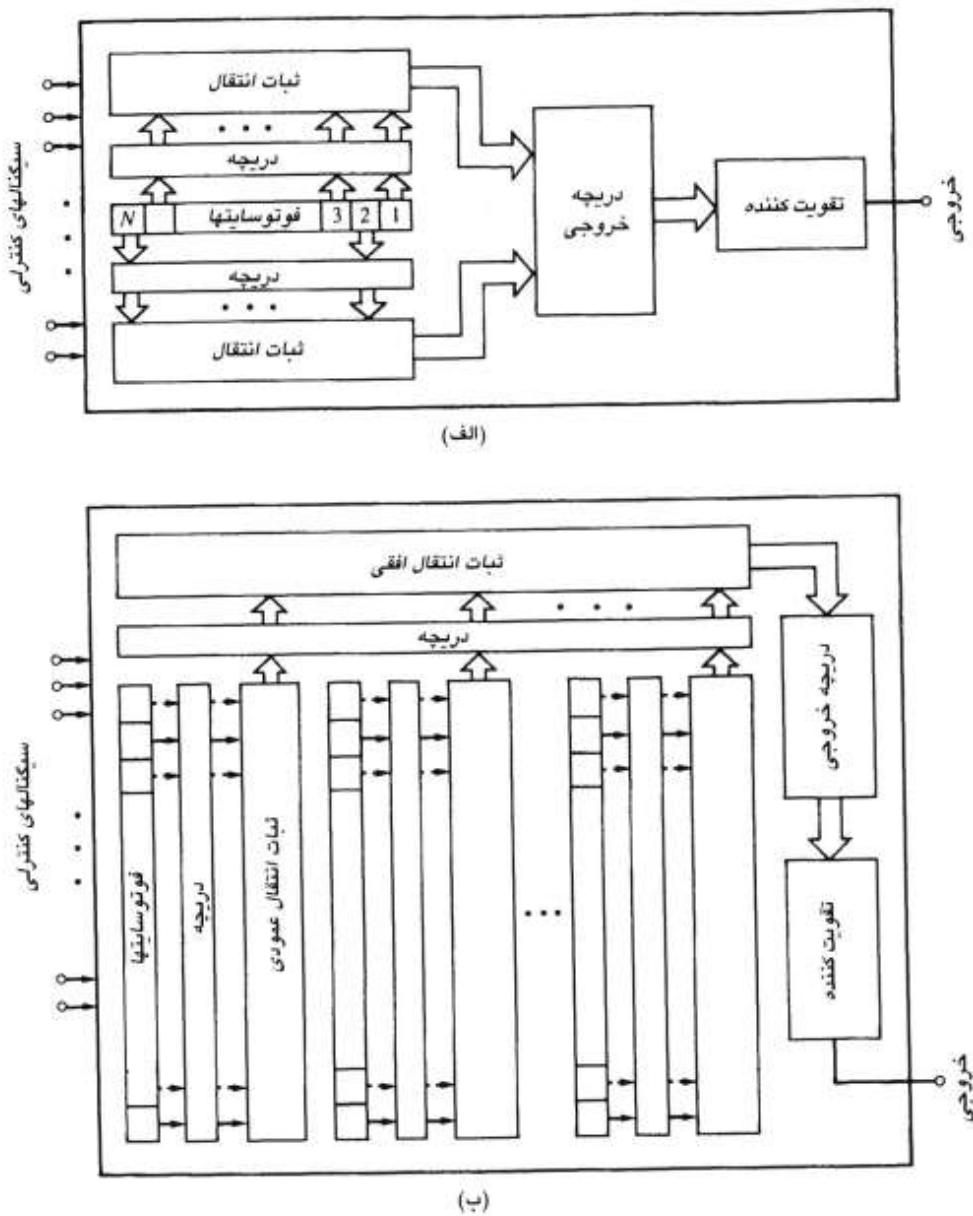
4) Area sensor

5) Shuttered

6) Freezing motion

7) Charge-coupled device (CCD)

۴.۱ عناصر سامانه های پردازش تصویر رقمنی ۳۳



شکل ۱.۸ (الف) حسگر سی سی دی پیماشگر خطی؛ (ب) حسگر سی سی دی ناحیه ای

دوم یک قاب تصویر تلویزیونی در هم بافته را تولید می کند؛ عمل پیمایش ۳۰ بار در ثانیه تکرار می شود. حسگرهای پیمایشگر خطی با تفکیک هایی بین ۲۵۶ تا ۴۰۹۶ عنصر متداول است. تفکیک حسگرهای ناحیه ای از 32×32 عنصر در کمترین تفکیک تا 256×256 عنصر برای یک حسگر با تفکیک متوسط تغییر می کند. وسایل با تفکیک بالاتر 480×640 عنصری به آسانی در دسترس است، و حسگرهای با تفکیک حدود 1240×1280 عنصر نیز به طور تجاری در قیمت های نسبتاً بالا اما غالباً تعديل پذیر در دسترس می باشد. حسگرهای ویژه با بالاترین تفکیک که برای دستیابی به تفکیک هایی از مرتبه 2048×2048 عنصر از حرکت مکانیکی یک تراشه سی سی دی استفاده می کند، نیز با قیمت های بالا قابل دسترسی است. آرایه های سی سی دی معمولاً در دوربین های تلویزیونی مورد استفاده قرار می گیرد. همانطور که قبل ذکر شد، با دادن سیگنال خروجی این دوربین ها به رقمنی ساز، تصویر رقمنی تولید می شود.

۴.۴.۱ ذخیره سازی

یک تصویر ۸ بیتی با ابعاد 1024×1024 پیکسل نیازمند یک میلیون بایت حافظه است. بنابراین فراهم کردن حافظه کافی معمولاً یکی از مشکلات سامانه های پردازش تصویر است. حافظه های رقمنی در کارهای پردازش تصویر به سه گروه اصلی تقسیم می شود: (۱) حافظه کوتاه مدت برای استفاده در زمان پردازش، (۲) حافظه در حال کار برای بازخوانی نسبتاً سریع، و (۳) حافظه بایگانی که از مشخصات آن فرکانس دستیابی کم است. مقدار حافظه بر حسب بایت (۸ بیت)، کیلو بایت (یک هزار بایت)، مگابایت (یک میلیون بایت)، گیگابایت (یک میلیارد بایت)، و ترابایت (یک تریلیون بایت) سنجیده می شود.

یک روش برای فراهم کردن حافظه کوتاه مدت، حافظه رایانه است. روش دیگر استفاده از بردهای مخصوص که ابافر قاب^۱ خوانده می شوند، است. این بردها می توانند یک یا چند تصویر را ذخیره کند و می توان با سرعت بالا، معمولاً در نرخ ویدیو (۳۰ تصویر کامل در ثانیه)، به آنها دستیابی داشت. روش دوم، تمرکز^۲، انتقال افقی^۳، و نیز انتقال عمودی^۴ مجازاً آنی روی تصویر را ممکن می سازد. مقدار حافظه روی یک کارت^۵ حافظه (بافر قاب) به ابعاد فیزیکی و چگالی حافظه تراشه های آن بستگی دارد.

1) Board

2) Frame buffer

3) Zoom

4) Pan

5) Scroll

6) Card

۴.۱ عناصر سامانه های پردازش تصویر رقی

۳۵

وجود ۳۲ مگابایت حافظه روی یک کارت با فرآب غیر معمول نیست.

حافظه های در حال کار به طور معمول به شکل لوح های مغناطیسی است. لوح های وینچستر^۱ با صدھا مگابایت ظرفیت رایج است. در یک فناوری جدیدتر، که حافظه مغناطیسی - نوری^۲ (MO) خوانده می شود، با استفاده از پرتو لیزر و مواد مخصوص دیگر، حافظه ای تزدیک یک گیگابایت روی یک لوح نوری^۳ ۵ اینچی به دست می آید. مشخصه اصلی حافظه در حال کار، دستیابی مکرر به داده ها است. بنابراین در کاربردهای پردازش تصویر بندرت از نوارهای مغناطیسی و سایر رسانه های ترتیبی به عنوان حافظه در حال کار استفاده می شود. گرامافون ها^۴ که ۳۰ تا ۱۰۰ برابر لوح نوری حافظه ذخیره سازی دارند، وسیله مناسبی است برای کاربردهایی که به حافظه بزرگ در حال کار با قابلیت خواندن - نوشتن نیاز دارند. گرامافون های نوری از این نظر که یک سامانه مکانیکی لوح های نوری را به (از) لوح گردانها وارد (خارج) می کند، مشابه گرامافون صوتی است.

درنهایت مشخصه حافظه بایگانی، نیاز به حافظه زیاد اما فرکانس دستیابی کم است. نوارهای مغناطیسی و لوح های نوری، رسانه متداول برای کاربردهای بایگانی است. نوارهای مغناطیسی با چگالی بالا (۶۴۰۰ بایت در اینچ) می توانند تصویر یک مگابایتی را در حدود ۱۳ فوت از نوار ذخیره کند. مشکل اصلی نوارهای مغناطیسی عمر بایگانی نسبتاً کوتاه آنها - حدود هفت سال - و نیاز به محیط نگهداری با مراقبت است. با استفاده از فناوری فعلی نوشتمن یکبار - خواندن فراوان^۵ (WORM) لوح نوری می توان در حدود یک گیگابایت را روی یک لوح^۶ ۵ اینچی ذخیره کرد. برخلاف لوح های MO، لوح های WORM با ضربیب شکل های^۷ بزرگتر با قابلیت ذخیره تقریباً ۶ گیگابایت روی لوح های ۱۲ اینچی و کمی بیش از ۱۰ گیگابایت روی لوح های ۱۴ اینچی موجود است. لوح های WORM گرچه پاک شدنی نیست، اما عمر بایگانی آنها بدون نیاز به شرایط ویژه بیش از سی سال است. وقتی لوح های WORM در یک گرامافون ذخیره شود، می توان از آنها به عنوان وسایل ذخیره سازی در حال کار، در کاربردهایی که اغلب فقط نیازمند خواندن است، استفاده کرد. اکنون در یک گرامافون که حجمی کمتر از ۱۵۰ فوت مکعب دارد، ایجاد یک تراپایت حافظه WORM امکان پذیر است. این ظرفیت معادل حافظه مورد نیاز برای ذخیره سازی یک میلیون تصویر^۸ $10^{24} \times 10^{24}$ بیتی است.

1) Winchester

2) Magnetic-optical

3) Juke box

4) Write-Once-Read-Many

5) Form factor

مطلوب قبلی درباره 'حافظه' بایگانی رقمی بود. در کاربردهایی که بازیابی تصویر به شکل رقمی مورد نیاز نیست، ذخیره تصاویر به شکل آنالوگ، معمولاً با استفاده از فیلم عکاسی یا نوار ویدیو، متداول است.

۳.۴.۱ پردازش

در پردازش^۱ تصاویر رقمی، معمولاً از شبوهایی که به شکل الگوریتم بیان می شود، استفاده می گردد. بنابراین غیر از تصویربرداری و نمایش تصویر، می توان اغلب عملیات پردازش تصویر را با نرم افزار اجرا کرد. تنها علت استفاده از سخت افزار ویژه پردازش تصویر، نیاز به سرعت بالا در بعضی کاربردها و یا غلبه بر بعضی محدودیت های اساسی رایانه است. مثلاً یک کاربرد مهم از تصویربرداری رقمی، ریزبینی در نورکم است. برای کاهش نویز تصویر باید چند متوسط گیری روی تصاویر متوالی با نرخ قاب^۲ (غالباً ۳۰ قاب در ثانیه) انجام شود. ساختار بزرگراه در غالب رایانه ها، جز چند رایانه بسیار کارآمد، نمی تواند به سرعت داده مورد نیاز برای اجرای این عمل دست یابد. بنابراین سامانه های پردازش تصویر امروزی ترکیبی از رایانه های متداول و سخت افزارهای ویژه پردازش تصویر است که کارهای آنها به وسیله نرم افزار در حال اجرا روی رایانه اصلی هدایت می شود.

مدل های متعدد سامانه های پردازش تصویر که در حدود نیمده^۳ دهه' ۸۰، در سراسر دنیا به فروش رسید، وسائل جانی نسبتاً بزرگی بود که به رایانه های میزبان با همان بزرگی متصل می شد. در انتهای دهه' ۸۰ و ابتدای دهه' ۹۰، سخت افزارهای تجاری پردازش تصویر به شکل برد های تکی که برای سازگاری با بزرگراه های استاندارد صنعتی و انطباق با کارگاه های کوچک مهندسی و رایانه های شخصی طراحی شده بود، تغییر یافت. این تغییر علاوه بر کاهش هزینه ها، یکی از عوامل تأسیس شرکت های زیادی با تخصص تولید نرم افزار پردازش تصویر بود.

گرچه هنوز هم سامانه های پردازش تصویر بزرگ برای کاربردهای نظری پردازش تصاویر ماهواره ای به فروش می رستند، حرکت به سوی کوچک سازی و ایجاد رایانه های کوچک همه منظوره که به

^۱ در شکل ۱. از یک رایانه 'کوچک' به عنوان نشانه پردازش استفاده شد. البته گاهی برای حل مسائل پردازش تصویر با مقیاس بالا رایانه های بزرگ و سامانه های ابر رایانه 'ویژه ای' مورد نیاز هستند.

² Frame rate

۴.۱ عناصر سامانه های پردازش تصویر رقمی

۳۷

سخت افزار پردازش تصویر مجهز است همچنان ادامه دارد. سخت افزار اصلی تصویربرداری که به این رایانه ها افزوده می شود، ترکیبی از یک «رقمی ساز و بافر قاب» برای رقمی سازی و ذخیره موقت تصویر، یک «واحد محاسبه و منطق» (ALU) برای اجرای عملیات حسابی و منطقی در نرخ قاب و یک یا چند «بافر قاب» برای دستیابی سریع به داده های تصویر در طول پردازش است. امروزه می توان نرم افزارهای پردازش تصویر پایه ای زیادی را در بازار تهیه کرد. وقتی این نرم افزارها با سایر نرم افزارهای کاربردی نظر صفحات گسترده و گرافیک ترکیب شود، نقطه شروع خوبی برای حل مسائل خاصی از پردازش تصویر فراهم می شود. وسایل نمایش و نرم افزارهای کارآمد پردازش کلمه و تولید گزارش، ارائه نتایج را تسهیل می کند. اغلب نتایج به دست آمده با چنین سامانه هایی به برد های پردازش تصویر سریع و خاصی که با بزرگراه مورد استفاده سازگارند، انتقال می یابد.

یکی از مشخصه های علم پردازش تصویر، عدم استفاده از یک راه حل برای کاربردهای متفاوت است. بنابراین فونی که در یک مورد خوب کار می کنند، ممکن است در دیگری کاملاً ضعیف باشند. تنها قایده وجود سخت افزار قوی و نرم افزار پایه در حال حاضر این است که نقطه شروع کار نسبت به یک دهه پیش بسیار پیشرفته تر (و بازای صرفاً کسری از هزینه آن موقع) می باشد. به طور کلی هنوز هم پیدا کردن راه حل واقعی برای یک مسئله خاص نیازمند تحقیق و توسعه فراوان است. مباحثی که در فصول بعدی مطرح می شود، فقط برای این نوع فعالیت ها ابزارهایی را ارائه می کند.

۴.۴.۱ مخابرات

استفاده از مخابرات در پردازش تصویر رقمی معمولاً در مورد ارسال داده های تصویری است و اغلب شامل مخابره محلی بین سامانه های پردازش تصویر و مخابره راه دور از یک نقطه به نقطه دیگر می باشد. برای اغلب رایانه ها، سخت افزار و نرم افزار مخابره محلی به آسانی در دسترس است. بیشتر کتب راجع به شبکه های رایانه ای به روشنی توافق های^۱ مخابره استاندارد را توضیح می دهند.

اگر داده های تصویری فشرده نشده در فواصل طولانی مخابره شوند، مشکل جدی تری ایجاد می شود. تاکنون باید این موضوع روشن شده باشد که تصاویر رقمی حاوی مقادیر بزرگی از داده ها هستند، اما یک خط تلفن صوتی حداقل می تواند ۹۶۰۰ بیت بر ثانیه را انتقال دهد؛ بنابراین ارسال یک تصویر

۸ بیتی 512×512 با این نرخ بیت، تقریباً پنج دقیقه طول می کشد. ارتباط بسیم با استفاده از ایستگاه های میانی نظری ما هواره ها بسیار سریعتر است، گرچه هزینه بیشتری دارد. نکته این است که در بسیاری مواقع، ارسال تصاویر کامل در فواصل طولانی مورد نیاز است. در فصل ۶ نشان خواهیم داد که فشرده سازی و افشرده سازی و (ابساط)^۱ داده ها نقش اصلی را در حل این مشکل بر عهده دارد.

۵.۴.۱ نمایش

نمایشگرهای تک رنگ و رنگی تلویزیونی ابزار اصلی نمایش در سیستم های پردازش تصویر امروزی است. همان طور که در بخش ۳.۴.۱ تشریح شد، نمایشگرها با خروجی (های) واحد سخت افزاری نمایش تصویر که در برداشی رایانه میزبان قرارداد و یا بخشی از سخت افزار پردازشگر تصویری می باشد، راه اندازی می شود. همچنین می توان سیگنالهای خروجی از واحد نمایش را به یک وسیله^۲ ثبت تصویر داد تا تصویری چاپی (اسلايد، عکس یا ورق شفاف)^۳ از آنچه روی نمایشگر دیده می شود، تولید کند. لامپ های اشعه کاتدی (CRT) با دستیابی تصادفی و دستگاه های چاپ نیز به عنوان وسائل نمایش استفاده می شود.

در نمایشگرهای CRT، رایانه با تولید سیگنال های راه انداز مورد نیاز، محل اشعه الکترونی را در هر لحظه معین می کند تا تصویر ایجاد شود. در هر نقطه، یک ولتاژ که مناسب با مقدار روشنایی نقطه متناظر در آرایه عددی است، شدت پرتو را مدوله می کند. مقدار شدت پرتو حاصل بین صفر برای نقاط سیاه تا مقدار بیشینه برای نقاط سفید متغیر است. یک دوربین عکاسی که روی صفحه CRT متمرکز است، الگوی نوری با شدت متغیر حاصل را ضبط می کند.

وسایل چاپ تصویر اغلب برای پردازش تصویر با تفکیک پایین مفید است. یک روش ساده برای تولید تصاویر خاکستری روی کاغذ، استفاده از قابلیت چند ضربه ای^۴ چاپگر خطی استاندارد است. می توان با تعداد و چگالی نویسه های روی هم چاپ شده در هر نقطه تصویر، سطح خاکستری آن نقطه را تنظیم نمود. با انتخاب مناسب مجموعه نویسه ها، می توان با یک برنامه رایانه ای ساده و نویسه های نسبتاً کم به توزیع های خوب سطح خاکستری دست یافت. ضمیمه الف حاوی مثالهای از این روش است. چند وسیله متداول دیگر برای ثبت تصویر روی کاغذ، چاپگرهای لیزری، وسایل دارای کاغذ حساس به

1) Decompression

2) Transparency

3) Overstrike capability

حرارت و دستگاه های جوهر افشان است.

۱. سازمان کتاب

مطالب این کتاب بیشتر حول مطالبی که در بخش ۳۰۱ تشریح شد، سازماندهی شده است. کتاب حاوی نه فصل است که به سه بخش اصلی تقسیم می شوند: (۱) زمینه، (۲) پیش پردازش، و (۳) تحلیل. در سه فصل اول مباحث پایه که برای فهم فضول بعدی مورد نیاز است، مطرح می شود. در فصل ۲ مبانی دریافت بصری، مطرح می شود و مباحثی درباره تفکیک تصویر، روابط هندسی پایه بین پیکسل ها، مبانی نظری هندسه تصویر پردازی و چند مفهوم مقدماتی مرتبط با خواص فیلم عکاسی بیان می شود. در فصل ۳ به تبدیل های گوناگون تصویر بویژه تبدیل فوریه و چند خاصیت آن پرداخت شده است. تبدیل ها، ابزارهای مهمی است که موارد استفاده زیادی در پردازش تصویر دارد.

در فصول ۴، ۵ و ۶ فتوئی از پردازش تصویر معرفی شده است. در فصل ۴، فتوئون ارتقای تصویر از کاهش نویز و ارتقای تمایز گرفته تا تیز کردن و پردازش رنگی مطرح می شود. در فصل ۵ به تفصیل به روش های بازیابی تصویر که نقش مهمی در ترمیم تصاویر آسیب دیده با خرابی هایی چون مات شدگی دارد، پرداخته شده است. در فصل ۶ مبحث مهم «فشرده سازی داده ها» توضیح داده شده است. همان طور که در بخش ۱. ۴ ذکر شد، تصاویر رقمی حاوی مقادیر زیادی داده است که این موضوع انگیزه ای قوی برای جستجوی روش های «کاهش داده» می باشد. روش های مژوی در فصل ۶ نمونه هایی از روش های موجود برای این منظور است.

در فصول ۷، ۸ و ۹ فتوئون مناسب برای تحلیل تصویر ارائه می شود. در فصل ۷ به بخش بندی که اولین مرحله در تحلیل تصویر خودکار است، پرداخته شده است. همان طور که در بخش ۱. ۳ بیان شد، بخش بندی فرآیندی است که تصویر را به قسمت های اصلی سازنده اش تقسیم می کند. در فصل ۸ نمایش و توصیف اجزایی بخش بندی شده توضیح داده شده است. در این فصل به موضوعات متنوعی، از توصیف گرهای ساده (نظیر گشتاورها و نمایه ها^۱) گرفته تا توصیف گرهای پیچیده تر مبتنی بر شکل شناسی تصویر پرداخته شده است. در نهایت در فصل ۹ «تشخیص و تعبیر تصویر» که مراحل نهایی «تحلیل تصویر» است، ارائه شده است. همچنین در این فصل، فتوئون متعددی برای تشخیص تصویر از

1) Signature

قواعد معمولی تصمیم آماری گرفته تاروش های جدیدتر میتوانی بر شبکه های عصبی شرح داده شده است. روش های تعبیر ارائه شده در بخش های آخر فصل ۹ براساس سامانه های خبره و سایر روابط مدل سازی تصویر است.

دو پیوست نیز در انتهای کتاب ضمیمه شده است. در پیوست الف شرحی از فنون نیم تن کردن ارائه شده است که می توان از آنها به عنوان مبنای برای تهیه برنامه های چاپ بازنگ های خاکستری استفاده کرد. اگر سامانه پردازش و نمایش تصویر در دسترس نباشد، می توان با استفاده از چنین برنامه هایی نتایج پردازش تصویر را با چاپگر خطی معمولی روی کاغذ چاپ کرد. در پیوست ب مجموعه ای از تصاویر رقمه تفکیک پایین رمز گذاری شده آورده شده است که می توان از آنها به همراه یک برنامه چاپ به عنوان ابزاری اصلی برای آزمودن روش های پردازش تصویر ارائه شده در کتاب استفاده نمود.

مراجع

مراجعی که در اینجا ذکر می شود، عمومی بوده و در آنها انواع روش های پردازش تصویر موجود و کاربردهایشان، بیان می شود. اما مراجعی که در انتهای هر یک از فصول بعدی کتاب آورده می شود، به مطالب مطرح شده در همان فصل مربوط است. تمام مراجع با مؤلف، نام کتاب یا نشریه و سال چاپ آن ذکر می شود. کتاب شناسی انتهای کتاب نیز حاوی تمام اطلاعات مورد نیاز برای مراجعه به هر یک از مراجع می باشد.

بعضی از نشریات مهم که مقالاتی درباره 'پردازش تصویر و موضوعات مرتبط چاپ می کنند، عبارتند

از:
Computer Vision, Graphics, and Image Processing, IEEE Transactions on Systems,
Man and Cybernetics, IEEE Transactions on Pattern Analysis and Machine Intelligence,
Pattern Recognition, IEEE Transactions on Medical Imaging, Journal of the Optical Society of America, IEEE Transactions on Information Theory, IEEE Transactions on Communications, IEEE Transactions on Acoustics, Speech and Signal Processing, Proceedings of the IEEE, Pattern Recognition Letters,

و شماره های *IEEE Transactions on Computers* در قبل از سال ۱۹۸۰.

کتبی که برای تکمیل مبحث پردازش تصویر پیشنهاد می شود، به ترتیب سال چاپ عبارتند از:

Andrews, H. C., *Computer Techniques in Image Processing*, Academic Press, New York, 1970.

۲۱ ۵.۱ سازمان کتاب

- Lipkin, S., and Rosenfeld, R., *Picture Processing and Psychopictorics*, Academic Press, New York, 1970.
- Duda, R. O., and Hart, P. E., *Pattern Classification and Scene Analysis*, Wiley-Interscience, New York, 1973.
- Young, T. C. and Calvert, T. N., *Classification, Estimation, and Pattern Recognition*, American Elsevier Publishing Co., New York, 1974.
- Tou, J. T., and Gonzalez, R. C., *Pattern Recognition Principles*, Addison-Wesley, Reading, Mass., 1974.
- Andrews, H. C., and Hunt, B. R., *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, N.J., 1977.
- Pavlidis, T., *Structural Pattern Recognition*, Springer-Verlag, New York, 1977.
- Pratt, W. K., *Digital Image Processing*, Wiley-Interscience, New York, 1978 (Second ed., 1991).
- Gonzalez, R. C., and Thomason, M. G., *Syntactic Pattern Recognition: An Introduction*, Addison-Wesley, Reading, Mass., 1978.
- Hall, E. L., *Computer Image Processing and Recognition*, Academic Press, New York, 1979.
- Castleman, K. R., *Digital Image Processing*, Prentice-Hall, Englewood Cliffs, N.J., 1979.
- Duff, M. J. B., and Levialdi, S., *Languages and Architectures for Image Processing*, Academic Press, New York, 1981.
- Fu, K. S., *Syntactic Pattern Recognition and Applications*, Prentice Hall, Englewood Cliffs, N.J., 1982.
- Nevatia, R., *Machine Perception*, Prentice-Hall, Englewood Cliffs, N.J., 1982.
- Pavlidis, T., *Algorithms for Graphics and Image Processing*, Computer Science Press, Rockville, Md., 1982.
- Ballard, D. H., and Brown, C. M., *Computer Vision*, Prentice-Hall, Englewood Cliffs, N.J., 1982.
- Rosenfeld, R., and Kak, A. C., *Digital Picture Processing*, 2nd ed., vols. 1 & 2, Academic Press, New York, 1982.
- Levine, M. D., *Vision in Man and Machine*, McGraw-Hill, New York, 1985.
- Dougherty, E. R. and Giardina, C. R., *Matrix Structured Image Processing*, Prentice Hall, Englewood Cliffs, N.J., 1987.
- Jain, A. K., *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, N.J., 1989.
- Schalkoff, R. J., *Digital Image Processing and Computer Vision*, John Wiley & Sons, New York, 1989.

فصل ۲

مبانی تصویر رقمی

«کسانی که آرزوی موفقیت دارند باید سؤالات
مقدماتی دقیقی بپرسند»

ارسطور

هدف این فصل معرفی چند مفهوم مربوط به تصاویر رقمی و بعضی از نمادگذاریهای استفاده شده در سراسر کتاب است. بخش اول به اختصار طرز کار سامانه بینایی انسان، از جمله تشکیل تصویر در چشم و توانایی چشم در تشخیص و تطابق روشنایی را بیان می کند. بخش ۲. ۲ مدلی تصویری بر اساس پدیده روشنایی- انعکاس^۱ ارائه می کند که باعث دریافت تصاویر در فعالیتهای عادی بینایی می شود. بخش ۲. ۳ مفاهیم نمونه برداری یکنواخت از تصویر و چندی کردن سطوح روشنایی تصویر را معرفی می کند. بخش ۲. ۴ درباره روابط بین پیکسلها نظری اتصال و انواع اندازه گیری فاصله آنها می باشد که به طور وسیع در فصول بعدی استفاده می شوند. بخش ۲. ۵ شرح مفصلی در مورد هندسه تصویر برداری و عنایون مربوطه دارد. در نهایت بخش ۲. ۶ شامل مقدمه ای در مورد فیلم عکاسی و بعضی از مهمترین مشخصات آن بر حسب ضبط نتایج پردازش تصویر می باشد.

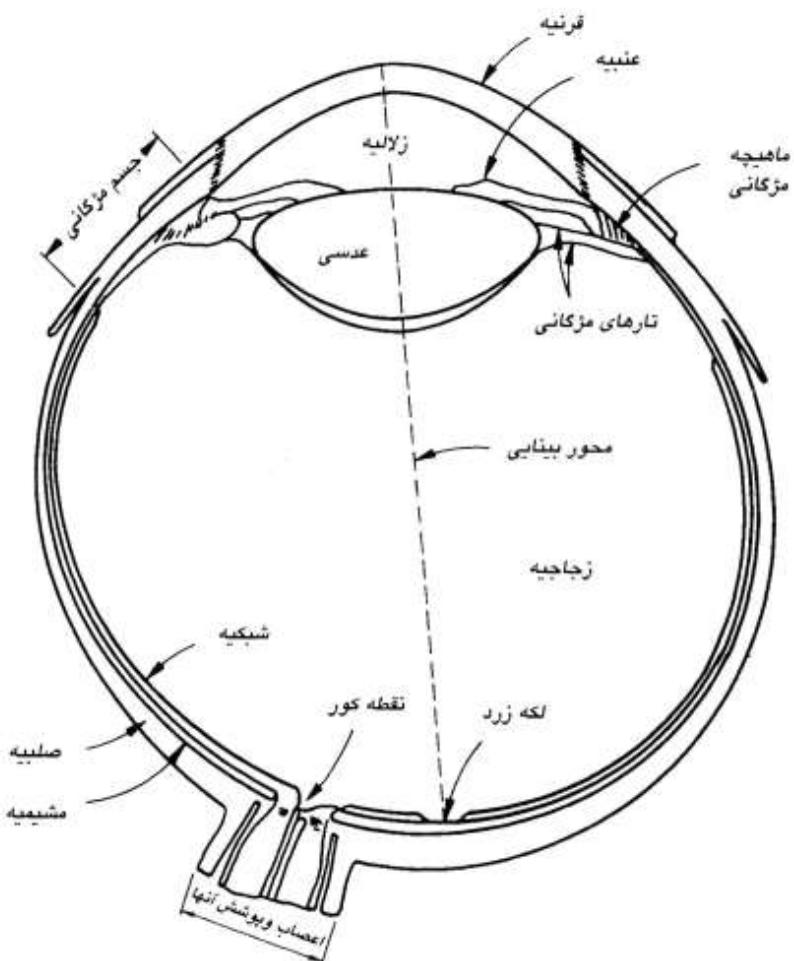
۱.۲ عناصر دریافت بصری

هدف نهایی بسیاری از فنون بیان شده در فصول بعدی کمک به بیننده در تفسیر محتویات تصویر است. بنابراین قبل از هر کار، ایجاد شناختی پایه ای از فرآیند دریافت بصری لازم است. در ادامه بررسی مختصراً در مورد طرز کار سامانه بینایی انسان با تأکید خاص بر مفاهیمی که بسیاری از موضوعات فصول بعد را زمینه سازی می کنند، خواهد آمد.

1) Illumination- reflection

۱۰.۱.۲ ساختهای چشم انسان

شکل ۱۰.۲ مقطعی افقی از چشم انسان را نشان می دهد. چشم تقریباً کره ای با قطر متوسط حدود ۲۰ میلی متر است. سه غشای قرنیه^۱ و صلیبه^۲ (برونی ترین لایه)، مشیمیه^۳ و شبکیه^۴ چشم انسان را احاطه کرده اند. قرنیه غشایی سفت و شفاف است که سطح داخلی چشم را می پوشاند. در بیرونی ترین لایه و در ادامه قرنیه، صلیبه قرار داد که غشای ماتی بوده و باقیمانده کره نوری را احاطه می کند.



شکل ۱۰.۲ نمودار ساده شده مقطع عرضی چشم انسان

1) Cornea

2) Sclera

3) Choroid

4) Retina

مشیمیه مستقیماً زیر صلیه قرار دارد. این غشا دارای شبکه ای از رگهای خونی است که نقش منبع تغذیه اصلی چشم را بر عهده دارند. پوشش مشیمیه کاملاً پررنگ است، بنابراین مقدار نور خارجی که وارد چشم می شود و نیز پراکندگی ناشی از کره نوری را کاهش می دهد. در قسمت جلویی، مشیمیه به جسم مژگانی^۱ و پرده عنیبه^۲ تقسیم می شود که دومی جهت کنترل مقدار نور ورودی به چشم باز و بسته می شود. قطر سوراخ مرکزی عنیبه (مردمک^۳) بین ۲ تا ۸ میلی متر تغییر می کند. قسمت رنگی چشم در جلوی عنیبه قرار دارد و عقب عنیبه سیاه رنگ است.

عدسی^۴ از لایه های هم مرکز از جنس سلولهای لنفی ساخته شده است و توسط رشته هایی که به جسم مژگانی متصل هستند، حرکت می کند. ۶۰ تا ۷۰ درصد عدسی آب و حدود ۶ درصد آن پوی (چربی) می باشد و نسبت به سایر قسمت های چشم پرتوین بیشتری دارد. عدسی رنگ زرد خفیفی دارد که با افزایش عمر زردرت می شود. عدسی تقریباً ۸ درصد طیف نور قابل رؤیت را جذب می کند که میزان جذب در طول موجه های کوتاهتر بیشتر است. نورهای مادون قرمز و ماورای بنفش با مقدار مجاز توسط پرتوین های درون عدسی جذب می شوند اما در مقادیر زیادتر ممکن است به چشم زیان برسانند.

داخلی ترین لایه چشم شبکیه است که تمام بخش پشتی دیواره داخلی چشم را می پوشاند. وقتی چشم روی شبکیه متعرکر شود، نور شبیه خارجی روی شبکیه می افتد. بینایی با توزیع گیرنده های نوری گسته در سطح شبکیه ممکن می شود. دو گروه از گیرنده ها وجود دارند: مخروطها^۵ و میله ها^۶. مخروط ها در هر چشم بین ۶ تا ۷ میلیون عدد هستند. آنها روی قسمت مرکزی شبکیه که لکه زرد^۷ نامیده می شود، قرار دارند و بسیار حساس به رنگ هستند. چشم انسان می تواند با استفاده از این مخروط ها تا حد زیادی جزئیات ریز را تشخیص دهد زیرا هر یک از آنها به یک پایانه عصبی متصل است. ماهیچه هایی که چشم را کنترل می کنند، تخم چشم را می چرخانند تا تصویر شی^۸ مورد نظر روی لکه زرد بیفتند. بینایی مخروطی را بینایی رنگی^۹ یا بینایی در نور روشن می نامند.

تعداد میله ها بسیار بیشتر است: بین ۷۵ تا ۱۵۰ میلیون میله در سطح شبکیه پخش شده اند. به علت ناحیه پخش وسیعتر و چون هر چند تا میله تنها به یک پایانه عصبی متصل شده اند، مقدار جزئیات قابل تشخیص توسط این نوع گیرنده ها کم است. کار میله ها تولید تصویری کلی و کامل از میدان دید

1) Ciliary body

2) Iris diaphragm

3) Pupil

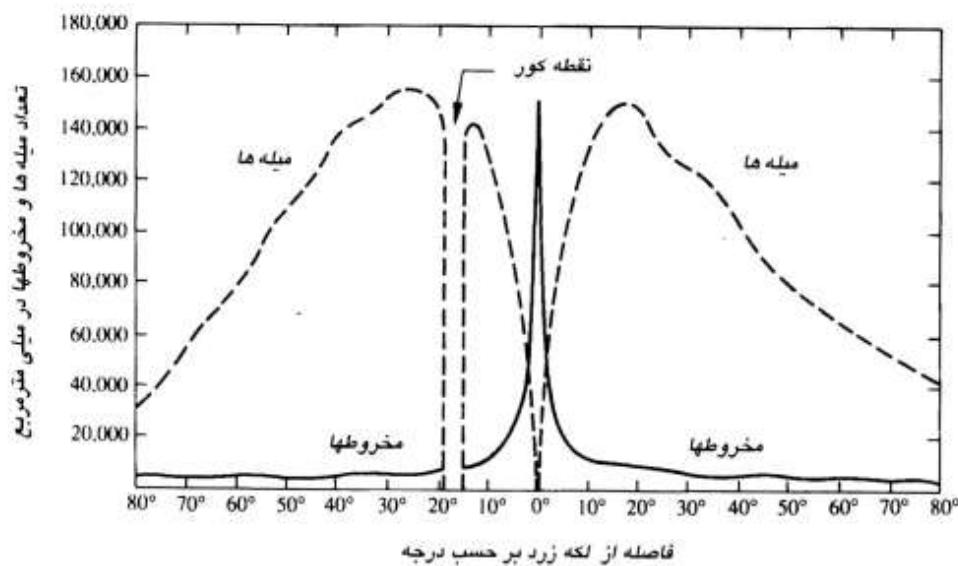
4) Lens

5) Cones

6) Rods

7) Fovea

8) Photopic vision



شکل ۲۰. توزیع میله ها و مخروطها در شبکیه (برگرفته از Graham ۱۹۶۵) [۱].

است. آنها در بینایی رنگی دخالتی ندارند و به سطوح پایین روشنایی حساس هستند. مثلاً اشیایی که در نور رنگی دیده می شوند، وقتی در نور ماه مشاهده شوند، مانند اشکال بی رنگ دیده می شوند، چرا که در نور ماه تنها میله ها تحریک می شوند. این پدیده را بینایی غیررنگی^۱ یا بینایی تاریکی می نامند.

شکل ۲. چگالی میله ها و مخروطها را در مقطع عرضی چشم راست که از ناحیه بسیار مهم نوری اعصاب نوری می گذرد، نشان می دهد. عدم وجود گیرنده ها در این ناحیه باعث ایجاد نقطه کور^۲ می شود (به شکل ۲. ۱ نگاه کنید). به جز این ناحیه، توزیع گیرنده ها حول لکه زرد متعارن شعاعی است. چگالی گیرنده ها بر حسب فاصله^۳ زاویه ای از لکه زرد سنجیده می شود (یعنی زاویه انحراف از محور که برابر است با زاویه بین محور بینایی و خطی که از مرکز عدسی می گذرد و با شبکیه برخورد می کند). در شکل ۲. به این نکته که مخروط ها در مرکز شبکیه (لکه زرد) چگالی بیشتری دارند، توجه کنید. همچنین توجه شود که چگالی میله ها از مرکز به سمت بیرون تا تقریباً ۲۰ درجه انحراف از محور افزایش و سپس تا حد جانبی شبکیه کاهش می یابد.

خود لکه زردیک تورنگی دایروی در شبکیه با قطر حدود ۱.۵mm است، گرچه از لحاظ توضیحات

1) Scotopic vision

2) Blind spot

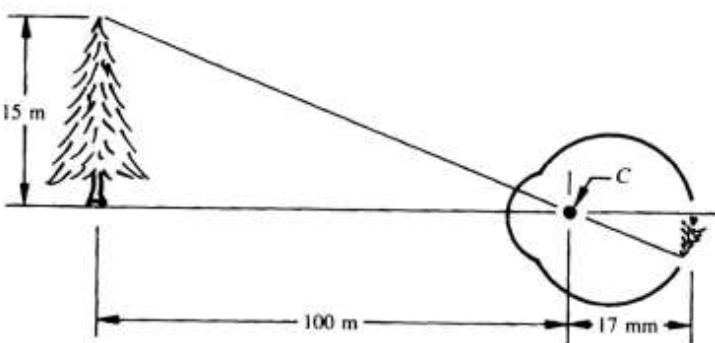
۱۰.۲ عناصر دریافت بصری

بعدی، صحبت درباره 'آرایه های مربعی' یا مثلى از عناصر حسگر مفیدتر است. بنابراین با قبول مقداری آزادی در تعبیر، لکه زرد را به عنوان یک آرایه مربعی از حسگرها با ابعاد $1/5 \times 1/5$ میلی متر در نظر می گیریم. در این ناحیه از شبکه چگالی مخروطها، حدوداً ۱۵۰۰۰۰ عنصر در میلی متر مربع است. بر اساس این تقریب، تعداد مخروط ها در ناحیه ای از چشم که بیشترین حساسیت^۱ را دارد، حدود ۳۳۷۰۰۰ عصر است. از نظر قدرت تفکیک، یک تراشه تصویربردار CCD با تفکیک متوسط می تواند این تعداد عناصر را در آرایه گیرنده ای که بزرگتر از 7×7 میلی متر نباشد، جای دهد. با این که توان بشر در داشتن هوشیاری و تجربه همراه با بینایی، این نوع مقایسه را مشکل می سازد، اما توجه داشته باشید که توان پایه ای چشم در تفکیک جزئیات، یقیناً در حدود حسگر های تصویربردار الکترونیکی فعلی است.

۲۰.۲ تشكيل تصوير در چشم

تفاوت اصلی بین عدسی چشم و یک عدسی نوری متداول این است که اولی انعطاف پذیر است. همانطور که در شکل ۱.۱ نشان داده شد، شعاع انحنای سطح جلویی عدسی چشم بزرگتر از شعاع انحنای سطح پشتی آن است. شکل عدسی با کشش تارهای جسم مژگانی کترل می شود. برای تمرکز روی اجسام دور، ماهیچه های کترل کننده باعث می شوند که عدسی نسبتاً تخت شود. به طور مشابه برای تمرکز چشم روی اشیای نزدیک این ماهیچه ها عدسی را ضخیم تر می کنند.

هنگامی که قدرت انکسار عدسی از مقدار حداقل به حداقل افزایش می یابد، فاصله بین مرکز کانونی عدسی و شبکه تقریباً از ۱۷ میلی متر تا حدود ۱۴ میلی متر تغییر می کند. وقتی چشم روی شیئی دورتر از ۳ متر تمرکز کند، عدسی کمترین قدرت انکسار و وقتی چشم روی یک جسم نزدیک تمرکز کند، بیشترین قدرت انکسار را دارد. این اطلاعات محاسبه اندازه تصویر شبکه ای شیئ را آسان می سازد. مثلاً در شکل ۲.۱ بینده به یک درخت با بلندی ۱۵ متر که در فاصله صدمتری قرار دارد، نگاه می کند. اگر λ اندازه تصویر شبکه ای برحسب میلی متر باشد، هندسه شکل ۲.۳، رابطه $x = 15/100 = 17/100$ را می دهد یعنی $x = 2.55\text{mm}$. همان طور که در بخش ۱.۱ بیان شد، تصویر شبکه ای روی لکه زرد منعکس می شود. آنگاه با تحریک نسبی گیرنده های نوری، این گیرنده ها انرژی تابشی را به تکانه های^۲ کوتاه الکتریکی تبدیل می کنند. در نهایت با رمزبرداری از این تکانه ها بوسیله 'مغز'، تصویر دریافت می شود.



شکل ۲ نمایش نوری چشمی که به یک درخت نگاه می کند. نقطه C مرکز نوری عدسی است

۳.۱۰.۲ تطابق و جداسازی روشنایی

چون تصاویر رقمه به شکل مجموعه گسته ای از نقاط روشنایی نمایش داده می شوند، توانایی چشم در جداسازی بین سطوح روشنایی مختلف، نکته مهمی در نمایش نتایج پردازش تصویر است. محدوده سطوح شدت نور که سامانه بینایی انسان می تواند با آن تطابق نماید، بی شمار است، یعنی در حدود 10^{10} سطح از آستانه غیررنگی تا حد زندگی. مشاهدات تجربی قابل ملاحظه ای نشان می دهد که روشنایی ذهنی (روشنایی که توسط سامانه بینایی انسان دریافت می شود) تابعی لگاریتمی از شدت نور تابشی به چشم است. شکل ۲ . ۴، نمودار شدت نور بر حسب روشنایی ذهنی، این مشخصه را نشان می دهد. منحنی ضخیم طولانی محدوده شدتهایی که سامانه بینایی انسان می تواند در آن تطابق کند، را نشان می دهد. در بینایی رنگی^۱، این محدوده تنها حدود 10^6 سطح است. آن طور که شاخه های دو گانه منحنی تطابق در این محدوده نشان می دهند، تغییر از بینایی غیررنگی^۲ به رنگی^۱ به تدریج و در محدوده تقریبی 10^0 تا 10^{-1} میلی لامبرت (۳-۱۰ میلی لامبرت در مقیاس لگاریتمی) رخ می دهد.

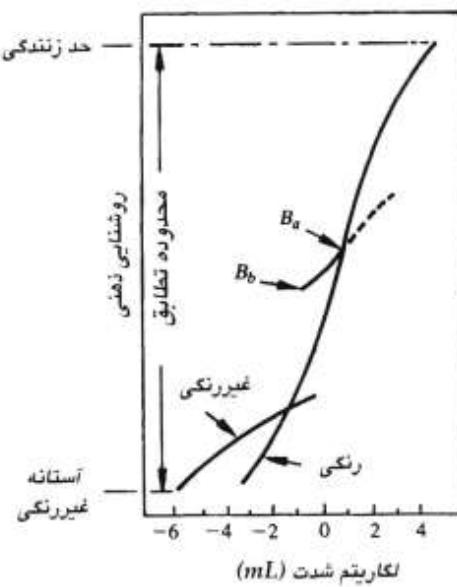
نکته اساسی در تعبیر محدوده پویای مؤثر^۳ ترسیم شده در شکل ۲ . ۴ این است که سامانه بینایی نمی تواند در یک زمان در تمام این محدوده کار کند. در عوض، تغییرات بزرگ را با تغییر در حساسیت کلی انجام می دهد که این پدیده تطابق روشنایی نامیده می شود. محدوده سطوح شدت که چشم در یک زمان می تواند از هم تشخیص دهد، در مقایسه با محدوده تطابق کل کوچک است. برای هر مجموعه ای

1) Photopic

2) Scotopic

3) Impressive

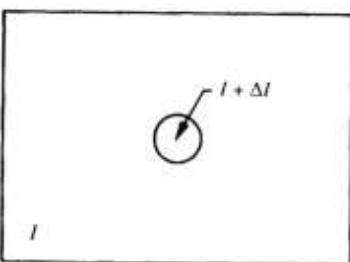
۱۰۲ عناصر دریافت بصری ۳۹



شکل ۲۰ محدوده احساس ذهنی روشنایی که سطح تطابق ویژه‌ای را نشان می‌دهد

از شرایط داده شده، سطح حساسیت فعلی سامانه بینایی، سطح تطابق روشنایی نامیده می‌شود که به عنوان مثال ممکن است روشنایی B در شکل ۲ . ۴ باشد. منحنی کوتاه متقاطع، محدوده روشنایی ذهنی که چشم هنگام تطابق با این سطح می‌تواند درک کند، را بیان می‌کند. این محدوده نسبتاً کوچک است، طوری که در سطح B و زیر آن، تمام محرك‌ها به عنوان رنگهای سیاه غیرقابل جداسازی هستند. بخش بالایی منحنی (خط چین) عملاً محدود نیست اما اگر خیلی گسترش یابد، بی معنی می‌شود زیرا شدت‌های خیلی بالاتر، سطح تطابق را نیز به بالاتر از B می‌برند.

توانایی چشم در جداسازی تغییرات روشنایی در هر سطح تطابق معین بسیار مورد توجه است. یک آزمایش کلاسیک که برای تعیین توان سامانه بینایی انسان در جداسازی روشنایی استفاده می‌شود، این است که فرد به ناحیه‌ای تخت با نور یکنواخت و به اندازه کافی بزرگ که تمام میدان دید را اشغال نماید، نگاه عمیقی داشته باشد. این ناحیه، معمولاً یک پخش کننده، نظیر شیشه مات، است که از طرف پشت توسط منبع نوری که شدتی، I ، قابل تغییر است، روشن می‌شود. همان طور که شکل ۲ . ۵ نشان می‌دهد، به این میدان افزایش روشنایی ΔI ، که به صورت چشمک کوتاه مدت دایره‌ای شکل در مرکز میدان با روشنایی یکنواخت است، افزوده می‌شود.

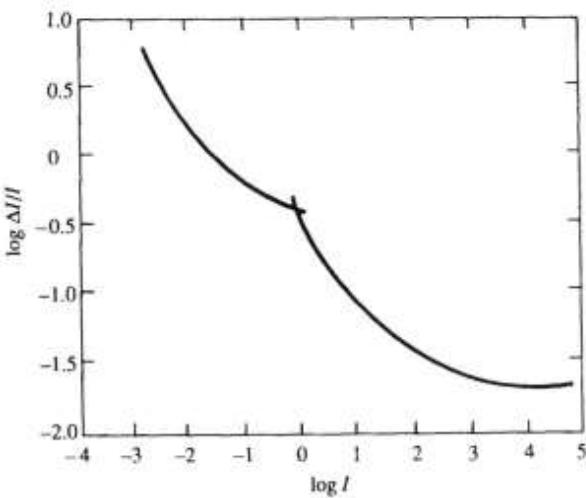


شکل ۲.۵ زمینه لازم برای آزمایش جداسازی روشنایی

اگر ΔI به اندازه کافی روشن نباشد، شخص جواب منفی می دهد که دلالت بر عدم تغییر قابل احساس می باشد. وقتی ΔI بزرگتر می شود، شخص ممکن است پاسخ مثبت «آری» بدهد که دلالت بر تغییر محسوس می باشد. در پایان وقتی ΔI به اندازه کافی بزرگ شود، شخص همواره پاسخ «آری» می دهد. مقدار $\Delta I_c / I$, که افزایش روشنایی قابل تشخیص در 5° درصد زمانها با روشنایی زمینه I است، نسبت وبر^{۱)} نامیده می شود. کوچک بودن $\Delta I_c / I$ دلالت بر این دارد که درصد کوچکی از تغییر شدت نیز قابل تشخیص است که نشانه جداسازی "خوب" روشنایی است. بر عکس بزرگ بودن مقدار $\Delta I_c / I$ بدین معنی است که درصد بزرگی از تغییر شدت برای تشخیص مورد نیاز است که دلالت بر جداسازی "ضعیف" روشنایی دارد.

در شکل ۲.۶ ترسیمی نوعی از نمودار لگاریتم $\Delta I_c / I$ به صورت تابعی از I log نشان داده شده است. این منحنی نشان می دهد که جداسازی روشنایی در سطوح پایین روشنایی ضعیف است (نسبت وبر بزرگ است) و وقتی روشنایی زمینه افزایش یابد، تا حد زیادی بهبود می یابد (نسبت وبر کاهش می یابد). وجود انشعاب در منحنی، نشان دهنده این واقعیت است که سطوح پایین بینایی با فعالیت میله ها انجام می شود، در حالی که سطوح بالای بینایی (که جداسازی بهتری دارند) وظیفه مخروط ها است. اگر روشنایی زمینه ثابت نگه داشته شود و شدت سایر منابع، به جای چشمک زدن، به طور تدریجی از حالت «کاملاً غیرقابل درک» تا حالت «کاملاً قابل درک» تغییر کند، بینته عادی می تواند جمعاً یک تا ۲۴ تغییر شدت مختلف را تشخیص دهد. این نتیجه مربوط به تعداد شدت های مختلفی است که تقریباً فرد می تواند در یک نقطه دلخواه از تصویر تکریگ بینند. این نتیجه به آن معنی نیست که می توان تصویر را با

1) Weber ratio



شکل ۲ . ۶ نسبت وبر متداول به صورت تابعی از شدت (گرفته شده از [Graham ۱۹۶۵])

این تعداد کم از مقادیر شدت بیان نمود، زیرا وقتی چشم در اطراف تصویر گردش کند، متوسط زمینه تغییر می کند و بنابراین امکان می دهد که مجموعه متفاوتی از تغییرات افزایشی در هر سطح تعابق جدید تشخیص داده شود. نتیجه نهایی این است که چشم قادر است تفکیک روشنایی کلی را در محدوده بسیار وسیع تری انجام دهد. در واقع در بخش ۳۰۲ نشان می دهیم که اگر تعداد سطوح شدت به کار رفته در نمایش یک تصویر تکرنگ به یک تا ۲۴ سطح محدود باشد، چشم کاملاً قادر به تشخیص اثرات آزاردهنده منحنی سازی^۱ است.

دو پدیده وجود دارد که به وضوح روشن می کند که روشنایی دریافت شده تابع ساده ای از شدت نیست. اوین پدیده این است که سامانه بینایی در مرز نواحی با شدت متفاوت زیرجهش^۲ یا فوق جهش^۳ می نماید. شکل ۲ . ۷ (الف) مثال بر جسته ای از این پدیده را نشان می دهد. گرچه شدت نوارها ثابت است، عملایک الگوی روشنایی را که شدیداً کنگره دار^۴ است، مخصوصاً در نزدیک مرزها، دریافت می کنیم. شکل ۲ . ۷ (ب) الگوی نواری مک^۵ نامیده می شود، که به نام ارنست مک^۶، اوین کسی که در سال ۱۸۶۵ این پدیده را بیان کرد، می باشد. این شکل توزیع واقعی شدت را نشان می دهد، اما الگوی

1) Contouring

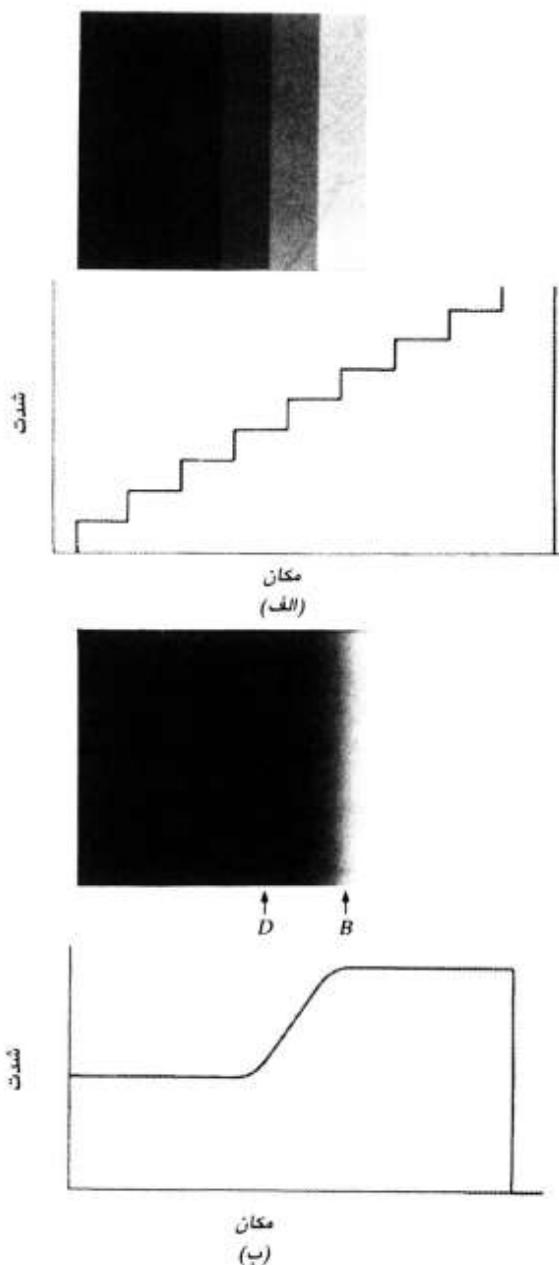
2) Undershoot

3) Overshoot

4) Scalloped

5) Mach band pattern

6) Ernst Mach



شکل ۷.۲ دو مثال که نشان می‌دهند روش‌نایابی دریافت شده تابع ساده‌ای از شدت نیست. (گرفته شده از [۱۹۷۰] Cornsweet)

۲.۲ یک مدل ساده تصویر ۵۳

روشنایی دریافتی، نوار تاریک تری در ناحیه D و نوار روشن تری در ناحیه B دارد.

پدیده دوم که تمایز همزمان^۱ نامیده می شود، همان طور که شکل ۲.۸ نشان می دهد، مربوط به این واقعیت است که روشنایی دریافتی یک ناحیه، ارتباط ساده ای باشدش ندارد. تمام مربعات وسطی شدت دقیقاً برابری دارند؛ هر چند وقتی زمینه روشن تر شود، برای چشم تبره تر به نظر می آیند. مثال واضح تر یک نکه کاغذ است که وقتی روی میز قرار دارد، سفید دیده می شود اما اگر هنگام نگاه مستقیم به آسمان از آن برای پوشاندن چشمها استفاده کنیم، کاملاً سیاه دیده می شود.

۲.۲ یک مدل ساده تصویر

عبارت تصویر به تابع دو بعدی شدت نور که به صورت $f(x, y)$ نوشته می شود، اشاره دارد که مقدار یا دامنه f در مختصات مکانی (x, y) شدت روشنایی تصویر در آن نقطه می باشد. چون نور صورتی از انرژی است، $f(x, y)$ باید بزرگتر از صفر و متناهی باشد، یعنی

$$0 < f(x, y) < \infty. \quad (1-2.2)$$

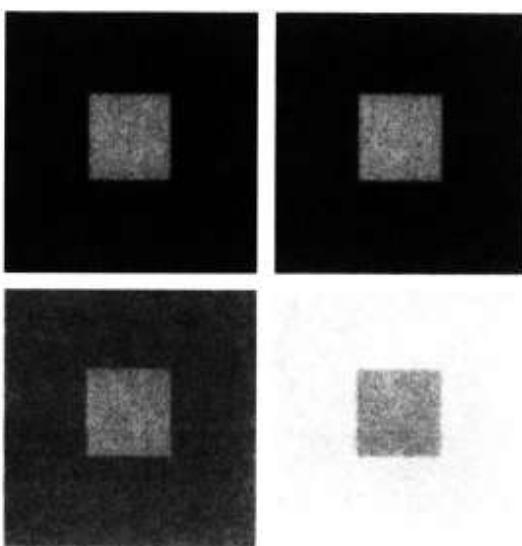
تصاویر دریافتی در فعالیت های روزانه معمولاً نور منعکس شده از اشیا است. طبیعت پایه ای $f(x, y)$ را می توان با دو عامل مشخص نمود: (۱) مقدار نور تابشی از منبع روی صحنه ای که دیده می شود و (۲) مقدار نور منعکس شده بوسیله اشیای صحنه. این دو عامل به ترتیب مؤلفه های روشنایی و انعکاس^۱ نامیده می شوند و به ترتیب با $i(x, y)$ و $r(x, y)$ نشان داده می شوند. توابع $i(x, y)$ و $r(x, y)$ به شکل حاصل ضرب ترکیب می شوند تا $f(x, y)$ ایجاد شود:

$$f(x, y) = i(x, y)r(x, y) \quad (2-2.2)$$

که

$$0 < i(x, y) < \infty \quad (3-2.2)$$

و



شکل ۲.۰۲ مثالی از تمایز همزمان. تمام مربعهای کوچک شدت دقیقاً برابر دارند، اما هرچه زمینه روشن تر شود، تپه تر دیده می شوند.

$$0 < r(x, y) < 1. \quad (4-2.2)$$

معادله (۴-۲.۲) بیان می کند که ضریب انعکاس به بازه 'عددی صفر(جذب کامل) و یک (انعکاس کامل) محدود می شود. طبیعت $(y, x)/z$ توسط منبع نور و $(x, y)/z$ توسط مشخصات اشیای صحنه معین می شود.

اعداد داده شده در معادلات (۲.۳-۲) و (۲.۴-۲) حدود نظری هستند. اعداد زیر بعضی محدوده های نامی $(u, v, x, y)/z$ را بیان می کنند. در یک روز صاف، خورشید می تواند بیش از ۹۰۰۰ فوت-شمع روشنایی روی سطح زمین تابش کند. این عدد به کمتر از ۱۰۰۰ فوت-شمع در یک روز ابری کاهش می یابد. در یک شب صاف، ماه کامل تقریباً ۱۰۰۰ فوت-شمع روشنایی دارد. روشنایی متدالو در یک دفتر تجاری حدود ۱۰۰ فوت-شمع است. همچنین چند مقدار متدالو $(u, v, x, y)/z$ عبارتند از: ۱۰۰۰ برای مخمل سیاه، ۶۵۰ برای استیل ضدزنگ، ۸۰۰ برای دیوار صاف بارنگ سفید، ۹۰۰ برای فلز با روکش نقره و ۹۳۰ برای برف.

در سراسر این کتاب، شدت تصویر تکرنگ f در مشخصات $(y, x)/z$ را سطح خاکستری (I) تصویر در

۳.۲ نمونه برداری و چندی کردن

۵۵

آن نقطه می نامیم. از معادلات (۲.۴-۲) تا (۲.۲-۲) واضح است که I در محدوده

$$L_{\min} \leq I \leq L_{\max}. \quad (5-2.2)$$

قرار دارد. از دید نظری تنها شرط این است که L_{\min} مشبّت و L_{\max} محدود باشد. در عمل $L_{\max} = i_{\max} r_{\max}$ و $L_{\min} = i_{\min} r_{\min}$. با استفاده از مقادیر قبلی روش‌نایابی و انعکاس به عنوان راه‌نمای، مقدار $L_{\max} = 100$ و $L_{\min} \approx 0.005$ در کارهای پردازش تصویر داخل اتاق مورد انتظار می باشد.

بازه $[L_{\min}, L_{\max}]$ محدوده خاکستری نامیده می شود. عموماً تلاش می شود که این بازه را به بازه $[0, L]$ که در آن $0 = I$ بیانگر سیاه و $L = I$ بیانگر سفید است، تغییر دهیم. تمام مقادیر میانی سایه‌های خاکستری هستند که به طور پیوسته از سیاه تا سفید تغییر می کنند.

۳.۲ نمونه برداری و چندی کردن^۱

۱.۳.۲ نمونه برداری و چندی کردن یکنواخت

تابع تصویر باید هم در مکان و هم در دامنه رقیق شود تا برای پردازش رایانه‌ای مناسب شود. رقیق‌سازی مختصات مکانی (y, x) را نمونه برداری تصویر^۲ و رقیق‌سازی دامنه را چندی کردن سطح خاکستری^۳ می نامند.

فرض کنید که تصویر پیوسته $(y, x) f(x, y)$ با نمونه های هم فاصله به شکل آرایه $M \times N$ ماتریس معادله زیر بیان شود که هر عنصر آرایه یک مقدار گستته است:

$$f(x, y) \approx \begin{bmatrix} f(0, 0) & f(0, 1) & \cdots & f(0, M-1) \\ f(1, 0) & f(1, 1) & \cdots & f(1, M-1) \\ \vdots & \vdots & & \vdots \\ f(N-1, 0) & f(N-1, 1) & \cdots & f(N-1, M-1) \end{bmatrix} \quad (1-3.2)$$

ماتریس سمت راست معادله فوق نشان دهنده یک تصویر رقیق است. همان طور که در بخش

۱. ۲. بیان شد، هر عنصر آرایه یک عنصر تصویر، عنصر عکس، پیکسل یا پل^۱ نامیده می شود.
عبارت تصویر و پیکسل ها در سراسر توضیحات بعدی به ترتیب دلالت بر تصویر رقمه و عناصر آن دارند.

بیان نمونه برداری و چندی کردن بر حسب عبارات ریاضی مفیدتر است. فرض کنید که Z و R به ترتیب بر مجموعه اعداد صحیع حقیقی و مجموعه اعداد حقیقی دلالت نمایند. فرآیند نمونه برداری را می توان تقسیم بندی صفحه xy به یک شبکه تصور کرد که مختصات مرکز هر شبکه یک زوج از عناصر حاصل از ضرب کارتزین $Z \times Z$ (که به صورت Z^2 نیز نوشته می شود) می باشد؛ Z^2 مجموعه تمام زوجهای مرتب (a, b) می باشد که a و b اعداد صحیحی از Z هستند. بنابراین اگر (x, y) اعداد صحیحی از $Z \times Z$ باشند و r تابعی باشد که به هر زوج مختصات (x, y) یک مقدار سطح خاکستری (یعنی، یک عدد حقیقی از مجموعه اعداد حقیقی، R) را نسبت می دهد، $(x, y) f(x, y)$ تصویر رقمه است. این انتساب در واقع همان فرآیند چندی کردن که قبلآ بیان شد، می باشد. همچنین اگر سطوح خاکستری اعداد صحیحی باشند (همان طور که معمولاً در این فصل و فصول بعدی این چنین است)، Z جایگزین R می شود و بدین ترتیب تصویر رقمه تابعی دو بعدی می شود که مختصات و مقادیر عناصرش صحیح هستند.

فرآیند رقمه سازی، نیازمند تصمیم گیری در مورد مقادیر N ، M و تعداد سطوح خاکستری مجاز برای هر پیکسل است. در پردازش تصویر رقمه سعی می شود که این مقادیر توانهای صحیحی از دو باشند؛ یعنی،

$$N = 2^n, \quad M = 2^k \quad (2-3. 2)$$

و

$$G = 2^m \quad (3-3. 2)$$

که G دلالت بر تعداد سطوح خاکستری دارد. در این بخش فرض می شود که سطوح گسته هم فاصله و در بازه 'خاکستری' بین ۰ و L باشند. تعداد بیت های مورد نیاز برای ذخیره تصویر رقمه، b ، با استفاده از معادلات (۲-۳. ۲) و (۳-۳. ۲) بدست می آید:

$$b = N \times M \times m. \quad (4-3. 2)$$

1) Pixel or pel

۳.۲ نمونه برداری و چندی کردن ۵۷

جدول ۲.۱ تعداد بیتهاي انباره ذخیره سازی تصویر برای مقادیر مختلف N و m

$N \backslash m$	1	2	3	4	5	6	7	8
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1,024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608

اگر $M = N$ باشد،

$$b = N^2m. \quad (5-3.2)$$

مثالاً تصویر 128×128 با 64×64 سطح خاکستری نیازمند 4^{98304} بیت انباره حافظه است. جدول

۱. مقادیر b حاصل از معادله (۵-۳.۲) را برای بعضی محدوده های نامی N و m نشان می دهد.

جدول ۲.۲ تعداد بایتهاي ۸ بیتی متناظر را می دهد.

چون معادله (۵-۳.۱) تقریبی از تصویر پیوسته است، سؤال منطقی در اینجا این است که: چند

نمونه و چند سطح خاکستری برای یک تقریب خوب نیاز است؟ تفکیک (درجه جزئیات قابل تشخیص) تصویر بشدت به این پارامترها بستگی دارد. هرچه این پارامترها افزایش یابند، آرایه رقمه تقریب نزدیکتری با تصویر اولیه است. با این حال معادله (۴-۳.۴) آشکارا به واقعیت نامطلوبی اشاره می کند که طبق آن

میزان حافظه و در نتیجه حجم پردازش مورد نیاز با افزایش N ، M و m به سرعت افزایش می یابد.

جدول ۲.۳ تعداد بایتهاي ۸ بیتی انباره ذخیره سازی تصویر برای مقادیر مختلف N و m

$N \backslash m$	1	2	3	4	5	6	7	8
32	128	256	512	512	1,024	1,024	1,024	1,024
64	512	1,024	2,048	2,048	4,096	4,096	4,096	4,096
128	2,048	4,096	8,192	8,192	16,384	16,384	16,384	16,384
256	8,192	16,384	32,768	32,768	65,536	65,536	65,536	65,536
512	32,768	65,536	131,072	131,072	262,144	262,144	262,144	262,144
1,024	131,072	262,144	393,216	524,288	655,360	786,432	917,504	1,048,576

در ادامه توضیحات قبلی - و طبق رسم متدالو بالفرض تصاویر مریعی - اجزاء دهید اکنون اثر تغییرات N و m را بر کیفیت تصویر بررسی کنیم. تعریف تصویر "خوب" مشکل است، زیرا کیفیت تصویر نه تنها بسیار ذهنی است، بلکه بسیار وابسته به نیازهای کاربرد مورد نظر است. این مسئله را در ارتباط با نمونه برداری تصویر در بخش ۳.۹. و در زمینه فشرده سازی داده های تصویر با تفصیل بسیار بیشتر در فصل ۶ بررسی می کنیم. در هر حال اکنون تنها علاقه مند به ایجاد ایده ای کلی این موارد هستیم که چگونه با کاهش تفکیک مکانی و چندی کردن سطوح خاکستری، کیفیت تصویر رقمه پایین می آید.

شکل ۲.۹(الف) تصویر رقمه 256×256 سطوح 1024×1024 از یک گل راشان می دهد. اشکال ۲.۹(ب) - (ج) نتایج کاهش تفکیک مکانی از $N=1024$ به ترتیب به N های 512 ، 256 ، 128 ، 64 و 32 را نشان می دهند. در تمام حالات حداقل تعداد مجاز سطوح خاکستری 256 است. چون وسعت صفحه نمایش مورد نیاز ثابت و میدان نمایش 1024×1024 است، پیکسل های تصویر با تفکیک پایین تر دو یا چندبار تکرار می شوند تا تمام صفحه نمایش را پر کنند. تکرار پیکسل ها^۱ اثر شطرنجی را به وجود می آورد که به ویژه در تصاویر با تفکیک کم قابل رویت است.

شکل ۲.۹(الف) را با تصویر 512×512 شکل ۲.۹(ب) مقایسه کرده و توجه کنید که عملاً تشخیص این دو تصویر از هم عملاً غیرممکن است، اگر دو عکس اولیه در این مقیاس را با هم مقایسه کنیم، تقریباً افزایش غیرمحسوسی در دانه ای بودن^۲ و کاهش کمی را در تیزی تصویر 512×512 ، بخصوص با حرکت به سمت مرکز غنچه، می بینیم. عموماً این نوع از جزئیات در بیشتر فرآیندهای چاپ از دست می رود و بسته به اندازه^۳ نسبی اشیا تشخیص آنها حتی در نمایش ها و عکس های اولیه نیز اغلب مشکل است. برای نمونه، بزرگ کردن شکل ۲.۹(ب) تکرار پیکسل هارا، که برای تولید این تصویر به کاررفته است، قابل رویت می کند. تصویر 256×256 طرح شطرنجی ریزی را در لبه ها و دانه ای بودن مشخص تری را در سراسر تصویر نشان می دهد. این اثر در تصویر 128×128 بسیار قابل رویت تر است و در تصویر 64×64 و 32×32 کاملاً مشخص می شود.

شکل ۲.۱۰ اثراتی را که با کاهش تعداد بیت های استفاده شده در نمایش تعداد سطوح خاکستری تصویر تولید می شوند، نشان می دهد. شکل ۲.۱۰(الف) همان تصویر 1024×1024 ، 128×128 بیتی استفاده شده در توضیح قبلی را نشان می دهد. اشکال ۲.۱۰(ب) - (ج) با کاهش تعداد بیت ها از $m=7$ به $m=1$

1) Pixel replication

2) graininess

۵۹

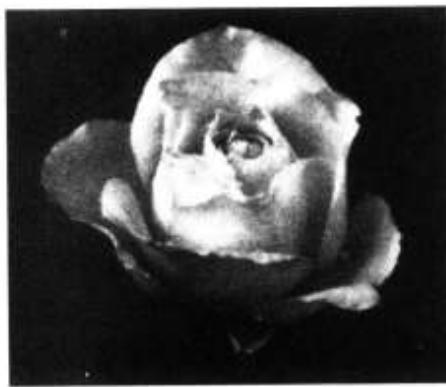
۳.۲ نمونه برداری و چندی کردن



(الف)



(ب)



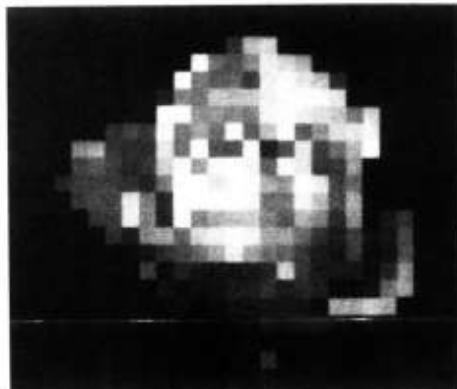
(پ)



(ت)

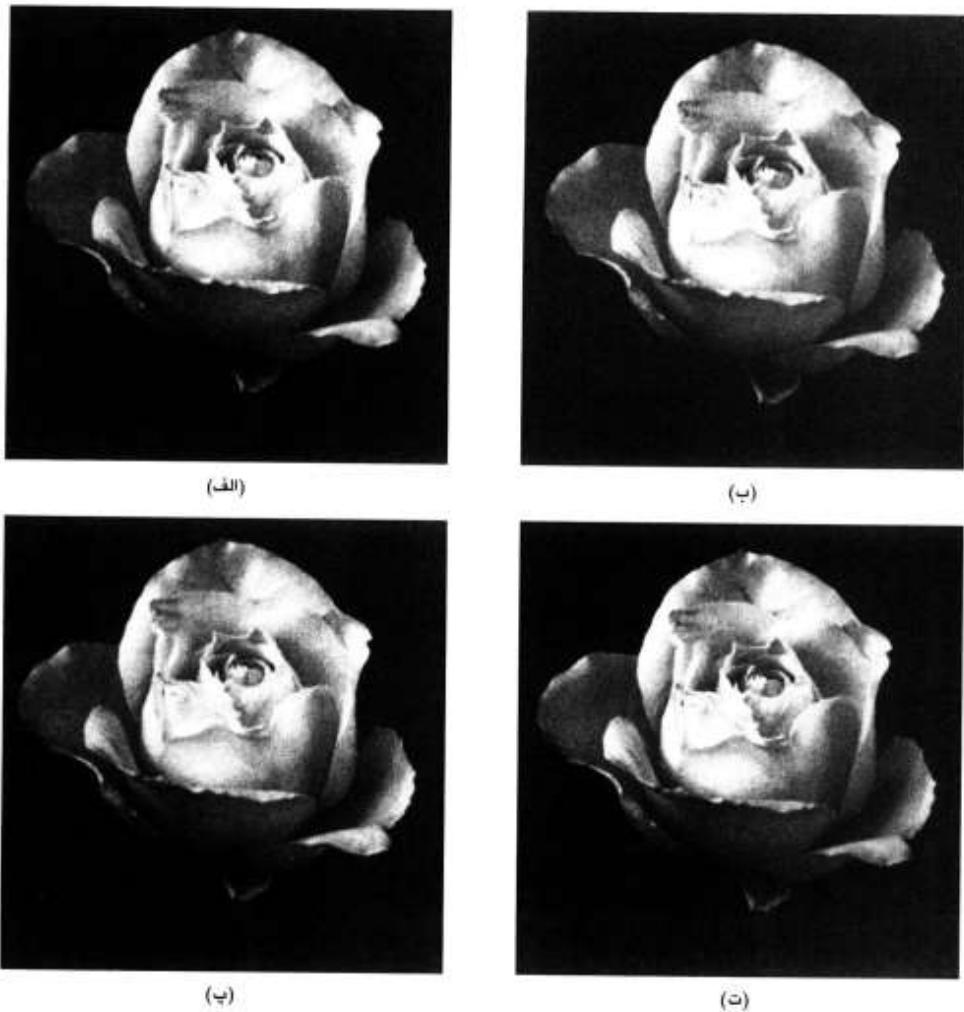


(ث)



(ج)

شکل ۲ اثر کاهش تفکیک مکانی

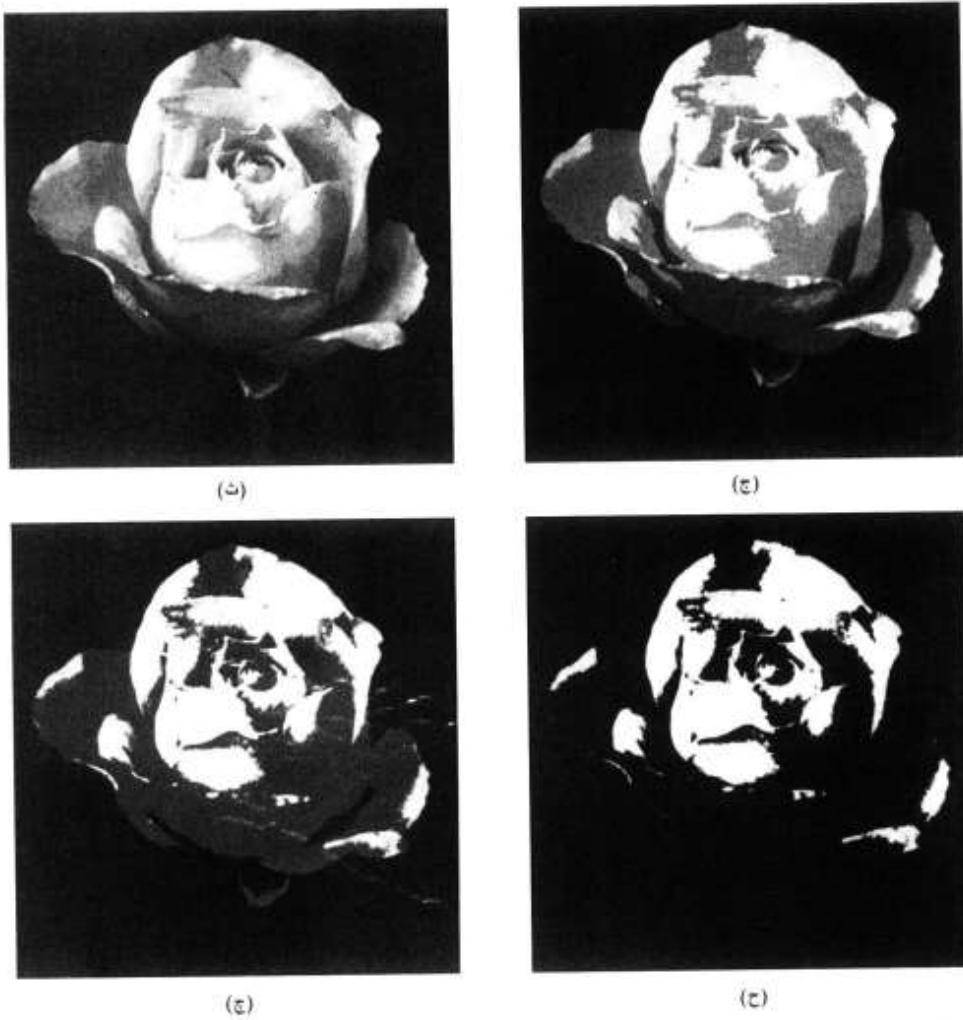


شکل ۲ ۱۰. ۲ یک تصویر 1014×1024 که به ترتیب با $256, 128, 64, 32, 16, 8, 4$ و ۲ سطح نمایش داده شده است

(به معادله ۲-۳ نگاه کنید) بدست آمدند، در حالی که تفکیک مکانی ثابت و برابر 1024×1024 پیکسل است. در تمام موارد عملی تصاویر $256, 128, 64$ و ۲ سطحی از نظرینیانی انسان یکسان هستند. تصویر ۳۲ سطحی شکل ۲. (ت) تقریباً مجموعه غیرقابل ادراکی از ساختارهای تیز بسیار ریز را در نواحی با سطوح خاکستری آرام ایجاد کرده است. این اثر که ناشی از کاربرد تعداد سطوح خاکستری ناکافی در نواحی آرام تصویر رقمه است، منحنی سازی اشتباه^۱ نامیده می شود. همان طور که

1) false contouring

۳۰۲ نمونه برداری و چندی کردن ۶۱



شکل ۲.۰۲ ادامه

در اشکال ۲.۰۲(ا)-(د) بوضوح دیده می شود، معمولاً این اثر در تصاویری که با استفاده از ۱۶ سطح روشنابی هم فاصله یا کمتر نمایش داده شوند، کاملاً قابل رویت است.

نتایج قبلی، اثراتی را که با تغییر مستقل N و m در کیفیت تصویر ایجاد می شود، نشان دادند. با این وجود این نتایج تنها به جزئی از سؤال مطروحه پاسخ می دهند، زیرا هنوز چیزی در مورد رابطه بین این پارامترها نگفته اند. Huang [۱۹۶۵] با آزمایش این مسأله را بررسی کرد تا اثرات حاصل از تغییر N و m بر کیفیت تصویر را به طور تجربی برآورد نماید. آزمایش وی شامل مجموعه ای از آزمایشات فرعی بود.



شکل ۲ ۱۱۰ تصاویر آزمون مورد استفاده در ارزیابی کیفیت ذهنی تصویر (از ۱۹۶۵) Huang

در شکل ۲ ۱۱۰ سه تا از تصاویر مورد استفاده در آزمایش نشان داده شده اند. صورت زن تصویری با جزئیات نسبتاً کم است؛ تصویر مرد عکاس دارای مقدار متوسطی از جزئیات است، و تصویر جمعیت در مقایسه با دو تصویر قبل جزئیات زیادی دارد.

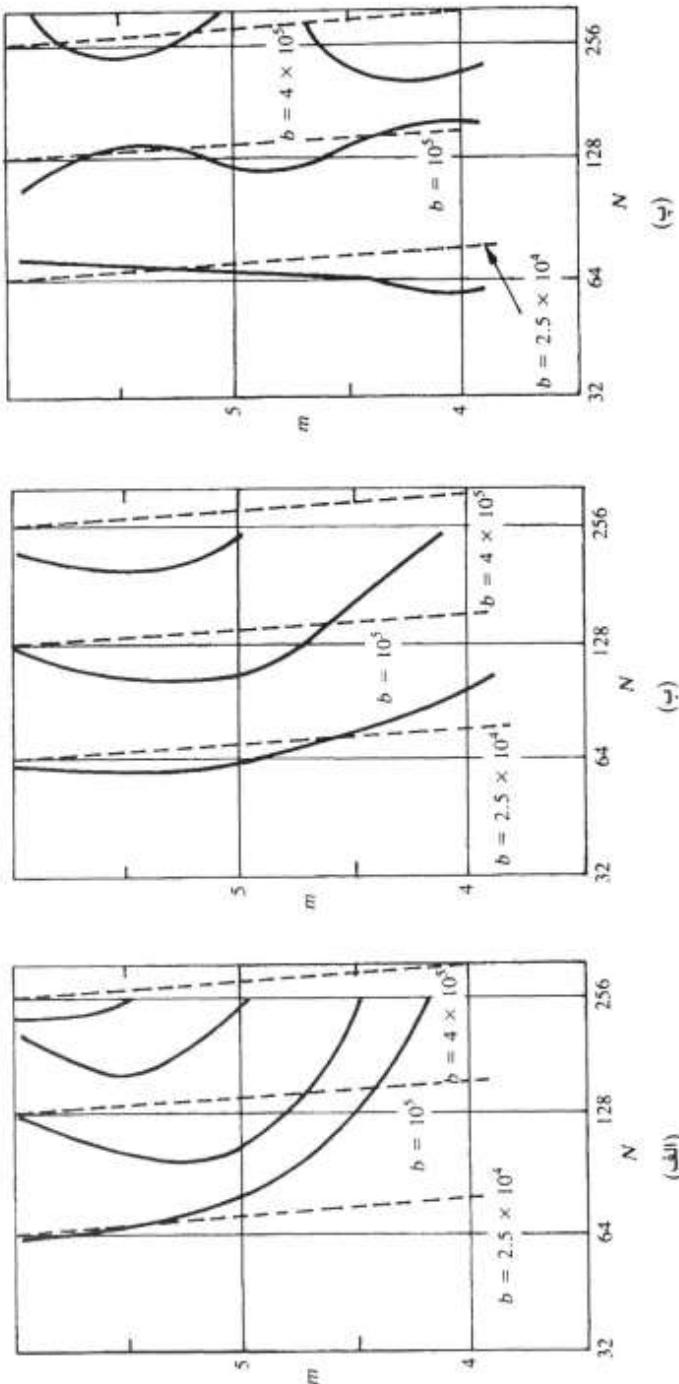
با تنظیر N و m مجموعه هایی از این سه تصویر تولید شدند و سپس از افراد خواسته شد تا آنها را بر حسب کیفیت ذهنی^۱ مرتب نمایند. نتایج آزمایش در شکل ۲ ۱۲۰ به صورت منحنی های هم ترجیحی^۲ در صفحه Nm ارائه شده اند. هر نقطه در این صفحه بیانگر تصویری است که مقادیر N و m آن مختصات

1) Subjective quality

2) Isopreference curves

۶۳

۳۰.۲ نمونه برداری و چنایی کردن



شکل ۳۰.۲۱ منحنی های هم زنجیری برای (الف) صورت انسان، (ب) مرد عکاس و (ب) جمعیت انسان (از Huang [54]).

نقطه برابر است. نقاطی که روی یک منحنی هم ترجیحی قرار دارند، متناظر با تصاویر با کیفیت ذهنی برابر هستند.

منحنی های هم ترجیحی شکل ۲.۱۲ از چپ به راست به ترتیب افزایش کیفیت ذهنی مرتب شده اند.

این نتایج تجربی بیان می کنند که: (۱) همان طور که انتظار می رفت، کیفیت تصاویر با افزایش N و m افزایش می باید. در تعداد کمی از حالات، بازای N ثابت، با کاهش m کیفیت بهبود می باید. محتمل ترین دلیل برای این نتیجه این است که کاهش m معمولاً تمايز قابل مشاهده تصویر را افزایش می دهد. (۲) با افزایش جزئیات، منحنی های عمودی تر می شوند. این بدان معناست که برای تصاویر با جزئیات زیاد ترها تعداد محدودی سطح روشنایی مورد نیاز است. مثلاً شکل ۲.۱۲(ب) نشان می دهد که برای N های برابر با ۶۴ یا ۱۲۸، کیفیت تصویر با افزایش m بهبود نمی باید. این نتیجه برای منحنی های دوشکل دیگر درست نیست. (۳) منحنی های هم ترجیحی به طور محسوسی از منحنی های با b ثابت (معادله ۲.۵-۳) را بیینید که در شکل ۲.۱۲ با خطوط نشان داده شده اند، جدا می شوند.

۲.۳.۴ نمونه برداری و چندی کردن غیربکنواخت

برای تفکیک مکانی ثابت، در بسیاری حالات می توان ظاهر تصویر را با استفاده از روشی وفقی که در آن فرآیند نمونه برداری به مشخصات تصویر بستگی دارد، بهبود داد. به طور کلی در نزدیک تغییرات تند روشنایی، نمونه برداری نزدیک^۱ مورد نیاز است؛ اما در نواحی نسبتاً آرام می توان از نمونه برداری دور^۲ نیز استفاده کرد. مثلاً تصویر صورت انسان روی زمینه یکنواخت، را در نظر بگیرید. به وضوح، زمینه اطلاعات کمی دارد و می توان آن را به طور کامل و کافی با نمونه برداری دور بیان کرد؛ اما صورت انسان جزئیات نسبتاً بیشتری دارد. اگر نمونه های اضافی استفاده نشده در زمینه در این ناحیه از تصویر استفاده شوند، نتیجه کلی بهبود خواهد یافت. هنگام توزیع نمونه ها، باید در مرزهای انتقال سطح خاکستری، نظیر مرز بین صورت انسان و زمینه در این مثال، از چگالی نمونه بیشتری استفاده شود.

نیاز به شناسایی مرزها، حتی به طور تقریبی، به یقین یکی از مشکلات روش نمونه برداری غیر یکنواخت است. همچنین این روش در تصاویری که نواحی یکنواخت نسبتاً کوچکی دارند، عملی نیست.

1) Fine

2) Coarse

٤٠.٢ چند رابطه پایه ای بین پیکسلها

به عنوان نمونه، اجرای نمونه برداری غیریکتواخت برای تصویر جمعیت متراکم مشکل خواهد بود. وقتی تعداد سطوح خاکستری باید کم باشد، معمولاً استفاده از سطوح با فاصله نامساوی در فرآیند چندی کردن مطلوب است. از روشی مشابه با نمونه برداری غیریکتواخت که قبلاً بیان شد، می‌توان در توزیع سطوح خاکستری تصویر نیز استفاده کرد. در هر حال چون چشم در تخمین سایه‌های خاکستری نزدیک تغییرات سریع سطح روشنایی ضعیف است، در نزدیکی مرزها تعداد کمی از سطوح روشنایی استفاده می‌شوند. پس سایر سطوح را می‌توان در نواحی دیگری که تغییرات سطح خاکستری آرام است، استفاده کرد و بدین صورت منحنی‌های اشتباه^۱ را که اغلب در صورت چندی کردن بسیار درشت^۲ در نواحی آرام بوجود می‌آیند، حذف کرد یا کاهش داد.

این روش نیازمند مشاهده قبلی تصویر برای آشکارسازی مرز و جزئیات است. یک روش جایگزین که بوسیله برای توزیع سطوح خاکستری مورد توجه است، شامل محاسبه فراوانی رخداد هریک از سطوح مجاز است. اگر شدتهاخاکستری از بازه^۳ معنی به طور فراوان رخ داده ولی سایر شدتها به ندرت رخ دهند، سطوح چندی کردن در آن بازه نزدیک هم و خارج از آن بازه دور از هم خواهد بود. این روش گاهی چندی کردن باریک شونده^۴ نامیده می‌شود. در فصل ۶ این مفاهیم بیشتر بیان می‌شوند.

٤٠.٣ چند رابطه پایه ای بین پیکسلها

در این بخش چند رابطه اولیه اما مهم بین پیکسل‌های تصویر رقمی را مطرح می‌کنیم. همان‌طور که قبلاً ذکر شد، تصویر با نماد $(y, x)f$ نشان داده می‌شود. برای اشاره به یک پیکسل معین از حروف کوچک نظیر p و q استفاده می‌کنیم. یک زیرمجموعه از پیکسل‌های $(y, x)f$ با S بیان می‌شود.

٤٠.٤ همسایه‌های یک پیکسل

پیکسل p در مختصات (y, x) چهار همسایه افقی و عمودی دارد که مختصات آنها عبارتند از:

$$(x + 1, y), (x - 1, y), (x, y + 1), (x, y - 1).$$

این مجموعه از پیکسل‌ها همسایه‌های چهارگانه^۱ p نامیده می‌شوند که با نماد $(p)N_4$ نشان داده

1) False contour

2) Coarsely

3) Tapered quantization

۶۶ پردازش تصویر رقمه

می شوند. هریک از این پیکسل ها به فاصله یک واحد از (y, x) قرار دارند و اگر (y, x) در مرز تصویر باشد، بعضی از همسایگان p خارج از تصویر رقمه قرار می گیرند.

چهار همسایه قطری p مختصات

$$(x + 1, y + 1), (x + 1, y - 1), (x - 1, y + 1), (x - 1, y - 1)$$

را دارند که با $N_p(p)$ نشان داده می شوند. این نقاط همراه با همسایه های چهارگانه، همسایه های هشتگانه p نامیده می شوند که با $N_8(p)$ نشان داده می شوند. ماتند $(N_8(p), N_p(p), y, x)$ روی مرز تصویر باشد، بعضی از نقاط همسایگی های $N_p(p)$ و $N_8(p)$ خارج از تصویر قرار می گیرند.

۲.۴.۲ اتصال

اتصال V بین پیکسل ها مفهوم مهمی است که در تعیین مرز اشیا یا اجزای نواحی تصویر استفاده می شود. برای تعیین این که آیا دو پیکسل متصل هستند، باید تعیین شود که آیا آنها نوعی همسایگی با هم دارند (مثلًا آیا همسایه های چهارگانه هستند) و آیا سطوح خاکستری آنها معیار شاباهت مشخص شده ای را برآورده می کنند (مثلًا آیا برابر هستند). به عنوان نمونه در تصویر دودویی با مقادیر ۰ و ۱، ممکن است دو پیکسل، همسایه چهارگانه باشند، اما متصل گفته نمی شوند مگر این که مقادیر یکسانی داشته باشند. اجازه دهید V مجموعه مقادیر سطوح خاکستری مورد استفاده در تعریف اتصال باشد؛ مثلًا در تصویر دودویی، $\{1\} = V$ برای اتصال پیکسل های با مقدار یک است. در تصویر خاکستری، برای اتصال پیکسل های دارای محدوده ای از مقادیر شدت، مثلًا $\{32, 33, \dots, 64\} = V$ تعیین می شود. سه نوع اتصال را در نظر می گیریم:

(الف) اتصال چهارگانه: دو پیکسل p و q با مقادیری از مجموعه V ، متصل چهارگانه اند اگر q در مجموعه $(N_4(p))$ باشد.

(ب) اتصال هشتگانه: دو پیکسل p و q با مقادیری از مجموعه V ، متصل هشتگانه اند اگر q در مجموعه $(N_8(p))$ باشد.

(پ) اتصال m گانه (اتصال مرکب): دو پیکسل p و q با مقادیری از مجموعه V ، در صورتی متصل

۶۷ ۴.۲ چند رابطه پایه ای بین پیکسلها

m گانه اند که:

(i) q در $N_4(p)$ باشد، یا

(ii) q در $N_D(p)$ باشد و مجموعه $(q) \cap N_4(p) \cap N_D(p)$ تهی باشد. (این مجموعه شامل پیکسل هایی است که همسایه های چهار گانه p و q هستند و مقادیرشان از V است.)

اتصال مرکب گونه ای از اتصال هشتگانه است و برای حذف اتصالات چند مسیره که اغلب هنگام استفاده از اتصال هشتگانه رخ می دهد، به کار می رود. مثلاً آرایش پیکسلی نشان داده شده در شکل ۲.۱۳(الف) را در نظر بگیرید: برای $\{V\} = 1$ مسیر های بین همسایه های هشتگانه پیکسل مرکزی در شکل ۲.۱۳(ب) با خط مقطع نشان داده شده اند. توجه شود که دو گانگی در مسیر اتصال ناشی از اتصال هشتگانه است. همان طور که در شکل ۲.۱۳(پ) دیده می شود، این ابهام (دو گانگی) با استفاده از اتصال m گانه رفع می شود.

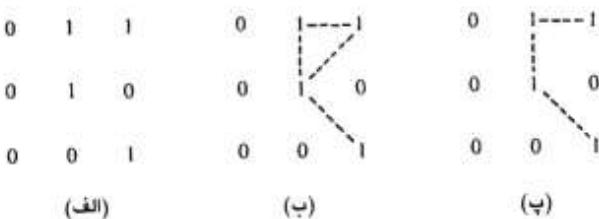
پیکسل p مجاور q است اگر آن دو متصل باشند. بسته به نوع اتصال مشخص شده، می توان مجاورت های ۴، ۸ یا m گانه را تعریف نمود. دو زیر مجموعه تصویری S_1 و S_2 از مجاور هستند اگر بعضی پیکسل های S_1 مجاور بعضی پیکسل های S_2 باشند.

یک مسیر از پیکسل p با مختصات (y, x) به پیکسل q با مختصات (t, s) دنباله ای از پیکسل های

جزا با مختصات

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

است که (x_i, y_i) و (x_{i+1}, y_{i+1}) مجاور (x_{i-1}, y_{i-1}) می باشد که



شکل ۲.۱۳ (الف) آرایش پیکسلها؛ (ب) همسایگان هشتگانه پیکسل مرکزی؛ (پ) همسایگان m گانه همان پیکسل. خطوط مقطع، مسیر بین آن پیکسل و همسایگانش هستند

$i \leq n$ و n طول مسیر است. می توان بسته به نوع مجاورت مشخص شده^۴، 8 ، 4 ، یا m مسیر را تعریف نمود.

پیکسل p به q در S متصل است اگر p و q هر دو عضو زیرمجموعه^۱ تصویری S باشند و مسیری که همه از پیکسلهای S باشند، بین p و q موجود باشد. برای هر پیکسل p در S ، مجموعه پیکسل هایی از S که به p متصل هستند، اجزای متصل S نامیده می شود. بنابراین هر دو پیکسل از یک جزء متصل به یکدیگر متصل هستند و جزء متصل مختلف، مجزا از هم هستند.

توانایی انتساب برچسبهای متفاوت به اجزای متصل گوناگون تصویر، اهمیتی اساسی در تحلیل تصویر خودکار دارد. در بخش بعدی روال ساده‌ای برای برچسب زنی ترتیبی اجزای متصل بیان می کنیم که در هر زمان روی دو سطر از تصویر دو دویی کار می کند. در بخش ۸.۴. روش دیگری را براساس شکل شناسی^۱ بیان می کنیم.

۳.۴.۲ برچسب زنی اجزای متصل

پیماش پیکسل به پیکسل تصویر از چپ به راست و از بالا به پایین را در نظر بگیرید و فعلاً فرض کنید که تنها به اجزای متصل چهارگانه علاقه مند هستیم. اجازه دهید در هر مرحله از فرآیند پیماش m بر پیکسل مرکزی و ω به ترتیب برهمسایگان بالایی و سمت چپ آن دلالت کند. طبیعت ترتیبی پیماش اطمینان می دهد که وقتی به p می رسیم، قبلاً با نقاط ω مواجه شده ایم (و اگر یک بوده اند، برچسب خورده اند). با توجه به مفاهیم قبل^۲ تعیین شده، اجازه دهید روند زیر را بکار گیریم. اگر مقدار m صفر باشد، بدون هیچ کاری روی محل بعدی پیماش بروید. اگر مقدار m یک باشد، ω را بررسی نمایید. اگر هر دو صفر هستند، یک برچسب جدید به p نسبت دهید. (یعنی براساس اطلاعات فعلی حاصل از پیماش، این اولین بار است که به این جزء پیوسته برخورد کرده ایم). اگر تنها یکی از دو همسایه فوق آباشد، برچسب آن را به p متنسب نمایید. اگر هر دو یک هستند و برچسب های متفاوتی دارند، یکی از این دو برچسب را به p متنسب کنید و یادداشت کنید که هر دو برچسب معادل هستند (یعنی نقاط ω از طریق p متصل هستند). در انتهای پیماش تمام نقاط با مقدار یک برچسب خورده اند، اما بعضی برچسبها ممکن است معادل باشند. کارهایی که اکنون باید انجام داد، به ترتیب عبارتند از: مرتب کردن تمام زوج برچسب های

۱) Morphology

٤٠.٢ چند رابطه پایه ای بین پیکسلها

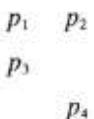
معادل در دسته های هم ارزی^۱ (بخش ٤٠.٢ را ببینید)، انتساب بر چسب به هر دسته، مرور مجدد تصویر با جایگزینی هر بر چسب با بر چسب متنسب شده به دسته هم ارزی آن.

برای بر چسب زنی اجزاء متصل هشتگانه نیز به همان طریق انجام می دهیم، اما دو پیکسل همسایه قطری بالای p که با q و r نشان داده می شوند، نیز باید بررسی شوند. طبیعت ترتیبی پیمایش اطمینان می دهد که این همسایه ها قبل از p ، پردازش شده اند. اگر p صفر باشد، به محل بعدی پیمایش بروید. اگر p یک و تمام چهار همسایه آن صفر باشد، یک بر چسب جدید به p متنسب کنید. اگر تنها یکی از همسایه های یک باشد، بر چسبش را به p متنسب کنید. اگر دو یا چند همسایه یک باشند، یکی از بر چسبها را به p متنسب کنید و هم ارزی های مناسب را یادداشت کنید. بعد از اتمام پیمایش تصویر زوج بر چسب های معادل را در دسته های هم ارزی مرتب کنید. به هر دسته یک بر چسب یکتا متنسب کنید و پیمایش دوم تصویر را انجام دهید که در این پیمایش هر بر چسب را با بر چسب متنسب شده به دسته هم ارزی آن جایگزین کنید.

٤٠.٣.٢ رابطه ها، هم ارزی و بسته بودن تراکذاری^۲

الگوریتم بر چسب زنی که در بخش قبل بیان شد، فایده ابزارهای رسمی برای کار با روابط و هم ارزی ها در پردازش پیکسل را بیان می کند. اجازه دهید به اختصار چند مفهوم مهم را که اساس چنین رابطه ها و هم ارزیهای هستند، مطرح کنیم.

رابطه دوتایی[†] R روی مجموعه A مجموعه ای از زوج عناصر A است. اگر زوج (a,b) در R باشد، اغلب به صورت aRb نمایش داده می شود که به معنای " a با b رابطه R را دارد" می باشد. مثلاً مجموعه نقاط $\{p_1, p_2, p_3, p_4\}$ را که به صورت زیر قرار دارند:



در نظر بگیرید و رابطه "اتصال چهارگانه" را تعریف کنید. در این حالت R مجموعه زوج نقطه هایی از A

1) Equivalence classes

2) Transitive closure

† در این مبحث، کلمه دوتایی به "دو" بر می گردد و هیچ ربطی با تصاویر دو دوی ندارد

است که متصل چهارگانه اند، یعنی $R = \{(p_1, p_2), (p_2, p_1), (p_1, p_3), (p_3, p_1)\}$ بنابراین p_1 با p_2 و p_3 رابطه دارد و بالعکس، اما p_4 با هیچ نقطه دیگری تحت رابطه اتصال چهارگانه مرتبط نیست.

رابطه^۱ دوتایی R روی مجموعه A ، گفته می شود که

(a) انعکاسی^۲ است اگر برای هر a از A ، aRa ؛

(b) متقارن^۳ است اگر برای هر a و b از A ، aRb نتیجه دهد که bRa ؛

(c) تراگذاری^۴ است اگر برای هر a و b و c از A ، aRb و bRc نتیجه دهنده باشد، aRc .

رابطه ای که این سه خاصیت را داشته باشد، رابطه هم ارزی^۵ نامیده می شود.

یک خاصیت مهم از روابط هم ارزی این است که اگر R رابطه هم ارزی روی مجموعه A باشد، آنگاه بازی بعضی k های بین یک و بی نهایت، می توان A را به تعداد k زیرمجموعه جدای از هم که دسته های هم ارزی^۶ نامیده می شوند، افزایش کرد. روش افزایش بدین صورت است که a و b هر دو در یک زیرمجموعه هستند، اگر و تنها اگر aRb باشد.

بیان یک رابطه با یک ماتریس دودویی مفید است. مثلاً رابطه^۷

$$R = \{(a, a), (a, b), (b, d), (d, b), (c, e)\}$$

با ماتریس

$$\mathbf{B} = \begin{bmatrix} a & b & c & d & e \\ a & 1 & 0 & 0 & 0 \\ b & 0 & 0 & 0 & 1 \\ c & 0 & 0 & 0 & 0 \\ d & 0 & 1 & 0 & 0 \\ e & 0 & 0 & 0 & 0 \end{bmatrix}$$

بیان می شود، که در محل متناظر بازوج های رابطه دار یک و در سایر محل ها صفر قرار داده می شود. اگر رابطه بررسی سؤال انعکاسی باشد، تمام اجزای قطر اصلی یک خواهد بود؛ اگر R متقارن باشد، \mathbf{B} نیز یک ماتریس متقارن خواهد بود.

همان طور که در بالا بیان شد، خاصیت تراگذاری موجب می شود که اگر aRb و aRc آنگاه bRc .

1) Reflexive

2) Symmetric

3) Transitive

4) Equivalence relation

5) Equivalence classes

۴.۲ چند رابطه پایه ای بین پکسلها ۷۱

درمثال اخیر، با a و b با d رابطه دارد زیرا (a, b) و (b, d) در R هستند. با این حال توجه می شود که (a, d) در مجموعه R نیست. مجموعه ای که دارای این روابط ضمنی باشد، بسته تراگذاری R نامیده شده و با R^+ نشان داده می شود که $R^+ = \{(a, a), (a, b), (a, d), (b, b), (b, d), (d, b), (d, d), (c, e)\}$ و وجود زوچهای (d, d) در مجموعه فوق، از تعریف تراگذاری نتیجه می شود (یعنی، اگر bRd و آنگاه aRb ، و نیز اگر bRb و dRd و dRb آنگاه dRb). در شکل ماتریسی داریم:

$$\mathbf{B}^+ = \begin{bmatrix} a & b & c & d & e \\ 1 & 1 & 0 & 1 & 0 \\ b & 0 & 1 & 0 & 1 \\ c & 0 & 0 & 0 & 0 \\ d & 0 & 1 & 0 & 1 \\ e & 0 & 0 & 0 & 0 \end{bmatrix}$$

که عناصر با مقدار یک، اعضای R^+ را که بسته تراگذاری است، معین می کنند. یک روند سرراست برای محاسبه ماتریس \mathbf{B}^+ از ماتریس داده شده \mathbf{B} به صورت زیر است:
فرض کنید \mathbf{B} یک ماتریس دودویی $n \times n$ باشد که رابطه R را روی الفبای n نمادی A نشان می دهد
ماتریس \mathbf{B}^+ را با رابطه

$$\mathbf{B}^+ = \mathbf{B} + \mathbf{BB} + \mathbf{BBB} + \dots + (\mathbf{B})^n \quad (1-4.2)$$

محاسبه می کنیم که $(\mathbf{B})^n = \mathbf{BBB} \dots \mathbf{B}$ (بار n). عناصر با مقدار یک از ماتریس \mathbf{B}^+ ، بیانگر R^+ ، بسته تراگذاری رابطه R ، می باشند (Gries [۱۹۷۱]). عملیات ماتریسی طبق روش معمول انجام می شوند، جز این که تمام ضربها با AND های منطقی و تمام جمع ها (از جمله آنهایی که در معادله ۱-۴.۱ نشان داده شده اند) با OR های منطقی جایگزین می شوند. عملیات معادله (۱-۴.۲) به ترتیب محاسبه \mathbf{B} ، $\mathbf{B}(\mathbf{B})$ ، $\mathbf{B}(\mathbf{B}\mathbf{B})$ و ... می باشد طوری که در هر مرحله تنها حاصل عملیات قبلی در \mathbf{B} ضرب می شود. به عنوان تمرین نشان دهید که معادله (۱-۴.۱) همان نتیجه را برای \mathbf{B}^+ ، مانند مثال فوق، می دهد.

پیاده سازی معادله (۱-۴.۱) باز محدود n^3 عمل AND و OR است. [Warshall ۱۹۶۲] روش

مفیدتری را به کاربرد که تنها نیازمند عمل OR روی عناصری از \mathbf{B} که مقدار یک دارند، بود. مراحل این

الگوریتم به ترتیب عبارتند از:

مرحله یک. قرار دهید $j = 1$.

مرحله دو. بازای $i = 1, 2, \dots, n$ و $k = 1, 2, \dots, n$ اگر $b(i, j) = 1$ آنگاه بازای $b(i, k) + b(j, k)$

$$.b(i, k) = b(i, k) + b(j, k)$$

مرحله سه. j را یک واحد افزایش دهید.

مرحله چهار. اگر $j \leq n$ باشد، به مرحله دو بروید؛ در غیر این صورت به مرحله ۵ بروید.

مرحله پنج. توقف کنید. مقدارنهایی به جای \mathbf{B}^+ است.

بررسی این موضوع آموزنده است که آیا انجام این فرآیند برای مثال قبلی در این بخش، همان نتیجه' معادله (۱-۴.۲) را می دهد. در عمل همانند روش ارائه شده در پایان بخش (۲-۴.۳) معمولاً فرض می شود که روابط هم ارزی هستند. که در این مورد ماتریس \mathbf{B} متقابران است و برای محاسبه' بسته بودن تراکنده ای قبل از کاربرد معادله' (۱-۴.۲) یا الگوریتم Warshall تمام عناصر قطر اصلی یک قرار داده می شوند. دسته های هم ارزی نمادهای مختلف الفبا که منجر به ماتریس \mathbf{B}^+ می شوند، را می توان با پیمایش این ماتریس از چپ به راست و از بالا به پایین معین نمود. مثلاً اگر در ردیف i و ستون j یک مواجه شویم، نماد مربوط به ستون j را برابر نماد مربوط به سطر i ام قرار می دهیم (آنها هم ارز هستند)، ستون j را صفر می کنیم و پیمایش ماتریس \mathbf{B}^+ را ادامه می دهیم.

۵.۴.۲ سنده های فاصله

برای پیکسلهای p ، q و z به ترتیب با مختصات (s, t) ، (x, y) و (u, v) ، D تابع فاصله یا متریک^۱ است اگر

. (الف) $D(p, q) = 0$ اگر $p = q$ باشد، $D(p, q) = 0$

(ب) $D(p, q) = D(q, p)$

(پ) $D(p, z) \leq D(p, q) + D(q, z)$

فاصله اقلیدسی بین p و q به صورت

۴.۲ چند رابطه پایه ای بین پیکسلها ۷۳

$$D_e(p, q) = [(x - s)^2 + (y - t)^2]^{1/2} \quad (2-4.2)$$

تعريف می شود. برای این سنجه فاصله، پیکسل هایی که فاصله آنها از (x, y) کمتر یا مساوی r است، نقاطی موجود روی قرص دواری به شعاع r با مرکز (x, y) هستند.

فاصله D_4 (که فاصله بلوک شهری^۱ هم نامیده می شود) بین p, q به صورت زیر تعریف می شود:

$$D_4(p, q) = |x - s| + |y - t| \quad (3-4.2)$$

در این حالت پیکسل هایی که فاصله D_4 آنها از (x, y) کمتر یا مساوی مقداری مانند r باشد، یک لوزی^۲ به مرکز (x, y) می سازند. مثلاً پیکسل هایی که فاصله D_4 آنها از نقطه (x, y) (نقطه مرکزی) کوچکتر یا مساوی 2 باشد، منحنی های^۳ فاصله^۴ ثابت مانند زیر می سازند:

$$\begin{matrix} & & 2 \\ & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ & 2 & 1 & 2 \\ & & 2 \end{matrix}$$

پیکسل های با $1 = D_4$ همسایه های چهارگانه (x, y) هستند.

فاصله D_8 (که فاصله شطرنجی^۵ هم نامیده می شود) بین p, q به صورت زیر تعریف می شود:

$$D_8(p, q) = \max(|x - s|, |y - t|). \quad (4-4.2)$$

در این حالت پیکسل هایی که فاصله D_8 آنها از (x, y) کمتر یا مساوی 2 باشد، مربعی به مرکز (x, y) می سازند. مثلاً پیکسل هایی که فاصله D_8 آنها از نقطه (x, y) (نقطه مرکزی) کوچکتر یا مساوی 2 باشد، منحنی های فاصله^۶ ثابت زیر را می سازند:

$$\begin{matrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{matrix}$$

پیکسل های با $1 = D_8$ همسایگان هشتگانه (x, y) هستند.

1) City-block distance

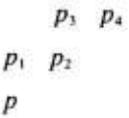
2) Diamond

3) Contour

4) Chessboard distance

۷۴ پردازش تصویر رقemi

فاصله D_4 بین دو نقطه p و q برابر طول کوتاهترین مسیر چهارگانه بین این دو نقطه است. در مورد فاصله D_8 نیز همین طور است. در واقع می توان فاصله های D_4 و D_8 بین p و q را بدون توجه به این که آیا مسیر متصلی بین آنها وجود دارد یا نه، بدست آورد؛ زیرا در تعریف این فاصله ها تنها مختصات این نقاط وارد می شوند. گرچه برای اتصال m گانه، مقدار فاصله (طول مسیر) بین دو پیکسل به مقادیر پیکسل ها و همسایگانشان در طول مسیر نیز بستگی دارد، برای نمونه، پیکسل ها با آرایش زیر را در نظر بگیرید و فرض کنید که p_1 و p_2 مقدار یک دارند و p_3 و p_4 می توانند مقادیر صفر یا یک را داشته باشند:



اگر تنها اتصال پیکسل ها با مقدار یک مجاز باشد و p_3 و p_4 صفر باشند، فاصله m گانه بین p_1 و p_2 ۲ است. اگر تنها یکی از p_3 یا p_4 یک باشد، فاصله ۳ است. اگر p_3 و p_4 هر دو یک باشند، فاصله ۴ است.

۶.۴.۲ اعمال حسابی/منطقی

در بسیاری از شاخه های پردازش تصویر از عملیات حسابی و منطقی بین پیکسل ها استفاده گستردۀ ای می شود. عملیات حسابی بین دو پیکسل p و q به صورت زیر نشان داده می شوند:

$p + q$	جمع:
$p - q$	تفريق:
$p * q$ (pq)	ضرب:
$p \div q$	تقسيم:

(همجنبن $p \times q$ و $p \div q$)

عملیات حسابی به صورت پیکسل به پیکسل روی کل تصویر انجام می شوند. کاربرد اصلی جمع تصویرها در متوسط گیری از تصویر برای کاهش نویز است. تفريق تصویرها ابزاری پایه ای در تصویربرداری پزشکی است، که برای حذف اطلاعات زمینه ثابت به کار می رود. یکی از استفاده های اصلی ضرب (یا تقسیم) تصویر، تصحیح سایه گرفتگی ناشی از غیریکنواختی روشنایی یا حسگر مورد استفاده در تصویربرداری است. عملیات حسابی در هر زمان تنها یک پیکسل را وارد محاسبات می کنند، بنابراین می توانند "درج" باشند.

۴.۲ چند رابطه پایه ای بین پیکسلها ۷۵

انجام شوند، بدین معنا که نتیجه انجام عملیات حسابی روی پیکسل (y, x) را می توان در همان محل ذخیره کرد، زیرا مجدداً به اطلاعات آن محل احتیاج نداریم.

عملیات منطقی اصلی مورد استفاده در پردازش تصویر AND، OR و متمم هستند که به صورت زیر نشان داده می شوند:

$$\begin{array}{ll} p \text{AND} q \quad (\text{همچنین } p \cdot q) & : \text{AND} \\ p \text{OR} q \quad (\text{همچنین } p + q) & : \text{OR} \\ \text{NOT} q \quad (\text{همچنین } \bar{q}) & : \text{متمم} \end{array}$$

چون از ترکیب این اعمال می توان هر عمل منطقی را پایه سازی کرد، این سه عمل روی هم کامل عملیاتی^۱ نامیده می شوند. اعمال منطقی تنها به تصویر دودویی اعمال می شوند، در حالی که اعمال حسابی تنها به پیکسل های چند مقداره اعمال می شوند. اعمال منطقی ابزارهای اصلی در پردازش تصویر دودویی هستند، که برای وظایفی نظیر نقاب زنی^۲، تشخیص ویژگی^۳ و تحلیل شکل^۴ استفاده می شوند. اعمال منطقی روی تصاویر کامل به صورت پیکسل به پیکسل انجام می شوند چون حاصل AND دو متغیر دودویی، "۱" است تنها اگر هر دو "۱" باشند، بنابراین نتیجه آن در هر محل تصویر "۱" است تنها اگر پیکسل های متناظر در دو تصویر "۱" باشند. چون عملیات منطقی در هر زمان تنها یک پیکسل را وارد محاسبات می کنند، پس می توانند مانند عملیات حسابی، "درجا"^۵ انجام شوند. شکل ۲۰.۲ مثالهای مختلفی از عملیات منطقی را نشان می دهد که سیاه دلالت بر "۱" و سفید دلالت بر "۰" می نماید. حاصل XOR، "۱" است تنها اگر یکی از دو پیکسل (اما نه هر دو) "۱" باشد، و در غیر این صورت "۰" است. حاصل این عمل با حاصل OR که "۱" است وقتی یکی از دو پیکسل یا هر دو "۱" باشند، متفاوت است.

علاوه بر پردازش پیکسل به پیکسل روی کل تصویر، عملیات حسابی و منطقی در اعمال همسایگی گرا^۶ نیز استفاده می شوند. پردازش همسایگی^۷ معمولاً با عنوان عملیات تقابی^۸ انجام می شود (همچنین در اغلب موارد عبارات الگو^۹، پنجره و فیلتر^{۱۰} به جای نقاب به کار می روند). نظریه پشت

1) Functionally complete

2) Masking

3) Feature detection

4) Shape analysis

5) In place

6) Neighborhood-oriented

7) Neighborhood processing

8) Mask operations

9) Template

10) Filter

اعمال نقابی این است که اجازه دهیم مقدار متنسب به یک پیکسل تابعی از سطح خاکستری آن پیکسل و همسایگانش باشد. به عنوان نمونه تصویر شکل ۲. ۱۵ (الف) را در نظر بگیرید و فرض کنید که می خواهیم مقدار \bar{z} را با متوسط مقادیر پیکسل ها در ناحیه ای 3×3 به مرکز پیکسل با مقدار \bar{z} جایگزین نماییم. برای انجام این کار باید یک عمل حسابی به صورت

$$\bar{z} = \frac{1}{9} (z_1 + z_2 + \dots + z_9) = \frac{1}{9} \sum_{i=1}^9 z_i$$

انجام شود و سپس مقدار \bar{z} به پیکسل با مقدار \bar{z} متنسب شود.

بارگذاری به نقاب نشان داده شده در شکل ۲. ۱۵ (ب) می توان با قرار دادن مرکز نقاب در \bar{z} ، ضرب هر پیکسل زیر نقاب در ضریب مربوطه و جمع حاصلضربها، عبارت کلی تری برای این عمل بدست آورده، یعنی

$$z = w_1 z_1 + w_2 z_2 + \dots + w_9 z_9 = \sum_{i=1}^9 w_i z_i \quad (5-4. 2)$$

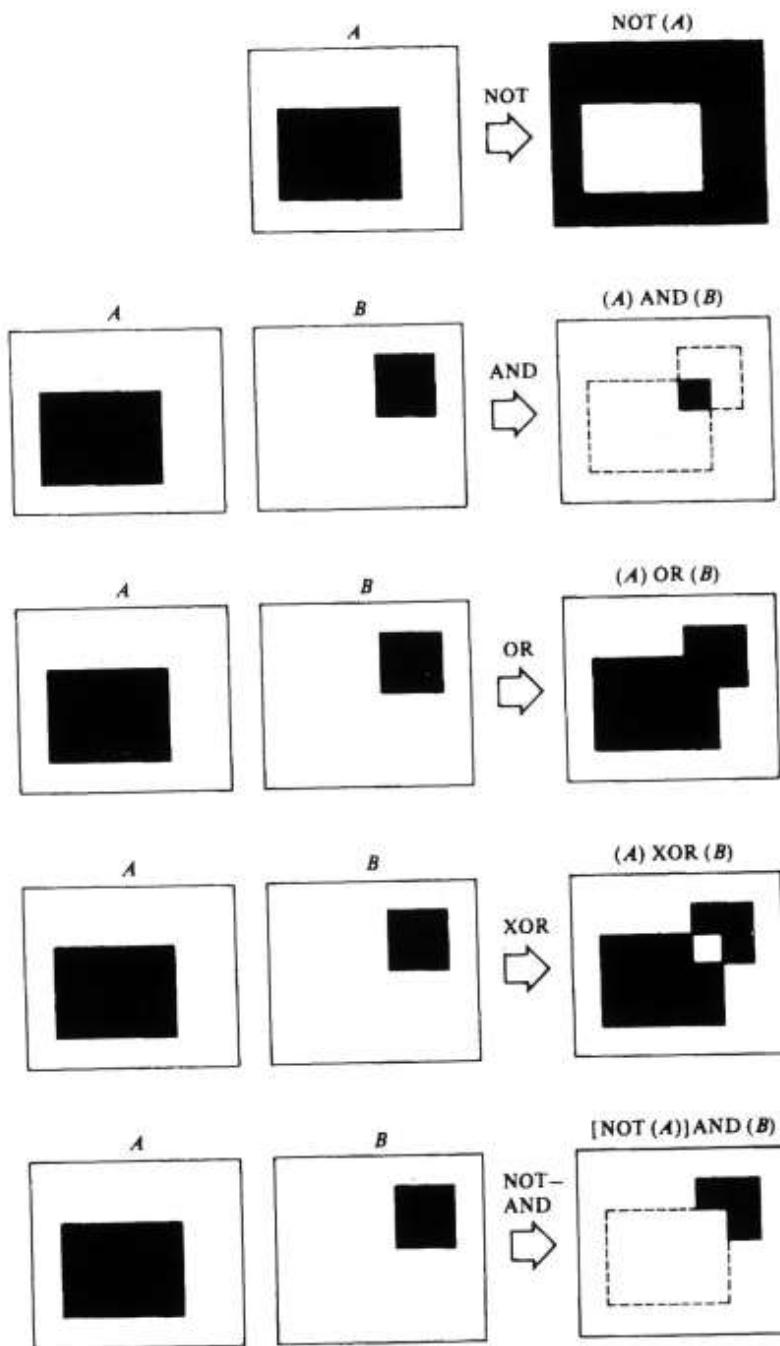
اگر اجازه دهیم که $w_i = 1/9$ $i = 1, 2, \dots, 9$ ، این عمل همان نتیجه عمل متوسط گیری را می دهد.

معادله (۵-۴. ۲) به طور گسترده ای در پردازش تصویر استفاده می شود. با انتخاب مناسب ضرایب و بکارگیری نقاب در هر نقطه تصویر، عملیات تصویری مفید و متنوعی نظیر کاهش نویز، نازکسازی ناحیه و آشکارسازی لبه ممکن می شود. در هر حال بکارگیری نقاب در تمام نقاط تصویر کاری پر محاسبه است. مثلاً کاربرد نقاب 3×3 در تصویر 512×512 نیازمند 9 ضرب و 8 جمع در هر نقطه است که جمماً 2359296 ضرب و 2097152 جمع مورد نیاز است.

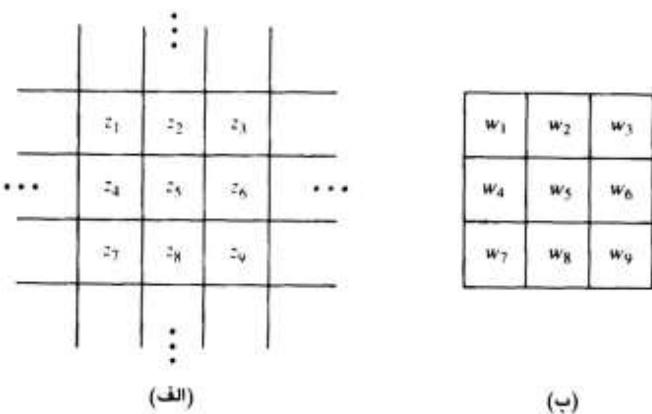
همان طور که در بخش ۱. ۴. ۳. بیان شد، اخیراً پردازشگرهای تصویر به واحد محاسبه و منطق (ALU) مجهر شده اند که وظیفه آن انجام اعمال حسابی و منطقی به طور موازی و نوعاً در نرخ قابهای ویدئو می باشد. برای ویدئو استاندارد ایالات متحده، ALU یک عمل حسابی یا منطقی بین دو تصویر 512×512 را در $1/30$ ثانیه (این بازه زمانی اغلب قاب^۱ یا زمان قاب تأمینه می شود) انجام دهد. با دانستن اهمیت عملیات نقابی در پردازش تصویر علاقه مندیم چگونگی استفاده از ALU برای شتاب دادن

۱) Frame

۴.۲ چند رابطه پایه ای بین پکسلها ۷۷



شکل ۲ ۱۴. چند مثال از عملیات منطقی روی تصاویر دودویی



شکل ۲.۱۵ (الف) زیر تابعه ای از یک تصویر که مقادیر پیکسل ها را نشان می دهد؛ (ب) یک نقاط 3×3 با ضرایب کلی

پردازش نقاطی را با کمی تفصیل برسی نماییم. برای روشن ساختن مطلب، نقاط 3×3 شکل ۲.۱۵. (b) و پیاده سازی آن را که توسط معادله (۲.۴-۵) بیان می شود، در نظر می گیریم. این روش به آسانی به نقاط $n \times m$ و سایر عملیات حسابی و منطقی قابل توسعه است.

روش بیان شده در بالا نیازمند دو بافر قاب با توانایی انتقال افقی و عمودی یک پیکسلی است (بخش ۲.۴. ۱ را ببینید). اجازه دهید بافر قاب A حاوی تصویری باشد که باید نقاط بر آن اعمال شود. در پایان پردازش بافر قاب B حاصل عملیات را در خود جای می دهد. به یاد آورید که عملیات ALU روی تمام پیکسل ها در یک زمان قاب انجام می شود، در حالی که تمام انتقال های قاب تقریباً آنی انجام می شوند. فرض می کنیم که تمام انتقال های یک پیکسلی هستند. با فرض این که در ابتدا $A = B$ ، و با استفاده از خط تیره به عنوان عدم انجام کار، از فرآیند ارائه شده در جدول ۲.۳ پیروی می کنیم. آخرین دو انتقال مورد نیاز هستند زیرا در پایان آخرین عمل روی B، تصویر در موقعیتی است که ضریب w نقاط روی محل z قرار دارد. این دو انتقال عدم انتباطی را تصحیح می کنند.

نکته اصلی در فهم اسلوب ارائه شده در جدول ۲.۳ این است که اعمال انجام شده در یک پیکسل منفرد از B را با توجه به چگونگی انتقالات نقاط طوری که حاصل معادله (۲.۴-۵) در آن محل تولید شود، بررسی نماییم. اولین عمل روی B، حاصل ضرب w در مقدار پیکسل آن محل را می دهد که حاصل این عمل z است. اولین انتقال به راست، همسایه با مقدار z (به شکل ۲.۱۵ (الف) نگاه کنید)

جدول ۳۰.۲ اعمال روی ALU

اعمال روی A	اعمال روی B
—	ضرب در w_5
انتقال به راست	—
—	جمع با
انتقال به پایین	$w_4 * A$
—	جمع با
انتقال به چپ	$w_3 * A$
—	جمع با
انتقال به چپ	$w_2 * A$
—	جمع با
انتقال به بالا	$w_1 * A$
—	جمع با
انتقال به بالا	$w_0 * A$
—	جمع با
انتقال به راست	$w_5 * A$
—	جمع با
انتقال به راست	$w_4 * A$
—	جمع با
انتقال به چپ	$w_3 * A$
انتقال به پایین	—

را روی مرکز نقاب می آورد. عمل بعدی، z را در w ضرب می کند و حاصل را به نتیجه' اولین مرحله می افزاید. بنابراین اکنون حاصل عملیات در نقطه مورد نظر $w_4 z_4 + w_5 z_5$ است. انتقال بعدی روی A و عمل ALU روی B، $w_1 z_5 + w_4 z_4 + w_5 z_1$ را در آن محل تولید می کند و بدین ترتیب ادامه می یابد. عملیات به طور موازی برای تمام نقاط B انجام می شود، بنابراین این عمل به طور همزمان در سایر نقاط بافر قاب انجام می شود. در بیشتر ALU ها، عمل ضرب تصویر در یک ثابت (مثلث $w_0 * A$) و سپس یک جمع در یک زمان قاب انجام می شود. بنابراین پیاده سازی معادله' (۴-۵) با ALU برای تصویر کامل در حدود ۹ زمان قاب ($\frac{9}{3}$ ثانیه) و برای نقاب $n \times m$ در حدود nm زمان قاب طول می کشد.

۵.۲ هندسه تصویربرداری

در توضیحات بعدی چند تبدیل مهم مورد استفاده در تصویربرداری مطرح می شود، مدلی برای دوربین بدست می آید و با مقداری تفصیل به مسأله تصویربرداری استریو^۱ پرداخته می شود.

1) Stereo imaging

۱.۵.۲ چند تبدیل پایه ای

موضوع این بخش درباره ایجاد نمایشی واحد برای مسائلی نظیر چرخش تصویر، تغییر مقیاس و انتقال است. تمام تبدیلات در سامانه مختصات سه بعدی کارترین که در آن هر نقطه مختصاتی به شکل (X, Y, Z) دارد، بیان می شوند. در مورد تصاویر دو بعدی، برای نمایش مختصات پیکسل از روش قبلی که نمایش با حروف کوچک (y, x) است، استفاده می کنیم. در اصطلاح عموم (Z) به عنوان مختصات جهانی^۱ نقطه می باشد.

انتقال

انتقال نقطه ای با مختصات (X_0, Y_0, Z_0) به یک محل جدید با استفاده از جابجایی های (X, Y, Z) را تصور کنید. انتقال به سادگی با معادلات زیر انجام می شود:

$$\begin{aligned} X^* &= X + X_0 \\ Y^* &= Y + Y_0 \\ Z^* &= Z + Z_0 \end{aligned} \quad (1-5.2)$$

که (X^*, Y^*, Z^*) مختصات نقطه جدید است. معادله ۱-۵.۲ را می توان در شکل ماتریسی به صورت زیر بیان کرد:

$$\begin{bmatrix} X^* \\ Y^* \\ Z^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2-5.2)$$

اغلب مفید است که با کاربرد چند تبدیل، یک تبدیل ترکیبی تولید کنیم، نظیر انتقالی که همراه تغییر مقیاس و سپس چرخش باشد. استفاده از ماتریس های مربعی بیان ریاضی این فرآیندر اتحاد زیادی آسان می کند. با دانستن این مطلب، معادله ۲-۵.۲ را می توان به شکل زیر نوشت:

$$\begin{bmatrix} X^* \\ Y^* \\ Z^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3-5.2)$$

از نظر مقادیر X^* ، Y^* و Z^* معادلات ۲-۵.۲ و ۳-۵.۲ یکسان هستند.

1) World coordinates

۸۱ ۵.۲ هندسه تصویربرداری

ما در سراسر این بخش از نمایش ماتریسی واحدی به صورت زیر استفاده می کنیم:

$$\mathbf{v}^* = \mathbf{A}\mathbf{v} \quad (4-5.2)$$

که \mathbf{A} یک ماتریس تبدیل 4×4 است، \mathbf{v} بردار ستونی است که مؤلفه های آن مختصات اولیه هستند:

$$\mathbf{v} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (5-5.2)$$

و \mathbf{v}^* بردار ستونی است که مؤلفه های آن مختصات تبدیل شده هستند:

$$\mathbf{v}^* = \begin{bmatrix} X^* \\ Y^* \\ Z^* \\ 1 \end{bmatrix} \quad (6-5.2)$$

با این طرز نمایش ماتریسی که برای انتقال به کار می رود، به صورت

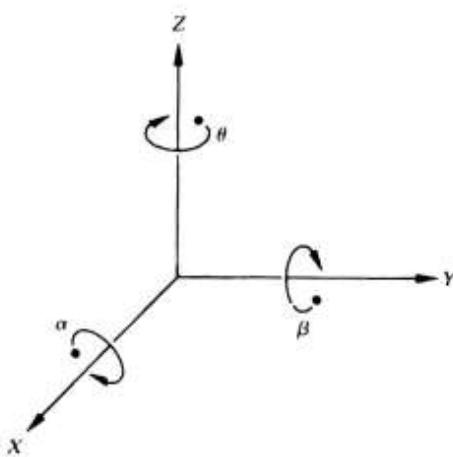
$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7-5.2)$$

است و فرآیند انتقال با استفاده از معادله (۷-۵.۲) انجام می شود بدین صورت که $\mathbf{v}^* = \mathbf{T}\mathbf{v}$.^{۱)}

تغییر مقیاس با ضرایب S_x, S_y و S_z در طول محورهای X, Y و Z ، با ماتریس تبدیل

$$\mathbf{S} = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (8-5.2)$$

1) Scaling



شکل ۱۶.۲ چرخش یک نقطه حول هر یک از محورهای مختصات. در صورتی که از روی محور چرخش به مبدأ نگاه کنیم، چرخش در جهت عقربه ساعت مثبت در نظر گرفته می شود.

انجام می شود.

چرخش

تبدیلاتی که برای چرخش سه بعدی استفاده می شوند، به طور طبیعی پیچیده‌تر از تبدیلاتی هستند که تاکنون بیان شدند. ساده‌ترین شکل این تبدیلات برای چرخش یک نقطه حول محورهای مختصات به کار می‌رود. چرخاندن نقطه دلخواه A حول نقطه دلخواه دیگر B در فضای نیازمند سه تبدیل متوالی است: اولین تبدیل طوری است که نقطه B را به مبدأ منتقل می‌کند، تبدیل دوم چرخش را انجام می‌دهد و تبدیل سوم نقطه B را به محل اولیه اش بر می‌گرداند.

با توجه به شکل ۱۶.۲، چرخش یک نقطه حول محور مختصات Z با زاویه θ با استفاده از ماتریس

تبدیل

$$\mathbf{R}_\theta = \begin{bmatrix} \cos \theta & \sin \theta & 0 & 0 \\ -\sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9-5.2)$$

بدست می‌آید. اگر از نقطه‌ای روی محور $Z +$ به مبدأ نگاه کنیم، زاویه چرخش θ مثبت است در صورتی

۸۳ ۵.۲ هندسه تصویربرداری

که چرخش در جهت عقربه های ساعت باشد. تبدیل فوق تنها بر مقادیر مختصات X و Y اثر می گذارد.

چرخش حول محور X با زاویه α با استفاده از تبدیل

$$\mathbf{R}_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10-5.2)$$

انجام می شود. همچنین چرخش حول محور Y با زاویه β با استفاده از تبدیل

$$\mathbf{R}_\beta = \begin{bmatrix} \cos \beta & 0 & -\sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11-5.2)$$

انجام می شود.

ترکیب چند تبدیل، و تبدیل معکوس

با یک ماتریس تبدیل 4×4 می توان چند تبدیل را اجرا کرد. مثلاً انتقال، تغییر مقیاس و چرخش نقطه^۷ حول محور Z با تبدیل ترکیبی

$$\mathbf{v}^* = \mathbf{R}_\theta(\mathbf{S}(\mathbf{T}\mathbf{v})) \\ = \mathbf{Av} \quad (12-5.2)$$

بیان می شود که می توان ماتریس تبدیل ترکیبی را از رابطه $\mathbf{R}_\theta \mathbf{ST} = \mathbf{A}$ بدست آورد که \mathbf{A} نیز یک ماتریس 4×4 است. در حالت کلی این ماتریس ها جایجا نمی شوند؛ بنابراین ترتیب تبدیلها مهم است.

گرچه توضیحات ارائه شده تاکنون در مورد تبدیل یک نقطه تکی بوده است، قابل تعمیم به تبدیل همزمان مجموعه m نقطه ای می باشد. با رجوع به معادله (۵-۵.۲) اجازه دهد که $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ بردارهای ستونی مختصات m نقطه مختلف باشند. برای ماتریس \mathbf{V} به ابعاد $m \times 4$ ، که ستونهایش بردارهای ستونی فوق باشند، تبدیل همزمان این نقاط با ماتریس \mathbf{A} به ابعاد 4×4 ، به صورت زیر

بیان می شود:

$$\mathbf{V}^* = \mathbf{AV} \quad (13-5.2)$$

ماتریس \mathbf{V}^* به ابعاد $m \times 4$ است. ستون نام آن، \mathbf{v}_1 ، مختصات نقطه تبدیل شده' متناظر با \mathbf{v}_1 است. بسیاری از تبدیلات بیان شده در بالا ماتریس های معکوسی دارند که عمل مخالف را انجام می دهند و می توان آنها را بدست آورد. مثلاً ماتریس انتقال معکوس به صورت زیر است:

$$\mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 & 0 & -X_0 \\ 0 & 1 & 0 & -Y_0 \\ 0 & 0 & 1 & -Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (14-5.2)$$

مشابهآ ماتریس چرخش معکوس \mathbf{R}_θ^{-1} به صورت زیر است:

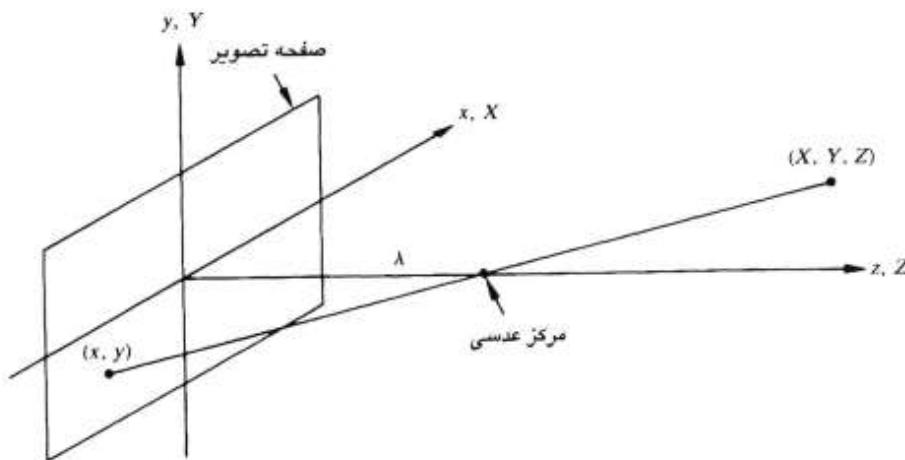
$$\mathbf{R}_\theta^{-1} = \begin{bmatrix} \cos(-\theta) & \sin(-\theta) & 0 & 0 \\ -\sin(-\theta) & \cos(-\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (15-5.2)$$

معمولآ ماتریس معکوس تبدیل های پیچیده تر با روش های عددی بدست می آیند.

۲.۵.۲ تبدیل پرسپکتیو

تبدیل پرسپکتیو (که تبدیل تصویر بردار هم نامیده می شود) نقاط فضای سه بعدی را روی یک صفحه تصویر می کند^۱. تبدیلات پرسپکتیو نقشی محوری در پردازش تصویر دارند زیرا تقریباً از چگونگی تشکیل تصویر در هنگام نظاره' دنیای سه بعدی فراهم می کنند. این تبدیلات تفاوت اساسی با تبدیلات مژروح در بخش ۲.۵.۱ هستند زیرا به دلیل وجود عمل تقسیم بر مقادیر متغیر (مقادیر مختصات)، غیرخطی هستند.

شکل ۲.۱۷. ۱۷ مدلی از فرآیند تشکیل تصویر را نشان می دهد. سامانه مختصات دوربین (x , y , z)



شکل ۱۷.۲ مدل پایه‌ای فرآیند تصویربرداری. سامانه مختصات دوربین (x, y, z) همراه است با سامانه مختصات جهانی (X, Y, Z) است

صفحه تصویری منطبق با صفحه yz دارد و محور نوری آن (توسط مرکز عدسی تعیین می‌شود) در طول محور z است. بنابراین مرکز صفحه تصویر در مبدأ و مرکز عدسی در مختصات $(0, 0, \lambda)$ است. اگر دوربین روی اشیای دور متمرکز شود، λ طول کانونی^۱ عدسی است. در اینجا فرض می‌شود که سامانه مختصات دوربین منطبق بر سامانه مختصات جهانی (X, Y, Z) است. در بخش ۲.۵.۳. این قید را بر می‌داریم.

اجازه دهید همان طور که در شکل ۱۷.۲ دیده می‌شود، (X, Y, Z) مختصات جهانی نقطه در صفحه^۲ بعدی باشد. در سراسر توضیحات بعدی فرض می‌کنیم که $\lambda > Z$ ؛ یعنی تمام نقاط مورد نظر، در جلوی مرکزی عدسی قرار دارند. مرحله بعدی به دست آوردن رابطه‌ای است که مختصات (x, y) مریبوط به تصویر^۳ نقطه (X, Y, Z) روی صفحه تصویر را بدهد. این کار به آسانی و با استفاده از تشابه مثلثها انجام می‌شود. با رجوع به شکل ۱۷.۲، داریم:

$$\frac{x}{\lambda} = -\frac{X}{Z - \lambda} = \frac{X}{\lambda - Z} \quad (16-5.2)$$

1) Focal length

2) Projection

3) Image plane

$$\begin{aligned} \frac{Y}{\lambda} &= -\frac{Y}{Z-\lambda} \\ &= \frac{Y}{\lambda-Z} \end{aligned} \quad (17-5. 2)$$

علامت های منفی در جلوی X و Y نیز همانند هندسه شکل ۱۷. نشان می دهند که در واقع نقاط تصویر معکوس شده اند.

مختصات صفحه تصویر مربوط به نقطه سه بعدی تصویر شده، مستقیماً از معادلات (۱۶-۵. ۲) و

(۱۷-۵. ۲) نتیجه می شود:

$$x = \frac{\lambda X}{\lambda - Z} \quad (18-5. 2)$$

$$y = \frac{\lambda Y}{\lambda - Z} \quad (19-5. 2)$$

این معادلات غیرخطی هستند، زیرا تقسیم بر متغیر Z را در خود دارند. گرچه می توان از آنها به همین صورت استفاده کرد، اغلب مرسوم است که آنها را به شکل ماتریس خطی مانند موارد چرخش، انتقال و تغییر مقیاس ارائه شده در بخش ۵. ۱ بیان کنیم. این کار به سادگی با استفاده از مختصات همگن اجرا می شود.

مختصات همگن نقطه ای با مختصات کارتزین (X, Y, Z) به صورت (kX, kY, kZ, k) تعریف می شود که k یک ثابت دلخواه غیر صفر است. به وضوح تبدیل مختصات همگن به کارتزین با تقسیم سه مختصه اول همگن بر مختصه چهارم انجام می شود. هر نقطه در سامانه مختصات جهانی کارتزین را می توان به شکل بارداری

$$\mathbf{w} = \begin{bmatrix} X \\ Y \\ Z \\ k \end{bmatrix} \quad (20-5. 2)$$

نمایش داد که متناظر همگن آن به صورت

$$\mathbf{w}_s = \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix} \quad (21-5. 2)$$

۸۷ ۵.۲ هندسه تصویربرداری

می باشد. اگر ماتریس تبدیل پرسپکتیو را به صورت

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{\lambda} & 1 \end{bmatrix} \quad (22-5.2)$$

تعريف کنیم، حاصل ضرب \mathbf{Pw}_k برداری است که با \mathbf{c}_k نشان داده می شود:

$$\begin{aligned} \mathbf{c}_k &= \mathbf{Pw}_k \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{\lambda} & 1 \end{bmatrix} \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix} \\ &= \begin{bmatrix} kX \\ kY \\ kZ \\ \frac{-kZ}{\lambda} + k \end{bmatrix} \end{aligned} \quad (23-5.2)$$

عناصر \mathbf{c}_k مختصات دوربینی در شکل همگن هستند. همان طور که بیان شد، این مختصات را می توان با تقسیم هر یک از سه مؤلفه اول \mathbf{c}_k بر مؤلفه چهارم آن به شکل کارتزین تبدیل نمود. بنابراین مختصات کارتزین هر نقطه در سیستم مختصات دوربین در شکل برداری با

$$\mathbf{c} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{\lambda X}{\lambda - Z} \\ \frac{\lambda Y}{\lambda - Z} \\ \frac{\lambda Z}{\lambda - Z} \end{bmatrix} \quad (24-5.2)$$

بیان می شود.

همان طور که قبل ا در معادلات (۲.۱۸-۵) و (۲.۱۹-۵) نشان داده شد، دوم مؤلفه اول، \mathbf{c}_1 ، مختصات (y, z) در صفحه تصویر مربوط به نقطه سه بعدی تصویر شده (X, Y, Z) هستند. از دیدگاه مدل موجود در شکل ۲.۱۷ مؤلفه سوم مورد توجه نیست. همان طور که بعداً نشان داده می شود، این مؤلفه مانند یک متغیر آزاد در تبدیل پرسپکتیو معکوس عمل می کند.

تبدیل پرسپکتیو معکوس هر نقطه تصویر را به حالت سه بعدی نگاشت می کند. بنابراین از معادله (۲.۲۳-۵) داریم،

$$\mathbf{w}_h = \mathbf{P}^{-1}\mathbf{c}_h \quad (25-5.2)$$

که ا \mathbf{P}^{-1} عبارتست از:

$$\mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\lambda} & 1 \end{bmatrix} \quad (26-5.2)$$

فرض کنید که یک نقطه تصویر مختصات $(0, y_0, z)$ را دارد، که عدد ۰ در محل z بیان می کند که صفحه تصویر در $z = 0$ قرار دارد. می توان این نقطه را در شکل برداری همگن به صورت

$$\mathbf{c}_h = \begin{bmatrix} kx_0 \\ ky_0 \\ 0 \\ k \end{bmatrix} \quad (27-5.2)$$

نمایش داد. آنگاه با کاربرد معادله (۲.۲۵-۵)، بردار مختصات جهانی همگن

$$\mathbf{w}_h = \begin{bmatrix} kx_0 \\ ky_0 \\ 0 \\ k \end{bmatrix} \quad (28-5.2)$$

یا معادل آن در مختصات کارتزین،

۱۹ ۵.۲ هندسه تصویربرداری

$$\mathbf{w} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ 0 \end{bmatrix} \quad (29-5.2)$$

بدست می آید. این نتیجه به وضوح غیرمنتظره است زیرا برای هر نقطه سه بعدی، $Z = 0$ بددست می آید. این مشکل ناشی از نگاشتن صحنه سه بعدی روی صفحه تصویر است که در واقع یک تبدیل چند-به-یک می باشد. نقطه تصویر (x_0, y_0) متناظر با مجموعه نقاط سه بعدی همراستا که روی خط گذرنده از دو نقطه $(x_0, y_0, 0)$ و $(0, 0, \lambda)$ قرار دارند، می باشد. معادلات این خط در سامانه مختصات جهانی، از معادلات $(18-5.2)$ و $(19-5.2)$ نتیجه می شوند؛ یعنی،

$$X = \frac{x_0}{\lambda} (\lambda - Z) \quad (30-5.2)$$

و

$$Y = \frac{y_0}{\lambda} (\lambda - Z). \quad (31-5.2)$$

معادلات $(30-5.2)$ و $(31-5.2)$ نشان می دهند که بدون اطلاع از نقطه سه بعدی که نقطه تصویر را ایجاد کرده است (مثلاً مختصه Z آن)، نمی توان به طور کامل نقطه سه بعدی را از نقطه تصویرش بازیابی کرد. با استفاده از این مطلب که به یقین غیرمنتظره نیست، می توان با به کار گیری مؤلفه \mathbf{Z} مربوط به بردار \mathbf{c}_k به عنوان متغیر آزاد، بجای این که صفر فرض شود، رابطه تبدیل پرسپکتیو معکوس را تعیین کرد. بنابراین با فرض

$$\mathbf{c}_k = \begin{bmatrix} kx_0 \\ ky_0 \\ kz \\ k \end{bmatrix} \quad (32-5.2)$$

از معادله $(25-5.2)$ نتیجه می شود که

$$\mathbf{w}_k = \begin{bmatrix} kx_0 \\ ky_0 \\ kz \\ \frac{kz}{\lambda} + k \end{bmatrix} \quad (33-5.2)$$

٤٠ پردازش تصویر رقمنی

که با تبدیل به مختصات کارتزین داریم

$$\mathbf{w} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \frac{\lambda x_0}{\lambda + z} \\ \frac{\lambda y_0}{\lambda + z} \\ \frac{\lambda z}{\lambda + z} \end{bmatrix} \quad (34-5.2)$$

به عبارت دیگر، با درنظر گرفتن z به عنوان متغیر آزاد، معادلات

$$\begin{aligned} X &= \frac{\lambda x_0}{\lambda + z} \\ Y &= \frac{\lambda y_0}{\lambda + z} \\ Z &= \frac{\lambda z}{\lambda + z} \end{aligned} \quad (35-5.2)$$

بدهست می‌آیند. در معادله فوق با نوشتن z بر حسب Z و قرار دادن در دو عبارت اول آن داریم:

$$X = \frac{x_0}{\lambda} (\lambda - Z) \quad (36-5.2)$$

و

$$Y = \frac{y_0}{\lambda} (\lambda - Z) \quad (37-5.2)$$

این نتیجه تأیید می‌کند که بازسازی یک نقطه سه بعدی از تصویرش با استفاده از تبدیل پرسپکتیو معکوس نیازمند اطلاع از حداقل یکی از مختصات جهانی آن نقطه می‌باشد. در بخش ۲.۵.۵. مجدد به این مسئله پرداخته می‌شود.

۳.۵.۴ مدل دورین

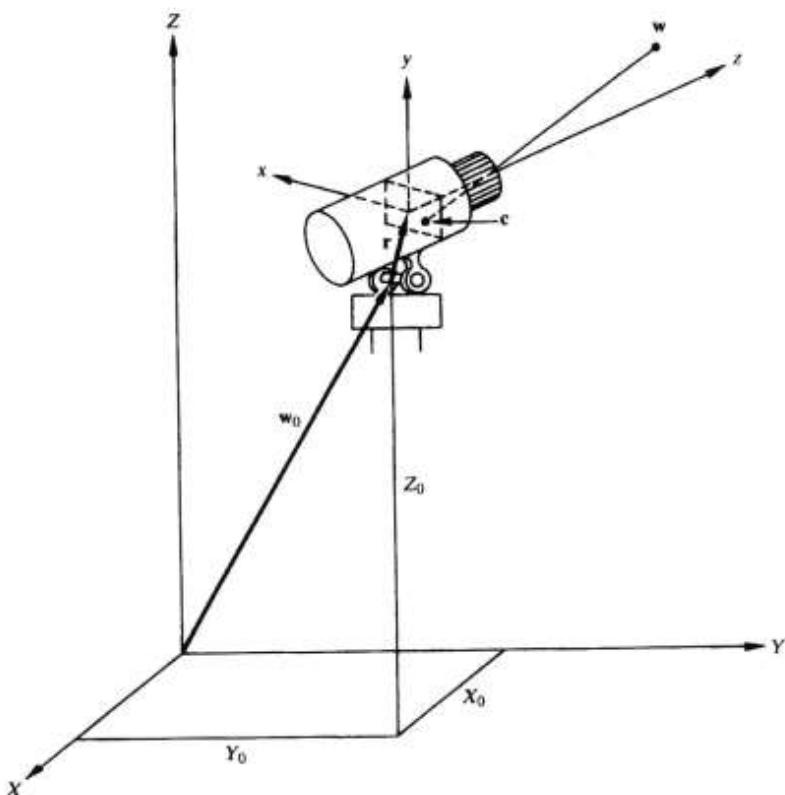
معادلات (۲.۵-۲۳) و (۲.۵-۲۴) تشکیل تصویر با تصویر کردن^۱ نقاط سه بعدی روی صفحه تصویر

۱) Projection

۵.۲ هندسه تصویربرداری ۹۱

راتوصیف می کنند. بنابراین این دو معادله یک مدل ریاضی پایه ای از دوربین تصویربرداری ارائه می کنند. این مدل بر پایه این فرض است که سامانه مختصات جهانی و سامانه مختصات دوربین منطبق هستند. در این بخش حالت کلی تری که در آن دو سامانه مختصات فوق می توانند منطبق نباشند، را در نظر می گیریم. در هر حال باز هم هدف اصلی بدست آوردن مختصات صفحه تصویر مربوط به هر نقطه جهانی مورد نظر است.

شکل ۱۸.۲ سامانه مختصات جهانی (X, Y, Z) را نشان می دهد که برای تعیین محل دوربین و نقاط سه بعدی (که بانداد w نشان داده می شوند) از آنها استفاده می کیم. همچنین شکل ۱۸.۲ سامانه مختصات دوربین (x, y, z) و نقاط تصویر (که با c نشان داده می شود) را نشان می دهد. فرض بر این است که دوربین روی پایه گردانی^۱ نصب شده است، که امکان می دهد دوربین به اندازه زاویه θ چرخش افقی و



شکل ۱۸.۲ هندسه تصویربرداری با دو سامانه مختصات (از [Lee و Fu, Gonzalez ۱۹۸۷]).

1) Gimbal

۴۲ پردازش تصویر رقمه

به اندازه زاویه α چرخش عمودی بنماید. در اینجا، پن^۱ زاویه بین دو محور x و X و تیلت^۲ زاویه بین دو محور z و Z است. انحراف مرکز پایه گردان از مبدأ سامانه مختصات جهانی با W_0 و انحراف مرکز تصویر برداری نسبت به مرکز صفحه پایه گردان با بردار r و با مؤلفه های (r_1, r_2, r_3) نشان داده می شود.

مفاهیم ارائه شده در بخش های ۲.۱.۵.۲ و ۲.۵.۲ تمام ابزارهای لازم برای بدست آوردن مدل دوربین، طبق آرایش هندسی شکل ۲.۱۸. را فراهم می آورند. راهکار این است که با به کار گیری مجموعه ای از تبدیلات سامانه مختصات جهانی و سامانه مختصات دوربین را هم راستا سازیم. آنگاه تنها با کاربرد تبدیل پرسپکتیو طبق معادله (۲.۲۲-۵) مختصات مربوط به هر نقطه جهانی در صفحه تصویر را بدست می آوریم. به عبارت دیگر، در ابتدا و قبل از کاربرد تبدیل پرسپکتیو مسأله را به صورت آرایش هندسی نشان داده شده در شکل ۲.۱۷. ساده می کنیم.

فرض کنید که در ابتداء دوربین در وضعیت عادی است، بدین مفهوم که مرکز پایه گردان و مبدأ صفحه تصویر در مبدأ سامانه مختصات جهانی بوده و تمام محورها هم راستا هستند. آنگاه آرایش هندسی شکل ۲.۱۸. را می توان به چند طریق به دست آورد. فرض کنید مراحل زیر اجرا شوند: (۱) جابجایی مرکز پایه گردان از مبدأ، (۲) چرخش افقی محور x ، (۳) چرخش عمودی محور z ، و (۴) جابجایی صفحه تصویر نسبت به مرکز پایه گردان.

روشن است که این دنباله مراحل مکانیکی بر نقاط جهانی اثر نمی گذارد زیرا مجموعه نقاطی که پس از انتقال دوربین از وضعیت عادیش توسط آن دیده می شوند، کاملاً متفاوت است. در هر حال، با اجرای دقیق همان دنباله از مراحل در تمام نقاط جهانی می توان دوباره به وضعیت عادی رسید. دوربین در وضعیت عادی، آرایش شکل ۲.۱۷. برای کاربرد تبدیل پرسپکتیو را برآورده می سازد. بنابراین مسأله به اجرای مجموعه ای از تبدیلات در هر نقطه جهانی، مطابق مراحلی که قبلاً بیان شد، ساده می شود.

انتقال مبدأ سامانه مختصات جهانی به محل مرکز پایه گردان با استفاده از ماتریس تبدیل

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & -X_0 \\ 0 & 1 & 0 & -Y_0 \\ 0 & 0 & 1 & -Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (۳۸-۵.۲)$$

1) Pan

2) Tilt

۵.۲ هندسه تصویربرداری ۹۳

انجام می شود. به عبارت دیگر، نقطه جهانی همگن \mathbf{w}_0 که در مختصات (X_0, Y_0, Z_0) قرار دارد، پس از تبدیل \mathbf{Gw}_0 در مبدأ سامانه مختصات جدید قرار می گیرد.

همان طور که قبلاً بیان شد، زاویه بین دو محور x و X است. در وضعیت عادی این دو محور هم راستا هستند. برای چرخش افقی محور x به اندازه زاویه مطلوب، تنها آن را به اندازه θ حول محور z می چرخانیم. این چرخش با استفاده از ماتریس تبدیل \mathbf{R} از معادله (۹-۵) انجام می شود. به بیان دیگر، کاربرد این ماتریس به تمام نقاط (از جمله نقطه \mathbf{Gw}_0) دقیقاً محور x را تا محل مطلوب می چرخاند. هنگام استفاده از معادله (۹-۵) لازم است که قرارداد تعیین شده در شکل ۲.۱۶ را در نظر داشته باشیم. یعنی وقتی نقاط در جهت عقربه ساعت می چرخند، زوایا مشبّت در نظر گرفته می شوند که به طور ضمنی به چرخش دوربین در خلاف عقربه ساعت حول محور z نیز دلالت دارد. وضعیت چرخانده نشده (صفر درجه) متناظر با موردي است که محورهای x و X هم راستا هستند.

در این لحظه هنوز محورهای z و Z هم راستا هستند. چون تیلت زاویه بین این دو محور است، با چرخاندن محور z به اندازه α ، دوربین را به اندازه زاویه α می چرخانیم. این چرخش حول محور x است و با کاربرد ماتریس تبدیل \mathbf{R} از معادله (۹-۵) بر تمام نقاط (از جمله نقطه \mathbf{Gw}_0) انجام می شود. در اینجا هم چرخش دوربین در خلاف عقربه ساعت به مفهوم چرخش مشبّت است و علامت صفر درجه (0°) برای وقتی است که محورهای z و Z هم راستا باشند.[†]

طبق توضیح بخش ۲.۵.۴، دو ماتریس چرخش را می توان به صورت ماتریس $\mathbf{R} = \mathbf{R}_\alpha \mathbf{R}_\theta$ با هم ترکیب کرد. آنگاه از معادلات (۹.۵.۲) و (۹.۵.۱۰) داریم:

$$\mathbf{R} = \begin{bmatrix} \cos \theta & \sin \theta & 0 & 0 \\ -\sin \theta \cos \alpha & \cos \theta \cos \alpha & \sin \alpha & 0 \\ \sin \theta \sin \alpha & -\cos \theta \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (39-5.2)$$

در نهایت جابجایی مبدأ صفحه تصویر با بردار \mathbf{t} با استفاده از ماتریس تبدیل

[†] یک راه مقید برای مشاهده این تبدیلات، ساخت سامانه محورها (مثلاً بالوله پاکن‌ها)، برچسب گذاری محورهای x ، y و z و اجرای دستی چرخش محورها به صورت تک به تک است.

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & -r_1 \\ 0 & 1 & 0 & -r_2 \\ 0 & 0 & 1 & -r_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (40-5.2)$$

اجرا می شود. بنابراین کاربرد دنباله تبدیلات \mathbf{CRGw}_h بر \mathbf{W}_h ، سامانه های مختصات دوربین و جهانی را منطبق می کند. در پایان، مختصات صفحه تصویر نقطه \mathbf{W}_h با استفاده از معادله (۲۳-۵.۲) بدست می آید. به عبارت دیگر، یک نقطه جهانی همگن که توسط دوربین دیده می شود و آرایش هندسی شکل ۲ را برآورده می کند، در سامانه مختصات دوربین نمایش همگن زیر را دارد:

$$\mathbf{c}_h = \mathbf{PCRGw}_h \quad (41-5.2)$$

معادله (۲۱-۵) تبدیل پرسپکتیوی را ارائه می کند که شامل دو سامانه مختصات است. همان طور که در بخش ۲.۵.۲ نشان داده شد، با تقسیم اولین و دومین مؤلفه، \mathbf{c}_h بر چهارمین مؤلفه، مختصات کارتزین (x, y) نقطه تصویر شده را بدست می آوریم. از بسط معادله (۲۱-۵) و تبدیل به مختصات کارتزین داریم

$$x = \lambda \frac{(X - X_0)\cos \theta + (Y - Y_0)\sin \theta - r_1}{-(X - X_0)\sin \theta \sin \alpha + (Y - Y_0)\cos \theta \sin \alpha - (Z - Z_0)\cos \alpha + r_3 + \lambda} \quad (42-5.2)$$

و

$$y = \lambda \frac{-(X - X_0)\sin \theta \cos \alpha + (Y - Y_0)\cos \theta \cos \alpha + (Z - Z_0)\sin \alpha - r_2}{-(X - X_0)\sin \theta \sin \alpha + (Y - Y_0)\cos \theta \sin \alpha - (Z - Z_0)\cos \alpha + r_3 + \lambda} \quad (43-5.2)$$

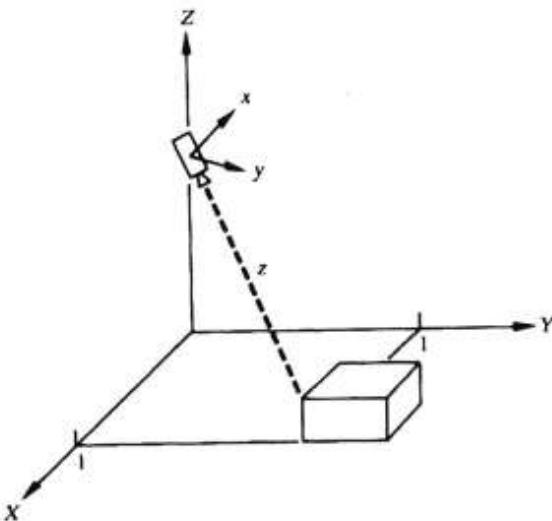
که x و y مختصات تصویر نقطه \mathbf{W}_h با مختصات جهانی (X, Y, Z) هستند. وقتی که $X_0 = Y_0 = Z_0 = 0$ باشند، دو معادله فوق به صورت معادلات (۲۱-۵) و (۱۹-۵.۲) ساده می شوند.

مثال: به منظور روشن ساختن مفاهیمی که اخیراً بیان شد، فرض کنید که می خواهیم مختصات تصویر

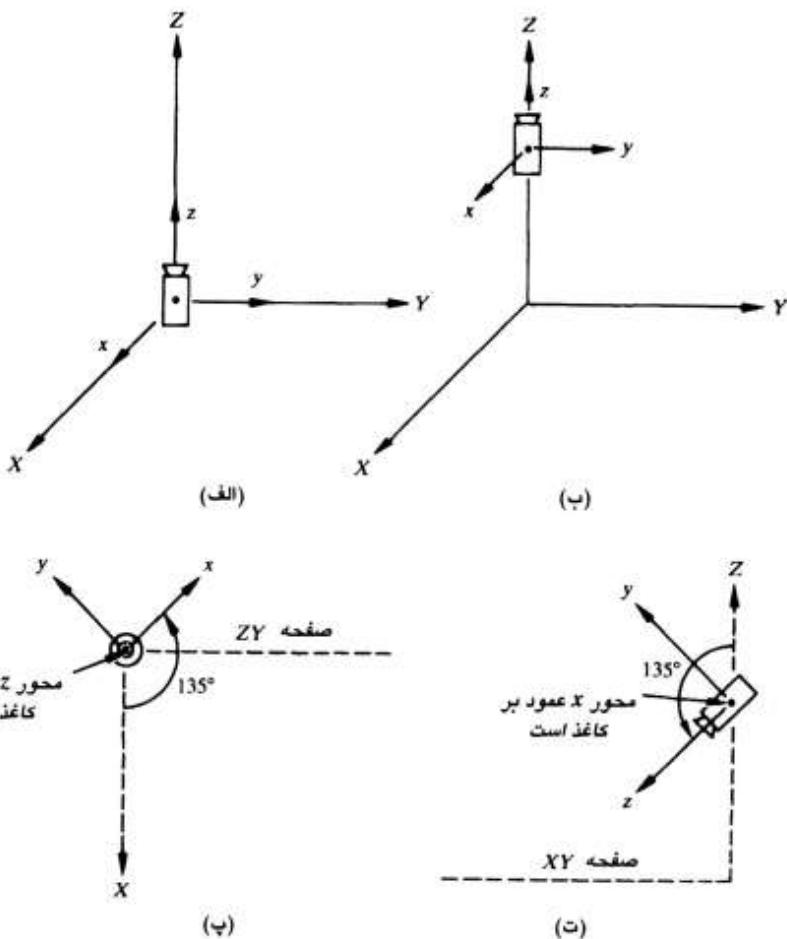
۵.۲ هندسه تصویربرداری ۴۵

گوشه' بلوک موجود در شکل ۲. ۱۹ را بیاییم. دوربین از مبدأ فاصله دارد و صحنه را با زوایای پن 135° و تیلت 135° می بیند. قرارداد می کنیم که وقتی در امتداد محور چرخش به مبدأ نگاه کنیم و دوربین خلاف عقربه ساعت بچرخد، زوایای تبدیل مشتبه است.

اجازه دهید مراحل مورد نیاز برای جابجا کردن دوربین از وضعیت عادی به هندسه نشان داده شده در شکل ۲. ۱۹ را به تفصیل بررسی کنیم. در شکل ۲. ۲۰ (الف) دوربین در وضعیت عادی است و در شکل ۲. ۲۰ (ب) از مبدأ جابجا شده است. توجه کنید که بعد از این مرحله، تنها برای تعیین مرجع زوایا از محورهای مختصات جهانی استفاده می شود. یعنی بعد از جابجایی مبدأ مختصات جهانی، تمام چرخشها حول محورهای جدید (دوربین) انجام می شوند. شکل ۲. ۲۰ (پ) نمایی را در امتداد محور z دوربین برای تعیین زاویه پن نشان می دهد. در این مورد چرخش دوربین حول محور z در خلاف عقربه ساعت است، بنابراین نقاط جهانی در جهت مخالف حول این محور چرخانده می شوند. در نتیجه θ زاویه ای مشتبه است. شکل ۲. ۲۰ (ت) نمایی را در امتداد محور x دوربین برای تعیین زاویه تیلت نشان می دهد. چرخش حول این محور خلاف عقربه ساعت است که باعث می شود α زاویه ای مشتب باشد. در دو تصویر بعدی محورهای مختصات جهانی با خطوط مقطع نشان داده می شوند تا بین وسیله تأکید کنند که تنها فایده آنها در اینجا تعیین مرجع صفر برای زوایای پن و تیلت است. مرحله آخر جابجایی صفحه تصویر از مرکز پایه' گردان را نمایش نداده ایم.



شکل ۲. ۱۹-۲ دوربینی که یک صحنه سه بعدی را می بیند (از Lee و Gonzalez [1987]).



شکل ۲۰.۲ (الف) دوربین در وضعیت عادی؛ (ب) مرکز پایه که از مبدأ جابجا شده است؛ (ج) نمای دید چرخش حول محور Z برای تعیین زاویه پن؛ (د) نمای دید چرخش حول محور X برای تعیین زاویه تیلت (از [۱۹۸۷] Lee, Gonzalez, Fu).

مقادیر پارامترها در این مسأله عبارتند از:

$$X_0 = 0 \text{ m} \quad Y_0 = 0 \text{ m} \quad Z_0 = 1 \text{ m};$$

$$\alpha = 135^\circ \quad \theta = 135^\circ;$$

$$r_1 = 0.03 \text{ m} \quad r_2 = r_3 = 0.02 \text{ m} \quad \lambda = 35 \text{ mm} = 0.035 \text{ m}$$

گوشه' مورد نظر در مختصات $(X, Y, Z) = (1, 1, 0.2)$ قرار دارد.

1) Observer view

۵.۲ هندسه تصویربرداری

برای محاسبه مختصات گوشه بلوک، کافیست مقادیر پارامترها در معادلات (۴۲-۵.۲) و (۴۳-۵.۲) جایگزین کنیم؛ یعنی،

$$x = \lambda \frac{-0.03}{-1.53 + \lambda}$$

به طور مشابه،

$$y = \lambda \frac{+0.42}{-1.53 + \lambda}$$

با قراردادن $\lambda = 0.035$ مختصات تصویر

$$x = 0.0007 \text{ m} \quad \text{و} \quad y = -0.009 \text{ m}.$$

بدست می آید.

توجه کنید که این مختصات کاملاً درون صفحه تصویربرداری 1×1 اینچ ($0.025 \times 0.025 \text{ m}$) قرار دارند. به آسانی تشخیص داده می شود که استفاده از عدسی با طول کانونی به عنوان مثال ۲۰۰mm گوشه بلوک را بیرون مرز صفحه ای با ابعاد فوق (یعنی، بیرون میدان دید مؤثر دوربین) تصویر می کند. در نهایت توجه کنید که تمام مختصات بدست آمده با معادلات (۴۲-۵.۲) و (۴۳-۵.۲) نسبت به مرکز صفحه تصویر هستند. برای استفاده از قرارداد قبلی که مبدأ تصویر در گوشه چپ بالای آن است، تغییری در مختصات مورد نیاز است.

۴.۵.۲ تنظیم^۱ دوربین

در بخش ۴.۵.۳ معادلاتی برای محاسبه مستقیم مختصات تصویر (y, x) از نقطه جهانی W بدست آورдیم. همان طور که در معادلات (۴۲-۵.۲) و (۴۳-۵.۲) دیده می شود، پیاده سازی این معادلات نیازمند اطلاع از طول کانونی، انحرافها و زوایای بن و تیلت است. گرچه می توان این پارامترها به طور مستقیم اندازه گرفت، اغلب مرسومتر است که یک یا چند عدد از این پارامترها با استفاده از خود دوربین به عنوان وسیله اندازه گیری تعیین شوند (مخصوصاً وقتی دوربین بارها حرکت می کند). انجام این کار

¹) Calibration

۹۸ پردازش تصویر رقمه

نیازمند مجموعه ای از نقاط تصویر با مختصات جهانی معلوم است و فرآیند محاسباتی که با استفاده از این نقاط معلوم پارامترهای دوربین را محاسبه می کند، اغلب تنظیم دوربین خوانده می شود.

با توجه به معادله (۴۱-۵. ۲)، اجازه دهد $\mathbf{A} = \mathbf{PCRG}$. ماتریس \mathbf{A} شامل تمام پارامترهای دوربین است و طبق معادله (۴۱-۵. ۲)، $\mathbf{c}_h = \mathbf{Aw}_h$ با قراردادن $1 = k$ در نمایش همگن رابطه زیر بدست می آید:

$$\begin{bmatrix} c_{h1} \\ c_{h2} \\ c_{h3} \\ c_{h4} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (44-5. 2)$$

براساس شرح موجود در بخش های ۲. ۵. ۳. ۵. ۲ و ۲. ۵. ۳. ۵. ۲، مختصات دوربین در شکل کارتزین

$$x = c_{h1}/c_{h4} \quad (45-5. 2)$$

و

$$y = c_{h2}/c_{h4}. \quad (46-5. 2)$$

هستند. با جایگزین کردن $c_{h4} = xc_{h4}$ و $c_{h2} = yc_{h4}$ در معادله (۴۴-۵. ۲) و بسط حاصلضرب ماتریسی داریم

$$\begin{aligned} xc_{h4} &= a_{11}X + a_{12}Y + a_{13}Z + a_{14} \\ yc_{h4} &= a_{21}X + a_{22}Y + a_{23}Z + a_{24} \\ c_{h4} &= a_{41}X + a_{42}Y + a_{43}Z + a_{44}, \end{aligned} \quad (47-5. 2)$$

که از بسط c_{h3} بدلیل بستگی آن به Z صرفنظر می شود.

با جایگذاری c_{h3} در دو معادله اول (۴۷-۵. ۲)، دو معادله با ۱۲ ضریب مجهول بدست می آید:

$$a_{11}X + a_{12}Y + a_{13}Z - a_{41}xX - a_{42}xY - a_{43}xZ - a_{44}x + a_{14} = 0 \quad (48-5. 2)$$

$$a_{21}X + a_{22}Y + a_{23}Z - a_{41}yX - a_{42}yY - a_{43}yZ - a_{44}y + a_{24} = 0. \quad (49-5. 2)$$

پس فرآیند کالیبراسیون مشتمل است بر (۱) بدست آوردن $m \geq 6$ نقطه جهانی (دو معادله وجود دارد)

۵.۲ هندسه تصویربرداری ۴۹

با مختصات معلوم (X_i, Y_i, Z_i) که $i = 1, 2, \dots, m$ ؛ (۲) تصویر کردن نقاط با دوربینی که در وضعیت معلوم است تا بدین وسیله نقاط تصویر متناظر آنها، (x_i, y_i) بازای $i = 1, 2, \dots, m$ ، بدست آیند؛ و (۳) به کارگیری این نتایج در معادلات $(2.48-5)$ و $(2.49-5)$ تابعی ترتیب ضرایب مجھول پیدا شود. روش های عددی بسیاری برای پیدا کردن جواب بهینه^۱ سامانه معادلات خطی نظری آنچه با این معادلات داده شده است، وجود دارد (به عنوان مثال به Noble [۱۹۶۹] مراجعه شود).

۵.۵.۲ تصویربرداری استریو^۲

به یاد آورید که نگاشت صحنه سه بعدی بر روی صفحه تصویر تبدیلی چند-به-یک است، یعنی نقطه تصویر به طور یکتا محل نقطه جهانی را معین نمی کند. با این حال اطلاعات عمتم^۳ از دست رفته را می توان با روش های تصویربرداری استریو بدست آورد.

همان طور که شکل ۲.۲۱ نشان می دهد، تصویربرداری استریو شامل بدست آوردن دو نمای مجزا از شیئی (در این شرح یک نقطه جهانی) می باشد. فاصله بین مراکز دو عدسی محور اصلی^۳ نامیده می شود و هدف پیدا کردن مختصات (X, Y, Z) نقطه W با داشتن نقاط تصویر (x_1, y_1) و (x_2, y_2) است. فرض می شود که دوربین های کسان و سامانه مختصات هر دو دوربین کاملاً هم راست هستند و تنها در محل مبدأ شان با هم تفاوت دارند، وضعیتی که معمولاً در عمل وجود دارد. به خاطر آورید که پس از این که سامانه های مختصات دوربین و جهانی منطبق شدند، صفحه XY مربوط به تصویر با صفحه XY مربوط به سامانه مختصات جهانی هم راست می شود. بنابراین با فرض فوق، مختصه Z از W برای سامانه های مختصات هر دو دوربین دقیقاً برابر است.

اجازه دهید که مانند شکل ۲.۲۲ دوربین اول را بر سامانه مختصات جهانی منطبق نماییم. از معادله $(2.50-5)$ نتیجه می شود که W روی خطی با مختصات (جزئی)

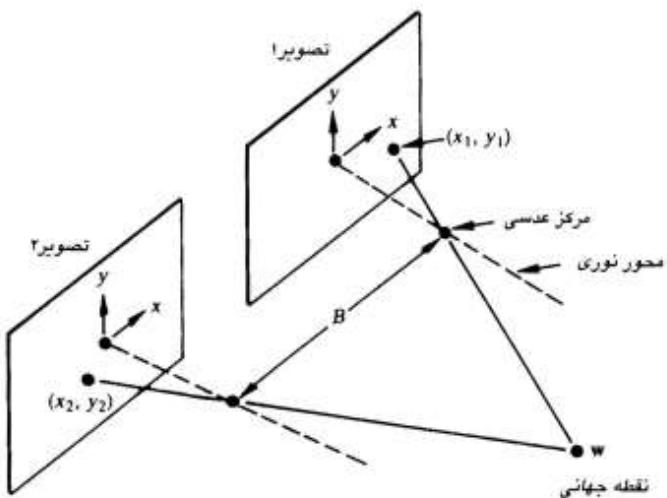
$$X_1 = \frac{x_1}{\lambda} (\lambda - Z_1) \quad (50-5.2)$$

قرار دارد. زیرنویس های مربوط به X و Z دلالت بر این دارند که دوربین اول به مبدأ سامانه مختصات جهانی رفته و دوربین دوم و W نیز به گونه ای به دنبال آن جایجا شده اند که آرایش نسبی شکل ۲.۲۱ را حفظ

1) Stereoscopic (or stereo) imaging

2) Depth

3) Baseline



شکل ۲۱.۲ مدل فرآیند تصویربرداری استریو. (از Lee [1987] و Gonzalez, Fu و

کرده‌اند. اگر در عوض، دوربین دوم به مبدأ سامانه مختصات جهانی آورده شود، w روی خطی با مختصات (جزئی) زیر قرار دارد:

$$X_2 = \frac{x_2}{\lambda} (\lambda - Z_2). \quad (51-5.2)$$

در هر حال به دلیل فاصله بین دوربینها و نیز به دلیل اینکه مختصه Z مربوط به w برای هردو سامانه مختصات دوربینی یکسان است، نتیجه می‌شود که

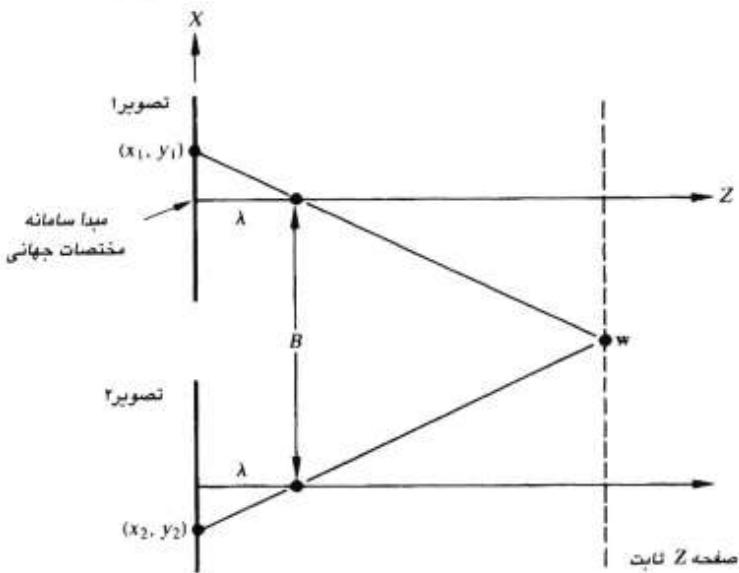
$$X_2 = X_1 + B \quad (52-5.2)$$

۶

$$Z_2 = Z_1 = Z \quad (53-5.2)$$

که B طول خط پایه است.
با جایگذاری معادلات (۵۱-۵.۲)، (۵۲-۵.۲) و (۵۳-۵.۲) در معادلات (۵۰-۵.۲) و (۵۱-۵.۲) نتیجه می‌شود:

۵.۲ هندسه تصویربرداری ۱۰۱



شکل ۲۰.۲ نمای بالای شکل ۲۱.۲ که دوربین اول بر سامانه مختصات جهانی منطبق شده است. (از *Fu, Lee [1987]* و *Gonzalez*)

$$X_1 = \frac{x_1}{\lambda} (\lambda - Z) \quad (54-5.2)$$

$$X_1 + B = \frac{x_2}{\lambda} (\lambda - Z). \quad (55-5.2)$$

با تفريح معادله (۵۴-۵.۲) از معادله (۵۵-۵.۲) و حل حاصل برای Z داریم:

$$Z = \lambda - \frac{\lambda B}{x_2 - x_1} \quad (56-5.2)$$

که نشان می دهد که اگر بتوان تفاصل بین دو مختصه تصویر مربوطه، x_1 و x_2 ، را تعیین کرد و طول خط پایه و فاصله کانونی معلوم باشد، محاسبه مختصه Z مربوط به نقطه w کار ساده ای است. آنگاه مختصات جهانی X و Y مستقیماً از معادلات (۳۰-۵.۲) و (۳۱-۵.۲) با استفاده از یکی از مختصات (x_1, y_1) یا (x_2, y_2) بدست آیند.

مشکل ترین کار در استفاده از معادله (۵۶-۵.۲) برای بدست آوردن Z، پیدا کردن دو نقطه متناظر در

تصاویر مختلف یک صحنه است. چون این نقاط معمولاً در نزدیکی هم هستند، یک روش متداول انتخاب یک نقطه از یک ناحیه^۱ کوچک در یکی از نماهای^۲ تصویر و سپس تلاش برای پیدا کردن بهترین ناحیه تطبیق کننده^۳ در نمای دیگر با استفاده از فنون همبستگی می‌باشد، که در فصل^۹ شرح داده می‌شود. وقتی صحنه، ویژگی‌های مجزایی نظری گوشه‌های مهم دارد، معمولاً روش تطبیق ویژگی راه حل سریعتری برای تعیین تراکم می‌دهد. فرآیند تنظیم که در بخش^۴ .^۵ ارائه شد، با تنظیم مستقل از هم دوربین‌ها بسادگی قابل کاربرد در تصویربرداری استریو است.

۶.۲ فیلم عکاسی

فیلم عکاسی یک عنصر مهم در سامانه‌های پردازش تصویر است. از فیلم عکاسی اغلب به عنوان رسانه ضبط تصاویر و رودی استفاده می‌شود و تاکنون پر مصرف‌ترین رسانه برای ضبط تابع خروجی بوده است. به این دلایل، این فصل را با شرحی از چند خاصیت پایه‌ای فیلم عکاسی تکریج و رابطه آنها با کاربردهای پردازش تصویر پایان می‌دهیم.

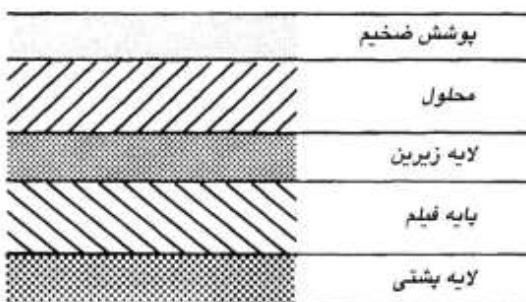
۱.۶.۲ ساختمان و نورخوردگی^۶ فیلم

شکل ۲۳. ۲۳. مقطع عرضی بزرگ نمایی شده یک فیلم عکاسی متداول را نشان می‌دهد که دارای لایه‌ها و اجزای زیر است (۱) یک پوشش ضخیم از رُلاتین برای محافظت در مقابل خراشها و سایش^۷؛ (۲) یک لایه محلول شفاف از بلورهای ریز هالاید نقره^۸؛ (۳) لایه زیرین جهت تقویت چسبندگی محلول به پایه فیلم^۹؛ (۴) پایه فیلم^{۱۰} یا تکیه گاه که از تری استات سلولز^{۱۱} یا پلی مری مرتبط با آن ساخته شده است؛ و (۵) لایه پشتی^{۱۲} برای جلوگیری از خم شدنگی.

وقتی که فیلم در معرض نور قرار داده شود، دانه‌های هالاید نقره انرژی را جذب کرده، تغییر فیزیکی پیچیده‌ای در آنها رخ می‌دهد. دانه‌هایی که به مقدار کافی انرژی جذب کرده‌اند، دارای تکه‌هایی ریز از نقره فلزی^{۱۰} که مراکز توسعه^{۱۱} نام دارند، هستند. وقتی فیلم نورخورده ظاهر شود، وجود تها یک مرکز توسعه در یک دانه هالاید نقره تغییر تمام دانه به نقره فلزی را تسریع می‌نماید. دانه‌هایی که مرکز توسعه

1) Views 2) Matching region 3) Exposure 4) Abrasion marks 5) Silver halide 6) Film base
7) Film base 8) Cellulose triacetate 9) Backing layer 10) Metallic silver 11) Development center

۶۰۲ فیلم عکاسی ۱۰۳



شکل ۲۳۰.۲ ساختمان یک فیلم سیاه-سفید امروزی

نداشته باشد، مشمول چنین تغییری قرار نمی گیرند. پس از ظهور با رفع شیمیایی باقیمانده دانه های هالید نقره، فیلم ثابت می شود. هر چه نور بیشتری به یک ناحیه فیلم برسد، هالید نقره بیشتری قابل توسعه می شود و در نتیجه نقره چگالتی در آن ناحیه نه تنین می شود. چون دانه های نقره در فرکانس های نوری بسیار کدر هستند، وقتی سطوح روشنایی معکوس شوند، تصویری با سطوح رنگی خاکستری بدست آید، بنابراین فیلم منفی تولید می گردد.

این فرآیند یک بار دیگر تکرار می شود تا عکس مثبت بدست آید. همانند فرآیند به کار رفته در مورد فیلم اولیه، فیلم منفی نیز روی کاغذ حساسی که آغشته به محلول هالید نقره است، تصویر می شود. با تابش یک منبع نور تصویری پنهان از فیلم منفی بدست می آید. پس از ظهور، این کاغذ دارای یک تصویر نقره ای مثبت است.

۶۰۴ مشخصات فیلم

از موضوعات مورد توجه برای عکاس، تمایز^۱، سرعت، دانه ای بودن^۲ و قدرت تفکیک^۳ هستند. مخصوصاً در کثر اثربخشی این عوامل در کاربردهای تخصصی نظیر عکسبرداری از نتایج حاصل از سامانه پردازش تصویر، اهمیت ویژه ای دارد.

تمایز

فیلمهای با تمایز بالا، تفاوت های روشنی موجود در صحنه را به صورت تفاوت های بزرگ چگالی

1) Contrast

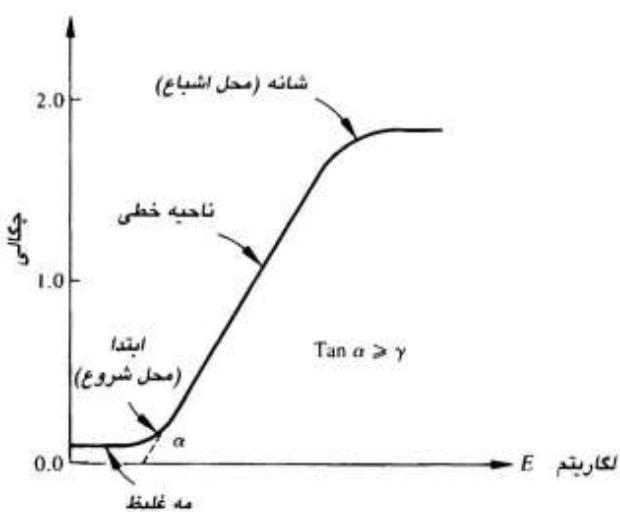
2) Graininess

3) Resolving power

به عکس منعکس می کنند؛ بر عکس فیلم های با تمایز پایین تفاوت های روشنی را به صورت تفاوت های کوچک چگالی به عکس منتقل می کنند. نورخوردگی E که فیلم با آن متأثر می شود، در هر نقطه از سطح حساس به نور به عنوان انرژی در واحد سطح تعریف می شود. مقدار نورخوردگی به I ، شدت تابشی، T ، طول دوره نورخوردگی، بستگی دارد. این عوامل رابطه زیر را با هم دارند:

$$E = IT. \quad (1-6.2)$$

متداول ترین توصیف در مورد خواص نوری فیلم عکاسی، نمودار چگالی نقره انباشته شده روی فیلم بر حسب لگاریتم E می باشد که به منحنی های مشخصه، منحنی های D -log- E (چگالی بر حسب لگاریتم نورخوردگی) و منحنی های $H & D$ (بعد از هارت و دریفلد¹) که این روش را ارائه کردند) نیز مشهورند. شکل ۲.۲۴. یک منحنی نوعی $H & D$ برای فیلم منفی را نشان می دهد. وقتی نورخوردگی زیر یک سطح معین باشد، چگالی مستقل از نورخوردگی و برابر مقدار کمینه که مه غلیظ² نامیده می شود، می باشد. در ابتدای منحنی³، با افزایش نورخوردگی، چگالی افزایش می یابد. پس از آن ناحیه ای از منحنی است که در آن چگالی به طور خطی بال لگاریتم نورخوردگی افزایش می یابد. شبیه این ناحیه خطی



شکل ۲.۲۴. یک نمونه معمولی از منحنی $H\&D$

1) Hurter & Driffield

2) Gross fog

3) Toe of curve

۶.۲ فیلم هاکاسی ۱۰۵

گاما های فیلم (۷) نامیده می شود. در نهایت منحنی در ناحیه ای موسوم به شانه^۱ اشباع می شود و دیگر با افزایش نور خوردگی چگالی تغییر نمی کند. مقدار ۷ معیاری از تمایز فیلم است: هرچه شبیب تندتر باشد، تمایز بیشتر خواهد بود. فیلم های همه منظوره با تمایز متوسط، گاما هایی در حدوده ۷،۰ تا ۱ دارند. فیلم های با تمایز بالا گاما هایی در حدوده ۱،۵ تا ۱۰ دارند. به عنوان یک قاعده، فیلم های با گاما های نسبتاً کوچک، برای بازسازی شدت پیوسته استفاده می شوند. فیلم های با تمایز بالا برای نسخه برداری^۲ از دست نوشته ها^۳ و سایر کارهای تخصصی استفاده می شوند.

سرعت

سرعت فیلم تعیین می کند که چه مقدار نور برای تولید مقدار معینی نقره در فرآیند توسعه مورد نیاز است. هرچه سرعت کمتر باشد، فیلم باید مدت طولانی تری در معرض نور قرار گیرد تا یک تصویر معین را ضبط نماید. متداول ترین استاندارد سرعت، معیار^۴ ASA است. این معیار، عددی^۵ است و مقدار سرعت تناسب مستقیم با حساسیت فیلم دارد. فیلم ASA200 دو برابر سریعتر از فیلم ASA100 است (و برای یک موضوع معین به نصف نور خوردگی آن نیاز دارد). چند معیار سرعت نظری سامانه DIN که در اروپا استفاده می شود، لگاریتمی هستند. هر افزایش^۳ واحدی در عدد سرعت DIN، سرعت واقعی را دو برابر می کند. فیلم ASA50 معادل DIN18 است، ASA100 معادل DIN21 است و به همین ترتیب تا آخر. فیلم های همه منظوره برای عکسبرداری در فضای باز و گاهی در فضای بسته سرعت هایی بین ASA80 و ASA160 دارند؛ فیلم های دانه ریز برای حداکثر دقیق تصویر سرعتی بین ASA20 و ASA64 دارند؛ فیلم های سریع برای نور کم و عکسبرداری در فضای بسته سرعتی بین ASA200 و ASA500 دارند؛ و فیلم های فوق سریع برای نور بسیار کم سرعت ASA650 و بیشتر دارند.

دانه ای بودن

تصویر حاصل از بلورهای هالید نقره ساختاری ناپیوسته دارد که در بزرگنمایی زیاد ظاهر دانه ای دارد. این اثر در فیلمهای سریع که نسبتاً بلورهای بزرگی دارند، بسیار مشخص است. بنابراین در کاربردهایی که جزئیات ریز مطلوب است یا وقتی بزرگ کردن فیلم های منفی مورد نیاز است، محلولهای دانه ریز و کندتر ترجیح داده می شوند.

1) Shoulder

2) Copying

3) Line originals

4) Scale

5) Arithmetic

۱۰۶ پردازش تصویر رقمه

قدرت تفکیک

ریزی جزئیاتی که یک فیلم می تواند تفکیک نماید، نه تنها به دانه ای بودنش بلکه به خواص پخش نور محلول و نیز به تمایزی که فیلم جزئیات ریز را بازسازی می کند، بستگی دارد. فیلم های دانه ریز با محلولهای رقیق بالاترین قدرت تفکیک را دارند.

۳.۶.۲ تنظیم های دیافراگم^۱ و دریچه^۲

بدون توجه به نوع فیلم مورد استفاده، تنظیم مناسب دوربین نیز برای بدست آوردن عکسهای مورد قبول ضروری است. تنظیمهای اصلی، دیافراگم عدسی و سرعت دریچه هستند.

در دیافراگم عدسی، برای کنترل مقدار نور عبوری از عدسی به فیلم، مجموعه ای از برگه ها^۳، اندازه سوراخ را کم و زیاد می نمایند. طوفه^۴ کنترل دیافراگم با معیاری معروف به اعداد f یا اعداد توقف تنظیم می شود که این اعداد از میان دنباله ای عددی نظیر ۱/۴، ۱/۸، ۲، ۲/۸، ۴، ۵/۶، ۸، ۱۱، ۱۶، ۲۲ و ۳۲ انتخاب می شوند. اعداد f با مقدار نور دریافتی تناسب معکوس دارند. بازای هر مقدار تنظیم f در دنباله فوق، دو برابر عدد f بزرگتر بعدی و نصف عدد f کوچکتر قبلی نور دریافت می شود (بنابراین به ترتیب دو برابر و نصف آنها نور خوردنگی دارد). همچنین تنظیم سرعت دریچه در دوربین های امروزی از یک استاندارد دنباله ای دو برابر - یا - نیم پیروی می کند. سرعت های متداول ۱، $\frac{1}{2}$ ، $\frac{1}{4}$ ، $\frac{1}{8}$ ، $\frac{1}{15}$ ، $\frac{1}{30}$ ، $\frac{1}{60}$ ، $\frac{1}{125}$ ، $\frac{1}{250}$ ، $\frac{1}{500}$ ، $\frac{1}{1000}$ ثانیه هستند. هرچه سرعت دریچه بیشتر باشد، زمان نور خوردنگی کمتر می شود.

دیافراگم و دریچه با تنظیم شدت و زمان عبور نور، مقدار نوری که به فیلم می رسد، را کنترل می کنند. بنابراین بعضی ترکیبات متفاوت روزنه^۵ - سرعت دریچه ممکن است نور خوردنگی برابری بدene. مثلاً دیافراگم $\frac{f}{2.8}$ با $\frac{1}{125}$ ثانیه، $\frac{f}{5}$ با $\frac{1}{25}$ ثانیه، $\frac{f}{8}$ با $\frac{1}{50}$ ثانیه، همه نور خوردنگی یکسانی می دهند. با این حال، انتخاب این دو عامل، مستقل از شرایط محیط عکاسی یا مشخصات فیلم نیست^۶. مثلاً

1) Diaphragm

2) Shutter

3) Leaves

4) Diaphragm control ring

5) Aperture

۶) وقتی T طولانی (مثلاً بزرگتر از یک ثانیه) باشد، باید پدیده ای با نام "خرابی متقابل"^۱ را ملاحظه آوریم که در آن معادله

(۲) صدق نمی کند و باید با استفاده از جداول ویژه فیلم، تنظیمهای صحیح برای نور خوردنگی را تعیین کرد.

1) Reciprocity failure

٧.٢ نکات پایانی ١٠٢

هنگام عکاسی از صحنه‌ای که در آن عمق تمرکز مورد نظر است، عکاس باید ^۱- توقف را تا حد ممکن بزرگ انتخاب کند تا این که به عدسی مشخصه سوراخ سنجاقی ^۲ بدهد. برای یک فیلم معین، این نیاز، بازه سرعت‌های دریچه که نور خوردنگی کافی می‌دهند، را محدود می‌کند. در کاربردهای دیگر، سرعت دریچه ملاحظه ضروری است. یک مثال که در پردازش تصویر کاربرد دارد، عکسبرداری از صفحه تلویزیون است. در این مورد، سرعت دریچه باید کمتر از نرخ تجدید تصویر در تلویزیون ($\frac{1}{3}$ ثانیه در قاب) باشد تا اثر ناهمزمانی دریچه با سیگنال‌های کنترل نمایش تلویزیونی را جبران نماید. معمولاً $\frac{1}{8}$ ثانیه کافی است، گرچه برای انگرال‌گیری از قابهای متواالی، اغلب از سرعت‌های کمتر استفاده می‌شود. به عنوان مثال بسیاری از عکس‌های این کتاب، در $\frac{1}{4}$ ثانیه با فیلم دانه‌ریز Kodak Panatomic-X (ASA32) عکسبرداری شده‌اند. تنظیمهای دیافراگم با استفاده از تابش سنج که شدت نور تصویر را اندازه می‌گیرد، تعیین شده‌اند.

٧.٣ نکات پایانی ^۳

مطلوب این فصل بیشتر اطلاعات مقدماتی لازم برای فصول بعدی هستند. گرچه بررسی ما در مورد سامانه بینایی انسان مختصر است، اطلاعات پایه‌ای از توانایی‌های چشم در دریافت اطلاعات تصویری را بدست می‌دهد. به طور مشابه مدل تصویر که در بخش ۲. ۲ ارائه شد، در فصل ۴ به عنوان مبنای یک روش ارتقای تصویر با نام فیلتر کردن هم شکلی ^۴ بکار می‌رود.

روش‌های نمونه برداری که در بخش ۲. ۳ معرفی شدند، را مجدد در بخش ۳. ۹. ۳. ۲ بعد از ارائه ابزار ریاضی لازم برای مطالعه تحلیلی عمیقتر آنها بررسی می‌کنیم. همچنین در فصل ۶ و در مبحث رمزگذاری تصویر که مسأله فشرده سازی مقادیر زیاد داده حاصل از رقمی کردن تصویر مطرح است، نمونه برداری و چندی کردن نقش اصلی را بازی می‌کنند.

مطلوب بخش ۲. ۴ در فهم روش‌های متعدد پردازش تصویر که در فصول بعدی بیان می‌شوند، پایه‌ای است. مفاهیم هندسه تصویربرداری که در بخش ۲. ۵ ارائه شدند، در موقعي که باید اطلاعات صحنه سه بعدی را با تصاویر حاصل از دوربین همبسته و سپس بوسیله رایانه پردازش کرد، نقش مهمی بازی می‌کنند.

1) f-stop

2) Pin-hole

3) Concluding remarks

4) Homomorphic filtering

مراجع

مطلوب ارائه شده در بخش های ۲.۱.۱ و ۲.۱.۲ بیشتر براساس تألیفات Connsweet [1970] و Graham [1965] می باشد. منابع مطالعه بیشتر در مورد بخش ۲.۱.۳ را می توان در تألیفات Stevens [1951]؛ Sheppard [1968]؛ Sheppard, Stratton [1970]؛ Gazley [1969] و Sheppard, Stratton [1970] پیدا کرد. مدل تصویری ارائه شده در بخش ۲.۰.۲ بوسیله' Oppenheim, Schafer و [1968] Stockham در ارتباط با کاربردهای ارتقای تصویر مورد بررسی قرار گرفته است. مراجع مورد استفاده در این بخش در مورد مقادیر انعکاس و روشنایی، [1961] Moon و [1972] IES Lighting Handbook است. بعضی از مطالب ارائه شده در بخش ۲.۳ براساس کار [1965] Huang است. مقالات Gaven, Tavitian و Scoville [1965]، Toriwaki [1970] Harabedian [1970] نیز قابل توجه هستند. مطالب اضافی در مورد بخش ۲.۴ را می توان در Kak [1982] و Rosenfeld [1989] پیدا کرد. به Jain [1989] و Schalkoff [1989] نیز و همکارانش [1979] در Lee [1987] Fu, Gonzalez و [1987] است. مراجع بخش ۲.۶ مراجعه کنید. بخش ۲.۵، از کتاب Kodak Plates and Films for Scientific، Nelson [1971]، Perrin [1960]، Mees [1966]، Langford [1984] و Photography [1973] هستند.

مسائل

۱.۲ با استفاده از اطلاعات زمینه ای ارائه شده در بخش ۲.۰ و تفکر هندسی محض، قطر کوچکترین نقطه چاپی را که چشم می تواند روی صفحه ای به فاصله ۲۰ متر از چشم تشخیص دهد، تخمین بزنید. برای سادگی فرض کنید که وقتی تصویر نقطه مذکور روی لکه زرد شبکیه، کوچکتر از قطر یک گیرنده در آن ناحیه (گیرنده مخروطی) بشود، سامانه بینایی نمی تواند آن را تشخیص دهد. همچنین فرض کنید که لکه زرد را می توان با یک آرایه مرتبی با ابعاد $1 \times 1, 5$ میلی متر مدل نمود که مخروطها و فاصله های بین آنها به طور یکنواخت در سراسر آن توزیع شده اند.

۲.۲ فرض کنید که سطح تختی به مرکز (x_0, y_0) به وسیله' یک منبع نور با توزیع شدت

$$i(x, y) = Ke^{-[(x-x_0)^2 + (y-y_0)^2]}$$

روشن شود. فرض کنید که ضریب انعکاس سطح یک و $K = 255$ است. اگر تصویر حاصل

١٠٩ مسائل

با m بیت رقمه شود، و چشم بتواند تغییرات ناگهانی با بلندی Δ سطح شدت را بین پیکسل های مجاور تشخیص دهد، چه مقداری از m باعث منحنی سازی اشتباه^۱ قابل مشاهده می شود؟

تصویر مسأله ۲ . ۲ را برای $m=2$ رسم کنید. ۳۰۴

یک معیار متدال برای ارسال داده های رقمه نرخ نماد^۲ است که به عنوان تعداد بیت های ارسال شده در یک ثانیه تعریف می شود. عمل ارسال معمولاً در بسته هایی که مشکل از یک بیت شروع، یک بایت (۸ بیت) اطلاعات و یک بیت پایان است، انجام می شود. با استفاده از این روش ارسال ، به سوالات زیر پاسخ دهید:

(الف) چند دقیقه طول می کشد که تصویر 256×512 با 256 سطح خاکستری در نرخ بیت 300 ارسال شود؟

ب) در نرخ بیت 9600 چقدر طول می کشد؟

پ) قسمتهای (الف) و (ب) را برای تصویر 1024×1024 با 256 سطح خاکستری تکرار کنید.

دو زیر مجموعه تصویری S_1 و S_2 که در شکل زیر دیده می شود، را در نظر بگیرید. ۵۰۴

	S_1	S_2	
0	0 0 0 0	0 0 1 1	0
1	0 0 1 0	0 1 0 0	1
1	0 0 1 0	1 1 0 0	0
0	0 1 1 1	0 0 0 0	0
0	0 1 1 1	0 0 1 1	1

برای $V=1$ ، تعیین کنید آیا S_1 و S_2 (الف) متصل چهارگانه، (ب) متصل هشتگانه، یا (پ) متصل m گانه اند؟

الگوریتمی برای تبدیل مسیر هشتگانه با ضخامت یک پیکسل به مسیر چهارگانه ارائه کنید. ۶۰۴

الگوریتمی برای تبدیل مسیر m گانه با ضخامت یک پیکسل به مسیر چهارگانه ارائه کنید. ۷۰۴

فرض کنید که رابطه R در یک کاربرد خاص به معنای "اتصال چهارگانه" باشد، طوری که

1) False contouring

2) baud rate

۱۱۰ پردازش تصویر رقمی

sRt بدین معنا باشد که s متصل چهارگانه به t است. تمام آرایشها هندسی که از ماتریس

$$\mathbf{B}^* = b \begin{bmatrix} a & b & c \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

استنباط می شوند، را رسم کنید.

ثابت کنید که الگوریتم بخش ۲.۴. برای محاسبه ماتریس \mathbf{B}^+ ، در واقع همان نتیجه معادله $(1-4.2)$ را می دهد.

قطعه تصویر زیر را در نظر بگیرید.

الف) اگر $V = \{1, 2, 3\}$ باشد، فواصل D_1, D_2, D_3 بین p و q را محاسبه کنید.

ب) برای $V = \{1, 2\}$ نیز تکرار نمایید.

$$\begin{array}{cccc} 3 & 1 & 2 & 1(q) \\ 2 & 2 & 0 & 2 \\ 1 & 2 & 1 & 1 \\ (p)1 & 0 & 1 & 2 \end{array}$$

الف) نشان دهید که فاصله D_i بین دو نقطه p و q برابر کوتاهترین مسیر چهارگانه بین این نقاط است.

ب) آیا این مسیر یکنائب؟

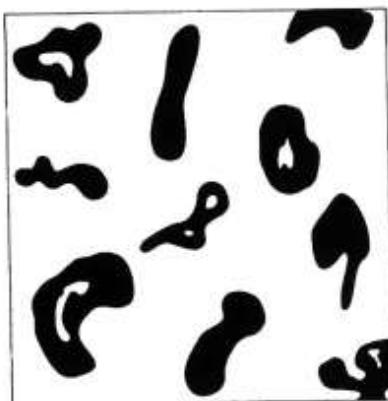
نقاب 3×3 زیر معمولاً برای محاسبه مشتق تصویر در هر نقطه درجهت محور x استفاده می شود.

$$\begin{array}{ccc} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{array}$$

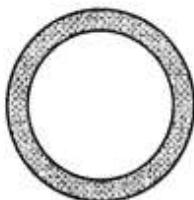
یک روال ALU برای اجرای این عمل ارائه کنید.

یک کاربرد فناوری زیستی¹ تصاویر دودویی لکه ها² با شکل کلی صفحه بعد را بدست می دهد. روشی ارائه کنید که (الف) تعیین کند آیا تصویر دارای لکه است و اگر پاسخ مثبت است، (ب) هر لکه ای که درونش سوراخ ندارد، در نوع A و در غیر این صورت در نوع B دسته بندی کند. روشن شما باید بتواند لکه هایی که با مرز تصویر تماس دارند، را تشخیص

داده و آنها را در نظر نگیرد.



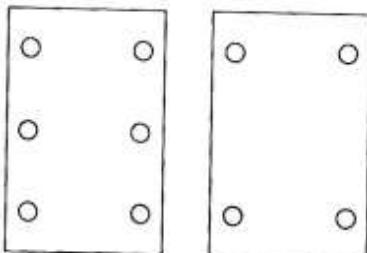
- ۱۴.۲ یک کارخانه^{۱)} صنعتی پیشرفته در یک مناقصه دولتی برای ساخت واشرهای خیلی دقیق با شکل زیر، برنده می شود.



طبق قرارداد باید شکل تمام واشرها بوسیله^{۱)} یک سامانه تصویربردار بازبینی شود. شما به عنوان مشاور استخدام شده اید تا در طراحی این سامانه کمک کنید. راه حلی براساس اعمال منطقی پیشنهاد کنید. فرض کنید که به سادگی می توان از هر واشر تصویری دودویی بدست آوردن و شخص دیگری در زمینه^{۱)} مکانیزم های مورد نیاز برای حرکت واشرها، دقیق قرار دادن واشرها در محل بازبینی و سایر امور کارمی کند.

- ۱۵.۲ تولید کننده ای دو نوع لولای درب که در زیر نشان داده شده اند، را تولید می کند. وی مشکلی جدی دارد. دو نوع لولای مذکور به جز از نظر تعداد سوراخ هایکسان هستند. کارگران گاهی به اشتباه لولاهای با تعداد سوراخ متفاوت را با هم بسته بندی می کنند. گاهی نیز لولاهایی که به طور صحیح تولید نشده اند یا به تعداد صحیح سوراخ ندارند، را تشخیص نمی دهند.

1) Plant



صاحب این کارگاه تصمیم می گیرد که فرآیند تفکیک را به طور کامل خودکار نماید و از موفقیت شما در حل مسأله واشر نیز خبر دارد. او شمارا به عنوان مشاور استخدام می کند تا یک سامانه بازبینی مبتنی بر تصویربرداری طراحی کنید. شخص دیگری نیز روباتی طراحی می کند که قادر است لولا را دقیقاً در محل مناسب برای تصویربرداری به وسیله دوربین تلویزیونی قرار دهد. تمام آنچه سامانه تصویربرداری باید انجام دهد عبارتند از: ۱- ضبط تصویر، ۲- تعیین این که آیا لولا از نوع A (سوراخه)، نوع B (۴ سوراخه) یا نوع C (لولای غیرقابل مصرف به دلیل تعداد غیر صحیح سوراخها) است و ۳- فرستادن یکی از این سه علامت به روبات. آنگاه روبات با توجه به نوع علامت تصمیم می گیرد که لولا را بسته بندی و یا حذف نماید. راه حلی براساس آنچه تاکنون در این کتاب یاد گرفته اید، پیشنهاد کنید. پیشنهادهای شما به صاحب شرکت باید به شکل نمودار بلوکی و همراه با شرح مختصری درباره وظیفه هر بلوك باشد.

۱۶.۴ الف) ماتریس تبدیل مورد استفاده در چرخاندن تصویر به اندازه ۴۵ درجه در جهت عقربه ساعت را بتویسید.

ب) چگونه از این تبدیل باید استفاده کرد تا چرخش مطلوب تصویر بدست آید؟

۱۷.۲ پ) از ماتریس بدست آمده در (الف) جهت چرخاندن نقطه $(x, y) = (1, 0)$ تصویر استفاده کنید. تعیین کنید آیا نقطه جهانی با مختصات $(\sqrt{2}/2, \sqrt{2}/2)$ روی محور نوری یک دوربین مستقر در نقطه $(0, 0)$ که دارای زوایای پن و تیلت 135° است، قرار دارد. فرض کنید که عدسی ۵۰ میلی متری است و $r_1 = r_2 = r_3 = 0$.

۱۸.۲ با استفاده از معادله $(2.41-5)$ ، معادلات $(2.42-5)$ و $(2.43-5)$ را بدست آورید.

۱۹.۲ روند ALU در بخش ۴.۶ را طوری تغییر دهید که هر پیکسل تصویر را با متوسط چهار همسایه اش جایگزین نماید. خود پیکسل را در محاسبه متوسط وارد نکنید.

۲۰.۴ کارخانه ای دارای خط تولید مربعات ریز پلیمری شفاف می باشد. بدلیل نیاز شدید به کیفیت، بازبینی چشمی ۱۰۰ درصد یک الزام است و مدیر کارخانه می داند که استفاده از تیروی انسانی برای بازبینی، هزینه^۱ مضاعفی دارد. لذا بازبینی نیمه خودکار اجرا می شود. در هر جایگاه بازبینی یک مکانیزم مکانیکی، تک تک پلیمرها را روی چراغی که بالای آن یک سامانه نوری قرار دارد، می گذارد و سامانه نوری تصویر بزرگ شده ای از مربع پلیمری تولید می کند. تصویر صفحه نظاره^۲ غیر انعکاسی با ابعاد 80×80 میلی متر را پر می کند. نواقص به صورت لکه های گردسیاه ظاهر می شوندو وظیفه بازبین نظاره صفحه و کنار گذاشتن نمونه هایی است که یک یا چند لکه با قطر 8 mm میلی متر یا بزرگتر (در مقیاس صفحه^۱ نمایش) دارند. مدیر می داند که اگر بتواند راهی برای خودکار کردن کامل فرآیند بیابد، بهره 50 درصد افزایش می باید. همچنین می داند که این موفقیت به او کمک می کند تا از نرده بان ترقی در صنف خودش بالا برسد. پس از بررسی زیاد، او مطمئن می شود که راه حل مسئله نظاره^۲ صفحه بازبینی با یک دوربین تلویزیونی CCD و دادن خروجی دوربین به سامانه پردازش تصویری است که قادر به تشخیص لکه ها، اندازه گیری قطر آنها و فعال کردن دکمه های قبول/رد که قبلاً توسط بازبین به کار می افتدند، باشد. او می تواند سامانه ای پیدا کند که این کار را انجام دهد، البته با این شرط که کوچکترین خرابی، یک ناحیه حداقل 2×2 پیکسلی را در تصویر رقمه اشغال نماید. مدیر، شمارا استفاده می کند تا با تعیین دوربین و عدسی به او کمک کنید؛ اما با این شرط که تنها از قطعات موجود در بازار^۳ استفاده کنید. در مورد عدسی ها، فرض کنید این شرط بدین معنی است که تنها اعداد صحیح مضرب 25 یا 35 میلی متر حداقل تا 200 میلی متر مجاز هستند. در مورد دوربین، نیز شرط بدین معنی است که تفکیک، 512×512 یا 1024×1024 یا 2048×2048 پیکسل است. عناصر تصویربردار در این دوربین ها مربعهایی با ابعاد 8×8 میکرون و فضای بین عناصر تصویربردار نیز مربع هایی به همان اندازه هستند. دوربین ها از عدسی ها گرانترند؛ بنابراین بهتر است که مسئله توسط دوربینی با کمترین تفکیک ممکن و بر اساس انتخاب عدسی ها حل شود. شما به عنوان مشاور باید پیشنهاد نامه ای بنویسید که به تفصیل کافی تحلیلی که شما را به این نتیجه رسانده است، رایان نماید.

فصل ۳

تبديل های تصوير

« خود را همشکل روزگار نسازید: اما
با تجدید افکار، خود را تغییر دهید... »
Romans 12:2

این فصل بیشتر به بیان تبدیل های دو بعدی و خواص آنها می پردازد. نظریه تبدیل سالهای است که نقشی کلیدی در پردازش تصویر بازی می کند و هم اکنون نیز چه از لحاظ نظری و چه عملی موضوع جذابی در این مبحث است. در فصول بعدی از تبدیل های دو بعدی برای ارتقا، بازیابی، رمزنگاری و توصیف تصویر استفاده می کنیم.

به دلیل کاربرد وسیع تبدیل فوریه در مسائل پردازش تصویر، در این فصل بر آن تأکید نموده و سایر تبدیلات را نیز با مقداری تفصیل بیان می کنیم. بخش ۱.۳ تبدیل فوریه با یک یادو متغیر پیوسته را معرفی می کند. آنگاه بخش ۲.۲ این مفاهیم را در شکل گسته بیان می کند. بخش ۳.۳ نیز چند خاصیت مهم تبدیل فوریه دو بعدی را بیان می کند. بخش ۳.۴ الگوریتم محاسبه سریع تبدیل فوریه را معرفی می کند. از این الگوریتم به منظور کاهش تعداد محاسبات به کسری از آنچه در پیاده سازی معمولی تبدیل فوریه نیاز است، استفاده می شود. بخش ۳.۵ درباره تبدیل های والش^۱، هادامارد^۲، کسینوسی گسته^۳، هaar^۴ و اسلنت^۵ است.

۱.۳ مقدمه ای بر تبدیل فوریه

اجازه دهید که $f(x)$ تابع پیوسته ای از متغیر حقیقی x باشد. تبدیل فوریه $f(x)$ که با $\{f(x)\}$ نشان

1) Walsh

2) Hadamard

3) Discrete cosine

4) Haar

5) Slant

۱۱۶ پردازش تصویر رقمه

داده می شود، به صورت زیر تعریف می شود:

$$\tilde{f}\{f(x)\} = F(u) = \int_{-\infty}^{\infty} f(x) \exp[-j2\pi ux] dx \quad (1-1.3)$$

$$j = \sqrt{-1}$$

با داشتن $F(u)$ ، می توان $f(x)$ را با استفاده از تبدیل فوریه معکوس بدست آورد:

$$\begin{aligned} \tilde{f}^{-1}\{F(u)\} &= f(x) \\ &= \int_{-\infty}^{\infty} F(u) \exp[j2\pi ux] du. \end{aligned} \quad (2-1.3)$$

معادلات (۱-۱.۳) و (۲-۱.۳)، که زوج تبدیل فوریه نامیده می شوند، در صورتی برقرار هستند که $f(x)$ پیوسته و انتگرال پذیر باشد. در عمل تقریباً همیشه این شرایط برقرار هستند.

در مراسر این کتاب، تنها روی توابع حقیقی کار می کنیم. با این حال تبدیل فوریه یکتابع حقیقی معمولاً مختلط است؛ یعنی،

$$F(u) = R(u) + jI(u) \quad (3-1.3)$$

که $R(u)$ و $I(u)$ به ترتیب مؤلفه های حقیقی و موهومی $F(u)$ هستند. اغلب مناسب تر است که معادله (۳-۱.۳) به صورت نمایی بیان شود، یعنی،

$$F(u) = |F(u)| e^{j\phi(u)} \quad (4-1.3)$$

که

$$|F(u)| = [R^2(u) + I^2(u)]^{1/2} \quad (5-1.3)$$

و

$$\phi(u) = \tan^{-1} \left[\frac{I(u)}{R(u)} \right]. \quad (6-1.3)$$

۱۱۷ ۱.۳ مقدمه ای بر تبدیل فوریه

طیف فوریه $f(x)$ و $\phi(u)$ زاویه فاز آن نامیده می شود. مربع طیف، $|F(u)|^2$

$$\begin{aligned} P(u) &= |F(u)|^2 \\ &= R^2(u) + I^2(u) \end{aligned} \quad (۷-۱.۳)$$

معمولًا طیف قدرت (x) نامیده می شود.

متغیر u که در تبدیل فوریه ظاهر می شود، اغلب متغیر فرکانس نامیده می شود، این نام از نمایش اول ر عبارت نمایی $\exp[-j2\pi ux]$ ، یعنی

$$\exp[-j2\pi ux] = \cos 2\pi ux - j \sin 2\pi ux. \quad (۸-۱.۳)$$

ناشی می شود. در معادله (۳.۱-۱)، تعبیر انتگرال به عنوان جمع حدی اجزای گسته، روشن می سازد که $F(u)$ از جمع تعداد نامحدودی مؤلفه های سینوسی و کسینوسی حاصل شده است و مقدار A فرکانس زوج سینوس - کسینوس متناظر را معین می کند.

مثال: تابع ساده شکل ۳.۱(الف) را در نظر بگیرید. تبدیل فوریه آن با استفاده از معادله (۳.۱-۱) به صورت زیر بدست می آید:

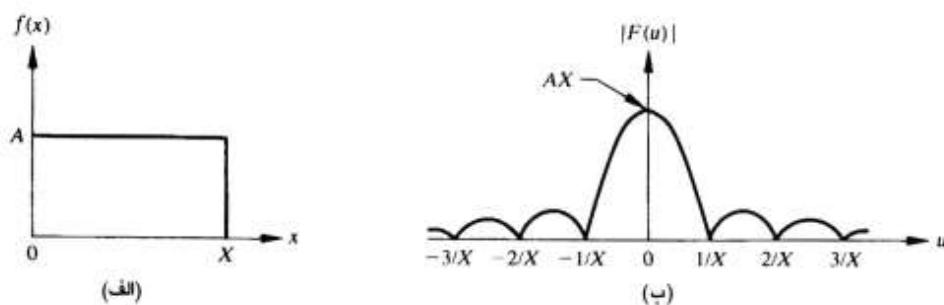
$$\begin{aligned} F(u) &= \int_{-\infty}^{+\infty} f(x) \exp[-j2\pi ux] dx \\ &= \int_0^X A \exp[-j2\pi ux] dx \\ &= \frac{-A}{j2\pi u} [e^{-j2\pi ux}]_0^X = \frac{-A}{j2\pi u} [e^{-j2\pi uX} - 1] \\ &= \frac{A}{j2\pi u} [e^{j2\pi uX} - e^{-j2\pi uX}] e^{-j2\pi uX} \\ &= \frac{A}{\pi u} \sin(\pi uX) e^{-j2\pi uX} \end{aligned}$$

تابعی مختلط است. طیف فوریه آن عبارتست از:

$$\begin{aligned} |F(u)| &= \left| \frac{A}{\pi u} \right| |\sin(\pi uX)| |e^{-j2\pi uX}| \\ &= AX \left| \frac{\sin(\pi uX)}{(\pi uX)} \right|. \end{aligned}$$

شکل ۳.۱(ب) ترسیمی از $|F(u)|$ را نشان می دهد.





شکل ۳.۱۰ یک تابع ساده و طیف فوریه آن

به آسانی می توان تبدیل فوریه را به تابع دو متغیره $f(x, y)$ نیز تعمیم داد. اگر $f(x, y)$ پیوسته و انتگرال پذیر و $F(u, v)$ نیز انتگرال پذیر باشد، زوج تبدیل فوریه زیر وجود خواهد داشت که u و v متغیرهای فرکانس هستند:

$$\tilde{\mathcal{F}}\{f(x, y)\} = F(u, v) = \iint_{-\infty}^{\infty} f(x, y) \exp[-j2\pi(ux + vy)] dx dy \quad (9-1.3)$$

$$\tilde{\mathcal{F}}^{-1}\{F(u, v)\} = f(x, y) = \iint_{-\infty}^{\infty} F(u, v) \exp[j2\pi(ux + vy)] du dv \quad (10-1.3)$$

مانند حالت یک بعدی، طیف فوریه، فاز و طیف قدرت به ترتیب عبارتند از:

$$|F(u, v)| = [R^2(u, v) + I^2(u, v)]^{1/2} \quad (11-1.3)$$

$$\phi(u, v) = \tan^{-1} \left[\frac{I(u, v)}{R(u, v)} \right] \quad (12-1.3)$$

$$P(u, v) = |F(u, v)|^2 = R^2(u, v) + I^2(u, v). \quad (13-1.3)$$

مثال: تبدیل فوریه تابع شکل ۳.۲.(الف) به صورت زیر بدست می آید:

$$\begin{aligned} F(u, v) &= \iint_{-\infty}^{\infty} f(x, y) \exp[-j2\pi(ux + vy)] dx dy \\ &= A \int_0^X \exp[-j2\pi ux] dx \int_0^Y \exp[-j2\pi vy] dy \end{aligned}$$

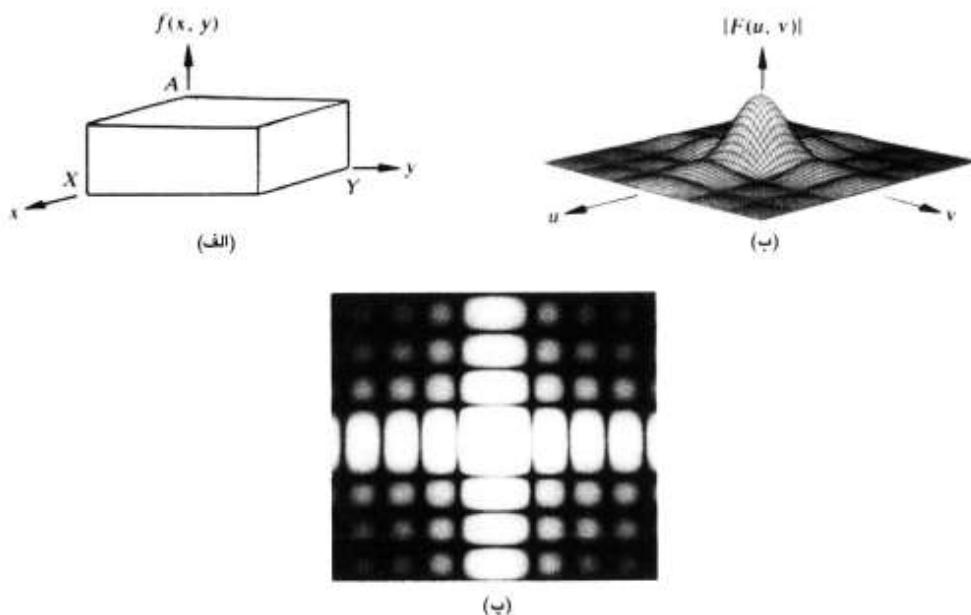
۱۱۴ ۱.۳ مقدمه ای بر تبدیل فوریه

$$\begin{aligned}
 &= A \left[\frac{e^{-j2\pi uX}}{-j2\pi u} \right]_0^X \left[\frac{e^{-j2\pi vY}}{-j2\pi v} \right]_0^Y \\
 &= \frac{A}{-j2\pi u} [e^{-j2\pi uX} - 1] \frac{1}{-j2\pi v} [e^{-j2\pi vY} - 1] \\
 &= AX Y \left[\frac{\sin(\pi uX)}{(\pi uX)} e^{-j\pi uX} \right] \left[\frac{\sin(\pi vX)}{(\pi vX)} e^{-j\pi vX} \right].
 \end{aligned}$$

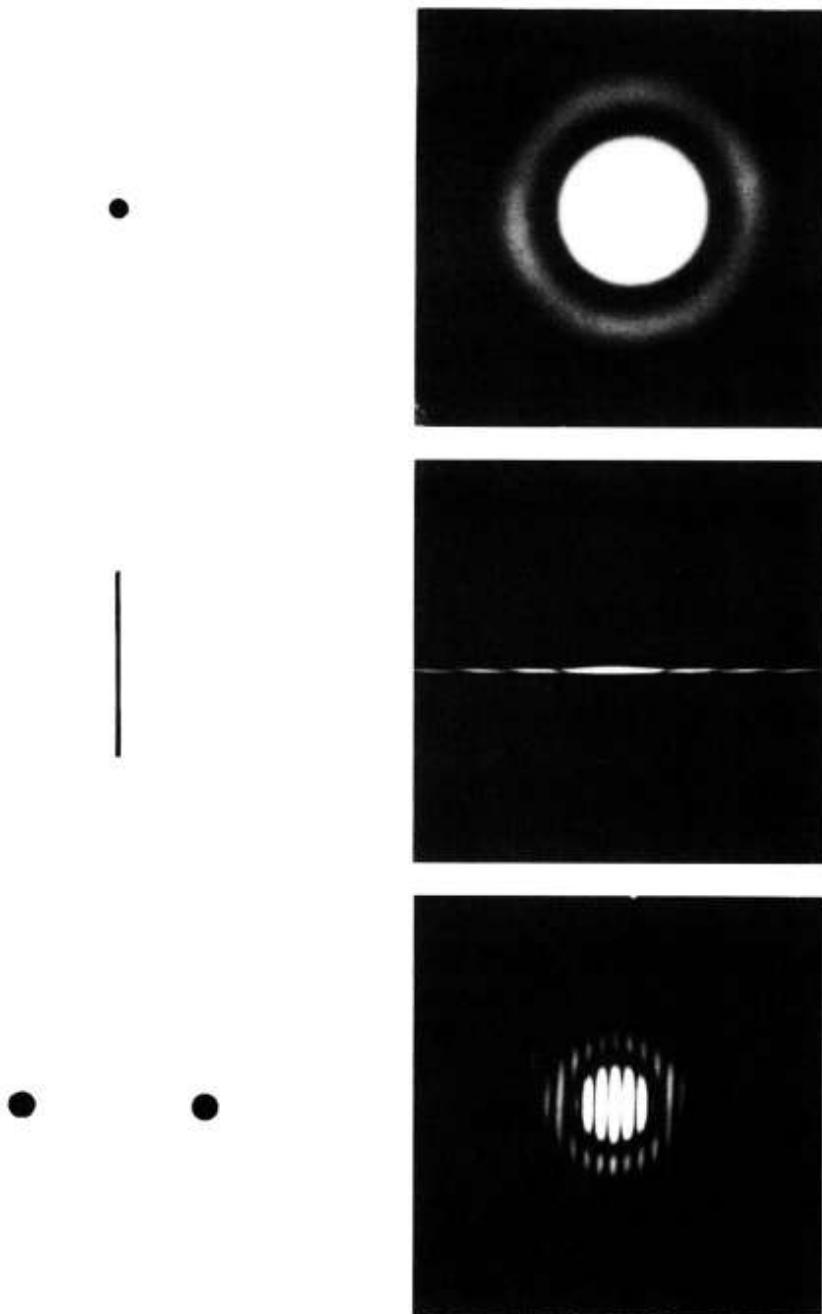
طیف فوریه آن عبارتست از:

$$|F(u, v)| = AX Y \left| \frac{\sin(\pi uX)}{(\pi uX)} \right| \left| \frac{\sin(\pi vX)}{(\pi vX)} \right|.$$

شکل ۲.۳ (ب) رسم این تابع در نمای سه بعدی را نشان می دهد. شکل ۲.۳ (ب) طیف را به صورت تابع شدت نمایش می دهد که روشنایی آن متناسب با $|F(u, v)|$ است. شکل ۲.۳ (ب) مثالهای دیگری از توابع دو بعدی و طیف آنها را نمایش می دهد. در این مورد $f(x, y)$ و $|F(u, v)|$ هر دو با تصویر نشان داده شده اند.

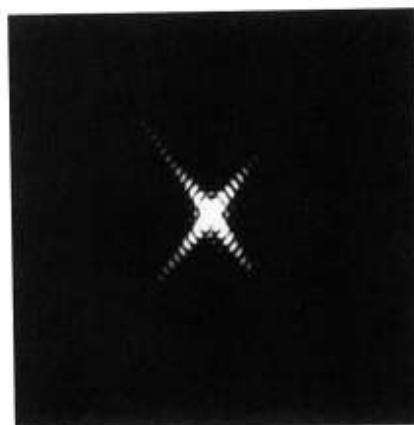
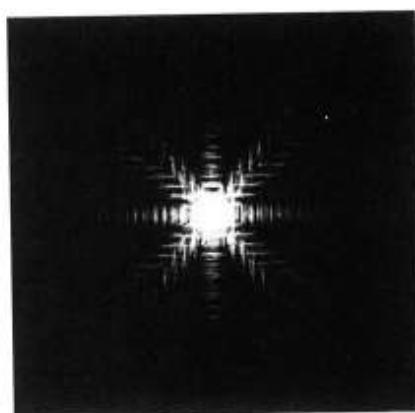
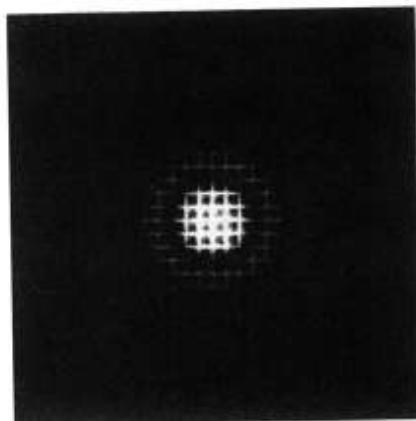
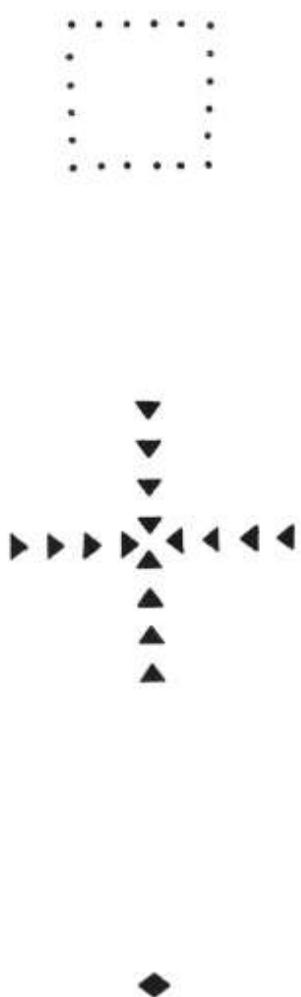


شکل ۲.۳ (الف) یک تابع دو بعدی؛ (ب) طیف فوریه آن؛ و (پ) همان طیف که به صورت تابع شدت نمایش داده شده است



شکل ۳۰-۳ چند تابع دو بعدی و طیف فوریه آنها

۱۰.۳ مقدمه ای بر تبدیل فوریه ۱۲۱



شکل ۱۰.۳ ادامه

۲.۳ تبدیل فوریه گستته

فرض کنید همانطور که در شکل ۳.۴ دیده می شود، تابع $f(x)$ با برداشتن N نمونه که به اندازه Δx با هم فاصله دارند، به دنباله'

$$\{f(x_0), f(x_0 + \Delta x), f(x_0 + 2\Delta x), \dots, f(x_0 + [N-1]\Delta x)\}$$

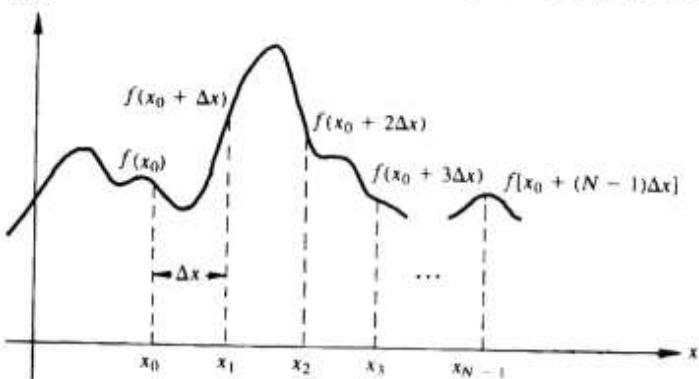
گستته شود. در ادامه جهت سهولت از x بسته به موضوع مورد بحث، به عنوان متغیر گستته یا پیوسته استفاده کنیم. انجام این کار نیازمند تعریف

$$f(x) = f(x_0 + x \Delta x) \quad (1-2.3)$$

است که در اینجا x مقادیر گستته $1, 2, \dots, N$ را می گیرد. به عبارت دیگر، دنباله N نمونه هم فاصله از تابع پیوسته مورد نظر دلالت دارد. با توجه به نمادگذاری فوق، زوج تبدیل فوریه گستته که برای توابع نمونه برداری شده به کار می روند، عبارتند از[†]:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \exp[-j2\pi ux/N] \quad (2-2.3)$$

$$u = 0, 1, 2, \dots, N-1 \quad \text{برای ۱}$$



شکل ۳.۴ نمونه برداری از تابع پیوسته

[†] ایات این روابط خارج از محدوده این کتاب است. برای اثبات روابط تبدیل فوریه پیوسته و گستته به Blackman و Lewis، Cooley و Tukey [1985]؛ Welch [1967] و Brigham [1974] مراجعه کنید.

۱۴۳ ۲.۲ تبدیل فوریه گسته

$$f(x) = \sum_{u=0}^{N-1} F(u) \exp[j2\pi ux/N] \quad (3-2.3)$$

برای $x = 0, 1, 2, \dots, N-1$

مقادیر $u = 0, 1, 2, \dots, N-1$ در تبدیل فوریه گسته، معادله (۲-۲.۳) متناظر با نمونه های تبدیل پیوسته در مقادیر $\Delta u = 0, \Delta u, 2\Delta u, \dots, (N-1)\Delta u$ هستند. به عبارت دیگر، $F(u)$ نمایانگر $F(u\Delta u)$ می باشد. این نمادگذاری شبیه آنچه در مورد $f(x)$ گسته استفاده شده می باشد، به جز این که نمونه های $F(u)$ از مبدأ محور فرکانس شروع می شوند. Δu و Δx به صورت زیر با هم رابطه دارند:

$$\Delta u = \frac{1}{N \Delta x} \quad (4-2.3)$$

در حالت دو متغیره، زوج تبدیل فوریه عبارتند از:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp[-j2\pi(ux/M + vy/N)] \quad (5-2.3)$$

برای $u = 0, 1, 2, \dots, M-1$ و $v = 0, 1, 2, \dots, N-1$

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \exp[j2\pi(ux/M + vy/N)] \quad (6-2.3)$$

برای $y = 0, 1, 2, \dots, N-1$ و $x = 0, 1, 2, \dots, M-1$

در اینجا نمونه برداری تابع پیوسته، در یک شبکه دو بعدی و با تقسیم بندی محورهای x و y به ترتیب با فاصله های Δx و Δy انجام می شود. مانند حالت یک بعدی، تابع گسته $f(x, y)$ ، نمونه های تابع $y = 0, 1, 2, \dots, N-1$ و $x = 0, 1, 2, \dots, M-1$ برای $f(x_0 + x\Delta x, y_0 + y\Delta y)$ می باشد. توضیحات مشابهی در مورد $F(u, v)$ صدق می کند. گام های نمونه برداری در میدان های مکان و فرکانس روابط زیر را با هم دارند:

$$\Delta u = \frac{1}{M \Delta x} \quad (7-2.3)$$

و

$$\Delta v = \frac{1}{N \Delta y} \quad (8-2.3)$$

وقتی تصاویر به صورت آرایه های مربعی نمونه برداری شوند، $N = M$ است و

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \exp[-j2\pi(ux + vy)/N] \quad (9-2. 3)$$

برای $u, v = 0, 1, 2, \dots, N-1$ و

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \exp[j2\pi(ux + vy)/N] \quad (10-2. 3)$$

برای $x, y = 0, 1, 2, \dots, N-1$ به وجود جزء $1/N$ در هر دو معادله (9-2. 3) و (9-2. 4) درجه شود. چون $f(x, y)$ یک زوج تبدیل فوریه هستند، محل قرار دادن آنها دلخواه است. در عمل معمولاً تصاویر به آرایه های مربعی رقمه می شوند، بنابراین غالباً از زوج تبدیل فوریه ارائه شده با معادلات (9-2. 3) و (10-2. 3) استفاده می کنیم. گاه گاهی وقتی بر کلی بودن ابعاد تصویر تأکید داشته باشیم، از معادلات (3. 5-2) و (3. 6-2) استفاده می کنیم.

طیف فوریه، فاز و طیف انرژی توابع یک بعدی و دو بعدی نیز به ترتیب با معادلات (3. 1. 3)-(3. 1. 5) و معادلات (3. 1. 11)-(3. 1. 13) بدست می آیند. تنها تفاوت این است که متغیرهای مستقل گسته اند.

برخلاف حالت پیوسته، برای تبدیل فوریه گسته شرایط وجود اصلاً بررسی نمی شود زیرا همواره وجود دارد. مثلاً در حالت یک بعدی می توان با قراردادن معادله (3. 2-2) در معادله (3. 2-2. 3) این مطلب را نشان داد:

$$\begin{aligned} F(u) &= \frac{1}{N} \sum_{x=0}^{N-1} \left[\sum_{r=0}^{N-1} F(r) \exp[j2\pi rx/N] \right] \exp[-j2\pi ux/N] \\ &= \frac{1}{N} \sum_{r=0}^{N-1} F(r) \left[\sum_{x=0}^{N-1} \exp[j2\pi rx/N] \exp[-j2\pi ux/N] \right] \quad (11-2. 3) \\ &= F(u). \end{aligned}$$

معادله (11-2. 3) از شرط تعامد نتیجه می شود، یعنی

$$\sum_{x=0}^{N-1} \exp[j2\pi rx/N] \exp[-j2\pi ux/N] = \begin{cases} N & \text{اگر } r = u \\ 0 & \text{در غیر این صورت} \end{cases} \quad (12-2. 3)$$

ذکر شود که در معادله (3. 3-2) تغییر متغیر از u به U نمادگذاری را واضح تر می کند.

۱۲۵ تبدیل فوریه گسته ۲.۳

همچنین با جایگذاری معادله (۳.۲-۲) در معادله (۳.۳-۲)، معادله ای در مورد $f(x)$ بدست می آید که نشان می دهد زوج تبدیل فوریه داده شده با این معادلات همواره وجود دارند. در مورد زوج تبدیل گسته دو بعدی نیز استنتاج مشابهی صدق می کند.

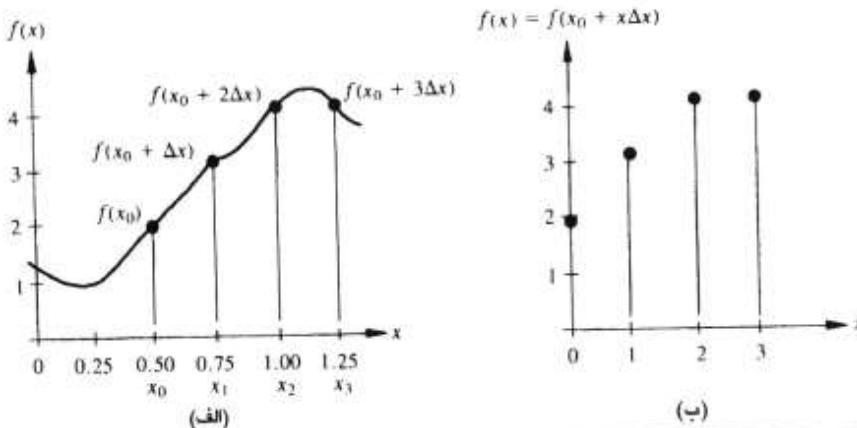
مثال: برای مصور ساختن معادلات (۳.۲-۲) و (۳.۳-۲)، تابع شکل ۳.۵(الف) را در نظر بگیرید. با نمونه برداری در مقادیر آرگومان $0.5 = x_0$ ، $x_1 = 0.75$ ، $x_2 = 1.0$ و $x_3 = 1.25$ و با تعریف مجلد آرگومان طبق توضیحات قبلی - تابع گسته شکل ۳.۵(ب) تولید می شود.

با کاربرد معادله (۳.۲-۲) روی چهار نمونه حاصل داریم:

$$\begin{aligned} F(0) &= \frac{1}{4} \sum_{x=0}^3 f(x) \exp[0] \\ &= \frac{1}{4} [f(0) + f(1) + f(2) + f(3)] \\ &= \frac{1}{4} (2 + 3 + 4 + 4) \\ &= 3.25 \end{aligned}$$

و

$$\begin{aligned} F(1) &= \frac{1}{4} \sum_{x=0}^3 f(x) \exp[-j2\pi x/4] \\ &= \frac{1}{4} (2e^0 + 3e^{-j\pi/2} + 4e^{-j\pi} + 4e^{-j3\pi/2}) \\ &= \frac{1}{4} (-2 + j) \end{aligned}$$



شکل ۳.۵ یک تابع ساده و نمونه های آن در میدان X . در (الف) متغیر پیوسته و در (ب) گسته است

مرحله آخر در محاسبه فوق از رابطه اول نتیجه می شود. با ادامه' این روند خواهیم داشت:

$$F(2) = -\frac{1}{4}[1 + j0]$$

و

$$F(3) = -\frac{1}{4}[2 + j].$$

تمام مقادیر $f(x)$ در هر یک از چهار مؤلفه تبدیل فوریه گسته سهیم هستند. همچنین طبق معادله' (۳-۲.۳) تمام مؤلفه های تبدیل در بدست آوردن تبدیل معکوس سهیم هستند. مراحل بدست آوردن تبدیل معکوس نیز شبیه مراحل محاسبه' $F(u)$ است.

طیف فوریه از دامنه' مؤلفه های تبدیل بدست می آید؛ یعنی،

$$|F(0)| = 3.25$$

$$|F(1)| = \left[\left(\frac{2}{4}\right)^2 + \left(\frac{1}{4}\right)^2 \right]^{1/2} = \frac{\sqrt{5}}{4}$$

$$|F(2)| = \left[\left(\frac{1}{4}\right)^2 + \left(\frac{0}{4}\right)^2 \right]^{1/2} = \frac{1}{4}$$

و

$$|F(3)| = \left[\left(\frac{0}{4}\right)^2 + \left(\frac{1}{4}\right)^2 \right]^{1/2} = \frac{\sqrt{5}}{4}.$$

□

۳.۳ چند خاصیت از تبدیل فوریه دو بعدی

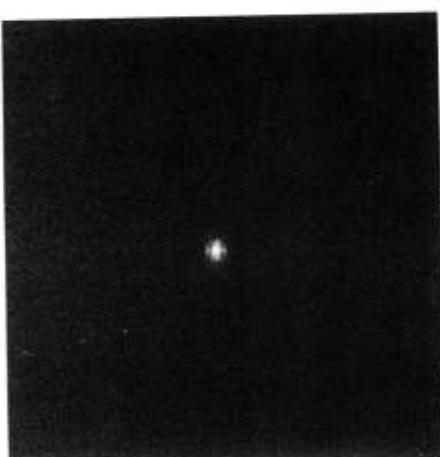
این بخش روی خواصی از تبدیل فوریه که در توضیحات بعدی مفید هستند، تمرکز می کند. گرچه توجه اصلی به تبدیلات گسته دو بعدی است، اما اگر این خواص در ابتدا در شکل پیوسته دو بعدی بیان شوند، مفاهیم پایه ای بعضی از آنها بهتر درک می شوند.

بانمایش تصاویر و طیف فوریه به صورت توابع شدت، تعدادی از مفاهیم مورد بررسی در این بخش روشن تر می شوند. بنابراین قبل از شرح خواص تبدیل فوریه ارائه برخی توضیحات در مورد این نمایش ها

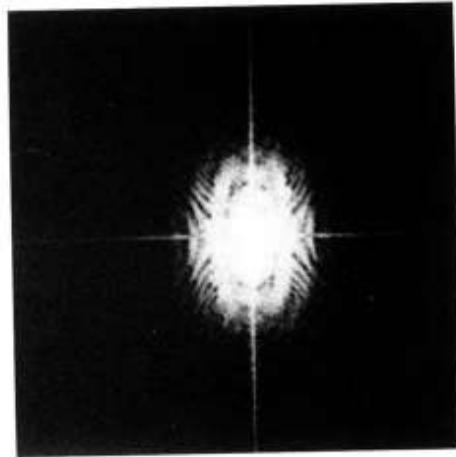
۳.۳ چند خاصیت از تبدیل فوریه دو بعدی ۱۴۲



(الف)



(ب)



(ب)

شکل ۳.۶ (الف) تصویری از سیاره کیوان؛ (ب) نمایش $|F(u, v)|$ ؛ (پ) نمایش $\log[1 + |F(u, v)|]$ که به بیت مقیاس شده است (یعنی، سطوح خاکستری آن بین ۰ و ۲۵۵ هستند)

مفید است. محدوده دینامیک^۱ طیف فوریه معمولاً بسیار بالاتر از حدی است که وسائل نمایش متداول قادر به بازسازی کاملاً صحیح آن هستند؛ بنابراین در این مورد تنها روش تنی بخش های تصویر را روی صفحه نمایشگر قابل رویت هستند. اغلب این مشکل در مورد ضبط تصویر روی فیلم نیز رخ می دهد. یک روش مفید که این مشکل را جبران می کند، شامل نمایش تابع

$$D(u, v) = c \log[1 + |F(u, v)|] \quad (1-3.3)$$

1) Dynamic range

به جای $|F(u, v)|$ است، که ثابت مقیاس می باشد و تابع لگاریتم، فشرده سازی مطلوب را انجام می دهد. همان طور که شکل ۳.۶ نشان می دهد، استفاده از معادله (۳.۱-۳) تا حد زیادی تحلیل چشمی طف فوریه را تسهیل می کند. شکل ۳.۶ (الف) یک تصویر رقemi از سیاره کیوان را نشان می دهد و شکل ۳.۶ (ب) طیف فوریه معمولی را که به صورت تصویر شدت نمایش داده شده است، نشان می دهد. این طیف ویژه، مقادیری در محدوده $[0, 2.5 \times 10^6]$ دارد. همان طور که برای چنین محدوده دینامیکی بزرگی انتظار می رود وقتی این تابع تصویر جهت نمایش روی سامانه ۸ بیتی مقیاس داده شود، روشن ترین مقادیر از حد بالای قابل نمایش بزرگتر می شوند. در این حالت مقادیر $|F(u, v)|$ در محدوده ۰ تا 4π قرار دارند که وقتی برای نمایش روی همان سامانه ۸ بیتی مقیاس داده شوند، نتیجه به صورت شکل ۳.۶ (ب) می شود که افزایش جزئیات قابل مشاهده در آن واضح است.

۱.۳.۳ جدابذیری

زوج تبدیل فوریه معادلات (۳.۲-۲)، (۳.۹-۲) و (۳.۱۰-۲) را می توان به صورت های جدابذیر

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \exp[-j2\pi ux/N] \sum_{y=0}^{N-1} f(x, y) \exp[-j2\pi vy/N] \quad (2-3.3)$$

برای $u, v = 0, 1, \dots, N-1$

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \exp[j2\pi ux/N] \sum_{v=0}^{N-1} F(u, v) \exp[j2\pi vy/N] \quad (3-3.3)$$

برای $x, y = 0, 1, \dots, N-1$ بیان کرد.

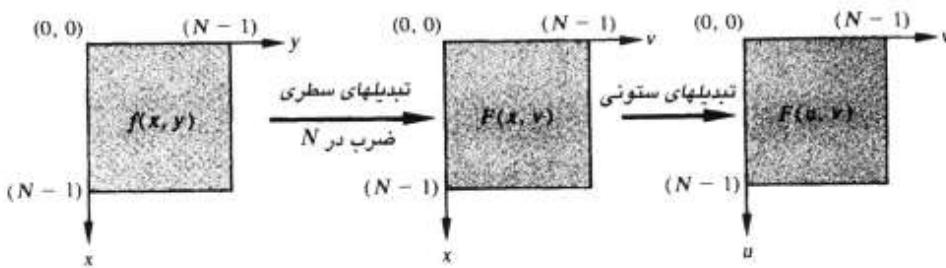
فاایده اصلی خاصیت جدابذیری این است که می توان (u, v) را با $f(x, y)$ را با کاربردهای متوالی تبدیل فوریه یک بعدی بدست آورد. با بیان معادله (۳.۲-۳) به صورت زیر فایده آن روشن تر می شود:

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} F(x, v) \exp[-j2\pi ux/N] \quad (4-3.3)$$

که

$$F(x, v) = N \left[\frac{1}{N} \sum_{y=0}^{N-1} f(x, y) \exp[-j2\pi vy/N] \right]. \quad (5-3.3)$$

۱۴۹ چند خاصیت از تبدیل فوریه دو بعدی



شکل ۷.۳ محاسبه تبدیل فوریه دو بعدی به صورت دنباله ای از تبدیلات یک بعدی

برای هر مقدار x ، عبارت درون کروشهای در معادله (۳.۵-۳)، تبدیلی یک بعدی با مقادیر فرکانس $v = 0, 1, \dots, N-1$ می باشد. بنابراین تابع دو بعدی $F(x, v)$ با گرفتن تبدیل در طول هر یک از سطرهای y و ضرب حاصل در عدد N بدست می آید. سپس همان طور که معادله (۳.۴-۳) نشان می دهد، با گرفتن تبدیل در طول هر ستون v ، نتیجه مطلوب یعنی $F(u, v)$ بدست می آید. این فرآیند در شکل ۳.۷ جمع بندی می شود. همان نتیجه را می توان با گرفتن تبدیل ابتداء در طول ستونهای y و سپس در طول سطرهای حاصل، نیز بدست آورد. این مورد بسادگی با معکوس کردن ترتیب مجموعهای موجود در معادله (۳.۳-۲) بررسی می شود. در پایه سازی معادله (۳.۳-۳) نیز توضیحات فوق صدق می کند.

۴.۳.۳ انتقال

خواص انتقال زوج تبدیل فوریه عبارتند از

$$f(x, y) \exp[j2\pi(u_0x + v_0y)/N] \Leftrightarrow F(u - u_0, v - v_0) \quad (6-3.3)$$

و

$$f(x - x_0, y - y_0) \Leftrightarrow F(u, v) \exp[-j2\pi(ux_0 + vy_0)/N] \quad (7-3.3)$$

مانند معادلات (۳.۱-۹) و (۳.۱-۱۰) یا معادلات (۳.۲-۹) و (۳.۲-۱۰)، پیکان دوسر، تناظر بین تابع و تبدیل فوریه اش (و بالعکس) را نشان می دهد.

معادله (۳.۶-۳) نشان می دهد که ضرب $f(x, y)$ در جزء نمایی نشان داده شده و گرفتن تبدیل از حاصل ضرب، باعث انتقال مبدأ صفحه فرکانس به نقطه (u_0, v_0) می شود. به طور مشابه ضرب $F(u, v)$ در عبارت نمایی نشان داده شده و گرفتن تبدیل معکوس، مبدأ صفحه مکانی را به (x_0, y_0) می برد.

در این فصل و فصل ۴، از معادله (۳.۶-۳) با مقادیر $u_0 = v_0 = N/2$ یا

$$\begin{aligned}\exp[j2\pi(u_0x + v_0y)/N] &= e^{j\pi(x+y)} \\ &= (-1)^{x+y}\end{aligned}$$

و

$$f(x, y)(-1)^{x+y} \Leftrightarrow F(u = N/2, v = N/2). \quad (8-3. 3)$$

استفاده قابل ملاحظه ای می شود. بنابراین بسادگی می توان با ضرب $f(x, y)$ در $(-1)^{x+y}$ مبدأ تبدیل فوریه $f(x, y)$ را به مرکز مربع فرکانسی $N \times N$ متناظر ش انتقال داد. در حالت یک متغیره این انتقال به صورت ضرب $f(x)$ در عبارت $(-1)^x$ ساده می شود.

از معادله (۳.۷-۳) نتیجه می شود که انتقال $f(x, y)$ بر بزرگی تبدیل فوریه اش اثر نمی گذارد، مانند

$$|F(u, v)\exp[-j2\pi(ux_0 + vy_0)/N]| = |F(u, v)|. \quad (9-3. 3)$$

لازم است که این نتیجه را در ذهن داشته باشیم، زیرا ارزیابی چشمی تبدیل معمولاً به نمایش بزرگیش محدود می شود.

۳.۳.۳ تناوب و تقارن مزدوج

تبدیل فوریه گستته و معکوس آن، متناوب با دوره N هستند؛ یعنی،

$$F(u, v) = F(u + N, v) = F(u, v + N) = F(u + N, v + N). \quad (10-3. 3)$$

درستی این خاصیت را می توان با جایگذاری مستقیم متغیرهای $(u + N)$ و $(v + N)$ در معادله (۳.۱۰-۲) نشان داد. گرچه معادله (۳.۱۰-۳) نشان می دهد که $F(u, v)$ برای تعداد نامحدودی از

۳.۳ چند خاصیت از تبدیل فوریه دو بعدی ۱۳۱

مقادیر $f(x, y)$ تکرار می شود، تنها N مقدار از هر متغیر در یک دوره تناوب لازم است تا $f(x, y)$ از $F(u, v)$ بدست آید. به عبارت دیگر، تنها یک دوره تناوب از تبدیل لازم است تا $F(u, v)$ به طور کامل در میدان فرکانس مشخص شود. توضیحات مشابهی در مورد $f(x, y)$ در میدان مکان صدق می کند.

اگر $f(x, y)$ حقیقی باشد، تبدیل فوریه خاصیت تقارن مزدوج نیز دارد:

$$F(u, v) = F^*(-u, -v) \quad (11-3. 3)$$

یا، به عبارت ساده تر

$$|F(u, v)| = |F(-u, -v)| \quad (12-3. 3)$$

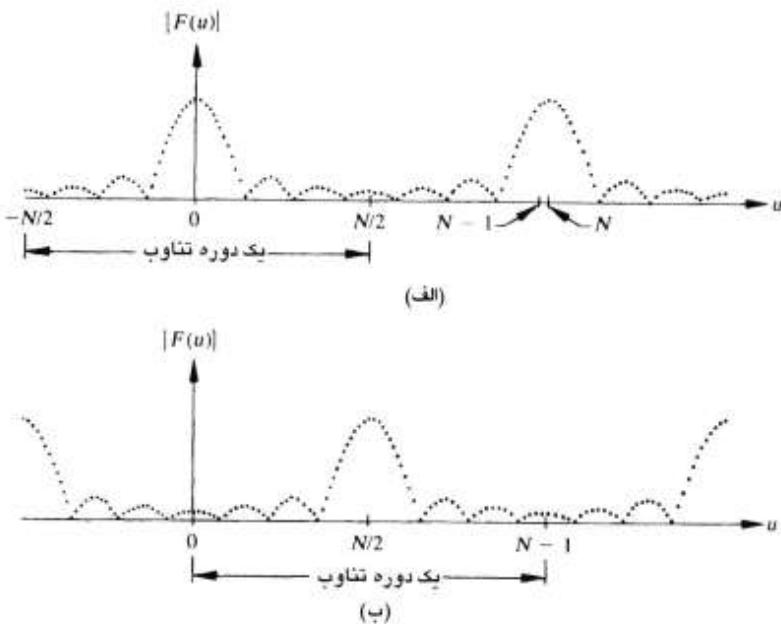
که $F^*(u, v)$ مزدوج مختلط $F(u, v)$ است. همان طور که قبلاً ذکر شد، در اغلب موارد برای تحلیل بیشتر، نمایش دامنه تبدیل فوریه مورد علاقه است. برای بررسی اثر معادلات (۱۰-۳. ۳) و (۱۲-۳. ۳) بر نمایش دامنه تبدیل، اجازه دهید در ابتدا حالت یک متغیره را در نظر بگیریم، که

$$\begin{aligned} F(u) &= F(u + N) \\ |F(u)| &= |F(-u)|. \end{aligned}$$

خاصیت تناوب نشان می دهد که $F(u)$ دوره تناوبی به طول N دارد و طبق شکل ۳. ۸ (الف) خاصیت تقارن نشان می دهد که مرکز دامنه تبدیل، مبدأ است. شکل ۳. ۸ (الف) و توضیحات قبلی روشن می سازند که دامنه مقادیر تبدیل در بازه $1 - N/2$ تا $1 + N/2$ تصویر مقادیر دامنه تبدیل در نیم دوره تناوب سمت چپ مبدأ هستند. چون روابط تبدیل فوریه گسته برای مقادیر $1 - N$ در بازه $[0, 1]$ [بیان شده اند، دو نیم دوره پشت به پشت را در این بازه نتیجه می دهند. طبق شکل ۳. ۸ (ب) برای نمایش یک دوره کامل کافیست مبدأ تبدیل به نقطه $N/2 = u$ جابجا شود. همان طور که قبلاً بیان شد، برای انجام این کار کافیست قبل از گرفتن تبدیل، $f(x)$ را در $(-1, 1)$ ضرب کنیم.

موارد فوق در مورد دامنه تبدیل فوریه دو بعدی نیز درست است، با این استثنای اگر مبدأ تبدیل به فرکانس $N/2, N/2$ منتقل نشود، تعبیر نتایج به طور قابل ملاحظه ای مشکل تر می شود. اشکال ۹. ۳ (ب) و (پ) این تفاوت را نشان می دهند؛ تصویر آخری با استفاده از خاصیت مرکزی کردن^۱ عبارت

1) Centering Property



شکل ۳ نمایش خواص تناوبی تبدیل فوریه: (الف) طیف قوریه که نیم دوره های چسبیده به هم در بازه $[0, N-1]$ را نشان می دهد؛ (ب) طیف انتقال یافته که یک دوره کامل را در همان بازه نشان می دهد.

۸-۳. ۳ بdst آمده است.

۴.۳.۳ چرخش

اگر مختصات قطبی را به صورت

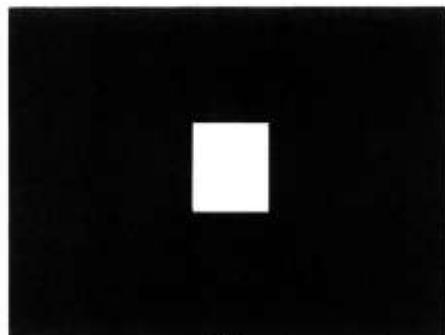
$$x = r \cos \theta \quad y = r \sin \theta \quad u = \omega \cos \phi \quad v = \omega \sin \phi$$

معرفی نماییم، آنگاه $f(x, y)$ و $F(u, v)$ بترتیب به $f(r, \theta)$ و $F(r, \theta)$ تبدیل می شوند. با جایگذاری مستقیم در زوج تبدیل فوریه گستته یا پیوسته داریم:

$$f(r, \theta + \theta_0) \Leftrightarrow F(r, \theta + \theta_0). \quad (13-3. 3)$$

به عبارت دیگر، با چرخش $f(x, y)$ به اندازه زاویه θ_0 ، $F(u, v)$ نیز به همان اندازه می چرخد. به طور

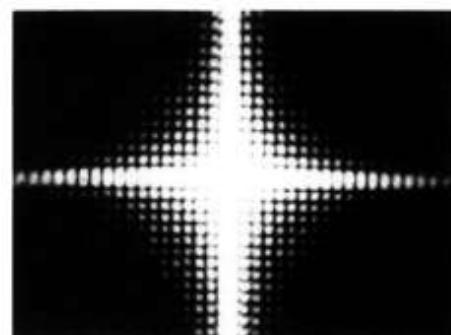
۳.۳ چند خاصیت از تبدیل فوریه دو بعدی ۱۳۳



(الف)



(ب)



(پ)

شکل ۹.۳ (الف) یک تصویر ساده؛ (ب) طیف فوریه آن بدون انتقال؛ (پ) طیف فوریه انتقال یافته به مرکز مربع فرکانسی

مشابه با چرخش $f(x, y)$ ، $F(u, v)$ نیز به همان اندازه می چرخد. شکل ۳.۱۰. این خاصیت را نمایش می دهد.

۵.۳.۳ توزیع بذیری^۱ و تغییر مقیاس^۲

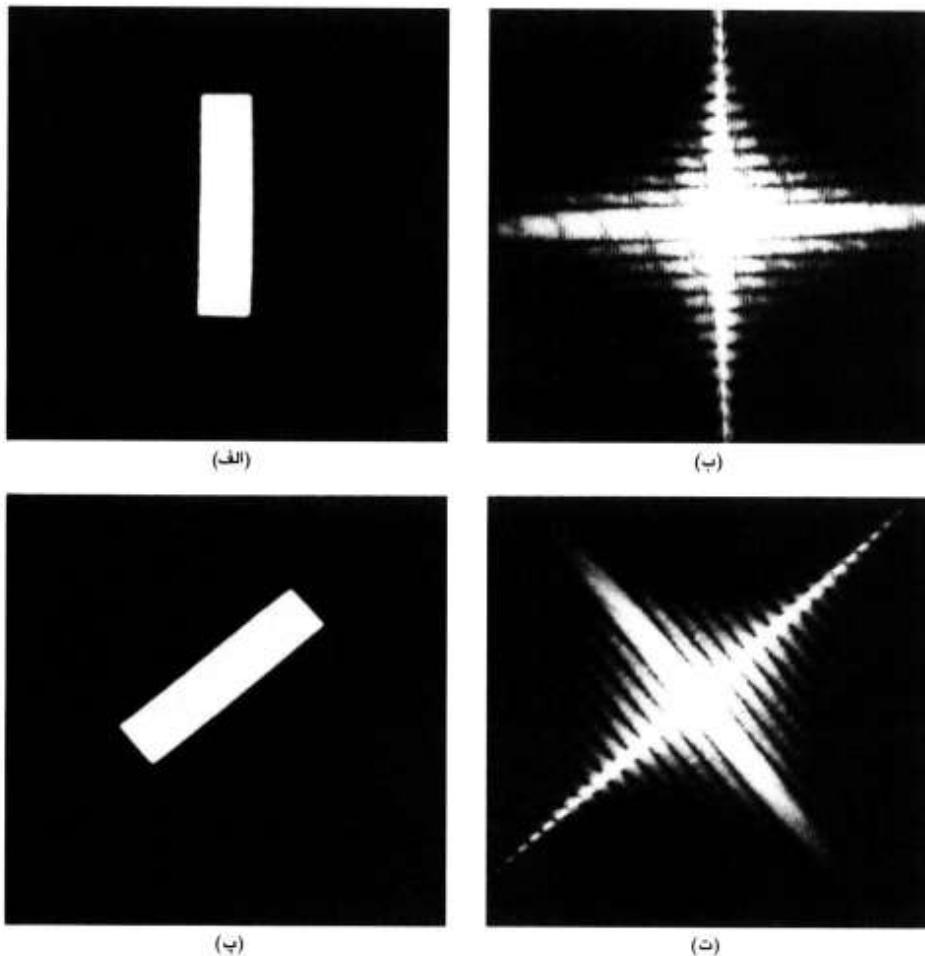
از تعریف زوج تبدیل پیوسته یا گسته داریم،

$$\tilde{\delta}\{f_1(x, y) + f_2(x, y)\} = \tilde{\delta}\{f_1(x, y)\} + \tilde{\delta}\{f_2(x, y)\} \quad (14-3.3)$$

و در حالت کلی می توان نوشت،

1) Distributivity

2) Scaling



شکل ۳.۰.۱۰ خواص چرخشی تبدیل فوریه: (الف) یک تصویر ساده؛ (ب) طیف آن؛ (پ) تصویر چرخانده شده؛ (ت) طیف حاصل

$$\mathcal{F}\{f_1(x, y) \cdot f_2(x, y)\} \neq \mathcal{F}\{f_1(x, y)\} \cdot \mathcal{F}\{f_2(x, y)\}. \quad (15-3. 3)$$

به عبارت دیگر، تبدیل فوریه و معکوس آن نسبت به عمل جمع توزیع پذیر هستند اما نسبت به عمل ضرب توزیع پذیر نیستند.

برای دو عدد اسکالر a و b ,

$$af(x, y) \Leftrightarrow aF(u, v) \quad (16-3. 3)$$

۱۳۵ ۳.۳ چند خاصیت از تبدیل فوریه دو بعدی

$$f(ax, by) \Leftrightarrow \frac{1}{|ab|} F(u/a, v/b). \quad (17-3. 3)$$

۶.۳.۳ مقدار متوسط

تعريف متداول مقدار متوسط یک تابع گسته دو بعدی عبارت است از:

$$\bar{f}(x, y) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y). \quad (18-3. 3)$$

با جایگذاری $u = v = 0$ در معادله (۹-۲) داریم:

$$F(0, 0) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y). \quad (19-3. 3)$$

بنابراین $\bar{f}(x, y)$ با تبدیل فوریه $f(x, y)$ رابطه زیر را دارد:

$$\bar{f}(x, y) = \frac{1}{N} F(0, 0). \quad (20-3. 3)$$

۷.۳.۳ لابلائسین^۱

لابلائسین تابع دو متغیره $f(x, y)$ به صورت زیر تعریف می شود:

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}. \quad (21-3. 3)$$

طبق تعریف تبدیل فوریه دو بعدی،

$$\hat{f}\{\nabla^2 f(x, y)\} \Leftrightarrow -(2\pi)^2(u^2 + v^2)F(u, v). \quad (22-3. 3)$$

عملگر لابلائسین برای تماش لبه های تصویر مفید است، که این مطلب در بخش ۱.۷ بیان می شود.

۸.۳.۳ گانولوشن^۲ و همبستگی^۳

در این بخش دورابطه از تبدیل فوریه را که پیوندی پایه ای بین میدانهای مکان و فرکانس بوجود می آورند،

۱۳۶ پردازش تصویر رقمه

بررسی می کنیم. این روابط که کانولوشن و همبستگی نامیده می شوند، در فهم فنون پردازش تصویر مبتنی بر تبدیل فوریه اهمیت اساسی دارند. به منظور تبیین مفاهیم فوق، کار را با در نظر گرفتن کانولوشن یک بعدی و با آرگومان پیوسته شروع می کنیم. آنگاه آن را به حالت گسته و در نهایت به حالات پیوسته و گسته دو بعدی توسعه می دهیم. مفهوم همبستگی را نیز به همین طریق ارائه می کنیم.

کانولوشن

کانولوشن دوتابع $f(x)$ و $g(x)$ ، که با $f(x) * g(x)$ نشان داده می شود، با انتگرال

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(\alpha)g(x - \alpha) d\alpha \quad (23-3. 3)$$

تعريف می شود که α متغیر کمکی انتگرال گیری است. طرز کار انتگرال کانولوشن مخصوصاً از نظر نمایش ساده نیست، بنابراین شرح آن را با دو مثال که به صورت گرافیکی کاربرد معادله (۲۳-۳. ۳) را نمایش می دهند، شروع می کنیم.

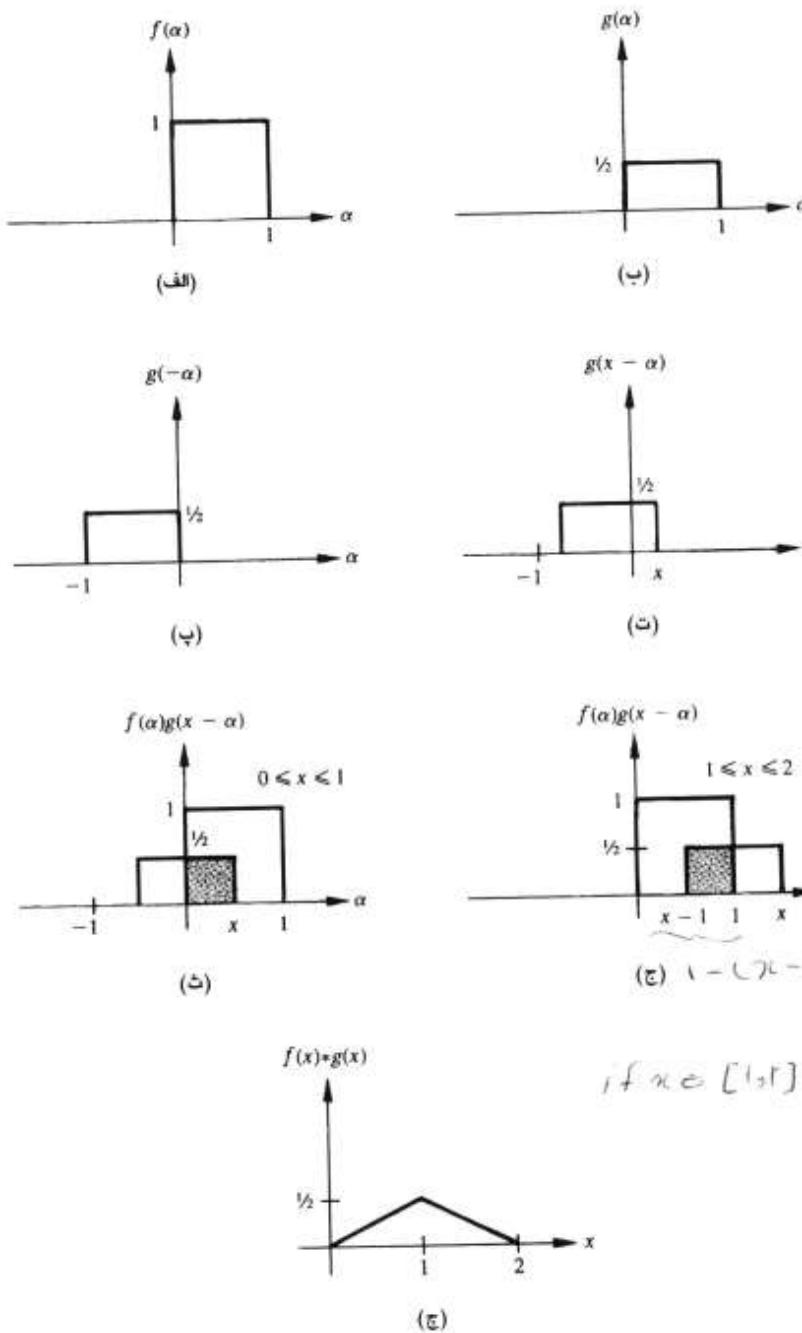
مثال: اولین مثال، کانولوشن توابع $f(x)$ و $g(x)$ را که به ترتیب در اشکال ۳. ۱۱(الف) و (ب) دیده می شوند، نشان می دهد. قبل از انتگرال گیری، باید تابع $(x - \alpha)g(x - \alpha)$ ایجاد شود. انجام این کار نیازمند دو مرحله نشان داده شده در اشکال ۳. ۱۱(ب) و (ت) می باشد. این عمل شامل چرخاندن $(\alpha)g$ حول مبدأ تا این که $(\alpha - x)g$ بدست آید و سپس جابجایی تابع به اندازه 'با x ' می باشد. آنگاه برای هر مقدار x ، $f(x - \alpha)g$ ضرب و از حاصل در بازه ' ∞ - x ' انتگرال گرفته می شود. حاصل ضرب $f(x - \alpha)g$ در $[0, x]$ بخش سایه خورده شکل ۳. ۱۱(ث) است. این شکل برای $0 \leq x \leq 1$ معتبر است. حاصل ضرب برای مقادیر α که خارج از بازه $[0, 1]$ باشند، صفر است، بنابراین $f(x) * g(x) = x/2$. کار برای سطح ناحیه سایه خورده شکل ۳. ۱۱(ت) می باشد. برای x در بازه $[1, 2]$ ، شکل ۳. ۱۱(ج) به کار می رود و $f(x) * g(x) = (1 - x/2)g(x - \alpha)$. بنابراین چون $f(x)g(x - \alpha)$ برای مقادیر x که خارج از بازه $[0, 2]$ هستند، صفر است، در نهایت داریم:

$$f(x) * g(x) = \begin{cases} x/2 & 0 \leq x \leq 1 \\ 1 - x/2 & 1 \leq x \leq 2 \\ 0 & \text{در غیر این صورت} \end{cases}$$

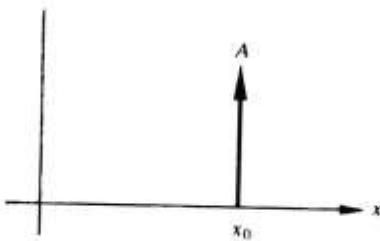
شکل ۳. ۱۱(ج) نتیجه نهایی را نمایش می دهد.



۳.۳ چند خاصیت از تبدیل فوریه دو بعدی ۱۳۷



شکل ۱۱.۳ نمایش گرافیکی کانولوشن. نواحی مایه خورده جاهایی هستند که حاصل ضرب صفر نیست.

شکل ۱۲۰.۳ نمایش گرافیکی $A \delta(x - x_0)$

یک جنبه معادله (۲۴-۳. ۳) که بعداً در این بخش مورد استفاده قرار می گیرد، کانولوشن تابع $f(x)$ با تابع ضربه $\delta(x - x_0)$ می باشد، که با رابطه

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0). \quad (24-3. 3)$$

تعريف می شود. تابع $\delta(x - x_0)$ در یک همسایگی بی نهایت کوچک حول x_0 مقدار دارد و در سایر نقاط صفر است و مساحت زیر آن نیز یک واحد سطح است؛ یعنی،

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx = \int_{x_0}^{x_0} \delta(x - x_0) dx = 1. \quad (25-3. 3)$$

در بیشتر کاربردها می توان گفت که $\delta(x - x_0)$ در $x_0 = x$ قرار دارد و قدرت ضربه برابر با مقدار $f(x_0)$ در $x = x_0$ است. به عنوان نمونه اگر $f(x) = A \delta(x - x_0)$ باشد، کانولوشن A در نقطه x_0 و با ارتفاع برابر با قدرت آن نمایش می دهد. شکل ۱۲. ۳ این نمایش را برابر با $A \delta(x - x_0)$ نشان می دهد.

مثال: در دومین نمایش کاربرد معادله (۲۳-۳. ۳)، تابع $(x)f$ در شکل ۱۳. ۳ (الف) با تابع $g(x) = \delta(x + T) + \delta(x) + \delta(x - T)$ می شود. با چرخاندن $(x)f$ ، لغزاندن آن بدنیال $(x)f$ و استفاده از معادلات (۲۳-۳. ۳) و (۲۴-۳. ۳) نتیجه نشان داده شده در شکل ۱۳. ۳ (ب) بدست می آید. کانولوشن در این مورد معادل تکرار $(x)f$ در محل هریک از ضربه ها است.

۱۳۹ ۳.۳ چند خاصیت از تبدیل فوریه دو بعدی

اهمیت کانولوشن در تحلیل میدان فرکانس دراین است که $(x) * g(x)$ و $f(x)G(u)$ یک زوج تبدیل فوریه را تشکیل می دهند. به عبارت دیگر، اگر تبدیل فوریه $(x) f(x)$ و $g(x)$ به ترتیب $F(u)$ و $G(u)$ باشد، $F(u)G(u)$ تبدیل فوریه $(x) f(x) * g(x)$ می باشد. این نتیجه که به صورت

$$f(x) * g(x) \Leftrightarrow F(u)G(u) \quad (26-3.3)$$

نمایش داده می شود، بیان می کند که می توان با گرفتن معکوس تبدیل فوریه از حاصل ضرب $F(u)G(u)$ حاصل کانولوشن در میدان x را بدست آورد. نتیجه مشابه دیگر این است که کانولوشن در میدان فرکانس به صورت ضرب در میدان x ساده می شود؛ یعنی،

$$f(x)g(x) \Leftrightarrow F(u) * G(u). \quad (27-3.3)$$

معمولًاً این دو نتیجه را قضیه کانولوشن^۱ می نامند.

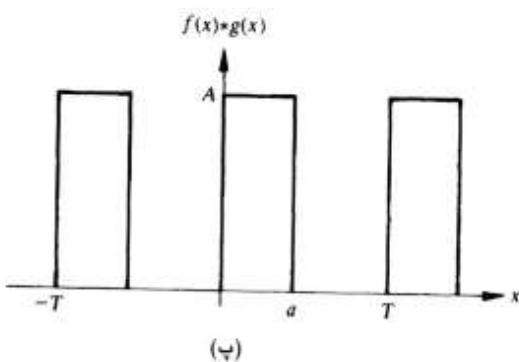
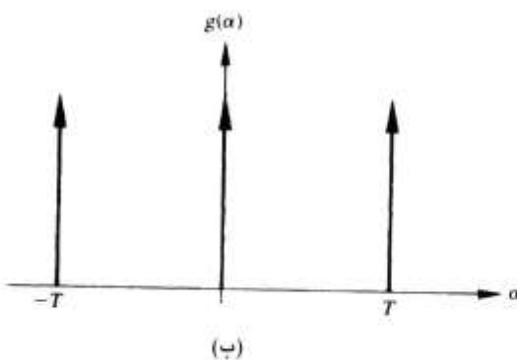
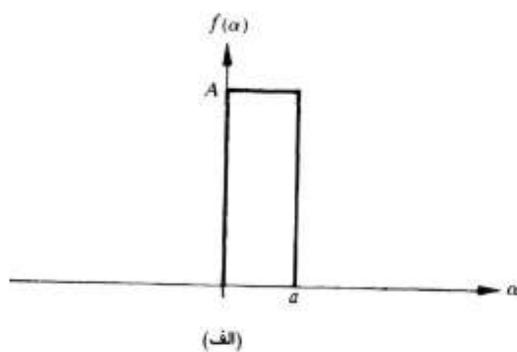
فرض کنید که $(x) f(x)$ و $(x) g(x)$ نمونه برداری شده و به آرایه هایی به ترتیب با اندازه های A و B گسته شده اند: $\{f(0), f(1), f(2), \dots, f(A-1)\}$ و $\{g(0), g(1), g(2), \dots, g(B-1)\}$ همان طور که در بخش ۳.۳ اشاره شد، تبدیل فوریه گسته و معکوس آن توابع متناوب هستند. برای نوشتן رابطه قضیه کانولوشن گسته به طوری که تناوب را در خود داشته باشد، باید فرض کرد که توابع گسته $f(x)$ و $(x) g$ متناوب با دوره تناوب M هستند. آنگاه حاصل کانولوشن، متناوب با همان دوره تناوب است. مسئله این است که چگونه مقدار M انتخاب شود. می توان نشان داد [Brigham 1974] که جز در حالی که

$$M \geq A + B - 1 \quad (28-3.3)$$

در سایر حالات دوره های تناوب متفاوت حاصل کانولوشن روی هم می افتد، که معمولاً این پدیده را خطای پیچشی^۲ می نامند. اگر $M = A + B - 1$ ، دوره ها چسبیده خواهند بود؛ اگر $M > A + B - 1$ ، دوره ها نامتصل خواهند بود که میزان فاصله آنها برابر با تفاضل $M - 1$ باشد. چون دوره در نظر گرفته شده باید بزرگتر از A و B باشد، باید طول دنباله های نمونه برداری

1) Convolution theorem

2) Wraparound error



شکل ۱۳۰.۳ کانولوشنی که شامل توابع ضربه است

۱۳۱ ۳.۳ چند خاصیت از تبدیل فوریه دو بعدی

شده افزایش یابد تا این که هر دو برابر M شوند. با افزودن صفر به نمونه ها، دنباله های توسعه یافته

$$f_r(x) = \begin{cases} f(x) & 0 \leq x \leq A - 1 \\ 0 & A \leq x \leq M - 1 \end{cases}$$

و

$$g_r(x) = \begin{cases} g(x) & 0 \leq x \leq B - 1 \\ 0 & B \leq x \leq M - 1. \end{cases}$$

ایجاد می شود. با این تغییرات، کانولوشن گسته $(f_r * g_r)(x)$ با عبارت

$$f_r(x) * g_r(x) = \frac{1}{M} \sum_{m=0}^{M-1} f_r(m) g_r(x - m) \quad (29-3.3)$$

برای $x = 0, 1, 2, \dots, M - 1$ تعریف می شود. تابع کانولوشن، آرایه گسته 'متناوبی' به طول M است که مقادیر $x = 0, 1, 2, \dots, M - 1$ بر یک دوره کامل آن دلالت می کنند.

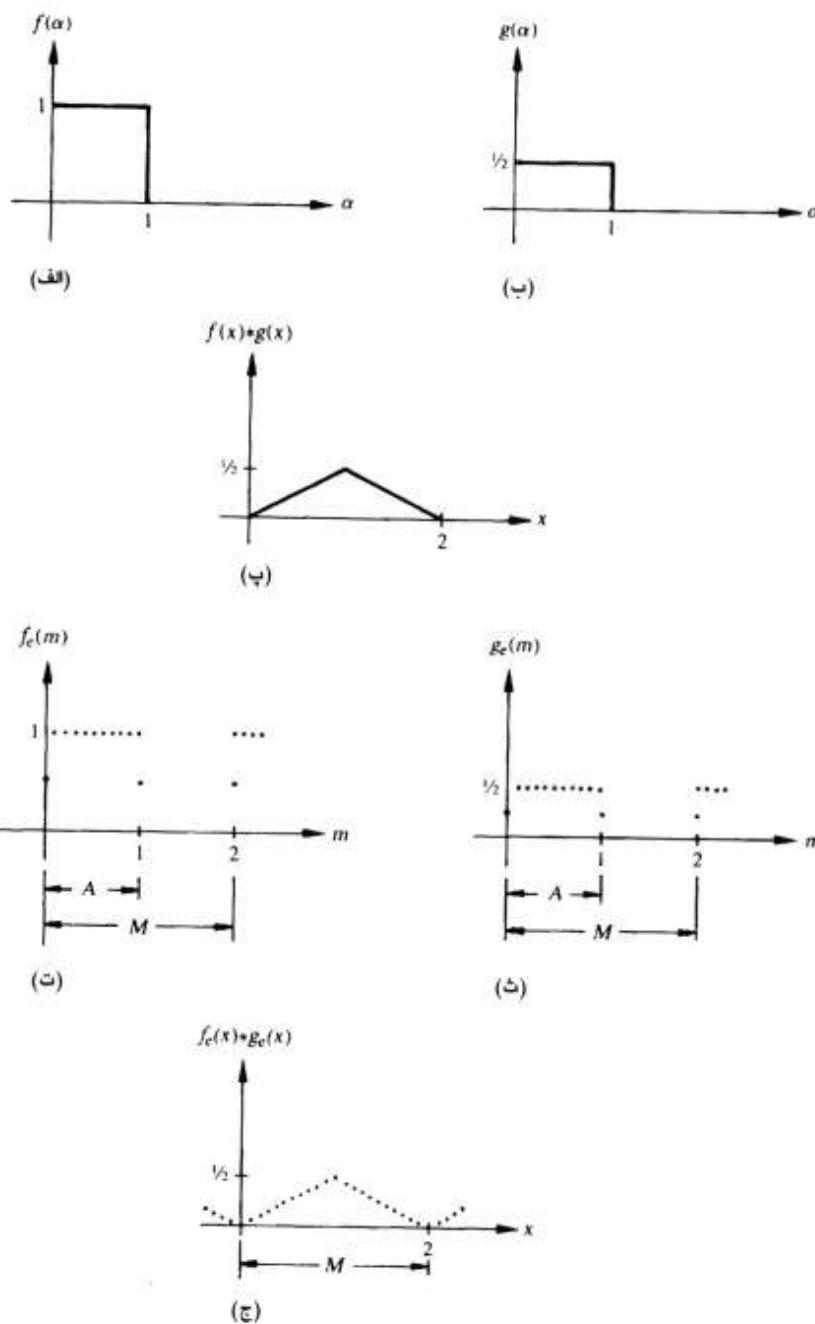
طرز انجام کانولوشن گسته در اساس مانند کانولوشن پیوسته است، تنها تفاوت این است که جابجایی ها در گامهای گسته که برابر فاصله بین نمونه ها است، انجام می شوند و دیگر این که عمل مجموع گیری جانشین انتگرال گیری می شود. به طور مشابه معادلات (۳.۲۶) و (۳.۲۷) در حالت گسته هم برقرارند؛ که در این حالت برای جلوگیری از خطای پیچشی از $(f_r * g_r)(x)$ استفاده می شود. متغیرهای گسته x و مقادیری در محدوده $1 - M, 0, 1, 2, \dots, M - 1$ می گیرند.

مثال: شکل ۱۴.۳ نتایج قبلی درباره 'کانولوشن پیوسته' و گسته رانمایش می دهد. نمودارهای حالت گسته، نمونه از هر یک از $f(x)$ و $g(x)$ در بازه $[0, 1]$ ، و یک دوره در نظر گرفته شده به طول $M = A + B - 1 = 2A - 1$ را نشان می دهند.

توجه کنید که تابع کانولوشن متناوب است و دوره های نیز به هم چسبیده اند زیرا $M = 2A - 1$. انتخاب $M > 2A - 1$ ، فاصله بزرگتری را بین این دوره ها ایجاد می کند. همچنین توجه کنید که نمونه یک دوره را به طور کامل بیان می کنند.

□

کانولوشن دو بعدی شبیه معادله (۳.۲۳) است. بنابراین برای دوتابع $f(x, y)$ و $g(x, y)$



شکل ۱۴.۳ مقایسه بین کانولوشن پیوسته و گسته

۱۴۳ ۳.۳ چند خاصیت از تبدیل فوریه دو بعدی

داریم،

$$f(x, y) * g(x, y) = \int_0^x \int_0^y f(\alpha, \beta) g(x - \alpha, y - \beta) d\alpha d\beta. \quad (30-3. 3)$$

پس قضیه کانولوشن دو بعدی با روابط زیر بیان می شود:

$$f(x, y) * g(x, y) \Leftrightarrow F(u, v) G(u, v) \quad (31-3. 3)$$

و

$$f(x, y) g(x, y) \Leftrightarrow F(u, v) * G(u, v). \quad (32-3. 3)$$

نمایش گرافیکی معادله (۳۰-۳. ۳) بسیار مشکل تر از معادله (۲۳-۳. ۳) است. شکل ۳. ۱۵. چند عمل پایه ای تاکردن^۱، جابجایی و ضرب را که در کانولوشن دو بعدی مورد نیاز هستند، نشان می دهد. حاصل تغییر متغیرهای جابجایی، x و y ، یک سطح کانولوشن دو بعدی خواهد بود که شکل آن به طبیعت توابع بکار رفته در فرآیند بستگی دارد.

با فرض اینکه (y, y) آرایه های گسته به ترتیب با اندازه های $B \times A$ و $C \times D$ باشند، روابط کانولوشن گسته دو بعدی بدست می آیند. مانند حالت یک بعدی این آرایه ها باید درجهات X و Y متناظر با دوره تناوب دلخواه M و N باشند. با انتخاب های

$$M \geq A + C - 1 \quad (33-3. 3)$$

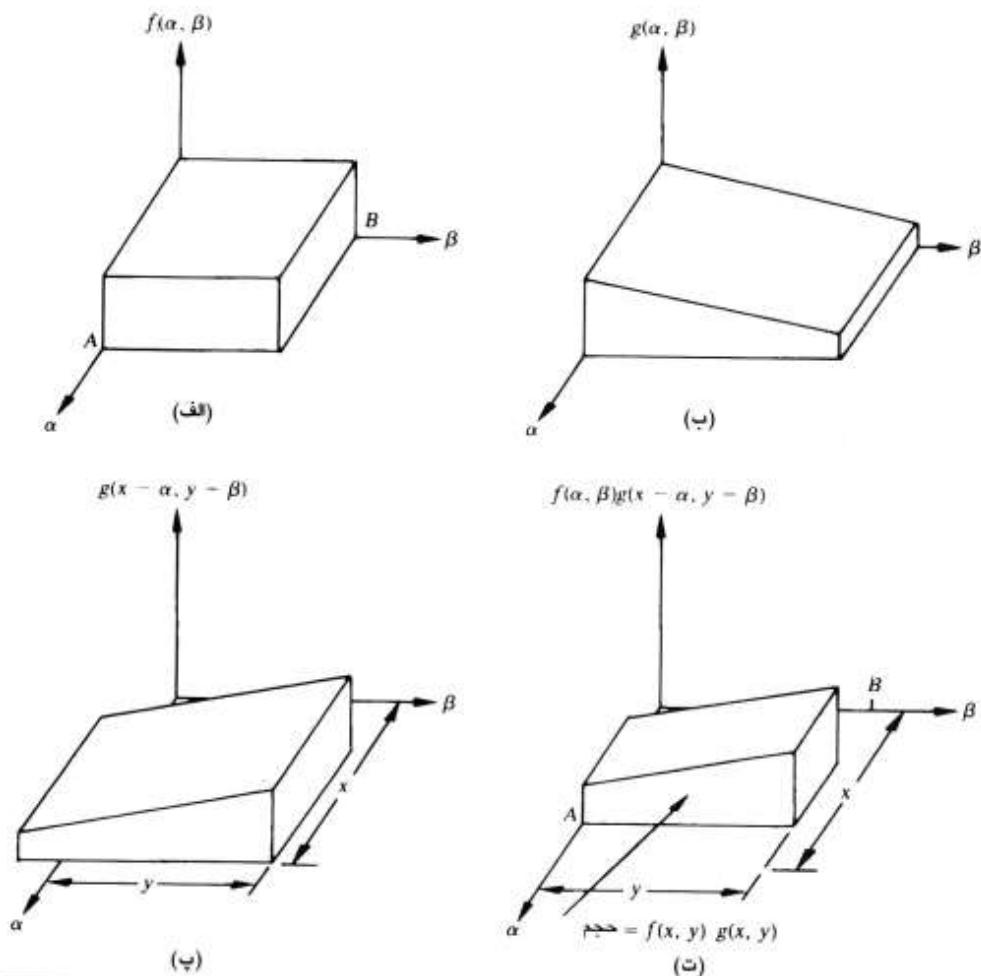
و

$$N \geq B + D - 1. \quad (34-3. 3)$$

از خطای پیچشی در دوره های کانولوشن متفاوت جلوگیری می شود. با گسترش $f(x, y)$ و $g(x, y)$ دنباله های متناظر به صورت زیر تشکیل می شوند:

$$f_r(x, y) = \begin{cases} f(x, y) & 0 \leq x \leq A - 1 \quad \text{و} \quad 0 \leq y \leq B - 1 \\ 0 & A \leq x \leq M - 1 \quad \text{یا} \quad B \leq y \leq N - 1 \end{cases}$$

1) *Folding*



شکل ۳.۱۵ نمایش مراحل مورد نیاز در اجرای کانولوشن دو بعدی، که عبارتند از: تاکردن، جابجا کردن،
و ضرب

$$g_r(x, y) = \begin{cases} g(x, y) & 0 \leq x \leq C-1 \quad \text{و} \quad 0 \leq y \leq D-1 \\ 0 & C \leq x \leq M-1 \quad \text{یا} \quad D \leq y \leq N-1. \end{cases}$$

کانولوشن دو بعدی $g_r(x, y)$ و $f_r(x, y)$ برای $y=0, 1, 2, \dots, N-1$ و $x=0, 1, 2, \dots, M-1$

با رابطه

$$f_r(x, y) * g_r(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_r(m, n) g_r(x-m, y-n) \quad (35-۳.۳)$$

۳.۳ چند خاصیت از تبدیل فوریه دو بعدی ۱۲۵

تعریف می شود. آرایه $N \times M$ حاصل از معادله (۳.۳-۳۵)، یک دوره از کانولوشن دو بعدی گسته است. اگر M و N طبق معادلات (۳.۳-۳۳) و (۳.۳-۳۴) انتخاب شوند، تضمین می شود که این آرایه عاری از تداخل با دوره های مجاور است. مانند حالت یک بعدی، قضیه کانولوشن پیوسته، یعنی معادلات (۳.۳-۳۱) و (۳.۳-۳۲)، در حالت گسته هم بازای مقادیر $M-1, 1, 2, \dots, N-1 = 0, 1, 2, \dots, N-1$ برقرار است. همه محاسبات روی توابع گسترش یافته (x, y) و (x, y) انجام می گیرند.

در بخش ۳.۹. وقتی قضیه نمونه برداری را توضیح می دهیم، قدرت نظری قضیه کانولوشن مشهود می شود. در عمل، محاسبه کانولوشن گسته در میدان فرکانس، اغلب با بازده تر از کاربرد مستقیم معادله (۳.۳-۳۵) است. روال این است که تبدیل فوریه های (x, y) و $f(x, y)$ با استفاده از تبدیل فوریه سریع^۱ (FFT) محاسبه شود (به بخش ۳.۴ مراجعه کنید). سپس دو تبدیل در هم ضرب می شوند و با گرفتن تبدیل فوریه معکوس از حاصلضرب، تابع کانولوشن بدست می آید. مقایسه ای که Brigham[1974] انجام داد، نشان می دهد که در مورد آرایه های یک بعدی، اگر تعداد نقاط بیشتر از ۳۲ باشد، روش FFT سریع تر است. گرچه این عدد به ماشین و الگوریتم مورد استفاده بستگی دارد، اما در بیشتر موارد کمتر از تعداد نقاط یک سطر یا ستون تصویر است.

همبستگی^۲

همبستگی[†] دوتابع پیوسته $f(x)$ و $g(x)$ ، که با $f(x) \circ g(x)$ نشان داده می شود، با رابطه

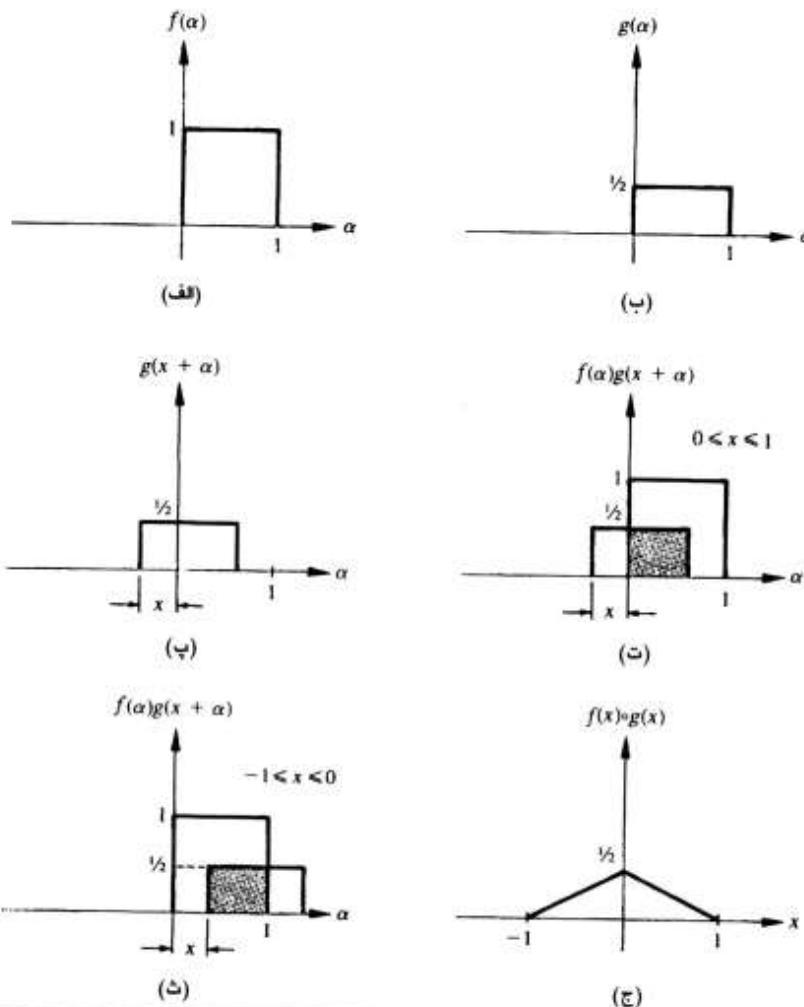
$$f(x) \circ g(x) = \int_{-\infty}^{\infty} f^*(\alpha)g(x + \alpha) d\alpha \quad (3.3-36)$$

تعریف می گردد که * نشانه مزدوج مختلط است. شکل معادلات (۳.۳-۳۶) و (۳.۳-۲۳) شبیه است و تنها تفاوت آنها این است که در معادله اخیر تابع (x) حول مبدأ چرخانده نمی شود. بنابراین برای اجرای همبستگی، $f(x)g(x)$ را روی x می لغزانیم و برای هر مقدار جابجایی x از حاصلضرب آنها در بازه $-\infty$ تا $+\infty$ انتگرال می گیریم. شکل ۳.۱۶.۱۱ و ۳.۱۶.۱۶ این روال را نمایش می دهد. اشکال ۳.۱۱ و ۳.۱۶ را

1) Fast Fourier Transform

2) Correlation

اگر $f(x)$ و $g(x)$ یکسان باشند، معادله (۳.۳-۳۶) معمولاً تابع خود همبستگی نامیده می شود؛ اگر $f(x)$ و $g(x)$ متفاوت باشند، معمولاً عبارت همبستگی متقابل استفاده می شود.



شکل ۳.۱۶ نمایش گرافیکی همبستگی. نواحی سایه خورده نشانگر جاهایی هستند که حاصل ضرب صفر نیست.

مقایسه کنید. معادل گستته معادله (۳.۳۶-۳) به صورت

$$f_r(x) * g_r(x) = \frac{1}{M} \sum_{m=0}^{M-1} f_r^*(m) g_r(x + m) \quad (3.37-3)$$

برای $1 \leq x = 0, 1, 2, \dots, M-1$ تعریف می شوند. توضیحات قبلی راجع به $f_r(x)$ و $g_r(x)$ ، تناوب فرضی این توابع و انتخاب مقادیر M ، در معادله (۳.۳۷-۳) نیز معتبر هستند. در حالت دو بعدی نیز توضیحات مشابهی صدق می کند. بنابراین اگر $(y, f(x))$ و $(y, g(x))$ توابعی از

۱۴۷ ۳.۳ چند خاصیت از تبدیل فوریه دو بعدی

متغیرهای پیوسته باشند، همبستگی آنها به صورت

$$f(x, y) \circ g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(\alpha, \beta) g(x + \alpha, y + \beta) d\alpha d\beta. \quad (38-3. 3)$$

تعریف می شود. در حالت گستته نیز برای $1 \leq M = x \leq N = y \leq 1$ داریم:

$$f_r(x, y) \circ g_r(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_r^*(m, n) g_r(x + m, y + n) \quad (39-3. 3)$$

همانند کانولوشن گستته، $f_r(x, y)$ و $g_r(x, y)$ توابعی گسترش یافته هستند و M و N با استفاده از معادلات $(33-3)$ و $(34-3)$ انتخاب می شوند تا از خطای پیچشی در دوره های تابع همبستگی جلوگیری شود.

برای هر دو حالت پیوسته و گستته، قضیه همبستگی به صورت زیر برقرار است:

$$f(x, y) \circ g(x, y) \Leftrightarrow F^*(u, v) G(u, v) \quad (40-3. 3)$$

و

$$f^*(x, y) g(x, y) \Leftrightarrow F(u, v) \circ G(u, v). \quad (41-3. 3)$$

هر یک از توابع وقتی برای متغیرهای گستته بکار روند، گسترش یافته و متناظر فرض می شوند. یکی از کاربردهای اصلی همبستگی در پردازش تصویر، در زمینه تطبیق الگو یا نمونه¹ است. همان طور که در فصل ۹ بیان خواهد شد، در این مبحث، مسأله پیدا کردن نزدیک ترین تطبیق بین یک تصویر ناشناخته و مجموعه ای از تصاویر معین است. یک روش این است که همبستگی بین تصویر ناشناخته و هر یک از تصاویر معین محاسبه شود. آنگاه با انتخاب تصویری که بزرگترین مقدار را برای تابع همبستگی نتیجه می دهد، نزدیکترین تطبیق پیدا می شود. چون همبستگی های حاصل، توابعی دو بعدی هستند، این کار شامل جستجوی بزرگترین دامنه هر تابع می باشد. مانند کانولوشن گستته، در اغلب موارد محاسبه $f_r(x, y) \circ g_r(x, y)$ در میدان فرکانس و با استفاده از الگوریتم FFT که برای بدست آوردن تبدیلات مستقیم

1) Template or prototype matching

۱۴۸ پردازش تصویر رقمه

و معکوس بکار می رود، پربازده تر است.

هنگام مقایسه نتایج کانولوشن یا همبستگی گسته با کانولوشن یا همبستگی پیوسته باید توجه داشت که در تعریف ما موارد گسته معادل انتگرال مستطیلی موارد پیوسته هستند. بنابراین برای مقایسه نتایج گسته و پیوسته در یک مبنای یکسان، باید معادلات $(3.29-3)$ و $(3.37-3)$ را در Δx و معادلات $(3.25-3)$ و $(3.39-3)$ را در $\Delta y \Delta x$ ضرب کرد، که طبق تعاریف بیان شده در بخش ۳.۲ همان فاصله بین نمونه ها هستند. مثلاً در شکل ۳.۱۴ نتایج کانولوشن پیوسته و گسته هم دامنه و به همان دامنه بین نمونه ها هستند. با این حال تنها در محاسبه و ارزیابی صور گسته، وارد هستند، زیرا نتیجه گسته در Δx ضرب شد. با این حال تنها در محاسبه و ارزیابی صور گسته، وارد کردن این ضرایب مقیاس ترجیح داده می شود. گذشته از این، اگر $f(x)$ و $g(x)$ - همراه تبدیلات متناظر شان - با هم عوض شوند، تمام عبارات کانولوشن و همبستگی برقرار هستند. این حالت در مورد نتایج دو بعدی نیز درست است.

۹.۳.۳ نمونه برداری

در بخش ۲.۳ مفهوم پایه ای نمونه برداری تصویر را به طور شهودی معرفی کردیم. تبدیل فوریه و قضیه کانولوشن، ابزارهای تحلیل عمیق تر این مسأله را فراهم می کنند. به ویژه می خواهیم به این سؤال پاسخ دهیم که چه میزان نمونه باید برداشته شود تا هیچ اطلاعاتی در فرآیند نمونه برداری از دست نرود. به عبارت دیگر، مسأله تعیین شرایط نمونه برداری است به طوری که تحت این شرایط بتوان تصویر پیوسته را به طور کامل از مجموعه مقادیر نمونه برداری شده بازسازی کرد. تحلیل رابا حالت یک بعدی شروع می کنیم.

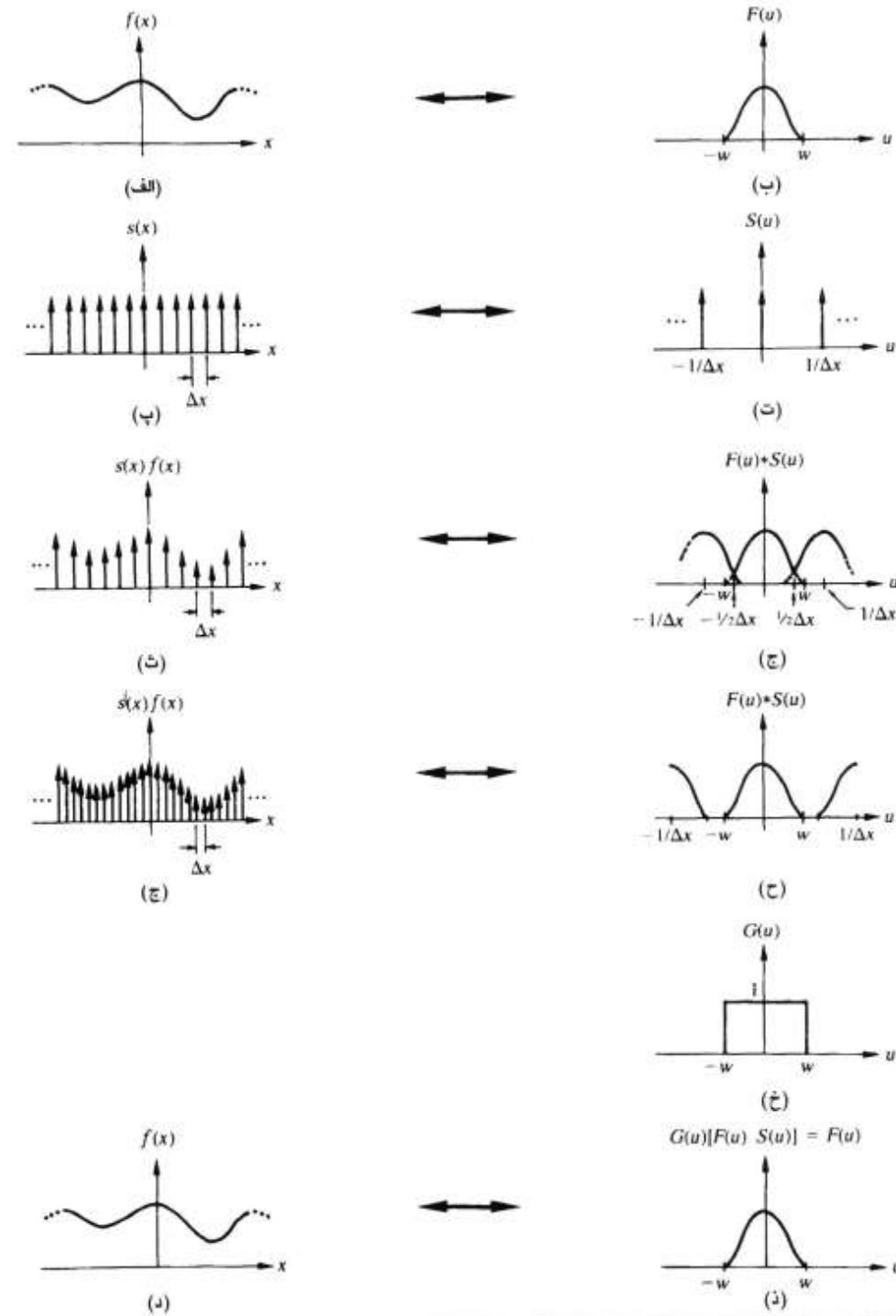
تابع یک بعدی

تابع شکل ۳.۱۷. (الف) را که فرض می شود از ∞ - ∞ گسترش یافته است، در نظر بگیرید. فرض کنید که تبدیل فوریه $f(x)$ برای مقادیر W خارج از بازه $[W_1, W_2]$ صفر باشد. ممکن است تبدیل به صورت شکل ۳.۱۷. (ب) دیده شود^۱. تابعی که تبدیلش بازی یک مقدار محدود W این خاصیت را داشته باشد، تابع عرض باند محدود^۱ نامیده می شود.

^۱ به یاد آورید که تبدیل فوریه تابعی مختلط است. در نمایش های گرافیکی بعدی، برای سادگی تنها دامنه تبدیلات را نمایش می دهیم. با این حال محور عمودی، برچسبهای $G(u), F(u)$ و غیره دارد تا نشان دهد که مقاهیم و نتایج بدست آمده نه تنها برای دامنه، بلکه برای تبدیل کامل نیز معتبر هستند.

1) Band- limited function

۳.۳ چند خاصیت از تبدیل فوریه دو بعدی ۱۴۹



شکل ۳۱۷.۳ نمایش گرافیکی مفاهیم نمونه برداری

بدست آوردن گونه نمونه برداری شده (x) ^۵ شامل ضرب این تابع درتابع نمونه برداری (x) ^۶ می باشد، که تابع نمونه برداری مانند شکل ۳.۱۷. (ب) از یک قطار ضربه با فواصل Δx از هم، تشکیل می شود. طبق قضیه کانولوشن، ضرب در میدان x معادل کانولوشن در میدان فرکانس است. بنابراین برای حاصل ضرب $(x)f(x)$ ^۷ تبدیل فوریه ای که در شکل ۳.۱۷. (ج) دیده می شود، بدست می آید. تبدیل متناوب دوره تناوب $\Delta x/1$ دارد و ممکن است دوره های متفاوت (u) همپوشانی داشته باشند. مثلاً در اولین دوره اگر مقدار $1/2\Delta x$ کمتر از W باشد، مرکز ناحیه همپوشانی در $x = 1/2\Delta x = u$ رخ می دهد. بنابراین برای جلوگیری از این مشکل فاصله نمونه برداری Δx را طوری انتخاب می کنیم که $W \geq 1/2\Delta x$ باشد.

$$\Delta x \leq \frac{1}{2W} \quad (42-3.3)$$

در اشکال ۳.۱۷. (ج) و (ح) نتیجه 'کاهش Δx ' دیده می شود. نتیجه 'اصلی'، جداسازی دوره ها است طوری که هیچ همپوشانی رخ ندهد. همان طور که در شکل ۳.۱۷. (ذ) دیده می شود، اهمیت این عمل در این است که با ضرب تبدیل موجود در شکل ۳.۱۷. (ح) در تابع

$$G(u) = \begin{cases} 1 & -W \leq u \leq W \\ 0 & \text{سایر جاهای} \end{cases} \quad (43-3.3)$$

جداسازی کامل (u) ممکن می شود. آنگاه تبدیل فوریه معکوس تابع پیوسته اولیه $(x)f(x)$ ^۸ را بدست می دهد. بازسازی کامل یک تابع باند محدود از نمونه هایی که فاصله شان شرط معادله (۳.۴۲-۳) را برآورده می کند، به عنوان قضیه نمونه برداری ویتکر-شانون^۹ شناخته می شود.

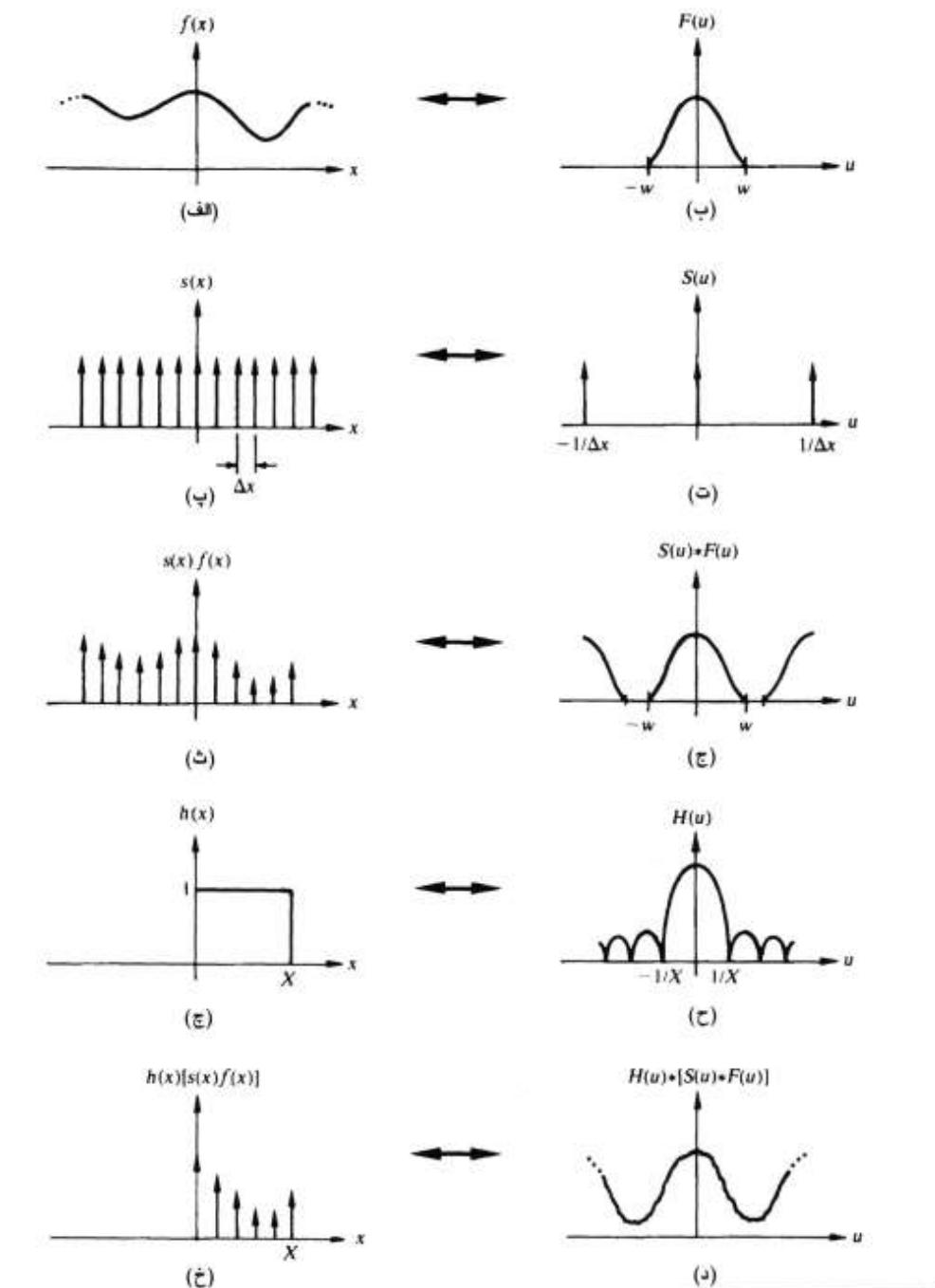
تمام اطلاعات میدان فرکانس یک تابع باند محدود در بازه $[W, -W]$ می باشد. با این حال اگر معادله (۳.۴۲-۳) برآورده نشود، حاصل تبدیل در این بازه بر اثر تداخل دوره های مجاور خراب می شود. این پدیده که روی هم افتادگی فرکانسی^{۱۰} نامیده می شود، بازسازی کامل یک تابع کم نمونه برداری شده را غیرممکن می سازد.

نتایج قبلی در مورد توابعی می باشد که دوره نامحدودی در میدان x دارند که مستلزم بازه نمونه برداری نامحدود می باشد. اما در بررسی عملی موضوع، تابعی که در یک ناحیه محدود نمونه برداری شده است، مورد نظر است. این حالت به طور گرافیکی در شکل ۳.۱۸. نمایش داده شده

۱) Whittaker- Shannon

۲) Aliasing

۱۵۱ ۳.۳ چند خاصیت از تبدیل فوریه دو بعدی



شکل ۱۸.۳ نمایش گرافیکی مفاهیم نمونه برداری محدود

1) Finite-sampling

است. قسمت های (الف) - (ج) شکل ۳.۱۸. ۱۷. هستند، جز این که فرض می شود فاصله بین نمونه ها قضیه نمونه برداری را برآورده می کند، طوری که هیچ روی هم افتادگی فرکانسی رخ ندهد، برای نمایش ریاضی بازه نمونه برداری محدود $[0, X]$ کافیست حاصل نمونه برداری که در شکل

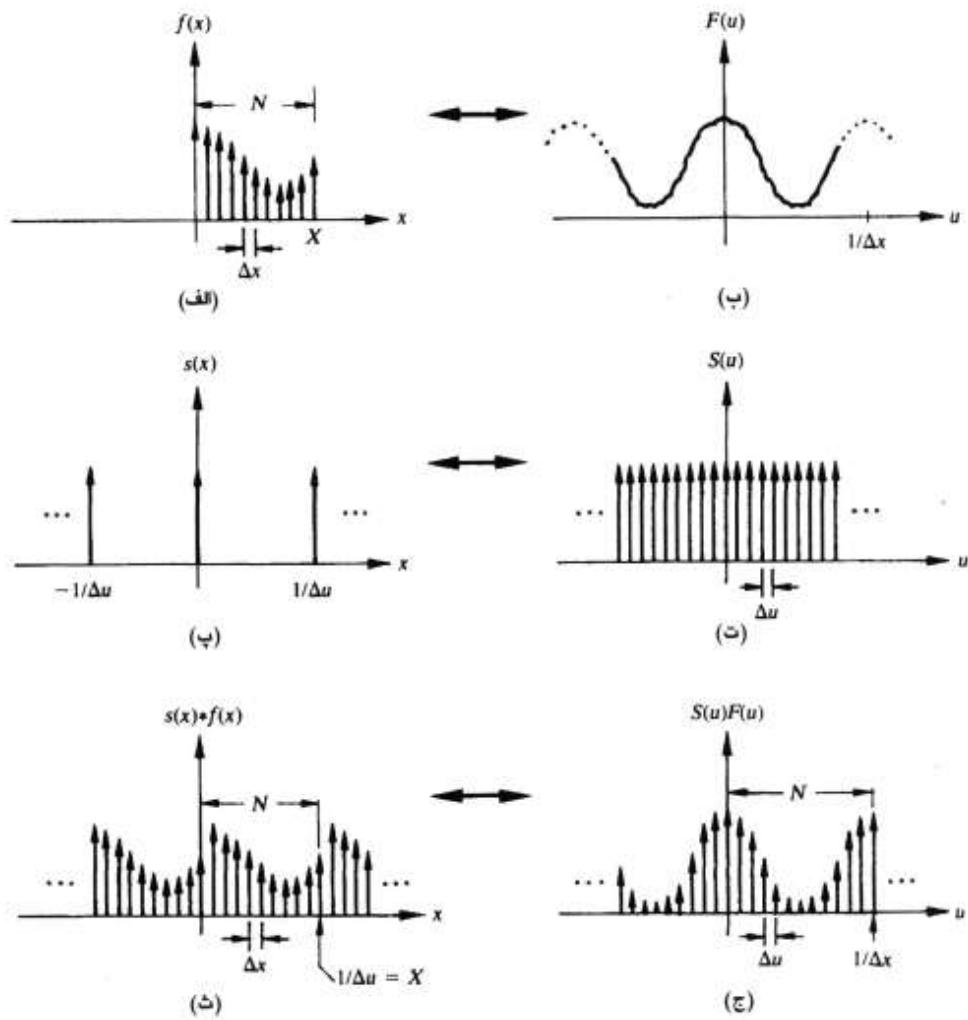
۳.۱۸. (ث) دیده می شود، را در تابع

$$h(x) = \begin{cases} 1 & 0 \leq x \leq X \\ 0 & \text{هر جای دیگر} \end{cases} \quad (44-3. ۳)$$

ضرب کنیم. این تابع، که اغلب پنجره خوانده می شود، و تبدیل فوریه اش به ترتیب در اشکال ۳.۱۸. (ج) و (ح) دیده می شوند. نتایج ضرب در اشکال ۳.۱۸. (خ) و (د) دیده می شوند. نتیجه نهایی در میدان فرکانس با کانولوشن تابع $S(u)*F(u)$ در $H(u)$ ، که تبدیل فوریه تابع پنجره ای $h(x)$ است، بدست می آید. چون مؤلفه های فرکانسی $H(u)$ تابی نهایت ادامه دارد، لذا کانولوشن این توابع، اعوجاجی در نمایش میدان فرکانس تابعی که نمونه برداری شده و به ناحیه $h(x)$ محدود است، وارد می کند که شکل ۳.۱۸. (د) نیز این مورد را نشان می دهد. بنابراین حتی اگر فاصله بین نمونه ها قضیه نمونه برداری را برآورده کند، بازسازی کامل تابعی که تنها در یک ناحیه محدود میدان x نمونه برداری شده است، معمولاً غیرممکن است. تحت این شرایط تنها وقتی (x) باشد محدود و متنابض با دوره X باشد، می توان تبدیل فوریه اولیه را بدست آورد. در این حالت خرابی های ناشی از $H(u)$ حذف می شود و اگر شرایط قضیه نمونه برداری برآورده شوند، بازسازی کامل (x) ممکن می گردد. تابع بازسازی شده در بازه $-\infty$ تا ∞ گسترده است و خارج بازه ای که $h(x)$ صفر است، صفر نیست. این نکات منجر به این نتیجه مهم می شوند که هیچ تابع (x) با دوره محدود نمی تواند باشد محدود باشد-بر عکس، تابعی که باشد محدود است، باید در میدان x از $-\infty$ تا ∞ گسترده باشد. این نتیجه، مهم و کاربردی است زیرا محدودیت اساسی کار با توابع رقمه را تعیین می کند.

قبل از ترک توضیح توابع یک بعدی، اجازه دهید با استفاده از نتایج قبلی، دلیل دیگری برای متنابض بودن تبدیل فوریه گسته ارائه دهیم. تاکنون تمام نتایج در میدان فرکانس، طبیعت پیوسته داشته اند. برای بدست آوردن تبدیل فوریه گسته کافی است که با یک قطار ضربه که ضربه ها به فاصله Δu از هم باشند، از تبدیل پیوسته نمونه برداریم. شکل ۳.۱۹. این وضعیت را بر اساس اشکال ۳.۱۸. (خ) و (د) رسم می کند. نمادهای (x) و $F(u)$ که در شکل ۳.۱۹. استفاده شده اند، مقایسه با توضیح بخش ۲.

۳.۳ چند خاصیت از تبدیل فوریه دو بعدی ۱۵۳



شکل ۳.۳ نمایش گرافیکی تبدیل فوریه گسته

را آسان می کنند. با این حال توجه کنید که فرض می شود اشکال ۳.۱۹ (الف) و (ب) حاصل دنباله عملیات نشان داده شده در شکل ۳.۱۸ هستند.

همان طور که قبلاً اشاره شد، می توان با ضرب قطار ضربه در تابع مورد نظر حاصل نمونه برداری را بدست آورد. در این حالت، ضرب $F(u)$ در $S(u)$ شکل ۳.۱۹ (ج) را نتیجه می دهد. عمل معادل در میدان x ، کانولوشن است که تابع شکل ۳.۱۹ (ث) را می دهد. این تابع متناظر با دوره $1/\Delta u$ است. اگر N نمونه از $f(x)$ و $s(x)$ برداشته شود و فاصله بین نمونه ها طوری انتخاب شوند که یک دوره در هر

میدان با N نمونه هم فاصله پوشانده شود، آن گاه در میدان $x, X = N\Delta x$ و در میدان فرکانس $\Delta u = 1/\Delta x$. همان طور که قبلآ نشان داده شد، معادله آخر براساس خاصیت تناوبی بودن تبدیل فوریه تابع نمونه برداری شده، با دوره $1/\Delta x$ ، است. بنابراین،

$$\Delta u = \frac{1}{N\Delta x} \quad (45-3.3)$$

که موافق با معادله (۳.۴-۲) است. انتخاب این فاصله تابع شکل ۳.۱۹(ث) را نتیجه می دهد که متناوب با دوره $1/\Delta u$ است. از معادله (۳.۴۵-۳)، $X = N\Delta x = 1/\Delta u$ که معادل کل بازه نمونه برداری در شکل ۳.۱۹(الف) است.

توابع دو بعدی

مفاهیم قبلی نمونه برداری را می توان با کمی تغییر در طرز نمایش، مستقیماً برای توابع دو بعدی بکاربرد. با استفاده از تابع ضربه دو بعدی $f(x, y)$ ، که به صورت

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x - x_0, y - y_0) dx dy = f(x_0, y_0). \quad (46-3.3)$$

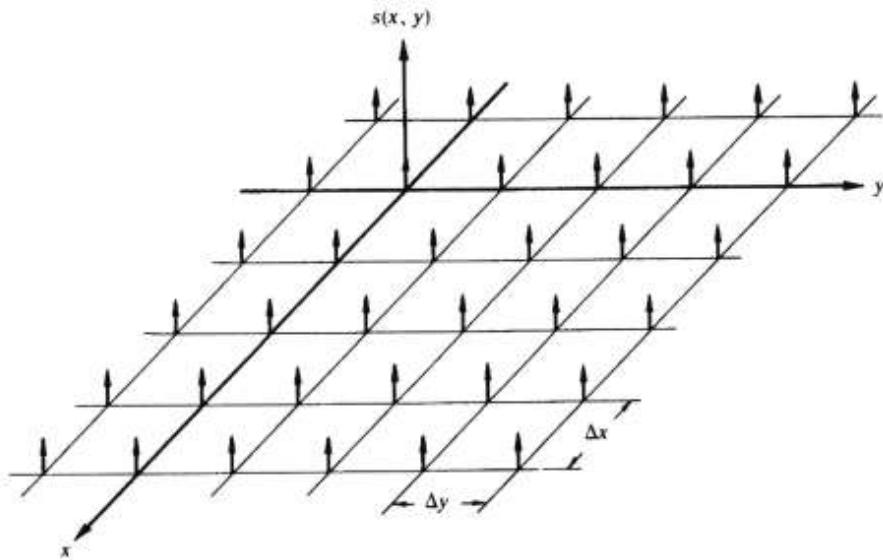
تعریف می شود، می توان نمونه برداری تابع دو بعدی را با روابط ریاضی بیان کرد. تعبیر معادله (۳.۴۶-۳) شبیه تعبیر معادلات (۳.۲۴-۳) و (۳.۲۵-۳) است. در شکل ۳.۲۰. یک تابع نمونه برداری دو بعدی دیده می شود که از یک قطار ضربه که فواصل ضربه ها در جهت محورهای x و y به ترتیب برابر Δx و Δy است، تشکیل می شود.

برای تابع $f(x, y)$ که x و y پیوسته هستند، تابع نمونه برداری شده با تشکیل حاصل ضرب $s(x, y) = f(x, y) \delta(x - x_0, y - y_0)$ می باشد که $s(x, y)$ بحسب می آید. عمل معادل در میدان فرکانس، کانولوشن $S(u, v)$ و $F(u, v)$ می باشد که $S(u, v) = F(u, v) * s(x, y)$ باشد. اگر $f(x, y)$ باشد محدود باشد (یعنی از ضربه ها با فواصل $1/\Delta x$ و $1/\Delta y$ به ترتیب درجهات W_x و W_y است. اگر $f(x, y)$ باشد محدود باشد (یعنی تبدیل فوریه اش خارج ناحیه محدود R ، صفر باشد)، نتیجه کانولوشن $S(u, v) = F(u, v) * s(x, y)$ ممکن است

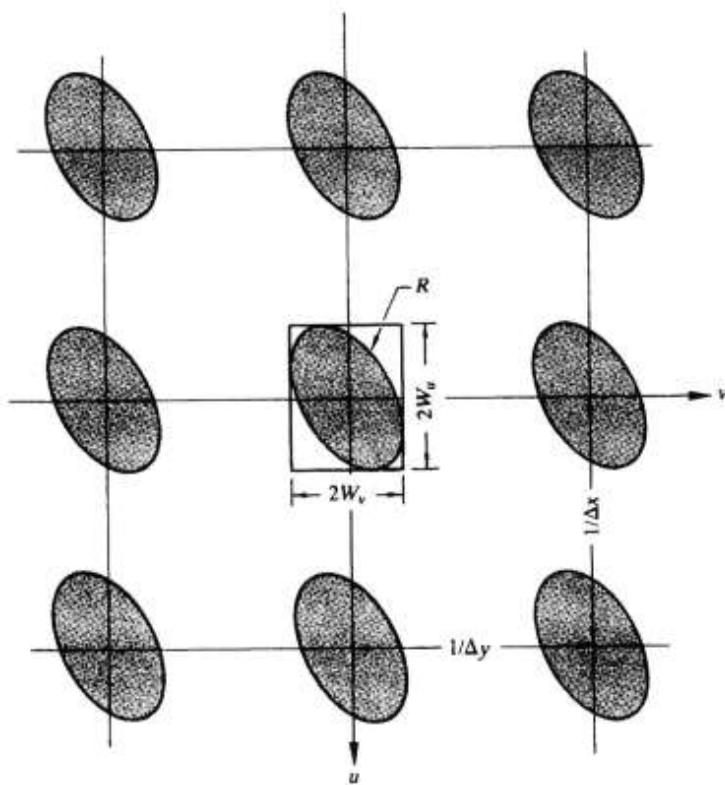
شبیه شکل ۳.۲۱ باشد. تابع نشان داده شده، در دو بعد متناوب است.

فرض کنید $W_x < 2$ و $W_y < 2$ به ترتیب بیانگر ابعاد درجهات W_x و W_y کوچکترین مستطیلی باشد که کاملاً ناحیه R را محصور می کند. آن گاه از شکل ۳.۲۱، اگر $W_x > 2$ و $W_y > 2$ (بدون روی هم افتادگی فرکانسی) باشد، یکی از دوره ها را می توان با ضرب $S(u, v) = F(u, v) * s(x, y)$ در تابع

۱۵۵ ۳.۳ چند خاصیت از تبدیل فوریه دو بعدی



شکل ۲۰.۳ یک تابع نمونه برداری دو بعدی



شکل ۲۱.۳ نمایش میدان فرکانس یک تابع عرض محدود دو بعدی نمونه برداری شده

$$G(u, v) = \begin{cases} 1 & \text{اگر } (u, v) \text{ درون یکی از مستطیلهای} \\ & \text{محصور کننده } R \text{ باشد} \\ 0 & \text{در سایر جاهای} \end{cases} \quad (47-3, 3)$$

به طور کامل بازسازی کرد. تبدیل معکوس فوریه $f(x, y) = G(u, v)[S(u, v)*F(u, v)]$ را نتیجه می دهد.

ملاحظات پیشین منجر به شکلی از قضیه نمونه برداری دو بعدی می شوند، که در آن می توان تابع عرض باند محدود $(y, f(x, y))$ را به طور کامل از نمونه هایش که فاصله آنها در روابط

$$\Delta x \leq \frac{1}{2W_x} \quad (48-3, 3\text{الف})$$

و

$$\Delta y \leq \frac{1}{2W_y} \quad (48-3, 3\text{ب})$$

صدق کند، بازسازی کرد.

وقتی $f(x, y)$ با استفاده از پنجره دو بعدی $h(x, y) = h(x, y)$ که مشابه تابع $h(x)$ مورد استفاده در شکل ۳ است، از نظر مکانی محدود شده باشد، با کانولوشن $S(u, v) = H(u, v)*F(u, v)$ تبدیل تابع نمونه برداری شده دارای اعوجاج می شود. این اعوجاج که بر اثر محدودیت مکانی تصاویر رقمه است، بازسازی کامل $(y, f(x, y))$ از نمونه هایش را غیرممکن می کند. مانند حالت یک بعدی، توابع متناوب مستثنی هستند، اما تصاویر دارای متناوب بندرت در عمل وجود دارند.

توضیحی مشابه آنچه برای حالت یک بعدی ارائه شد، را می توان در مورد چگونگی ایجاد متناوب در تبدیل فوریه گستته دو بعدی نیز ارائه کرد. این تحلیل برای تصویر $N \times N$ نتیجه می دهد که:

$$\Delta u = \frac{1}{N\Delta x} \quad (49-3, 3\text{الف})$$

و

$$\Delta v = \frac{1}{N\Delta y} \quad (48-3, 3\text{ب})$$

برقراری این روابط بین فاصله نمونه ها تضمین می کند که در هر یک از میدانهای مکان و فرکانس یک دوره

٤.٣ تبدیل فوریه سریع ۱۵۷

تناوب کامل (دو بعدی) با $N \times N$ نمونه هم فاصله پوشش داده می شود.

٤.٣ تبدیل فوریه سریع

تعداد ضرب ها و جمع های مختلط مورد نیاز در پیاده سازی معادله (۲-۲. ۳) متناسب با N^2 است. یعنی، برای هریک از N مقدار x ، بسط مجموع، نیازمند N ضرب مختلط $f(x)$ با $[f(Nx)]$ است. $\exp[-j2\pi ux/N]$ را می توان یکبار محاسبه و برای $0 \leq u \leq N-1$ جمع حاصل ضربها است. عبارات $\exp[-j2\pi ux/N]$ را می توان یکبار محاسبه و برای استفاده های بعدی در جدول ذخیره کرد. بدین دلیل معمولاً در این عبارات ضرب x در $\exp[-j2\pi ux/N]$ در بخش مستقیم از پیاده سازی در نظر گرفته نمی شود.

با تفکیک مناسب معادله (۲-۲. ۳) می توان تعداد اعمال ضرب و جمع را متناسب با $N \log_2 N$ نمود. روش تفکیک، الگوریتم تبدیل فوریه سریع^۱ (FFT) نامیده می شود. همان طور که جدول ۱. ۳ نشان می دهد، کاهش نسبی تعداد اعمال محاسباتی از N^2 به $N \log_2 N$ بیانگر صرف جویی چشمگیر در حجم محاسبات است. واضح است که روش FFT به ویژه وقتی N نسبتاً بزرگ باشد، بهره محاسباتی

جدول ۱.۳ مقایسه N^2 و $N \log_2 N$ بازای مقادیر مختلف N

N	N^2	$N \log_2 N$	بهره محاسباتی ($N/\log_2 N$)
2	4	2	2.00
4	16	8	2.00
8	64	24	2.67
16	256	64	4.00
32	1,024	160	6.40
64	4,096	384	10.67
128	16,384	896	18.29
256	65,536	2,048	32.00
512	262,144	4,608	56.89
1024	1,048,576	10,240	102.40
2048	4,194,304	22,528	186.18
4096	16,777,216	49,152	341.33
8192	67,108,864	106,496	630.15

1) Fast Fourier Transform

قابل توجهی نسبت به پیاده سازی مستقیم تبدیل فوریه دارد. به عنوان مثال فرض کنید که محاسبه 'FFT آرایه^۱ ۸۱۹۲ نقطه ای با رایانه' معمولی ۵ ثانیه طول بکشد. همان ماشین حدود ۶۰۰ بار بیشتر (۵۰ دقیقه) زمان لازم دارد تا تبدیل فوریه همان آرایه را با استفاده از معادله (۳-۲) محاسبه کند. در بخش بعدی، الگوریتم FFT یک متغیره را ارایه می کنیم. همان طور که در بخش ۳. ۳. ۱ نشان داده شد، می توان با دو تبدیل یک بعدی متوالی، تبدیل فوریه^۲ دو بعدی را بدست آورد.

۱۰.۴.۳ FFT الگوریتم

الگوریتم FFT که در این بخش بیان شد، براساس روش مشهور به دو برابر کردن متوالی^۱ است. برای سهولت، معادله (۳-۲-۲) را به صورت

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) W_N^{ux} \quad (1-4. 3)$$

بیان می کنیم که در آن

$$W_N = \exp[-j2\pi/N] \quad (2-4. 3)$$

و فرض می شود N به صورت

$$N = 2^n \quad (3-4. 3)$$

باشد که n یک عدد صحیح مثبت است. بنابراین N را می توان به صورت

$$N = 2M \quad (4-4. 3)$$

بیان کرد که M نیز عدد صحیح مثبت است. جایگذاری معادله (۴-۴. ۳) در معادله (۱-۴. ۳) نتیجه می دهد که:

$$\begin{aligned} F(u) &= \frac{1}{2M} \sum_{x=0}^{2M-1} f(x) W_{2M}^{ux} \\ &= \frac{1}{2} \left[\frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_{2M}^{u(2x)} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_{2M}^{u(2x+1)} \right]. \end{aligned} \quad (5-4. 3)$$

1) Successive doubling

۴.۳ تبدیل فوریه سریع ۱۵۹

از معادله (۳.۴-۲)، $W_{2M}^{2u} = W_M^{u*}$ ، بنابراین معادله (۳.۴-۵) را می‌توان به صورت

$$F(u) = \frac{1}{2} \left[\frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_M^{ux} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_M^{ux} W_{2M}^u \right]. \quad (6-4.3)$$

بیان کرد. با تعریف

$$F_{\text{even}}(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_M^{ux} \quad (7-4.3)$$

برای $u = 0, 1, 2, \dots, M-1$

$$F_{\text{odd}}(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_M^{ux} \quad (8-4.3)$$

برای $u = 0, 1, 2, \dots, M-1$ ، معادله (۶-۴.۳) به صورت

$$F(u) = \frac{1}{2} [F_{\text{even}}(u) + F_{\text{odd}}(u) W_{2M}^u]. \quad (9-4.3)$$

(۹-۴.۳)-(۷-۴.۳)، از معادلات

نتیجه می‌شود:

$$F(u+M) = \frac{1}{2} [F_{\text{even}}(u) - F_{\text{odd}}(u) W_{2M}^u]. \quad (10-4.3)$$

تحلیل دقیق معادلات (۳.۴-۷)-(۹-۴.۳) چند خاصیت جالب این عبارات را آشکار می‌کند.

همان‌طور که در معادلات (۳.۴-۹) و (۱۰-۴.۳) نشان داده شد، می‌توان با تقسیم عبارت اولیه به دو

قسمت، تبدیل N نقطه‌ای را محاسبه نمود. محاسبه نیمه اول $F(u)$ نیازمند تعیین مقادیر دو تبدیل $(N/2)$

نقطه‌ای داده شده با استفاده از معادلات (۳.۴-۷) و (۳.۴-۸) می‌باشد. آن گاه مقادیر حاصل از $F_{\text{even}}(u)$

و $F_{\text{odd}}(u)$ در معادله (۹-۴.۳) جایگذاری می‌شوند تا $F(u)$ برای $u = 0, 1, 2, \dots, (N/2)-1$ بدست آید.

بسی نیمه دیگر تبدیل مستقیماً از معادله (۳.۴-۱۰) بدون محاسبات اضافه بدست می‌آید.

به منظور بررسی جنبه‌های محاسباتی این روش، اجازه دهید $m(n)$ و $a(n)$ به ترتیب بیانگر تعداد

ضرب‌ها و جمع‌های مختلط مورد نیاز در پیاده‌سازی آن باشد. مانند قل، تعداد نمونه‌ها n^2 است که n

یک عدد صحیح مثبت است. ابتدا فرض کنید که $n = 1$. تبدیل دو نقطه‌ای نیازمند محاسبه $F(0)$ است؛

۱۶۰ پردازش تصویر رقمه

آن گاه (۱) از معادله $(3 \cdot ۴ - ۱۰)$ بدست می آید. برای تعیین $F(0)$ ابتدا باید که $F_{\text{even}}(0)$ و $F_{\text{odd}}(0)$ محاسبه شوند. در این حالت $M = 1$ و معادلات $(3 \cdot ۴ - ۷)$ و $(3 \cdot ۴ - ۸)$ تبدیلات یک نقطه‌ای هستند. چون تبدیل فوریه یک نقطه تکی، خود آن نقطه است؛ پس هیچ ضرب و جمعی در بدست آوردن (0) و $F_{\text{even}}(0)$ لازم نیست. طبق معادله $(3 \cdot ۴ - ۹)$ ، با ضرب $F_{\text{odd}}(0)$ در W_2^0 و انجام یک جمع، $F(0)$ بدست می آید. آنگاه (۱) با یک جمع بیشتر، از معادله $(3 \cdot ۴ - ۱۰)$ بدست می آید (تفريق همانند جمع در نظر گرفته می شود). چون $W_2^0 F_{\text{odd}}(0)$ قبلًا محاسبه شده بود، تعداد کل اعمال مورد نیاز در تبدیل دو بعدی $= 1$ ضرب و $2 = (1)$ جمع می باشد.

مقدار بعدی مجاز برای $n = 2$ است. طبق مطالب فوق، تبدیل $\frac{1}{2}$ نقطه‌ای را می توان به دو قسمت تقسیم نمود. نیمه اول $(u) F(u)$ نیازمند محاسبه 'دو تبدیل دونقطه‌ای'، مانند معادلات $(3 \cdot ۴ - ۷)$ و $(3 \cdot ۴ - ۸)$ برای $M = 2$ است. تبدیل دو نقطه‌ای نیازمند (1) ضرب و (1) جمع است، بنابراین محاسبه 'این دو معادله نیازمند جمماً (1) ضرب و (1) جمع است. برای بدست آوردن (0) و (1) از $F(0)$ و $F(1)$ از معادله $(3 \cdot ۴ - ۹)$ ، دو جمع و ضرب دیگر مورد نیاز است. چون $W_{2M}^0 F_{\text{odd}}(u)$ قبلًا برای $\{0, 1\}$ می باشد، با دو جمع بیشتر (2) و (3) بدست می آیند. پس در مجموع $m(2) = 2m(1) + 2 + 4 = 2a(1) + 2$ است.

وقتی n برابر 3 باشد، دو تبدیل $\frac{1}{2}$ نقطه‌ای در محاسبه $(u) F_{\text{even}}(u)$ و $(u) F_{\text{odd}}(u)$ در نظر گرفته می شوند. آن دو به (2) ضرب و (2) جمع نیاز دارند. با چهار ضرب و 8 جمع دیگر تبدیل کامل بدست می آید. پس در مجموع $m(3) = 2m(2) + 4 = 2a(2) + 8$ است.

با ادامه' این روتند برای هر مقدار صحیح مثبت n ، عبارات بازگشتی زیر را برای تعداد ضربها و جمعهای مورد نیاز در پیاده سازی FFT بدست می آوریم:

$$m(n) = 2m(n-1) + 2^{n-1} \quad n \geq 1 \quad (11-۴.۳)$$

و

$$a(n) = 2a(n-1) + 2^n \quad n \geq 1 \quad (12-۴.۳)$$

که $m(0) = 0$ و $a(0) = a$ ، زیرا تبدیل یک نقطه تکی به هیچ ضرب و جمعی نیاز ندارد.

٤.٣ تبدیل فوریه سریع ١٦١

با پیاده سازی معادلات (۳.۷-۴) و (۳.۱۰-۴)، الگوریتم FFT دو برابر سازی متوالی تشکیل می شود. این نام از روش محاسبه تبدیل دو نقطه ای از حاصل دو تبدیل یک نقطه ای، تبدیل ۴ نقطه ای از دو تبدیل دو نقطه ای و الی آخر، برای هر مقدار N که توان صحیح از ۲ باشد، گرفته شده است.

٤.٢.٣ تعداد عملیات

با استقرار، تعداد ضرب ها و جمع های مختلط مورد نیاز در پیاده سازی الگوریتم FFT، به ترتیب عبارتند از:

$$\begin{aligned} m(n) &= \frac{1}{2} 2^n \log_2 2^n \\ &= \frac{1}{2} N \log_2 N \\ &= \frac{1}{2} Nn \quad n \geq 1 \end{aligned} \quad (13-4.3)$$

$$\begin{aligned} a(n) &= 2^n \log_2 2^n \\ &= N \log_2 N \\ &= Nn \quad n \geq 1 \end{aligned} \quad (14-4.3)$$

اثبات استقراری با نشان دادن این که معادلات (۳.۱۳-۴) و (۳.۱۴-۴) برای $n = 1$ برقرارند، شروع می شود. به یادآورید که

$$m(1) = \frac{1}{2}(2)(1) = 1 \quad \text{و} \quad a(1) = (2)(1) = 2.$$

سپس فرض می شود که عبارات برای n برقرارند. آنگاه باید ثابت کنیم که آنها برای $n + 1$ هم درست هستند.

از معادله (۳.۱۱-۴) داریم

$$m(n + 1) = 2m(n) + 2^n.$$

با جایگذاری مقدار $m(n)$ در معادله (۳.۱۳-۴)، که فرض می شود برای n درست است، داریم:

$$\begin{aligned}m(n+1) &= 2\left(\frac{1}{2}Nn\right) + 2^n \\&= 2\left(\frac{1}{2}2^n n\right) + 2^n \\&= 2^n(n+1) \\&= \frac{1}{2}(2^{n+1})(n+1).\end{aligned}$$

بنابراین معادله (۳.۴-۱۳) برای تمام مقادیر صحیح مثبت n درست است.

از معادله (۳.۴-۱۲) داریم

$$a(n+1) = 2a(n) + 2^{n+1}$$

با جایگذاری مقدار $a(n)$ در معادله (۳.۴-۱۴) داریم

$$\begin{aligned}a(n+1) &= 2Nn + 2^{n+1} \\&= 2(2^nn) + 2^{n+1} \\&= 2^{n+1}(n+1)\end{aligned}$$

که اثبات را کامل می کند.

۳.۴.۳ معکوس FFT

تاکنون درباره تبدیل فوریه معکوس مطالب کمی گفته ایم. علت این است که هر الگوریتم پیاده سازی تبدیل مستقیم گسته، رامی توان (با کمی تغییرات) برای محاسبه معکوس نیز به کاربرد. برای نشان دادن این واقعیت، اجازه دهید به معادلات (۳.۲-۲) و (۳.۲-۳) باز گردیم، که در زیر تکرار شده اند:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \exp[-j2\pi ux/N] \quad (15-4.3)$$

و

$$f(x) = \sum_{u=0}^{N-1} F(u) \exp[j2\pi ux/N]. \quad (16-4.3)$$

با مزدوج مختلط کردن معادله (۱۶-۴.۳) و تقسیم طرفین آن بر N ، داریم

۴.۳ تبدیل فوریه سریع ۱۶۳

$$\frac{1}{N} f^*(x) = \frac{1}{N} \sum_{u=0}^{N-1} F^*(u) \exp[-j2\pi ux/N]. \quad (17-4.3)$$

مقایسه این نتیجه با معادله (۳.۱۵-۴) نشان می دهد که طرف راست معادله (۳.۱۷-۴) به صورت تبدیل فوریه مستقیم است. بنابراین با دادن (u) F^* به الگوریتمی که برای محاسبه تبدیل مستقیم طراحی شده است، $f^*(x)/N$ بدست می آید، که با مزدوج مختلط کردن آن و ضرب در N ، تبدیل معکوس مطلوب، $f(x)$ ، بدست می آید.

برای آرایه های مربعی دو بعدی مزدوج مختلط معادله (۳.۱۰-۲۰) را می گیریم، یعنی،

$$f^*(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F^*(u, v) \exp[-j2\pi(ux + vy)/N] \quad (18-4.3)$$

که به شکل تبدیل دو بعدی مستقیم معادله (۳.۹-۲۰) است. بنابراین با دادن (u, v) F^* به الگوریتمی که برای محاسبه تبدیل مستقیم طراحی شده است، $(x, y) f^*$ بدست می آید، که با مزدوج مختلط کردن حاصل، $(y, x) f$ بدست می آید. وقتی $f(x, y)$ یا $f(y, x)$ حقیقی باشد، عمل مزدوج مختلط کردن لازم نیست، زیرا برای توابع حقیقی $(x) f^* = f^*(x)$ و $f(x, y) = f^*(y, x)$ می باشد.

وقتی از روش قبلی جهت بدست آوردن تبدیل معکوس استفاده شود، محاسبه تبدیل دو بعدی با مرورهای متوالی تبدیل یک بعدی، یک منبع متداول خطای باشد. روایی که در بخش ۳.۱ بیان شد، را به یاد داشته باشید و مراقب باشید توسط معادله (۳.۱۷-۴) گمراه نشود. به بیان دیگر، وقتی از الگوریتم یک بعدی برای محاسبه تبدیل معکوس دو بعدی استفاده می شود، پس از پردازش هر سطر یا ستون مزدوج مختلط گرفته نمی شود. در عوض، با تابع (۴.۷) $F^*(u, v)$ طوری رفتار می شود که گویا همان $(y, x) f$ در روای تبدیل دو بعدی مستقیم می باشد که در شکل ۳.۷ خلاصه شده می باشد. مزدوج مختلط حاصل (اگر نیاز باشد)، تبدیل معکوس، $(x, y) f$ ، را بدست می دهد.

۴.۴.۳ پیاده سازی

پیاده سازی رایانه ای الگوریتم FFT که در بخش ۳.۴.۱ بیان شد، کار ساده ای است. نکته اصلی که باید در ذهن داشت این است که داده های ورودی باید در ترتیب مورد نیاز برای کاربرد متوالی معادلات (۳.۷-۴) و (۳.۸-۴) مرتب شوند. روای مرتب سازی را می توان با یک مثال ساده واضح نمود.

فرض کنید که می خواهیم از الگوریتم دو برابر کردن متوالی جهت محاسبه FFT تابع ۸ نقطه ای $\{f(0), f(1), \dots, f(7)\}$ استفاده کنیم. معادله (۳.۷-۴) از نمونه های با آرگومان زوج، $\{f(1), f(3), f(5), f(7)\}$ و معادله (۳.۸-۴) از نمونه های با آرگومان فرد، $\{f(0), f(2), f(4), f(6)\}$ استفاده می کند. در هر حال هر تبدیل ۴ نقطه ای به صورت دو تبدیل دو نقطه ای محاسبه می شود که نیازمند استفاده از معادلات (۳.۷-۴) و (۳.۸-۴) نیز می باشد. بنابراین برای محاسبه FFT اولین مجموعه فوق، باید آن را به دو بخش زوج $\{f(0), f(4), f(2), f(6)\}$ و فرد $\{f(1), f(3), f(5), f(7)\}$ تقسیم نمود. به طور مشابه مجموعه دوم به بخش $\{f(1), f(5), f(3), f(7)\}$ برای معادله (۳.۷-۴) و بخش $\{f(0), f(2), f(4), f(6)\}$ برای معادله (۳.۸-۴) تقسیم بندی می شود. هیچ تغییر ترتیب دیگری لازم نیست، زیرا هر مجموعه 'دو عنصری دارای یک عنصر زوج و یک عنصر فرد در نظر گرفته می شود. برای ترکیب این نتایج باید آرایه ورودی به صورت $\{f(0), f(4), f(2), f(6), f(1), f(5), f(3), f(7)\}$ بیان شود. الگوریتم دو برابر کردن متوالی با روش نشان داده شده در شکل ۳.۲۲ روی این آرایه عمل می کند. در اولین سطح محاسبات، چهار تبدیل دو نقطه ای بر روی $\{f(0), f(4)\}, \{f(2), f(6)\}, \{f(1), f(5)\}$ و $\{f(3), f(7)\}$ انجام می شود. در سطح بعدی با استفاده از این نتایج، دو تبدیل چهار نقطه ای تشکیل می شود و در آخرین سطح با استفاده از این دو نتیجه، تبدیل مطلوب تولید می شود.

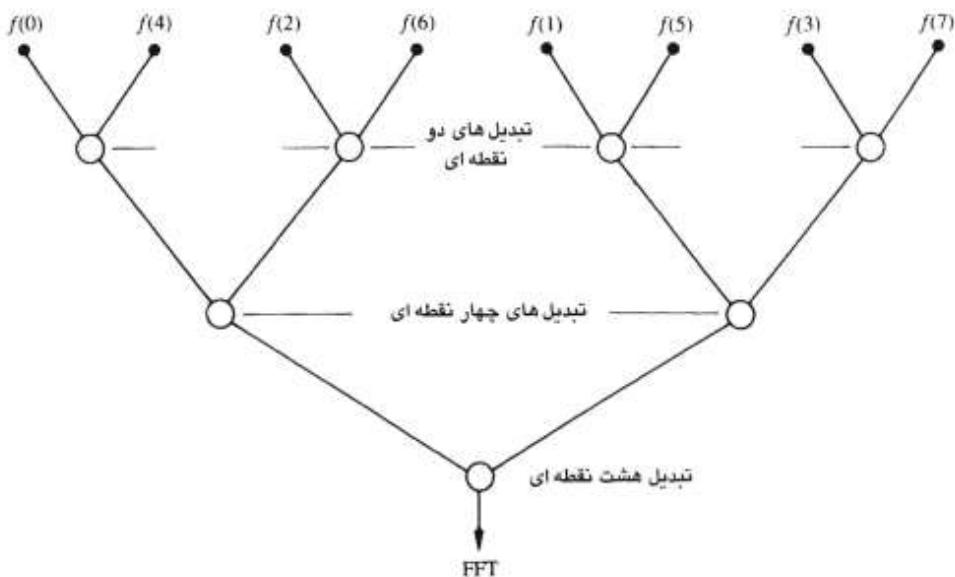
خوبیختانه، روال کلی برای بازمرتب سازی آرایه ورودی از قاعده 'ساده' بیت-برگردان^۱ پیروی می کند. اگر x یک مقدار معتبر برای آرگومان (x) باشد، آرگومان مربوطه در آرایه بازمرتب شده را با نمایش دودویی x و برگردان کردن بیت ها بدست می آوریم. مثلاً اگر $N = 2^3 = 8$ ، عنصر هفتم آرایه اولیه، یعنی $f(6)$ ، عنصر چهارم در آرایه بازمرتب شده می باشد، زیرا $110_2 = 6$ که وقتی بیتها برگردان شوند، $011_2 = 3$ می شود. این عمل برگردان چپ به راست عدد دودویی است و نباید با متمم دودویی اشتباه شود. جدول ۳.۲ این روال را برای $N = 8$ نشان می دهد. اگر در محاسبه FFT از آرایه بازمرتب شده استفاده شود، پاسخ بدست آمده، عناصر تبدیل فوریه در ترتیب صحیح خواهد بود. بر عکس اگر از آرایه در ترتیب اصلی آن استفاده شود، ترتیب پاسخ به صورت بیت-برگردان خواهد بود. این توضیحات در محاسبه تبدیل معکوس نیز صادق است.

شکل ۳.۲۳ یک زیر برنامه فرترن را برای محاسبه FFT با روش دو برابر کردن متوالی نشان می دهد.

۱) bit-reversal rule

۴.۳ تبدیل فوریه سریع

۱۶۵



شکل ۲۰.۳ آرایه ورودی مرتب و کاربرد آن در روش دوبرابر سازی متوالی

عاملهای آرگومان زیر برنامه به صورت زیر هستند. در ورودی، F آرایه ای است که تبدیلش مطلوب است و برابر n است. در خروجی، آرایه F حاوی تبدیل فوریه است. چون F آرایه ای مختلط است، اگر ورودی تابعی حقیقی باشد، باید قبل از فرآخوانی زیر برنامه قسمت موهومی F صفر شود.

اولین قسمت برنامه که شامل حلقه "DO 3" است، داده های ورودی را مرتب می کند. قسمت دوم که شامل حلقه "DO 5" است، محاسبات دو برابر کردن متوالی را انجام می دهد. حلقه "DO 6" که شامل حلقه "DO 6" است، برگردان و باز مرتب کردن آرایه برای دادن آن به الگوریتم FFT می باشد.

جدول ۲۰.۳ مثالی در مورد بیت-برگردان و باز مرتب کردن آرایه برای دادن آن به الگوریتم FFT

آرگومان اولیه			آرایه اولیه	آرگومان بیت برگردان شده			آرایه باز مرتب شده
0	0	0	$f(0)$	0	0	0	$f(0)$
0	0	1	$f(1)$	1	0	0	$f(4)$
0	1	0	$f(2)$	0	1	0	$f(2)$
0	1	1	$f(3)$	1	1	0	$f(6)$
1	0	0	$f(4)$	0	0	1	$f(1)$
1	0	1	$f(5)$	1	0	1	$f(5)$
1	1	0	$f(6)$	0	1	1	$f(3)$
1	1	1	$f(7)$	1	1	1	$f(7)$

```

SUBROUTINE FFT(F,LN)
COMPLEX F(1024),U,W,T,CMPLX
PI=3.141593
N=2**LN
NV2=N/2
NM1=N-1
J=1
DO 3 I=1,NM1
    IF(I.GE.J) GO TO 1
    T=F(J)
    F(J)=F(I)
    F(I)=T
1   K=NV2
2   IF(K.GE.J) GO TO 3
    J=J-K
    K=K/2
    GO TO 2
3   J=J+K
    DO 5 L=1,LN
        LE=2**L
        LE1=LE/2
        U=(1.0,0.0)
        W=CMPLX(COS(PI/LE1),-SIN(PI/LE1))
        DO 5 J=1,LE1
            DO 4 I=J,N,LE
                IP=I+LE1
                T=F(IP)*U
                F(IP)=F(I)-T
4            F(I)=F(I)+T
5            U=U*W
        DO 6 I=1,N
6        F(I)=F(I)/FLOAT(N)
        RETURN
END

```

شکل ۲۳.۳ پیاده سازی فرترن الگوریتم FFT دو برابر کردن متوالی (گرفته شده از Cooley و همکاران [1969]).

نتایج را بر N تقسیم می کند. برای $1024 = N$ ، این برنامه ساده تنها ۱۲ درصد کم بازده تر از برنامه فرترنی است که برای ذخیره مقادیر W_{2^M} از جدول استفاده می کند.

معادلات $(3-3)-(3-6)$ به همراه اشکال ۲۳.۳ و ۲۳.۷، اطلاعات لازم برای پیاده سازی FFT دو بعدی مستقیم را فراهم می کنند. همان طور که در بخش ۳.۴ نشان داده شد، اگر مزدوج مختلط تبدیل فوریه به عنوان ورودی زیر برنامه FFT استفاده شود، همان مفاهیم به تبدیل معکوس اعمال می شوند.

می توان از مبنایهای صحیح بزرگتر از ۲ نیز در روابط FFT استفاده نمود. در واقع نمایش مبنای ۳ نسبت به هر مبنای دیگر به اعمال جزئی کمتری نیاز دارد (Welch[1969] و Cooley, Lewis) اما مشکلات آن از نظر برنامه نویسی، آن را انتخاب نامطلوبی می سازد. از نظر تعداد عملیات مورد نیاز، پیاده سازی در

۵.۳ سایر تبدیلهای جداپذیر تصویر ۱۶۷

مبنای ۴، معادل پیاده سازی در مبنای ۲ است. سایر مبنای های بزرگتر به طور تصاعدی به تعداد عملیات بیشتری نیاز دارند و کم بازده تر هستند. الگوریتم های تبدیل فوریه سریع، معمولاً در مبنای ۲ بیان می شوند زیرا پیاده سازی آن در زبان اسکلی آسان تر است.

۵.۳ سایر تبدیلهای جداپذیر تصویر

تبدیل فوریه گستته یک بعدی از دسته 'تبدیلات مهم' است که می توان آن را بر حسب رابطه 'کلی'

$$T(u) = \sum_{x=0}^{N-1} f(x)g(x, u) \quad (1-5.3)$$

بیان کرد که در آن $T(u)$ تبدیل $f(x)$ و $g(x, u)$ پنجره تبدیل مستقیم و "مقادیری در محدوده $0, 1, \dots, N-1$ دارد. به طور مشابه، رابطه 'تبدیل معکوس' به صورت

$$f(x) = \sum_{u=0}^{N-1} T(u)h(x, u) \quad (2-5.3)$$

می باشد که $h(x, u)$ پنجره تبدیل معکوس است و x مقادیری در محدوده $0, 1, \dots, N-1$ دارد. خواص پنجره تبدیل، طبیعت آن تبدیل را معین می کند.

برای آرایه های دو بعدی مربعی، تبدیلات مستقیم و معکوس عبارتند از

$$T(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)g(x, y, u, v) \quad (3-5.3)$$

و

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} T(u, v)h(x, y, u, v) \quad (4-5.3)$$

که باز هم $h(x, y, u, v)$ و $g(x, y, u, v)$ به ترتیب پنجره های تبدیل مستقیم و معکوس نامیده می شوند. پنجره ها تنها به اندیس های x, y, u و v و نه به مقادیر (y, u) یا (x, v) تبدیل می شوند. بنابراین می توان $h(x, y, u, v)$ را به عنوان توابع پایه بسط سری^۱ با استفاده از معادله (۳-۵.۳) یا (۴-۵.۳) در نظر گرفت. این نکته را در ادامه این بخش با تفصیل بیشتر شرح می دهیم.

پنجره 'مستقیم'، جداپذیر خوانده می شود اگر

1) Series expansion

$$g(x, y, u, v) = g_1(x, u)g_2(y, v). \quad (5-5. 3)$$

به علاوه، پنجره متقارن است اگر g از نظر عملکرد معادل g باشد. در این حالت می توان معادله (5-5. 3) را به صورت

$$g(x, y, u, v) = g_1(x, u)g_2(y, v). \quad (6-5. 3)$$

بیان کرد. اگر در معادلات (5-5. 3) و (6-5. 3) $g(x, y, u, v)$ جایگزین $h(x, y, u, v)$ شود، همان توضیحات در مورد پنجره معکوس نیز صادق است. تبدیل فوریه دو بعدی حالت خاصی از معادله (3-5. 3) است. پنجره آن

$$g(x, y, u, v) = \frac{1}{N} \exp[-j2\pi(ux + vy)/N]$$

می باشد که متقارن و جداپذیر است؛ زیرا

$$\begin{aligned} g(x, y, u, v) &= g_1(x, u)g_2(y, v) \\ &= \frac{1}{\sqrt{N}} \exp[-j2\pi ux/N] \frac{1}{\sqrt{N}} \exp[-j2\pi vy/N]. \end{aligned} \quad (7-5. 3)$$

پنجره فوریه معکوس نیز جداپذیر و متقارن است. یک تبدیل با پنجره جداپذیر را می توان در دو مرحله محاسبه نمود که در هر مرحله یک تبدیل دو بعدی مورد نیاز است. ابتدا با گرفتن تبدیل یک بعدی در طول هر سطر از $f(x, y)$ ، برای

$$x, v = 0, 1, 2, \dots, N-1$$

$$T(x, v) = \sum_{y=0}^{N-1} f(x, y)g_2(y, v) \quad (8-5. 3)$$

سپس با گرفتن تبدیل دو بعدی در طول هر ستون $T(x, v)$ ، عبارت

$$T(u, v) = \sum_{x=0}^{N-1} T(x, v)g_1(x, u) \quad (9-5. 3)$$

را برای $u, v = 0, 1, 2, \dots, N-1$ به عنوان حاصل تبدیل دو بعدی بدست می آوریم. این روال

۵.۳ سایر تبدیلهای جدابزیر تصویر ۱۶۹

موافق روش ارائه شده در بخش ۳.۳.۱ برای تبدیل فوریه است. می‌توان ابتدا با گرفتن تبدیل فوریه در طول هرستون (y , $f(x, y, u)$, $T(y, u)$) را بدست آورد و سپس با گرفتن تبدیل در طول هر سطر ($T(y, u)$, نتیجه نهایی ($T(u, v)$) را بدست آورد. اگر $h(x, y, u, v)$ جدابزیر باشد، توضیحات مشابهی برای تبدیل معکوس صدق می‌کند.

اگر پنجه ($g(x, y, u, v)$) متقارن و جدابزیر باشد، می‌توان معادله (۳.۵) را به صورت ماتریسی

$$\mathbf{T} = \mathbf{A}\mathbf{F}\mathbf{A} \quad (10-5.3)$$

نیز بیان نمود که \mathbf{F} ماتریس تصویر $N \times N$, \mathbf{A} ماتریس تبدیل متقارن $N \times N$ با عناصر (i, j) و \mathbf{T} تبدیل $N \times N$ حاصل شده برای مقادیر u و v در محدوده $0, 1, 2, \dots, N - 1$ می‌باشد.

برای بدست آوردن تبدیل معکوس، معادله (۱۰-۵.۳) را از چپ و راست در ماتریس تبدیل معکوس ضرب می‌کنیم:

$$\mathbf{B}\mathbf{T}\mathbf{B} = \mathbf{B}\mathbf{A}\mathbf{F}\mathbf{A}\mathbf{B}. \quad (11-5.3)$$

$$\text{اگر } \mathbf{B} = \mathbf{A}^{-1}$$

$$\mathbf{F} = \mathbf{B}\mathbf{T}\mathbf{B} \quad (12-5.3)$$

که نشان می‌دهد می‌توان تصویر رقمی \mathbf{F} را به طور کامل از تبدیلش بازسازی نمود. اگر \mathbf{B} برابر \mathbf{A}^{-1} نباشد، استفاده از معادله (۱۱-۵.۳) تقریبی از \mathbf{F} را می‌دهد:

$$\hat{\mathbf{F}} = \mathbf{B}\mathbf{A}\mathbf{F}\mathbf{A}\mathbf{B}. \quad (13-5.3)$$

چند تبدیل-از جمله تبدیلهای فوریه، والش، هادامارد، کسینوسی گستته، ها آر و اسلت-را می‌توان به شکل معادلات (۳.۱۰-۵) و (۳.۱۲-۵) بیان کرد. یک خاصیت مهم از ماتریسهای تبدیل فوق، این است که می‌توان آنها را به ضرب ماتریس هایی که عناصر غیر صفرشان کمتر از ماتریس اولیه است، تجزیه نمود. این نتیجه که در ابتداء توسط [Good 1958] برای تبدیل فوریه بیان شد، افزونگی و در نتیجه تعداد اعمال مورد نیاز در پیاده سازی تبدیل دو بعدی را کاهش می‌دهد. میزان کاهش برابر کاهشی

است که توسط الگوریتم FFT به آن دست یافته باشد. گرچه ما بر روش های محاسباتی که با تکرار متوالی الگوریتم های یک بعدی تصویر $N \times N$ می باشد. تأکید می کنیم، اما از طریق بیان ماتریسی مسئله نیز تبدیل مستقیم و معکوس تصویر را بدست می آورند، تأکید می کنیم، اما از طریق بیان ماتریسی مسئله نیز می توان به همان ترتیب محاسباتی دست یافت. برای جزئیات بیشتر درباره این موضوع به Andrews[1970] مراجعه کنید.

١ . ٥ . ٣ تبدل والش

و قته $N = 2^n$ ، تبدیل والش گسته تابع $f(x)$ ، که با (ii) W نشان داده می شود، با جایگذاری پنجره

$$g(x, u) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)} \quad (14-5, r)$$

در معادله $(3-1-5)$ پدست می‌آید. به بیان دیگر،

$$W(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \prod_{i=0}^{n-1} (-1)^{h_i(x)b_{n-1-i}(u)} \quad (15-5.3)$$

که $b_k(z)$ بیت k ام نمایش دودویی z است. مثلاً اگر $3 = 110_2$ و $n = 6$ باشد، آن‌گاه

مقادیر $(u, g(x, u))$ ، با کنارگذاشتن عبارت ثابت $N/1$ ، برای $N = 8$ در جدول ۳. فهرست شده اند.
 ماتریس تشکیل شده با پنجره 'تبديل والش'، ماتریس مقاینه است که سطرها و ستون های آن متعامدند.
 این خواص، که در حالت کلی برقرار هستند، پنجره 'معکوسی' را نتیجه می دهند که جز در مورد ضریب ثابت $N/1$ ، با پنجره مستقیم برابر است؛ یعنی،

$$h(x, u) = \prod_{i=0}^{n-1} (-1)^{h_i(x)h_{n-1-i}(u)}, \quad (19-5, 3)$$

بنابر این تبدیل والش معکوس عبارتست از

$$f(x) = \sum_{u=0}^{N-1} W(u) \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n+i}(u)}. \quad (17-5.3)$$

ر خلاف تدبیا، فوریه که بر اساس عبارات مثلثاتی است، تبدیل والش از بسط سری توابع پایه‌ای که

۱۷۱ ۵.۳ سایر تبدیلهای جداپذیر تصویر

جدول ۳.۳ مقادیر پنجره تبدیل والش یک بعدی برای $N = 8$

$u \setminus x$	0	1	2	3	4	5	6	7
0	+	+	+	+	+	+	+	+
1	+	+	+	+	-	-	-	-
2	+	+	-	-	+	+	-	-
3	+	+	-	-	-	-	+	+
4	+	-	+	-	+	-	+	-
5	+	-	+	-	+	-	-	+
6	+	-	-	+	+	-	-	-
7	+	-	-	+	-	+	+	-

مقادیر شان $+1$ یا -1 است، تشکیل می شود.

درستی معادله (۳.۱۷-۵) به آسانی با جایگذاری معادله (۳.۱۵-۵) برای $W(u)$ و استفاده از شرط تعامل که قبلاً ذکر شد، اثبات می شود. در معادلات (۳.۱۵-۵) و (۳.۱۷-۵) توجه کنید که تبدیلهای والش مستقیم و معکوس تنها در عبارت $N/1$ با هم تفاوت دارند. بنابراین هر الگوریتم محاسبه تبدیل مستقیم را می توان بسادگی با ضرب حاصل الگوریتم در N ، برای بدست آوردن تبدیل معکوس نیز بکار برد.

پنجره های والش مستقیم و معکوس دو بعدی با روابط

$$g(x, y, u, v) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{[b_i(x)b_{n-i}(u) + b_i(y)b_{n-i}(v)]} \quad (18-5.3)$$

$$h(x, y, u, v) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{[b_i(x)b_{n-i}(u) + b_i(y)b_{n-i}(v)]}. \quad (19-5.3)$$

ارائه می شوند. گرچه گذاشتن هر دو عبارت $N/1$ در جلوی یکی از $g(x, y, u, v)$ یا $h(x, y, u, v)$ نیز صحیح است، در کاربردهای پردازش تصویر که بستگی برابری به تبدیلهای مستقیم و معکوس دارد، معادلات (۳.۱۸-۵) و (۳.۱۹-۵) ترجیح داده می شوند. چون از معادلات فوق پنجره های یکسانی حاصل می شود، معادلات (۳.۳-۵) و (۳.۴-۵) نیز به تبدیلهای والش مستقیم و معکوسی که شکل برابری دارند، منجر می شوند؛ یعنی،

۱۷۲ پردازش تصویر رقمه

$$W(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \prod_{i=0}^{n-1} (-1)^{[b_i(x)b_{n-1-i}(u) + b_i(y)b_{n-1-i}(v)]} \quad (20-5. 3)$$

و

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} W(u, v) \prod_{i=0}^{n-1} (-1)^{[b_i(x)b_{n-1-i}(u) + b_i(y)b_{n-1-i}(v)]}. \quad (21-5. 3)$$

بنابراین هر الگوریتمی که در محاسبه تبدیل والش مستقیم استفاده شود، را می توان بدون تغییر برای محاسبه تبدیل معکوس آن نیز بکار برد.

پنجره های تبدیل والش جدا پذیر و متقاض هستند، زیرا

$$\begin{aligned} g(x, y, u, v) &= g_i(x, u)g_i(y, v) \\ &= h_i(x, u)h_i(y, v) \\ &= \left[\frac{1}{\sqrt{N}} \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)} \right] \left[\frac{1}{\sqrt{N}} \prod_{i=0}^{n-1} (-1)^{b_i(y)b_{n-1-i}(v)} \right]. \end{aligned} \quad (22-5. 3)$$

بنابراین با کاربرد متوالی تبدیل والش یک بعدی که با معادله (15-5. 3) بیان شد، می توان $W(u, v)$ و تبدیل معکوس آن را محاسبه کرد. روایی که در محاسبات بکار گرفته می شود، همانند روش ارائه شده در بخش ۳. ۳. ۱ و شکل ۳. ۷ برای تبدیل فوریه است.

تبدیل والش را می توان با الگوریتم سریعی که تقریباً هم شکل روش دوبرابر سازی متوالی بیان شده در بخش ۴. ۱ در مورد FFT است، محاسبه نمود. تنها تفاوت این است که در تبدیل والش سریع (FWT) تمام اجزای نمایی W برابر ۱ قرار داده می شوند.[†] آن گاه معادلات (۳. ۹-۴) و (۳. ۱۰-۴)، که روابط اساسی FFT هستند، به صورت زیر تبدیل می شوند

$$W(u) = \frac{1}{2} [W_{\text{even}}(u) + W_{\text{odd}}(u)] \quad (23-5. 3)$$

و

$$W(u + M) = \frac{1}{2} [W_{\text{even}}(u) - W_{\text{odd}}(u)] \quad (24-5. 3)$$

[†] کاربرد W در این بخش که برای نشان دادن تبدیل والش به کار می رود، نباید با کاربرد آن در بخش ۳. ۴. ۱ که برای نشان دادن اجزای نمایی به کار رفته، اشتباه گرفته شود.

۱۷۳ ۵.۳ سایر تبدیلهای جدالپذیر تصویر

که در معادلات فوق $N/2 = M = N$ و $u = 0, 1, \dots, M-1$ ، $M = N$ نیز برای نشان دادن تبدیل والش یک بعدی به کار می رود. به جای اثبات کلی این نتیجه، می توان کاربرد معادله (۳.۵-۱۵) و درستی معادلات (۳.۵-۲۳) و (۳.۵-۲۴) را با یک مثال نشان داد. برای تفصیل بیشتر در این موضوع، به [Shanks 1969] مراجعه کنید.

مثال: اگر $N = 4$ ، کاربرد معادله (۳.۵-۱۵) دنباله مراحل زیر را نتیجه می دهد:

$$\begin{aligned} W(0) &= \frac{1}{4} \sum_{x=0}^3 \left[f(x) \prod_{i=0}^1 (-1)^{b_i(x)b_{1-i}(0)} \right] \\ &= \frac{1}{4} [f(0) + f(1) + f(2) + f(3)] \\ W(1) &= \frac{1}{4} \sum_{x=0}^3 \left[f(x) \prod_{i=0}^1 (-1)^{b_i(x)b_{1-i}(1)} \right] \\ &= \frac{1}{4} [f(0) + f(1) - f(2) - f(3)] \\ W(2) &= \frac{1}{4} \sum_{x=0}^3 \left[f(x) \prod_{i=0}^1 (-1)^{b_i(x)b_{1-i}(2)} \right] \\ &= \frac{1}{4} [f(0) - f(1) + f(2) - f(3)] \\ W(3) &= \frac{1}{4} \sum_{x=0}^3 \left[f(x) \prod_{i=0}^1 (-1)^{b_i(x)b_{1-i}(3)} \right] \\ &= \frac{1}{4} [f(0) - f(1) - f(2) + f(3)]. \end{aligned}$$

تقسیم بندی این نتایج به دو گروه، درستی معادلات (۳.۵-۲۳) و (۳.۵-۲۴) را نشان می دهد:

$$W_{\text{even}}(0) = \frac{1}{2} [f(0) + f(2)] \quad , \quad W_{\text{odd}}(0) = \frac{1}{2} [f(1) + f(3)]$$

$$W_{\text{even}}(1) = \frac{1}{2} [f(0) - f(2)] \quad , \quad W_{\text{odd}}(1) = \frac{1}{2} [f(1) - f(3)].$$

از معادله (۳.۵-۲۳) داریم،

$$W(0) = \frac{1}{2} [W_{\text{even}}(0) + W_{\text{odd}}(0)]$$

$$= \frac{1}{4} [f(0) + f(1) + f(2) + f(3)]$$

$$\begin{aligned} W(1) &= \frac{1}{2} [W_{\text{even}}(1) + W_{\text{odd}}(1)] \\ &= \frac{1}{4} [f(0) + f(1) - f(2) - f(3)]. \end{aligned}$$

با محاسبه دو جزء بعدی این نتایج با استفاده از معادله (۳.۵-۲۴) داریم

$$\begin{aligned} W(2) &= \frac{1}{2} [W_{\text{even}}(0) - W_{\text{odd}}(0)] \\ &= \frac{1}{4} [f(0) - f(1) + f(2) - f(3)] \end{aligned}$$

$$\begin{aligned} W(3) &= \frac{1}{2} [W_{\text{even}}(1) - W_{\text{odd}}(1)] \\ &= \frac{1}{4} [f(0) - f(1) - f(2) + f(3)]. \end{aligned}$$

بنابراین محاسبه $W(u)$ با معادله (۳.۱۵) یا با معادلات (۳.۵-۲۳) و (۳.۵-۲۴) نتایج یکسانی خواهد داد.

همان طور که قبلاً ذکر شد، بسادگی می‌توان با برابر "۱" قرار دادن تمام عبارات مثلثاتی در الگوریتم محاسبه FFT با روش دوبرابر سازی متوالی، آن را به الگوریتمی برای محاسبه تبدیل والش سریع تبدیل کرد. شکل ۳.۲۴. تغییرات مورد نیاز برنامه FFT موجود در شکل ۳.۲۳ را نشان می‌دهد. تبدیل والش حقیقی است، بنابراین برای یک مسئله یکسان نسبت به تبدیل فوریه که در حالت کلی مختلط است، به حافظه رایانه‌ای کمتری نیاز دارد.

همان طور که در مورد معادلات (۳.۵-۳) و (۳.۵-۴) بیان شد، هر تبدیل و معکوسش را می‌توان بر حسب بسط سری با استفاده از پنجره‌های مناسب بیان کرد. پنجره‌ها تنها به اندیشهای x, v, u و y -نه به مقادیر تصویر یا تبدیلش-بستگی دارند، بنابراین پنجره‌های مذکور به عنوان مجموعه‌ای از توابع پایه (پکار می‌روند که تا وقتی ابعاد تصویر ثابت باشد، تغییر نمی‌کنند. مثلاً شکل ۳.۲۵. توابع پایه (پنجره‌ها) را به صورت تابعی از v و u (با کنارگذاشتن عبارت $1/N$) برای محاسبه تبدیل والش وقتی $N = 4$ را به صورت تابعی از v و u (با کنارگذاشتن عبارت $1/N$) برای محاسبه تبدیل والش وقتی $N = 16$ نشان می‌دهد. در هر بلوک، x و y از ۰ تا ۳ تغییر می‌کنند (یعنی از ۰ تا $N-1$)، در حالی که v و u

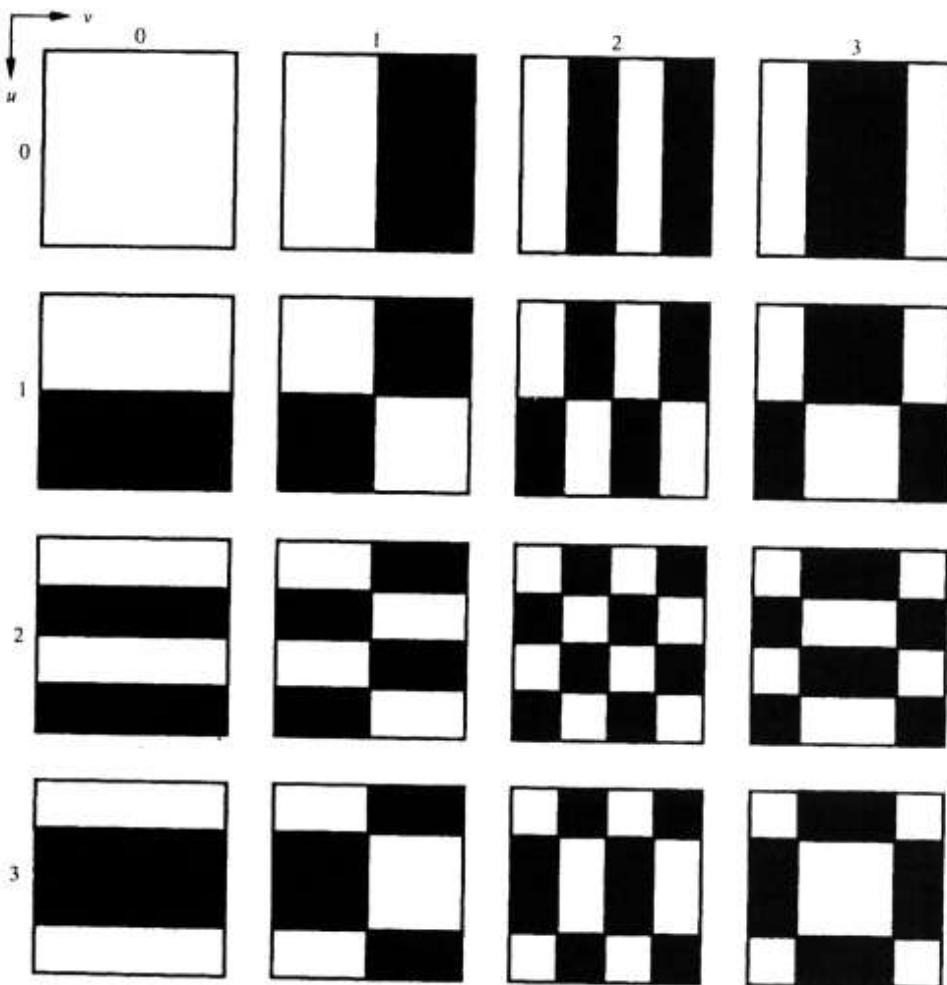
```

SUBROUTINE FWT(F,LN)
REAL F(1024),T
N=2**LN
NV2=N/2
NM1=N-1
J=1
DO 3 I=1,NM1
  IF(I.GE.J) GO TO 1
  T=F(J)
  F(J)=F(I)
  F(I)=T
  K=NV2
  IF(K.GE.J) GO TO 3
  J=J-K
  K=K/2
  GO TO 2
3   J=J+K
  DO 5 L=1,LN
    LE=2**L
    LE1=LE/2
    DO 5 J=1,LE1
      DO 4 I=J,N,LE
        IP=I+LE1
        T=F(IP)
        F(IP)=F(I)-T
        F(I)=F(I)+T
4   CONTINUE
  DO 6 I=1,N
6   F(I)=F(I)/FLOAT(N)
  RETURN
END

```

شکل ۳ ۲۴۰ تغییر الگوریتم FFT دو برابر کردن متواالی برای محاسبه تبدیل والش سریع

مریبوط به آن بلوک، ثابت هستند. بنابراین هر بلوک شامل آرایه ای 4×4 از عناصر دودویی می باشد. در بلوک متاظر با $v = v = 0$ ، تمام مقادیر پنجره "۱" هستند (که بارنگ سفید نشان داده می شوند). وقتی $v = 2$ ، حداقل تغییرات رخ می دهد که "۱" ها و "۱"-های متناوب ("۱"-بارنگ سیاه نشان داده می شود) به بلوک 4×4 الگوی شطرنجی می دهند. برای استفاده از توابع پایه شکل ۳.۲۵ جهت محاسبه تبدیل والش یک تصویر 4×4 ، مثلاً محاسبه $W(0,0)$ ، آرایه تصویری را به صورت نقطه به نقطه در بلوک پایه 4×4 متاظر با $v = 0$ ضرب کرده و پس از جمع نتایج با تقسیم حاصل بر 4 ، $W(0,0)$ بدست می آید. برای بدست آوردن $W(0,1)$ باید از بلوک متاظر با $v = 1$ استفاده کرد و به همین ترتیب برای تمام 16 بلوک عمل می شود. در تبدیل والش چون پنجره' معکوس با پنجره مستقیم یکسان است، توابع پایه شکل ۳.۲۵ در تبدیل معکوس هم به کارمی روند، با این تفاوت که در این حالت در هر بلوک x



شکل ۲۵.۳ ۲۵ توابع پایه والش برای $N = 4$. هر بلوک دارای 4×4 عنصر متناظر با تغییر زوج متغیرهای x و y از ۰ تا ۳، می باشد. مبدأ هر بلوک در سمت چپ بالای آن است. سفید و سیاه به ترتیب نمایانگر +1 و -1 هستند.

و y ثابت هستند و ii و jj در محدوده ۰ تا $N-1$ تغییر می کنند.

۲۰.۵.۳ تبدیل هادامارد

یکی از چند رابطه شناخته شده برای پنجره هادامارد مستقیم یک بعدی رابطه

$$g(x, u) = \frac{1}{N} (-1)^{\sum_{i=0}^{n-1} h_i(x)h_i(u)} \quad (20.5.3)$$

۵.۳ سایر تبدیلهای جدایلیر تصویر ۱۷۷

می باشد که عمل جمع در نمای آن در مبنای دو انجام می شود و مانند معادله (۳.۵-۱) $b_i(z), b_i(y), b_i(x)$ بیت kام در نمایش دودویی z است. با جایگذاری معادله (۳.۵-۲) در معادله (۳.۵-۱) عبارت زیر برای تبدیل هادامارد یک بعدی بدست می آید

$$H(u) = \frac{1}{N} \sum_{i=0}^{N-1} f(x)(-1)^{\sum_{j=0}^{n-1} b_j(x)b_j(u)} \quad (26-5.3)$$

که در آن $N = 2^n$ و u مقادیری در محدوده $0, 1, 2, \dots, N-1$ دارد. مانند تبدیل والش، پنجره هادامارد تیز ماتریسی با سطرها و ستون های متعامد است. این حالت پنجره معکوسی را نتیجه می دهد که بجز در مورد جزء $1/N$ ، برابر با پنجره هادامارد مستقیم است؛ یعنی،

$$h(x, u) = (-1)^{\sum_{j=0}^{n-1} b_j(x)b_j(u)} \quad (27-5.3)$$

با جایگذاری این پنجره در معادله (۳.۵-۲) عبارت زیر برای تبدیل هادامارد معکوس بازی $x = 0, 1, 2, \dots, N-1$ نتیجه می شود:

$$f(x) = \sum_{u=0}^{N-1} H(u)(-1)^{\sum_{j=0}^{n-1} b_j(x)b_j(u)} \quad (28-5.3)$$

به طور مشابه پنجره های دو بعدی با روابط

$$g(x, y, u, v) = \frac{1}{N} (-1)^{\sum_{j=0}^{n-1} [b_j(x)b_j(u) + b_j(y)b_j(v)]} \quad (29-5.3)$$

و

$$h(x, y, u, v) = \frac{1}{N} (-1)^{\sum_{j=0}^{n-1} [b_j(x)b_j(u) + b_j(y)b_j(v)]} \quad (30-5.3)$$

تعیین می شوند که باز هم عمل جمع موجود در نما در مبنای دو انجام می شود. مانند تبدیل والش، پنجره های هادامارد دو بعدی معکوس و مستقیم تیز یکسان هستند.

با جایگذاری معادلات (۳.۵-۲۹) و (۳.۵-۳۰) در معادلات (۳.۵-۳) و (۳.۵-۴) زوج تبدیل

هادامارد دو بعدی زیر بدست می آید:

$$H(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)(-1)^{\sum_{j=0}^{n-1} [b_j(x)b_j(u) + b_j(y)b_j(v)]} \quad (31-5.3)$$

۵

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} H(u, v) (-1)^{\sum_{i=0}^{n-1} [b_i(x)b_i(u) + b_i(y)b_i(v)]} \quad (32-5, 3)$$

چون تبدیل های مستقیم و معکوس یکسان هستند، می توان الگوریتم مورد استفاده در محاسبه $H(u, v)$ را بدون تغییر برای بدست آوردن $f(x, y)$ نیز استفاده کرد، و بر عکس. با توجه به جمع مبنای دو، می توان نشان داد که پنجره های هادامارد، جدایذیر و متقارن هستند. بنابراین

(33-5, 3)

$$\begin{aligned} g(x, y, u, v) &= g_1(x, u)g_1(y, v) \\ &= h_1(x, u)h_1(y, v) \\ &= \left[\frac{1}{\sqrt{N}} (-1)^{\sum_{i=0}^{n-1} b_i(x)b_i(u)} \right] \left[\frac{1}{\sqrt{N}} (-1)^{\sum_{i=0}^{n-1} b_i(y)b_i(v)} \right]. \end{aligned}$$

به استثنای عبارت $\frac{1}{\sqrt{N}}$ ، g_1 و h_1 با معادله (25-5, 3) یکسان هستند. علاوه بر آن، چون پنجره های هادامارد دو بعدی جدایذیر هستند، زوج تبدیل دو بعدی را می توان با کاربرد متوالی الگوریتم تبدیل هادامارد یک بعدی بدست آورد.

جدول ۳. ۴ ماتریس مقادیر تولید شده با پنجره هادامارد یک بعدی معادله (3. 25-5) را برای $N=8$ نشان می دهد که جهت سهولت، جزء ثابت $N/1$ حذف شده است. با این که عناصر، همان عناصر تبدیل والش هستند، ترتیب سطرها و ستون ها متفاوت است. در واقع وقتی $N=2^m$ ، مورد فوق تنها تفاوت بین دو تبدیل والش و هادامارد است. وقتی N بر ابر توان صحیحی از ۲ نباشد، تفاوت مهمتر است. رابطه تبدیل والش را می توان برای هر مقدار صحیح مثبت N نوشت، اما وجود تبدیل هادامارد برای مقادیر N غیر از توان های صحیح دو، تنها تا $200=N$ نشان داده شده است.

از آنجا که اکثر تبدیلات تصویر رقمه بر اساس $= N$ نمونه در هر سطر یا ستون تصویر می باشند، در متون پردازش تصویر کاربرد (ونام) تبدیلهای والش و هادامارد با هم مخلوط شده اند. در اغلب موارد عبارت تبدیل والش-هادامارد برای اشاره به هر یک از آن دو تبدیل استفاده می شود.

دو ویژگی مهم که ممکن است در انتخاب یکی از این دو تبدیل بر دیگری مؤثر باشند، قابل ذکر هستند؛ همان طور که در بخش ۳. ۵. ۲ اشاره شد، یک مزیت معادله (۳. 15-۵) این است که می توان آن را مستقیماً در قالب دو برابر سازی متوالی بیان نمود. این خاصیت محاسبه FWT با تغییر الگوریتم FFT را،

۱۷۹ ۵.۳ سایر تبدیلهای جدابزیر تصویر

جدول ۴.۳ مقادیر پنجره تبدیل هادامارد یک بعدی برای $N = 8$

$u \backslash x$	0	1	2	3	4	5	6	7
0	+	+	+	+	+	+	+	+
1	+	-	+	-	+	-	+	-
2	+	+	-	-	+	+	-	-
3	+	-	-	+	+	-	-	+
4	+	+	+	+	-	-	-	-
5	+	-	+	-	-	+	-	+
6	+	+	-	-	-	-	+	+
7	+	-	-	+	-	+	+	-

طبق روش سرراست بیان شده در بخش ۳.۴.۱، ممکن می‌سازد. در محاسبه تبدیل هادامارد سریع (FHT) برای در نظر گرفتن تفاوت موجود در ترتیب، تغییرات بیشتر این الگوریتم مورد نیاز خواهد بود. روش دیگر، استفاده از الگوریتم FWT شکل ۲۴.۳ و سپس باز مرتب کردن نتایج است تا تبدیل هادامارد بدست آید.

گرچه ترتیب هادامارد از نظر دو برابر کردن متوالی معایبی دارد، اما برای تولید ماتریس‌های تبدیل موردنیاز در پیاده‌سازی معادلات (۳.۵-۵) و (۳.۱۰-۵) به رابطه‌ای بازگشتی منجر می‌شود. ماتریس هادامارد با کمترین مرتبه ($N = 2$)،

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (34-5.3)$$

می‌باشد. آنگاه با فرض این که \mathbf{H}_N بیانگر ماتریس مرتبه N باشد، رابطه بازگشتی به صورت

$$\mathbf{H}_{2N} = \begin{bmatrix} \mathbf{H}_N & \mathbf{H}_N \\ \mathbf{H}_N & -\mathbf{H}_N \end{bmatrix} \quad (35-5.3)$$

وجود دارد که \mathbf{H}_{2N} ماتریس هادامارد از مرتبه $2N$ است و $2^N = N$ فرض می‌شود. ماتریس تبدیل برای استفاده در معادله (۳.۱۰-۵) با تراز سازی¹ ماتریس هادامارد مربوطه بوسیله ریشه دوم مرتبه ماتریس بدست می‌آید. بنابراین در حالت $N \times N$ ، این دو ماتریس رابطه زیر را با هم دارند:

$$\mathbf{A} = \frac{1}{\sqrt{N}} \mathbf{H}_N \quad (36-5.3)$$

1) Normalizing

عبارات برای ماتریس هadamard معکوس با معادلات (۳.۵-۳) و (۳-۵.۴) یکسان هستند.

مثال: با استفاده از معادلات (۳.۵-۳) و (۳-۵.۴) ماتریس های hadamard از مرتبه ۴ و ۸ به صورت

زیربدهست می آیند:

$$\mathbf{H}_4 = \begin{bmatrix} \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & -\mathbf{H}_2 \end{bmatrix}$$

$$= \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{bmatrix}$$

$$\mathbf{H}_8 = \begin{bmatrix} \mathbf{H}_4 & \mathbf{H}_4 \\ \mathbf{H}_4 & -\mathbf{H}_4 \end{bmatrix}$$

$$= \begin{bmatrix} + & + & + & + & + & + & + & + \\ + & - & + & - & + & - & + & - \\ + & + & - & - & + & + & - & - \\ + & - & - & + & + & - & - & + \\ + & + & + & + & - & - & - & - \\ + & - & + & - & - & + & - & + \\ + & + & - & - & - & - & + & + \\ + & - & - & + & - & + & + & - \end{bmatrix}$$

که $+/-$ به ترتیب نشانه $+1$ و -1 هستند.

همان طور که معادلات (۳.۵-۲۵) و (۳-۵.۳) نشان می دهند، (x, u) و $g(x, u)$ تنها در یک ضریب ثابت با هم تفاوت دارند. چون $(i, j) = g_{ij}$ ، درایه های ماتریس \mathbf{A} شکلی مانند بسط $g(x, u)$ دارند. در این مثال صحت مطلب فوق با مقایسه جدول ۳.۴ و عبارت

$$\mathbf{A} = \frac{1}{\sqrt{8}} \mathbf{H}_8.$$



بررسی می شود.

۵.۳ سایر تبدیلهای جدا پذیر تصویر ۱۸۱

جدول ۵.۳ مقادیر پنجره تبدیل هادامارد یک بعدی مرتب شده برای $N = 8$

$x \backslash u$	0	1	2	3	4	5	6	7
0	+	+	+	+	+	+	+	+
1	+	+	+	+	-	-	-	-
2	+	+	-	-	-	-	+	+
3	+	+	-	-	+	+	-	-
4	+	-	-	+	+	-	-	+
5	+	-	-	+	-	+	+	-
6	+	-	+	-	-	+	-	+
7	+	-	+	-	+	-	+	-

تعداد تغییر علامت در طول هر ستون ماتریس هادامارد اغلب فرکانس مکانی^۱ آن ستون نامیده می شود. چون عناصر این ماتریس از مقادیر پنجره بدست می آیند، مفهوم فرکانس مکانی در مورد بسط (x, u) بازای $g(x, u) = (-1)^{b_i(x)p_i(u)}$ به کار می رود. برای نمونه، فرکانس مکانی ۸ ستون و جدول ۳.۴ به ترتیب $0, 1, 2, 3, 4, 5, 6, 7$ هستند.

بیان پنجره های هادامارد طوری که فرکانس مکانی به صورت تابعی صعودی از u باشد، مشابه تبدیل فوریه است که در آن هم فرکانس به صورت تابعی صعودی از x افزایش می یابد. پنجره هادامارد یک بعدی که این ترتیب ویژه را داشته باشد، با رابطه^۲

$$g(x, u) = \frac{1}{N} \sum_{i=0}^{n-1} b_i(x)p_i(u) \quad (37-5.3)$$

بدست می آید که در آن

$$\begin{aligned} p_0(u) &= b_{n-1}(u) \\ p_1(u) &= b_{n-1}(u) + b_{n-2}(u) \\ p_2(u) &= b_{n-2}(u) + b_{n-3}(u) \\ &\vdots \\ p_{n-1}(u) &= b_0(u) + b_1(u). \end{aligned} \quad (38-5.3)$$

1) *Sequency*

^۱ مانند تبدیل فوریه که نا متغیر فرکانس است، به طور عادی مفهوم فرکانس مکانی به این متغیر محدود می شود. بنابراین مرتب نمودن فرکانس مکانی با ستونهای ماتریس هادامارد مبتنی بر این فرض است که ستونهایا به صورت تابعی از x و سطرها به صورت تابعی از u تغییر می کنند. در جدول ۳.۴ از این قرارداد استفاده می شود.

مانند گذشته، اعمال جمع در معادلات (۳.۵-۳۷) و (۳.۵-۳۸) در حساب مدول دو^۱ انجام می شوند.
در جدول ۳.۵ بسط معادله (۳.۵-۳۷) بازای $N=8$ نشان داده می شود. در این جدول برای سهولت، ضریب ثابت، حذف شده است و عناصر + و - به ترتیب نمایانگر +۱ و -۱ هستند. ستون ها و نیز سطرها، به علت تقارن، به ترتیب افزایش فرکانس مکانی هستند.
پنجره هادامارد معکوس مرتب شده،

$$h(x, u) = (-1)^{\sum_{i=0}^{n-1} b_i(x)p_i(u)} \quad (39-5.3)$$

می باشد که $p_i(u)$ با استفاده از معادله (۳.۵-۳۸) محاسبه می گردد. با جایگذاری پنجره های مستقیم و معکوس در معادلات (۳.۵-۱) و (۳.۵-۲) زوج تبدیل هادامارد مرتب شده زیر نتیجه می شود:

$$H(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x)(-1)^{\sum_{i=0}^{n-1} b_i(x)p_i(u)} \quad (40-5.3)$$

و

$$f(x) = \sum_{u=0}^{N-1} H(u)(-1)^{\sum_{i=0}^{n-1} b_i(x)p_i(u)} \quad (41-5.3)$$

مانند حالت مرتب نشده، پنجره های دوبعدی جدابذیر و یکسان هستند:

$$\begin{aligned} g(x, y, u, v) &= h(x, y, u, v) \\ &= \frac{1}{N} (-1)^{\sum_{i=0}^{n-1} [b_i(x)p_i(u) + b_i(y)p_i(v)]} \end{aligned} \quad (42-5.3)$$

جایگذاری این پنجره ها در معادلات (۳.۵-۳) و (۳.۵-۴) زوج تبدیل هادامارد مرتب دوبعدی زیر را نتیجه می دهد:

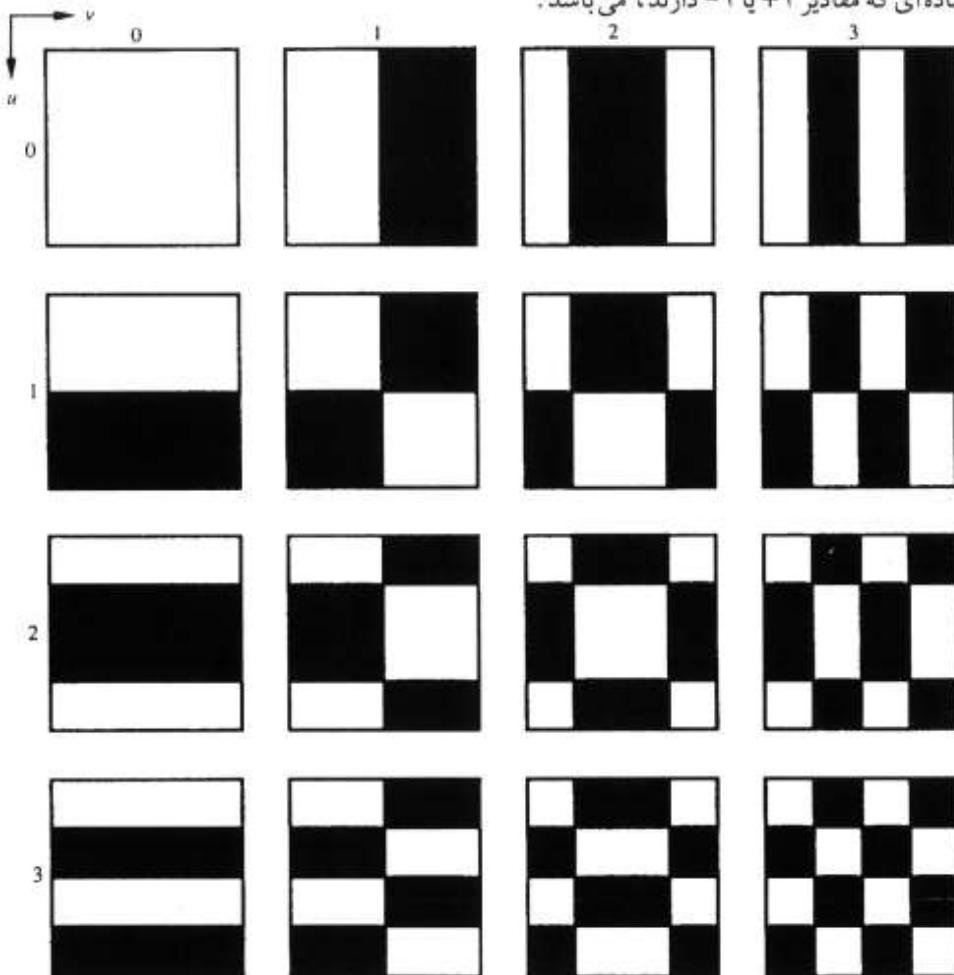
$$H(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)(-1)^{\sum_{i=0}^{n-1} [b_i(x)p_i(u) + b_i(y)p_i(v)]} \quad (43-5.3)$$

و

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} H(u, v)(-1)^{\sum_{i=0}^{n-1} [b_i(x)p_i(u) + b_i(y)p_i(v)]} \quad (44-5.3)$$

۵.۳ سایر تبدیلهای جدایی بر تصویر ۱۸۳

شکل ۳.۲۶ توابع (پنجره های) پایه هادامارد مرتب دو بعدی را برای $N=4$ نشان می دهد. توابع پایه موجود در اشکال ۳.۲۶، ۳.۲۵، ۳.۲۶ تنها از این نظر متفاوت هستند که توابع شکل ۳.۲۶ به ترتیب افزایش فرکانس مکانی مرتب شده اند و بنا بر این تعبیر طبیعی تری دارند. شکل ۳.۲۷، ۳.۲۸ یک تصویر ساده و لگاریتم دامنه تبدیل هادامارد مرتب آن را نشان می دهد. برخلاف طیف فوریه که در آن مفهوم فرکانس نقش مهمی دارد، شکل ۳.۲۷، ۳.۲۸ (ب) چنان تعبیر فیزیکی مفیدی ندارد. در هر حال فرکانس مکانی به صورت تابعی از ۰ و ۱ افزایش می یابد. گذشته از این، شکل ۳.۲۷، ۳.۲۸ (ب) تجزیه تابع تصویر اولیه و براساس توابع پایه ساده ای که مقادیر +1 یا -1 دارند، می باشد.



شکل ۳.۲۶ توابع پایه هادامارد مرتب برای $N=4$. هر بلوک شامل 4×4 عنصر متناظر با x و y که از $+1$ و -1 تغییر می کنند، می باشد. مبدأ هر بلوک در سمت چپ بالای آن است. سفید و سیاه به ترتیب نشانه +1 و -1 هستند.

۳.۵.۳ تبدیل کسینوسی گستته^۱

تبدیل کسینوسی گستته یک بعدی (DCT) برای $1 \leq u \leq N-1$ به صورت

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \quad (45-5.3)$$

تعریف می شود. به طور مشابه DCT معکوس برای $1 \leq x \leq N-1$ به صورت

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \quad (46-5.3)$$

تعریف می شود. در هر دو معادله (45-5.3) و (46-5.3)، α عبارت است از:

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & u = 0 \\ \sqrt{\frac{2}{N}} & u = 1, 2, \dots, N-1 \end{cases} \quad \text{برای } (47-5.3)$$

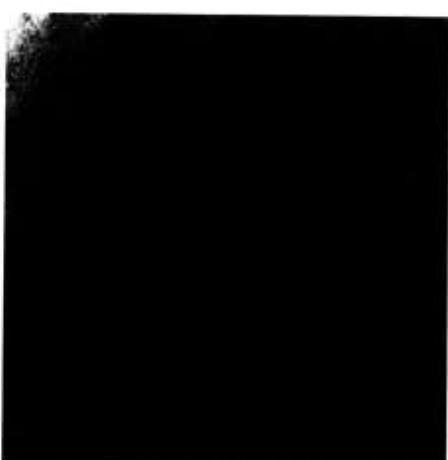
زوج DCT دو بعدی متاظر عبارتند از:

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (48-5.3)$$

برای $u, v = 0, 1, 2, \dots, N-1$



(الف)

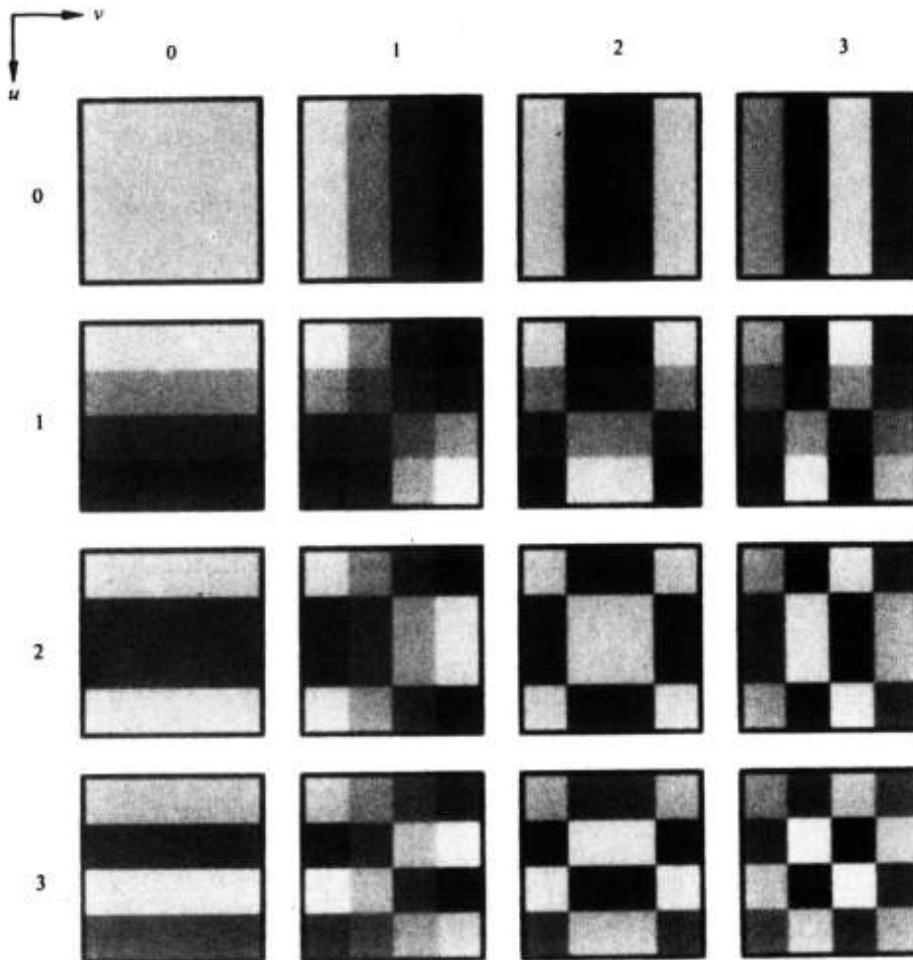


(ب)

شکل ۳.۲۷. یک تصویر ساده و لگاریتم دامنه تبدیل هادامارد آن.

1) Discrete Cosine Transform

۵.۳ سایر تبدیلهای جداپذیر تصویر ۱۸۵



شکل ۲۸.۳ توابع پایه تبدیل کسینوسی گسته برای $N=4$. هر بلوک شامل 4×4 عنصر متناظر با x و y که از تأثیر می کنند، می باشد. مبدأ هر بلوک در سمت چپ بالای آن است. بزرگترین مقدار با سفید نشان داده شده است. سایر مقادیر با خاکستری نشان داده می شوند، که رنگ تیره تر بیانگر مقدار کوچکتر است.

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (49-5, 3)$$

برای $1 - N - 1$ ، $x, y = 0, 1, 2, \dots, N - 1$ داده می شوند.

در سالهای اخیر تبدیل کسینوسی گسته به روش منتخب برای فشرده سازی داده های تصویری گردیده

۱۸۶ پردازش تصویر رقمه

است، که دلایل آن در فصل ۶ شرح داده شده است. شکل ۳.۲۸. توابع پایه DCT برای $N=4$ را نشان می دهد و شکل ۳.۲۹. یک تصویر ساده و لگاریتم دامنه تبدیل کسینوسی گسته آن را نمایش می دهد.

۴.۵.۳ تبدیل ها آر^۱

تبدیل هایی که تاکنون تشریح شدند، ابزارهای مفید و معروف پردازش سیگنال و تصویر رقمه هستند. تبدیل هایی که در این بخش و بخش ۳.۵. بیان می شوند، تا حد قابل ملاحظه ای ناشناخته هستند و در عمل هم آن قدر مفید نیستند که فقط برای تکمیل بحث آنها را آورده ایم.

تبدیل ها آر براساس توابع پایه ها آر، $(z)_k$ ، است که توابع مذکور در بازه 'بسته و پیوسته' $[0, 1]$ و برای $k = 0, 1, 2, \dots, N-1$ تعریف می شوند که 2^k . مرحله اول در تولید تبدیل ها آر، توجه به این واقعیت است که می توان عدد صحیح k را به طور یکتا به صورت

$$k = 2^p + q - 1 \quad (50-5.3)$$

تجزیه نمود که $1 \leq p \leq n-1$ ، که برای $p=0$ ، $q=0$ برابر ۰ یا ۱ است و برای $p \neq 0$ ، $1 \leq q \leq 2^p$. مثلاً اگر $N=4$ باشد:

k	p	q
0	0	0
1	0	1
2	1	1
3	1	2

با این زمینه، توابع ها آر به صورت زیر تعریف می شوند:

$$h_0(z) \triangleq h_{00}(z) = \frac{1}{\sqrt{N}} \quad z \in [0, 1] \quad \text{برای } (51-5.3 \text{ الف)}$$

و

$$h_k(z) \triangleq h_{pq}(z) = \frac{1}{\sqrt{N}} \begin{cases} 2^{pq} & \frac{q-1}{2^p} \leq z < \frac{q-1/2}{2^p} \\ -2^{pq} & \frac{q-1/2}{2^p} \leq z < \frac{q}{2^p} \\ 0 & z \in [0, 1] \end{cases} \quad (51-5.3 \text{ ب})$$

در غیر این صورت برای

1) Haar transform

۵.۳ سایر تبدیلهای جداپنیر تصویر ۱۸۷



(الف)



(ب)

شکل ۲۹.۳ یک تصویر ساده و لگاریتم دامنه تبدیل کسینوسی گستته آن.

این نتایج، بدست آوردن ماتریس های تبدیل ها آر مرتبه $N \times N$ را با تشکیل سطر نام ماتریس ها آر از عناصر $(z)h_i$ بازی $N/N, 1/N, 2/N, \dots, (N-1)/N$ ممکن می سازد. برای نمونه، وقتی $z = 0/N, 1/N, 2/N, \dots, (N-1)/N$ بازی $h_0(z) = 0/2, 1/2, 1/2, \dots, (N-1)/2$ محاسبه می شود. از معادله ۲۹.۵-۱، اولین سطر ماتریس ها آر با استفاده از $h_0(z) = 0/2, 1/2, 1/2, \dots, (N-1)/2$ برابر $\sqrt{2}/1$ و مستقل از z است، بنابراین هر دو عنصر سطر اول ماتریس برابر با $1/\sqrt{2}$ هستند. سطر دوم با محاسبه $h_1(z) = 0/2, 1/2, 1/2, \dots, (N-1)/2$ بدست می آید که در این حالت $k=1$ از معادله ۲۹.۵-۲، $h_1(0) = 2^0/\sqrt{2} = 1/\sqrt{2}$ و $h_1(1) = 2^1/\sqrt{2} = -1/\sqrt{2}$ می باشد. آنگاه ماتریس ها آر 2×2 عبارتست از:

$$\mathbf{A}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

با پیروی از روشی مشابه، ماتریس ها آر بازی $N = 4$ به صورت زیر بدست می آید:

$$\mathbf{A}_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}$$

ماتریس های ها آر متعامد هستند و بنابراین خواص لازمی که پیاده سازی الگوریتم ها آر سریع براساس

روابط ماتریسی تشریح شده در رابطه با معادلات (۳.۵-۱۰) و (۳.۵-۱۲) را ممکن می سازند، دارا هستند.

۵.۵.۳ تبدیل اسلنت^۱

ماتریس تبدیل اسلنت از مرتبه $N \times N$ با عبارت بازگشتی

$$\mathbf{S}_N = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & & & & \\ a_N & b_N & 0 & & & \\ & & I_{(N/2)-2} & & & \\ \hline 0 & 1 & & 0 & & \\ & & 0 & -1 & & \\ -b_N & a_N & & b_N & a_N & \\ \hline 0 & & & 0 & & \\ & & & & -I_{(N/2)-2} & \\ \end{bmatrix} \begin{bmatrix} 1 & 0 & & & & \\ -a_N & b_N & 0 & & & \\ & & I_{(N/2)-2} & & & \\ \hline 0 & -1 & & 0 & & \\ & & & & 0 & \\ & & & & & \\ \hline 0 & & & & & \\ & & & & & \\ \end{bmatrix} \begin{bmatrix} S_{N/2} & 0 & & & & \\ & & & & & \\ & & & & & \\ \hline & & & & & \\ & & & & & \\ 0 & & & & & S_{N/2} \\ & & & & & \end{bmatrix} \quad (52-5.3)$$

بیان می شود که \mathbf{I}_M ماتریس همانی از مرتبه $M \times M$ است، و

$$\mathbf{S}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (53-5.3)$$

برای $1 < N$ ضرایب عبارتند از:

$$a_N = \left[\frac{3N^2}{4(N^2 - 1)} \right]^{1/2} \quad (54-5.3)$$

$$b_N = \left[\frac{N^2 - 4}{4(N^2 - 1)} \right]^{1/2} \quad (54-5.3)$$

یک مثال از کاربرد معادلات (۳.۵-۱۰) و (۳.۵-۱۲)، ماتریس اسلنت مرتبه ۴ می باشد:

$$\mathbf{S}_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ \frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{-1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} \\ 1 & -1 & -1 & 1 \\ \frac{1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} & \frac{3}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{bmatrix}$$

1) Slant transform

۱۸۹ ۶.۳ تبدیل هتلینگ

ماتریس های اسلنت متعامد هستند و بنابراین خواص لازمی که پیاده سازی الگوریتم تبدیل اسلنت سریع براساس روابط ماتریسی بیان شده در ارتباط با معادلات (۳.۵-۱۰) و (۳.۵-۱۲) را ممکن می سازند، دارا هستند.

۶.۳ تبدیل هتلینگ^۱

برخلاف تبدیلات قبلی، تبدیل هتلینگ^۲ که در این بخش بیان می شود، براساس خواص آماری نمایش برداری می باشد. تبدیل هتلینگ چند خاصیت مفید دارد که آن را ابزار مهمی برای پردازش تصویر می سازد.

جامعه ای از بردارهای تصادفی به شکل

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (۱-۶.۳)$$

را در نظر گیرید. بردار متوسط جامعه به صورت

$$\mathbf{m}_* = E\{\mathbf{x}\} \quad (۲-۶.۳)$$

تعريف می شود که $\{x_i\}_{i=1}^n$ مقدار امید آرگومان است و زیرنویس \mathbf{x} نشان می دهد که \mathbf{m} مربوط به جامعه بردارهای \mathbf{x} است. بیاد آورید که مقدار امید بردار یا ماتریس، با گرفتن امید هر عنصر آن بدست می آید.

ماتریس کوواریانس^۳ جامعه برداری به صورت

$$\mathbf{C}_* = E\{(\mathbf{x} - \mathbf{m}_*)(\mathbf{x} - \mathbf{m}_*)^T\} \quad (۳-۶.۳)$$

تعريف می شود که \mathbf{T} نشانه عمل ترانهاده گیری^۴ برداری است. چون \mathbf{x} n بعدی است، \mathbf{C}_* و

^۱ عموماً این تبدیل با عنوان های تبدیل بردار ویژه، تبدیل مؤلفه اصلی یا تبدیل کارهاتون-لو^۴ گسته نیز اشاره می شود.

1) Hotelling transform 2) Covariance matrix 3) Transposition 4) Karhunen-Loeve

۱۴۰ پردازش تصویر رقمه

$(\mathbf{x} - \mathbf{m}_*)^T (\mathbf{x} - \mathbf{m}_*)$ ماتریسهای از مرتبه $n \times n$ هستند. عنصر C_{ij} از \mathbf{C}_* ، واریانس x_i ، مؤلفه آن بردارهای \mathbf{x} جامعه و عنصر C_{ji} ، کوواریانس بین عناصر x_i و x_j این بردارها می باشد. ماتریس \mathbf{C}_* حقیقی و متقارن است. اگر عناصر x_i و x_j ناهمبسته باشند، کوواریانس آنها صفر است و بنابراین $C_{ij} = C_{ji} = 0$. برای M بردار نمونه از یک جامعه تصادفی، می توان با روابط زیر، بردار میانگین و ماتریس کوواریانس را از روی نمونه ها تقریب زد:

$$\mathbf{m}_* = \frac{1}{M} \sum_{k=1}^M \mathbf{x}_k \quad (4-6.3)$$

۵

$$\mathbf{C}_* = \frac{1}{M} \sum_{k=1}^M \mathbf{x}_k \mathbf{x}_k^T - \mathbf{m}_* \mathbf{m}_*^T. \quad (5-6.3)$$

مثال: به عنوان مثالی که طرز کار معادلات (۳.۴-۶) و (۵-۶.۳) را روشن می سازد، چهار بردار ستونی $\mathbf{x}_1 = (0,0,0)^T$, $\mathbf{x}_2 = (1,0,0)^T$, $\mathbf{x}_3 = (1,1,0)^T$, $\mathbf{x}_4 = (1,0,1)^T$ را در نظر بگیرید، که ترانهاده بدین منظور استفاده شده است که بتوان بردارهای ستونی را روی یک خط افقی نیز نمایش داد. با کاربرد معادله (۳.۴-۶)، بردار میانگین بدست می آید:

$$\mathbf{m}_* = \frac{1}{4} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

به طور مشابه، کاربرد معادله (۳.۵-۶) ماتریس کوواریانس زیر را می دهد:

$$\mathbf{C}_* = \frac{1}{16} \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}$$

تمام عناصر قطر اصلی برابر هستند که نشان می دهد هر سه مؤلفه بردارهای جامعه، واریانس برابری دارند. همچنین، عناصر x_1 و x_2 نیز x_3 و x_4 همبسته مثبت و عناصر x_1 و x_3 همبسته منفی هستند. □

چون \mathbf{C}_* حقیقی و متقارن است، همواره پذاکردن یک مجموعه n تابی از بردارهای ویژه متعامد ممکن می باشد (Noble [1969]). فرض کنید e_i و λ_i ، که $i = 1, 2, \dots, n$ ، به ترتیب بردارهای ویژه

٤٤١ تبدیل هتلینگ

و مقادیر ویژه مربوط به \mathbf{C}_y باشند،^۱ که (برای سهولت) به صورت نزولی مرتب شده اند طوری که برای j داریم، $\lambda_{j+1} \geq \lambda_j \geq \dots \geq \lambda_1 \geq 0$. همچنین فرض کنید \mathbf{A} ماتریسی باشد که سطراهایش بردارهای ویژه \mathbf{C}_y هستند و این بردارهای ویژه، طوری مرتب شده اند که اولین سطر \mathbf{A} بردار ویژه متناظر با بزرگترین مقدار ویژه و آخرین سطر آن بردار ویژه متناظر با کوچکترین مقدار ویژه می باشد. فرض کنید که \mathbf{A} ماتریس تبدیلی است که مانند زیر بردارهای \mathbf{x} را به بردارهایی که با \mathbf{y} نشان داده می شوند، نگاشت می کند:

$$\mathbf{y} = \mathbf{A}(\mathbf{x} - \mathbf{m}_s). \quad (6-6.3)$$

معادله (۶-۶.۳) تبدیل هتلینگ نامیده می شود. میانگین بردارهای \mathbf{y} حاصل از این تبدیل، صفر است؛ یعنی،

$$\mathbf{m}_s = \mathbf{0} \quad (7-6.3)$$

و با معادله ذیل می توان ماتریس کوواریانس \mathbf{y} را بر حسب \mathbf{A} و \mathbf{C}_y بدست آورد:

$$\mathbf{C}_y = \mathbf{AC}_x\mathbf{A}^T. \quad (8-6.3)$$

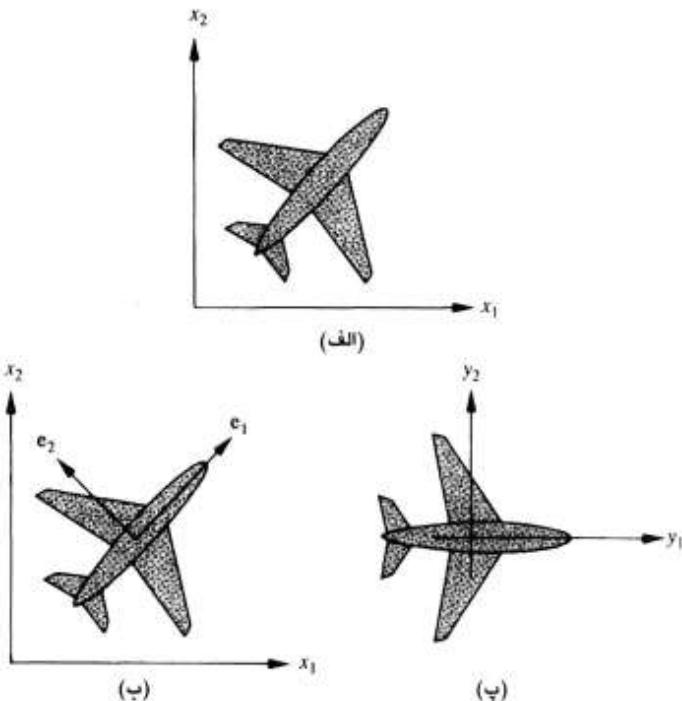
گذشته از این، \mathbf{C}_y ماتریسی قطری است که عناصر قطر اصلی آن مقادیر ویژه \mathbf{C}_y هستند؛ یعنی،

$$\mathbf{C}_y = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \quad (9-6.3)$$

عناصر غیر قطری ماتریس کوواریانس صفر هستند، بنابراین عناصر بردارهای \mathbf{y} ناهمبسته هستند. در ذهن داشته باشید که λ_i ها مقادیر ویژه \mathbf{C}_y هستند و عناصر قطر اصلی یک ماتریس قطری، مقادیر ویژه آن ماتریس هستند ([1969] Noble). بنابراین \mathbf{C}_y و \mathbf{C}_x مقادیر ویژه و نیز بردارهای ویژه یکسانی دارند.

مثال: شکل ۳۰. مفاهیمی که اخیراً بیان شدند، را مصور می کند. با شیوه دودویی نشان داده شده به

^۱ طبق تعریف، بردارها و مقادیر ویژه ماتریس \mathbf{C} با ابعاد $n \times n$ بازای $i = 1, 2, \dots, n$ در رابطه $\mathbf{Ce}_i = \lambda_i e_i$ مصدق می کنند.



شکل ۳.۳۰ (الف) یک شیئ دودویی؛ (ب) محورهای اصلی (بردارهای ویژه) آن؛ (پ) شیئ چرخانده شده با استفاده از معادله (۳.۶.۳)

عنوان یک جامعه دو بعدی رفتار می شود. به عبارت دیگر، با هر پیکسل شیئ به عنوان یک بردار دو بعدی $x = (a, b)^T$ بروخورد می شود که a و b مقادیر مختصات آن پیکسل نسبت به محورهای x_1 و x_2 هستند. از این بردارها برای محاسبه بردار میانگین و ماتریس کوواریانس جامعه (شیئ) استفاده می شود.

اثر خالص استفاده از معادله (۳.۶-۶) تعیین سامانه مختصات جدیدی است که مبدأش در مرکز ثقل جامعه و محورهایش درجهت بردارهای ویژه C هستند. این سامانه' جدید در شکل ۳.۳۰ (ب) دیده می شود. همانطور که در شکل ۳.۳۰ (ب) دیده می شود، تبدیل با معادله (۳.۶-۶) تبدیلی چرخشی است که داده ها را با بردارهای ویژه همراستا می کند. در واقع، این همراستاسازی است که داده ها را ناهمبسته می کند. گذشته از این، چون مقادیر ویژه در طول قطر اصلی C ظاهر می شوند، λ واریانس مؤلفه y در طول بردار ویژه e است.

مفهوم همراستاسازی یک شیئ دو بعدی با بردار ویژه اصلیش، نقش مهمی در تحلیل تصویر دارد. همان طور که در فصول ۸ و ۹ نشان می دهیم، پس از آن که یک شیئ از تصویر استخراج شد، روش های

۱۹۳ تبدیل هتلینگ ۶.۳

رایانه ای تشخیص شیئ عموماً حساس به چرخش شیئ هستند. چون به وضوح یکسان بودن شیئ قبل از تشخیص معلوم نیست، توانایی هماراستاسازی شیئ با محور اصلیش راه مطمئنی را برای حذف اثرات چرخش از فرآیند تحلیل تصویر فراهم می سازد.

□

یک ویژگی مهم دیگر تبدیل هتلینگ درباره 'بازسازی x از y می باشد. چون سطرهای A بردارهای متعامد و یکه هستند، آن گاه $A^T = A^{-1}$ ، و هر بردار x را می توان با استفاده از رابطه

$$x = A^T y + m, \quad (10-6.3)$$

از بردار y متاظرش بازسازی کرد. با این حال فرض کنید که به جای استفاده از تمام مقادیر ویژه C_x ماتریس A_K را از K بردار ویژه متاظر با K مقدار ویژه بزرگتر تشکیل دهیم. بنابراین A_K ماتریس تبدیلی از مرتبه $n \times K$ می باشد. پس بردارهای y, K بعدی خواهند بود و بازسازی ارائه شده در معادله (۱۰-۶.۳) دیگر دقیق نخواهد بود. بردار بازسازی شده با استفاده از A_K

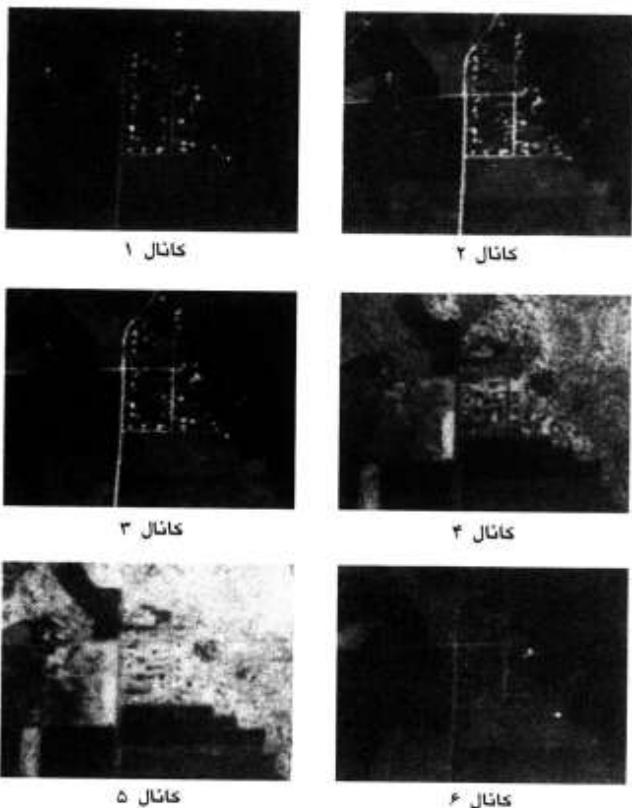
$$\hat{x} = A_K^T y + m, \quad (11-6.3)$$

می باشد. می توان نشان داد که خطای میانگین مرربع بین \hat{x} و x با عبارت زیر داده می شود:

$$\begin{aligned} e_m &= \sum_{j=1}^n \lambda_j - \sum_{j=1}^K \lambda_j \\ &= \sum_{j=K+1}^n \lambda_j \end{aligned} \quad (12-6.3)$$

قسمت اول معادله (۱۲-۶.۳) نشان می دهد که اگر $n = K$ (یعنی اگر تمام بردارهای ویژه در تبدیل استفاده شوند)، خطای صفر است. همچنین چون λ ها به طور یکنواکا کاهش می یابند، معادله (۱۲-۶.۳) نشان می دهد که می توان با انتخاب K بردار ویژه متاظر با بزرگترین K مقدار ویژه خطا را به حداقل رساند. بنابراین تبدیل هتلینگ از این نظر بهینه است که خطای میانگین مرربع بین بردارهای x و تقریبهای \hat{x} آنها را به حداقل می رساند.

مثال: این بخش را با مثال دیگری از استفاده های فراوان تبدیل هتلینگ در پردازش تصویر پایان می دهیم. شکل ۳.۳۱، شش تصویر تولید شده با یک پیمایشگر رنگی ۶ نواری که در طول موجهای ارائه شده در جدول ۳.۶ کار می کند، را نشان می دهد. با نمایش تصاویر به صورت شکل ۳.۳۲، تشكیل بردار ۶



شکل ۳۱.۳ شش تصویر طیفی از یک پیماشگر هوایی. (با اجازه آزمایشگاه کاربردهای سنجش از راه دور، دانشگاه Purdue)

بعدی $x^T = (x_1, x_2, \dots, x_n)$ از هر مجموعه پیکسل های متناظر در تصاویر ممکن می گردد. در این کاربرد ویژه، تفکیک تصاویر 384×239 است، بنابراین جامعه دارای $91,776$ بردار است که باید بردار میانگین و ماتریس کوواریانس از روی آنها محاسبه شوند. جدول ۳.۷ مقادیر ویژه C را نشان می دهد؛ به غالب بودن دو مؤلفه اول توجه کنید.

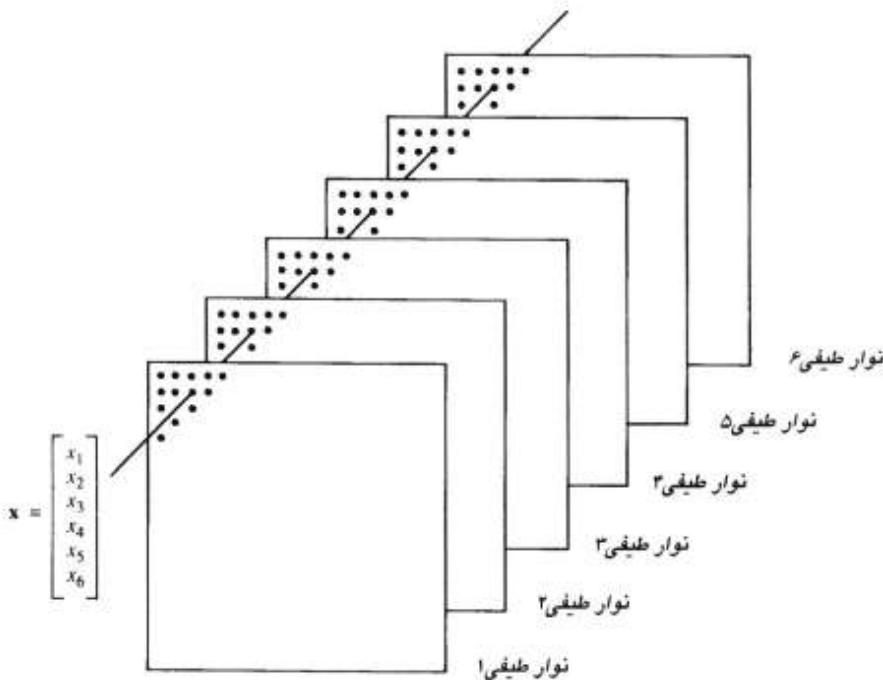
با استفاده از معادله (۳.۶-۶)، مجموعه ای از بردارهای تبدیل z متناظر با بردارهای x تولید گردید. آن گاه با استفاده از آنها و با معکوس کردن فرآیند نشان داده شده در شکل ۳۲.۳، شش تصویر مزبور با هم ترکیب می شوند. شکل ۳.۳۳ نتایج را نشان می دهد. مؤلفه ۱ بر تصویر حاصل از تمام مؤلفه های z بردارهای تبدیل دلالت می کند و به همین ترتیب برای ۵ تصویر دیگر. براساس نظریه اساسی ماتریس، z

1) air-borne Scanner

جدول ۳.۰ ع شماره کانال ها و طول موج ها

کانال	نوار طول موج (μm)
1	0.40–0.44
2	0.62–0.66
3	0.66–0.72
4	0.80–1.00
5	1.00–1.40
6	2.00–2.60

با ضرب داخلی (نقطه ای) سطر اول \mathbf{A} در بردار ستونی $(\mathbf{m} - \mathbf{x})^T$ بدست می آید. سطر اول \mathbf{A} بردار ویژه متناظر با بزرگترین مقدار ویژه ماتریس کوواریانس جامعه است و این مقدار ویژه، واریانس سطوح خاکستری اولین تصویر تبدیل را می دهد. بنابراین براساس اعداد جدول ۳.۷، این تصویر باید بالاترین تمایز^۱ را داشته باشد که چنین حالتی در شکل ۳.۳۳ کاملاً واضح است. چون دو تصویر اول حدود ۹۴ درصد واریانس کل را دارند، دور از انتظار نیست که چهار تصویر دیگر تمایز پایینی داشته باشند. بنابراین



شکل ۳.۲۰.۳ تشکیل بردار از پیکسلهای متناظر در ع تصویر

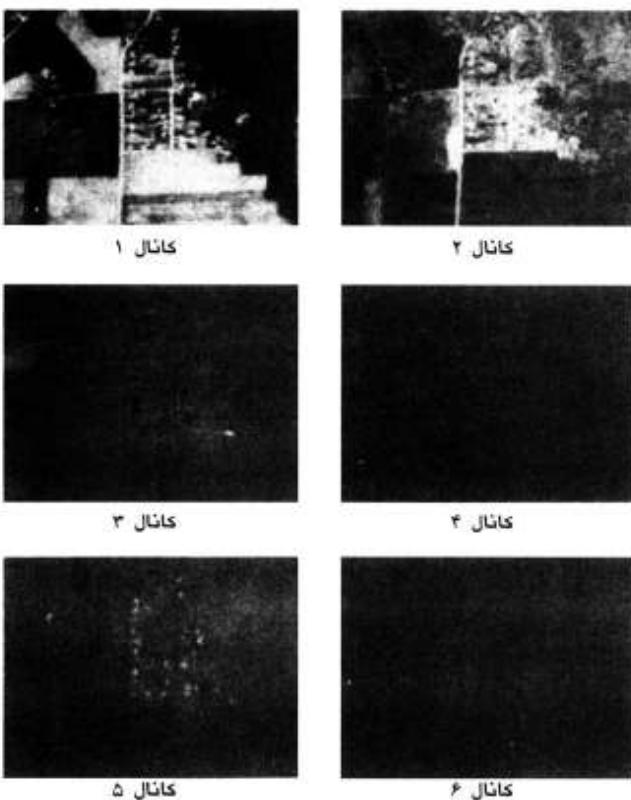
1) Contrast

۱۹۶ پردازش تصویر رفعی

جدول ۷.۳ مقادیر ویژه ماتریس کروواریانس تصاویر نشان داده شده در شکل ۳۱.۳

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
3210	931.4	118.5	83.88	64.00	13.40

اگر بجای ذخیره همه ۶ تصویر برای آینده، تنها اولین دو تصویر تبدیل را به همراه m و دو سطر اول، A ذخیره کنیم، در مرحله 'بعد می توان تقریب معنبری برای شش تصویر اولیه بازسازی کرد. این توانایی برای فشرده سازی داده ها محصول فرعی مفیدی از تبدیل هتلینگ است، اگرچه در استانداردهای فعلی فشرده سازی مورد استفاده قرار نمی گیرد (به فصل ۶ مراجعه کنید). □



شکل ۳۳.۳ شش تصویر مؤلفه ای اصلی که از داده های شکل ۳۱.۳ محاسبه شده اند. (با اجازه از آزمایشگاه کاربردهای سنجش از راه دور، دانشگاه *Purdue*).

۷.۳ نکات پایانی

منظور اصلی این فصل، ارائه یک مبنای نظری از تبدیل های تصویر و خواص آنها است. با توجه به این

هدف، نکات اساسی لازم برای فهم پایه ای این مفاهیم ارائه شدند.

تأکیدی که روی تبدیل فوریه شد، محدوده وسیع کاربرد آن در مسائل پردازش تصویر را نشان می دهد.

به دلیل جنبه های محاسباتی، مبحث تبدیل فوریه سریع اهمیت ویژه ای دارد. از جدایزیری، مرکزی کردن^۱ و خواص کانولوشن تبدیل فوریه نیز به طور گسترده در فصول بعدی استفاده می شود.

همان طور که در توضیحات بعدی روشن می شود، نظریه 'تبدیل نقشی' محوری در گسترش پردازش تصویر ایفا کرده است. در چند فصل بعدی، چند کاربرد تبدیل فوریه برای ارتقا و بازیابی تصویر را مورد توجه قرار می دهیم. توضیح بیشتر در مورد سایر تبدیل های عنوان شده در بخش ۳، ۵، ۶، در فصل ۶ خواهد آمد، جایی که سودمندی آن تبدیل ها در فشرده سازی داده با تفصیل بیشتری مورد بحث قرار می گیرد. تبدیل هتلینگ دوباره در فصول ۸ و ۹ در ارتباط با تراز سازی چرخش شیئی مورد استفاده قرار خواهد گرفت.

مراجع

معرفی مراجعت به تبدیل فوریه در سطحی مقدماتی صورت می گیرد. متنون اولیه توسط Papoulis[1962] و Titchmarsh[1948] برخورد نظری جامع تری با تبدیل فوریه پیوسته و خواص آن داشته اند. بیشتر کتب مهندسی مدار و مخابرات توضیحات مختلفی درباره 'تبدیل فوریه ارائه می کنند. کتب [Thomas[1969]، Carlson[1968]، VanValkenburg[1955] و Cooley, Lewis, Tukey[1988] نمونه ای از آنها هستند. بدست آوردن تبدیل فوریه گستره از شکل پیوسته آن نیز به طور گسترده در نوشته ها پوشش داده شده است. سه مرجع خوب در این مبحث Welch[1967] و Blackman[1955] هستند. Brigham[1974] بویژه مرجع اول و آخر برای مطالعه اولیه مناسب است.

روابط تبدیل فوریه سریع اغلب به Cooley[1965] و Tukey[1965] متنسب می شوند. در هر حال، FFT تاریخچه جالبی دارد که مطرح کردن آن در اینجا مفید است. در باسخ به مقاله 'Cooley-Tukey' Rudnick[1966] خبر داد که وی از روش مشابهی استفاده می کند که تعداد عملیاتش مناسب با $N \log_2 N$ بوده و براساس روش منتشر شده توسط Lanczos[1942] و Danielson[1942] می باشد. آن دو نیز به Runge[1903, 1905] به عنوان منبع روش خود ارجاع دادند. دو مقاله اخیر همراه با متن سخنرانی های Köring[1924]، فواید محاسباتی اصلی الگوریتم های FFT موجود را بیان می کنند.

روش های مشابهی نیز توسط Thomas[1958]، Good[1958]، Stumpff[1939]، Yates[1937] و Welch[1967a] تاریخچه و مقایسه 'جالبی از نتایج بدست منتشر شد. مقاله ای از Cooley, Lewis و Cooley-Tukey 1965 ارائه می کند.

الگوریتم FFT ارائه شده در این فصل به هیچ وجه منحصر بفرد نیست. مثلاً الگوریتم مشهور به Sande-Tukey (Sande[1966] و Gentleman[1966]) براساس روش دیگری برای تجزیه داده های ورودی است. کتاب Brigham[1974] حاوی شرح جامعی از این الگوریتم و سایر روابط FFT، از جمله روش های مبتنی بر مبنای غیر ۲ می باشد.

گرچه مافقط روی روش های رقمه تمرکز کرده ایم، لازم است اشاره کنیم که با روش های نوری نیز می توان تبدیل فوریه دو بعدی را بدست آورد (به مساله ۳. ۷. نگاه کنید). کتب Papoulis[1968]، Zajac[1937] و Hech Goodman[1968] جنبه های نظری و کاربردی نور و تبدیلات نوری را در سطحی مقدماتی پوشش می دهند.

برای مطالعه بیشتر در مورد روابط ماتریسی تبدیلات تصویر، کتاب Andrews[1970] را ببینید که مفهوم تجزیه ماتریسی را ارائه داده و سایر تبدیلات تصویر را نیز علاوه بر آنها که در این فصل معرفی شدند، شرح می دهد. مقالات Good[1958]، Genglman[1968]، Elliott Kahaner[1970] و کتاب Rao[1983] نیز مورد توجه هستند.

مطالعه اولین مقاله درباره 'تبدیل والش' (Walsh[1923]) از دیدگاه تاریخی ارزشمند است. مراجع دیگر درباره این تبدیل، Henderson[1964]، Johnson[1962]، Hammond[1949]، Fine[1949]، ۱۹۵۰ و Andrews[1970] و Shanks[1969] هستند.

برای مطالعه بیشتر در مورد تبدیل هادامارد، مقاله اولیه از Hadamard[1893] را ببینید. مطالعه Andrews[1970] و Whelchel[1968] نیز توصیه می گردد. دونوشه 'جالب دیگر که درباره' ماتریس های هادامارد مبتنی بر توان های اعداد صحیح غیر از ۲ باشد، Baumert[1962] و Golomb[1963] هستند. به نظر می رسد که مفهوم فرکانس مکانی توسط Harmuth[1970] معرفی شده باشد. مراجع برای تبدیل فوریه گسته Ahmed[1974] و همکاران[1974] و Rao[1975] و Ahmed[1975] هستند. مراجع آخر حاوی شرح گسترده ای درباره 'سایر تبدیلات متعامد نیز می باشد. مراجع درباره 'تبدیل های هار و اسلنت، Pratt[1991]، Jain[1989]، Shore[1973]، Harmuth[1970] و Jain[1989] هستند.

Hotelling[1933] اولین فردی بود که تبدیل را بدست آورده و منتشر کرد که آن تبدیل متغیرهای گسته را به ضرایب ناهمبسته تبدیل می نماید. او این روش را روش مؤلفه های اصلی^۱ نامید. مقاله وی نگاه قابل توجهی به این روش داشته و مطالعه آن مفید است. تبدیل هتلینگ مجددآسالها بعد توسط Schulthesis, Huang[1963] و Kramer[1963] مراجعت کرد. برای شرحی کلی درباره این موضوع به Lawley و Maxwell[1963] مراجعه کنید. تبدیل مشابه برای تبدیل داده های پیوسته به مجموعه ای از ضرایب ناهمبسته توسط Karhunen[1947] و Loeve[1948] کشف گردید که بسط کارهانن-لو^۲ نامیده می شود. برای دیدن شرح بسیار خوبی در این مورد به Selin[1965] مراجعت کنید. این نتیجه که بسط کارهانن-لو میانگین مربعات خطای قطع کردن^۳ را حداقل می کند، در ابتدا توسط Koschman[1954] منتشر شد و مجددآ توسط Brown[1960] نیز کشف گردید.

مسائل

۱.۱ الف) متغیرهای پیوسته فرض کنید و نشان دهید که تبدیل فوریه تابع ثابت $f(x, y) = 1$ ، تابع ضریبه واحد $\delta(u, v)$ است. با این تعریف که در $u = v = 0$ $\delta(u, v) = \infty$ و در سایر جاها $\delta(u, v) = 0$ است. (طبق معادله (۳.۴۶-۳.۴۷)، انتگرال تابع ضریبه واحد ۱ است).

ب) اگر $f(x, y) = 1$ تصویر رقمه $N \times N$ باشد، نتیجه چه خواهد بود؟

۱.۲ الف) با شروع از معادله (۳.۱-۹) نشان دهید که تبدیل فوریه تابع سینوسی دو بعدی عبارتست از:

$$n(x, y) = A \sin(u_0 x + v_0 y)$$

$$N(u, v) = -jA/2[\delta(u - u_0/2\pi, v - v_0/2\pi) - \delta(u + u_0/2\pi, v + v_0/2\pi)].$$

ب) طیف فوریه $N(u, v)$ را بدست آورید.

۱.۳ یک تابع حقیقی را می توان به صورت مجموع یک تابع زوج و یک تابع فرد تجزیه کرد.

الف) نشان دهید که $f_{\text{even}}(x) = \frac{1}{2}[f(x) + f(-x)]$ و $f_{\text{odd}}(x) = \frac{1}{2}[f(x) - f(-x)]$.

ب) نشان دهید که $\Re[f_{\text{even}}(x)] = \text{Re}\{\tilde{f}[f(x)]\}$ و $\Im[f_{\text{odd}}(x)] = j \text{Im}\{\tilde{f}[f(x)]\}$.

۱.۴ نشان دهید که تبدیل فوریه تابع خود همبستگی $f(x)$ ، طیف قدرت (چگالی

1) The method of principal components

2) Karhunen- Loeve expansion

3) Truncation

- | | |
|---|--|
| <p>۱۱.۳ با رجوع به بخش ۴.۱، نشان دهید که ضرب نیاز خواهد داشت؟</p> | <p>۱۰.۳ همان طور که در بخش ۴.۲ ذکر شد، در محاسبه $N \log N$ نقطه ای، N جمع و $N \log N$ ضرب مورد نیاز است. محاسبه FFT دو بعدی تصویر $N \times N$ به چند جمع و نشان دهید که تبدیل فوریه گسته و معکوسش توابعی متنابع هستند. برای سهولت، توابع را یک متغیره فرض کنید.</p> |
| <p>۸.۳ نشان دهید که تبدیل فوریه گسته و معکوسش توابعی متنابع هستند. برای سهولت، توابع را یک متغیره فرض کنید.</p> | <p>۷.۳ در بخش مراجع این فصل، این مطلب گفته شد که می‌توان با روش نوری نیز تبدیل فوریه را محاسبه کرد. فرآیند پایه‌ای در این روش شامل بدست آوردن عکسی شفاف^۱ از تصویر مورد نظر، قرار دادن آن عکس در مقابل عدسی محدب و تاباندن (عبور) اشعه‌ای همراستا از یک منبع نور همدوس^۲ (مثلاً از یک لیزر) از طریق عکس شفاف و عدسی می‌باشد. طبق علم نورشناسی، تبدیل فوریه دو بعدی تصویر روی صفحه کانونی عدسی تشکیل می‌شود. در اصل تبدیلی بیکران روی آن صفحه تشکیل می‌شود. در واقع طیف فوریه نشان داده شده در شکل ۳.۳ با استفاده از این روش و قرار دادن یک صفحه فیلم در صفحه کانونی عدسی بدست آمده است. با رجوع به شکل ۳.۳، طیف یک دایره سیاه کوچک، مجموعه‌ای از دایره‌های هم مرکز است. اگر طیف فوریه یک تصویر رقمه‌ی که دایره سیاه کوچکی روی یک زمینه روشن دارد، با استفاده از تبدیل فوریه گسته بدست آید، طیف حاصل شبیه چه خواهد بود؟ آن طیف را رسم کرده و استنتاج خود را توضیح دهید.</p> |
| <p>۶.۳ تبدیل فوریه زیر را بدست آورید. متغیرهای x و y را پیوسته فرض کنید.</p> | <p>۵.۳ درستی معادلات (۳.۷-۳.۷) و (۳.۷-۳.۷) را نشان دهید.</p> |

$$\alpha, W_{2M}^{2m} = W_M^m \quad (\text{الف})$$

$$\beta, W_M^{m+N} = W_N^m \quad (\text{ب})$$

$$\gamma, W_{2M}^{m+N} = -W_{2M}^m \quad (\text{ج})$$

۱۲۰.۳ بسیاری از برنامه های فعلی محاسبه FFT منحصر به داده های ورودی حقیقی و یک بعدی هستند.

(الف) نشان دهد چگونه می توان از چنین برنامه ای برای محاسبه FFT داده های مختلط یک بعدی استفاده نمود.

(ب) چگونه می توان با استفاده از این برنامه، FFT دو بعدی تصویری که پیکسل هایش اعداد حقیقی هستند، را محاسبه کرد؟

۱۳۰.۳ با رجوع به جدول ۳.۲، چگونه آرایه ۱۶ نقطه ای را برای استفاده در الگوریتم FFT دو برابر سازی متواالی، مرتب می کنید؟

۱۴۰.۳ نشان دهد که معادلات (۳.۱۵-۵) و (۳.۱۷-۵) یک زوج تبدیل را تشکیل می دهند. یعنی، ثابت کنید که این دو معادله معکوس های یکدیگرند.

۱۵۰.۳ یک تولیدکننده سوزن های تزریق مشکلی در تولید دارد که باعث می شود بعضی سوزنها نوک های بد شکل داشته باشند. با اینکه در صد سوزن های خراب بسیار کم است، هر سوزن خراب موقع تزریق درد زیادی را در مريض ایجاد می کند که باعث می شود به اعتبار تولیدکننده لطمہ بخورد. فرض کنید تولیدکننده شمارا استفاده می کند تا سامانه ای طراحی کنید که قادر به بازبینی تک تک سوزن ها باشد طوری که بتواند سوزن های خراب را جدا کند. با استفاده از تبدیل فوریه راه حلی پیشنهاد نماید. مسأله آنقدر برای تولیدکننده مهم است که می توانید فرض کنید هر وسیله ای که در حل مشکل موردنیاز باشد، تهیه می شود.

۱۶۰.۳ نشان دهد که معادلات (۳.۴۵-۵) و (۳.۴۶-۵) یک زوج تبدیل کسینوسی گسته تشکیل می دهند. یعنی ثابت کنید که این دو معادله معکوس یکدیگرند.

۱۷۰.۳ چون پنجه تبدیل کسینوسی گسته دو بعدی جدایذیر است، DCT دو بعدی تصویر را می توان با پیمایش های سطحی و سنتوی آن بوسیله الگوریتم DCT یک بعدی محاسبه کرد. در واقع، یک خاصیت جالب DCT یک بعدی این است که می توان آن را با الگوریتم FFT نیز

محاسبه کرد. چگونگی اجرای این محاسبه را به تفصیل نشان دهد.

۱۸.۳ با رجوع به توابع پایه تبدیل کسینوسی گسته بازای $N = 4$ (شکل ۳.۲۸)، بدون ماشین

حساب با مقادیر کار کرده، بلوک متناظر با $v = u$ را بسازید، از (u) α صرفنظر نماید.

۱۹.۳ ماتریس تبدیل ها آر بازای $8 = N$ را بدست آورید.

۲۰.۳ ماتریس تبدیل اسلشت بازای $8 = N$ را بدست آورید.

۲۱.۳ درستی معادلات $(3.7-6)$ ، $(3.8-6)$ و $(3.9-6)$ را اثبات کنید.

۲۲.۳ جمله‌ای در انتهای آخرین مثال بحث 3.6 آمده مبنی بر این که تنها با استفاده از دو

تصویر مؤلفه‌ای اصلی مربوط به دو مقدار ویژه بزرگتر می‌توان تقریب‌های معتبری را برای

شش تصویر اولیه بازسازی کرد. میانگین مربعات خطای در این عمل وارد می‌شود، چه

خواهد بود؟ پاسخ خود را به صورت درصدی از حداقل خطای ممکن بیان کنید.

۲۳.۳ فرض کنید که برای مجموعه‌ای از تصاویر 64×64 ثابت شود که ماتریس کوواریانس داده

شده با معادله $(3.9-6)$ ماتریس همانی است. میانگین مربعات خطای بین تصاویر اولیه و

تصاویری که طبق معادله $(3.11-6)$ و تنها بانیمی از بردارهای ویژه اولیه بازسازی شده‌اند،

چه خواهد بود؟

فصل ۴

ارتقای تصویر

«کلی تفاوت وجود دارد بین این که شخص تاریکی را از میان نور ببیند یا روشنی را از میان تاریکی.»

دیوید لینتسی

هدف اصلی روش های ارتقا^۱، پردازش تصویر است به صورتی که تصویر حاصل برای یک کاربرد خاص نسبت به تصویر اولیه مناسب تر باشد. کلمه 'خاص در این جمله مهم است، زیرا در همین ابتدای کار مشخص می کند که روش های تشریح شده در این فصل بسیار مساله گرا^۲ هستند. بنابراین روشی که به عنوان مثال برای ارتقای تصاویر اشعه^۳ ایکس کاملاً مفید است، ممکن است بهترین روش برای ارتقای عکس های مریخ که کاوشگر فضایی آنها را ارسال کرده است، نباشد.

روش های تشریح شده در این فصل در دو گروه وسیع طبقه بندی می شوند: روش های میدان مکان و روش های میدان فرکانس. میدان مکان^۴ به خود تصویر اشاره دارد و روش های این گروه براساس کار مستقیم با پیکسل های تصویر می باشند. روش های میدان فرکانس^۵ مبتنی بر تبدیل فوریه هستند. استفاده از ترکیب های گوناگون دو نوع روش فوق نیز معمول است.

روش پایه ای که زیر بنای مطالب این فصل است، در بخش ۴.۱ ارائه می شود. بخش ۴.۲ درباره روش هایی است که مبتنی بر پردازش نقطه ای بوده و سطح خاکستری هر پیکسل را مستقل از همسایگانش تغییر می دهند. بخش ۴.۳ روش های ارتقای مبتنی بر پردازش نقاطی^۶ را پوشش می دهد. نقاط ها که به صورت زیر تصاویر کوچکی تعریف می شوند، برای تغییر هریک از پیکسل های تصویر در پردازش محلی بکار می روند. بخش ۴.۴ روش های گوناگون ارتقای تصویر در میدان فرکانس با استفاده از تبدیل فوریه

1) Enhancement

2) Problem-oriented

3) Spatial domain

4) Frequency domain

5) Mask processing

۴۰۳ پردازش تصویر رقمه

رامعرفی می کند. بخش ۴.۵ نشان می دهد که چگونه با استفاده از مشخصات میدان فرکانس، نقاب های مکانی کوچکی تولید کنیم تا بین وسیله پیوند مفهومی مفیدی بین بخش های ۴.۳ و ۴.۴ ایجاد شود. در نهایت بخش ۴.۶ در مورد ارتقای تصاویر رنگی می باشد. این بخش اصول تولید و احساس رنگ را توضیح می دهد، مدل های رنگی مختلف را معرفی می کند، و نیز روش های گوناگون ارتقای تصویر رنگی را با مقداری تفصیل بیان می کند.

۱۰.۴ زمینه

روش های ارتقای تصویر ارائه شده در این فصل، مبتنی بر فنون میدان های فرکانس یا مکان هستند. منظور از این بخش ارائه مفاهیم اصلی که این دوروش را زمینه سازی کرده و به هم مربوط ساخته اند، می باشد.

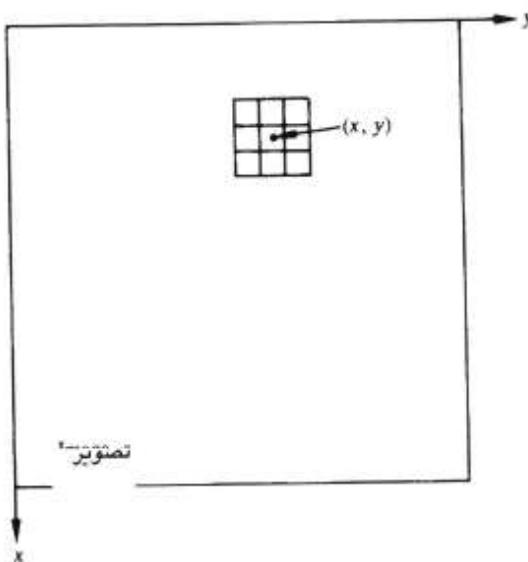
۱۰.۱.۲ روش های میدان مکان

اصطلاح میدان مکان به کل پیکسل های تشکیل دهنده تصویر اشاره دارد و روش های میدان مکان روش هایی هستند که به طور مستقیم بر روی پیکسل ها کار می کنند. توابع پردازش تصویر در میدان مکان را می توان به صورت

$$g(x, y) = T[f(x, y)] \quad (1-1.4)$$

بیان کرد که $f(x, y)$ تصویر ورودی، $g(x, y)$ تصویر پردازش شده، و T نیز عملگری روی f است که در یک همسایگی پیکسل (x, y) تعریف می شود. به علاوه T می تواند روی مجموعه ای از تصاویر هم عمل کند، نظری جمع پیکسل به پیکسل M تصویر برای کاهش نویز که این مورد در بخش ۴.۲.۴ مطرح می شود.

همان طور که شکل ۱.۱ نشان می دهد، روش اصلی برای تعریف یک همسایگی حول (x, y) ، استفاده از زیر تصویر های کوچک مربعی یا مستطیلی به مرکز (x, y) می باشد. مرکز زیر تصویر، مثلاً با شروع از گوشه 'چپ بالای تصویر'، پیکسل به پیکسل جابجا می شود و در هر نقطه (x, y) با استفاده از عملگر T مقدار g در آن نقطه تعیین می شود. گرچه گاهی سایر اشکال هندسی نظری تقریب های دایره، نیز



شکل ۱۰.۳ یک همسایگی 3×3 حول نقطه (x, y) تصویر

به عنوان همسایگی استفاده می شوند، تاکنون آرایه های مربعی و مستطیلی به علت سهولت پیاده سازی شان بیشترین کاربرد را داشته اند.

ساده ترین شکل T وقتی است که همسایگی 1×1 باشد. در این مورد، g تنها به مقدار f در (x, y) بستگی دارد و T یک تابع تبدیل سطح خاکستری^۱ (که نگاشت^۲ نیز خوانده می شود) است به صورت

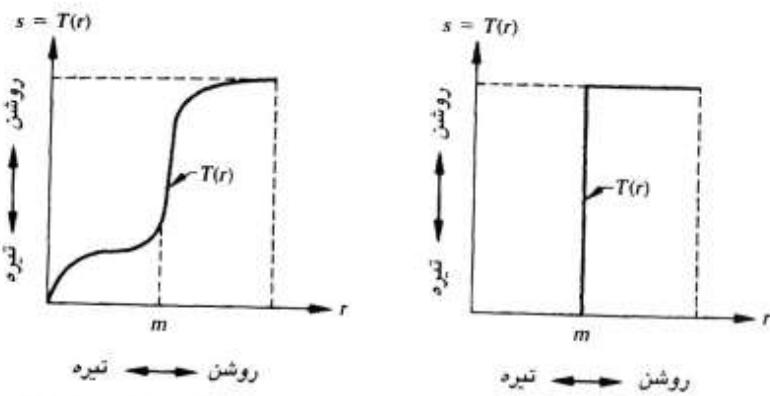
$$g = T(r) \quad (2-1.4)$$

بیان می شود که برای سهولت، در هر نقطه (x, y) متغیرهای s و r به ترتیب نشانه سطح خاکستری $f(x, y)$ و $g(x, y)$ در نظر گرفته می شوند. مثلاً اگر $T(r)$ به صورت شکل ۲.۴ (الف) باشد، با تاریک تر کردن سطح خاکستری پایین تر از m و روشن تر کردن سطح بالای m در تصویر اولیه، تصویری با تمایز بالاتر تولید می شود. در این روش که کشش تمایز^۳ نامیده می شود، مقادیر r کمتر از m با تابع تبدیل فوق به محدوده باریک s و به سمت سیاه فشرده می شوند؛ برای مقادیر r بزرگتر از m اثر مخالف رخ می دهد. در حالت حدی ارائه شده در شکل ۲.۴ (ب)، $T(r)$ تصویر دو سطحی (دودویی) تولید می کند. بعضی روش های پردازش نسبتاً ساده اما قوی را می توان با تبدیل های سطح خاکستری بیان کرد. چون ارتقا در هر نقطه تصویر تنها به سطح خاکستری آن نقطه بستگی دارد، روش های این گروه اغلب با عنوان پردازش

1) Gray-level transformation

2) Mapping

3) Contrast stretching



شکل ۲۰.۴ توابع تبدیل سطح خاکستری برای ارتقای تمایز

نقطه‌ای^۱ اشاره می‌شوند.

همسایگی‌های بزرگتر، پردازش‌های متنوعی را ممکن می‌سازند که خارج از محدوده 'ارتقای تصویر' می‌باشد. در هر حال با صرف نظر از کاربردهای خاص، روش کلی این است که مقادیر f در یک همسایگی پیش تعریف شده از نقطه (x, y) ، مقدار g در آن نقطه را تعیین نماید. یکی از روش‌های اصلی مبتنی بر استفاده از نقاب‌ها^۲ (که با عنوانین نمونه^۳، پنجره، یا فیلتر نیز اشاره می‌شوند) می‌باشد. در اصل، نقاب یک آرایه^۴ دو بعدی کوچک (مثلاً 3×3) نظیر شکل ۴.۱ است که مقادیر ضرایب آن ماهیت فرآیند، نظیر تیزکردن تصویر را تعیین می‌کند. فنون ارتقایی که مبتنی بر این نوع رهیافت هستند، اغلب با عنوان تیزکردن یا فیلتر کردن نقابی^۵ شناخته می‌شوند.

۲۰.۱.۴ روش‌های میدان فرکانس

اساس فنون میدان فرکانس، قضیه کانولوشن^۶ است. فرض کنید $(y, x) g(x, y)$ تصویری باشد که از کانولوشن تصویر $(y, x) f(x, y)$ با یک عملگر خطی مستقل از مکان $(y, x) h(x, y)$ بدست می‌آید،^۷ یعنی،

- | | | |
|---------------------------------|----------------|-------------|
| 1) Point processing | 2) Mask | 3) Template |
| 4) Mask processing or filtering | 5) Convolution | |

^۱ عملگر مستقل از مکان، عملگری است که خروجی آن تنها به مقدار $(y, x) f(x, y)$ در آن نقطه (یا حول آن) بستگی دارد، اما به محل آن نقطه بستگی ندارد. استقلال از مکان یک نیاز ضمیمی در تعریف انگرال‌های کانولوشن موجود در معادلات (۲۳-۳) و (۳۰-۳) است. مفاهیم خطی بودن و استقلال از مکان را در بخش ۱.۱ بانفصل بیشتر شرح می‌دهیم.

$$g(x, y) = h(x, y) * f(x, y). \quad (۴-۱. ۴)$$

آن گاه طبق قضیه کانولوشن (بخش ۳. ۸. ۳)، رابطه زیر در میدان فرکانس صدق می کند

$$G(u, v) = H(u, v)F(u, v) \quad (۴-۱. ۴)$$

که G و H به ترتیب تبدیل فوریه های g و f هستند. در نظریه سامانه های خطی، تبدیل $H(u, v)$ تابع انتقال^۱ فرآیند نامیده می شود. در اپتیک، $H(u, v)$ تابع انتقال نوری^۲ و بزرگی آن تابع انتقال مدولاسیون^۳ نامیده می شود.

مسائل بی شماری از ارتقای تصویر را می توان به صورت معادله^۴ (۴-۱. ۴) بیان کرد. در یک کاربرد ارتقای تصویر، معمولاً $f(x, y)$ داده می شود و هدف، پس از محاسبه $F(u, v)$ ، انتخاب $H(u, v)$ است طوری که تصویر مطلوب، یعنی

$$g(x, y) = \mathcal{F}^{-1}[H(u, v)F(u, v)] \quad (۵-۱. ۴)$$

چند ویژگی بر جسته $(y, f(x))$ را نمایش دهد. برای نمونه، لبه های $(y, f(x))$ را می توان با استفاده از تابع $H(u, v)$ که مؤلفه های فرکانس بالای $F(u, v)$ را تقویت می کند، بر جسته نمود. در مورد این موضوع در بخش ۴. ۴ بیشتر شرح می دهیم.

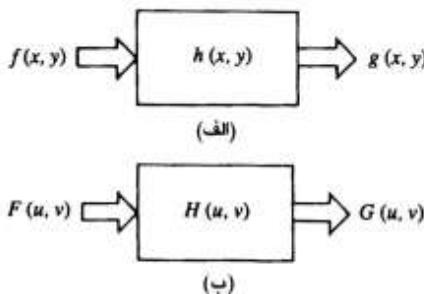
در شکل ۴. ۳ (الف)، سامانه ای را مشخص می کند که وظیفه اش تولید تصویر خروجی $(y, g(x, y))$ از تصویر ورودی $(y, f(x))$ است. این سامانه $(y, h(x, y))$ را با تصویر ورودی کانولوشن نموده، نتیجه را در خروجی بدست می دهد. قضیه کانولوشن نگاه دیگری به فرآیند را ممکن می سازد، چرا که با ضرب $F(u, v)$ در $H(u, v)$ نیز همان نتیجه^۵ قبلی، $G(u, v)$ ، بدست آمد. با گرفتن تبدیل فوریه معکوس از حاصل ضرب، تصویر مطلوب بدست می آید.

برای یک لحظه فرض کنید که $(y, h(x, y))$ نامعلوم است و تابع ضربه^۶ واحد (یعنی یک نقطه نور) را به سامانه اعمال می کنیم. روشن است که تبدیل فوریه^۷ ضربه واحد^۸ است، بنابراین از معادله^۹ (۴-۱. ۴) نتیجه می شود که $G(u, v) = H(u, v)$. با این شرایط تبدیل معکوس خروجی $(y, G(u, v))$ را بدست می دهد. این نتیجه که یک سامانه خطی مستقل از مکان کاملاً با پاسخ ضربه اش مشخص می شود، در

1) Transfer function

2) Optical transfer function

3) Modulation transfer function



شکل ۳.۰.۴ عملکرد یک سامانه خطی. در (الف)، خروجی سامانه، کانولوشن (y, x, h) با ورودی است. در (ب)، خروجی، حاصل ضرب (v, u, H) با ورودی است.

نظریه سامانه های خطی کاملاً مشهور است. یعنی اگر ضربه' واحد به چنین سامانه ای وارد شود، خروجی آن دقیقاً $H(u, v)$ ، تابع انتقال سامانه، است. به عبارت دیگر با ورود ضربه' تها، (y, x, h) در خروجی بدست می آید. به همین دلیل در نظریه سامانه های خطی، تبدیل معکوس تابع انتقال سامانه را پاسخ ضربه' می نامند. در اپتیک، (y, x, h) که معکوس تابع انتقال نوری است، تابع پخش نقطه ای^۱ خوانده می شود. این نام به خاطر این پدیده' نوری مشهور است که ضربه متناظر با یک نقطه' نور است و پاسخ سامانه نوری معمولاً به صورت مات کردن (پخش) نقطه می باشد. درجه' مات کردن نیز توسط کیفیت وسائل نوری معین می شود. بنابراین تابع انتقال نوری و تابع پخش نقطه ای یک سامانه خطی، تبدیل فوریه های یکدیگر هستند. این رابطه کاربرد قابل ملاحظه ای در بخش ۴.۳ دارد.

معادله (۴.۱-۳) فرآیندی مکانی را نشان می دهد که تقریباً مشابه کاربرد نقاب های بیان شده در بخش ۴.۱.۱ می باشد. در واقع بیان کانولوشن گسته به صورت معادله (۳.۳۵-۳)، در اصل نمایش ریاضی مراحلی است که اغلب در پیاده سازی فرآیند انتقال نقاب، تشریح شده در بخش ۴.۱.۱، استفاده می شوند. به این دلیل در اغلب موارد به (y, x, h) با عنوان نقاب کانولوشن مکانی^۲ اشاره می شود. اغلب همان عنوان در مورد نقاب های مکانی تشریح شده در بخش ۴.۱.۱ نیز بکار می رود. در حالت کلی به بیان دقیق، این عبارت صحیح نیست، زیرا همان طور که در بخش ۳.۳.۸ بیان شد، کانولوشن شامل برگردان کردن^۳ یکی از تصاویر حول مبدأ است. استفاده از این نام در ارتباط با نقاب های تشریح شده در بخش ۴.۱.۱ تنها وقتی صحیح است که نقاب حول مبدأ خود متقابن باشد.

1) Impulse response

2) Point spread function

3) Spatial convolution mask

4) Flipping (چرخش ۱۸۰ درجه)

۲۰۴ ارتقا با پردازش نقطه ای

گرچه احتمالاً می دانید، اما باز هم یادآور می شویم که هیچ نظریه 'کلی در مورد ارتقای تصویر وجود ندارد. وقتی تصویر برای تعبیر چشمی پردازش می شود، بینته، قضاوت کننده' نهایی در مورد کیفیت روش مورد استفاده می باشد. ارزیابی چشمی کیفیت تصویر، یک فرآیند کاملاً ذهنی است و به همین دلیل ارائه تعریفی دقیق و استاندارد از "تصویر خوب"، که برای مقایسه 'کارآیی الگوریتم ها ضروری است، بسیار مشکل می باشد. وقتی پردازش تصویر به منظور درک ماشینی باشد، کار ارزیابی مقداری آسانتر است. مثلاً در تشخیص نویسه، بهترین روش پردازش تصویر رو شی خواهد بود که بهترین نتایج تشخیص ماشینی را بدهد، اما حتی در حالاتی که معیار کارآیی کاملاً معین است، معمولاً تحلیلگر در انتخاب روش با مقدار معینی از سعی و خطأ مواجه است.

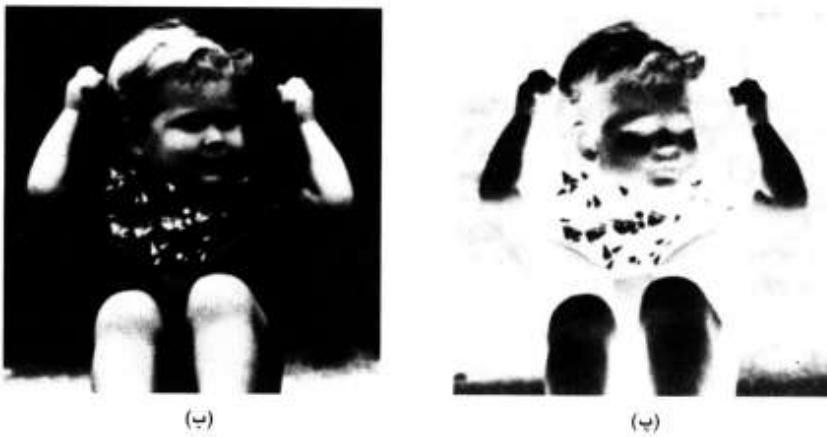
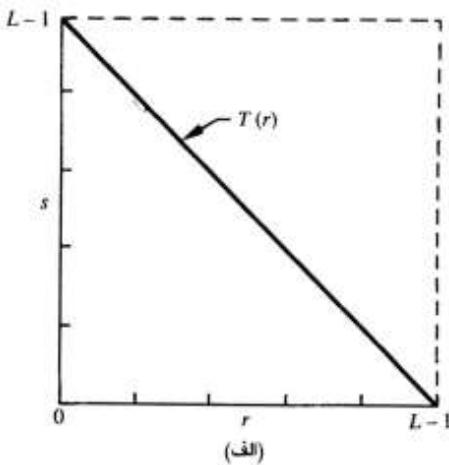
۲۰۴ ارتقا با پردازش نقطه ای

مطالعه 'فنون ارتقای تصویر را با بررسی روش های پردازشی که تنها مبتنی بر شدت یک نقطه' نکی هستند، شروع می کنیم. همان طور که در بخش ۱.۱ بیان شد، فرآیندهای تک نقطه ای از ساده ترین فنون ارتقای تصویر هستند. در ادامه 'این بحث، شدت پیکسل ها در قبل و بعد از پردازش را به ترتیب با ۲ و ۳ نشان می دهیم.

۱۰۴.۴ چند تبدیل شدت ساده

تصاویر منفی

تصاویر منفی^۱ رقمی در بسیاری کاربردها نظر نمایش عکس های پزشکی، و عکسبرداری از صحته با فیلم مثبت تکرنگ با این هدف که از تصاویر منفی حاصل به عنوان اسلایدهای^۲ عادی استفاده شود، مفید هستند. تصویر منفی رنگی با استفاده از تابع تبدیل $T(r) = r^{\alpha}$ که در شکل ۴.۴ (الف) دیده می شود، بدست می آید که r تعداد سطوح خاکستری است. هدف این است که ترتیب سیاه به سفید عکس شود طوری که با افزایش شدت ورودی شدت تصویر خروجی کاهش یابد. اشکال ۴.۴ (ب) و (پ) کاربرد این تبدیل ساده را نشان می دهند.



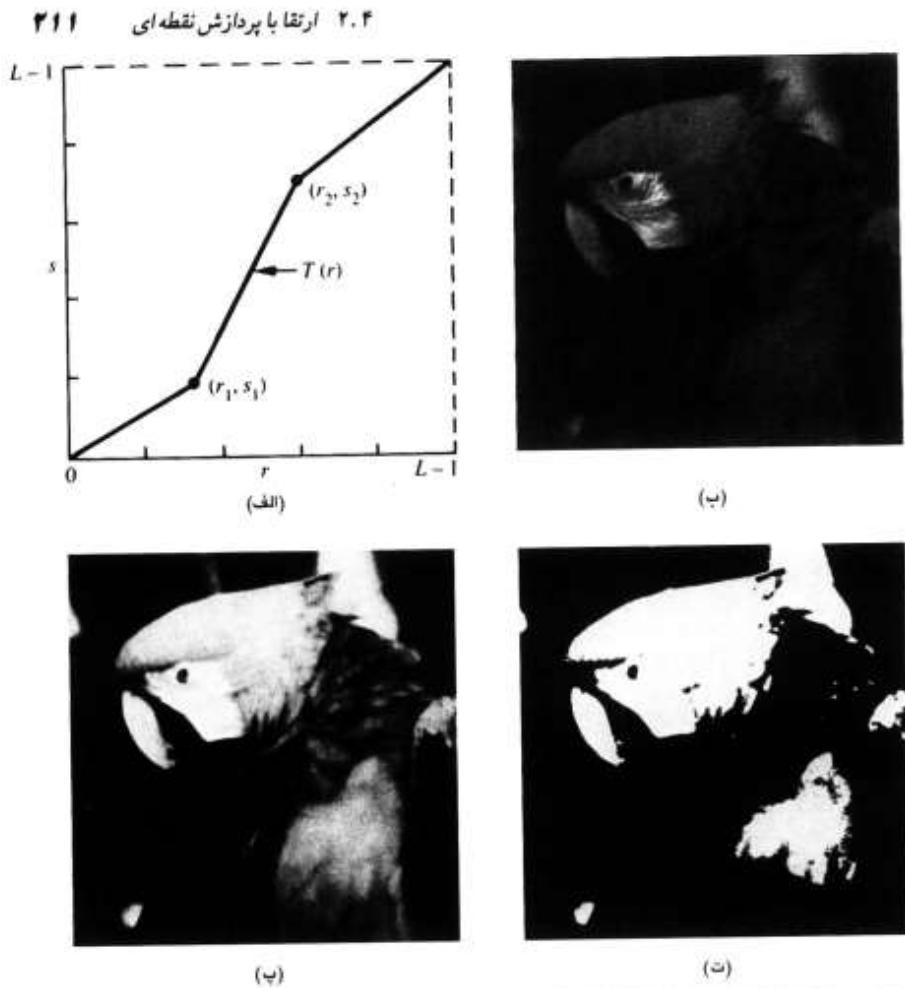
شکل ۴.۳ بدست آوردن تصویر منفی: (الف) تابع تبدیل سطوح خاکستری: (ب) یک تصویر؛ و (ج) منفی آن. در (الف) r و s به ترتیب نشانه سطوح خاکستری ورودی و خروجی هستند.

کنش تمایز

تصاویر با تمایز پایین ممکن است ناشی از روشنایی کم، کوچکی محدوده دینامیکی حسگر تصویربرداری یا حتی تنظیم اشتباہ دهانه عدسی¹ در طول مدت تصویربرداری² باشند. هدف از کنش تمایز، افزایش محدوده دینامیکی سطوح خاکستری تصویر تحت پردازش می‌باشد. شکل ۴.۵(الف) تبدیلی را که برای کنش تمایز بکار می‌رود، نشان می‌دهد. محل نقاط (r_1, s_1) و (r_2, s_2) شکل تابع تبدیل را کنترل می‌کند. برای نمونه اگر $r_1 = r_2$ و $s_1 = s_2$ باشد، تبدیل تابعی خطی است که هیچ تغییری

1) Lens aperture

2) Image acquisition



شکل ۲۰.۵ کشش تمایز: (الف) شکل تابع تبدیل؛ (ب) یک تصویر کم تمایز؛ (پ) حاصل کشش تمایز؛ (ت) حاصل آستانه گیری

در سطوح خاکستری ایجاد نمی‌کند. اگر $r_1 = r_2$ و $s_1 = s_2 = L - 1$ تبدیل یک تابع آستانه گیر^۱ است که تصویری دودویی تولید می‌کند. مقادیر میانی (r_1, s_1) و (r_2, s_2) درجات گوناگونی از گستردگی را در سطوح خاکستری تصویر خروجی تولید می‌کنند، بنابراین تمایز آن را تحت تأثیر قرار می‌دهند. به طور کلی فرض می‌شود که $r_1 \leq r_2 \leq s_1 \leq s_2$ تا این که تابع، تک مقداره و کاملاً صعودی باشد. این وضعیت ترتیب سطوح خاکستری را حفظ می‌کند، بنابراین از ایجاد نواقص در شدت تصویر پردازش شده ممانعت می‌نماید. در شکل ۲۰.۵ (ب) تصویری با تمایز پایین، در شکل ۲۰.۵ (پ) حاصل کشش تمایز آن، و

1) Thresholding function

۲۱۲ پردازش تصویر رقمه‌ی

در شکل ۴.۵ (ت) حاصل آستانه گیری از آن دیده می‌شود. سطح آستانه گیری $r = 128$ است، بدین معنا که خروجی برای هر سطح خاکستری در تصویر ورودی که ۱۲۸ یا بالاتر باشد، ۲۵۵ (سفید) و برای سایر مقادیر صفر (سیاه) قرارداده می‌شود.

فشرده سازی محدوده دینامیکی

گاهی اوقات محدوده دینامیکی تصویر پردازش شده به میزان زیادی از توان وسیله نمایش تجاوز می‌کند. در این شرایط تنها روشن ترین بخش های تصویر روی صفحه نمایش قابل مشاهده هستند. همین مشکلات در مورد ضبط تصویر روی فیلم نیز وجود دارد. همان طور که در بخش ۳.۳ تشریح شد، یکی از بروزهای متداول این مسئله در نمایش طیف فوریه تصویر است. یک راه مؤثر برای فشرده سازی محدوده دینامیکی مقادیر پیکسل ها، اجرای تبدیل شدت با رابطه:

$$s = c \log(1 + |r|) \quad (1-2.4)$$

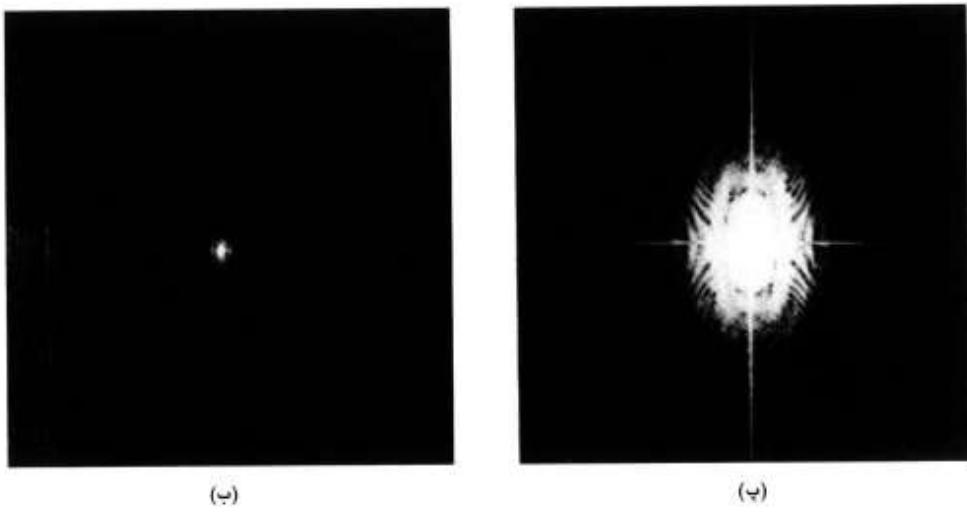
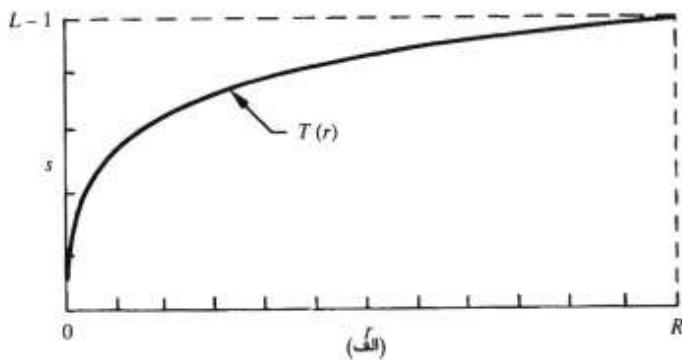
است که c ثابت مقیاس بوده و تابع لگاریتم فشرده سازی مطلوب را انجام می‌دهد. شکل ۴.۶ (الف) منحنی این تابع را نشان می‌دهد و شکل ۴.۶ (ب) طیف فوریه آن را با مقادیری در محدوده $[0, 10^6] \times [0, 2.5]$ نشان می‌دهد. همان طور که از چنین محدوده دینامیکی بزرگی انتظار می‌رود، وقتی این تابع برای نمایش بر روی سامانه ۸ بیتی، به طور خطی مقیاس شود، روشن ترین مقادیر از حد نمایشگر تجاوز می‌کنند. در این مورد $(|r| + 1) \log(|r| + 1)$ مقادیری در محدوده ۰ تا ۶.۴ دارد. ما خواستیم این محدوده را برای نمایش روی همان سامانه ۸ بیتی به محدوده بزرگتر $[0, 255] = [0, L - 1]$ مقیاس کنیم. بنابراین ضریب مقیاس $c = 255/6.4$ را انتخاب کردیم. شکل ۴.۶ (پ) حاصل را بعد از تبدیل و تغییر مقیاس نشان می‌دهد. به افزایش قابل رؤیت تصویر توجه کنید.

بخش بندی سطح خاکستری^۱

در بسیاری مواقع برجسته کردن محدوده مشخصی از سطوح خاکستری تصویر مطلوب است. از جمله کاربردهای این عمل ارتقای ویژگی هایی نظری توده های آب در تصاویر ماهواره ای و نیز بهبود نواقص

1) Gray-level slicing

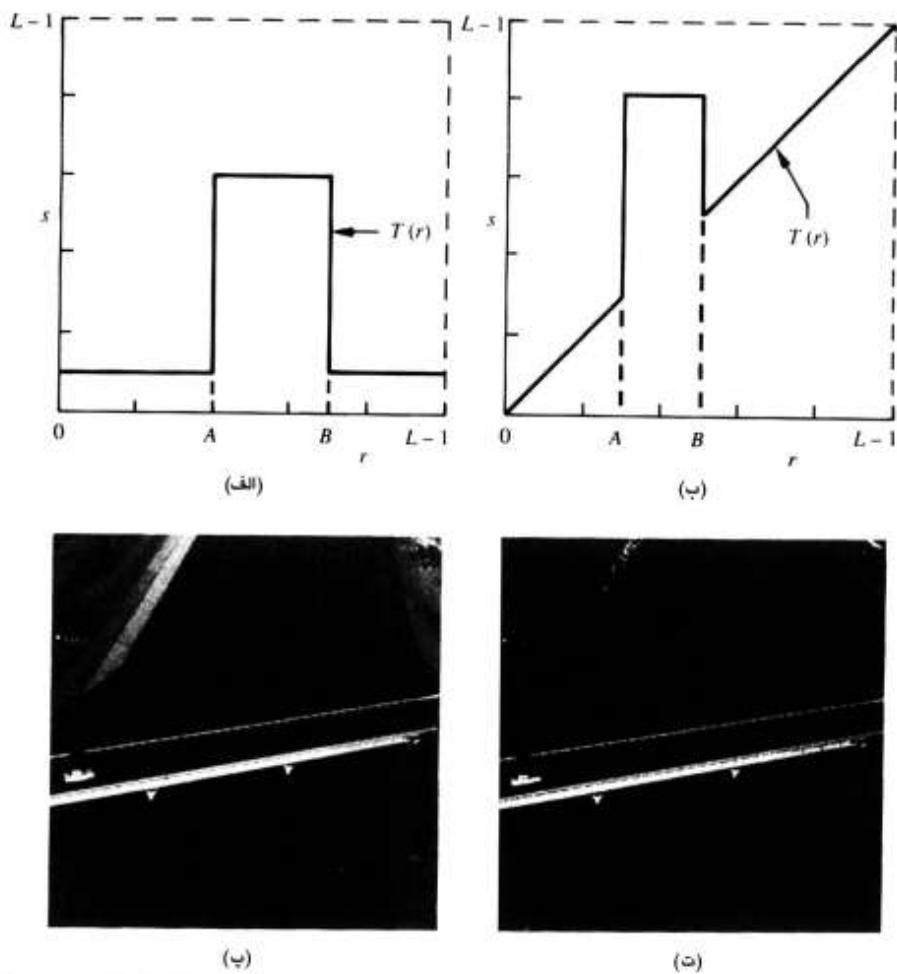
۲۱۳ ۲۰.۴ ارتقا با پردازش نقطه ای



شکل ۳.۶ فشرده کردن محدوده دینامیکی: (الف) تابع تبدیل لگاریتمی؛ (ب) تصویری با محدوده دینامیکی بزرگ (مقادیر پیکسل محدوده ای از $0 \text{ تا } 10^6 \times 2.5$ دارد)؛ (ب) حاصل پس از تبدیل

تصاویر اشعه ایکس می باشد. چندین راه برای اجرای قطعه بندی سطح وجود دارد، اما اغلب آنها گونه های مختلف دو روش پایه ای هستند. در روش اول تمام سطوح خاکستری محدوده 'مورد علاقه با یک مقدار بالا و سایر سطوح با یک مقدار پایین جایگزین می شود. این تبدیل که در شکل ۴.۷ (الف) نشان داده شده است، یک تصویر دودویی تولید می کند. روش دوم که براساس تبدیل نشان داده شده در شکل ۴.۷ (ب) است، محدوده 'مطلوب از سطوح خاکستری را روشن تر اما سایر سطوح را حفظ می کند. شکل ۴.۷ (پ) یک تصویر خاکستری و شکل ۴.۷ (ت) حاصل استفاده از تبدیل ارائه شده در شکل ۴.۷ (الف) را نشان می دهد. به آسانی می توان گونه های مختلفی را از دو تبدیل شدت که در شکل ۴.۷ آمده است، ارائه نمود.

پردازش تصویر رقیعی ۴۱۴



شکل ۴.۳ بخش بندی سطح شدت: (الف) تابع تبدیلی که شدت های بازه $[A, B]$ را برجسته^۱ می کند، در حالی که سایر شدت ها را به سطح ثابت پایینی کاهش می دهد؛ (ب) یک تبدیل که بازه شدت $[A, B]$ را برجسته کرده اما سایر شدت ها را حفظ می کند؛ (پ) یک تصویر؛ (ج) حاصل استفاده از تبدیل قسمت (الف).

بخش بندی بیت صفحه‌ای^۲

ممکن است به جای برجسته کردن محدوده های شدت، برجسته کردن سهم بعضی بیت های مشخص در نمایش تصویر مطلوب تر باشد. فرض کنید که هر پیکسل تصویر با ۸ بیت بیان شود. بنابراین تصویر دارای ۸ بیت صفحه شامل صفحه "۰"، مربوط به کم ارزش ترین بیت و صفحه "۷" مربوط به با ارزش ترین بیت هر پیکسل است. شکل ۴.۸، ۹.۴ بیت-صفحه های مربوط به

1) highlight

2) Bit-plane slicing

۲۰.۴ ارتقا با پردازش نقطه ای

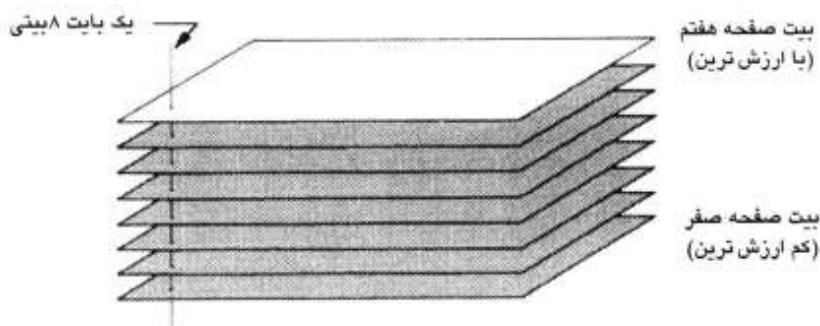
۲۱۵

تصویر موجود در شکل ۴.۵(پ) را نشان می دهد. توجه کنید که تنها ۵ بیت صفحه با ارزشتر حاوی داده های چشمی با اهمیت هستند. سایر بیت صفحه ها در جزئیات ریزتر سهیم هستند. با مقایسه بیت صفحه "۷" در شکل ۴.۹ و شکل ۴.۵(ت) توجه کنید که صفحه "۷" دقیقاً متناظر با تصویر آستانه گیری شده در سطح خاکستری ۱۲۸ می باشد. به عنوان تمرین محدوده های سطوح روشنایی را که توسط سایر بیت صفحه ها پوشش داده می شوند، تعیین کنید.

۲۰.۴.۴ پردازش بافت نگار^۱

هیستوگرام تصویر رقیمی با سطوح خاکستری در محدوده $[0, L-1]$ ، تابع گستته $p(r_k) = n_k/n$ می باشد که r_k امین سطح روشنایی، n_k تعداد پیکسل های تصویر با آن سطح روشنایی، n مجموع پیکسل های تصویر، و $L-1 = 0, 1, 2, \dots, k$ است.

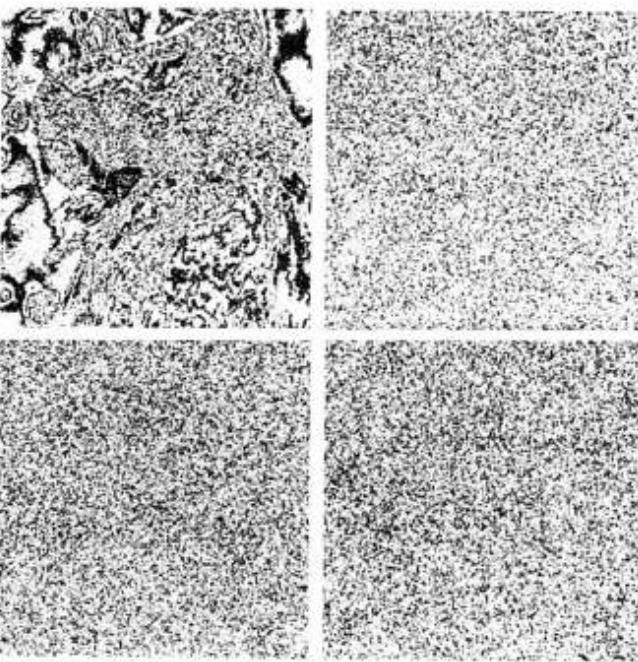
$p(r_k)$ تخمینی از احتمال وقوع سطح خاکستری r_k را می دهد. ترسیم این تابع برای تمام مقادیر k توصیفی کلی در مورد ظاهر تصویر فراهم می آورد. مثلاً شکل ۴.۱۰ بافت نگارهای چهار نوع پایه ای از تصاویر را نشان می دهد. بافت نگار شکل ۴.۱۰(الف) نشان می دهد که سطوح خاکستری به سمت انتهای تیره 'محدوده' خاکستری متتمرکز شده اند. بنابراین این بافت نگار متناظر با تصویری با مشخصات غالباً تیره می باشد. در شکل ۴.۱۰(ب) دقیقاً عکس مطلب فوق صدق می کند. بافت نگار شکل ۴.۱۰(پ) باریک است که نشانه 'محدوده' دینامیکی کوچک می باشد و بنابراین مربوط به تصویری با تمایز پایین



شکل ۴.۱۰ نمایش بیت-صفحه ای تصویر رقیمی ۸ بیتی



7	6
5	4



3	2
1	0

شکل ۴.۳ - صفحه ها برای تصویر شکل ۴. (پ). اعداد درون مریع های کوچک شماره صفحات مختلف هستند. صفحه ۷ حاوی با ارزش ترین بیت ها و صفحه ۰ حاوی کم ارزش ترین بیت های مربوط به پیکسل های تصویر اولیه است.

۲۱۷ ۲.۴ ارتقا با پردازش نقطه ای

است. وقتی تمام سطوح خاکستری در نزدیک وسط محدوده' خاکستری رخ دهند، تصویر تار دیده می شود. در نهایت شکل ۴ . ۱۰ (ت) بافت نگاری با گسترش قابل توجه را نشان می دهد که متناظر تصویری با تمایز بالا می باشد.

گرچه خواصی که اخیراً بیان شد، توصیفاتی کلی هستند که چیز خاصی در مورد محتوای تصویر نمی گویند، شکل بافت نگار تصویر اطلاعات مفیدی در مورد امکان ارتقای تمایز به دست می دهد. توضیحات بعدی روش هایی را جهت تغییر هماهنگ و معنادار هیستوگرامها معرفی می کند.

تعديل بافت نگار

فرض کنید متغیر r نمایانگر آن دسته از سطوح خاکستری تصویر باشد که باید ارتقا بابند. در ابتدا، فرض می کنیم که مقادیر پیکسل ها کمیاتی پیوسته و تراز شده در بازه' $[0, 1]$ هستند، که $r = r$ نمایانگر سیاه و $1 - r$ نمایانگر سفید است. بعد احالت گستره را بررسی کرده و اجازه می دهیم مقادیر پیکسل ها در بازه' $[0, L - 1]$ باشند.

برای r های موجود در بازه' $[0, 1]$ ، توجه را روی تبدیلاتی به شکل

$$s = T(r) \quad (2-2 . ۴)$$

من مرکز می کنیم که این تبدیلات برای هر مقدار پیکسل r در تصویر اولیه، یک سطح s تولید می کنند.

فرض می شود که تابع تبدیل داده شده با معادله' (۲-۲ . ۴) شرایط زیر را برآورده کند:

(الف) $T(r)$ تک مقداره بوده و در بازه' $1 \leq r \leq 0$ اکیداً صعودی است؛ و

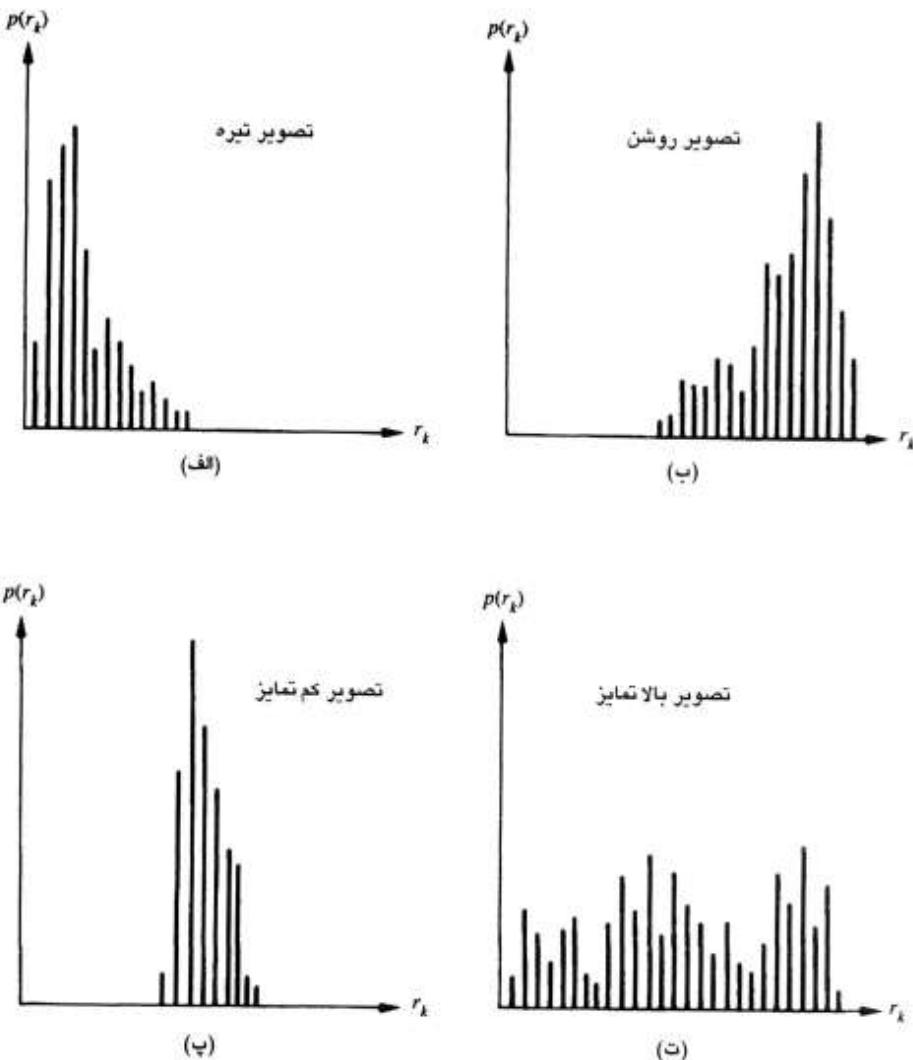
(ب) برای $0 \leq r \leq 1$ داریم $0 \leq T(r) \leq 1$.

شرط (الف) ترتیب سیاه به سفید در محدوده' خاکستری را حفظ می کند، در حالی که شرط (ب) تضمین می کند که تبدیل در محدوده' مجاز مقادیر پیکسل است. شکل ۴ . ۱۱ تابع تبدیلی که این شرایط را برآورده می کند، نمایش می دهد.

تبدیل معکوس از s به r با رابطه'

$$r = T^{-1}(s) \quad 0 \leq s \leq 1 \quad (3-2 . ۴)$$

نشان داده می شود که فرض می شود $(s)^{-1} T$ نیز نسبت به متغیر s شرایط (الف) و (ب) را برآورده می کند.

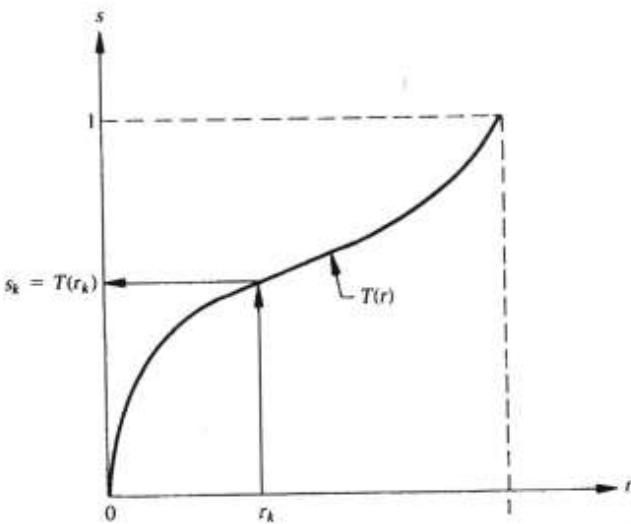


شکل ۴.۱۰. بافت نگارهای متناظر با چهار نوع پایه ای تصویرها

مقادیر خاکستری تصویر را می توان به عنوان کمیت هایی تصادفی درباره '۱، ۰، ۰' تصور کرد. اگر آنها متغیرهای پیوسته ای باشند، سطوح خاکستری اولیه و تبدیل شده را می توان به ترتیب با توابع چگالی احتمال $p_r(r)$ و $p_s(s)$ توصیف کرد، که زیرنویسهای r و s نشان می دهند که توابع p_r و p_s متفاوت هستند.

طبق نظریه احتمال مقدماتی، اگر $p_r(r)$ و $p_s(s)$ معلوم باشند و T شرط (الف) را برآورده کند،

۲۱۹ ارتقا با پردازش نقطه ای ۲.۴



شکل ۱۱۰.۴ یک تابع تبدیل سطح خاکستری

تابع چگالی احتمال سطوح خاکستری تبدیل شده عبارتست از :

$$p_r(s) = \left[p_r(r) \frac{dr}{ds} \right]_{r=T^{-1}(s)} \quad (4-2.4)$$

روش های ارتقای بعدی مبتنی بر بھبود ظاهر تصویر از طریق کنترل تابع چگالی احتمال سطوح خاکستری آن به وسیله تابع تبدیل $T(r)$ می باشند.

تابع تبدیل

$$s = T(r) = \int_0^r p_r(w) dw \quad 0 \leq r \leq 1 \quad (5-2.4)$$

را در نظر بگیرید که w متغیر مجازی انگرایی است. طرف راست معادله '۴-۲.۵' به عنوان تابع توزیع تجمعی (CDF) r شناخته می شود. شرایط (الف) و (ب) که قبلًا بیان شدند، با این تابع تبدیل برآورده می شوند، زیرا CDF به طور یکنواز از ۰ به ۱ به عنوان تابعی از r افزایش می یابد. از معادله '۴-۲.۵'، مشتق s نسبت به r عبارتست از

$$\frac{ds}{dr} = p_r(r). \quad (6-2.4)$$

جایگذاری dr/ds در معادله '۴-۲.۴' نتیجه می دهد که

$$\begin{aligned}
 p_r(s) &= \left[p_r(r) \frac{1}{p_r(r)} \right]_{r=T^{-1}(s)} \\
 &= [1]_{r=T^{-1}(s)} \\
 &= 1 \quad 0 \leq s \leq 1
 \end{aligned} \tag{۷-۲ . ۴}$$

که به صورت یک تابع چگالی یکنواخت در بازه 'تعریف متغیر تبدیل s می باشد. اهمیت این نتیجه به دلیل استقلال آن از تابع تبدیل معکوس می باشد، زیرا مواردی وجود دارد که بدست آوردن $(s)^T$ به صورت تحلیلی آسان نیست.

طبق توضیحات قبلی، استفاده از تابع تبدیلی که برابر با توزیع تجمعی r باشد، تصویری تولید می کند که چگالی سطوح خاکستری آن یکنواخت است. از نظر ارتقا، این نتیجه به معنای افزایش در محدوده دینامیکی پیکسل ها است که می تواند اثر قابل توجهی در ظاهر تصویر داشته باشد.

مثال: قبل از بحث در مورد متغیرهای گستته، نمایشی ساده از کاربرد معادلات (۴-۲ . ۴) و (۵-۲ . ۴) را بررسی می نماییم. فرض کنید که سطوح r دارای تابع چگالی احتمال به صورت شکل ۴ . ۱۲ (الف) باشند. در این مورد $p_r(r)$ عبارتست از:

$$p_r(r) = \begin{cases} -2r + 2 & 0 \leq r \leq 1 \\ 0 & \text{بازی مقادیر } r \end{cases}$$

با جایگذاری عبارت فوق در معادله (۴ . ۲ . ۵) تابع تبدیل زیر بدست می آید:

$$\begin{aligned}
 s &= T(r) = \int_0^r (-2w + 2)dw \\
 &= -r^2 + 2r.
 \end{aligned}$$

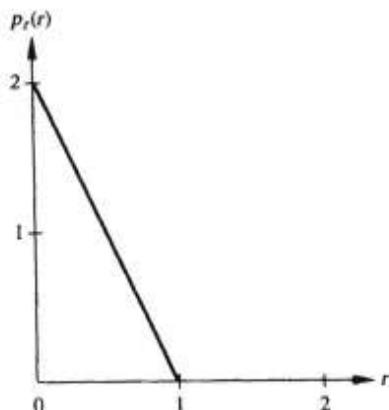
گرچه T کاملاً نیاز تعديل بافت نگار را برآورده می کند، آموزنده خواهد بود که نشان دهیم چگالی $p_s(s)$ بدست آمده یکنواخت است. در عمل این مرحله مورد نیاز نیست، زیرا معادله (۴ . ۷-۲) مستقل از تابع تبدیل معکوس است. با حل r بر حسب s داریم:

$$r = T^{-1}(s) = 1 \pm \sqrt{1-s}$$

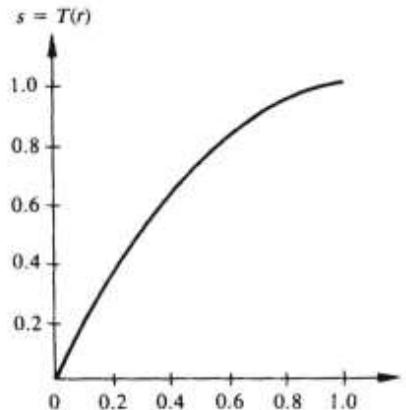
از آنجا که r در بازه $[0, 1]$ قرار دارد، تنها جواب زیر معتبر است:

$$r = T^{-1}(s) = 1 - \sqrt{1-s}$$

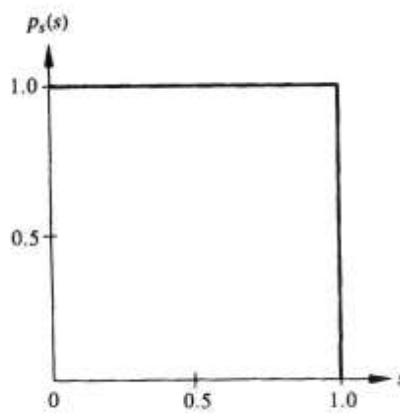
۲۴۱ ارتقا با پردازش نقطه ای



(الف)



(ب)



(پ)

شکل ۱۲۰۴ نمایش روش تبدیل چگالی یکنواخت: (الف) تابع چگالی احتمال اولیه؛ (ب) تابع تبدیل؛ (پ) چگالی یکنواخت حاصل.

تابع چگالی احتمال s با استفاده از معادله ۱۲.۴-۲ بددست می آید:

$$\begin{aligned} p_s(s) &= \left[p_r(r) \frac{dr}{ds} \right]_{r=T^{-1}(s)} \\ &= \left[(-2r + 2) \frac{dr}{ds} \right]_{r=1-\sqrt{1-s}} \\ &= \left[(2\sqrt{1-s}) \frac{d}{ds} (1 - \sqrt{1-s}) \right] \\ &= 1 \quad 0 \leq s \leq 1 \end{aligned}$$

که یک چگالی یکنواخت در بازه مطلوب است. شکل ۱۲.۴ (ب) تابع تبدیل T و شکل ۱۲.۴ (پ)،

□

 $p_r(s_k)$ را نشان می دهد.

به منظور استفاده در پردازش تصویر رقمنی، باید مقاهمی که قبلاً بیان شد را به صورت روابط گسته بیان کرد. برای سطوح خاکستری که مقادیر گسته می گیرند، با احتمالهای زیر سروکار داریم:

$$p_r(r_k) = \frac{n_k}{n} \quad 0 \leq r_k \leq 1 \quad \text{و} \quad k = 0, 1, \dots, L-1 \quad (8-2.4)$$

که همان طور که در ابتدای این بخش بیان شد، L تعداد سطوح، $p_r(r_k)$ احتمال k امین سطح خاکستری، n_k تعداد رخدادهای این سطح در تصویر، و n تعداد کل پیکسل های تصویر است. نمودار $p_r(r_k)$ برحسب r_k ، بافت نگار^۱ نامیده شده و روش مورد استفاده برای بدست آوردن بافت نگار یکنواخت نیز با عنوان تعديل بافت نگار^۲ یا خطی سازی بافت نگار^۳ شناخته می شود.

شكل گسته' معادله' (۸-۲.۵) به صورت زیر بیان می شود:

$$\begin{aligned} s_k &= T(r_k) = \sum_{j=0}^k \frac{n_j}{n} \\ &= \sum_{j=0}^k p_r(r_j) \quad 0 \leq r_k \leq 1 \quad \text{و} \quad k = 0, 1, \dots, L-1. \end{aligned} \quad (9-2.4)$$

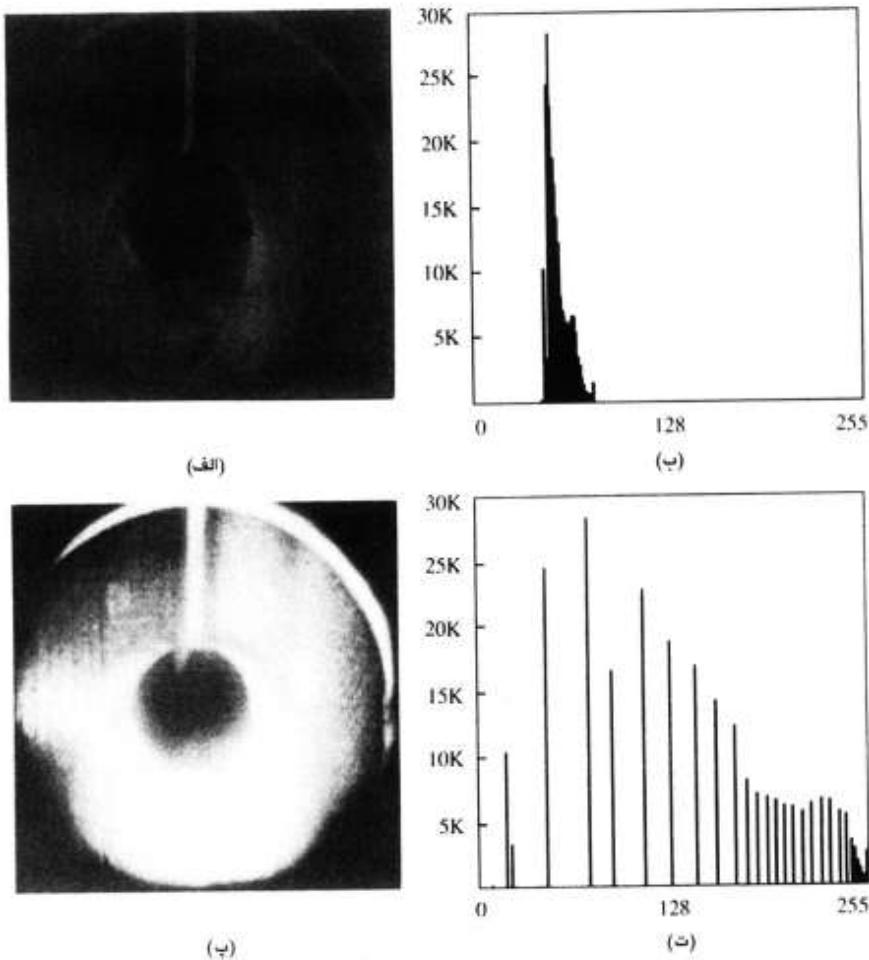
'تبديل معکوس با رابطه'

$$r_k = T^{-1}(s_k) \quad 0 \leq s_k \leq 1$$

نشان داده شده و فرض می شود که هر دو تبدیل $T(r_k)$ و (s_k) شرایط (الف) و (ب) را که قبلاً در این بخش بیان شدند، برآورده می کنند. تابع تبدیل $T(r_k)$ را می توان مستقیماً از تصویر با استفاده از معادله' (۹-۲.۴) محاسبه نمود. گرچه تابع معکوس $T^{-1}(s_k)$ در تعديل بافت نگار استفاده نمی شود، نقشی محوری در روش تشریح شده در بخش بعد بازی می کند.

مثال: برای روشن ساختن سودمندی تعديل بافت نگار، شکل ۱۳.۴ (الف) را که یک تصویر ۸ بیتی ۵۱۲ × ۵۱۲ (از جوشکاری) است و تیره بوده و محدوده' دینامیکی خیلی کوچکی دارد، در نظر بگیرید. براساس توضیح قبلی، بافت نگار این تصویر باید نسبتاً باریک بوده و در انتهای تیره محدوده' خاکستری قرار داشته باشد. این در واقع همان حالتی است که بافت نگار موجود در شکل ۱۳.۴ (ب) دارد. محورافقی

۲۰.۴ ارتقا با پردازش نقطه‌ای ۲۲۳



شکل ۱۳.۰.۴ (الف) تصویر اولیه و (ب) بافت نگار آن؛ (پ) تصویری که تحت عمل تبدیل بافت نگار قرار گرفته است و (ت) بافت نگار آن.

باft نگار محدوده [0, 255] را شامل می شود که بزرگترین محدوده، ممکن از مقادیر سطح خاکستری برای تصویر ۸ بیتی است. همان طور که قبلًا ذکر شد، محور عمودی به جای احتمال، فراوانی پیکسل های دارای هر یک از سطوح خاکستری را نشان می دهد. در عمل شکل ۱۳.۰.۴ (ب) بیشتر استفاده می شود، زیرا طبیعی تر بوده و تعبیر ساده تری دارد. تبدیل محور عمودی به احتمال آسان خواهد بود: کافیست که مقادیر را برابر 2^{12} تقسیم کنیم. به طور مشابه می توان محور افقی را با تقسیم تمام مقادیر بر 255 به محدوده [0, 1] تراز کرد. نکته مهم در مورد همه این تغییر مقیاسها این است که ما بیشتر راجع به تعبیر صحبت می کنیم و شکل و معنی باft نگار تحت تأثیر تغییر مقیاس محورها قرار

نمی گیرد.

شکل ۱۳. ۴ (ب) حاصل تعديل بافت نگار را نشان می دهد. بهبود نسبت به تصویر اولیه کاملاً آشکار است. شکل ۱۳. ۴ (ت) بافت نگار تعديل شده را نشان می دهد. توجه کنید که بافت نگار، تخت نیست- این نتیجه باید باعث تعجب شود زیرا در تقریب گسته مربوط به حاصل پیوسته ای که قبلًا بدست آمد، چیزی در موری این که باید بافت نگار تخت باشد، گفته نشد (به مسئله ۲. ۴ مراجعه کنید). با این حال توجه کنید که سطوح خاکستری تصویری که تحت فرآیند تعديل بافت نگار قرار گرفته است، گستردۀ تر شده و همواره به سفید می رساند. این فرآیند محدوده^{۱۰} دینامیکی سطوح خاکستری و در نتیجه تمایز تصویر را افزایش می دهد. در تصاویری که بافت نگار باریک و سطوح خاکستری نسبتاً کمی دارند، افزایش در محدوده^{۱۱} دینامیکی معمولاً باعث دو اثر مخرب دانه ای بودن^۱ و افزایش غیریکنواختی کیفیت^۲ می شود که هر دو اثر در شکل ۱۳. (ب) آشکار است. نوارهای افقی مربوط به خطای حاصل از رقمی سازی نیز ارتقا یافته اند. به هر حال تعديل بافت نگار به میزان زیادی ظاهر تصویر را بهبود داد. با استفاده از روش کشش تمایز که در بخش قبلی بیان شد، نیز می توان به نتایج مشابهی دست یافت. البته مزیت تعديل بافت نگار نسبت به فنون دستی تغییر تمایز این است که اولی کاملاً خودکار است.

مشخص سازی بافت نگار

گرچه روش تعديل بافت نگار کاملاً مفید است، اما مناسب کاربردهای ارتقای تعاملی تصویر نیست. علت این است که این روش تنها می تواند تقریبی از بافت نگار یکنواخت را تولید کند. گاهی اوقات مطلوب است که بتوانیم شکل های بافت نگار ویژه ای که برای بر جسته کردن محدوده های معینی از سطح خاکستری تصویر مورد نیاز است، تعیین کنیم. برای مشاهده چگونگی انجام این کار، اجازه دهید برای لحظه ای به سطوح خاکستری پیوسته باز گردیم. فرض کنید (r) و (z) به ترتیب توابع چگالی احتمال اولیه و مطلوب باشند. همچنین فرض کنید که تصویر اولیه ابتدا طبق معادله^۳ (۵-۲. ۴) تعديل بافت نگار می شود؛ یعنی،

$$s = T(r) = \int_0^r p_s(w) dw. \quad (10-2. 4)$$

1) Graininess

2) Patchiness

۲.۴ ارتقا با پردازش نقطه ای ۲۲۵

اگر تصویر مطلوب در دست باشد، سطوح آن را نیز می توان با استفاده ازتابع تبدیل زیر تعديل کرد:

$$v = G(z) = \int_0^z p_s(w) dw. \quad (11-2. 4)$$

آنگاه فرآیند معکوس، $v = G^{-1}(z)$ ، سطوح تصویر مطلوب را می دهد. البته این مطلب فرضی است، زیرا سطوح z دقیقاً همان مطلوب هستند. با این حال، $p_s(s)$ و $p_v(v)$ چگالی های یکنواخت یکسان خواهند بود، زیرا نتیجه نهایی معادله '۴. ۵-۲' مستقل از چگالی درون انتگرال است. بنابراین اگر در فرآیند معکوس اگر به جای v از سطوح یکنواخت s حاصل از تصویر اولیه استفاده کنیم، سطوح حاصل، $v = G^{-1}(s)$ ، تابع چگالی احتمال مطلوب را خواهند داشت. با فرض این که $G^{-1}(s)$ تک مقداره باشد، می توان این روال را به صورت زیر جمع بندی کرد:

(۱) سطوح تصویر اولیه را با استفاده از معادله '۴. ۵-۲' تعديل کنید.

(۲) تابع چگالی مطلوب را تعیین کرده و با استفاده از معادله '۴. ۱۱-۲' تابع تبدیل $G(z)$ را بدست آورید.

(۳) تابع تبدیل معکوس، $v = G^{-1}(z)$ ، را بر سطوح حاصل از مرحله '۱' اعمال کنید.

این روال یک نسخه پردازش شده از تصویر اولیه را می دهد که سطوح خاکستری جدید با چگالی معین $p_z(z)$ توصیف می شوند.

گرچه روش مشخص سازی بافت نگار از دو تابع تبدیل $T(r)$ و به دنبال آن $G^{-1}(s)$ استفاده می کند، ترکیب آن دوتابع در یک تابع تبدیل که مستقیماً از پیکسل های اولیه، پیکسل های مطلوب را تولید کند، کار ساده ای است. طبق توضیحات قبلی داریم:

$$z = G^{-1}(s). \quad (12-2. 4)$$

با جایگذاری معادله '۴. ۵-۲' در معادله '۴. ۲-۵' فوق، تابع تبدیل مرکب

$$z = G^{-1}[T(r)] \quad (13-2. 4)$$

بدست می آید که r را به z مربوط می کند. وقتی $[T(r)] = G^{-1}[T(r)]$ باشد، معادله '۴. ۲-۱۳' همان معادله 'تعديل بافت نگار' است.

ضمیمن معادله^۱ (۴. ۲-۱۳) این است که نیازی نیست که به طور صریح تصویر را تعديل باف نگار کرد. تنها کافی است که $T(r)$ تعیین شود و با تابع تبدیل معکوس^۲ G^{-1} ترکیب گردد. مشکل استفاده از روش قبلی در مورد متغیرهای پیوسته بدست آوردن تحلیلی تابع معکوس است. در حالت گسته این مشکل با کمک این واقعیت بر طرف می شود که معمولاً تعداد سطوح خاکستری نسبتاً کم بوده و بنابراین محاسبه و ذخیره نگاشت برای هر مقدار ممکن پکسل، عملی است. در مورد متغیرهای گسته، روش مشخص سازی بافت نگار طبق معادلات (۴. ۸-۲) و (۴. ۹-۲) مشابه سازی می گردد.

در عمل، در اغلب موارد تبدیل معکوس از S به \mathbb{R}^2 تک مقداره نیست. این مشکل وقتی رخ می دهد که در بافت نگار مشخص شده سطوح پرنشده وجود داشته باشد (که باعث می شود CDF در بازه های پرنشده ثابت بماند) و یا این که در فرآیند گردکردن (G^{-1}) به نزدیکترین سطح خاکستری مجاز رخ می دهد. معمولاً ساده ترین راه حل این است که سطوح خاکستری را به گونه ای منتب کنیم که بهتر با بافت نگار مورد نظر تطبیق نماید.

مشکل اصلی در به کار گیری روش مشخص سازی بافت نگار جهت ارتقای تصویر، به توانایی ساخت بافت نگار معنادار برمی گردد. دو راه حل برای این مشکل وجود دارد: راه حل اول انتخاب یک تابع چگالی احتمال خاص (نظری چگالی گوسی) و سپس تشکیل بافت نگار با رقمی سازی تابع فوق می باشد. راه حل دوم شامل مشخص کردن شکل بافت نگار با یک وسیله 'گرافیکی (مثلاً صفحه' تعاملی^۳ یا تابلوی طراحی^۴) می باشد که خروجی آن به عنوان ورودی پردازشگری که الگوریتم مشخص سازی بافت نگار را اجرا می کند، مورد استفاده قرار می گیرد.

مثال: شکل ۴. (الف) را که منظر یک اتاق نیمه تاریک از سمت در می باشد، در نظر بگیرید. شکل ۴. (ب) تصویر را پس از تعديل بافت نگار نشان می دهد و شکل ۴. (پ) حاصل مشخص سازی تعاملی بافت نگار است. در شکل ۴. (ت) به ترتیب از پایین به بالا بافت نگارهای اولیه، تعديل شده، مشخص شده و نهایی دیده می شود.

دقت کنید که تعديل بافت نگار تصویری تولید کرد که تمایز تقریباً بالایی داشت، در حالی که نتیجه موجود در شکل ۴. (پ) ظاهر بسیار متوزن تری دارد. به علت انعطاف پذیری موجود در روش

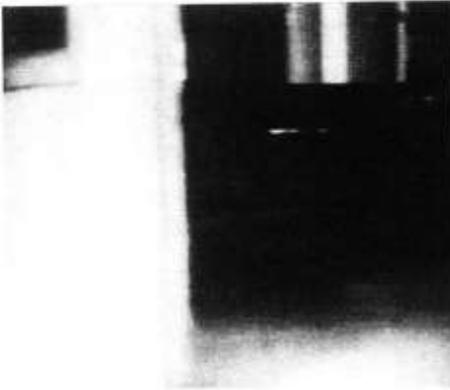
1) Interactive screen

2) Drawing tablet

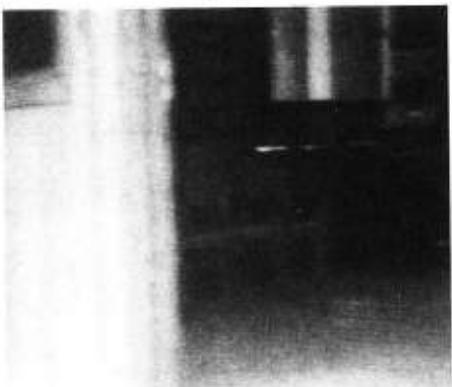
۲۰.۴ ارتقا با پردازش نقطه ای ۲۲۷



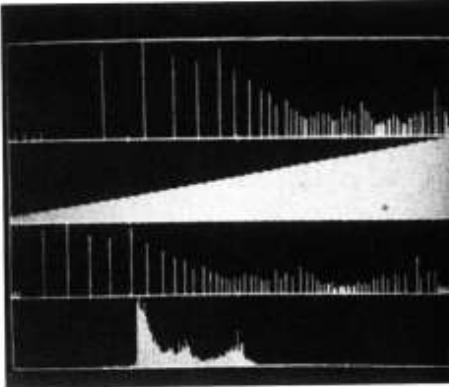
(الف)



(ب)



(پ)



(ت)

شکل ۱۰.۴ نمایش روش مشخص سازی بافت نگار؛ (الف) تصویر اولیه؛ (ب) تصویر پس از تغییر بافت نگار؛ (پ) تصویر ارتقا یافته به وسیله مشخص سازی بافت نگار؛ (ت) بافت نگارها

مشخص سازی بافت نگار، معمولاً این روش نتایجی بهتر از تغییر هیستوگرام دارد.

ارتقای محلی

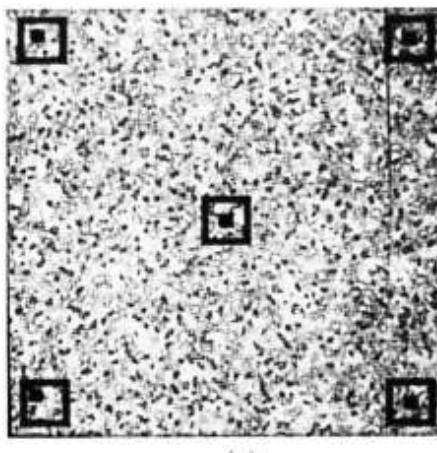
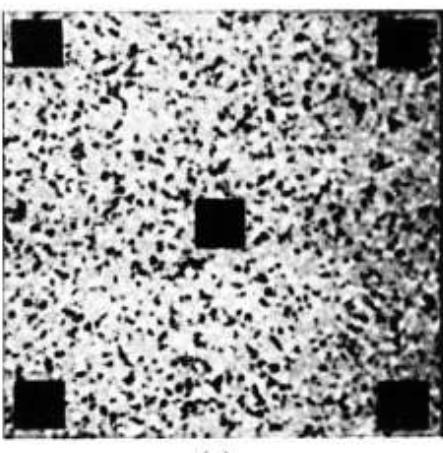
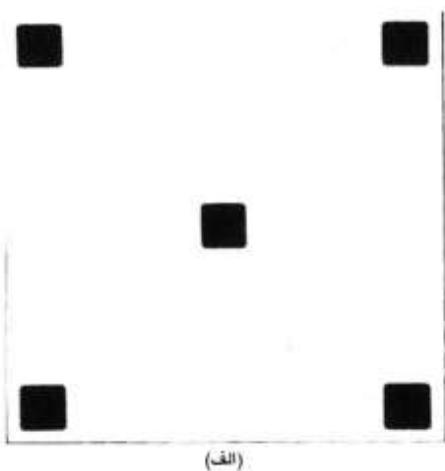
دو روش پردازش بافت نگار تشریح شده در دو بخش قبلی سراسری^۱ هستند، از این نظر که پیکسل ها بوسیله تابع تبدیلی که براساس توزیع سطح خاکستری کل تصویر است، تغییر داده می شوند. گرچه این روش های سراسری برای ارتقای کل تصویر مناسب هستند، در اغلب موارد لازم است که جزئیات درون نواحی کوچکی را ارتقا دهیم. تعداد پیکسل های درون این نواحی ممکن است اثر ناچیزی بر محاسبه تبدیل

سراسری داشته باشد، بنابر این استفاده از تبدیل های سراسری ارتقای محلی مطلوب را تضمین نمی کند. راه حل این است که توابع تبدیلی براساس توزیع سطح خاکستری تصویر - یا دیگر خواص - در همسایگی هر پیکسل ابداع نماییم. اگرچه روش های پردازش مبتنی بر همسایگیها موضوع بخش ۳. هستند، برای روشن شدن موضوع و حفظ پیوستگی، در این بخش پردازش بافت نگار محلی را شرح می دهیم.

روش های پردازش بافت نگار که قبلاً بیان شدند، با کمی تغییر برای ارتقای محلی مناسب هستند. روال این است که یک همسایگی مربعی یا مستطیلی تعریف کرده و مرکز آن را پیکسل به پیکسل حرکت دهیم. در هر پیکسل، بافت نگار نقاط درون همسایگی را تشکیل داده و تابع تبدیل تعديل بافت نگار یا تابع تبدیل مشخص سازی بافت نگار را محاسبه می کنیم. این تابع در نهایت برای نگاشت سطح خاکستری پیکسل مرکز همسایگی استفاده می شود. در مرحله بعد، مرکز ناحیه همسایگی به یک پیکسل مجاور جابجا شده و این عمل تکرار می گردد. از آنجایی که در طول انتقال پیکسل به پیکسل تنها یک سطر یا یک ستون از همسایگی تغییر می کند، می توان بافت نگار بدست آمده در محل قبل را با داده های جدیدی که در هر حرکت وارد می شوند، بهنگام کرد (مسئله ۴.۵ را ببینید). این روش نسبت به روش محاسبه 'کامل بافت نگار پیکسل های همسایگی در هر حرکت یک پیکسلی مزیت آشکاری دارد. روش دیگری که اغلب برای کاهش محاسبات استفاده می شود، به کارگیری همسایگی های بدون همپوشانی می باشد، اما این روش معمولاً اثر شترنجی نامطلوبی را تولید می کند.

مثال: شکل ۴.۱۵.۱ تبدیل بافت نگار محلی را با یک همسایگی که پیکسل به پیکسل جابجا می شود، نشان می دهد. شکل ۴.۱۵.۲ (الف) تصویری را نشان می دهد که به منظور کاهش نویز، کمی مات شده است (بخش ۴.۳.۲ را ببینید). شکل ۴.۱۵.۲ (ب) حاصل تبدیل بافت نگار سراسری را نشان می دهد. همان طور که در اغلب موارد دیده می شود، وقتی این روش بر نواحی آرام نویزدار اعمال شود، شکل ۴.۱۵.۲ (ب) ارتقای قابل ملاحظه نویز و کمی افزایش در تمایز را نشان می دهد. با این حال توجه کنید که با این روش هیچ جزئیات ساختاری جدیدی آشکار نشد. از طرف دیگر، تبدیل بافت نگار محلی با استفاده از همسایگی ۷×۷، وجود مربع هایی کوچک را درون مربع های تاریک بزرگتر آشکار کرد. نزدیکی سطوح خاکستری و ابعاد کوچک مربعها اثر قابل ملاحظه ای بر تبدیل بافت نگار سراسری ندارند. به نویز زمینه ریزتر در مشکل ۴.۱۵ (ب)، که نتیجه پردازش محلی در همسایگی های نسبتاً کوچک است، توجه کنید. □

۲۰۴ ارتقا با پردازش نقطه ای ۴۴۹



شکل ۱۵.۴ (الف) تصویر اولیه؛ (ب) حاصل تبدیل بافت نگار سراسری؛ (پ) حاصل تبدیل بافت نگار محلی با استفاده از همسایگی 7×7 حول هر پیکسل. (از Lee [1987] و Fu, Gonzalez [1987].)

به جای استفاده از بافت نگارها، ارتقای محلی را می‌توان براساس سایر خواص همسایگی انجام داد. به دلیل ارتباط میانگین و واریانس (یا انحراف معیار) شدت با ظاهر تصویر، این دو ویژگی به طور مکرر استفاده می‌شوند. در واقع میانگین معیاری از متوسط روشنایی و واریانس معیاری از تمایز است.

یک تبدیل محلی نوعی که مبتنی بر این مفاهیم باشد، با اجرای تبدیل زیر، تصویر ورودی $f(x, y)$ را به تصویر جدید $g(x, y)$ می‌نگارد:

$$g(x, y) = A(x, y) \cdot [f(x, y) - m(x, y)] + m(x, y) \quad (14-2. 4)$$

که

$$A(x, y) = k \frac{M}{\sigma(x, y)} \quad 0 < k < 1. \quad (15-2. 4)$$

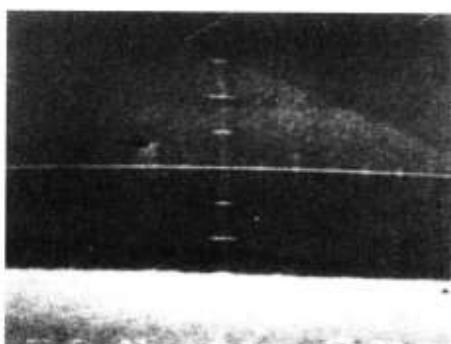
در این رابطه $m(x, y)$ و $\sigma(x, y)$ به ترتیب میانگین و انحراف معیار سطح خاکستری هستند که در یک همسایگی به مرکز (x, y) محاسبه شده اند، M میانگین سراسری $f(x, y)$ و k ثابتی در بازه 'ذکر شده در معادله' (۱۵-۲. ۴) است.

مقادیر متغیرهای A و m به همسایگی تعریف شده حول (x, y) بستگی دارند. اعمال ضرب بهره ' محلی $A(x, y)$ بر تفاصل بین (x, y) و میانگین محلی، تغییرات محلی را تقویت می کند. چون (x, y) با انحراف معیار شدت، تناسب معکوس دارد، نواحی با تمایز پایین با بهره 'بزرگتری تقویت می شوند. حضور دوباره 'میانگین در معادله' (۱۴-۲. ۴) به صورت جمع، مقدار متوسط سطح شدت در ناحیه ' محلی تصویر را به آن بازمی گرداند. در اغلب موارد عملی، جمع مجدد بخشی از میانگین محلی و همچنین محدود کردن تغییرات (y) بین دو حد (A_{\min}, A_{\max}) ، به منظور متوازن کردن تغییرات بزرگ شدت در نواحی مجزا از هم، مطلوب است.

مثال: در شکل ۴. ۱۶ مثالی از قابلیت معادلات (۱۴-۲. ۴) و (۱۵-۲. ۴) با استفاده از یک ناحیه ' محلی 15×15 پیکسلی دیده می شود. به ارتقای جزئیات در مرز بین دو ناحیه با سطوح خاکستری متفاوت و همچنین جزئیات سطح خاکستری هر ناحیه توجه کنید.



(الف)



(ب)

شکل ۴. ۱۶ تصاویر قبل و بعد از ارتقای محلی. (از Fitch[1981] و Narendra)

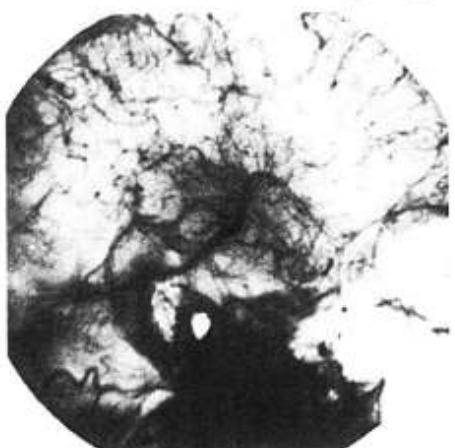
۳.۲.۴ تفريقي تصوير

تفاضل دو تصوير y و $f(x, y)$ که به صورت

$$g(x, y) = f(x, y) - h(x, y) \quad (16-2.4)$$

بيان می شود، با محاسبه 'تفاضل' بین هر يك از زوج پيكسل هاي متناظر در f و h بدست می آيد. تفريقي تصوير کاربردهای مهم متعددی در بخش بندی (فصل ۷) و ارتقا دارد.

يکي از کاربردهای کلاسيک معادله (۱۶-۲.۴) در مسأله 'ارتقا'، مبحثي از تصويربرداری پژشكى به نام پرتونگاري حالت نقاب^۱ می باشد. در اين مورد نقاب y خود، تصوير اشعه اينكس قسمتی از بدن بيمار است که با قرارگرفتن يك تشديدكشنه^۲ و يك دوربين تلویزیونی (به جای فilm اشعه اينكس معمولی) در مقابل منبع مولد اشعه اينكس، برداشته شده است. تصوير y نمونه ای از مجموعه تصاوير تلویزیونی مشابه از يك ناحيه بدن انسان است که پس از تزریق ماده رنگی به شريان آن ناحيه برداشته شده است. تفريقي نقاب از هر نمونه در دنباله 'تصاوير تلویزیونی ورودی'، باعث می شود که تنها جزئيات آن دسته از نواحي که بين y و $f(x, y)$ متفاوت هستند، در تصوير خروجی به صورت جزئيات ارتقا يافته دیده شوند. چون در يافت تصاوير در نرخهای تلویزیونی امكان پذير است، اين روش تصوير متحرکی از چگونگي انتشار ماده رنگی در رگها، توليد می کند.



(الف)



(ب)

شكل ۱۷.۴ ارتقا به وسیله تفريقي تصوير: (الف) تصوير نقاب؛ (ب) تصوير (پس از تزریق ماده رنگی درون شريان) که تصوير نقاب از آن تفريقي شده است.

مثال: شکل ۴.۱۷ (الف) یک تصویر اشعه ایکس از بالای سر بیمار است که قبل از تزریق ماده رنگی بد^۱ به شریان برداشته شده است. دوربینی که این تصویر را تولید کرده است، بالای سر بیمار و رو به پایین جاسازی شده است. به عنوان مرجع مقایسه، نقطه روشن در یک سوم پایین تصویر، هسته ستون فقرات بیمار است. شکل ۴.۱۷ (ب) تفاضل بین نتاب (شکل ۴.۱۷ الف) و تصویری که مدت کوتاهی پس از ورود ماده رنگی به شریان گرفته شده است، را نشان می دهد. رگ های روشن که ماده رنگی را حمل می کنند، به صورت کاملاً بارزی در شکل ۴.۱۷ (ب) ارتفا یافته اند. این رگ ها کاملاً روشن ظاهر می شوند، زیرا تغیر نمی شوند (چرا که بخشی از تصویر اولیه نیستند). زمینه در کل بسیار تیره تر از شکل ۴.۱۷ (الف) است، زیرا حاصل تفاضل در نواحی کم تغییر، مقادیر کوچکی را می دهد، که این مقادیر به صورت سایه های خاکستری تیره در تصویر تفاضل ظاهر می شوند. به عنوان مثال توجه کنید که همان نخاع که در شکل ۴.۱۷ (الف) روشن است، در شکل ۴.۱۷ (ب) بر اثر تغیرنامه کاملاً تیره به نظر می رسد.



۴.۲.۴ متوسطگیری تصویر

تصویر نویزدار $(x, y)g(x, y)$ را در نظر بگیرید که از جمع نویز $(x, y)\eta(x, y)$ با تصویر اولیه $(x, y)f(x, y)$ تشکیل شده است؛ یعنی،

$$g(x, y) = f(x, y) + \eta(x, y) \quad (17-2.4)$$

فرض شود که در هر مختصات (x, y) ، نویز ناهمبسته و مقدار متوسط آن صفر است. هدف روش زیر کاهش اثرات نویز با جمع مجموعه تصاویر نویزی $\{g_i(x, y)\}$ است.

اگر نویز این شرایط را برآورده نماید، به سادگی می توان نشان داد (به Papoulis [1965] مراجعه

کنید) که اگر تصویر $(x, y)\bar{g}$ با متوسط گیری از M تصویر نویزدار بدست آید،

$$\bar{g}(x, y) = \frac{1}{M} \sum_{i=1}^M g_i(x, y) \quad (18-2.4)$$

آنگاه نتیجه می شود که

$$E\{\bar{g}(x, y)\} = f(x, y) \quad (19-2.4)$$

۳.۴ فیلتر کردن مکانی ۲۳۳

$$\sigma_{\bar{g}(x,y)}^2 = \frac{1}{M} \sigma_{g(x,y)}^2 \quad (20-2.4)$$

که $\{E\{\bar{g}(x,y)\}$ مقدار امید \bar{g} است، و $\sigma_{\bar{g}(x,y)}^2$ به ترتیب واریانسها \bar{g} و η در مختصات (x, y) هستند. انحراف معیار در هر نقطه از تصویر متوسط عبارتست از

$$\sigma_{\bar{g}(x,y)} = \frac{1}{\sqrt{M}} \sigma_{g(x,y)} \quad (21-2.4)$$

معادلات (۲۰-۲.۴) و (۲۱-۲.۴) نشان می دهند که با افزایش M میزان تغییرات^۱ مقادیر پیکسل کاهش می یابد. چون $E\{\bar{g}(x,y)\} = f(x,y)$ ، این شرط بدین معناست که وقتی تعداد تصاویر نویزدار مورد استفاده در فرآیند متوسط گیری افزایش یابد، $f(x,y)$ به $g(x,y)$ نزدیک می شود. در عمل به منظور دوری از مات شدن تصویر خروجی، باید تصاویر (x,y) ضبط شوند.

مثال: شکل ۴.۱۸ (الف) تصویری میکروسکوپی نویزدار یک سلول را نشان می دهد و اشکال ۴.۱۸ (ب)-(ج) به ترتیب حاصل متوسط گیری از ۲، ۸، ۱۶، ۳۲ و ۱۲۸ عدد از چنین تصاویر نویزداری را نشان می دهند. از دیدگاه تحلیل چشمی، تصویر حاصل از $M=32$ ، به صورت قابل قبولی تمیز است و تصویر بدست آمده با $M=128$ اصولاً برای هر کاربردی بدون نویز تلقی می شود.
□

۳.۴ فیلتر کردن مکانی

۱.۳.۴ زمینه

استفاده از ناقاب های مکانی برای پردازش تصویر معمولاً فیلتر کردن مکانی^۲ (در مقابل فیلتر کردن میدان فرکانس^۳ با استفاده از تبدیل فوریه) نامیده می شود و خود ناقاب ها فیلتر های مکانی^۴ خوانده می شوند. در این بخش فیلتر های مکانی خطی و غیرخطی ارتقای تصویر را بررسی می کنیم. فیلتر های خطی بر اساس مفاهیم معرفی شده در بخش ۱.۱ هستند. طبق این مفاهیم،تابع انتقال و تابع ضربه (یا تابع پخش نقطه ای) هر سامانه خطی، تبدیل فوریه معکوس یکدیگر هستند.

فیلتر های معروف به پایین گذر^۵، مؤلفه های فرکانس بالا در میدان فوریه را تضعیف یا حذف می کنند

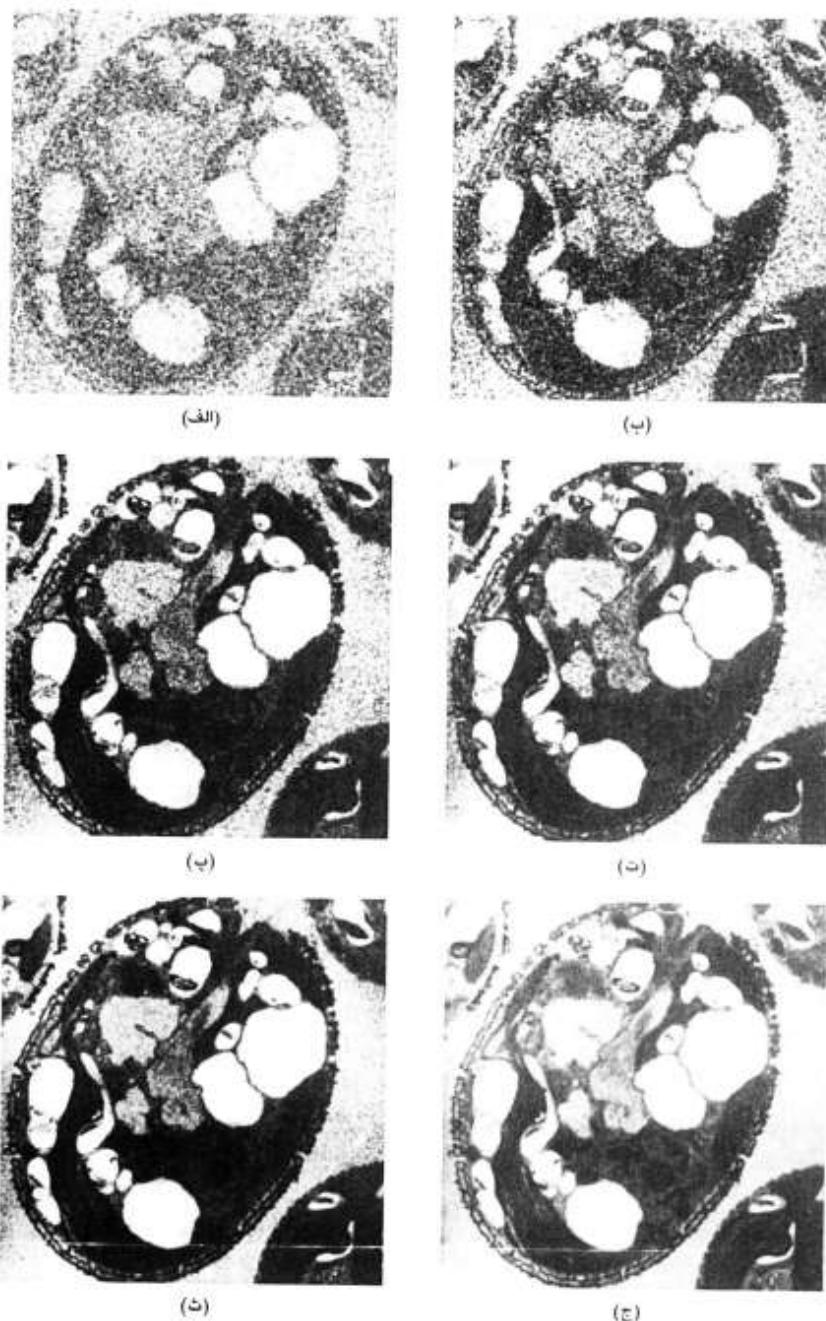
1) Variability

2) Spatial filtering

3) Frequency domain filtering

4) Spatial filter

5) Lowpass



شکل ۱۸.۴ مثالی از کاهش نویز بوسیله متوسطگیری: (الف) یک تصویر نویز دار نوعی؛ (ب)-(ج) حاصل متوسطگیری از ۲، ۸، ۱۶، ۳۲، ۱۲۸ تصویر نویز دار.

۳.۴ فیلتر کردن مکانی

۲۳۵

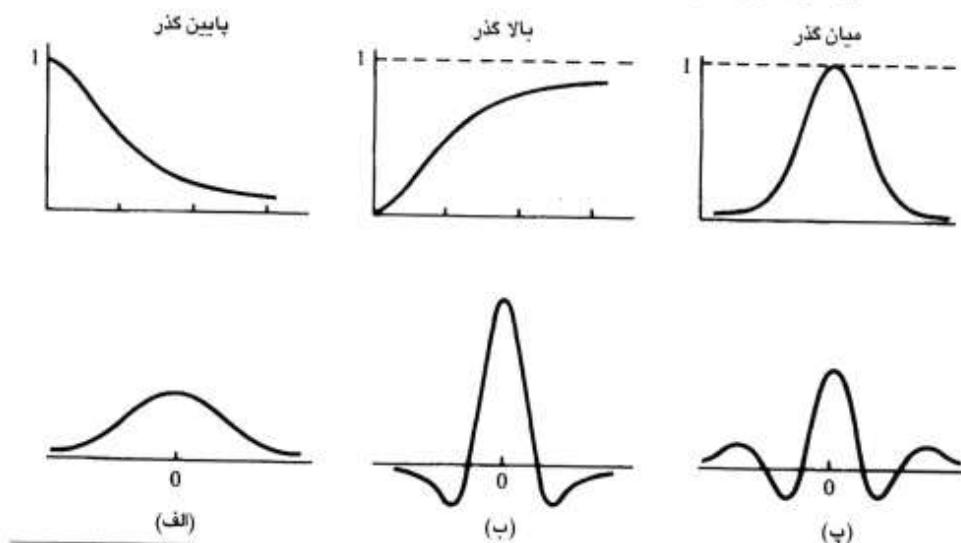
در حالی که فرکانس های پایین را دست نخورد باقی می گذارند (یعنی فرکانس های پایین را "عبور می دهند"). مؤلفه های فرکانس بالا، لبه ها و سایر جزئیات تیز را توصیف می کنند، بنابراین اثر خالص فیلتر کردن گذر مات کردن تصویر است. به طور مشابه، فیلتر های بالا گذر^۱ مؤلفه های فرکانس پایین را تضعیف یا حذف می کنند. چون این مؤلفه ها مربوط به مشخصات با تغییر آرام تصویر، نظیر تمایز کلی و شدت متوسط، هستند، نتیجه خالص فیلتر کردن بالا گذر کاهش این ویژگی ها و در نتیجه تیز کردن لبه ها و سایر جزئیات تیز می باشد. نوع سوم فیلتر کردن که فیلتر کردن میان گذر^۲ نامیده می شود، نواحی فرکانسی انتخاب شده ای بین فرکانس های پایین و بالا را حذف می کند. این فیلتر ها برای بازیابی تصویر (بخش ۵.۸) بکار می روند و بnderت در ارتقای تصویر مورد استفاده واقع می شوند.

شکل ۴.۱۹. برش های عرضی فیلتر های پایین گذر، بالا گذر و میان گذر متقاضان چرخشی در میدان فرکانس و فیلتر های مکانی متناظر شان را نشان می دهد. محور های افقی در اشکال ردیف بالا متناظر با فرکانس و در اشکال ردیف پایین متناظر با مشخصات مکانی هستند. اشکال ردیف پایین به عنوان خطوط راهنمای مشخص کردن فیلتر های مکانی خطی بکار می روند. در هر حال بدون توجه به نوع فیلتر خطی مورد استفاده، روش اصلی پیاده سازی آنها جمع حاصل ضرب های ضرایب نقاط با شدت پیکسل های زیر نقاط در هر محل تصویر می باشد. شکل ۴.۲۰. یک نقاط عمومی 3×3 را نشان می دهد. اگر سطوح خاکستری پیکسل های زیر نقاط را با $z_1, z_2, z_3, \dots, z_n$ نشان دهیم، پاسخ نقاط خطی عبارتست از :

$$R = w_1 z_1 + w_2 z_2 + \dots + w_n z_n \quad (4-3)$$

با مراجعه به شکل ۴.۱، اگر مرکز نقاط نقطه (y, x) باشد، سطح خاکستری پیکسلی که در (y, x) قرار دارد با R جایگزین می شود. آنگاه مرکز نقاط به پیکسل بعدی تصویر جا بجا شده و این فرآیند تکرار می گردد تا این که تمام پیکسل های تصویر مرور شوند. برای پیکسل هایی که در مرز تصویر قرار دارند، مقدار R با استفاده از همسایگی های جزئی محاسبه می شود. معمولاً تلاش می شود به جای جایگزین کردن مقادیر پیکسل در تصویر اولیه، تصویر جدیدی برای ذخیره مقادیر R ایجاد گردد. این عمل باعث می شود چنانچه قبل از معادله (۴.۳-۱) استفاده شده باشد، باز هم امکان دسترسی به

۴۳۶ پردازش تصویر رقمنی



شکل ۴.۱۹. ردیف بالا: مقاطع عرضی اشکال پایه ای مربوط به توابع میدان فرکانس متقارن چرخشی. ردیف پایین: مقاطع عرضی فیلترهای میدان مکان منتظر

مقادیر سطح خاکستری اولیه وجود داشته باشد.

فیلترهای مکانی غیرخطی نیز روی همسایگی‌ها کار می‌کنند. با این حال در حالت کلی این فیلترها مستقیماً با مقادیر پیکسل‌ها در همسایگی موردنظر کار می‌کنند و به صورت معادله^۱ (۴.۳-۱)، از ضرایب استفاده نمی‌کنند. همان‌طور که در بخش بعدی نشان داده می‌شود، با یک فیلتر غیرخطی که میانه^۲ مقادیر سطح خاکستری در همسایگی محل فیلتر را محاسبه کند، می‌توان به طور مؤثری نویز را کاهش داد. فیلتر بیشینه^۱ (با پاسخ $R = \max\{z_k | k = 1, 2, \dots, 9\}$) که برای پیدا کردن روش‌ترین نقاط همسایگی و همچنین فیلتر کمینه^۲ که به منظور مخالف استفاده می‌شود، مثال‌های دیگری از فیلترهای غیرخطی هستند.

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

شکل ۴.۲۰. یک تابع 3×3 با ضرایب (وزنهای) دلخواه

1) Max filter

2) Min filter

۲.۳.۴ فیلترهای آرام کننده^۱

فیلترهای آرام کننده برای مات کردن و کاهش نویز بکار می روند. از مات کردن در مراحل پیش پردازش، نظری حذف جزئیات کوچک تصویر قبل از استخراج شیبی (بزرگ) و متصل کردن فواصل کوچک درون خطوط و منحنی ها استفاده می شود. کاهش نویز را می توان با مات کردن به وسیله یک فیلتر خطی یا غیرخطی انجام داد.

فیلتر کردن مکانی پایین گذر

شکل پاسخ ضربه^۲ مورد نیاز برای پیاده سازی یک فیلتر مکانی پایین گذر (آرام کننده) نشان می دهد که باید تمام ضرایب فیلتر مثبت باشند (شکل ۴.۱۹.الف را بینید). گرچه شکل فیلتر مکانی نشان داده شده در شکل ۴.۱۹.(الف) را می توان مثلاً با یک تابع گوسی نمونه برداری شده مدل کرد، اما بدون تردید باید تمام ضرایب مثبت باشند. برای یک فیلتر مکانی 3×3 ساده ترین آرایش، نقاب خواهد بود که تمام ضرایب آن 1^* باشند. با این حال طبق معادله (۴.۳-۱)، پاسخ برابر جمع سطوح خاکستری ۹ پیکسل خواهد بود، که ممکن است باعث شود R خارج از محدوده^۳ معتبر سطح خاکستری قرار گیرد. راه حل مشکل، مقیاس کردن مجموع با تقسیم R بر ۹ می باشد. شکل ۴.۲۱.(الف) نقاب حاصل را نشان می دهد. همان طور که اشکال ۴.۲۱.(ب) و (پ) نشان می دهند، نقاب های بزرگتر نیز از همان مفهوم پیروی می کنند. توجه کنید که در تمام موارد، پاسخ R برابر با متوسط پیکسل های درون ناحیه^۴ نقاب خواهد بود. به همین دلیل کاربرد نقاب هایی که به صورت شکل ۴.۲۱. هستند، اغلب با عنوان متوسط گیری همسایگی^۵ اشاره می شود. شکل ۴.۲۲. مثالی از مات کردن به وسیله نقاب های آرام کننده با ابعاد بزرگتر را نشان می دهد. به ویژه به از دست رفتن تیزی در فیلامنت لامپ وقتی که نقاب آرام کننده بزرگتر می شود، دقت نمایید.

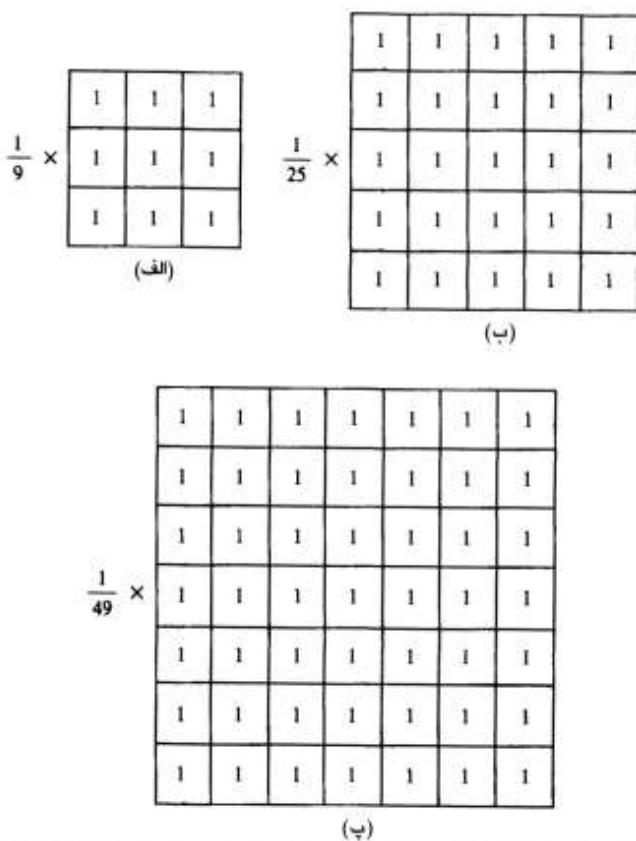
فیلتر کردن عیانه ای

یکی از مشکلات اصلی روش آرام کردن که در بخش قبل بیان شد، این است که لبه ها و سایر جزئیات تیز را مات می کند. اگر هدف کاهش نویز باشد نه مات کردن، یک راهکار استفاده از فیلترهای میانه ای^۶ است.

1) Smoothing filter

2) Neighborhood averaging

3) Median filter



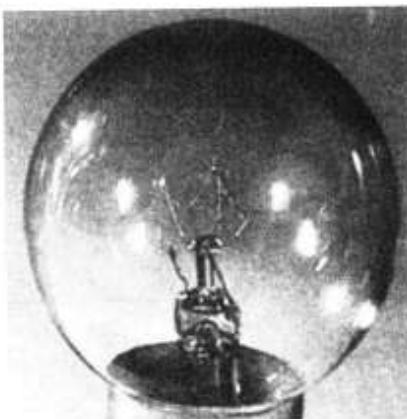
شکل ۲۱.۴ فیلترهای پایین گذار مکانی با ابعاد مختلف

در این روش، سطح خاکستری هر پیکسل به جای متوسط، با میانه سطوح خاکستری در همسایگی آن پیکسل جایگزین می شود. این روش بویژه وقتی الگوی نویز شامل مؤلفه های ضربه ای قوی بوده و می باید تیزی لبه حفظ شود، مؤثر می باشد. همان طور که قبل بیان شد، فیلترهای میانه ای غیرخطی هستند. تیزی مجموعه ای از مقادیر چنان است که نیمی از مقادیر مجموعه کوچکتر از m و نیمی بزرگتر از m میانه m مجموعه ای از مقادیر چنان است که نیمی از مقادیر همسایگی یک پیکسل، ابتدا مقادیر پیکسل های همسایگی را باشند. برای اجرای فیلتر کردن میانه ای در همسایگی یک پیکسل، ابتدا مقادیر پیکسل های همسایگی را مرتب کرده، میانه را تعیین نموده و آن را به پیکسل مرکز همسایگی نسبت می دهیم. مثلاً در همسایگی 3×3 ، میانه پنجمین مقدار بزرگتر است، در همسایگی 5×5 میانه سیزدهمین مقدار بزرگتر است و الی آخر. وقتی چند مقدار در یک همسایگی برابر باشند، تمام آن مقادیر باید در یک گروه قرار گیرند. مثلاً فرض کنید که یک همسایگی 3×3 مقادیر $3, 10, 20, 25, 20, 20, 15, 20, 10$ را داشته

۳۰.۳ فیلتر کردن مکانی



(الف)



(ب)



(ب)



(ت)



(ت)



(ج)

شکل ۲۲۰.۴ (الف) تصویر اولیه؛ (ب)- (ج) نتایج فیلتر کردن پایین گذر مکانی با مقابس با ابعاد $n \times n$ ، که $n = 3, 5, 7, 15, 25$

باشد. این مقادیر به صورت دنباله^{۱)} $(100, 10, 15, 20, 20, 20, 20, 25, 20, 20, 20, 20, 20, 20)$ مرتب می شوند، که مقدار میانه 20 را نتیجه می دهد. بنابراین نتیجه اصلی فیلتر میانه ای این است که نقاط با شدت متفاوت را بیشتر به همسایگانشان شبیه می نماید. این کار ضربه های شدتی را که در ناحیه 'ماسک فیلتر' و به صورت تک تک رخ می دهند، حذف می نماید.

مثال: شکل ۴.۲۳ (الف) تصویر اولیه، شکل ۴.۲۳.۲ (ب) همان تصویر که تقریباً 20 درصد پیکسل های آن توسط نویز ضربه ای خراب شده اند، شکل ۴.۲۳.۲ (پ) حاصل متوسط گیری همسایگی در ناحیه 5×5 ، و شکل ۴.۲۳.۲ (ت) حاصل فیلتر میانه 5×5 را نشان می دهد. در این مورد برتری فیلتر میانه نسبت به متوسط گیری همسایگی به هیچ شرحی نیاز ندارد. نقاط روشن باقی مانده در شکل ۴.۲۳.۲ (ت)، ناشی از تمرکز زیاد نویز در حول آن نقاط است؛ بنابراین محاسبه 'میانه را بایاس می کنند'^{۱)}. دو یا چند بار مرور تصویر با فیلتر میانه آن نقاط را نیز حذف خواهد کرد.



۳.۳.۴ فیلترهای تیزکننده

هدف اصلی تیزکردن، تقویت جزئیات ریز تصویر با ارتقای جزئیاتی است که بر اثر خطای در نتیجه یک روش تصویربرداری خاص مات شده اند. کاربردهای تیزکردن تصویر گوناگون است و کاربردهایی از چاپ الکترونیکی و تصویربرداری پزشکی گرفته تا بازرسی صنعتی و تشخیص خودکار هدف در سلاحهای هوشمند را شامل می شوند.

فیلتر کودن مکانی بالاگذر پایه ای

شکل پاسخ ضربه مورد نیاز برای پیاده سازی یک فیلتر مکانی بالاگذر (تیزکننده) یا ان می کند که فیلتر باید در نزدیک مرکز ضرایب ثابت و در حاشیه ها ضرایب منفی داشته باشد (شکل ۴.۱۹ ب را بینید). برای ماسک 3×3 ، انتخاب مقدار ثابت در نقطه 'مرکز' و مقادیر منفی در سایر نقاط این شرط را برآورده می کند.

شکل ۴.۲۴ پیاده سازی کلاسیک یک فیلتر تیزکننده 3×3 را نشان می دهد. توجه کنید که جمع ضرایب برابر با صفر است. بنابراین وقتی ماسک روی ناحیه ای با سطح خاکستری ثابت یا با تغییرات آرام

1) Biasing

۳۰.۴ فیلتر کردن مکانی

۲۲۱



(الف)



(ب)



(پ)



(ت)

شکل ۳۰.۴ (الف) : تصویر اولیه؛ (ب) تصویر تخریب شده با نویز ضربه ای؛ (پ) حاصل متوسط گیری در همسایگی 5×5 (ت) حاصل فیلتر میانه 5×5 . با مجوز از Texas Instruments, Inc., Martin Connor, Lewisville, Tex.

قرار داشته باشد، خروجی ماسک صفر یا بسیار کوچک خواهد بود (معادله ۴.۱-۳ را ببینید). این نتیجه با آنچه که از فیلتر میدان فرکانس متناظر موجود در شکل ۴.۱۹ (ب) انتظار می‌رود، سازگار است. همچنین توجه کنید که این فیلتر مؤلفه' با فرکانس صفر را حذف می‌کند. طبق توضیحات بخش ۳.۶.۳، حذف این مؤلفه مقدار متوسط سطح خاکستری تصویر را به صفر کاهش می‌دهد و تمایز سراسری تصویر را تا حد زیادی کاهش می‌دهد. شکل ۴.۲۵ (ب) حاصل کاربرد فیلتر شکل ۴.۲۴

$$\frac{1}{9} \times \begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline -1 & 8 & -1 \\ \hline -1 & -1 & -1 \\ \hline \end{array}$$

شکل ۴.۳۰ یک فیلتر مکانی بالاگذر پایه ای

روی تصویر شکل ۴.۲۵ (ب) را نشان می دهد. تصویر اولیه شامل جزئیات نسبتاً ریزی بر روی بخش قابل توجهی از سطوح زمینه ای که تغییرات آرام دارند، می باشد. همانطور که انتظار می رفت، نتیجه فیلتر کردن دارای لبه های تقریباً ارتقا یافته روی زمینه ای نسبتاً تاریک می باشد. با استفاده از فیلتر کردن تقویت بالا^۱ که در بخش بعدی توضیح داده می شود، می توان به نتایج بسیار بهتری دست یافت. کاهش مقدار متوسط تصویر به صفر بدین معناست که تصویر تعدادی سطوح خاکستری منفی دارد. وقتی تنها با سطوح مشت سروکار داشته باشیم، با استفاده از یک یا چند نوع مقیاس کردن و / یا پرش زدن بر روی نتایج فیلتر کردن بالاگذر، سطوح خاکستری نتیجه^۲ نهایی در بازه $[1-L, 0]$ قرار می گیرند. قدر مطلق گرفتن از تصویر فیلتر شده به منظور مشت کردن تمام مقادیر، فکر خوبی نیست، زیرا مقادیر منفی بزرگ به صورت کاملاً روشن در تصویر ظاهر می شوند.

فیلتر کردن تقویت بالا

تصویر فیلتر بالاگذر شده را می توان به صورت تفاضل تصویر اولیه و نسخه^۳ فیلتر پایین گذر شده^۴ آن تصویر کرد؛ یعنی،

$$\text{پایین گذر} - \text{اولیه} = \text{بالاگذر} \quad (4-3)$$

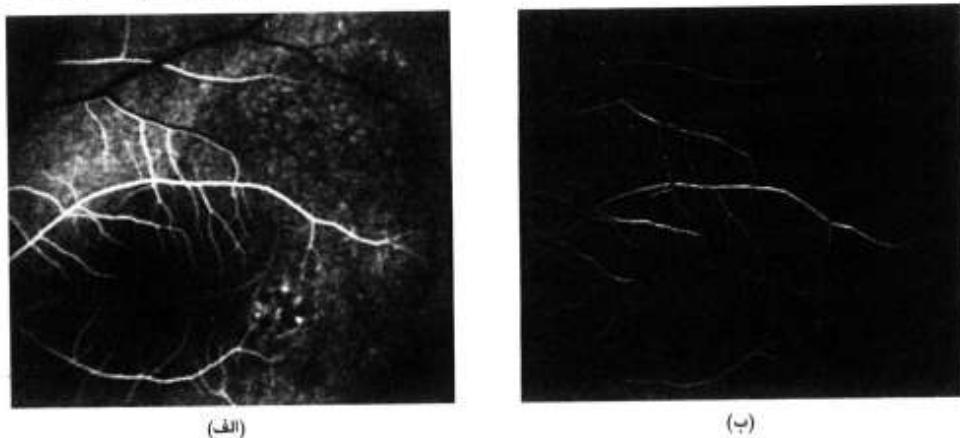
به عنوان تمرین توصیه می شود که درستی این معادله را با استفاده از معادله^۴ ۴.۱-۳ در مورد اشکال ۴.۲۱ (الف) و ۴.۲۴ بررسی کنید. ضرب تصویر اولیه در یک ضریب تقویت، که با A نشان داده می شود، تعریف فیلتر تقویت بالا یا فیلتر تأکید فرکانس بالا^۵ را بدست می دهد:

$$\begin{aligned} \text{پایین گذر} - (\text{اولیه})(A) &= \text{تقویت بالا} \\ \text{پایین گذر} - \text{اولیه} + (\text{اولیه})(A-1) &= (A-1)(\text{اولیه}) \\ \text{بالاگذر} + (\text{اولیه})(A-1) &= (A-1)(\text{اولیه}) \end{aligned} \quad (3-3)$$

1) High-boost filtering

2) High-frequency-emphasis filter

۳۰.۴ فیلتر کردن مکانی



شکل ۳۰.۴ (الف) تصویر شبکیه چشم انسان؛ (ب) حاصل بالاگذر فیلتر شده با استفاده از ناقاب شکل ۲۴.۴

مقدار $A = 1$ ، فیلتر بالاگذر استاندارد را می دهد. وقتی $1 > A$ باشد، بخشی از تصویر اولیه نیز به خروجی بالاگذر افزوده می شود که در نتیجه بخشی از مؤلفه های فرکانس پایین حذف شده در عمل فیلتر کردن بالاگذر را جبران می کند. نتیجه این است که تصویر تقویت بالا با مقداری ارتقای لبه که به مقدار A بستگی دارد، بیشتر به تصویر اولیه شباهت نشان می دهد. به فرآیند کلی تفربیق تصویر مات شده از اولیه، آنطور که در سطر اول معادله^{۱)} (۴.۳-۳) آمده است، حذف غیر تیزی^{۱)} می گویند. این روش یکی از ابزارهای پایه ای برای کاربردهای پردازش تصویر در صنعت چاپ و نشر است. از دیدگاه پاده سازی، با فرض این که وزن مرکز ناقاب در شکل ۲۶.۴ برابر

$$w = 9A - 1 \quad (4-3.4)$$

بازای $1 \geq A$ باشد، می توان نتایج قبلی را با هم ترکیب کرد. مقدار A طبیعت فیلتر را تعیین می کند.

$$\frac{1}{9} \times \begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline -1 & w & -1 \\ \hline -1 & -1 & -1 \\ \hline \end{array}$$

شکل ۲۶.۴ ناقاب استفاده شده برای فیلتر کردن مکانی تقویت بالا. مقدار وزن مرکزی $w = 9A - 1$ است که $A \geq 1$ باشد.

1) Unsharp masking

مثال: شکل ۴.۲۷، تصویر اولیه^{۱)} شکل ۴.۲۵.(الف) و نتایج فیلتر کردن تقویت بالا بازای $A = 1.1$ و ۱.۲ را نشان می دهد. با مقایسه نتایج شکل ۴.۲۷ و شکل ۴.۲۵ مزیت استفاده از فیلتر کردن تقویت بالا نسبت به فیلتر کردن بالاگذر کلاسیک روشن می شود. مخصوصاً توجه کنید که طبق توضیحات قبلی، حاصل فیلتر کردن تقویت بالا بازای $A = 1.1$ که معادل جمع ۰.۱ تصویر اولیه با خروجی بالاگذر پایه است، بهبود قابل توجهی نسبت به نتیجه^{۱)} شکل ۴.۲۵.(ب) نشان می دهد. وقتی A افزایش یابد، زمینه خروجی فیلتر تقویت بالا روشنتر می شود. در واقع خروجی فیلتر بازای $A = 1.2$ در مرز مقبولیت است. همچنین توجه کنید که نویز نقش مهمی در ظاهر چشمی تصویر خروجی^{۱)} نداشت بالا بازی است. این نتیجه غیرمنتظره نیست، زیرا فیلتر کردن بالاگذر، نویز رانیز همراه سایر تغییرات سریع شدت می کند. این نتیجه غیرمنتظره نیست، زیرا فیلتر کردن بالاگذر، نویز رانیز همراه سایر تغییرات سریع شدت افزایش می دهد. بوضوح فیلتر کردن تقویت بالا این اثر را افزایش می دهد. □

مانند مورد فیلترهای مکانی پایین گذر، تعیین فیلترهای مکانی بالاگذر با ابعاد بزرگتر از آنهایی که تاکنون تشریح شدند، نیز ممکن است. برای نمونه، یک فیلتر بالاگذر 7×7 پایه ای وزن مرکزی ۴۸ را خواهد داشت، در حالی که سایر ضرایب آن -1 و ضریب تراز سازی $\frac{1}{49}$ خواهد بود. با این حال در عمل بندرت فیلترهای بالاگذر بزرگتر از 3×3 مورد نیاز است.

فیلترهای مشتقی^{۱)}

متوسط گیری از پیکسل ها در یک ناحیه تصویر منجر به مات شدن جزئیات تصویر می شود. چون متوسط گیری نوعی انتگرال گیری است، انتظار می رود که مشتق گیری اثر مخالف را داشته باشد و بنابراین تصویر را تیز نماید.

عمومی ترین روش مشتق گیری در کاربردهای پردازش تصویر، گرادیان است. برای تابع $(y, f(x, y))$ ، گرادیان f در مختصات (x, y) به صورت بردار زیر تعریف می شود:

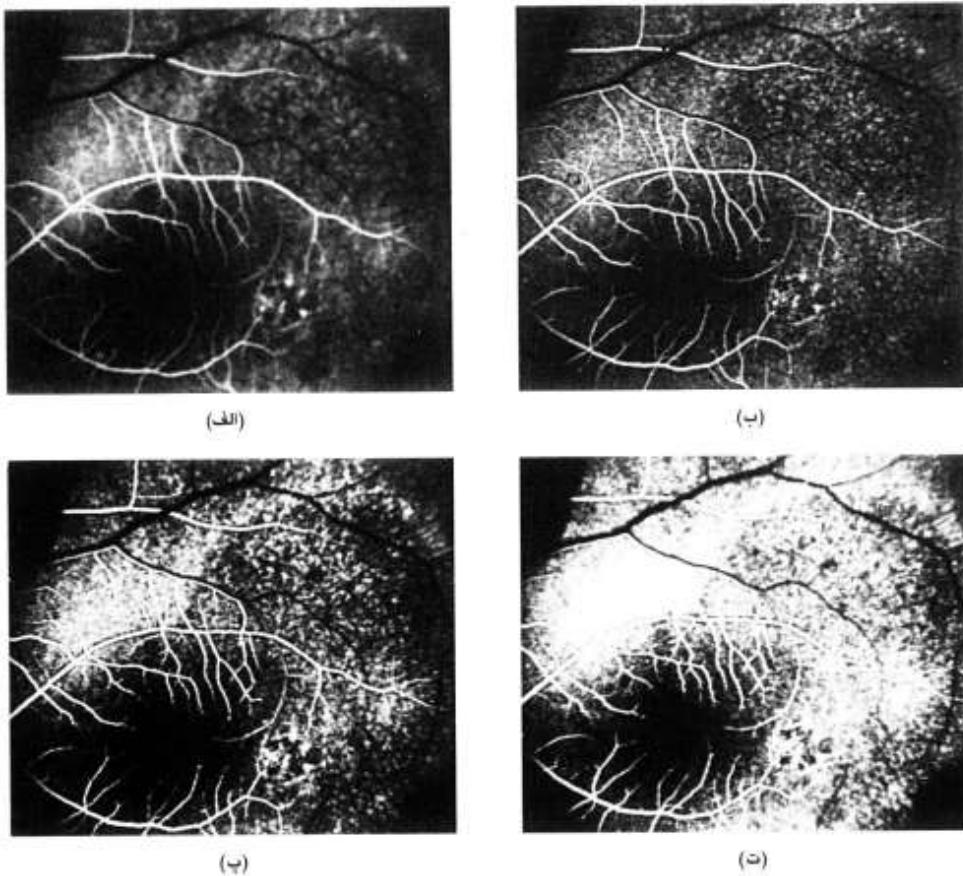
$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (5-3.4)$$

بزرگی این بردار،

$$\nabla f = \text{mag}(\nabla f) = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{1/2} \quad (6-3.4)$$

1) Derivative filters

۳.۴ فیلتر کردن مکانی ۲۲۵



شکل ۳.۲۷. (الف) تصویر اولیه؛ (ب)- (ت) حاصل فیلتر کردن تقویت بالا با استفاده از نقاب شکل ۳.۲۶ به ترتیب با $A = 1.1, 1.15, 1.2$. این نتایج را با نتایج شکل ۳.۲۵ مقایسه کنید.

مبنا برای روش های گوناگون مشتق گیری تصویر می باشد. ناحیه' تصویری نشان داده شده در شکل ۳.۲۸. (الف) را در نظر بگیرید، که جهانشانه' مقادیر سطوح خاکستری هستند. به چند روش می توان معادله' $(z_6 - z_3)^2 + (z_5 - z_4)^2$ را در نقطه' z تقریب زد. ساده ترین روش، استفاده از تفاضل $(z_5 - z_6)$ درجهت x و تفاضل $(z_5 - z_6)$ درجهت y می باشد که به صورت

$$\nabla f = [(z_5 - z_6)^2 + (z_5 - z_6)^2]^{1/2}. \quad (3.4-۷)$$

ترکیب می شوند. به جای استفاده از توان ها و ریشه های دوم، می توان با استفاده از قدر مطلق نیز به نتایج مشابهی دست یافت:

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

(الف)

1	0	0	1
0	-1	-1	0

(ب) رابرتز

-1	-1	-1	-1	0	1
0	0	0	-1	0	1
1	1	1	-1	0	1

(ب) پرویت

-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1

(ت) سوبل

شکل ۲۸.۴ یک ناحیه 3×3 از تصویر (zها مقادیر سطح خاکستری هستند) و نقاط های مختلفی که برای محاسبه مشتق در نقطه z_5 استفاده می شوند. توجه کنید که مجموع ضرایب هر نقاط صفر است، که همان طور که در مورد یک عملگر مشتقی انتظار می رود، نشانه پاسخ "0" در نواحی ثابت است.

$$\nabla f = |z_5 - z_8| + |z_5 - z_6| \quad (28.3.b)$$

روش دیگر برای تقریب معادله (28.3.c) استفاده از تفاضلهای متقاطع به صورت

$$\nabla f \approx [(z_5 - z_9)^2 + (z_6 - z_8)^2]^{1/2} \quad (28.3.c.الف)$$

یا استفاده از قدر مطلق هایی به صورت

۳.۴ فیلتر کردن مکانی ۲۴۷

$$\nabla f \approx |z_5 - z_9| + |z_6 - z_8|. \quad (4-3.4)$$

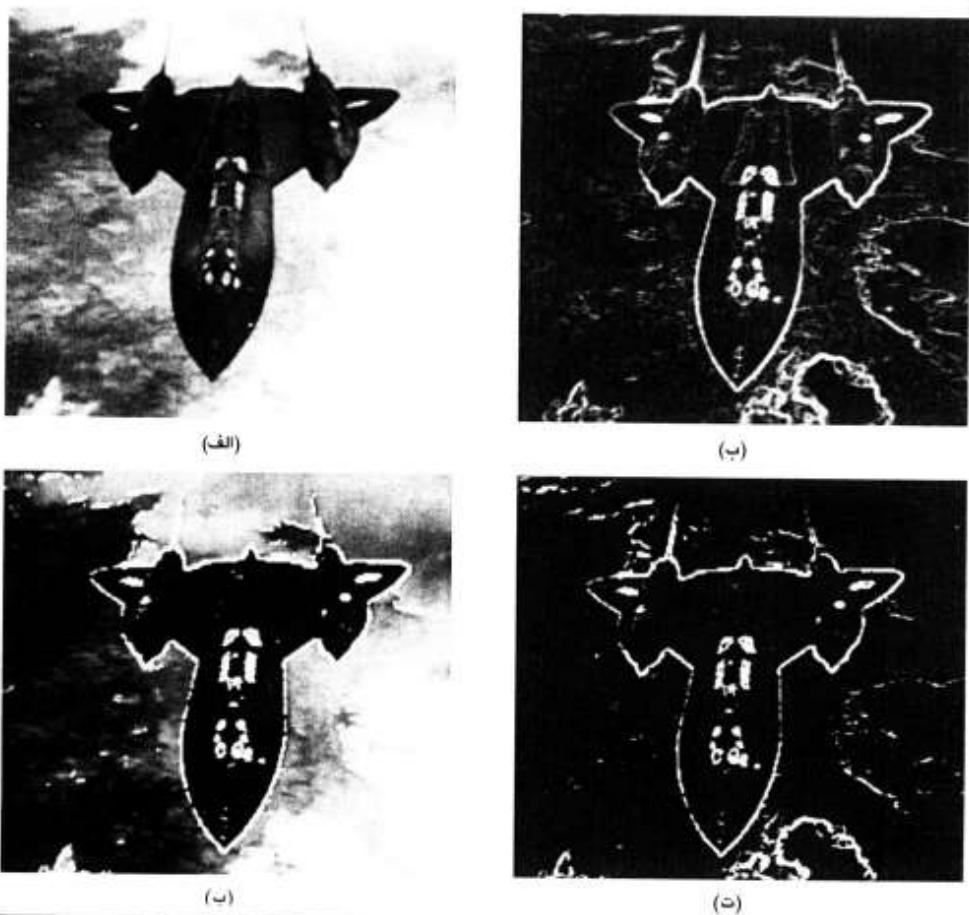
می باشد. معادلات (۴.۷-۳) و (۴.۸-۳) را می توان با استفاده از ناقاب های 2×2 پیاده سازی کرد. به عنوان مثال معادله (۴.۸-۳) را می توان با گرفتن قدر مطلق پاسخ دو ناقاب موجود در شکل ۴.۲۸. (ب) و جمع نتایج، پیاده سازی کرد. این ناقاب ها عملگرهای گرادیان متقاطع رابرتز^۱ نامیده می شوند.

ماسک های با ابعاد زوج، پیاده سازی مشکلی دارند. یک تقریب دیگر برای معادله (۴.۶-۳)، باز هم در نقطه^۲، اما این بار با استفاده از همسایگی 3×3 عبارتست از

$$\begin{aligned} \nabla f \approx & |(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)| \\ & + |(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)|. \end{aligned} \quad (4-3.4)$$

تفاضل بین سطر سوم و اول ناحیه 3×3 ، مشتق در جهت x . را تقریب می زند و تفاضل بین ستون سوم و اول، مشتق در جهت y را تقریب می زند. ناقاب های موجود در شکل ۴.۲۸. (پ) که عملگرهای پرویت^۲ نامیده می شوند، را می توان برای پیاده سازی معادله (۴.۹-۳) بکار برد. نهایتاً شکل ۴.۲۸. (ت) زوج دیگری از ناقاب ها (که عملگرهای سوبیل^۳ نامیده می شوند) را نشان می دهد که جهت تقریب زدن بزرگی گرادیان بکار می روند. در بخش ۷.۱ عملگرهای سوبیل را بیشتر شرح می دهیم.

مثال: شکل ۴.۲۹. (الف) یک تصویر اولیه و شکل ۴.۲۹. (ب) حاصل محاسبه بزرگی گرادیان با استفاده از ناقاب های پرویت را نشان می دهد (معادله ۴.۹-۳). شکل ۴.۲۹. (پ) با کاملاً سفید (۲۵۵) قرار دادن هر مقداری از گرادیان که بیش از ۲۵ (یعنی تقریباً بزرگتر یا مساوی ده درصد بزرگترین مقدار سطح خاکستری ممکن در تصویر) بود، بدست آمد. هر نقطه ای که مقدار گرادیان در آن این معیار را برآورده نکند، برابر مقدار اولیه اش در تصویر قرارداده می شود. بدین ترتیب، این روش، زمینه را بدون تغییر نگه می دارد، در حالی که لبه های مهم را ارتقا می دهد. در نهایت شکل ۴.۲۹. (ت) نیز به همان روش شکل ۴.۲۹. (پ) بدست آمد، جز این که نقاطی که گرادیان در آنها بیش از ۲۵ نبود، برابر صفر (سباه) قرار داده شدند. نتیجه این فرآیند یک تصویر دودویی می باشد. در تمام حالات، لبه های اصلی به طور قابل ملاحظه ای ارتقا می یابند.



شکل ۴.۲۹ ارتقای لبه با روش‌های گرادیانی (متن را بینید)

۴.۴ ارتقا در میدان فرکانس

طبق توضیحات بخش ۴.۲، ارتقا در میدان فرکانس در اصل ساده است. کافی است تبدیل فوریه تصویری که باید ارتقا یابد، را محاسبه کرده، حاصل را در تابع انتقال فیلتر ضرب نموده و از نتیجه تبدیل معکوس بگیریم تا تصویر ارتقا یافته بدست آید.

ایده های مات کردن با تضعیف محتوای فرکانس بالا و تیز کردن با تقویت مؤلفه های فرکانس بالا نسبت به مؤلفه های فرکانس پایین، از مفاهیمی استخراج می گردند که مستقیماً با تبدیل فوریه مرتبط هستند. در واقع، ایده فیلتر کردن خطی در میدان فرکانس به طور قابل توجهی جذابتر و شهودی تر است. البته در عمل نقاب های مکانی کوچک، به علت پیاده سازی آسان و سرعت عملشان به طور قابل ملاحظه ای بیشتر

۴.۴ ارتقا در میدان فرکانس

از تبدیل فوریه استفاده می شوند. با این حال فهم مفاهیم میدان فرکانس در حل بسیاری از مسائل که بسادگی با فتون مکانی قابل حل نیستند، ضروری است. روش فیلتر کردن همشکلی^۱ که در این بخش تشریح شد و چند روش بازیابی تصویر که در فصل ۵ بیان شدند، مثالهایی از این دسته مفاهیم هستند.

۱.۴.۴ فیلتر کردن پایین گذرهای

همان طور که قبل^۲ بیان شد، لبه ها و سایر تغییرات سریع در سطح خاکستری تصویر (نظیر نویز) به میزان زیادی در محتوای فرکانس بالای تبدیل فوریه تصویر سهیم هستند. بنابراین مات کردن (آرام کردن) در میدان فرکانس، با تضعیف محدوده^۳ مشخصی از مؤلفه های فرکانس بالای تبدیل فوریه^۴ تصویر حاصل می گردد.

از معادله^۱ (۴-۱)،

$$G(u, v) = H(u, v)F(u, v) \quad (1-4.4)$$

که تبدیل فوریه تصویری که باید آرام شود، می باشد. هدف، انتخاب تابع انتقال فیلتر $H(u, v)$ است که با تضعیف مؤلفه های فرکانس بالای $F(u, v)$ ، $G(u, v)$ را بدهد. آنگاه تبدیل معکوس، تصویر آرام شده مطلوب (x, y) را خواهد داد. در توضیح بعدی دسته ای از توابع انتقال فیلتر را در نظر می گیریم که به طور دقیقاً یکسانی بر قسمت های حقیقی و موهومی $F(u, v)$ اثر می گذارند. چنین فیلتر هایی با عنوان فیلتر های با انتقال فاز صفر^۵ اشاره می شوند، زیرا فاز تبدیل را تغییر نمی دهند.

فیلتر ایده آل^۶

یک فیلتر پایین گذرهای دو بعدی^۷ (ILPF) فیلتری است که تابع انتقالش در رابطه

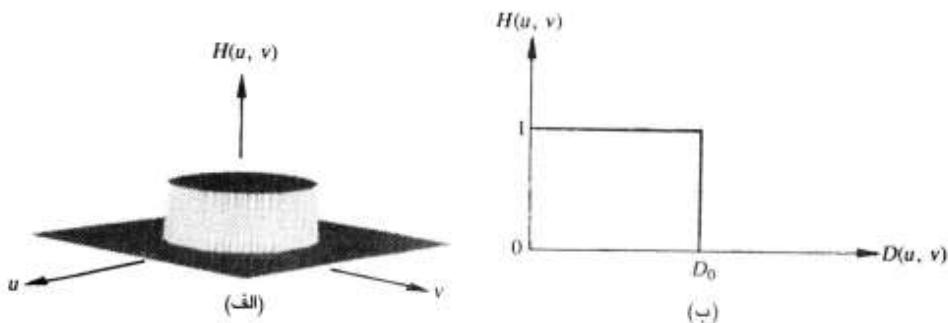
$$H(u, v) = \begin{cases} 1 & D(u, v) \leq D_0 \\ 0 & D(u, v) > D_0 \end{cases} \quad \begin{array}{l} \text{اگر} \\ \text{اگر} \end{array} \quad (2-4.4)$$

1) Homomorphic filtering

2) Zero- phase- shift filters

3) Ideal filter

4) Ideal lowpass filter (ILPF)



شکل ۴.۳۰ (الف) نمای پرسپکتیو تابع انتقال فیلتر پایین گذر ایده‌آل؛ (ب) مقطع عرضی فیلتر

صدق کند. در این رابطه D_0 یک کمیت غیر منفی معین است، و (u, v) فاصله نقطه (u, v) تا مبدأ صفحه فرکانسی می‌باشد؛ یعنی،

$$D(u, v) = (u^2 + v^2)^{1/2}. \quad (4-4.4)$$

شکل ۴.۳۰ (الف) پرسپکتیو سه بعدی $H(u, v)$ را به صورت تابعی از u و v نشان می‌دهد. ایده‌آل بودن فیلتر بدین معناست که تمام فرکانس‌های درون دایره با شعاع D_0 بدون تضعیف عبور داده می‌شوند، در حالی که تمام فرکانس‌های خارج آن دایره کاملاً تضعیف می‌شوند.

فیلترهای پایین گذر که در این فصل بررسی می‌شوند، متقارن شعاعی حول مبدأ هستند. همان‌طور که شکل ۴.۳۰ (ب) نشان می‌دهد، برای این نوع فیلتر تعیین یک مقطع عرضی که به صورت تابعی از فاصله تا مبدأ در طول یک خط شعاعی باشد، کافی است. تابع انتقال کامل فیلتر را می‌توان با چرخش ۳۶۰ درجه این مقطع عرضی حول مبدأ تولید کرد. همان‌طور که در بخش ۳.۲.۲ تشریح شد، تعیین فیلترهای متقارن شعاعی که در مرکز یک مربع فرکانسی $N \times N$ قرار دارند، براساس این فرض است که مبدأ تبدیل فوریه در مرکز مربع قرار دارد.

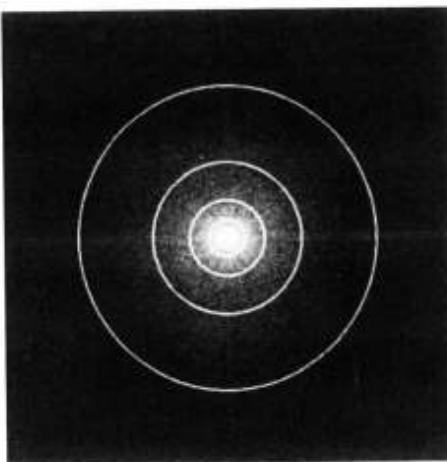
برای مقطع عرضی فیلتر پایین گذر ایده‌آل، اغلب نقطه تغییر از $H(u, v) = 0$ به $H(u, v) = 1$ قطع^{۱)} می‌نامند. به عنوان مثال در شکل ۴.۳۰ (ب) فرکانس قطع D_0 است. وقتی مقطع عرضی حول مبدأ چرخانده شود، نقطه D_0 یک منحنی دایروی از فرکانس‌های قطع که همه آنها به فاصله D_0 از مبدأ

1) Cutoff frequency

۴.۴ ارتقا در میدان فرکانس ۲۵۱



(الف)



(ب)

شکل ۳۱.۴ (الف) یک تصویر 512×512 و (ب) طیف فوریه آن. دایره های قرار داده شده روی طیف که شعاع های ۸، ۱۸، ۴۳، ۷۸ و ۱۵۲ دارند، به ترتیب ۹۰، ۹۳، ۹۵، ۹۹ و ۹۹/۵ درصد توان تصویر را محصور می کنند.

هستند، بدست می دهد. مفهوم فرکانس قطع در تعیین مشخصات فیلتر بسیار مفید است. این مفهوم به عنوان معیاری عمومی جهت مقایسه رفتار انواع فیلترها بکار می رود.

یک فیلتر پایین گذر ایده آل با فرکانس قطع تیز رانمی توان با قطعات الکترونیکی پیاده سازی کرد، گرچه یقیناً می توان آن را در رایانه شبیه سازی کرد. پس از مثال بعدی، اثرات استفاده از این فیلتر های "غیر واقعی" بر روی تصاویر رقمی تشریح می شوند.

مثال: در شکل ۴.۳۱ (الف) تصویری از یک زنبور درشت اروپایی که قسمت نرم میوه را می خورد، نشان می دهد. این شکل بدلیل تنوع جزئیات آن، از ریز (موهای سر) تا درشت (شاخکها و سایر ساختارهای تبره' مشابه روی سر)، انتخاب شد تا بدینوسیله عملکرد فیلتر های پایین گذر بیان شده در این بخش را نمایش دهد. به ویژه به نقطه' سفید در بخش میانی بالای سر، انعکاس شاخکها و تورفتگی های کوچک در بخش جلویی سر توجه کنید.

کارآئی فیلتر های پایین گذر معرفی شده در این بخش با استفاده از مکان هندسی فرکانس های قطع یکسان مقایسه می شوند. یک طریق برای تعیین مجموعه ای از مکان هندسی های استاندارد، محاسبه 'دوایری است که مقادیر مختلف P_r ، توان کل سیگنال، را شامل می شوند. این کمیت با جمع توان در هر نقطه' (،،

v بازای $1, 0, 1, \dots, N - u$ بحسبت می آید؛ یعنی،

$$P_T = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} P(u, v)$$

که $P(u, v)$ با معادله (۳-۱-۱۳) داده می شود. اگر تبدیل در مرکز قرار گرفته باشد، دایره ای به شعاع r که مبدأ آن در مرکز مربع فرکانسی قرار دارد، β درصد توان را شامل می شود، که

$$\beta = 100 \left[\sum_u \sum_v P(u, v) / P_T \right]$$

و مجموع روی مقادیری از (u, v) که درون یا روی محیط دایره قرار دارند، محاسبه می شود.
شکل ۴.۳۱ (ب) طیف فوریه تصویر موجود در شکل ۴.۳۱ (الف) را نشان می دهد. دایره هایی که روی طیف مشخص شده اند، شعاعهای $r = 8, 18, 43, 78, 152$ دارند. این دایره ها به ترتیب طوری که $90, 93, 95, 99, 99.5 = \beta$ درصد توان را محصور می کنند. دامنه طیف به سرعت کاهش می یابد به طوری که 90 درصد کل توان توسط دایره 'نسبتاً کوچکی با شعاع 8 محصور شده است.

شکل ۴.۳۲ نتایج کاربرد فیلتر های پایین گذر ایده آل با فرکانس های قطع در شعاع های گفته شده را نشان می دهد. شکل ۴.۳۲ (ب) از هر نظر بی فایده است. ماتی شدید این تصویر بدین دلیل است که بیشترین اطلاعات مربوط به جزئیات تیز عکس، در 10 درصد توانی که به وسیله 'فیلتر حذف شده است، قرار داشته اند. وقتی شعاع فیلتر افزایش یابد، انرژی فرکانس بالا کمتر حذف می شود، که نتیجه آن کاهش مات شدگی تصویر است. با این حال حتی وقتی 5 درصد انرژی حذف شود، تصویر مات شده باز هم مبتلا به حلقه زدن^۱ شدید که خاصیتی از فیلتر های ایده ال است، می گردد. همان طور که در مقایسه 'انعکاس شاخک ها و نقطه' بالای سر با موارد مشابه در شکل ۴.۳۲ (الف) آشکار است، با حذف تنها یک درصد انرژی میزانی از مات شدگی در تصویر بوجود می آید. در نهایت، خروجی فیلتر بازای $99.5 = \beta$ اساساً همانند تصویر اولیه است. این موضوع نشان می دهد که در این مورد خاص، اطلاعات کمی از لبه ها و سایر جزئیات تیز تصویر در 5% درصد بالایی طیف توان قرار دارد.

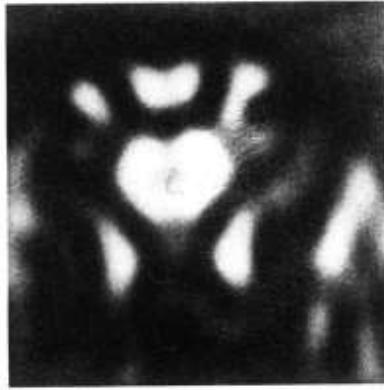
خواص مات کردن و حلقه زدن ILPF را بسادگی می توان با رجوع به قضیه کانولوشن شرح داد. چون

1) Ringing

۴۵۳ ۴.۴ ارتقا در میدان فرکانس



(الف)



(ب)



(ب)



(ت)



(ث)



(ج)

شکل ۴.۳۲۰ (الف) تصویر اولیه؛ (ب)- (ج) نتایج فیلتر کردن پایین گذر ایده آل با فرکانس های قطع که در شعاع های تسان داده شده در شکل ۴.۳۱ (ب) تعیین شده اند.

تبديل های فوريه' تصاویر اوليه و مات شده در ميدان فركانس با معادله'

$$G(u, v) = H(u, v) F(u, v)$$

به هم مربوط می شوند، استفاده از قضيه' کانولوشن، عبارت

$$g(x, y) = h(u, v) * f(x, y)$$

را در ميدان مكان نتيجه می دهد که $h(x, y)$ تبدل فوريه' معکوس تابع انتقال فیلتر $H(u, v)$ است.

كليدهم مات کردن به عنوان يك فرآيند کانولوشن در ميدان مكان، در طبیعت $h(x, y)$ قرار دارد. برای $h(x, y)$ صورت کلي نشان داده شده در شکل ۴.۳۳ (الف) را دارد^۱. همان طور که شکل ۴.۳۳، ILPF (ب) نشان می دهد، فرض کنید که $f(x, y)$ تصویر ساده ای باشد که از دو پیکسل روشن روی زمینه ای سیاه تشکیل شده است. دو نقطه' روشن را می توان به عنوان تقریب هایی از دو ضربه که قدرتشان بستگی به روشنایی آن نقاط دارد، در نظر گرفت. همان طور که در بخش ۳.۸.۳ تشریح شد، کانولوشن $h(x, y)$ معادل فرآيند "کمی کردن"^۱ $h(x, y)$ در محل هر ضربه است. نتيجه' این عمل که در شکل ۴.۳۳ (ب) دیده می شود، نشان می دهد که چگونه دو نقطه' اولیه بر اثر کانولوکردن $f(x, y)$ با تابع مات کننده $h(x, y)$ مات می شوند. اگر هر پیکسل را به عنوان ضربه ای در نظر بگیریم که قدرت آن متناسب با سطح خاکستری آن پیکسل است، این مقاهم به تصاویر پیچیده تر قابل تعمیم هستند.

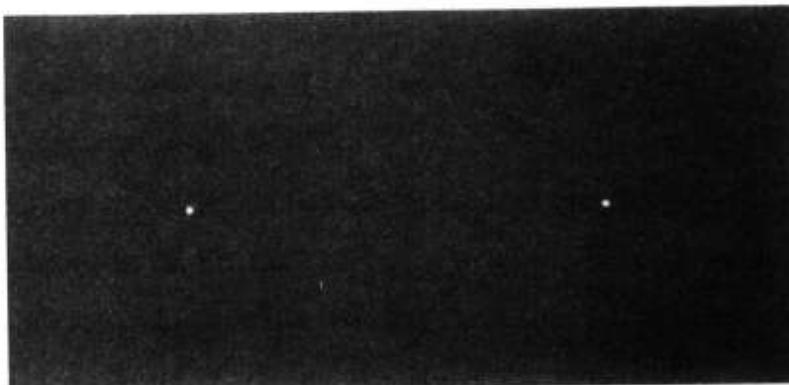
شکل (y) به شعاع تابع فیلتر در ميدان فركانس بستگی دارد. محاسبه' تبدل معکوس $H(u, v)$ مربوط به يك ILPF نشان می دهد که شعاع حلقه های هم مرکز در (y) با مقدار D_0 در معادله' (۴.۲-۴) تناسی معکوس دارند. بنابراین فیلتر کردن شدید در ميدان فركانس (يعني انتخاب D_0 کوچک) تعداد نسبتاً کمی از حلقة های وسیع در (y) تولید کرده و در نتيجه میزان مات شدگی (y) را افزایش می دهد. وقتی D_0 افزایش یابد، تعداد حلقة ها در يك ناحیه نیز افزایش یافته و در نتيجه حلقة ها نزدیکتر و مات شدگی کمتر می گردد. این اثر را می توان با مقایسه اشکال ۴.۳۲ (ت) و (ث) مشاهده کرد. اگر خارج از حوزه' تعریف $F(u, v)$ باشد، (y) در ناحیه' مکانی متناظرش برابر "۱" می شود و در نتيجه D_0

^۱ درستی این صورت کلي به سادگي با گرفتن تبدل فوريه' معکوس از معادله' (۴.۲-۴) تأييد می شود.

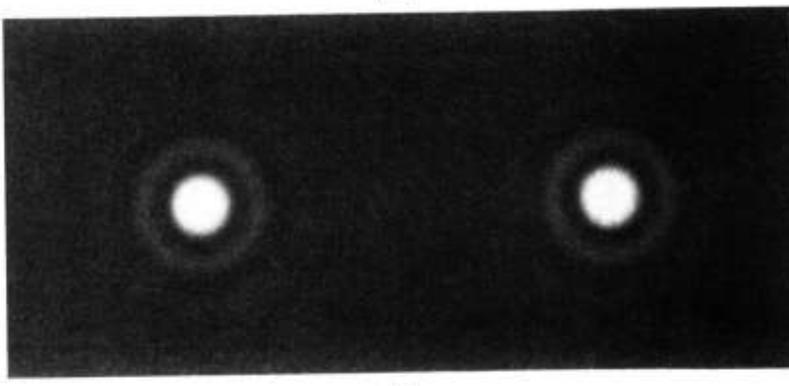
۴۵۵ ۴.۴ ارنقا در میدان فرکانس



(الف)



(ب)



(پ)

شکل ۴.۳۳ نمایش فرآیند مات کردن در میدان مکان: (الف) تابع مات کننده $h(x, y)$ برای یک فیلتر پایین گذر ایده‌آل؛ (ب) یک تصویر ساده که از دو تصویر روشن تشکیل شده است؛ (پ) کانولوشن $f(x, y)$ و $h(x, y)$ و

۲۵۶ پردازش تصویر رقمه

کانولوشن ($h(x, y)$) با $f(x, y)$ است. البته این وضعیت معادل با عدم فیلتر کردن است. اثرات میدان مکان فیلتر باترورث^۱ که در بخش بعدی توضیح داده می شود، را نیز می توان به همین روش بیان کرد.

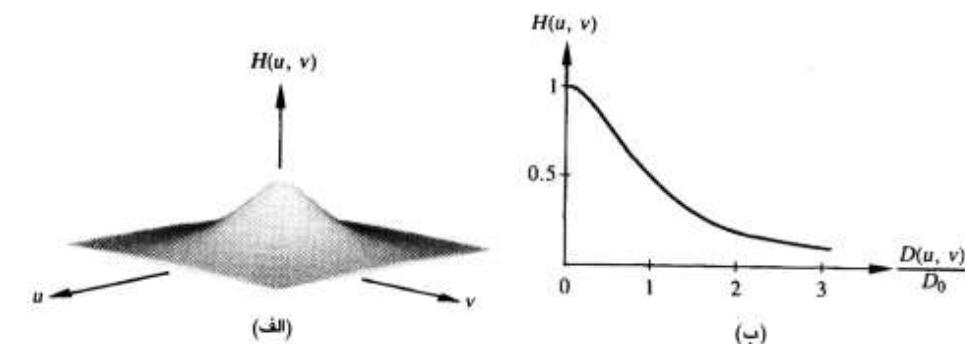
فیلتر باترورث

تابع انتقال فیلتر پایین گذر باترورث (BLPF) از مرتبه n و با مکان هندسی فرکانس قطع در فاصله D_0 از مبدأ با رابطه^۲

$$H(u, v) = \frac{1}{1 + [D(u, v)/D_0]^{2n}} \quad (4-4.4)$$

تعریف می گردد که $D(u, v)$ با معادله^۳ (۴.۳-۴) داده می شود. نمای پرسپکتیو و مقطع عرضی تابع BLPF در شکل ۴.۳۴. دیده می شوند.

برخلاف ILPF، تابع انتقال BLPF ناپوستگی تیز که باعث تعیین فرکانس قطع مشخصی بین فرکانس های عبور داده شده و فیلتر شده گردد، ندارد. برای فیلترهایی که توابع انتقال با تغییرات آرام دارند، معمولاً منحنی فرکانس قطع در نقاطی که $H(u, v)$ برابر کسر معینی از مقدار بیشینه اش باشد، تعریف می شود. در مورد معادله^۴ (۴.۴-۴)، وقتی $D(u, v) = D_0$ باشد، $H(u, v) = 0.5$ در صد کوچکتر از مقدار بیشینه) می باشد. مقدار دیگری که عموماً استفاده می شود، $\sqrt{2}/1$ برابر مقدار بیشینه^۵ $H(u, v)$



شکل ۴.۳۴ (الف) یک فیلتر پایین گذر باترورث؛ (ب) مقطع عرضی شعاعی بازای $n = 1$.

¹) Butterworth filter

است. برای معادله^۱ (۴.۴)، با تغییر ساده زیر و بازای $D_0 = D(u, v)$ مقدار فوق بدست می آید:

$$\begin{aligned} H(u, v) &= \frac{1}{1 + [\sqrt{2} - 1][D(u, v)/D_0]^{2n}} \\ &= \frac{1}{1 + 0.414[D(u, v)/D_0]^{2n}}. \end{aligned} \quad (5-4. ۴)$$

مثال: شکل ۴. ۳۵ حاصل کاربرد BLPF ها (معادله^۱ ۴.۵-۴. ۳۱) را نشان می دهد. در این شکل، $n = 1$ برابر پنج شعاعی است که در شکل ۴. ۳۱ (ب) دیده می شوند. برخلاف نتایجی که در شکل ۴. ۳۲ دیده می شود، در اینجا به تغییر آرامی در میزان مات شدگی به عنوان تابعی از مقدار توان حذف شده از طیف برخورد می کنیم. گذشته از آن، در هیچ یک از تصاویر پردازش شده با BLPF پدیده حلقه زدن دیده نمی شود، که ناشی از تغییرات آرام فیلتر بین فرکانس های پایین و بالا می باشد.

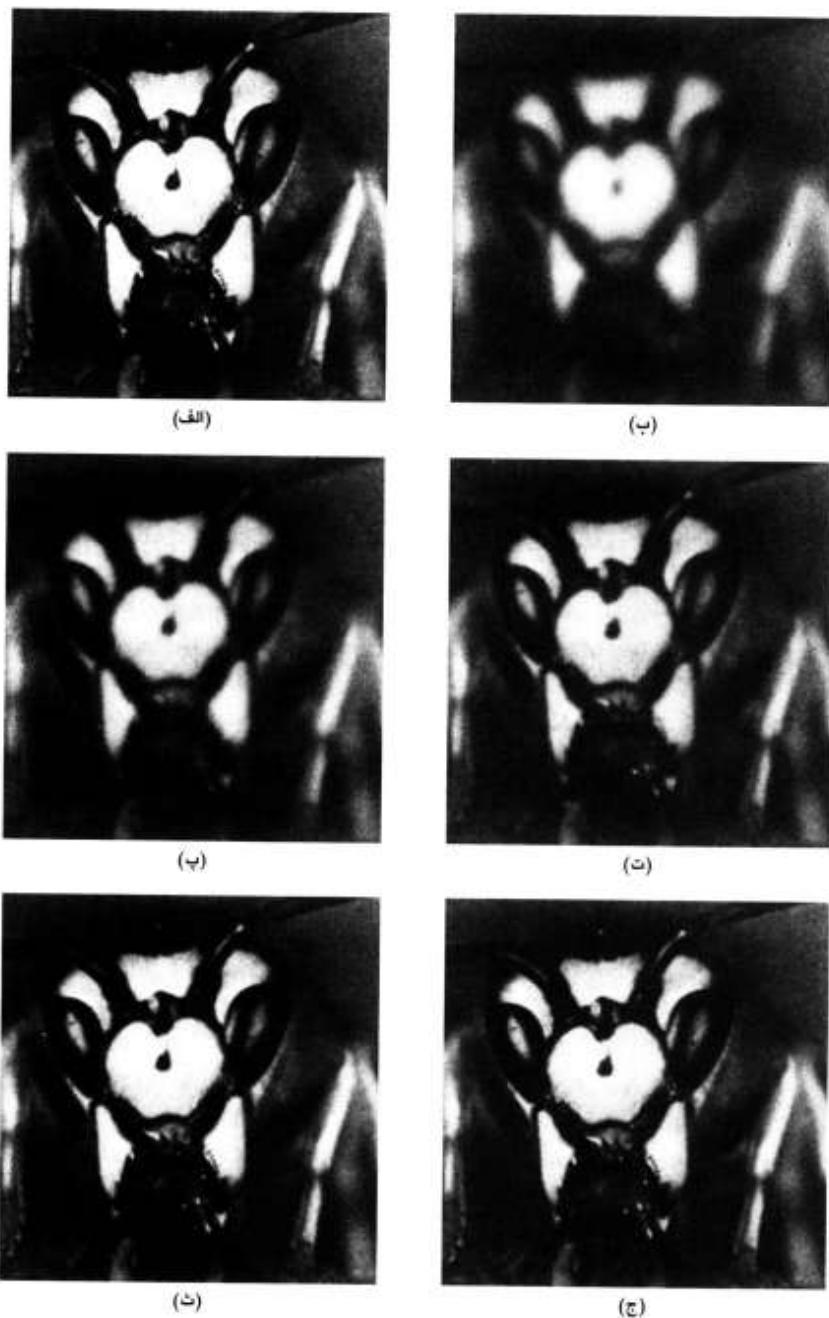
□

نتایج فیلتر کردن پایین گذر که تاکنون ارائه شدند، با استفاده از تصاویری با کیفیت خوب بدست آمده اند تا بدين وسیله بتوان اثرات فیلتر را نمایش داده و آنها را مقایسه نمود. شکل ۴. ۳۶ دو کاربرد عملی فیلتر پایین گذر برای آرام کردن تصویر را نشان می دهد. تصویر شکل ۴. ۳۶ (الف) تنها با ۱۶ سطح خاکستری رقمی شد، و در نتیجه مقدار قابل توجهی از منحنی سازی اشتباه را در خود دارد. شکل ۴. ۳۶ (ب) حاصل آرام کردن این تصویر با یک فیلتر با تورورث پایین گذر مرتبه یک می باشد. به طور مشابه شکل ۴. ۳۶ (ت) اثر کاربرد یک BLPF بر تصویر نویزدار شکل ۴. ۳۶ (ب) را نشان می دهد. از این مثال ها نتیجه می گیریم که فیلتر کردن پایین گذر یک فرآیند ارتقا دهنده است که بازای کاهش تیزی تصویر، اثرات مخرب^۱ را کاهش می دهد.

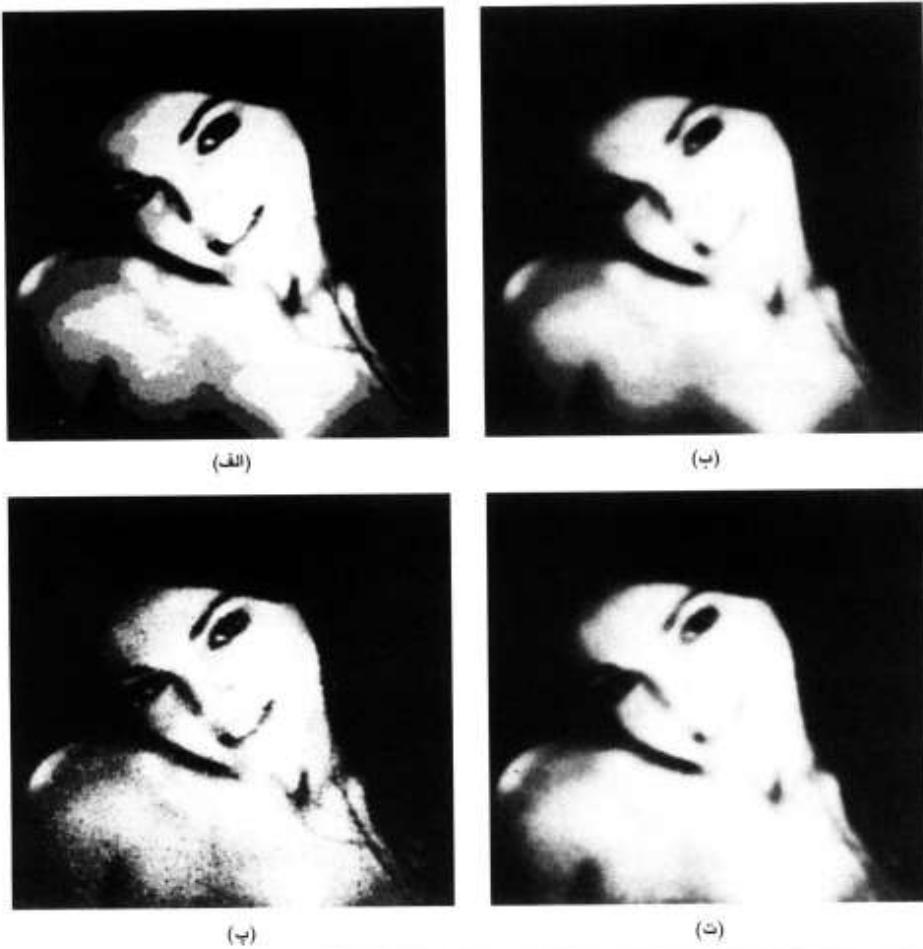
۴.۴.۳ فیلتر کردن بالا گذر

در بخش ۴. ۱ نشان دادیم که می توان با تضعیف مؤلفه های فرکانس بالا در تبدیل فوریه^۱ تصویر، آن را مات کرد. چون لبه ها و سایر تغییرات سریع در سطوح خاکستری با مؤلفه های فرکانس بالا مرتبط هستند، می توان با یک فرآیند فیلتر کردن بالا گذر که بدون تغییر اطلاعات فرکانس بالای تبدیل فوریه،

1) Spurious.



شکل ۳۵. ۳ (الف) تصویر اولیه؛ (ب)- (ج) نتایج فیلتر کردن پایین گذر با ترورث با فرکانس های قطع که در شعاع های نشان داده شده در شکل ۴. ۳۱(ب) تعیین شده اند.



شکل ۴.۳۶ دو مثال از آرام کردن تصویر با فیلتر کردن پایین گذر (متن را بینید)

مؤلفه های فرکانس پایین را تضعیف می کند، تصویر را تیز کرد. در شرح همتاها های فرکانس بالای فیلتر هایی که بخش ۴.۱.۱ بیان شدند، تنها فیلتر هایی با انتقال فاز صفر را برای توضیح بعدی در نظر می گیریم. این فیلتر ها منقارن شعاعی هستند و می توان آنها را به طور کامل با یک مقطع عرضی، که به صورت تابعی از فاصله تا مبدأ تبدیل فوریه مرکزی شده^۱ می باشد، مشخص کرد.

فیلتر ایده آل

فیلتر بالاگذر ایده آل^۲ (IHPF) دو بعدی، فیلتری است که تابع انتقال آن در رابطه

1) Centered fourier transform

2) Ideal highpass filter (IHPF)

۴۶۰ پردازش تصویر رقمه

$$H(u, v) = \begin{cases} 0 & D(u, v) \leq D_0 \\ 1 & D(u, v) > D_0 \end{cases} \quad (6-4. 4)$$

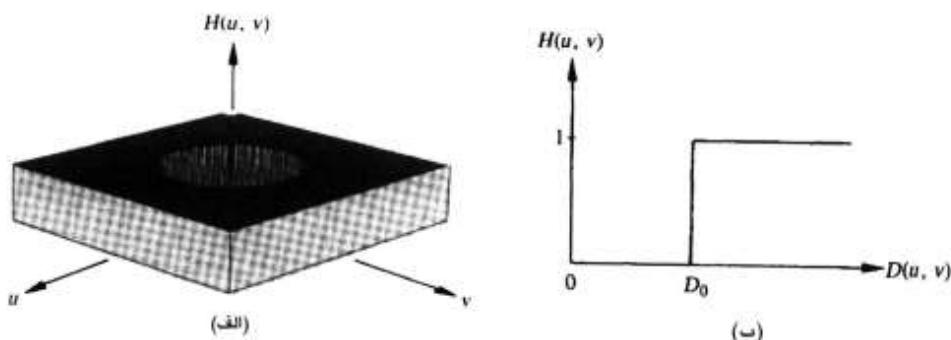
صدق نماید. در این معادله، D_0 فاصله قطع از مبدأ صفحه فرکانس بوده و $D(u, v)$ با معادله '۴. ۳' داده می شود. شکل ۴. ۳۷. نمای پرسپکتیو و مقطع عرضی تابع IHPF را نشان می دهد. این فیلتر متضاد فیلتر پایین گذر ایده اال که در بخش ۴. ۱. تشریح شد، می باشد؛ زیرا تمام فرکانس های درون دایره ای به شعاع D_0 را تضعیف می کند، در حالی که تمام فرکانس های خارج دایره را بدون تضعیف عبور می دهد. مانند فیلتر پایین گذر ایده اال، IHPF را نیز نمی توان به طور فیزیکی تحقق بخشد.

فیلتر بالاترورث

تابع انتقال فیلتر بالاگذر بالاترورث^۱ (BHPF) از مرتبه n و با منحنی فرکانس قطع به فاصله D_0 از مبدأ با رابطه'

$$H(u, v) = \frac{1}{1 + [D_0/D(u, v)]^{2n}} \quad (7-4. 4)$$

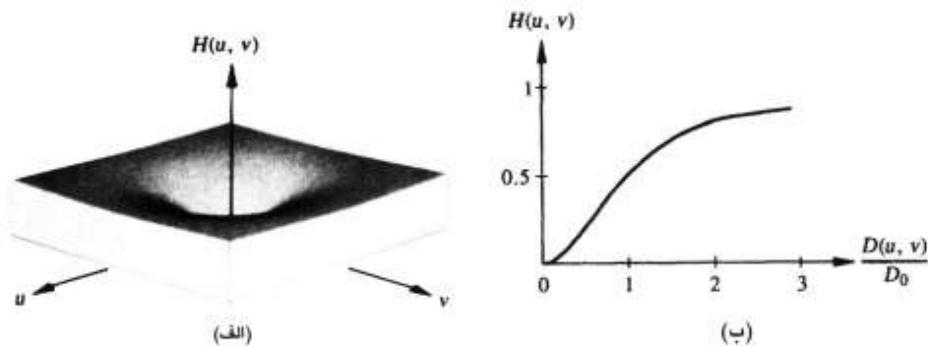
تعريف می شود که $D(u, v)$ با معادله '۳-۴. ۴' داده می شود. شکل ۴. ۳۸. نمای پرسپکتیو و مقطع عرضی یک تابع BHPF را نشان می دهد.



شکل ۴. ۳۷. نمای پرسپکتیو و مقطع عرضی شعاعی مربوط به فیلتر بالاگذر ایده اال

1) Butterworth highpass filter (BHPF)

۴.۴ ارتقا در میدان فرکانس ۲۶۱

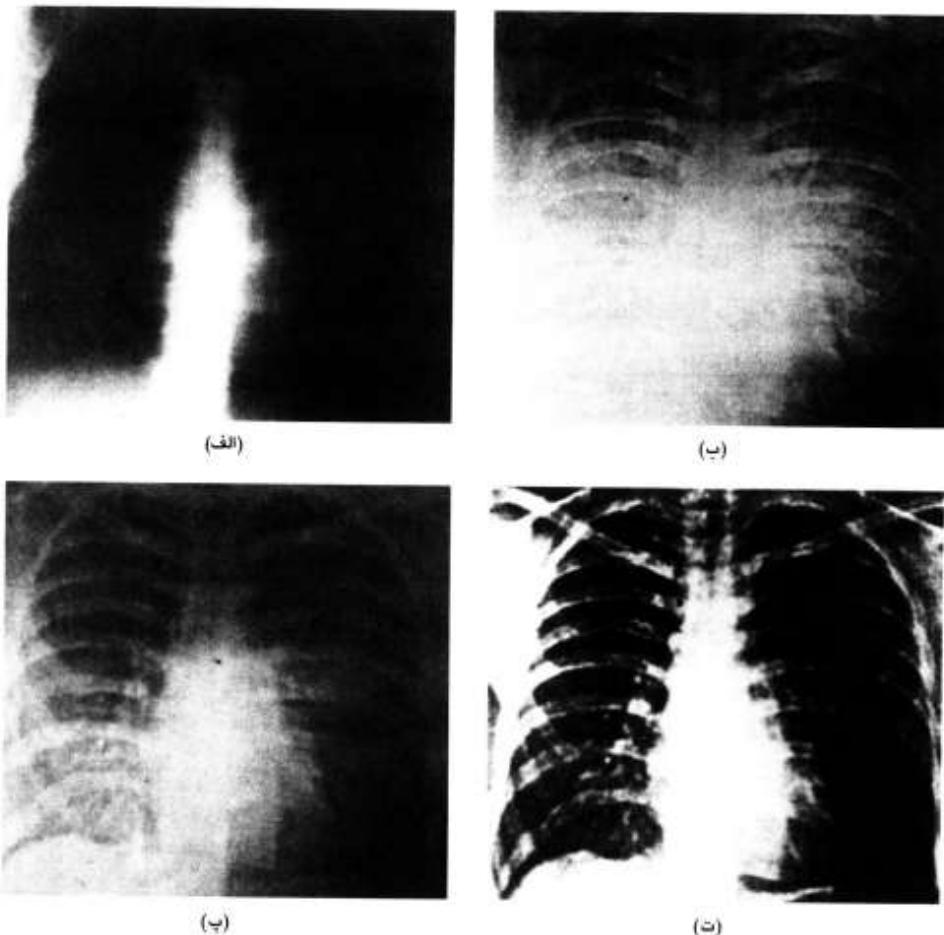
شکل ۴.۳۸ نمای پرسپکتیو و مقطع عرضی شعاعی فیلتر بالاگذر با ترورث بازای $I = \pi/2$.

توجه کنید که وقتی $D(u, v) = D_0$ باشد، $H(u, v) = 1/2$ مقدار بیشینه اش کاهش می‌یابد. مانند فیلتر پایین‌گذر با ترورث، روش عمومی این است که منحنی فرکانس قطع را در نقاطی انتخاب کنیم که در آنها $H(u, v)$ به $1/\sqrt{2}$ برابر مقدار بیشینه اش کاهش یافته است. برای برآوردن این شرط معادله (۴.۷-۴) به صورت زیر تغییر مقیاس می‌یابد:

$$\begin{aligned} H(u, v) &= \frac{1}{1 + [\sqrt{2} - 1][D_0/D(u, v)]^{2n}} \\ &= \frac{1}{1 + 0.414[D_0/D(u, v)]^{2n}}. \end{aligned} \quad (4.4-8)$$

مثال: شکل ۴.۳۹ (الف) یک تصویر اشعه ایکس از قضیه 'سینه را که با کیفیت پایینی تولید شده است، و شکل ۴.۳۹ (ب) آن تصویر را پس از این که با فیلتر بالاگذر مرتبه یک پردازش شده است، نشان می‌دهد. در این تصویر تنها لبه‌ها قابل توجه هستند زیرا مؤلفه‌های فرکانس پایین شدیداً تضعیف شدند، که باعث گردید نواحی با سطح خاکستری متفاوت (اما با تغییر آرام) کاملاً مانند هم دیده شوند. یک روش که اغلب برای کاهش این مشکل استفاده می‌شود، شامل جمع یک عدد ثابت با تابع انتقال فیلتر بالاگذر است تا بدین وسیله مؤلفه‌های فرکانس پایین حفظ گردد. البته این ثابت، مؤلفه‌های فرکانس بالا را نیز تقویت می‌کند. این روش که تأکید فرکانس بالا¹ نامیده می‌شود، در شکل ۴.۳۹ (ب) به نمایش گذاشده می‌شود. توجه کنید که در این مورد تصویر سطوح رنگی نسبتاً بهتری، بویژه در قسمت سمت چپ پایین عکس، دارا می‌باشد. این فرآیند مشابه فیلتر تقویت بالا که در بخش ۴.۳.۰ تشریح شد، می‌باشد.

1) High-frequency emphasis



شکل ۳۹.۴ مثالهایی از فیلتر کردن بالاگذر: (الف) تصویر اولیه؛ (ب) تصویر پردازش شده با یک فیلتر بالاگذر باترورث؛ (ب) حاصل تأکید فرکانس بالا؛ (ت) حاصل تأکید فرکانس بالا و تعدیل بافت نگار. (از Hall و همکاران [۱۹۷۱])

همان طور که با توجه به افزایش کوچک کیفیت از شکل ۳۹.۴ (ب) به شکل ۳۹.۴ (پ) دیده می شود، گرچه تأکید فرکانس بالا، مؤلفه های فرکانس پایین را حفظ می کند، اما اجزای با فرکانس نسبتاً بالاتر را مبهم می کنند. روشنی که اغلب جهت جبران این مشکل استفاده می شود، پردازش با پس-فیلتر کردن است که بوسیله آن سطوح خاکستری دوباره توزیع گردند. روشن تعدیل بافت نگار به دلیل توانایی اش در افزایش تمایز برای این منظور ایده اال است. شکل ۳۹.۴ (ت) بهبود قابل توجهی را نشان می دهد. این بهبود را می توان با تعدیل بافت نگار تصویری که تأکید فرکانس بالا

شده است، بدست آورد.

۳۰.۴.۴ فیلتر کردن همشکلی

مدل روشنایی- انعکاس معرفی شده در بخش ۲.۰ را می توان به عنوان مبنای برای معرفی یک روش جدید میدان فرکانس استفاده کرد. این روش برای بهبود ظاهر تصویر بوسیله 'انجام همزمان فشرده سازی محدوده' روشنایی و ارتقای تمایز مفید است. طبق توضیحات بخش ۲.۰، تصویر $f(x, y)$ را می توان بر حسب مؤلفه های روشنایی و انعکاس آن با رابطه'

$$f(x, y) = i(x, y)r(x, y). \quad (9-4.4)$$

بیان کرد. معادله '۹-۴.۴' را نمی توان مستقیماً برای ایجاد تغییر در مؤلفه های فرکانسی روشنایی و انعکاس به طور جداگانه، بکار برد، زیرا تبدیل فوریه 'حاصلضرب دوتابع' جداپذیر نیست؛ به عبارت دیگر،

$$\tilde{Y}\{f(x, y)\} \neq \tilde{Y}\{i(x, y)\}\tilde{Y}\{r(x, y)\}.$$

در هر حال فرض کنید که تعریف نماییم

$$\begin{aligned} z(x, y) &= \ln f(x, y) \\ &= \ln i(x, y) + \ln r(x, y). \end{aligned} \quad (10-4.4)$$

آنگاه،

$$\begin{aligned} \tilde{Y}\{z(x, y)\} &= \tilde{Y}\{\ln f(x, y)\} \\ &= \tilde{Y}\{\ln i(x, y)\} + \tilde{Y}\{\ln r(x, y)\} \end{aligned} \quad (11-4.4)$$

با این که

$$Z(u, v) = I(u, v) + R(u, v) \quad (12-4.4)$$

که $R(u, v)$ و $I(u, v)$ به ترتیب تبدیل فوریه های $\ln r(x, y)$ و $\ln i(x, y)$ هستند. اگر $Z(u, v)$ را بوسیله 'فیلتر پردازش نماییم، آنگاه از معادله '۴-۱.۱' داریم،

$$\begin{aligned} S(u, v) &= H(u, v)Z(u, v) \\ &= H(u, v)I(u, v) + H(u, v)R(u, v) \end{aligned} \quad (13-4. 4)$$

که $S(u, v)$ تبدیل فوریه' حاصل است. در میدان مکان،

$$\begin{aligned} s(x, y) &= \mathcal{F}^{-1}\{S(u, v)\} \\ &= \mathcal{F}^{-1}\{H(u, v)I(u, v)\} + \mathcal{F}^{-1}\{H(u, v)R(u, v)\}. \end{aligned} \quad (14-4. 4)$$

با فرض این که

$$i'(x, y) = \mathcal{F}^{-1}\{H(u, v)I(u, v)\} \quad (15-4. 4)$$

و

$$r'(x, y) = \mathcal{F}^{-1}\{H(u, v)R(u, v)\} \quad (16-4. 4)$$

می توان معادله (14-4. 4) را به صورت

$$s(x, y) = i'(x, y) + r'(x, y). \quad (17-4. 4)$$

بیان کرد. در نهایت وقتی با گرفتن لگاریتم از تصویر اولیه' $z(x, y) = f(x, y)$ ، $f(x, y)$ تشکیل شد، عمل معکوس، تصویر ارتقا یافته مطلوب $g(x, y)$ را می دهد؛ یعنی،

$$\begin{aligned} g(x, y) &= \exp[s(x, y)] \\ &= \exp[i'(x, y)] \cdot \exp[r'(x, y)] \\ &= i_0(x, y)r_0(x, y) \end{aligned} \quad (18-4. 4)$$

که

$$i_0(x, y) = \exp[i'(x, y)] \quad (19-4. 4)$$

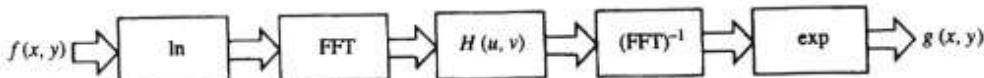
و

$$r_0(x, y) = \exp[r'(x, y)] \quad (20-4. 4)$$

مؤلفه های روشنایی و انعکاس تصویر خروجی هستند.

در شکل ۴. ۴۰ روش ارتقا بر اساس مفاهیم بیان شده در بالا، جمع بندی می شود. اساس این

۴۰.۴ ارتقا در میدان فرکانس ۲۶۵

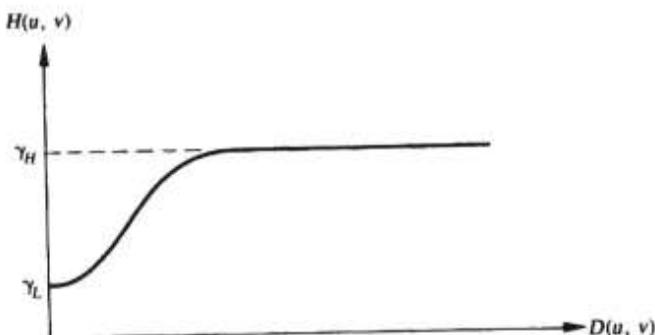


شکل ۴۰.۴ روش فیلتر کردن همشکلی جهت ارتقای تصویر

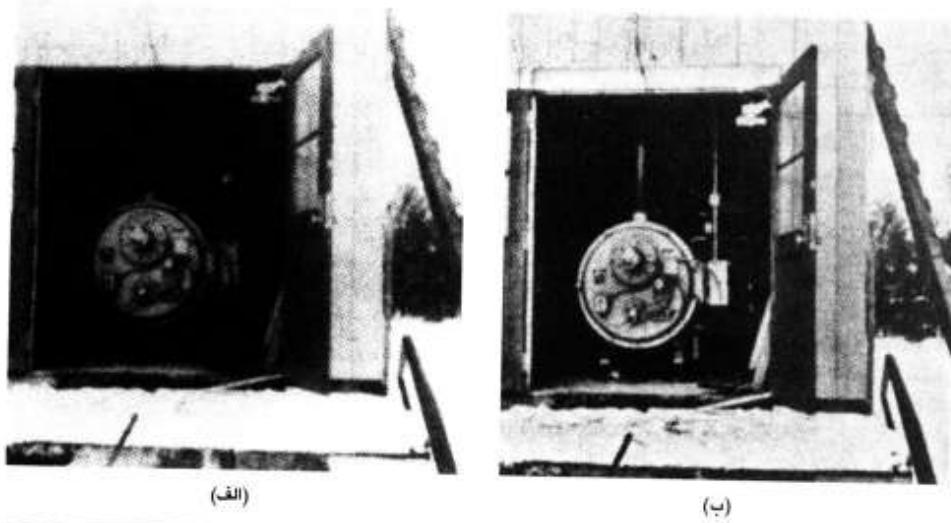
روش، مورد خاصی از یک دسته سامانه ها که سامانه های همشکلی^۱ نامیده می شوند، می باشد. در این کاربرد خاص، نکته مهم این است که جداسازی مؤلفه های روشنایی و انعکاس به صورت معادله^۲ (۴.۴-۱۲) بدست می آید. آن گاه همان طور که در معادله^۲ (۴.۴-۱۳) بیان شد، فیلتر همشکلی^۲ $H(u, v)$ می تواند به طور جداگانه روی هر یک از این مؤلفه ها عمل نماید.

مؤلفه^۲ روشنایی تصویر عموماً تغییرات مکانی آرامی دارد، در حالی که مؤلفه^۲ انعکاس دارای تغییرات سریع، بویژه در مرز اشیای غیر مشابه، است. این مشخصات باعث می شود که فرکانس های فرکانس های پایین تبدیل فوریه^۲ لگاریتم تصویر با روشنایی و فرکانس های بالای آن با انعکاس مرتبط شوند. گرچه این ارتباطات تقریب های نادقيقی هستند، می توان از آنها در ارتقای تصویر بهره برد.

با یک فیلتر همشکلی به میزان خوبی می توان کنترل مؤلفه های روشنایی و انعکاس را در دست داشت. این کنترل نیازمند تعیین تابع فیلتر $H(u, v)$ است آن چنان که اثرات متفاوتی روی مؤلفه های فرکانس پایین و فرکانس بالای تبدیل فوریه داشته باشد. شکل ۴۱.۴ مقطع عرضی چنین تابعی را نشان می دهد. با چرخاندن ° ۳۶۰ مقطع عرضی حول محور عمودی، $H(u, v)$ به طور کامل مشخص می گردد. اگر



شکل ۴۱.۴ مقطع عرضی تابع یک فیلتر متقارن چرخشی برای استفاده در فیلتر کردن همشکلی. $D(u, v)$ فاصله از مبدأ است.



شکل ۴۲.۳ (الف) تصویر اولیه؛ (ب) تصویر پردازش شده به وسیله فیلتر کردن همشکلی برای حصول همزمان فشرده سازی محدوده دینامیکی و ارتقای تمایز [از [Stockham ۱۹۷۲]] .

پارامترهای $γ_1$ و $γ_2$ طوری انتخاب شوند که $1 < γ_2 < γ_1$ باشد، تابع فیلتر موجود در شکل ۴۱.۴ فرکانس های پایین را کاهش داده و فرکانس های بالا را تقویت می نماید. نتیجه 'حالص' این عمل، انجام همزمان فشرده سازی محدوده دینامیکی و ارتقای تمایز می باشد.

مثال: شکل ۴۲.۴ از نوع نتایجی است که می توان با تابع همشکلی موجود در شکل ۴۱.۴ بدست آورد. در تصویر اولیه، شکل ۴۲.۴ (الف)، جزئیات درون اتاق براثر نور زننده 'منعکس شده از دیوارهای بیرونی و ضوح خود را از دست داده اند. شکل ۴۲.۴ (ب) حاصل پردازش این تصویر بوسیله 'فیلتر کردن همشکلی با تابع فیلتر فوق بازای $0.5 = γ_2$ و $2.0 = γ_1$ را نشان می دهد. با مقداری کاهش در محدوده 'دینامیکی روشنی، همراه با افزایشی در تمایز، جزئیات اشیای درون اتاق واضح شده و سطوح خاکستری دیوار بیرونی متعادل گردید.

□

۵.۴ تولید نقاب های مکانی از مشخصات میدان فرکانس

همان طور که در بخش ۴.۱ نیز بیان شد، سرعت و سادگی پیاده سازی، از ویژگی های مهم نقاب های مکانی پردازش تصویر هستند. در این بخش روشنی را برای تولید نقاب های مکانی که یک فیلتر میدان فرکانس داده شده را (در مفهوم حداقل مربعات خط) تقریب می زند، بیان می کنیم.

۴.۵. تولید نقاب های مکانی از مشخصات میدان فرکانس ۲۶۷

از بخش ۴.۱ به یاد آورید که فرآیند فیلتر کردن در میدان فرکانس براساس معادله'

$$G(u, v) = H(u, v)F(u, v) \quad (4-5.4)$$

می باشد که در آن $G(u, v)$ و $F(u, v)$ به ترتیب تبدیل فوریه های تصاویر ورودی و خروجی بوده و $H(u, v)$ تابع انتقال فیلتر است. از قضیه کانولوشن (بخش ۳.۸.۳) می دانیم که معادله (۴.۵.۱) را می توان در میدان مکان با عبارت

$$g(x, y) = \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} h(x - i, y - k) f(i, k) \quad (4-5.5)$$

بازای $x, y = 0, 1, 2, \dots, N-1$ پیاده سازی کرد. جهت سهولت بیان، فرض می شود که از آرایه های تصویری مربعی استفاده می کنیم. گذشته از آن همان طور که در بخش ۳.۸.۳ راجع به کانولوشن بیان شد، فرض می کنیم که تمام توابع به میزان مناسبی گسترش یافته اند.

در معادله (۴.۵.۲)، h نمایش مکانی فیلتر (یعنی تبدیل فوریه معکوس ($H(u, v)$)) تصویر ورودی و g تصویر فیلتر شده است. همان طور که در بخش ۴.۱ بیان شد، تمام h اغلب با عنوان نقاب کانولوشن مکانی^۱ اشاره می شود. اگر این نقاب به ابعاد $N \times N$ باشد، حاصل داده شده در معادله (۴.۵.۲)، با تبدیل فوریه^۱ معکوس $G(u, v)$ از معادله (۴.۵.۱) یکسان است.

وقتی H تبدیل فوریه^۱ باشد، از معادله (۳.۹-۲) نتیجه می شود که بازای $u, v = 0, 1, 2, \dots, N-1$ داریم:

$$H(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} h(x, y) \exp[-j2\pi(ux + vy)/N] \quad (4-5.6)$$

در هر حال فرض کنید که برای مقادیر $u < N$ و $v < N$ که $h(x, y)$ باشد، h محدود به صفر است. اثر این محدودیت ایجاد نقاب کانولوشن \hat{h} به ابعاد $n \times n$ و با تبدیل فوریه

$$\hat{H}(u, v) = \frac{1}{N} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} \hat{h}(x, y) \exp[-j2\pi(ux + vy)/N] \quad (4-5.7)$$

بازای $u, v = 0, 1, 2, \dots, N-1$ می باشد. در توضیح بعدی، هدف پیدا کردن ضرایب $\hat{h}(x, y)$ است

(1) Spatial convolution mask

طوری که معیار خطای

$$e^2 = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} |\hat{H}(u, v) - H(u, v)|^2 \quad (5-5. 4)$$

حداقل شود، که $|.|$ نشانه بزرگی کمیت مختلط است.

معادله (4-5. 4) را می توان به شکل ماتریسی

$$\hat{\mathbf{H}} = \mathbf{C}\hat{\mathbf{h}} \quad (6-5. 4)$$

یافتن کرد که $\hat{\mathbf{H}}$ بردارستونی از مرتبه N^2 حاوی عناصر $\hat{H}(u, v)$ با هر ترتیب دلخواه، $\hat{\mathbf{h}}$ بردارستونی از مرتبه n^2 حاوی عناصر $\hat{h}(x, y)$ با هر ترتیب دلخواه، و \mathbf{C} یک ماتریس $n^2 \times N^2$ از جملات نمایی است که محل آنها با ترتیب موجود در $\hat{\mathbf{H}}$ و $\hat{\mathbf{h}}$ تعیین می شود. یک روال ساده برای تولید عناصر

$\hat{H}(u, v)$ مربوط به بردار $\hat{\mathbf{H}}$ از ماتریس $\hat{H}(u, v)$ به صورت

$$\hat{H}(u, v) \Rightarrow \hat{H}(i) \quad (7-5. 4)$$

با زای $v = 0, 1, 2, \dots, N-1$ و $i = uN + v$ می باشد. حرکت در طول سطرهای $\hat{H}(u, v)$ با ترتیب $u = 1, v = 0, 1, 2, \dots, N-1$ ، $u = 1, v = 0, 1, 2, \dots, N-1$ ، $u = 0, 1, 2, \dots, N-1$ و الی آخر متناظر با تشکیل N عنصر اول $\hat{\mathbf{H}}$ از سطر اول $\hat{H}(u, v)$ ، N عنصر بعدی از سطر دوم و الی آخر می باشد. به طور مشابه عناصر $\hat{\mathbf{h}}$ که با $\hat{h}(k)$ ماتریس \mathbf{C} نشان داده می شوند، با نگاشت

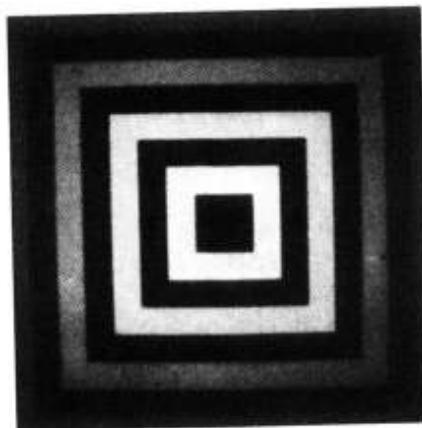
$$\hat{h}(x, y) \Rightarrow \hat{h}(k) \quad (8-5. 4)$$

با زای $x, y = 0, 1, 2, \dots, n-1$ و $k = xn + y$ می باشد. نهایتاً عناصر متناظر در ماتریس \mathbf{C} که با $C(i, k)$ نشان داده می شوند، از عبارات نمایی

$$\frac{1}{N} \exp[-j2\pi(ux + vy)/N] \Rightarrow C(i, k) \quad (9-5. 4)$$

با زای $u, v = 0, 1, 2, \dots, N-1$ و $k = xn + y$ می باشد. $x, y = 0, 1, 2, \dots, n-1$

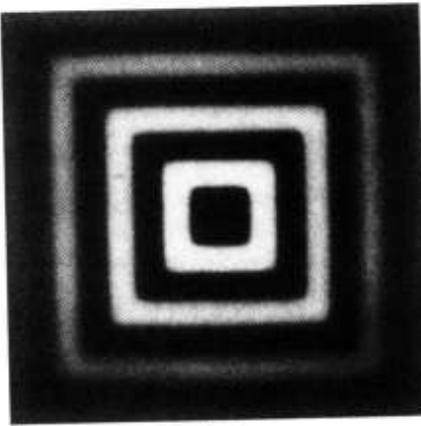
۴.۵. تولید نقاب های مکانی از مشخصات میدان فرکانس ۲۶۹



(الف)



(ب)



(پ)

شکل ۴.۳۰. (الف) تصویر اولیه؛ (ب) تصویر مات شده حاصل از یک فیلتر پایین گذر با ترورث مرتبه یک در میدان فرکانس؛ (پ) تصویر مات شده در میدان مکان بوسیله نقاب کانولوشن 9×9 حاصل از معادله (۴-۵). [۱۹۸۳] Gonzalez و Meyer از (۱۲-۵).

در نمایش ماتریسی، معادله (۴-۵) به صورت

$$\begin{aligned} e^2 &= (\hat{\mathbf{H}} - \mathbf{H})^* (\hat{\mathbf{H}} - \mathbf{H}) \\ &= \|\hat{\mathbf{H}} - \mathbf{H}\|^2 \\ &= \|\mathbf{Ch} - \mathbf{H}\|^2 \end{aligned} \quad (4-5)$$

می شود، که طبق توضیحات قبلی * نشانگر ترانهاده مزدوج، || نشانه نرم اقلیدسی مختلط، و \mathbf{H} برداری است که از $H(u, v)$ تشکیل شده است. با اگر قرن مشتق جزئی و برابر قرار دادن آن، با بردار

صفر کمینه^۱ e^2 بر حسب \hat{h} بدست می آید:

$$\frac{\partial e^2}{\partial \hat{h}} = 2C^*(\hat{h} - H) = 0 \quad (11-5. 4)$$

$$\begin{aligned} \hat{h} &= (C^*C)^{-1}C^*H \\ &= C^*H \end{aligned} \quad (12-5. 4)$$

که ماتریس $C^*C^{-1}C^*$ اغلب معکوس تعیین یافته^۲ (Noble [1969])¹ Moore-Penrose است. خوانده می شود.

معادله^۳ (12-5. 4) ضرایب حداقل خطای که برای تشکیل ماسک کانولوشن $(y, x)\hat{h}$ با ابعاد $n \times n$ از فیلتر مشخص شده^۴ $H(u, v)$ در میدان فرکانس لازم است، را بدست می دهد. در حالت کلی، عناصر $(y, x)\hat{h}$ کمیات مختلط هستند. با این حال اگر تابع انتقال فیلتر در میدان فرکانس، حقیقی و متقارن باشد (مانند تمام فیلترهای تشریح شده در این فصل)، $(y, x)\hat{h}$ نیز حقیقی و متقارن خواهد بود. مثال: به عنوان نمایشی از روشی که اخیراً بیان شد، الگوی آزمون موجود در شکل ۴ (الف) با استفاده از یک فیلتر پایین گذر با تورورث فیلتر شد تا تصویر مات شده^۵ موجود در شکل ۴ (ب) بدست آمد. یک ناقاب کانولوشن 9×9 با استفاده از معادله^۶ (12-5. 4) تولید گردید و بر تصویر اولیه اعمال شد. نتیجه ای که در شکل ۴ (پ) دیده می شود، کمی واضحتر از خروجی فیلتر کامل در میدان فرکانس است. این نتیجه مورد انتظار است، زیرا فرآیند میدان مکان با $N < n$ ، تنها تقریبی در مفهوم حداقل مربعات خطای است. سایر مثال های این روش در بخش ۵. ۸ ارائه می شوند. □

۴. ۴ پردازش تصویر رنگی

استفاده از رنگ در پردازش تصویر دو انگیزه^۷ اصلی دارد: اولاً در تحلیل تصویر خودکار، رنگ توصیف گر توانایی است که در اغلب موارد شناسایی و استخراج شیئی از صحته را ساده می سازد. ثانیاً در مواردی که تحلیل تصویر بوسیله^۸ انسان انجام می شود، چشم انسان قادر است هزاران سایه و شدت رنگی را، در مقایسه با حدود ۲۴ سایه خاکستری، تشخیص دهد.

پردازش تصویر رنگی به دو حوزه^۹ اصلی تقسیم می شود: پردازش تمام رنگی^{۱۰} و پردازش

1) Moore-Penrose generalized inverse

2) Full color processing

شبیه رنگی^۱. در گروه اول، تصاویر مورد نظر معمولاً بایک حسگر تمام رنگی نظیر دوربین تلویزیون رنگی یا پیمایش گر رنگی^۲ برداشته می شوند. در گروه دوم، به هر شدت تکریگ خاص یا محدوده ای از شدت هایک سایه رنگی متنسب می شود. تقریباً تا همین اواخر، بیشتر پردازش تصویر رنگی به صورت شبیه رنگی انجام می شد. پیشرفت قابل توجهی که در دهه ۱۹۸۰ انجام شد، باعث گردید حسگرهای رنگی و سخت افزار لازم برای پردازش تصویر رنگی با قیمت های قابل قبولی در دسترس قرار گیرند. در نتیجه این پیشرفت ها استفاده از روش های پردازش تصویر تمام رنگی در محدوده وسیعی از کاربردها در حال افزایش است.

۱.۶.۴ مبانی رنگ

گرچه فرآیند ادراک رنگ در مغز انسان، پدیده ای روانی- جسمی^۳ است که تاکنون به طور کامل شناخته نشده است، اما ماهیت فیزیکی رنگ را می توان براساس مجموعه ای از نتایج آزمایشی و نظری بیان کرد.

در سال ۱۶۶۶ میلادی اسحاق نیوتون کشف کرد که وقتی نور خورشید از میان منشور شیشه ای عبور می کند، خروجی آن، نور سفید نیست، بلکه شامل طیف پیوسته ای از رنگها در محدوده 'بنفش تا قرمز' می باشد. همان طور که پلیت^۴ I نشان می دهد، طیف رنگی را می توان به ۶ ناحیه وسیع تقسیم کرد: بنفش، آبی، سبز، زرد، نارنجی و قرمز. هیچ رنگی در طیف یکباره قطع نمی شود، بلکه هر رنگ به آرامی با رنگ بعدی مخلوط می گردد (پلیت II).

به طور کلی رنگ هایی که انسان از شبیه دریافت می کند، بوسیله 'ماهیت نور منعکس شده از آن شبیه معین می شود. همان طور که در پلیت II دیده می شود، نور مرئی نوار نسبتاً باریکی در طیف انرژی الکترو مغناطیسی می باشد. جسمی که در تمام طول موج های مرئی نور نسبتاً متوازنی را منعکس کند، برای ناظر سفید ظاهر می شود، اما جسمی که تنها در محدوده 'کوچکی از طیف مرئی به خوبی انعکاس داشته باشد، برای ناظر رنگی به نظر می رسد. مثلاً اشیای سبز، طول موجهای نوری در محدوده ۵۰۰ تا ۵۷۰ نانومتر (m^{-9}) را منعکس می کنند، در حالی که بیشتر انرژی موجود در سایر طول موجهای را جذب می کنند.

1) Pseudo-color processing

2) Color scanner

3) Physiosychological

4) Plate

تصویف نور، اساس علم رنگ است. اگر نور غیررنگی (بدون رنگ) باشد، تنها مشخصه آن شدت^۱ یا مقدارش می باشد. نور غیررنگی همانند نوری است که بینندگان روی تلویزیون سیاه و سفید می بینند و تاکنون نیز به طور ضمیمی موضوع بحث ما درباره 'پردازش تصویر بوده است. بنابراین اصطلاح سطح خاکستری^۲ به اندازه عددي شدت که در محدوده 'سیاه تا خاکستری ها و در نهایت تا سفید تغییر می کند، اشاره دارد.

نور رنگی طیف انرژی الکترومغناطیسی تقریباً از ۴۰۰ تا ۷۰۰ نانومتر را می پوشاند. سه کمیت اصلی برای توصیف کیفیت منبع نور رنگی استفاده می شوند: تشعشع^۳، لومنانس^۴، و روشنی^۵. تشعشع مقدار کل انرژی است که از منبع نور صادر گردیده و اغلب با وات (W) سنجیده می شود. لومنانس که بر حسب لومن^۶ (lm) اندازه گرفته می شود، مقدار انرژی است که ناظر از منبع نور دریافت می کند. مثلاً نوری که از منبعی در ناحیه انتهایی مادون قرمز طیف صادر می شود، ممکن است انرژی (تشعشع) قابل ملاحظه ای داشته باشد اما ناظر به سختی آن را احساس می کند؛ بنابراین لومنانس آن تقریباً صفر است. در نهایت روشنی توصیف گری ذهنی است که اندازه گیری آن عملای غیرممکن است. روشنی بخش غیررنگی شدت را مجسم می کند و یکی از عوامل مهم در توصیف احساس رنگ است.

بدلیل ساختمان چشم انسان، تمام رنگ ها به صورت ترکیبات مختلف سه رنگ اولیه 'قرمز (R)، سبز (G) و آبی (B) دیده می شوند. به منظور استانداردسازی، CIE^۷ (جمع جهانی درباره روشنایی) در سال ۱۹۳۱ مقادیر طول موج های مشخص زیر را به سه رنگ اولیه متسب کرد: آبی = ۴۳۵/۸ nm، سبز = ۵۴۶ nm، قرمز = ۵۷۰ nm. توجه کنید که طبق پلیت II، هیچ رنگی را به تنهایی نمی توان قرمز، سبز، یا آبی نامید. بنابراین داشتن سه طول موج رنگی مشخص و استاندارد بدین معنا نیست که سه مؤلفه ثابت RGB به تنهایی می توانند تمام رنگ های طیف را تولید نمایند. این موضوع نکته مهمی است، زیرا در بسیاری موارد به هنگام استفاده از کلمه 'اولیه^۸، این گمان رفته است که اگر سه رنگ اولیه 'استاندارد در نسبت های مختلف با هم مخلوط شوند، می توانند تمام رنگ های مرئی را تولید نمایند. این تعبیر درست نیست مگر این که طول موج هم امکان تغییر داشته باشد.

با مخلوط کردن رنگ های اولیه می توان رنگهای ثانویه^۹ بنفس روشن (قرمز به اضافه 'آبی)، آبی

1) Intensity	2) Gray level	3) Radiance	4) Luminance	5) Brightness	6) Lumen
7) Commission Internationale de L'Eclairage		8) Primary		9) Secondary	

فیروزه ای (سبز به اضافه آبی) ، وزرد (قرمز به اضافه سبز) را تولید نمود . با مخلوط کردن سه رنگ اولیه با هم ، یا یک رنگ ثانویه با اولیه متصادش ، البته با شدت های صحیح ، نور سفید تولید می شود . این نتیجه در پلیت III (الف) دیده می شود . همچنین این پلیت سه رنگ اولیه و ترکیبات آنها برای تولید رنگ های ثانویه را نمایش می دهد .

تفکیک بین رنگ های اولیه نوری و رنگ های اولیه مادی^۱ یا کلرانت ها^۲ مهم است . در مورد اخیر رنگ اولیه به عنوان رنگی تعریف می شود که رنگ اولیه نوری را جذب یا تفرقی می کند و دور رنگ دیگر را انعکاس یا عبور می دهد . بنابراین رنگ های اولیه مادی بنفس ، آبی فیروزه ای ، وزرد بوده و رنگ های ثانویه مادی قرمز ، سبز و آبی هستند . این رنگ ها در پلیت III (ب) دیده می شوند . ترکیب مناسب سه رنگ اولیه مادی ، یا یک ثانویه با اولیه متصادش ، رنگ سیاه تولید می کند .

تصویر تلویزیون رنگی مثال مناسبی از ترکیب جمعی رنگ های نوری است . بسیاری از لامپ های تلویزیون رنگی از آرایه بزرگی از الگوهای نقطه ای مثلثی از جنس فسفر حساس به الکترون تشکیل شده اند . هر نقطه در مثلث^۳ ، اگر تحریک شود یکی از رنگ های اولیه را تولید می نماید . شدت نور نقاط فسفری مولد نور قرمز بوسیله تفنجک الکترونی داخل لامپ مadolه می شود . این تفنجک تکانه هایی متناظر با "توان نور قرمز" که توسط دوربین تلویزیونی احساس می شود ، تولید می کند . نقاط فسفر سبز و آبی در هر تریاد نیز به همین روش مadolه می شوند . تصویری که روی گیرنده تلویزیون دیده می شود ، حاصل فرآیند مخلوط شدن سه رنگ اولیه تولید شده با هر مثلث فسفری ، دریافت نور حاصل بوسیله مخروط های حساس به رنگ در چشم ، و در نهایت درک تصویر تمام رنگی می باشد . مشاهده^۴ تصویر متواالی در هر ثانیه در سه رنگ ، تصویر پیوستگی نمایش تصویر روی صفحه را کامل می کند .

مشخصاتی که برای تشخیص یک رنگ از سایر رنگ هاییشت استفاده می شوند ، روشنی^۵ ، اصل رنگ^۶ و اشباع^۷ هستند . همان طور که قبلاً بیان شد ، روشنی بخش رنگی شدت را مجسم می کند . اصل رنگ صفت مرتبط با طول موج غالب در ترکیب امواج نوری است . بنابراین اصل رنگ یا بانگر رنگ غالب که بیننده دریافت می کند ، است ! وقتی شبیه را قرمز ، نارنجی یا زرد می خوانیم ، در واقع اصل رنگ آن را مشخص می کنیم . اشباع به خلوص نسبی یا مقدار نور سفید مخلوط با اصل رنگ مربوط است . رنگ های

1) Pigment

2) Colorant

3) Triad

4) Brightness

5) Hue

6) Saturation

طیفی خالص کاملاً اشباع شده هستند. رنگ هایی نظیر صورتی (قرمز به اضافه سفید) و بنفش کمرنگ^۱ (بنفش به اضافه سفید) کم اشباع تر هستند. درجه اشباع با مقدار نور سفید اضافه شده تناسب معکوس دارد.

اصل رنگ و اشباع روی هم، رنگینگی نامیده می شوند، و بنابراین هر رنگ را می توان با روشنی و رنگینگی آن توصیف کرد. مقادیر قرمز، سبز و آبی مورد نیاز برای تشکیل یک رنگ خاص، مقادیر حرکه سه گانه^۲ نامیده شده و به ترتیب با X , Y و Z نشان داده می شوند. هر رنگ با ضرایب سه گانه رنگینگی^۳ آن که به صورت

$$x = \frac{X}{X + Y + Z} \quad (1-6.4)$$

$$y = \frac{Y}{X + Y + Z} \quad (2-6.4)$$

و

$$z = \frac{Z}{X + Y + Z} \quad (3-6.4)$$

تعریف می شوند، مشخص می گردد. به وضوح از این معادلات نتیجه می شود که

$$x + y + z = 1 \quad (4-6.4)$$

برای هر طول موج نور در طیف مرئی می توان مقادیر حرکه سه گانه مورد نیاز برای تولید رنگ متناظر با آن طول موج را به طور مستقیم از روی منحنی های جداولی که نتیجه آزمایشات گسترده هستند، بدست آورد (Walsh[1959], Kiver[1965]).

روش دیگر برای توصیف رنگ ها، نمودار رنگینگی^۵ (پلت IV) است، که ترکیب رنگ را به صورت تابعی از x (قرمز) و y (سبز) نشان می دهد. برای هر مقدار x و y ، مقدار z (آبی) متناظر طبق معادله^۶ $z = 1 - x - y$ بودست می آید. برای مثال نقطه ای در پلت IV که به عنوان سبز مشخص شده است، حاوی تقریباً ۶۲ درصد سبز و ۲۵ درصد قرمز می باشد. بنابراین طبق معادله^۷ $z = 1 - 0.62 - 0.25 = 0.13$ درصد است.

1) Lavender

2) Chromaticity

3) Tristimulus

4) Trichromatic coefficients

5) Chromaticity diagram

محل رنگ های گوناگون طیف - از بنفس در ۳۸۰ نانومتر تا قرمز در ۷۸۰ نانومتر - حول مرز زبانه ای شکل نمودار رنگینگی نشان داده شده اند. اینها رنگ های خالصی هستند که در طیف پلیت II دیده می شوند. هر نقطه که روی مرز نباشد اما درون نمودار باشد، بیانگر مخلوطی از رنگ های طیف است. در نقطه 'تساوی انرژی ها'^۱ که در پلیت IV دیده می شود، سهم سه رنگ اولیه با هم برابر است. این نقطه در استاندارد CIE نمایانگر نور سفید است. هر نقطه ای که روی مرز نمودار رنگینگی باشد، کاملاً اشباع شده خوانده می شود. هرچه از مرز دور شده و به نقطه 'تساوی انرژی ها' نزدیک شویم، نور سفید بیشتری به رنگ افزوده شده و اشباع کمتر می شود. اشباع در نقطه 'تساوی انرژی ها' (نور سفید)، صفر است.

نمودار رنگینگی برای ترکیب رنگ های نیز مفید است، زیرا هر پاره خط راست که دو نقطه 'دلخواه در نمودار را به هم متصل کند، تمام رنگ های متفاوتی که می توان با ترکیب جمعی آن دو رنگ بدست آورد، را شامل می شود. برای مثال خط راستی را که از نقطه 'قرمز' به نقطه 'سبز' کشیده شده است و در پلیت IV دیده می شود، در نظر بگیرید. اگر در رنگی نسبت نور قرمز به سبز بیشتر باشد، نقطه 'متناظر با آن رنگ' روی پاره خط خواهد بود، اما به نقطه 'قرمز' نزدیک تر است. به طور مشابه، خطی که از نقطه 'تساوی انرژی ها' به نقطه 'دلخواهی' روی مرز نمودار کشیده شود، تمام سایه های آن رنگ طیفی خاص را شامل می شود.

تعییم این روال به سه رنگ، ساده است. برای تعیین محدوده 'رنگ هایی' که می توان از هر سه رنگ معین شده در نمودار رنگینگی، بدست آورد، کافی است خطوطی بین آن سه نقطه رسم کنیم. حاصل یک مثلث است و هر رنگ درون این مثلث را می توان با ترکیب جمعی سه رنگ اولیه تولید کرد. یک مثلث دلخواه که رئوس آن سه نقطه 'رنگی ثابت'^۲ در پلیت IV باشند، نمی تواند تمام ناحیه 'رنگی' را پوشاند. مشاهده این موضوع به طور گرافیکی نکته 'از پیش گفته شده ای را که نمی توان تنها با سه رنگ اولیه تمام رنگها را بدست آورد، تأیید می کند.

۴.۶.۴ مدل های رنگ

هدف از انتخاب مدل رنگ، تسهیل مشخص سازی رنگ ها در یک استاندارد است که معمولاً روش مورد

1) Point of equal energies

2) Fixed

قابلی می باشد. در اصل، مدل رنگ، تعیین یک سامانه مختصات سه بعدی و زیرفضایی درون آن سامانه است که در آن سامانه هر رنگ تنها با یک نقطه بیان می شود.

بیشتر مدل های رنگ که اکنون استفاده می شوند، به سمت سخت افزار (مانند نمایشگرهای چاپگر های رنگی) یا کاربردهایی گرایش دارند که هدف آنها کار با رنگ های می باشد (نظیر تولید گرافیک های رنگی برای پرینت) . عمومی ترین مدل های سخت افزار گرای اعبارند از: مدل RGB (قرمز، سبز، آبی) برای نمایشگرهای رنگی و یک گروه وسیع از دوربین های ویدیویی؛ مدل CMY (آبی فیروزه ای، بنفش، زرد) برای چاپگر های رنگی؛ و مدل YIQ که یک استاندارد پخش تلویزیون رنگی است. در مدل سوم، Y (بنفش)، I (قرمز) و Q (دو مؤلفه رنگ هستند که به ترتیب هم فاز^۱ و متعادل^۲ خوانده می شوند. متناظر با لومینانس است، و این دو مؤلفه رنگ هستند که به ترتیب هم فاز^۳ و متعادل^۴ خوانده می شوند. مدل HSI (اصل رنگ، اشباع، شدت) و مدل HSV (اصل رنگ، اشباع، مقدار) از مدل هایی هستند که به طور مکرر برای کار با تصویر رنگی استفاده می شوند.

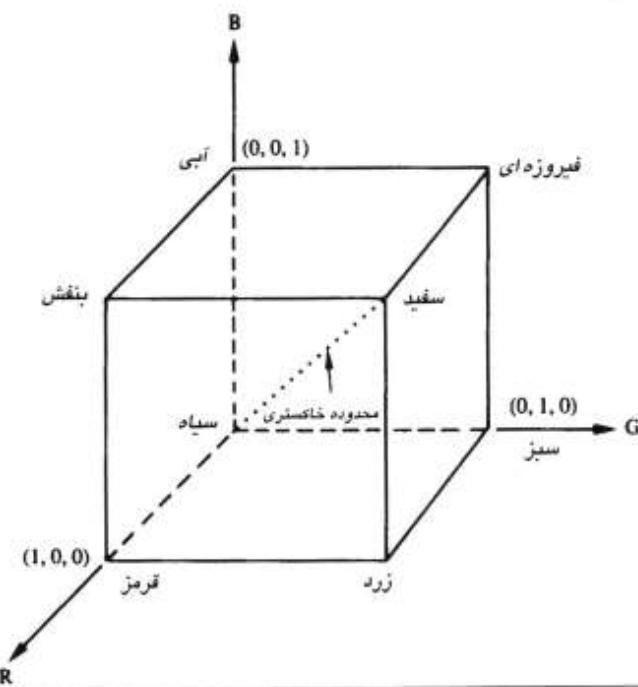
مدل های رنگی RGB، YIQ و HSI^۵ متدائلترین مدل هایی هستند که برای پردازش تصویر استفاده می شوند. در بخش های بعدی، ویژگی های اصلی این سه مدل را معرفی کرده، تفاوت ها و فواید آنها را در کاربردهای پردازش تصویر رقمه شرح می دهیم. گرچه مدل CMY، به جای استفاده در پردازش تصویر واقعی، در چاپ استفاده می شود، به دلیل اهمیت آن در بدست آوردن خروجی های چاپی در اینجا نیز بررسی می گردد.

مدل رنگ RGB

در مدل RGB هر رنگ به صورت مؤلفه های طیفی اولیه قرمز، سبز و آبی ظاهر می شود. این مدل بر اساس سامانه مختصات کارتزین است. زیرفضای رنگی مورد علاقه، مکعب شکل ۴۴.۴ می باشد که در آن مقادیر RGB در سه گوشه؛ آبی فیروزه ای، بنفش و زرد در سه گوشه دیگر؛ سیاه در مبدأ؛ و سفید در دورترین گوشه از مبدأ قرار دارد. در این مدل محدوده خاکستری از سیاه تا سفید در طول خط واصل این دو نقطه قرار دارد و سایر رنگ های نیز نقاطی روی یا درون مکعب هستند که با بردارهایی که از مبدأ

^۱ گرچه برای یک دست بودن با بیشتر مtron، در هر دو مدل YIQ و HSI، از ااستفاده می کنیم، اما دقت کنید که این نساد معنایهای بسیار متفاوتی در این دو مدل دارد.

۴۰۶. پردازش تصویر رنگی ۲۷۷



شکل ۴۰۶ مکعب رنگی RGB، نقاط در طول قطر اصلی، مقادیر خاکستری از سیاه در مبدأ تا سفید در نقطه $(1, 1, 1)$ دارند

می گذرند، تعریف می شوند. برای تسهیل مدل، فرض بر این است که تمام مقادیر رنگ تراز شده اند طوری که مکعب شکل ۴۰۶ مکعب واحد باشد. یعنی فرض می شود که تمام مقادیر R, G و B در محدوده $[0, 1]$ قرار دارند.

هر تصویر در مدل رنگ RGB سه صفحه 'تصویر مستقل'، هر صفحه برای یک رنگ اولیه، دارد. وقتی این سه صفحه به نمایشگر RGB داده شوند، روی صفحه 'فسفری' ترکیب می شوند تا یک تصویر رنگی را تولید نمایند. بنابراین وقتی خود تصاویر به طور طبیعی بر حسب سه صفحه 'رنگی' بیان شوند، استفاده از مدل RGB جهت پردازش تصویر معنا پیدا می کند. همچنین اغلب دوربین های رنگی که برای تصویربرداری رقمی به کار می روند، از قالب RGB استفاده می کنند که این موضوع خود به تنها مدل RGB را مدل مهمی در مبحث پردازش تصویر می سازد.

یکی از بهترین مثال های کاربرد مدل RGB، پردازش داده های تصویری چند طیفی هوایی یا ماهواره ای است. تصاویر بوسیله 'حسگرهای تصویربرداری' که در محدوده های طیفی مختلف کار

می کنند، گرفته می شوند. برای نمونه، هر قاب خروجی تصویربردار LANDSAT[†] دارای چهار تصویر رقemi است. همه تصویرها از یک صحته هستند که در محدوده یا پنجره^۱ طبیعی متفاوتی گرفته می شوند. دو پنجره از چهار پنجره^۲ فوق در بخش مرئی طیف هستند که تقریباً متناظر با سبز و قرمز می باشند؛ دو پنجره^۳ دیگر در بخش مادون قرمز طیف هستند. بنابراین هر صفحه تصویر^۴ معنای فیزیکی دارد، و ترکیبات رنگی که با استفاده از مدل RGB برای پردازش و نمایش بدست می آید، معمولاً وقتی معنا پیدا می کنند که روی یک صفحه رنگی دیده شوند، یا مانند فصل ۷، هنگام بخش بندی تصویر رنگی براساس مؤلفه های طبیعی آن معنا دارند.

فرض کنید که مسأله ارتقای تصویر رنگی صورت انسان که بخشی از آن در سایه مخفی است، باشد. همان طور که در بخش ۲.۴ بیان شد، تعديل بافت نگار ابزار ایده آلی برای حل این نوع مسائل است. بدلیل وجود سه تصویر و به دلیل آن که تعديل بافت نگار تنها با مقادیر شدت کار می کند، روال این است که هر یک از صفحه تصویرها به طور مستقل تحت تعديل بافت نگار قرار گیرد. به احتمال زیاد، بخشی از تصویر که در سایه مخفی است، ارتقا می یابد. در هر حال شدت ها در سه صفحه تصویر به طور متفاوتی تغییر داده می شوند که این عمل باعث تغییر شدت های نسبی آنها می شود. نتیجه^۵ تهابی این خواهد بود که خواص رنگی مهم در تصویر، نظیر تنهای نرم^۶، روی نمایشگر RGB به طور طبیعی ظاهر نمی شوند. بعضی از مدل های رنگ مشروح در بخش های بعدی برای چنین مسائلی مفیدتر هستند.

مدل رنگ CMY

همان طور که قبلاً بیان شد، فیروزه ای، بنفش و زرد رنگ های ثانویه^۷ نوری یا رنگهای اولیه^۸ مادی هستند. برای مثال، وقتی بر سطح پوشیده از ماده^۹ رنگی فیروزه ای، نور سفید تاییده شود، هیچ نور قرمزی از آن منعکس نمی شود. یعنی فیروزه ای نور قرمز را از نور سفید تفرق می کند. بیشتر وسایلی که مواد رنگی را بر روی کاغذ می نشانند، نظیر چاپگرها و کپی بردارهای^{۱۰} رنگی، به داده های CMY نیاز دارند، یا این که در داخل خود داده های RGB را به CMY تبدیل می کنند. این

[†] مخفف عبارت Land Sattellite است که این نام را NASA به ماهواره هایی داده است که جهت نظارت سطح زمین ساخته می شوند.

1) Window

2) Image plane

3) Flesh tone

4) Copier

تبديل با استفاده از عمل ساده^{*}

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5-6. 4)$$

انجام می شود. دوباره فرض بر این است که تمام مقادیر رنگی به محدوده [0, 1] تراز شده اند.

معادله ۴. ۶-۵) بیان می کند که نور منعکس شده از سطح پوشیده با ماده 'فیروزه ای خالص، حاوی نور قرمز نیست (زیرا طبق معادله داریم، $C = 1 - R$)؛ به طور مشابه، بنفش خالص، سبز را وزرد خالص، آبی را منعکس نمی کند. همچنین معادله ۴. ۶-۵) بیان می کند که می توان با تفریق تک تک مقادیر CMY از "۱" مقادیر RGB را بدست آورد. همان طور که قبلاً بیان شد، در پردازش تصویر این مدل در ارتباط با تولید خروجی چاپی استفاده می شود؛ بنابراین عموماً تبدیل معکوس از CMY به RGB کاربردی ندارد.

مدل رنگ YIQ

مدل YIQ در پخش عمومی تلویزیون رنگی تجاری استفاده می شود. در واقع YIQ تغییر شکل RGB به منظور افزایش بازده انتقال و حفظ سازگاری با استانداردهای تلویزیون تکرنسگ می باشد. در واقع مؤلفه 'Y در سامانه YIQ، تمام اطلاعات ویدیویی مورد نیاز تلویزیون تکرنسگ را دارد. تبدیل RGB به YIQ صورت رابطه^{*}

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (6-6. 4)$$

تعريف می شود. جهت بدست آوردن مقادیر RGB از مقادیر YIQ، بسادگی عمل ماتریسی معکوس را انجام می دهیم.

مدل YIQ برای بهره گیری از حساسیت بیشتر سامانه بینایی انسان به تغییرات لومینانس نسبت به تغییرات اصل رنگ یا اشباع طراحی گردید. بنابراین استاندارد YIQ به عرض باند (یا در حالت رقمی، بیت های) بیشتری برای نمایش Y و عرض باند (یا بیت های) کمتری برای نمایش I و Q نیاز دارد. برای

جزئیات بیشتر در مورد این خاصیت مهم، به Smith[1978] و Pritchard[1977] مراجعه کنید.

علاوه بر این که مدل YIQ استاندارد پرکاربردی است، مزیت اصلی آن در پردازش تصویر ناشی از این خاصیت است که اطلاعات لومینانس (L) و اطلاعات رنگی (I و Q) ناهمبسته^۱ هستند. توجه دارید که لومینانس متناسب با مقدار نور دریافتی چشم است. بنابراین اهمیت این ناهمبستگی در این است که مؤلفه لومینانس تصویر را می‌توان بدون اثرگذاری بر محتواهای رنگ پردازش کرد. برای نمونه، برخلاف مشکل مدل RGB که قبلًا ذکر شد، می‌توان تصویر رنگی با قالب YIQ را بسادگی با کاربرد تبدیل بافت نگار بر مؤلفه L آن، تبدیل هیستوگرام نمود. با این فرآیند رنگ‌های نسبی تصویر متأثر نمی‌شوند.

مدل رنگ HSI

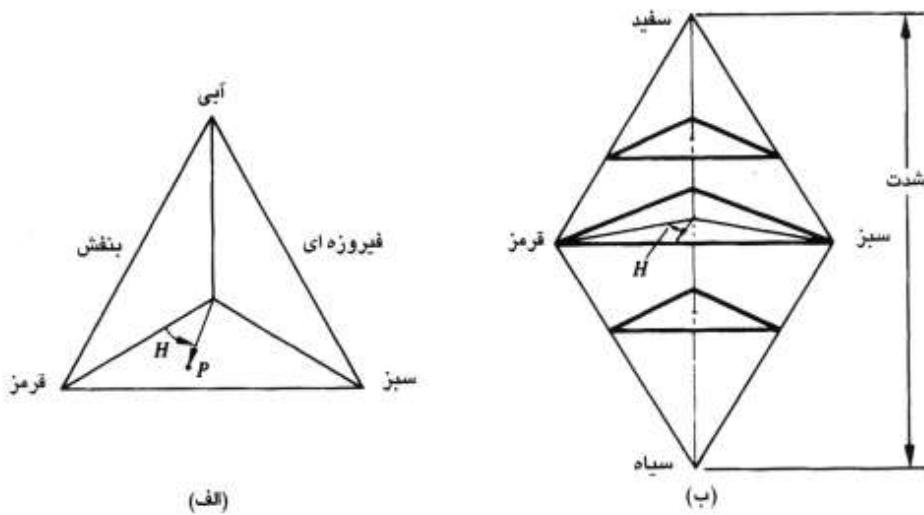
طبق توضیحات بخش ۴.۶.۱، اصل رنگ صفتی از رنگ خالص را توصیف می‌کند (زرد، نارنجی یا قرمز خالص)، در حالی که اشباع میزانی که یک رنگ خالص با نور سفید ترقیق شده است، را می‌دهد. سودمندی مدل رنگ HSI مدیون دو عامل اصلی است. اولاً، مؤلفه شدت، I، از اطلاعات رنگ تصویر مجزا است. ثانیاً، مؤلفه‌های اصل رنگ و اشباع رابطه زدیکی با روش دریافت رنگ توسط انسان دارند. این ویژگی‌ها مدل HSI را به ابزار ایده‌آلی برای تولید الگوریتمهای پردازش تصویر که مبتنی بر بعضی خواص احساس رنگ سامانه بینایی انسان هستند، تبدیل می‌کند.

کاربردهای مدل HSI، محدوده‌ای از طراحی سامانه تصویرپردازی برای تعیین خودکار رسیده بودن میوه‌ها و سبزیها گرفته تا سامانه‌هایی برای تطبیق نمونه‌های رنگ یا بازرسی نهایی کیفیت محصولات رنگی را شامل می‌شود. در این کاربردها و مشابه آنها، مهم این است که کار سامانه براساس آن دسته‌ای از خواص رنگی که سامانه بینایی انسان برای انجام کار مورد نظر از آنها استفاده می‌کند، طراحی شود. روابط تبدیل برای رفتن از فضای RGB به HSI و بر عکس به طور قابل ملاحظه‌ای پیچیده‌تر از تبدیل مدل‌های قبلی هستند. بنابراین به جای این که تنها این روابط را ذکر کنیم، تلاش می‌کنیم تا با به دست آوردن آنها، فهم عمیق‌تری در مورد تغییر رنگ به خواننده بدheim.

تبدیل از RGB به HSI. همان‌طور که قبلًا بیان شد، مدل RGB بر حسب مکعب واحد تعریف می‌شود. در هر حال، مؤلفه‌های رنگی مدل HSI (اصل رنگ و اشباع) بر حسب مثلث رنگ موجود در

(1) Decoupled

۶.۴ پردازش تصویر رنگی ۲۸۱



شکل ۴.۳۵ (الف) مثلث رنگی HSI؛ (ب) هرم گونه رنگی HSI.

شکل ۴.۴۵ (الف) تعریف می شوند. (از توضیحات موجود در بخش ۴.۶.۱ در مورد نمودار رنگینگی به یاد آورید که تمام رنگ های قابل حصول با ترکیب سه رنگ داده شده، درون مثلثی قرار دارند که رتوشن سه رنگ ابتدایی هستند.) توجه کنید که در شکل ۴.۴۵ (الف)، مؤلفه اصل رنگ H ، مربوط به نقطه رنگ P برابر با زاویه بین بردار نقطه رنگ P و محور قرمز است. بنابراین وقتی $0^\circ = H$ باشد، اصل رنگ قرمز است، وقتی $60^\circ = H$ باشد، اصل رنگ زرد است و الی آخر. S ، اشباع نقطه رنگ P ، میزانی است که هنوز به رنگ مربوطه، سفید اضافه نشده است و متناسب با فاصله P تا مرکز مثلث می باشد. هرچه P دورتر از مرکز مثلث باشد، رنگ آن بیشتر اشباع شده است.

در مدل HSI، شدت بر حسب خطی که بر مثلث عمود است و از مرکزش می گذرد، سنجیده می شود. شدت هایی که در طول این خط و زیر مثلث قرار دارند، از تیره به سیاه تغییر می کنند. بر عکس شدت های بالای مثلث از روشن به سفید تغییر می کنند.

ترکیب اصل رنگ، اشباع و شدت در فضای رنگی سه بعدی، ساختار سه وجهی هرم گونه موجود در شکل ۴.۴۵ (ب) را بدست می دهد. هر نقطه روی سطح این ساختار، نمایانگر یک رنگ کاملاً اشباع شده می باشد. اصل رنگ بر حسب زاویه آن رنگ نسبت به محور قرمز، و شدت رنگ بر حسب فاصله عمودی آن از نقطه سیاه تعیین می شود (یعنی هرچه فاصله از سیاه بیشتر باشد، مقدار شدت رنگ بزرگتر است). در مورد نقاط درون ساختار نیز همین توضیحات صدق می کند، تنها تفاوت این است که وقتی

رنگ ها به محور عمودی نزدیک تر شوند، میزان اشباع آنها کمتر می شود.

رنگ ها در مدل HSI نسبت به مقادیر قرمز، سبز و آبی تراز شده تعریف می شوند و بر حسب مؤلفه های

اولیه RGB با روابط

$$r = \frac{R}{(R + G + B)} \quad (7-6. ۴)$$

$$g = \frac{G}{(R + G + B)} \quad (8-6. ۴)$$

و

$$b = \frac{B}{(R + G + B)} \quad (9-6. ۴)$$

بدست می آیند. مانند گذشته فرض براین است که R ، G و B طوری تراز شده اند که در محدوده [0, 1] قرار گیرند. معادلات (4. ۶-۷) - (4. ۶-۹) نشان می دهند که r ، g و b نیز در بازه [0, 1] قرار دارند و

$$r + g + b = 1. \quad (10-6. ۴)$$

توجه کنید که وقتی R و B هر سه برابر "1" باشند، باز هم متغیرهای تراز شده باید در معادله (4. ۶-۱۰) صدق کنند. در واقع معادله فوق، معادله 'صفحه' شامل مثلث HSI می باشد. برای هر سه مؤلفه رنگ R ، G و B که در محدوده [0, 1] باشند، مؤلفه شدت در مدل HSI به صورت

$$I = \frac{1}{3}(R + G + B) \quad (11-6. ۴)$$

تعریف می شود که مقادیری در محدوده [0, 1] دارد.

مرحله' بعد به دست آوردن H و S است. به دست آوردن H به ساختار هندسی مثلث HSI که در اشکال

۴. ۶-۴ (الف)، (ب) و (پ) دیده می شود، نیاز دارد. در مورد آن اشکال شرایط زیر را ذکر می کنیم:

(الف) نقطه' W دارای مختصات $(1/3, 1/3, 1/3)$ می باشد.

(ب) هر نقطه' رنگ دلخواه P مختصات (r, g, b) دارد.

(پ) برداری که از مبدأ به W وصل شود، با w نشان داده می شود. به طور مشابه، بردارهایی که از مبدأ به P_R و P_g کشیده شوند، به ترتیب با p_R و p_g نشان داده می شوند.

۴۸۳ ۶. پردازش تصویر رنگی

(ت) خطوط $i = R, G, B, P_i Q_i$ در W با هم تلاقی می کنند.

(ث) با فرض $r_0 = R/I$ ، $b_0 = G/I$ و $g_0 = B/I$ که طبق معادله (۱۱-۶) به دست می آید، در شکل ۴.۴۶ (الف) می بینیم که مکان هندسی نقاط (r_0, g_0, b_0) است که برای آنها $g_0 = b_0$ می باشد. به طور مشابه، در طول خط $P_B Q_B$ ، $g_0 = r_0$ و در طول خط $P_G Q_G$ ، $r_0 = b_0$ است.

(ج) در هر نقطه از ناحیه 'مسطح محصور در مثلث $P_R Q_R P_G$ '، $g_0 \geq b_0$ است. در هر نقطه از ناحیه 'مسطح محصور در مثلث $P_R Q_R P_B$ '، $b_0 \geq g_0$ است. بنابراین خط $P_R Q_R$ ناحیه 'را از ناحیه ' $b_0 > g_0$ جدا می کند.

(د) برای $i = R, G, B$ ، $|WP_i|/|P_i Q_i| = 2/3$ و $|WQ_i|/|P_i Q_i| = 1/3$ ، که نماد آرگومان نشانه 'طول آرگومان' است.

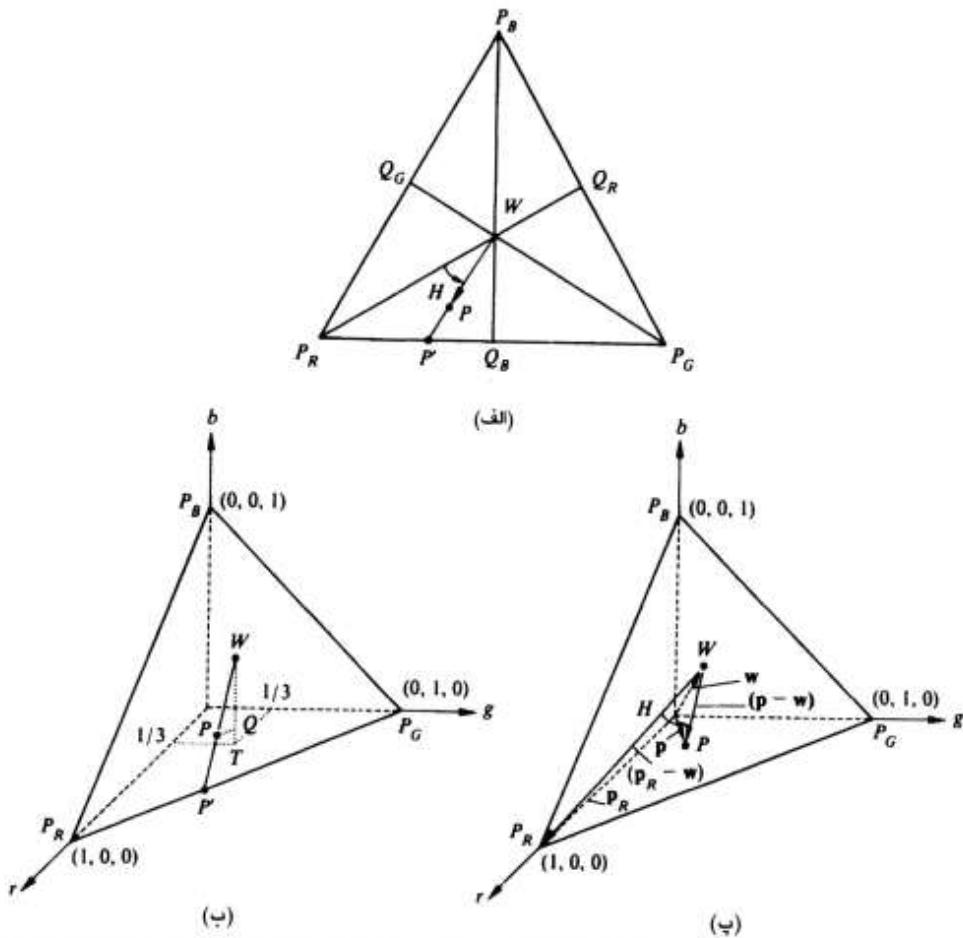
(ه) طبق تعریف، بخش RG ناحیه 'محصور در $WP_R P_G$ '، بخش GB ناحیه 'محصور در $WP_G P_B$ ' و بخش BR ناحیه 'محصور در $WP_B P_R$ ' می باشد.

با مراجعه به شکل ۴.۴۶ (الف)، مؤلفه 'اصل رنگ' هر رنگ دلخواه با زاویه 'بین پاره خطهای WP_R و WP '، یا در شکل برداری (شکل ۴.۴۶ ب) با زاویه 'بین بردارهای $(\mathbf{p}_R - \mathbf{w})$ و $(\mathbf{p} - \mathbf{w})$ '، تعریف می شود. مثلاً همان طور که قبلًا بیان شد، $H = 0^\circ$ متناظر با قرمز است، $H = 120^\circ$ متناظر با سبز است، و الی آخر. گرچه زاویه H را می توان نسبت به هر خطی که از W می گذرد، اندازه گرفت، اما اندازه گیری مؤلفه 'اصل رنگ' نسبت به محور قرمز مرسوم تر است. در حالت کلی معادله زیر برای $0^\circ \leq H \leq 180^\circ$ برقرار است:

$$(\mathbf{p} - \mathbf{w}) \cdot (\mathbf{p}_R - \mathbf{w}) = \|\mathbf{p} - \mathbf{w}\| \|\mathbf{p}_R - \mathbf{w}\| \cos H \quad (12-6.4)$$

که $\|\mathbf{x}\| \|\mathbf{y}\| \cdot (\mathbf{x} \cdot \mathbf{y}) = \mathbf{x}^T \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\|$ (یانگر ضرب نقطه ای یا داخلی دو بردار است، و زوج خطوط کوتاه عمودی نماد نرم^۱ (طول) آرگومان برداری است. مسأله فعلی، بیان این نتیجه بر حسب مجموعه ای از مؤلفه های اولیه 'RGB' است.

از شرایط (الف) و (ب) نتیجه می شود،



شکل ۴.۳۶. ۳ جزئیات مثلث رنگی HSI که در بدست آوردن عباراتی برای اصل رنگ و اشباع مورد نیاز هستند.

$$\|p - w\| = \left[\left(r - \frac{1}{3}\right)^2 + \left(g - \frac{1}{3}\right)^2 + \left(b - \frac{1}{3}\right)^2 \right]^{1/2} \quad (13-6.4)$$

زیرا طول بردار \mathbf{a} با مؤلفه های a_1, a_2 و a_3 برابر $\|\mathbf{a}\| = [a_1^2 + a_2^2 + a_3^2]^{1/2}$ است. در معادله (۱۳-۶.۴) با جایگذاری معادلات (۷-۶.۴)-(۹-۶.۴) داریم:

$$\|p - w\| = \left[\frac{9(R^2 + G^2 + B^2) - 3(R + G + B)^2}{9(R + G + B)^2} \right]^{1/2}. \quad (14-6.4)$$

چون بردارهای \mathbf{p}_R و \mathbf{w} به ترتیب از مبدأ به نقاط $(1, 0, 0)$ و $(1/3, 1/3, 1/3)$ وصل می شوند، داریم:

۴.۶ پردازش تصویر رنگی ۲۸۵

$$\|\mathbf{p}_R - \mathbf{w}\| = \left(\frac{2}{3}\right)^{1/2} \quad (15-6.4)$$

توجه داشته باشید که برای دو بردار \mathbf{a} و \mathbf{b} داریم، آن گاه نتیجه می شود:

$$\begin{aligned} (\mathbf{p} - \mathbf{w}) \cdot (\mathbf{p}_R - \mathbf{w}) &= \frac{2}{3}\left(r - \frac{1}{3}\right) - \frac{1}{3}\left(g - \frac{1}{3}\right) + \frac{1}{3}\left(b - \frac{1}{3}\right) \\ &= \frac{2R - G - B}{3(R + G + B)}. \end{aligned} \quad (16-6.4)$$

از معادله' (12-6.4) نتیجه می شود:

$$H = \cos^{-1} \left[\frac{(\mathbf{p} - \mathbf{w}) \cdot (\mathbf{p}_R - \mathbf{w})}{\|\mathbf{p} - \mathbf{w}\| \|\mathbf{p}_R - \mathbf{w}\|} \right]. \quad (17-6.4)$$

با جایگذاری معادلات (4.6.14)-(4.6.16) در معادله' (4.6.17) و ساده سازی آنها، عبارت زیر برای H بر حسب R , G و B بدست می آید:

$$H = \cos^{-1} \left[\frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{[(R - G)^2 + (R - B)(G - B)]^{1/2}}} \right]. \quad (18-6.4)$$

معادله' (4.6.18) مقادیر H در بازه' $0^\circ \leq H \leq 180^\circ$ را بدست می دهد. اگر $g_0 > b_0$ باشد، آن گاه باید H بزرگتر از 180° باشد. بنابراین هر گاه $b_0 > g_0$ باشد، کافی است قرار دهیم $H = 360^\circ - H$. گاهی معادله' اصل رنگ با استفاده از اتحاد مثلثاتی $\cos^{-1}(x) = 90^\circ - \tan^{-1}(x/\sqrt{1-x^2})$ ، بر حسب تائزات بیان می شود. با این حال معادله' (4.6.18) نه تنها نمایش ساده تری دارد، بلکه از نظر پیاده سازی سخت افزاری نیز بهتر است.

قدم بعدی به دست آوردن عبارتی برای S بر حسب مقادیر مؤلفه های RGB است. برای انجام این کار باز هم به اشکال ۴.46 (الف) و (ب) نیاز داریم. چون اشباع رنگ، میزان عدم اضافه شدن رنگ سفید به آن می باشد، طبق شکل ۴.46 (الف)، S , اشباع نقطه رنگی P با نسبت $|WP|/|WP'|$ تعیین می شود، که در آن P' محل تلاقی امتداد خط WP با نزدیکترین ضلع مثلث است.

بارجوع به شکل ۴.46 (ب)، فرض کنید T تصویر W به موازات محور b بر روی صفحه' rg ، و Q تصویر P به موازات صفحه' rg بر روی WT باشد. آنگاه داریم

$$S = \frac{|WP|}{|WP'|} = \frac{|WQ|}{|WT|} = \frac{|WT| - |QT|}{|WT|} \quad (19-6. ۴)$$

که تساوی دوم معادله فوق از مثلث های مشابه نتیجه می شود. چون در بخش نشان داده شده $\frac{1}{3}|WT| = b|QT| = b$ است، نتیجه می شود:

$$\begin{aligned} S &= 3\left(\frac{1}{3} - b\right) \\ &= 1 - 3b \\ &= 1 - b_0 \end{aligned} \quad (20-6. ۴)$$

که تساوی آخر از معادله $(10-6. ۴)$ و شرط (ث) نتیجه می شود. همچنین ذکر می کنیم که در بخش می باشد. در واقع، توضیحی مشابه توضیح اخیر نشان خواهد داد که رابطه $b_0 = \min(r_0, g_0, b_0), RG$

$$\begin{aligned} S &= 1 - \min(r_0, g_0, b_0) \\ &= 1 - \frac{3}{(R + G + B)} [\min(R, G, B)] \end{aligned} \quad (21-6. ۴)$$

به طور کلی برای هر نقطه ای که روی مثلث HSI باشد، صحیح است.
نتایجی که تاکنون بدست آمد، عبارات زیر را برای بدست آوردن مقادیر HSI در محدوده $[0, 1]$ از مقادیر RGB در همان محدوده نتیجه می دهد:

$$I = \frac{1}{3}(R + G + B) \quad (22-6. ۴)$$

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)] \quad (23-6. ۴)$$

$$H = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\} \quad (24-6. ۴)$$

که طبق توضیحات قبلی، اگر $(B/I) > (G/I)$ باشد، قرار می دهیم $H = 360^\circ - H$. به منظور این که مؤلفه اصل رنگ به محدوده $[0, 1]$ تراز شود، قرار می دهیم $H = H/360^\circ$. در نهایت اگر $S = 0$

و

۴.۶.۴ پردازش تصویر رنگی ۲۸۷

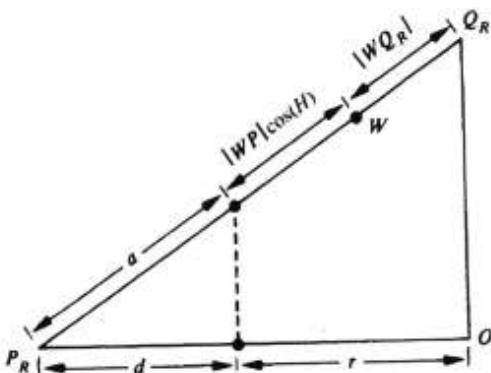
باشد، از معادله (۴.۱۹-۶) نتیجه می شود که باید $|WP|$ صفر باشد، این بدان معناست که نقاط W و P برهم منطبق شده اند و بنابراین تعریف زاویه H بی معنای گردد. به همین دلیل، وقتی اشباع صفر باشد، مؤلفه اصل رنگ تعریف نمی شود. به طور مشابه با توجه به معادلات (۴.۲۲-۶) و (۴.۲۳-۶)، اگر $I = 0$ باشد، اشباع تعریف نشده است.

تبدیل از RGB به HSI. اکنون می خواهیم بازی مقادیر HSI داده شده در بازه $[0, 1]$ ، مقادیر RGB متناظر در همان محدوده را پیدا کنیم. تحلیل ما بستگی دارد به این که کدام یک از بخش های تعریف شده در شرط (ج) ، مقدار داده شده H را شامل است. کار را با قراردادن $(H) = 360^\circ$ شروع می کنیم تا اصل رنگ به محدوده $[0^\circ, 360^\circ]$ باز گردد.

برای بخش RG طبق معادله (۴.۲۰-۶) داریم،

$$b = \frac{1}{3}(1 - S). \quad (4.25-6)$$

سپس با توجه به شکل ۴.۴۶ (الف) و این واقعیت که r تصویر نقطه P روی محور قرمز است، مقدار r را می پاییم. مثلث $P_R Q_R O$ در شکل ۴.۴۷ را در نظر بگیرید، که O مبدأ سامانه مختصات rgb است. وتر این مثلث پاره خط $P_R Q_R$ در شکل ۴.۴۶ (الف) است، و خطی که از O به P_R وصل شده است، محور قرمز می باشد که شامل r است. خط مقطع محل تلاقی مثلث $P_R Q_R O$ با صفحه ای است که شامل P بوده و بر محور قرمز عمود است. این دو شرط، این معنارا می رسانند که صفحه شامل r نیز می باشد. به علاوه، نقطه ای که در آن $P_R Q_R$ با صفحه تلاقی می کند، شامل تصویر P روی خط $P_R Q_R$ بوده و با توجه



شکل ۴.۴۷ آرایش مورد استفاده برای بدست آوردن معادلاتی برای تبدیل از RGB به HSI

به شکل ۴. ۴۶ (الف)، برابر $|WP| \cos H$ است. از مثلثهای متشابه نتیجه می شود،

$$\frac{|P_R Q_R|}{|P_R O|} = \frac{a}{d}. \quad (26-6. 4)$$

که $a = |P_R Q_R| - (|WP| \cos H + |WQ_R|)$ و $d = 1 - r$ ، $|P_R O| = 1$ با قرار دادن این نتایج در معادله (۲۶-۶. ۴) و ساده سازی داریم:

$$\begin{aligned} r &= \frac{|WQ_R|}{|P_R Q_R|} + \frac{|WP|}{|P_R Q_R|} \cos H \\ &= \frac{1}{3} + \frac{|WP|}{|P_R Q_R|} \cos H \end{aligned} \quad (27-6. 4)$$

که در رابطه 'فوق براساس شکل ۴. ۴۶ (الف)، از تساوی $|P_R Q_R| = 3|WQ_R|$ استفاده کردیم. تنها مجهول در رابطه 'فوق $|WP|$ است که طبق معادله (۱۹-۶. ۴) برابر $S|WP'|$ می باشد. در شکل ۴. ۴۶ (الف) زاویه بین پاره خطهای $P_R Q_R$ و WQ_B در نقطه 'W برابر 60° است؛ بنابراین $|WP'| = |WQ_B|/\cos(60^\circ - H)$ ، یا $|WQ_B| = |WP'| \cos(60^\circ - H)$ می باشد. با توجه به این که $|WQ_B| = |WQ_R|$ و با جایگذاری نتایج فوق در معادله (۲۷-۶. ۴) داریم:

$$\begin{aligned} r &= \frac{1}{3} + \frac{S|WQ_R|\cos H}{|P_R Q_R|\cos(60^\circ - H)} \\ &= \frac{1}{3} \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right] \end{aligned} \quad (28-6. 4)$$

که باز هم از تساوی $|P_R Q_R| = 3|WQ_R|$ استفاده کردیم. نهایتاً از معادله (۱۰-۶. ۴) نتیجه می گیریم که $(r + b)g = 1 - (r + b)$. بنابراین نتایج به ازای $0^\circ < H \leq 120^\circ$ عبارتند از:

$$b = \frac{1}{3}(1 - S) \quad (29-6. 4)$$

$$r = \frac{1}{3} \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right] \quad (30-6. 4)$$

$$g = 1 - (r + b). \quad (31-6. 4)$$

۴.۶.۴ پردازش تصویر رنگی

۲۸۹

مؤلفه های رنگ که اخیراً بدست آمدند، در مفهوم معادله^{۱)} (۴.۶-۱۰) تراز شده هستند. با توجه به معادلات (۴.۶-۱۱) که در آنها $R = 3Ib$ ، $G = 3Ig$ و $B = 3Ir$ مولفه های RGB را بازیابی می کنیم.

برای بخش $GB < H \leq 240^\circ$ (۱۲۰) نیز، با روایی مشابه شرح اخیر، نتیجه می گیریم که:

$$H = H - 120^\circ \quad (32-6.4)$$

$$r = \frac{1}{3}(1 - S) \quad (33-6.4)$$

$$g = \frac{1}{3} \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right] \quad (34-6.4)$$

و

$$b = 1 - (r + g). \quad (35-6.4)$$

مقادیر R ، G و B به رویی که قبلاً تشریح شد، از روی r ، g و b بدست می آیند.

برای بخش $BR < H \leq 360^\circ$ (۲۴۰) نیز داریم:

$$H = H - 240^\circ \quad (36-6.4)$$

$$g = \frac{1}{3}(1 - S) \quad (37-6.4)$$

$$b = \frac{1}{3} \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right] \quad (38-6.4)$$

و

$$r = 1 - (g + b). \quad (39-6.4)$$

مقادیر R ، G و B به رویی که قبلاً بیان شد، از روی r ، g و b بدست می آیند.

مثال های پردازش تصویر با استفاده از مدل HSI، در بخش ۶.۶.۴ ارائه می شوند.

۴.۶.۴ پردازش تصویر شبه رنگی^{۱)}

در این بخش برای انتساب رنگ به تصاویر تکریج، چند روش را که براساس خواص گوناگون موجود در محیط ای سطح خاکستری تصاویر شکل گرفته اند، ارائه می کنیم.

1) Pseudo-color

بخش بندی شدت

روش بخش بندی شدت^۱ (که گاهی بخش بندی چگالی^۲ نامیده می شود) و سپس رنگ-رمزنگاری^۳ از ساده ترین مثالهای پردازش تصویر شب رنگی است. اگر به تصویر به عنوان یکتابع دو بعدی شدت نگاه کنیم (به بخش ۱. ۲. مراجعه کنید)، این روش را می توان به عنوان روشی که صفحه هارا موازی صفحه مختصات تصویر قرار می دهد، تعبیر کرد؛ آنگاه هر صفحه تابع را در ناحیه 'تلاقی' "بخش می کند". شکل ۴۸. ۴۸ مثالی از کاربرد یک صفحه در $I = f(x, y)$ به منظور "بخش بندی" تابع به دو سطح را نشان می دهد.

اگر به هر طرف صفحه شکل ۴۸. ۴۸ رنگ متفاوتی متنسب شود، هر پیکسلی که سطح خاکستریش بالای صفحه باشد با یک رنگ و هر پیکسلی که زیر صفحه باشد، با رنگ دیگری رمز می شود. حاصل این فرآیند، یک تصویر دو رنگ است که می توان بالا و پایین بردن صفحه 'بخش کننده' نسبت به محور سطح خاکستری، ظاهر نسبی آن را کترنل نمود.

در حالت کلی می توان روش فوق را به این صورت جمع بندی کرد. فرض کنید که M صفحه در سطوح I_1, I_2, \dots, I_L تعریف شوند و اجازه دهید I_0 بیانگر سیاه $[f(x, y) = 0]$ و I_L بیانگر سفید $[f(x, y) = L]$ باشد. آنگاه با فرض این که $L < M < 0$ باشد، M صفحه، محدوده خاکستری را به $1 + M$ ناحیه بخش می کند و انتسابات رنگ طبق رابطه^۴

$$f(x, y) = c_k \quad f(x, y) \in R_k \quad \text{اگر } (40-6. ۴)$$

انجام می شود که رنگ مربوط به R_k ، امین ناحیه ای که بوسیله صفحات بخش کننده تعریف می شود، است.

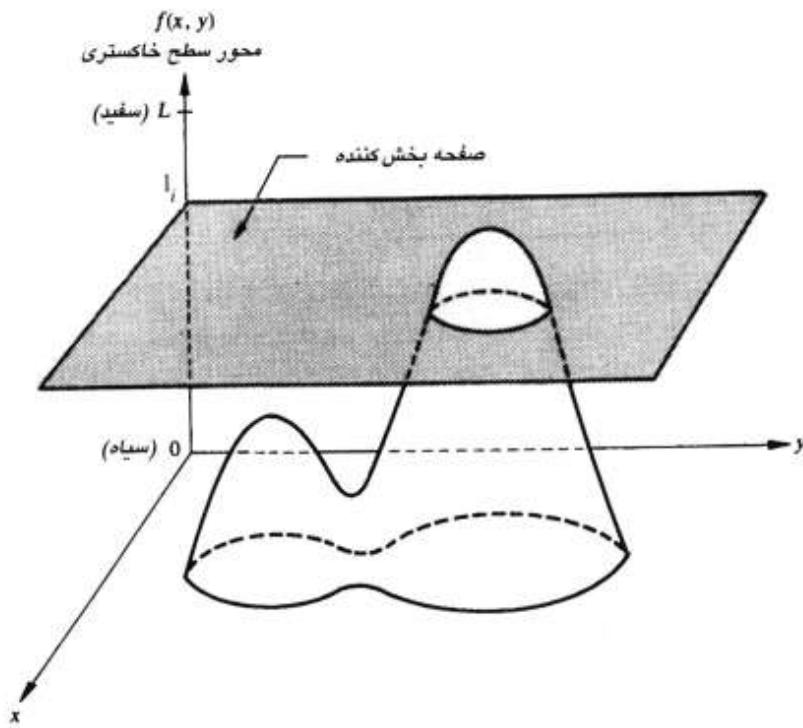
ایده' صفحات بیش از هر چیز برای تعبیر هندسی روش بخش بندی شدت مفید است. شکل ۴۹. ۴۹. ۴ نمایش دیگری از همان نگاشت شکل ۴۸. ۴۸ را نشان می دهد. برطبق تابع نگاشت شکل ۴۹. ۴، به هر سطح خاکستری ورودی بسته به این که بیشتر یا کمتر از مقدار I باشد، یکی از دو رنگ متنسب می شود. وقتی سطوح خاکستری بیشتری استفاده شوند، تابع نگاشت به شکل پله ای می گردد. این نوع نگاشت حالت خاصی از روش مشروح در بخش بعدی است.

1) Intensity slicing

2) Density slicing

3) Color- coding

۴.۶ پردازش تصویر رنگی ۲۹۱



شکل ۴۱۰۴ تغییر هندسی روش بخش بندی شدت

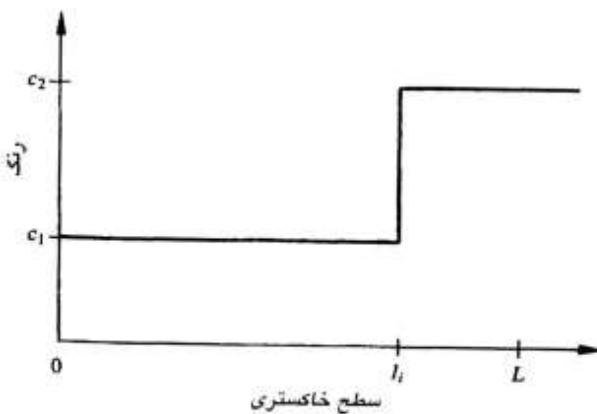
مثال: در پلیت V مثالی از بخش بندی شدت دیده می شود. قسمت (الف) تصویر تک رنگی از Picker Thyroid Phantom (الگوی آزمون تشعشع^۱) و پلیت V (ب) حاصل بخش بندی شدت این تصویر به نواحی دارای هشت رنگ مختلف است. بسیاری نواحی که در تصویر تک رنگ با شدت ثابت ظاهر می شوند، در واقع کاملاً متغیر هستند که در تصویر بخش بندی شده با رنگهای مختلف نشان داده می شوند. برای نمونه، قله^۲ چپ، در تصویر تک رنگ غیر واضح بوده و تشخیص تغییرات شدت در آن مشکل است. بر عکس، تصویر رنگی، هشت ناحیه متفاوت با شدت ثابت، هر ناحیه برای یکی از رنگ های مورد استفاده، را نشان می دهد.

تبديلات سطح خاکستری به رنگ

سایر انواع تبدیلات کلی تر هستند و بنابراین نسبت به روش بخش بندی ساده که در بخش قبلی بیان

1) Radiation test pattern

2) Lobe



شکل ۴.۴۹ نمایش دیگری از روش بخش‌بندی شد

شد، قادر به دستیابی به محدوده 'وسیعتری از نتایج ارتقای شبه رنگی' هستند. یک روش که جذابیت ویژه‌ای دارد، در شکل ۴.۵۰ دیده می‌شود. اندیشه پشت این روش انجام سه تبدیل مستقل روی سطح خاکستری هر پیکسل ورودی می‌باشد. آنگاه سه خروجی به طور جداگانه به تفنگ‌های قرمز، سبز و آبی نمایشگر تلویزیون رنگی وارد می‌شوند. این روش تصویر مرکبی تولید می‌کند که محتوای رنگش بوسیله 'طیعت توابع تبدیل مدوله شده' اند. توجه کنید که این تبدیل‌ها، مقادیر سطح خاکستری تصویر را تبدیل می‌کنند، اما تابع مکان نیستند.

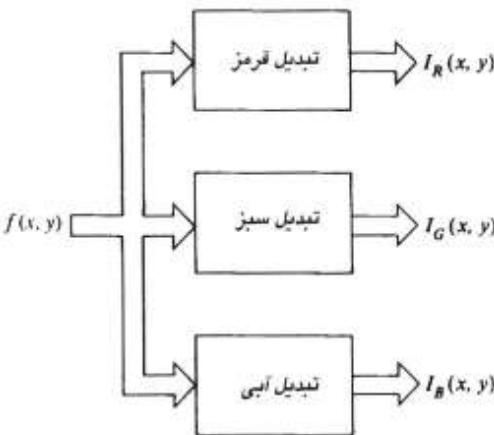
همان طور که در بخش قبلی بیان شد، روش موجود در شکل ۴.۴۹ مورد خاصی از روش اخیر است. در آنجا توابع خطی قطعه‌ای^۱، سطوح خاکستری را به رنگ تبدیل می‌کنند. با این حال روشی که در این بخش بیان شد، را می‌توان براساس توابع غیرخطی آرام^۲ انجام داد. همانطور که انتظار می‌رود، این کار انعطاف‌پذیری روش را افزایش می‌دهد.

مثال: پلیت VI (الف) تصویر تک رنگ مرکبی را که شامل دو تصویر از یک چمندان است، نشان می‌دهد. این تصاویر بوسیله 'سامانه پیمایش گر اشعه ایکس مخصوص فرودگاه گرفته شده' اند. تصویر سمت چپ حاوی وسائل معمولی است. تصویر سمت راست حاوی همان اشیای تصویر سمت چپ به اضافه 'جعبه‌ای از مواد منفجره' پلاستیکی شبیه سازی شده می‌باشد. هدف از این مثال، نمایش کاربرد تبدیل‌های سطح خاکستری به رنگ برای بدست آوردن درجات مختلفی از ارتقای تصویر می‌باشد.

1) Piecewise linear function

2) Smooth

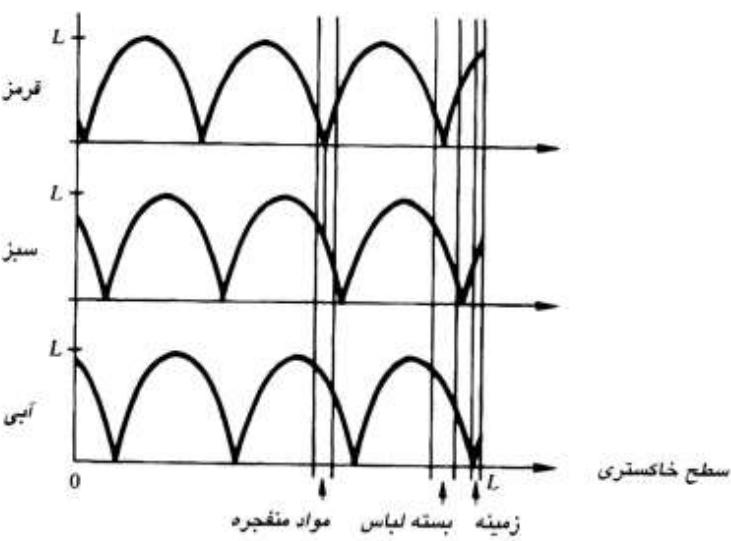
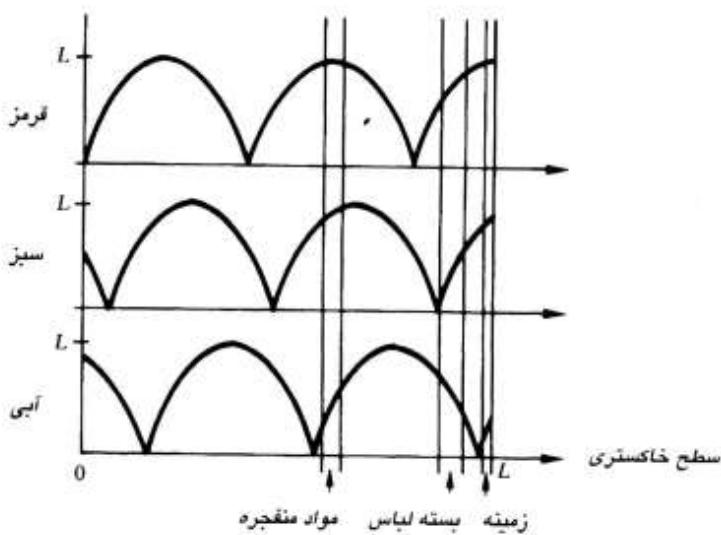
۴.۶. پردازش تصویر رنگی ۲۹۳



شکل ۴.۵۰ نمودار بلوکی عملکردی مربوط به پردازش تصویر شبه رنگی I_R ، I_G و I_B به ترتیب به ورودیهای قرمز، سبز و آبی یک نمایشگر رنگی RGB وارد می شوند.

شکل ۴.۵۱ توابع تبدیل استفاده شده را نشان می دهد. این توابع سینوسی شامل نواحی با مقدار نسبتاً ثابت در اطراف قله ها^۱ و نیز نواحی با تغییر تند در نزدیک دره ها^۲ می باشند. تغییر فاز و فرکانس هر سینوس می تواند محدوده های سطح خاکستری را (توسط رنگ) مشخص نماید. برای نمونه، اگر هر سه تبدیل، فاز و فرکانس یکسانی داشته باشند، تصویر خروجی تک رنگ خواهد بود. یک تغییر کوچک در فاز بین سه تبدیل، تغییر کوچکی را در پیکسل هایی که سطوح خاکستری شان متناظر با قله های سینوس هاست، تولید می کند، مخصوصاً اگر سینوس ها دوره^۳ وسیعی (فرکانس های پایینی) داشته باشند. در نتیجه تفاوت زیاد بین دامنه های سه سینوس که ناشی از جابجایی فاز بین آنها می باشد، به پیکسل هایی که مقادیر خاکستری شان در بخش پرشیب^۴ سینوس ها قرار دارد، محتوای رنگ بسیار قویتری نسبت داده می شود.

تصویری که در پلیت VI(b) دیده می شود، با توابع تبدیل شکل ۴.۵۱ (الف) که نوارهای^۵ سطح خاکستری متناظر با مواد منفجره، بسته لباس و زمینه را نیز نشان می دهند، بدست آمد. توجه کنید که ماده منفجره و زمینه، سطوح خاکستری کاملاً متفاوتی دارند، اما بر اثر تناوبی بودن موج های سینوسی هر دو بار نگ تقریباً یکسانی رمز شده اند. تصویر موجود در پلیت VI(b) با توابع تبدیل موجود در شکل ۴.۵۱ (ب) بدست آمد. در این مورد نوارهای شدت مربوط به ماده منفجره و بسته لباس با تبدیل های



شکل ۴۱۰۵ توابع تبدیل مورد استفاده در بدست آوردن تصاویر پلت VI

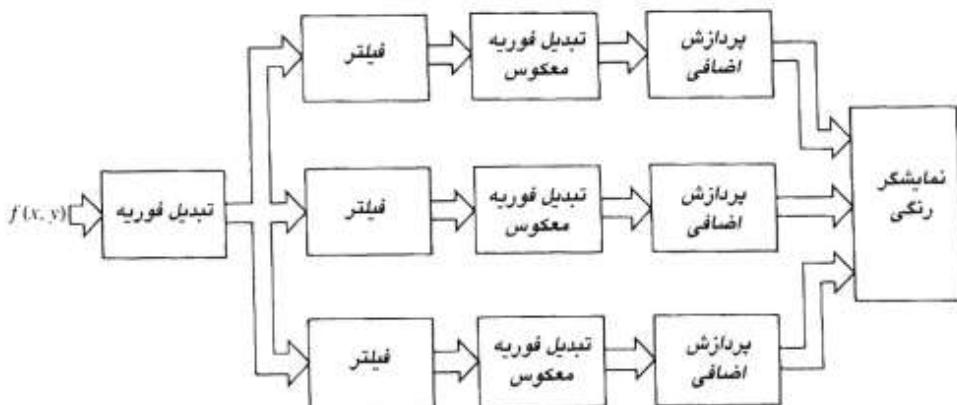
۴.۶.۴ پردازش تصویر رنگی ۲۹۵

مشابهی نگاشت شدند و بنابراین اساساً رنگ یکسانی برای آنها در نظر گرفته شد. توجه کنید که این نگاشت امکان می دهد که ناظر از میان مواد منفجره "بینند". نگاشت های زمینه تقریباً همان بود که برای پلیت VI (ب) بکار گرفته شدند، و بنابراین انتسابهای رنگ تقریباً یکسانی را نتیجه دادند.

بلک روشن فیلتر کردن

شکل ۴.۵۲. یک شیوه رنگ- رمزگاری^۱ را که براساس اعمال میدان فرکانس است، نشان می دهد. روش طراحی شده شبیه روش فیلتر کردن عادی که قبلاً در این فصل توضیح داده شد، می باشد، با این تفاوت که در اینجا برای تولید سه تصویری که باید به ورودیهای قرمز، سبز و آبی یک نمایشگر رنگی وارد شوند، تبدیل فوریه تصویر، به طور مستقل با سهتابع فیلتر تغییر داده می شود. به عنوان مثال مراحلی را که برای بدست آوردن تصویر برای کانال قرمز دنبال شد، در نظر بگیرید. تبدیل فوریه تصویر ورودی با استفاده از یکتابع فیلتر مشخص، تغییر داده می شود. آنگاه تصویر پردازش شده با استفاده از تبدیل فوریه معکوس بدست می آید. حال قبل از این که تصویر به ورودی قرمز نمایشگر داده شود، می توان نوعی پردازش اضافی (نظیر تعديل بافت نگار) را نیز اجرا کرد. برای دو مسیر دیگر موجود در شکل ۴.۵۲، نیز توضیحات مشابهی صدق می کند.

هدف این روش رنگ- پردازش^۲، آن است که نواحی تصویر را توسط رنگ و براساس محتوای



شکل ۴.۵۲. یک مدل فیلتر کردن برای پردازش تصویر شبیه رنگی

1) Color- coding

2) Color- processing

۴۹۶ پردازش تصویر رقمه

فرکانسی آنها رمز نماید. یک روش نوعی فیلتر کردن، استفاده از فیلترهای پایین گذر، میان گذر (یا حذف باند) و بالاگذر برای بدست آوردن سه محدوده از مؤلفه های فرکانسی است. فیلترهای حذف باند و میان گذر تعمیمی از فیلترهای پایین گذر و بالاگذر که قبل توضیح داده شد، هستند. یک روش ساده برای تولید فیلترهایی که فرکانس های یک همسایگی دایروی حول نقطه 'دلخواه' (u_0, v_0) را حذف یا تضعیف می کنند، انتقال مختصات فیلترهای بالاگذر تشریح شده در بخش ۴.۲ می باشد. روال تولید فیلتر ایده آل به صورت زیر است.

فیلتر حذف باند ایده ال^۱ (IBRF) که تمام فرکانسها در یک همسایگی به شعاع D_0 حول نقطه ' (u_0, v_0) ' را حذف می کند، با رابطه^۲

$$H(u, v) = \begin{cases} 0 & D(u, v) \leq D_0 \\ 1 & D(u, v) > D_0 \end{cases} \quad (41-6. ۴)$$

داده می شود که

$$D(u, v) = [(u - u_0)^2 + (v - v_0)^2]^{1/2}. \quad (42-6. ۴)$$

توجه کنید که معادله '۴.۶-۴۱' از نظر شکل با معادله '۴.۶-۴' یکسان است، اما تابع فاصله ' $D(u, v)$ ' به جای مبدأ حول نقطه ' (u_0, v_0) ' محاسبه می شود.
به علت تقارن تبدیل فوریه، هر حذف باندی که حول مبدأ نباشد، به صورت زوج های^۳ متقاضان انجام می شود تا این که نتایج معناداری بدست آید. در مورد فیلتر ایده ال، معادله '۴.۶-۴۱' به صورت

$$H(u, v) = \begin{cases} 0 & D_1(u, v) \leq D_0 \quad \text{یا} \quad D_2(u, v) \leq D_0 \\ 1 & \text{در غیر این صورت} \end{cases} \quad (43-6. ۴)$$

در می آید که در آن

$$D_1(u, v) = [(u - u_0)^2 + (v - v_0)^2]^{1/2} \quad (44-6. ۴)$$

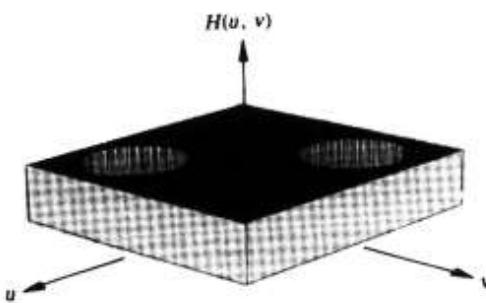
و

$$D_2(u, v) = [(u + u_0)^2 + (v + v_0)^2]^{1/2}. \quad (45-6. ۴)$$

1) Ideal bandreject filter (IBRF)

2) In pairs

۴۹۷ ۶۰.۴ پردازش تصویر رنگی



شکل ۴.۵۳. فیلتر حذف باند ایده‌آل

به طور مشابه این روال را می‌توان به چهار ناحیه یا بیشتر تعمیم داد. فیلتر باترورث تشریح شده در بخش ۴.۲.۴ را می‌توان با تبعیت از روشی که اخیراً برای فیلتر ایده‌آل بیان شد، مستقیماً برای حذف باند نیز بکاربرد. شکل ۴.۵۳. رسم پرسپکتیوتابع انتقال IBRF نوعی را نشان می‌دهد.

فیلتری که در بالا بیان شد، حول نقطه‌ای غیر از مبدأ تبدیل فوریه قرار گرفته است. برای حذف یک باند فرکانسی در حول مبدأ، می‌توان از فیلترهای متقارن مشابه فیلترهای پایین گذر و بالا گذری که قبله بیان شدند، استفاده کرد. این روال برای فیلترهای ایده‌آل و باترورث در زیر بیان می‌شود.

یک فیلتر حذف باند ایده‌آل متقارن شعاعی که باندی از فرکانس‌های حول مبدأ را حذف می‌کند، با

رابطه'

$$H(u, v) = \begin{cases} 1 & D(u, v) < D_0 - \frac{W}{2} \\ 0 & D_0 - \frac{W}{2} \leq D(u, v) \leq D_0 + \frac{W}{2} \\ 1 & D(u, v) > D_0 + \frac{W}{2} \end{cases} \quad (46-6.4)$$

داده می‌شود که W عرض باند و D_0 مرکز شعاعی آن است. مانند تمام فیلترهای متقارن شعاعی این فیلتر را می‌توان با یک مقطع عرضی به طور کامل مشخص کرد. مثلاً یک فیلتر حذف باند باترورث¹ (BBRF) متقارن شعاعی از مرتبه n دارای تابع انتقال

$$H(u, v) = \frac{1}{1 + \left[\frac{D(u, v)W}{D^2(u, v) - D_0^2} \right]^{\frac{2n}{m}}} \quad (47-6.4)$$

1) Butterworth bandreject filter (BBRF)

می باشد که W عرض باند و D مرکز آن باند می باشد.

فیلتر های میان گذر، فرکانس هایی در یک باند یا ناحیه، مشخص را عبور می دهند، در حالی که سایر فرکانس ها را تضعیف یا کاملاً حذف می کنند. بنابراین آنها دقیقاً عکس فیلتر های حذف باند هستند. با این شرایط اگر $H_R(u, v)$ تابع انتقال یکی از فیلتر های حذف باندی که اخیراً آیان شد، باشد، تابع میان گذر متناظر را می توان بسادگی با تغییر^۱ $H_R(u, v)$ به صورت زیر بدست آورد:

$$H(u, v) = -[H_R(u, v) - 1]. \quad (48-6)$$

مثال: پلیت VII (الف) یک تصویر نک رنگ و پلیتهای VII (ب) و (پ) نتایج کاربرد فیلتر های باترورث را نشان می دهند. پلیت VII (ب) (در تفنگ^۲ قرمز نمایشگر رنگی) حاصل استفاده از فیلتر بالاگذری را نشان می دهد که نقطه 'قطع آن دایره شامل ۹۰ درصد انرژی تصویر است، (بخش ۴.۰.۴ را ببینید). پلیت VII (پ) نسخه 'بالاگذر بر روی تفنگ قرمز، و نیز نسخه 'پایین گذر (تفنگ آبی) و نسخه 'میان گذر (تفنگ سبز) پلیت VII (الف) را نشان می دهد. نقطه 'قطع فیلتر پایین گذر دایره ای بود که ۹۸ درصد انرژی تصویر را شامل بود؛ و محدوده 'فیلتر میان گذر نیز بین دایره هایی که ۲۰ و ۹۵ درصد انرژی را شامل بودند، قرار داشت. ارتقای اصلی حاصل از این فرآیند، مرئی تر شدن حلقه 'بیرونی است که در تصویر اولیه تقریباً نامرئی بود.

□

۴.۶.۴ پردازش تصویر تمام رنگی

مبحث پردازش تصویر رنگی را با مقداری شرح نقش روش های تمام رنگی برای ارتقای تصویر پایان دهیم. بویژه، به دلایلی که در بخش ۴.۶.۲. بیان شد، به مدل HSI علاقه مند هستیم، که آن دلایل عبارتند از: (۱) در این مدل اطلاعات شدت و رنگ از هم مجزا شده اند؛ و (۲) اصل رنگ و اشباع رابطه نزدیکی با روش احساس رنگ در سامانه بینایی انسان دارند.

به دست آوردن مؤلفه های تصویری HSI از تصویر RGB

به دلیل اهمیت مدل RGB در نمایش تصاویر رنگی، شرح پردازش تمام رنگی را با نمایش تفاوت ها و

1) Flipping

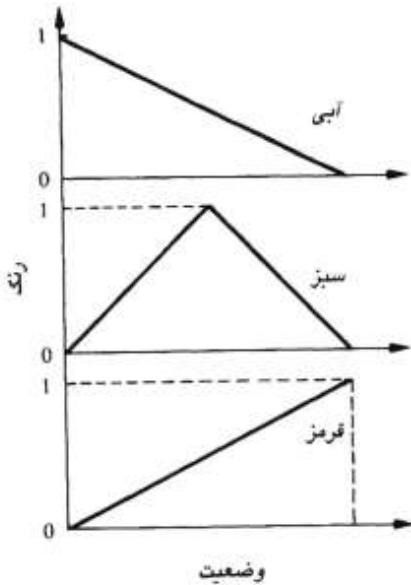
2) Gun

۴.۶ پردازش تصویر رنگی ۴۹۹

مشابهت های موجود بین تصاویری که در مدل های RGB و HSI بیان می شوند، آغاز می کنیم. این مسأله دانش ما درباره خود مدل HSI را نیز افزایش می دهد. به یاد آورید که تصویر پردازش شده در فضای HSI، جهت نمایش باید دوباره به RGB تبدیل شود.

پلیت VIII (الف) یک الگوی آزمون رنگی RGB را نشان می دهد که در بالای آن هشت نوار نازک، ابتداء سیاه، سپس رنگهای اولیه و ثانویه خالص (ترتیب نشان داد شده اهمیت خاصی ندارد)، و در نهایت سفید وجود دارد. پس از این هشت نوار، یک نوار چند رنگ پهن قرار دارد که از آبی، تاسیز، و در نهایت تا قرمز تغییر می کند. پس از این نوار، دو نوار^۱ خاکستری از سیاه به سفید اما در جهات مخالف، قرار دارند. سپس با تکرار الگوهای رنگی، تصویر مربعی می شود. شکل ۵۴.۴ ترکیب رنگ های قرمز، سبز و آبی مورد استفاده برای تولید نوار بار رنگ متغیر در پلیت VIII (الف) را نشان می دهد. توجه کنید که رنگ نوار در انتهای چپ، آبی خالص است، در انتهای راست، قرمز خالص است، و در مرکز نوار مجموع قرمز و آبی هر کدام یک سهم و سبز دو سهم می باشد.

معمولًا در عنوانی که برای اشاره به تصویر تمام رنگی یک کار می رود، تعداد کل بیت های مورد استفاده



شکل ۵۴.۴ توابع رنگ که برای تولید نوار بار رنگ متغیر موجود در تصویر RGB پلیت VIII استفاده شدند

1) Wedge

در نمایش رنگ ذکر می شود. بنابراین تصویر رنگی ۲۴ بیتی بدین معناست که ۲۴ بیت برای نمایش هر پیکسل آن استفاده می شود. معمولاً (اما نه همیشه)، بیت ها به طور برابر بین مؤلفه های رنگی تصویر توزیع می شوند. به عنوان مثال در پلیت VIII (الف) ۸ بیت در هر مؤلفه رنگی استفاده شده اند. بنابراین هر پیکسل در هر مؤلفه تصویری مقادیری در محدوده [۰، ۲۵۵] دارد. در بخش ۴.۶. فرض کردیم که پیکسل های RGB مقادیری در محدوده [۰، ۱] دارند. این دو محدوده ناسازگار نیستند، زیرا می توان فرض کرد که یک نمایش ۸ بیتی، مقادیری بین ۰ و ۲۵۵ با نمو "۱" یا به طور معادل، مقادیری بین ۰ و ۱ با نمو_{255/1} دارد.

وضعیت تصاویر مؤلفه ای HSI کمی متفاوت است. تصویر شدت روشنایی دقیقاً شبیه هر یک از سه تصویر RGB که اخیراً تشریح شد، می باشد. هر پیکسل در تصویر اصل رنگ مقادیری بر حسب درجه دارد. با ۸ بیت، می توان ۲۵۶ مقدار مختلف را در بازه $[0^{\circ}, 360^{\circ}]$ به صورت $\frac{360}{255}$ n، برای $n = 0, 1, \dots, 255$ ، بیان کرد. به طور مشابه، پیکسل های تصویر اشباع مقادیری از ۰ (بدون اشباع) تا ۱ (شباع کامل) با نمونه های $\frac{1}{255}$ دارند.

پلیت VIII (ب) تصویر مؤلفه اصل رنگ است که با استفاده از معادله $(4.6-24)$ از تصویر RGB پلیت VIII (الف) به دست آمده است. برای نواحی که مؤلفه های تک رنگ تصویر RGB اشباع صفر داشته باشند، اصل رنگ تعریف نمی شود (توضیح بعد از معادله 4.6 را بینید)؛ به طور دلخواه این نواحی در پلیت VIII (ب) بارگذاری شان داده شده اند. چون مقادیر اصل رنگ بر حسب درجه هستند، سطوح خاکستری پلیت VIII (ب) باید به عنوان زاویه تعبیر شوند (که طبق شکل ۴.۵. نسبت به محور قرمز اندازه گرفته می شوند). بنابراین رنگ های خاکستری روشن تر در پلیت VIII (ب) متناظر با مقادیر زاویه بزرگتر هستند. چون قرمزها کوچکترین مقادیر زاویه را دارند (شکل ۴.۵. را بینید)، قرمز های موجود در پلیت VIII (الف) در پلیت VIII (ب) به صورت تاریکترین خاکستریها ظاهر می شوند، زرده به صورت رنگ خاکستری روشن تر بعدی ظاهر می شود، و به همین ترتیب برای سبز، فیروزه ای، آبی و بنفش. می توان با مقایسه رنگ های پلیت VIII (الف) با خاکستری های پلیت VIII (ب) این نتیجه را مشاهده کرد. بویژه به تغییرات خاکستری متناظر نوار با رنگ متغیر توجه کنید.

پلیت VIII (ب) تصویر اشباع بدست آمده از تصویر RGB با استفاده از معادله $(4.6-23)$ را نشان می دهد. در تصویر اشباع، رنگ های اولیه و ثانویه خالص کاملاً اشباع شده، به صورت سفید (حداکثر)

۴۰۱ پردازش تصویر رنگی

ظاهر می شوند. چون وقتی شدت صفر باشد، اشباع تعریف نمی شود (توضیح پس از معادله ۴. ۲۴-۶) را ببینید)، تمام مؤلفه های سیاه تصویر RGB در پلیت VIII (پ) با رنگ سفید (بدلخواه) نشان داده می شوند. همچنین توجه کنید که سفیدهای موجود در پلیت VIII (الف) در پلیت VIII (پ) به صورت سیاه ظاهر می شوند، زیرا سفید متناظر با اشباع صفر است. خاکستری های مربوط به نوار بارنگ متغیر، هرچه به سمت دو انتهای نوار پیش رویم، روشن تر می شوند، که نشان دهنده اشباع بیشتر در آن نواحی می باشد. خاکستری های مربوط به بخش مرکزی (سبز) نوار، تیره تر از بخش های انتهایی نوار هستند که نشانه اشباع کمتر در ناحیه سبز است. این شرایط با روشن تولید بخش سبز غالب در الگوی آزمون سازگاری دارد (شکل ۴. ۵۴. را ببینید).

در نهایت پلیت VIII(ت) مؤلفه شدت تصویر HSI را نشان می دهد که با استفاده از معادله ۴. (۲۲-۶) به دست آمد. همان طور که انتظار می رفت، نوارهای مربوط به سیاه، سفید و خاکستری ها به صورت نوارهای متناظر شان در تصویر RGB ظاهر می شوند. مؤلفه های متناظر با رنگ های اولیه خالص همه مقدار یکسانی دارند و بنابراین در پلیت VIII (ت) به صورت خاکستری ثابت ظاهر می شوند. رنگ های ثانویه خالص نیز مقادیر شدت ثابتی دارند، اما مقدار شدت آنها دو برابر مقدار شدت مربوط به اولیه ها می باشد و بنابراین به صورت خاکستری روشن تری ظاهر می شوند. این نتیجه نیز مورد انتظار بود، زیرا هر نور رنگی ثانویه با جمع دو اولیه بدست می آید. شدت های متناظر با نوار بارنگ متغیر نیز همان طور که انتظار می رفت، از تیره به روشن و بر عکس تغییر می کنند.

ارتقا با استفاده از مدل HSI

همان طور که قبلاً بیان شد، مدل HSI برای ارتقای تصویر ایده اال است، زیرا در آن مؤلفه شدت نور از اطلاعات رنگ مجزا شده است. بنابراین می توان روش های ارتقای تصویر تک رنگ را که در همین فصل بیان شدند، برای ارتقای تصاویر تمام رنگی به کاربرد. این عمل نیازمند تبدیل تصویر به قالب HSI، پردازش مؤلفه شدت و تبدیل حاصل به RGB برای نمایش می باشد. این فرآیند بر محضوای رنگ تصویر تأثیر نمی گذارد.

پلیت IX این عمل را نشان می دهد. پلیت IX (الف) یک تصویر رنگی RGB را که جزئیات زمینه اش تا حد زیادی مبهم شده اند، نشان می دهد. این تصویر به HSI تبدیل شد و مؤلفه شدتیش با استفاده از

۳۰۴ پردازش تصویر رقمه

روشی که در بخش ۴.۲.۲ بیان گردید، تحت عمل تبدیل بافت نگار قرار گرفت. آنگاه دوباره تصویر به RGB تبدیل شد که حاصل در پلیت IX (ب) آمده است. بهبود در جزئیات مرئی آشکار است. چون تبدیل بافت نگار تعاییل زیادی به روشن تر کردن تصاویر دارد، مؤلفه های رنگ به میزانی متفاوت با تصویر اولیه ظاهر می شوند. گرچه اصل رنگ و اشباع همان هستند، اما به دلیل افزایش شدت، رنگ ها روشن تر ظاهر می شوند. کاربرد این روش ارتقا بر هر مؤلفه تصویر RGB، جزئیات مرئی و روشنی را افزایش خواهد داد، اما در نتیجه تغییرات ایجاد شده در مقادیر نسبی پیکسل های متناظر در سه مؤلفه تصویری RGB، رنگ های حاصل دارای مؤلفه های اصل رنگ نامفهومی خواهند بود.

۷.۴ نکات پایانی

موضوع این فصل روش های متداول برای ارتقای تصویر رقمه است. در هر حال این بخش از مبحث پردازش تصویر بسیار فعال بوده و همواره گزارش هایی در مورد روش ها و کاربردهای جدید مشاهده می شود. به همین دلیل، عناوین این فصل بیشتر با ملاک پایه و پیش نیاز بودن برای مطالعه بیشتر در این مبحث، انتخاب شدند.

مراجع

مطلوب بخش ۴.۱ از Gonzalez[1986] است. متنون تکمیلی برای موضوع بخش ۴.۲.۱ را می توان در [1983] Schowengerdt و Jain[1989] در مورد روش های پردازش بافت نگار (بخش ۴.۲.۲) براساس مقالات Hall و همکارانش[1971]، Hall[1974]، Gonzalez، Fitts[1977] و Woods, Gonzalez[1981] می باشد. برای جزئیات بیشتر در مورد ارتقای محلی به Ketcham[1976]، Narendra, Fitch[1981] و Fitts[1977] مراجعه کنید. تفریق تصویر (بخش ۴.۲.۳) یک دیگر از این روش برای آشکارسازی تغییر، در [1989] Schalkoff ارائه می شوند. روش کاهش نویز بوسیله متوسط گیری تصویر (بخش ۴.۲.۴) اولین بار بوسیله' Howell[1963] Kohler، Levine[1985] داده شده است. کتاب های شرح جامعی از فیلتر کردن مکانی (بخش ۴.۳) توسط Rosenfeld, kak[1982] و Schowengerdt[1983] نیز در این رابطه قابل توجه هستند. برای جزئیات

بیشتر در مورد پیاده سازی فیلترهای میانه ای، به Huang و همکارانش [1979]، Mannos Wolfe [1979]، و Chaudhuri [1983] نگاه کنید. کتابهای Pitas و Venetsanopoulos [1990] نیز به فیلترهای میانه ای و سایر فیلترهای مکانی غیرخطی می پردازنند. مطالب مطرح شده در مورد فیلتر کردن تقویت بالا از Schowengerdt [1983] است.

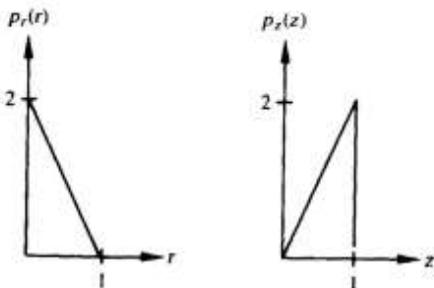
مراجع اولیه در مورد تیزکردن تصویر با تفاضل گیری، Goldmard و Hollywood [1951]، و Kovaszany Roberts [1965] پیشنهاد Joseph [1953, 1955] هستند. گردیان رایتر نیز بوسیله' Davis [1975] انجام شد. مروری بر روش هایی که یک دهه بعد در این مبحث استفاده می شدند، توسط Chen [1977] نیز قابل توجه هستند. همانند مثال هایی که توسط Lee [1983] و Chaudhuri [1983] ارائه گردیده است، کارهای جدیدتر در این زمینه بیشتر بر سرعت محاسباتی تأکید دارند.

مفاهیم فیلتر کردن میدان فرکانس که در بخش ۴.۴ معرفی شدند، براساس تعیین مستقیم فیلترهای یک بعدی هستند. در این تعیین به منظور بدست آوردن توابع فیلتر متقابران چرخشی، به جای استفاده از یک متغیر یک بعدی، از فاصله تا مبدأ صفحه' فرکانس به عنوان متغیر استفاده کردیم. برای توضیحات گسترده درباره' فیلترهای یک بعدی، به عنوان مثال به کتب Budak [1974] و Weinberg [1962] مراجعه کنید. توضیح درباره' تأکید فرکانس بالا از Hall و همکارانش [1971] می باشد. مطالب مطرح شده در مورد فیلتر کردن همشکلی، براساس مقاله Stockham [1972] است؛ کتب Oppenheim و Schafer و Pitas [1975] و Venetsanopoulos [1972] را نیز ببینید.

مطالب بخش ۴.۵ از Gonzalez [1983] و Meyer [1980] و Schutten [1983] و Vermij [1980] است. مطالب پایه ای درباره' مبانی رنگ (بخش ۴.۵) در کتب Walsh [1958] و Kiver [1965] وجود دارند. مقاله' Smith [1978] و Pritchard [1977] را نیز ببینید. به دست آوردن مدل HSI در بخش ۴.۶ براساس مقاله' Foley و Gerald و Pokorny و VanDam [1982] می باشد. مراجع دیگر برای مبانی و مدل های رنگ، Smith [1963]، Andrews [1970] و Billingsley، Roth [1968] هستند. روش های پردازش شبه رنگی در بخش ۴.۶.۳، براساس مقالات [1989]، Smith [1978] و همکاران [1972] می باشند. کتاب Green [1983] نیز قابل توجه است.

مسائل

- ۱.۴** مجموعه ای از تبدیلات بخش کننده سطح خاکستری پیشنهاد کنید که قادر به تولید تمام صفحات بینی جدآگانه یک تصویر تکرنگ ۸ بینی باشند. (برای مثال یکتابع تبدیل با خاصیت $0 = T(r) = 255$ برای r در محدوده $[0, 127]$ ، و $T(r) = 255$ برای r در محدوده $[128, 255]$ صفحه 'بینی هفتم یک تصویر ۸ بینی را تولید می کند).
- ۲.۴** توضیح دهید که چرا در حالت کلی، روش تعديل بافت نگار گسته، بافت نگار تخت نمی دهد.
- ۳.۴** فرض کنید که یک تصویر رقمه تحت تعديل بافت نگار قرار دارد. نشان دهید که گذر دوم تعديل بافت نگار دقیقاً همان نتیجه گذر اول را تولید خواهد کرد.
- ۴.۴** تصویری دارای PDF سطح خاکستری، $p_r(r)$ ، به صورت شکل زیر، است. مطلوب است که سطوح خاکستری تصویر را طوری تبدیل نماییم که PDF حاصل، $p_z(z)$ ، نیز به صورت زیر نشان داده شده باشد. کمیت ها را پسونه فرض کنید و تبدیلی (بر حسب r و z) باید که این کار را انجام دهد.



- ۵.۴** روشی برای بهنگام کردن بافت نگار محلی برای استفاده در روش ارتقای محلی تشریح شده در بخش ۴.۲ پیشنهاد نماید.
- ۶.۴** درستی معادلات (۴.۲.۱۹-۲۰) و (۴.۲.۲۰) را ثابت کنید.
- ۷.۴** در یک کاربرد صنعتی برای بازبینی درون مذابهای آهن، باید از تصویربرداری اشعه ایکس استفاده شود. هدف جستجوی حباب های درون مذاب است که معمولاً در تصویر به صورت حباب های کوچک ظاهر می شوند. در اغلب موارد، نویز زیاد، بازبینی را مشکل می سازد. بنابراین تصمیم گرفته می شود با استفاده از متوسط گیری تصویر، نویز آن کاهش داده شود تا

در نتیجه تمایز مرئی ببینید. برای محاسبه 'متوسط در این کاربرد خاص، لازم است که تعداد تصاویر تا حد ممکن کم باشد تا بدین ترتیب زمانی که باید قطعات به منظور تصویربرداری ساکن باشند، کاهش یابد. پس از آزمایش های زیاد، این نتیجه به دست آمد که در این کاربرد کاهش واریانس نویز به یک دهم کافی است. اگر وسیله 'تصویربرداری بتواند ۳۰ قاب بر ثانیه تولید کند، چه مدت باید مذاهبها هنگام تصویربرداری ساکن بمانند تا به کاهش مطلوب در واریانس نویز دست یابیم؟ فرض کنید که نویز ناهمبسته و دارای میانگین صفر است.

اثر محدودکننده 'کاربرد مکرر یک فیلتر مکانی پایین گذر 3×3 بر تصویر رقمی را شرح دهید.
۸.۴ می توانید از اثرات مرزی^۱ صرفنظر نمایید.

پیاده سازی فیلتر های مکانی نیازمند حرکت دادن مرکز یک نقاب در تمام تصویر و سپس محاسبه 'جمع حاصل ضرب های ضرایب نقاب با پیکسل های متناظر در هر محل است (معادله '۴-۳. ۱-۳. را ببینید). در مورد فیلتر کردن پایین گذر، تمام ضرایب "۱" هستند که استفاده از فیلتر جعبه ای^۲ یا الگوریتم متوجه^۳ را ممکن می سازد. در این الگوریتم تنها در بخشی از محاسبه که از یک محل به محل بعدی تغییر می کند، بهنگام سازی اجرا می شود.

(الف) رابطه 'چنین الگوریتمی برای یک فیلتر $n \times n$ را بیان کنید طوری که نه تنها ماهیت محاسبات بلکه دنباله 'پیمایش مورد استفاده برای حرکت دادن نقاب حول تصویر را نیز نشان دهد.

(ب) نسبت تعداد محاسبه های انجام شده در پیاده سازی ساده^۴ به تعداد محاسبات انجام شده با الگوریتم فیلتر جعبه ای را بهره 'محاسباتی^۵ می نامند. بهره 'محاسباتی را در این مورد به دست آورید و آن را به عنوان تابعی از n برای $1 < n$ رسم کنید. ضریب مقیاس $1/n^2$ برای هر دو روش مشترک است، بنابراین در به دست آوردن بهره 'محاسباتی، لازم نیست آن را در نظر بگیرید. فرض کنید که تصویر یک حاشیه' بیرونی از صفحه های دارد که به اندازه 'کافی پهن بوده و به شما امکان می دهد در تحلیل آن از اثرات حاشیه ای صرف نظر کنید.

۹.۰ (الف) یک روش عملی برای محاسبه 'میانه در یک همسایگی $n \times n$ ارائه نمایید.

(ب) روشی برای بهنگام سازی میانه وقتی مرکز همسایگی از یک پیکسل به پیکسل بعدی جا به جا می شود، پیشنهاد کنید.

1) Border effect 2) Box filter 3) Moving-average 4) Brute-force 5) Computational advantage

- ۱۱.۴ نشان دهد که طبق توضیح بخش ۳.۳.۴ می توان تصویر فیلتر بالاگذر شده را در میدان مکان با رابطه' ریاضی، پایین گذر - اولیه=بالاگذر، بدست آورد. برای سهولت فیلترها را 3×3 فرض کنید.
- ۱۲.۴ ستون دوم شکل ۴.۱۹. مقاطع عرضی تابع انتقال یک فیلتر بالاگذر و پاسخ ضربه متناظر آن را نشان می دهد. برای فیلتر تقویت بالا ، تابع انتقال و پاسخ ضربه شبیه چیست؟
- ۱۳.۴ فرض کنید که فیلتر مکانی پایین گذری می سازید که از چهار همسایه' نقطه' (y, x) متوسط می گیرد (بخش ۲.۴.۱ را بینید)، اما خود نقطه' (y, x) را وارد متوسط گیری نمی کند.
- (الف) $H(u, v)$ ، فیلتر معادل آن در میدان فرکانس، را باید.
- (ب) نشان دهد که نتیجه' قسمت (الف) یک فیلتر پایین گذر است.
- ۱۴.۴ روش پایه ای مورد استفاده برای محاسبه' گرادیان رقمه (بخش ۳.۳.۴)، محاسبه' تفاضل به شکل رابطه' $f(x, y) - f(x + 1, y)$ است.
- (الف) $H(u, v)$ ، تابع انتقال فیلتر برای انجام فرآیند معادل در میدان فرکانس، را باید.
- (ب) نشان دهد که این فیلتر بالاگذر است.
- ۱۵.۴ در چه شرایطی فیلتر پایین گذر با تورث داده شده در معادله' (۴.۴-۴)، یک فیلتر پایین گذر ایده ال می شود؟
- ۱۶.۴ نشان دهد که تصویر بالاگذر فیلتر شده در میدان فرکانس را می توان با تفریق تصویر پایین گذر فیلتر شده از تصویر اصلی بدست آورد.
- (ب) به عنوان یک مثال بارز ، با معادله' فیلتر با تورث پایین گذر شروع کنید و از مفهوم قسمت (الف) برای تولید متناظر بالاگذر آن استفاده کنید.
- ۱۷.۴ یک روال متدائل برای ارتقای تصویر، ترکیب دوروش تأکید فرکانس بالا و تعدیل بافت نگار می باشد تا بدین وسیله تیزکردن لبه و ارتقای تمایز به طور همزمان حاصل شوند.
- (الف) آیا ترتیب این دو فرآیند اهمیت دارد؟
- (ب) اگر ترتیب مهم است، دلیل اصلی برای استفاده از یک ترتیب خاص را بیان کنید.
- ۱۸.۴ فرض کنید که مجموعه ای از تصاویر تولید شده در آزمایش تجربی که به تحلیل و قایع نجومی می پردازد، به شما داده می شود. هر تصویر دارای مجموعه ای از نقاط روشن دور از هم،

متناظر با ستاره های دور از هم در بخشی از آسمان، می باشد. مشکل این است که بر اثر روش نابی ناشی از پخش اتمسفری، ستاره ها به سختی قابل رویت هستند. اگر این تصاویر به صورت حاصل ضرب یک مؤلفه 'روشنابی' ثابت در مجموعه ای از ضربه ها مدل شوند، یک روال ارتقا براساس فیلتر کردن همشکلی پیشنهاد کنید که بتواند بخش های مربوط به ستاره ها را از تصویر جدا نماید.

۱۹.۴ بارجoue به توضیع بخش ۴.۵، نشان دهید که اگر $H(u, v)$ حقیقی و متقارن باشد، $(y(x, y))$

نیز حقیقی و متقارن خواهد بود.

۲۰.۴ در یک سامانه خودکار مونتاژ، سه گروه از قطعات باید توسط رنگ رمز شوند تا بین وسیله تشخیص آنها آسان گردد. از طرفی، تنها یک دوربین تلویزیونی تک رنگ موجود است. روشی برای استفاده از این دوربین در تشخیص سه رنگ مختلف پیشنهاد کنید.

۲۱.۴ نشان دهید که برای هر نقطه P که روی مثلث رنگی HSI قراردارد، معادله $(۴.۶-۲۱)$ معتبر است.

۲۲.۴ یک پزشک متخصص، مسؤول بازرگانی از دسته ای تصاویر که بوسیله 'میکروسکوپ الکترونی تولید شده' اند، است. برای راحتی کار، او تصمیم می گیرد که از ارتقای تصویر رقمه بهره بگیرد و با این هدف مجموعه ای از تصاویر نمونه را بررسی می کند و با این مشکل ها مواجه می شود: (۱) وجود نقاط روشن جدا از هم که مورد علاقه نیستند؛ (۲) تیزی کم؛ (۳) تمايز ناکافی در بعضی تصاویر؛ و (۴) تغییراتی در مقدار متوسط سطح خاکستری، در حالی که این مقدار باید K باشد تا بتوان سنجش های معینی از شدت را انجام داد. او می خواهد این مشکلات را حل کرده و سپس تمام سطوح خاکستری بین $1/\omega_1$ و $1/\omega_2$ را به رنگ قرمز ثابت تبدیل نماید و در همان حال سایر سطوح خاکستری را بدون تغییر باقی گذارد. دنباله ای از مراحل پردازش را پیشنهاد کنید که این پزشک بتواند برای وصول به نتیجه 'مطلوب' از آن استفاده نماید.

۲۳.۴ توضیع دهید چرا نقطه 'وسط نوار چند رنگ' از پلیت VIII (الف) سبز خالص به نظر می رسد، در حالی که طبق شکل ۴.۵، آن نقطه ترکیبی از مؤلفه های قرمز و آبی با مقدار مساوی و

مؤلفه سبز با مقدار دو برابر آنها می باشد.

۴۴.۴ از شما خواسته شده تا یک سامانه پردازش تصویر طرح کنید که قادر به تشخیص بین صفحاتی

از بدنه اتومبیل^۱ که شکل یکسان اما رنگ متفاوت دارند، باشد. قطعات روی نقاله 'خط مونتاژ حرکت می کنند. هدف شناسایی هر قطعه بحسب رنگ آن است تا این که رویات آن را بردارند و روی سایر خطوط مونتاژ بگذارند. رنگ های قطعات مختلفی که روی خط مونتاژ حرکت می کنند، قرمز، زرد، سبز و آبی هستند. شماتجهیزات زیر را در اختیار دارید: یک دوربین ویدیویی رنگی با خروجی های RGB، یک رقمی ساز رنگی که سیگنال های ویدیویی آنالوگ RGB را دریافت می کند و قادر به تولید تصاویر رقمی RGB یا HSI در نرخ قاب متدداول (یعنی هر ثانیه ۳۰ تصویر رنگی کامل) می باشد، سه بافر قاب که قادر به دریافت تصاویر در نرخ قاب ویدیویی هستند، و یک واحد سخت افزاری که قادر به محاسبه بافت نگار تصویر رقمه در نرخ قاب است. همه این سخت افزارهای پردازش تصویر در یک PC مجهر مجتمع شده اند. هدف، طراحی یک سامانه نرم افزاری است که با استفاده از سخت افزار موجود بتواند با سرعت بالارنگ قطعات را تعیین کند. چگونه می توان تنها با استفاده از مفاهیم بیان شده در این فصل، چنین سامانه ای را طراحی کرد؟ می توانید فرض کنید که قطعات به اندازه کافی کُند حرکت می کنند طوری که ماتی تصاویر رقمی شده ناچیز است. طرح خود را به صورت نمودار جریان^۲ ارائه دهید و در آن منظور از هر عمل و دلیل انتخاب آن را به طور مفصل توضیح دهید.

فصل ۵

بازیابی تصویر

«اشیا تنها آنچه می بینیم نیستند... واقعیت وجودی اشیا کاملاً برای ما مجھول باقی می ماند ما در حد حواس خود اشیا را درک می کنیم.»
امانوئل کانت

همانند ارتقای تصویر، هدف نهایی روش های بازیابی، بهبود تصویر از بعضی جهات است. برای تفکیک قابل شدن بین این دو عمل، بازیابی را فرآیندی در نظر می گیریم که می خواهد با استفاده از اطلاعات موجود درباره 'پدیده' خراب کننده تصویر^۱، تصویر خراب شده را بازسازی یا اصلاح نماید. بنابراین روش های بازیابی برای مدل کردن پدیده^۲ تخریب و اجرای فرآیند معکوس آن برای بازیابی تصویر اولیه، به کار می روند. این فرآیند معمولاً یک معیار کیفیت را به کار می گیرد تا تخمین نسبتاً بهینه ای از نتیجه^۳ مطلوب را تیجه دهد. بر عکس، روش های ارتقا، مبتنی بر مشاهده^۴ بوده و برای ایجاد تغییر در تصویر در راستای بهره گیری از جنبه های جسمی- روانی^۵ دستگاه بینایی انسان طراحی می شوند. مثلاً کشش تمایز^۶ یک روش ارتقا محسوب می شود، زیرا به دلیل جنبه های خواشیدی که برای بیننده فراهم می کند، به کار می رود، در حالی که حذف ماتی تصویر با استفاده از تابع مات زدا^۷ یک فرآیند بازیابی محسوب می شود.

روش های اولیه^۸ بازیابی تصویر رقمی، غالباً از مفاهیم حوزه^۹ فرکانس به دست آمدند. این فصل بر یک روش جبری نو، که به دست آوردن روش های بازیابی متعددی را از اصول پایه ای مشابه ممکن می سازد، متمرکز است. گرچه برای حل مستقیم مسأله پاروش های جبری معمولاً با سامانه های بزرگی از معادلات همزمان، سروکار داریم، در این فصل نشان داده می شود که با وجود شرایط معینی، می توان

1) Degradation phenomenon

2) Heuristic procedure

3) Psychophysical

4) Contrast stretching

5) deblurring function

پیچیدگی محاسباتی را به سطح مناسب برای روش های بازیابی متداول در حوزه 'فرکانس، کاهش داد. مطالب بیان شده در این فصل کاملاً مقدماتی است. در اینجا مسأله بازیابی را، تنها از نقطه ای که تصویررقemi^۱ خراب شده، در دست است، بررسی می کنیم؛ و در مورد خرابی های ناشی از حسگر، رقمی ساز و نمایشگر سخن گفته نمی شود. گرچه این مباحث در برخورد کلی با کاربردهای بازیابی تصویر، مهم است، اما خارج از بحث فعلی ما، می باشند. مراجع آورده شده در انتهای بخش راهنمای مناسبی در مورد این موضوع ها و مباحث مرتبط هستند.

۱.۵ مدل تخریب

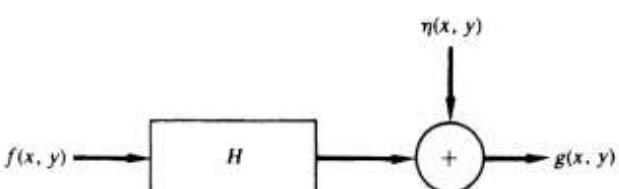
همان طور که شکل ۱.۱ نشان می دهد، فرآیند تخریب در این فصل به صورت عملگر یا سامانه H مدل می شود که همراه نویز جمع شونده^۲ (y, η) بر تصویر ورودی $(x, f(x, y))$ اعمال می شود تا تصویر تخریب شده^۳ $(y, g(x, y))$ را ایجاد نماید. بازیابی تصویر رقمی را می توان فرآیندی در نظر گرفت که با داشتن $(y, g(x, y))$ و اطلاع از تخریب به شکل عملگر H ، تقریبی از تصویر $(y, f(x, y))$ را بدست می آورد. فرض می کنیم که آگاهی ما از (y, η) تنها به طبیعت آماری آن محدود است.

۱.۱.۵ چند تعريف

رابطه ورودی-خروجی در شکل ۱.۱ به صورت

$$g(x, y) = H[f(x, y)] + \eta(x, y). \quad (1-1.5)$$

بیان می شود. اکنون اجازه دهید فرض کنیم که $\eta(x, y) = H[f(x, y)]$ ، بنابراین $g(x, y) = H[f(x, y)] + H[f(x, y)]$. آنگاه H خطی است اگر:



شکل ۱.۵ مدلی از فرآیند تخریب تصویر

۳۱۱ ۱.۵ مدل تغیریب

$$H[k_1 f_1(x, y) + k_2 f_2(x, y)] = k_1 H[f_1(x, y)] + k_2 H[f_2(x, y)] \quad (۲-۱. ۵)$$

که k_1 و k_2 دو مقدار ثابت و $f_1(x, y)$ و $f_2(x, y)$ دو تصویر ورودی دلخواه هستند.

اگر $k_1 = k_2 = 1$ باشد، معادله (۲-۱. ۵) به صورت

$$H[f_1(x, y) + f_2(x, y)] = H[f_1(x, y)] + H[f_2(x, y)] \quad (۳-۱. ۵)$$

در می آید که خاصیت جمع پذیری^۱ خوانده می شود؛ این خاصیت می گوید که اگر H یک عملگر خطی باشد، پاسخ آن به جمع دو ورودی برابر جمع پاسخ های آن به هریک از ورودی ها می باشد.

با $f_2(x, y) = 0$ ، معادله (۲-۱. ۵) به صورت

$$H[k_1 f_1(x, y)] = k_1 H[f_1(x, y)] \quad (۴-۱. ۵)$$

در می آید که خاصیت همگنی^۲ خوانده می شود. این خاصیت می گوید که پاسخ به یک مضرب ثابت از هر ورودی برابر با حاصلضرب پاسخ به آن ورودی در همان ثابت می باشد. بنابراین یک عملگر خطی هر دو خاصیت جمع پذیری و همگنی را دارد.

یک عملگر با رابطه ورودی-خروجی $g(x, y) = H[f(x, y)]$ مستقل از مکان^۳ گفته می شود اگر

$$H[f(x - \alpha, y - \beta)] = g(x - \alpha, y - \beta) \quad (۵-۱. ۵)$$

برای هر $f(x, y)$ و هر α و β . این تعریف بیان می کند که پاسخ در هر نقطه از تصویر تنها به مقدار ورودی در آن نقطه بستگی دارد و به مکان آن نقطه بستگی ندارد.

۲. ۱. ۵ مدل تغیریب برای توابع پیوسته

با یک تغییر جزئی (اما معادل) در طرز نمایش تعریف تابع ضربه، می توان معادله (۳. ۴۶-۳) را به صورت

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha, \beta) \delta(x - \alpha, y - \beta) d\alpha d\beta. \quad (۶-۱. ۵)$$

1) Additivity

2) Homogeneity

3) Position (or space) invariant

بیان کرد. حال اگر در معادله (۱-۱. ۵) $\eta(x, y) = 0$ باشد، داریم:

$$g(x, y) = H[f(x, y)] = H\left[\iint_{-\infty}^{\infty} f(\alpha, \beta) \delta(x - \alpha, y - \beta) d\alpha d\beta\right]. \quad (V-1. 5)$$

اگر H یک عملگر خطی باشد و خاصیت جمع پذیری را به انتگرالها نیز تعمیم دهیم، خواهیم داشت:

$$g(x, y) = \iint_{-\infty}^{\infty} H[f(\alpha, \beta) \delta(x - \alpha, y - \beta)] d\alpha d\beta. \quad (A-1. 5)$$

از آنجاکه $f(\alpha, \beta)$ مستقل از x و y است، از خاصیت همگنی داریم که

$$g(x, y) = \iint_{-\infty}^{\infty} f(\alpha, \beta) H[\delta(x - \alpha, y - \beta)] d\alpha d\beta. \quad (A-1. 5)$$

عبارت

$$h(x, \alpha, y, \beta) = H[\delta(x - \alpha, y - \beta)] \quad (10-1. 5)$$

پاسخ ضربه^۱ H خوانده می شود. به زیان دیگر، اگر در معادله (۱-۱. ۵) $\eta(x, y) = 0$ باشد، آنگاه $h(x, \alpha, y, \beta)$ به ضربه ای با قدرت واحد در مختصات (α, β) است. طبق توضیع بخش ۲. ۱. در نورشناسی^۲، ضربه یک نقطه نور است و به $h(x, \alpha, y, \beta)$ معمولاً با عنوان تابع پخش نقطه ای^۳ اشاره می شود.

جایگذاری معادله (۱۰-۱. ۵) در معادله (۹-۱. ۵) عبارت

$$g(x, y) = \iint_{-\infty}^{\infty} f(\alpha, \beta) h(x, \alpha, y, \beta) d\alpha d\beta \quad (11-1. 5)$$

را بدست می دهد که انتگرال جمع آثار نوع اول^۴ خوانده می شود. این عبارت در نظریه سامانه خطی اهمیت اساسی دارد. این عبارت بیان می کند که اگر پاسخ H به یک ضربه معلوم باشد، پاسخ به هر ورودی دلخواه $f(\alpha, \beta)$ را می توان با معادله (۱۱-۱. ۵) محاسبه نمود. به عبارت دیگر، سامانه خطی

1) Impulse response

2) Optics

3) Point spread function (PSF)

4) Superposition (or Fredholm) integral of the first kind

۱.۵ مدل تخریب ۳۱۳

H را می توان با پاسخ ضربه اش کاملاً توصیف کرد.

اگر H مستقل از مکان باشد، از معادله (۱۱-۱) داریم

$$H[\delta(x - \alpha, y - \beta)] = h(x - \alpha, y - \beta). \quad (12-1. 5)$$

معادله (۱۱-۱) در این حالت به صورت

$$g(x, y) = \iint_{-\infty}^{\infty} f(\alpha, \beta) h(x - \alpha, y - \beta) d\alpha d\beta \quad (13-1. 5)$$

ساده می شود که همان انتگرال کاتولوشن تعریف شده در معادله (۳-۳۰) می باشد.

با وجود نویز جمع شونده، عبارتی که یک مدل تخریب خطی را بیان می کند، به صورت:

$$g(x, y) = \iint_{-\infty}^{\infty} f(\alpha, \beta) h(x, \alpha, y, \beta) d\alpha d\beta + \eta(x, y). \quad (14-1. 5)$$

می باشد. اگر H مستقل از مکان باشد، معادله (۱۴-۱) به صورت

$$g(x, y) = \iint_{-\infty}^{\infty} f(\alpha, \beta) h(x - \alpha, y - \beta) d\alpha d\beta + \eta(x, y). \quad (15-1. 5)$$

در می آید. البته در هر دو حالت فرض می شود که نویز تصویر، مستقل از مکان است.

بسیاری از انواع تخریب را می توان با فرآیندهای خطی مستقل از مکان تقریب زد. مزیت این کار این است که ابزارهای گسترشده ای از نظریه سامانه های خطی برای حل مسائل بازیابی تصویر در دسترس قرار می گیرند. روش های غیرخطی و متغیر با مکان، گرچه کلی تر (و معمولاً دقیق تر) هستند، اما مشکلاتی را نیز به همراه دارند که اغلب یا هیچ حل شناخته شده ای ندارند و یا حل آنها محاسبات خیلی مشکلی دارد. این فصل روی روش های بازیابی خطی مستقل از مکان تمرکز می کند. در هر حال حتی این ساده سازی نیز می تواند مشکلات محاسباتی بوجود آورد که اگر مستقیماً با آنها مقابله شود، خارج از توانایی های عملی اغلب رایانه های امروزی هستند.

۳۰.۱.۵ بیان گستره

استنتاج یک مدل تخریب گستره مستقل از مکان، با شروع از حالت یک بعدی و حذف موقت

عبارت نویز ساده می شود. فرض کنید که دوتابع $f(x)$ و $h(x)$ به طور یکتاخت نمونه برداری شوند تا آرایه هایی به ترتیب با ابعاد $A \times B$ را تشکیل دهند. در این مورد، x یک متغیر گسته در محدوده $0, 1, 2, \dots, A-1$ برای $f(x)$ و در محدوده $0, 1, 2, \dots, B-1$ برای $h(x)$ است.

رابطه کانولوشن گسته موجود در بخش ۳.۳، براساس این فرض است که توابع نمونه برداری شده، متناوب و با دوره تناوب M هستند. با انتخاب $1 = A + B - M \geq 0$ و گسترش توابع با صفر طوری که طولشان برابر M شود، از همپوشانی دوره های مجزا در حاصل کانولوشن جلوگیری می شود. با فرض این که $f(x)$ و $h(x)$ توابع گسترش یافته هستند، با استفاده از معادله (۳.۲۹-۳) کانولوشن آنها به صورت $x = 0, 1, 2, \dots, M-1$ است:

$$g_r(x) = \sum_{m=0}^{M-1} f_r(m)h_r(x-m) \quad (16-1.5)$$

محاسبه می گردد. چون فرض می شود که $f_r(x)$ و $h_r(x)$ هردو دوره تناوب M دارند، $g_r(x)$ نیز همین دوره تناوب را دارد.

با استفاده از نمایش ماتریسی، می توان معادله (۱۶-۱.۵) را به شکل

$$\mathbf{g} = \mathbf{H}\mathbf{f} \quad (17-1.5)$$

یکان کرد که \mathbf{f} و \mathbf{g} بردارهای ستونی M بعدی هستند:

$$\mathbf{f} = \begin{bmatrix} f_r(0) \\ f_r(1) \\ \vdots \\ f_r(M-1) \end{bmatrix} \quad (18-1.5)$$

$$\mathbf{g} = \begin{bmatrix} g_r(0) \\ g_r(1) \\ \vdots \\ g_r(M-1) \end{bmatrix} \quad (19-1.5)$$

و \mathbf{H} یک ماتریس $M \times M$ به شکل زیر است:

و

۱.۵ مدل تحریب ۳۱۵

$$\mathbf{H} = \begin{bmatrix} h_r(0) & h_r(-1) & h_r(-2) & \cdots & h_r(-M+1) \\ h_r(1) & h_r(0) & h_r(-1) & \cdots & h_r(-M+2) \\ h_r(2) & h_r(1) & h_r(0) & \cdots & h_r(-M+3) \\ \vdots & & & & \\ h_r(M-1) & h_r(M-2) & h_r(M-3) & \cdots & h_r(0) \end{bmatrix} \quad (20-1.5)$$

از فرض تناوبی بودن $h_r(x) = h_r(M+x)$. با استفاده از این خاصیت می توان معادله (20-1.5) را به شکل زیرنوشت:

$$\mathbf{H} = \begin{bmatrix} h_r(0) & h_r(M-1) & h_r(M-2) & \cdots & h_r(1) \\ h_r(1) & h_r(0) & h_r(M-1) & \cdots & h_r(2) \\ h_r(2) & h_r(1) & h_r(0) & \cdots & h_r(3) \\ \vdots & & & & \\ h_r(M-1) & h_r(M-2) & h_r(M-3) & \cdots & h_r(0) \end{bmatrix} \quad (21-1.5)$$

ساختار این ماتریس نقشی اساسی در ادامه این فصل بازی می کند. در معادله (21-1.5) سطرها با یک انتقال چرخشی^۱ به راست باهم مرتبط هستند؛ یعنی، راست ترین عنصر در یک سطر برابر چپ ترین عنصر در سطر بعدی آن است. این انتقال، چرخشی خوانده می شود زیرا هر عنصر که از انتهای راست یک سطر به خارج انتقال یابد، دوباره در انتهای چپ سطر بعدی ظاهر می شود. گذشته از این، در معادله (21-1.5) چرخشی بودن \mathbf{H} از این نظر که سطر آخر نیز همان رابطه را با سطر اول دارد، کامل می باشد. یک ماتریس مربعی که هر سطر آن انتقال چرخشی سطر قبلی باشد، و سطر اول نیز انتقال چرخشی سطر آخر باشد، ماتریسی چرخشی^۲ خوانده می شود. به خاطر بسپارید که رفتار چرخشی \mathbf{H} نتیجه مستقیمی از فرض تناوبی بودن $h_r(x)$ است.

مثال: فرض کنید که $A=4$ و $B=3$. می توان انتخاب کرد که $M=6$ باشد و سپس دو صفر به نمونه های $f(x)$ و $g(x)$ افزود. در این حالت f و g بردارهای ۶ بعدی هستند و \mathbf{H} یک ماتریس 6×6 است:

1) Circular shift

2) Circulant matrix

$$\mathbf{H} = \begin{bmatrix} h_r(0) & h_r(5) & h_r(4) & \cdots & h_r(1) \\ h_r(1) & h_r(0) & h_r(5) & \cdots & h_r(2) \\ h_r(2) & h_r(1) & h_r(0) & \cdots & h_r(3) \\ \vdots & & & & \\ h_r(5) & h_r(4) & h_r(3) & \cdots & h_r(0) \end{bmatrix}$$

در هر حال وقتی بازای $h_r(x) = h(x)$ ، $x = 0, 1, 2$ و بازای $h_r(x) = 0$ ، $x = 3, 4, 5$ باشد، داریم:

$$\mathbf{H} = \begin{bmatrix} h(0) & & h(2) & h(1) \\ h(1) & h(0) & & h(2) \\ h(2) & h(1) & h(0) & \end{bmatrix}$$

□

که تمام عناصر غیرمذکور در ماتریس، صفر هستند.

تعیین بحث به مدل تخریب گسته دو بعدی ساده است. برای دو تصویر رقمه شده $f(x, y)$ و $h(x, y)$ که به ترتیب با ابعاد $C \times D$ و $A \times B$ هستند، می‌توان با افزودن صفر به آنها تصاویری با ابعاد $M \times N$ را بدست آورد. همانطور که در بخش ۳.۸.۳. بیان شد، یک روش انجام این کار بدین صورت

است که قرار دهیم

$$f_r(x, y) = \begin{cases} f(x, y) & 0 \leq x \leq A - 1 & 0 \leq y \leq B - 1 \\ 0 & A \leq x \leq M - 1 & B \leq y \leq N - 1 \end{cases}$$

و

$$h_r(x, y) = \begin{cases} h(x, y) & 0 \leq x \leq C - 1 & 0 \leq y \leq D - 1 \\ 0 & C \leq x \leq M - 1 & D \leq y \leq N - 1. \end{cases}$$

با فرض این که توابع گسترش یافته $(y, f_r(x, y))$ و $(y, h_r(x, y))$ درجهات x و y متناوب به ترتیب با دوره های تناوب M و N هستند، معادله (۳.۳۵-۳) کاتولوشن این دو تابع را به صورت

۳۱۷ ۱.۵ مدل تخریب

$$g_r(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_r(m, n) h_r(x - m, y - n) \quad (22-1.5)$$

برای $1 \leq y \leq N-1$ و $0 \leq x \leq M-1$ بحسب مدل تخریب می دهد. تابع کانولوشن (x, y) نیز متناظر با همان دوره تناوب (x, y) و $f_r(x, y)$ است. با انتخاب $N \geq B + D - 1$ و $M \geq A + C - 1$ می شود. برای تکمیل مدل تخریب گستته، یک عبارت نویز گسته گسترش یافته با ابعاد $N \times M$ ، $\eta_r(x, y)$ ، به معادله (۲۲-۱.۵) افزوده می شود طوری که برای $0 \leq x \leq M-1$ و $0 \leq y \leq N-1$ داریم:

$$g_r(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_r(m, n) h_r(x - m, y - n) + \eta_r(x, y) \quad (23-1.5)$$

اجازه دهید، \mathbf{f} و \mathbf{g} و \mathbf{n} بیانگر بردارهای ستونی MN بعدی تشکیل شده از پشت سطرهای توابع $M \times N$ عنصری $f_r(x, y)$ ، $g_r(x, y)$ و $\eta_r(x, y)$ باشند. مثلاً عنصر اول \mathbf{f} ، عنصر اول \mathbf{g} و عنصر اول \mathbf{n} هستند، عنصر بعدی از سطر دوم بوده و به همین ترتیب الی آخر. این قرارداد امکان می دهد که معادله (۲۳-۱.۵) به شکل برداری-ماتریسی^۱ بیان شود:

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n} \quad (24-1.5)$$

که \mathbf{f} و \mathbf{n} دارای ابعاد $1 \times (MN)$ و \mathbf{H} دارای ابعاد $MN \times MN$ می باشد. این ماتریس دارای M^2 بخش است که هر بخش به ابعاد $N \times N$ است و طبق

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_{M-1} & \mathbf{H}_{M-2} & \cdots & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & \mathbf{H}_{M-1} & \cdots & \mathbf{H}_2 \\ \mathbf{H}_2 & \mathbf{H}_1 & \mathbf{H}_0 & \cdots & \mathbf{H}_3 \\ \vdots & & & & \\ \mathbf{H}_{M-1} & \mathbf{H}_{M-2} & \mathbf{H}_{M-3} & \cdots & \mathbf{H}_0 \end{bmatrix} \quad (25-1.5)$$

مرتب شده اند. هر بخش \mathbf{H}_i به صورت زیر از سطر زام تابع گسترش یافته (x, y) و $h_r(x, y)$ تشکیل می شود:

1) Vector- matrix form

$$\mathbf{H}_j = \begin{bmatrix} h_r(j, 0) & h_r(j, N-1) & h_r(j, N-2) & \cdots & h_r(j, 1) \\ h_r(j, 1) & h_r(j, 0) & h_r(j, N-1) & \cdots & h_r(j, 2) \\ h_r(j, 2) & h_r(j, 1) & h_r(j, 0) & \cdots & h_r(j, 3) \\ \vdots & & & & \\ h_r(j, N-1) & h_r(j, N-2) & h_r(j, N-3) & \cdots & h_r(j, 0) \end{bmatrix} \quad (26-1.5)$$

که مانند معادله (۵.۲۱-۱)، از تناوبی بودن $(y, h_r(x))$ استفاده شده است. در اینجا \mathbf{H} یک ماتریس چرخشی است و بلوکهای \mathbf{H} به روشنی چرخشی اندیس گذاری شده‌اند. به این دلایل، ماتریس \mathbf{H} در معادله (۵.۲۵-۱) اغلب ماتریس بلوک-چرخشی^۱ خوانده می‌شود.

بیشتر توضیحات بخش‌های بعدی درباره مدل تخریب گسته داده شده در معادله (۵.۲۴-۱) می‌باشد. به خاطر بسیاری‌د که بدست آوردن این عبارت براساس فرض فرآیند تخریب نامتفاوت با مکان خطی بود. همان طور که قبلاً ذکر شد، هدف ما تخمین تصویر $(y, f(x))$ با داشتن $(y, g(x))$ و داشتن اطلاعاتی در مورد $(y, h(x))$ است. از نظر معادله (۵.۲۴-۱)، این هدف نیازمند تخمین f با داشتن g و داشتن مقداری اطلاع در مورد \mathbf{H} و n است.

گرچه معادله (۵.۲۴-۱) مفهوم ظاهرآ ساده‌ای دارد، حل مستقیم آن برای به دست آوردن عناصر f برای تصاویر با ابعاد عملی نیازمند پردازش‌های زیادی است. برای مثال اگر $M = N = 512$ باشد، \mathbf{H} دارای ابعاد $262,144 \times 262,144$ است. پنابراین بدست آوردن مستقیم f نیازمند حل یک دستگاه حاوی ۲62,144 معادله خطی همزمان است. خوشبختانه می‌توان با بهره گیری از خواص چرخشی \mathbf{H} تا حد زیادی پیچیدگی مسأله را کاهش داد.

۲.۵ قطری‌سازی ماتریسهای چرخشی و بلوک-چرخشی

در این بخش نشان می‌دهیم که راه حل‌هایی که از نقطه نظر محاسباتی قابل تحقق هستند، را می‌توان از مدل معادله (۵.۲۴-۱) و با قطری کردن ماتریس \mathbf{H} بدست آورد. جهت ساده‌سازی، بحث را با در نظر گرفتن ماتریسهای چرخشی شروع می‌کنیم و سپس آن را به ماتریسهای بلوک-چرخشی تعمیم می‌دهیم.

1) Block-circulant matrix

۲.۵ قطعی سازی ماتریسهای چرخشی و بلوك-چرخشی ۳۱۹

۱.۴.۵ ماتریسهای چرخشی

ماتریس چرخشی \mathbf{H} با ابعاد $M \times M$ را در نظر بگیرید

$$\mathbf{H} = \begin{bmatrix} h_r(0) & h_r(M-1) & h_r(M-2) & \cdots & h_r(1) \\ h_r(1) & h_r(0) & h_r(M-1) & \cdots & h_r(2) \\ h_r(2) & h_r(1) & h_r(0) & \cdots & h_r(3) \\ \vdots & & & & \\ h_r(M-1) & h_r(M-2) & h_r(M-3) & \cdots & h_r(0) \end{bmatrix} \quad (1-2.5)$$

اجازه دهید تابع اسکالر (k) و بردار (k) \mathbf{w} را به صورت

$$\begin{aligned} \lambda(k) = h_r(0) + h_r(M-1) \exp\left[j \frac{2\pi}{M} k\right] + h_r(M-2) \exp\left[j \frac{2\pi}{M} 2k\right] \\ + \cdots + h_r(1) \exp\left[j \frac{2\pi}{M} (M-1)k\right] \end{aligned} \quad (2-2.5)$$

تعريف کنیم که $j = \sqrt{-1}$ است، و برای $1, 2, \dots, M-1$ داریم $k = j$

$$\mathbf{w}(k) = \begin{bmatrix} 1 \\ \exp\left[j \frac{2\pi}{M} k\right] \\ \exp\left[j \frac{2\pi}{M} 2k\right] \\ \vdots \\ \exp\left[j \frac{2\pi}{M} (M-1)k\right] \end{bmatrix} \quad (3-2.5)$$

با ضرب ماتریسی می توان نشان داد که

$$\mathbf{H}\mathbf{w}(k) = \lambda(k)\mathbf{w}(k). \quad (4-2.5)$$

این عبارت نشان می دهد که $\mathbf{w}(k)$ یک بردار ویژه ماتریس چرخشی \mathbf{H} و $\lambda(k)$ مقدار ویژه متناظر با آن است (بخش ۳.۶ را بینید).

سپس اجازه دهید ماتریس \mathbf{W} با ابعاد $M \times M$ را با استفاده از M بردار ویژه \mathbf{H} به عنوان ستونهای آن

تشکیل دهیم:

$$\mathbf{W} = [\mathbf{w}(0) \quad \mathbf{w}(1) \quad \mathbf{w}(2) \quad \cdots \quad \mathbf{w}(M-1)]. \quad (5-2.5)$$

عنصر $k^{\text{ام}} W$ ، که با (k, i) نشان داده می شود، برای $i = 0, 1, 2, \dots, M-1$ با رابطه

$$W(k, i) = \exp \left[j \frac{2\pi}{M} ki \right] \quad (6-2.5)$$

بدست می آید. خواص تعامل جملات نمایی مختلط، نوشتن ماتریس معکوس \mathbf{W}^{-1} را ممکن می سازد؟

عنصر $k^{\text{ام}} W^{-1}$ که با نماد (k, i) نشان داده می شود، عبارتست از:

$$W^{-1}(k, i) = \frac{1}{M} \exp \left[-j \frac{2\pi}{M} ki \right]. \quad (7-2.5)$$

از معادلات (5.6-2) و (5.7-2) داریم:

$$\mathbf{W}\mathbf{W}^{-1} = \mathbf{W}^{-1}\mathbf{W} = \mathbf{I} \quad (8-2.5)$$

که \mathbf{I} ماتریس همانی $M \times M$ است.

اهمیت وجود ماتریس معکوس \mathbf{W}^{-1} این است که استقلال خطی¹ ستون های \mathbf{W} (بردارهای ویژه \mathbf{H}) را تضمین می کند. سپس از نظریه ماتریس مقدماتی (Noble [1969]) می توان \mathbf{H} را به شکل

$$\mathbf{H} = \mathbf{WDW}^{-1} \quad (9-2.5)$$

یا

$$\mathbf{D} = \mathbf{W}^{-1}\mathbf{HW} \quad (10-2.5)$$

بیان کرد که \mathbf{D} یک ماتریس قطری است که عناصر قطر آن مقادیر ویژه \mathbf{H} هستند؛ یعنی،

$$D(k, k) = \lambda(k). \quad (11-2.5)$$

معادله (5.10-2) بیان می کند که \mathbf{H} با استفاده از \mathbf{W}^{-1} و \mathbf{W} و به ترتیب فوق قطری می شود.

1) Linearly independent

۲.۵ قطری سازی ماتریس های چرخشی و بلوک- چرخشی ۳۲۱

۲.۴.۵ ماتریس های بلوک- چرخشی

ماتریس تبدیل برای قطری سازی بلوک های چرخشی^۱ به صورت زیر تشکیل می شود. اجازه دهید که

$$w_M(i, m) = \exp \left[j \frac{2\pi}{M} im \right] \quad (12-2. 5)$$

و

$$w_N(k, n) = \exp \left[j \frac{2\pi}{N} kn \right]. \quad (13-2. 5)$$

براساس این نحوه نمایش، ماتریس \mathbf{W} با ابعاد $MN \times MN$ را که دارای M^2 بخش با ابعاد $N \times N$ است، تعریف می کنیم. بخش $\lim_{i, m=0, 1, 2, \dots, M-1}$ را برای \mathbf{W} به صورت

$$\mathbf{W}(i, m) = w_M(i, m) \mathbf{W}_N \quad (14-2. 5)$$

تعریف می کنیم. آنگاه \mathbf{W} یک ماتریس $N \times N$ با عناصر

$$W_N(k, n) = w_N(k, n) \quad (15-2. 5)$$

برای $k, n = 0, 1, 2, \dots, N-1$ است.

ماتریس معکوس \mathbf{W}^{-1} نیز به ابعاد $MN \times MN$ است که دارای M^2 بخش با ابعاد $N \times N$ می باشد.

بخش $\lim_{i, m=0, 1, 2, \dots, M-1}$ که با نامad $(i, m)^{-1}$ \mathbf{W}^{-1} نشان داده می شود، عبارتست از

$$\mathbf{W}^{-1}(i, m) = \frac{1}{M} w_M^{-1}(i, m) \mathbf{W}_N^{-1} \quad (16-2. 5)$$

که $w_M^{-1}(i, m)$ برای $i, m = 0, 1, 2, \dots, M-1$ عبارتست از:

$$w_M^{-1}(i, m) = \exp \left[-j \frac{2\pi}{M} im \right] \quad (17-2. 5)$$

ماتریس \mathbf{W}^{-1} عناصر

$$W_N^{-1}(k, n) = \frac{1}{N} w_N^{-1}(k, n) \quad (18-2. 5)$$

1) Block circulant

را دارد که برای $k, n = 0, 1, 2, \dots, N-1$ به صورت زیر است:

$$w_n^{-1}(k, n) = \exp\left[-j \frac{2\pi}{N} kn\right] \quad (19-2.5)$$

با جایگذاری مستقیم عناصر \mathbf{W} و \mathbf{W}^{-1} می توان صحت معادله

$$\mathbf{WW}^{-1} = \mathbf{W}^{-1}\mathbf{W} = \mathbf{I} \quad (20-2.5)$$

را ابررسی کرد که \mathbf{I} ماتریس همانی $MN \times MN$ است.

با استفاده از نتایج بخش ۵.۲.۱. و اگر \mathbf{H} یک ماتریس بلوک-چرخشی باشد، می توان نشان داد

: (Hunt[1973]) که:

$$\mathbf{H} = \mathbf{WDW}^{-1} \quad (21-2.5)$$

یا

$$\mathbf{D} = \mathbf{W}^{-1}\mathbf{HW} \quad (22-2.5)$$

که \mathbf{D} یک ماتریس قطری است که عناصر قطر آن با تبدیل فوریه گستته تابع گسترش یافته $h_r(x, y)$ که در بخش ۵.۳.۱ تشریح شد، مرتبط هستند. گذشته از آن، ترانهاده \mathbf{H} ، که با \mathbf{H}^T نشان داده می شود، عبارتست از

$$\mathbf{H}^T = \mathbf{WD}^*\mathbf{W}^{-1} \quad (23-2.5)$$

که \mathbf{D}^* مزدوج مختلط \mathbf{D} است.

۳.۲.۵ انوات قطری سازی بر مدل تغیری

ماتریس \mathbf{H} در مدل یک بعدی گستته' معادله (۵.۱-۱۷) چرخشی است، بنابراین می توان آن را به شکل معادله (۵.۱-۱۷) بیان کرد. بنابراین معادله (۵.۱-۱۷) به صورت

$$\mathbf{g} = \mathbf{WDW}^{-1}\mathbf{f} \quad (24-2.5)$$

در می آید. با تغییر شکل معادله فوق داریم:

۲.۵ تقطیری سازی ماتریسها چرخشی و بلوك-چرخشی ۳۴۴

$$\mathbf{g} = \mathbf{W} \mathbf{D} \mathbf{W}^{-1} \mathbf{f} \quad (25-2.5)$$

حاصل ضرب $\mathbf{W}^{-1} \mathbf{f}$ یک بردارستونی M بعدی است. از معادله (۵.۷-۲) و تعریف \mathbf{f} در بخش ۱.۵، عنصر k ام حاصل ضرب $\mathbf{W}^{-1} \mathbf{f}$ ، که با $F(k)$ نشان داده می شود، برای $k = 0, 1, 2, \dots, N-1$ عبارتست از:

$$F(k) = \frac{1}{M} \sum_{i=0}^{M-1} f_r(i) \exp\left[-j \frac{2\pi}{M} ki\right] \quad (26-2.5)$$

این عبارت به عنوان تبدیل فوریه گستته دنباله گسترش یافته $(x)_r$ شناخته می شود. به عبارت دیگر، ضرب \mathbf{f} در \mathbf{W}^{-1} برداری بدست می دهد که عناصرش تبدیل فوریه های \mathbf{f} هستند. به طریق مشابه $\mathbf{g} = \mathbf{W}^{-1} \mathbf{f}$ تبدیل فوریه عناصر \mathbf{g} را می دهد که با بردار $G(k)$ بازای $k = 0, 1, 2, \dots, M-1$ نشان داده می شود. حال ماتریس \mathbf{D} در معادله (۵.۲۵-۲) را بررسی می کنیم. شرح بخش ۲.۱ نشان داد که عناصر قطر اصلی \mathbf{D} مقادیر ویژه ماتریس چرخشی \mathbf{H} هستند. با استفاده از این واقعیت که

$$\exp\left[j \frac{2\pi}{M}(M-i)k\right] = \exp\left[-j \frac{2\pi}{M}ik\right] \quad (27-2.5)$$

می توان مقادیر ویژه داده شده در معادله (۵.۲.۲) را به شکل

$$\begin{aligned} \lambda(k) &= h_r(0) + h_r(1) \exp\left[-j \frac{2\pi}{M}k\right] + h_r(2) \exp\left[-j \frac{2\pi}{M}2k\right] \\ &\quad + \dots + h_r(M-1) \exp\left[-j \frac{2\pi}{M}(M-1)k\right]. \end{aligned} \quad (28-2.5)$$

نیز نوشته. از معادله (۱۱-۲) و (۵.۲۸-۲) برای $k = 0, 1, 2, \dots, M-1$ داریم:

$$D(k, k) = \lambda(k) = \sum_{i=0}^{M-1} h_r(i) \exp\left[-j \frac{2\pi}{M}ki\right] \quad (29-2.5)$$

طرف راست این معادله $MH(k)$ است، که $H(k)$ تبدیل فوریه گستته دنباله گسترش یافته $(x)_r$ است. بنابراین

$$D(k, k) = MH(k). \quad (30-2.5)$$

این تبدیلات را می توان دریک ترکیب کرد. چون \mathbf{D} یک ماتریس قطری است، حاصل ضرب

D با هر بردار، هر عنصر آن بردار را با تنها یک عنصر قطری **D** ضرب می کند. در نتیجه می توان رابطه ماتریسی داده شده در معادله (۵. ۲۵-۲) را به صورت ضرب عنصر به عنصر دنباله های تبدیل فوریه یک بعدی ساده کرد. به عبارت دیگر، برای $k = 0, 1, 2, \dots, M-1$ داریم:

$$G(k) = MH(k)F(k) \quad (31-2. 5)$$

که $G(k)$ عناصر بردار $\mathbf{g}^{-1}\mathbf{W}$ و $MH(k)F(k)$ عناصر بردار $\mathbf{D}\mathbf{W}^{-1}\mathbf{f}$ هستند. طرف راست معادله (۳۱-۲. ۵) کانتولوشن $f_r(x)$ و $h_r(x)$ در میدان فرکانس است (بخش ۳. ۳. ۸ را ببینید). از نظر محاسباتی، این نتیجه، ساده سازی زیادی را به همراه دارد، زیرا $G(k)$ ، $H(k)$ و $F(k)$ تبدیل های گستته M نمونه ای هستند، که می توان آنها را با یک الگوریتم تبدیل فوریه سریع بدست آورد. با ضرب با روشی نظیر کارهای قبلی، نتایج معادلی برای مدل تحریب دو بعدی بدست می آید. با ضرب دو طرف معادله (۵. ۱. ۲۴) در \mathbf{W}^{-1} و استفاده از معادلات (۵. ۲۰-۲) و (۵. ۲۱-۲)، خواهیم داشت:

$$\mathbf{W}^{-1}\mathbf{g} = \mathbf{D}\mathbf{W}^{-1}\mathbf{f} + \mathbf{W}^{-1}\mathbf{n} \quad (32-2. 5)$$

که \mathbf{W}^{-1} یک ماتریس $MN \times MN$ است که عناصرش با معادله (۵. ۱۶-۲) داده می شوند، **D** یک ماتریس قطری $MN \times MN$ است، **H** ماتریس بلوک-چرخشی تعریف شده در معادله (۵. ۲۵-۱) است، و \mathbf{f} و \mathbf{g} بردارهایی به طول MN هستند که به ترتیب از دنبال هم گذاردن سطرهای تصاویر گسترش یافته $f_r(x)$ و $g_r(x, y)$ تشکیل شده اند.

طرف چپ معادله (۵. ۳۲-۲) برداری با ابعاد $J \times MN$ است. اجازه دهید عناصرش را با ترتیب $G(0, 0), G(0, 1), \dots, G(0, N-1); G(1, 0), G(1, 1), \dots, G(1, N-1); \dots; G(M-1, 0), G(M-1, 1), \dots, G(M-1, N-1)$ نشان دهیم. می توان تشان داد (Hunt[1973]) که برای $u = 0, 1, 2, \dots, M-1$ و $v = 0, 1, 2, \dots, N-1$

$$G(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} g_r(x, y) \exp\left[-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right] \quad (33-2. 5)$$

معادله (۵. ۳۳-۲) رابطه تبدیل فوریه دو بعدی (x, y) و (u, v) است. به عبارت دیگر عناصر $\mathbf{g}^{-1}\mathbf{W}$ متناظر با

۲.۵ قطری سازی ماتریس‌های چرخشی و بلوك-چرخشی ۳۲۵

سطرهای دنیال هم گذاشته شده ماتریس تبدیل فوریه با عناظر $G(u, v)$ بازای $1 \leq u \leq M-1$ و $1 \leq v \leq N-1$ هستند. به طور مشابه بردارهای $\mathbf{W}^{-1}\mathbf{n}$ و $\mathbf{W}^{-1}\mathbf{f}$ بعدی بوده و دارای عناظر $N(u, v)$ و $F(u, v)$ هستند. بنابراین بازای $1 \leq u \leq M-1$ و $1 \leq v \leq N-1$ می‌توان نوشت:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f_r(x, y) \exp\left[-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right] \quad (34-2.5)$$

و

$$N(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \eta_r(x, y) \exp\left[-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right] \quad (35-2.5)$$

نهایتاً عناظر ماتریس قطری D مربوط به تبدیل فوریه تابع پاسخ ضربه گسترش یافته $h_r(x, y)$ هستند؛ یعنی، برای $1 \leq u \leq M-1$ و $1 \leq v \leq N-1$ داریم:

$$H(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} h_r(x, y) \exp\left[-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right] \quad (36-2.5)$$

عنصر قطری D به صورت زیر تشکیل می‌شوند: N عنصر اول عناظر سطر اول H بوده و MN عنصر بعدی عناظر سطر دوم H می‌باشد، و بدین ترتیب، N عنصر آخر عناظر سطر آخر H هستند. البته عناظر غیرقطری D صفر می‌باشند. سپس ماتریس قطری مشکل از عناظر فوق در MN ضرب می‌شوند تا D بدست آید. یک روش کوتاهتر، بیان این ماتریس به صورت

$$D(k, i) = \begin{cases} MNH\left(\left[\frac{k}{N}\right], k \bmod N\right) & \text{اگر } i = k \\ 0 & \text{اگر } i \neq k \end{cases} \quad (37-2.5)$$

است که $[c]$ برای دلالت بر بزرگترین عدد صحیح ناگزیرگتر از c (جزء صحیح c) بکار می‌رود، و $k \bmod N$ باقی مانده تقسیم k بر N است.

با استفاده از معادلات $(33-2.5) - (36-2.5)$ می‌توان نشان داد که هریک از عناظر معادله

$$(32-2.5) \quad \text{برای } 1 \leq u \leq M-1 \text{ و } 1 \leq v \leq N-1 \quad \text{با عبارت}$$

$$G(u, v) = MNH(u, v)F(u, v) + N(u, v) \quad (38-2.5)$$

به هم مرتبط هستند.

جزء MN تنها یک ضریب مقیاس است که برای سادگی می توان آن را در $H(u, v)$ ادغام نمود. با این

نحوه نمایش، می توان معادلات (۵. ۳۷-۲) و (۵. ۳۸-۲) را به صورت

$$D(k, i) = \begin{cases} H\left(\frac{k}{N}\right), k \bmod N & i = k \\ 0 & i \neq k \end{cases} \quad (5. ۳۹-۲)$$

برای $k, i = 0, 1, 2, \dots, MN-1$ و به صورت

$$G(u, v) = H(u, v)F(u, v) + N(u, v) \quad (5. ۴۰-۲)$$

برای $u=0, 1, 2, \dots, N-1$ و $v=0, 1, 2, \dots, M-1$ بیان کرد.

اهمیت معادله (۵. ۳۸-۲) یا (۵. ۴۰-۲) این است که دستگاه بزرگ معادلات نهفته در مدل معادله (۵. ۲۴-۱) را می توان به محاسبه تعداد اندکی تبدیل فوریه گسته با ابعاد $N \times M$ ساده کرد. به عنوان مثال اگر M و N توان های صحیح ۲ باشند، در صورت استفاده از الگوریتم FFT مسئله ساده خواهد شد. در هر حال همان طور که قبلاً ذکر شد، اگر مستقیماً با مدل معادله (۵. ۱-۱) با این مسئله برخورد کنیم، حل آن به یک کارمحاسباتی تقریباً غیرممکن تبدیل می شود.

در بخش های بعدی از مدل معادله (۵. ۱-۱) به عنوان مبنای برای بدست آوردن چند روش بازیابی تصویر استفاده می کنیم. آنگاه با استفاده از مفاهیم معرفی شده در این بخش نتایج ماتریسی حاصل را ساده می کنیم. به خاطر داشته باشید که ساده سازی های بدست آمده نتیجه فرض های زیر هستند: (۱) تخریب یک فرآیند نامتغیر با مکان و خطی است، و (۲) با تمام تصاویر به عنوان توابع متناوب گسترش یافته رفتار می شود.

معادله (۵. ۴۰-۲) را می توان مستقیماً با استفاده از قضیه کانولوشن نوشت. در هر حال هدف ما این بود که نشان دهیم، با روابط ماتریسی نیز می توان به همان نتیجه دست یافت. بدین ترتیب، تعدادی از خواص ماتریسی مهم را جهت استفاده در بخش ۳. ۵ بدست آوردهیم تا این که به روش واحدی^۱ برای بازیابی دست یابیم.

1) Unified

۳.۵ رهیافت جبری بازیابی

همان طور که در بخش ۱.۳ ذکر شد، هدف بازیابی تصویر، تخمین تصویر اولیه \hat{f} از تصویر خراب شده g با داشتن اطلاعات یا فرضهای در مورد H است. با فرض این که این مقادیر طبق مدل معادله (۵.۱-۲۴) با هم رابطه دارند، بیان گروهی از مسائل بازیابی تصویر را در یک قالب جبری خطی واحد ممکن می شود.

محور رهیافت جبری، مفهوم جستجوی تخمینی از \hat{f} ، که با \hat{f} ذکر می شود، است که یک معیار کارآمدی پیش تعریف شده را حداقل سازد. این فصل روی توابع معیار حداقل مربعات، به علت سادگی شان، تمرکز می کند. مزیت این انتخاب دستیابی به یک رهیافت محوری برای به دست آوردن چند روش بازیابی کاملاً شناخته شده می باشد. این روش ها حاصل استفاده از یک رهیافت نامقید یا مقید برای مسئله بازیابی حداقل مربعات (با حداقل مجموع مربعات خطأ) هستند.

۱.۳.۵ بازیابی نامقید

طبق معادله (۵.۱-۲۴)، جزء نویز در مدل تخریب عبارتست از:

$$\mathbf{n} = \mathbf{g} - \mathbf{H}\hat{\mathbf{f}} \quad (1-3.5)$$

بدون هیچ اطلاعی در مورد n ، یکتابع معیار معنادار این است که با فرض تاحد ممکن کوچک بودن $\|\mathbf{n}\|^2$ جزء نویز، در جستجوی \hat{f} باشیم چنانکه $\mathbf{H}\hat{\mathbf{f}}$ ، \mathbf{g} را در مفهوم حداقل مربعات تقریب بزند. به عبارت دیگر می خواهیم \hat{f} را چنان بیابیم که

$$\|\mathbf{n}\|^2 = \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2 \quad (2-3.5)$$

حداقل شود، که طبق تعریف،

$$\|\mathbf{n}\|^2 = \mathbf{n}^T \mathbf{n} \quad \text{و} \quad \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2 = (\mathbf{g} - \mathbf{H}\hat{\mathbf{f}})^T (\mathbf{g} - \mathbf{H}\hat{\mathbf{f}})$$

به ترتیب مجدوز نرم های $\mathbf{H}\hat{\mathbf{f}} - \mathbf{g}$ و \mathbf{n} هستند. معادله (۵.۲-۳) دید معادلی از این مسئله را به دست می دهد. با این دید، تابعی از \hat{f} به عنوان تابع معیار که باید حداقل شود، انتخاب گردیده و به صورت زیر

تعريف می شود:

$$J(\hat{\mathbf{f}}) = \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2 \quad (3-3.5)$$

غیر از شرایط لازمی که معادله (5.3-3) را حداقل می کند، $\hat{\mathbf{f}}$ از هیچ نظر دیگری مقید نیست. حداقل سازی معادله (5.3-3) ساده است. تنها از J نسبت به $\hat{\mathbf{f}}$ مشتق می گیریم و نتیجه را برابر بردار صفر قرار می دهیم؛ یعنی،

$$\frac{\partial J(\hat{\mathbf{f}})}{\partial \hat{\mathbf{f}}} = \mathbf{0} = -2\mathbf{H}^T(\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}). \quad (4-3.5)$$

از حل معادله (4-3.5) برای $\hat{\mathbf{f}}$ داریم:

$$\hat{\mathbf{f}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{g}. \quad (5-3.5)$$

با فرض $N = M$ ، ماتریس \mathbf{H} مربعی می شود و با فرض وجود \mathbf{H}^{-1} ، معادله (5.5-3) به صورت زیر ساده می شود:

$$\begin{aligned} \hat{\mathbf{f}} &= \mathbf{H}^{-1}(\mathbf{H}^T)^{-1}\mathbf{H}^T\mathbf{g} \\ &= \mathbf{H}^{-1}\mathbf{g}. \end{aligned} \quad (6-3.5)$$

۲۰.۳.۵ بازیابی مقید

در این بخش مسأله بازیابی حداقل مربعات رامانند کمینه سازی توابع به شکل $\|\mathbf{Q}\hat{\mathbf{f}}\|^2$ در نظر می گیریم، که \mathbf{Q} یک عملگر خطی روی \mathbf{f} با قید $\|\mathbf{n}\|^2 = \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2$ است. این روش انعطاف پذیری زیادی در فرآیند بازیابی ایجاد می کند، زیرا بازی انتخابهای متفاوت \mathbf{Q} جواب های متفاوتی را نتیجه می دهد. قید اعمال شده بر جواب، با مدل معادله (5.1-24) هماهنگ¹ است.

با استفاده از روش ضرایب لاگرانژ² (Elsgolc[1961]) بدون مشکل می توان قید برابری را به مسأله کمینه سازی اضافه کرد. برای این کار باید قید به شکل $(\|\mathbf{n}\|^2 - \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2)^\alpha$ بیان و سپس به تابع $\|\mathbf{Q}\hat{\mathbf{f}}\|^2$ افزوده شود. به عبارت دیگر، در جستجوی یک $\hat{\mathbf{f}}$ هستیم که تابع معیار

$$J(\hat{\mathbf{f}}) = \|\mathbf{Q}\hat{\mathbf{f}}\|^2 + \alpha(\|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2 - \|\mathbf{n}\|^2) \quad (7-3.5)$$

1) Consistent

2) Lagrange multipliers

۴.۵ فیلتر کردن معکوس ۳۴

را حداقل کند. در معادله^{۱)} فوق α ثابتی است که ضریب لاگرانژ نامیده می شود. پس از افزودن قيد مذکور، کمینه سازی به روش معمول انجام می شود.

با مشتق گیری از معادله (۵. ۷-۳) نسبت به $\hat{\mathbf{f}}$ و برابر صفر قرار دادن بردار حاصل داریم:

$$\frac{\partial J(\hat{\mathbf{f}})}{\partial \hat{\mathbf{f}}} = \mathbf{0} = 2\mathbf{Q}^T \mathbf{Q}\hat{\mathbf{f}} - 2\alpha \mathbf{H}^T(\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}). \quad (۸-۳. ۵)$$

با حل معادله (۸-۳) برای $\hat{\mathbf{f}}$ جواب بدست می آید؛ یعنی،

$$\hat{\mathbf{f}} = (\mathbf{H}^T \mathbf{H} + \gamma \mathbf{Q}^T \mathbf{Q})^{-1} \mathbf{H}^T \mathbf{g} \quad (۹-۳. ۵)$$

که $\alpha = 1/\gamma$ است. برای برآورده شدن قيد مذکور باید کمیت γ تنظیم شود، که این مسأله را در جای دیگری در همین فصل بررسی می کنیم. معادلات (۶-۳. ۵) و (۹-۳. ۵) اساس تمام شیوه های بازیابی که در بخش های بعدی تشریح می شوند، هستند. مثلاً در بخش ۴. ۵، نشان می دهیم که معادله (۶-۳. ۵) روش متداول بازیابی با فیلتر معکوس را نتیجه می دهد. به طور مشابه از رابطه^{۲)} کلی معادله (۹-۳. ۵) می توان برای بدست آوردن نتایجی نظری فیلتروینر کلاسیک و سایر روش های بازیابی استفاده نمود. انجام چنین کاری تنها نیازمند انتخاب یک ماتریس تبدیل مناسب \mathbf{Q} و استفاده از ساده سازی های بدست آمده در بخش ۲. ۵ می باشد.

۴.۵ فیلتر کردن معکوس

۱.۴.۵ بیان روابط

استنتاج روش های بازیابی تصویر را با در نظر گرفتن نتیجه^{۱)} نامقید در معادله (۶-۳. ۵)، شروع می کنیم. با فرض $M = N$ و استفاده از معادله (۶-۳. ۵)، معادله (۶-۳. ۵) به صورت زیر می شود:

$$\begin{aligned} \hat{\mathbf{f}} &= \mathbf{H}^{-1} \mathbf{g} \\ &= (\mathbf{W} \mathbf{D} \mathbf{W}^{-1})^{-1} \mathbf{g} \\ &= \mathbf{W} \mathbf{D}^{-1} \mathbf{W}^{-1} \mathbf{g}. \end{aligned} \quad (۱-۴. ۵)$$

با پیش - ضرب کردن^{۱)} دو طرف معادله (۱-۴. ۵) در \mathbf{W}^{-1} داریم

$$\mathbf{W}^{-1} \hat{\mathbf{f}} = \mathbf{D}^{-1} \mathbf{W}^{-1} \mathbf{g}. \quad (۲-۴. ۵)$$

1) Premultiplying

از شرح بخش ۵. ۲. ۳. عناصر تشکیل دهنده' معادله (۵. ۴-۲) را می توان به شکل

$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)} \quad u, v = 0, 1, 2, \dots, N - 1 \quad (۵-۴. ۵)$$

نوشت. طبق معادله (۵. ۴-۲) فرض می شود که $H(u, v)$ در N^2 ضرب شود. چون D یک ماتریس قطری است، معکوس آن بسادگی بدست می آید.

رهیافت بازیابی تصویر معرفی شده با معادله (۵. ۴. ۳)، معمولاً با عنوان روش فیلتر معکوس اشاره می شود. این اصطلاح ناشی از این است که $(v, u) H(u, v)$ به عنوان تابع "فیلتری" در نظر گرفته می شود که در $F(u, v)$ ضرب می شود تا تبدیل تصویر تخریب شده $(y, x) g(x, y)$ بدست آید. آنگاه تقسیم $G(u, v)$ بر $H(u, v)$ مذکور در معادله (۵. ۴-۴)، یک عمل فیلتر کردن معکوس را تشکیل می دهد. البته تصویر بازیابی شده با استفاده از رابطه

$$\begin{aligned} \hat{f}(x, y) &= \hat{\mathcal{F}}^{-1}[\hat{F}(u, v)] \\ &= \hat{\mathcal{F}}^{-1}[G(u, v)/H(u, v)] \end{aligned} \quad (۴-۴. ۵)$$

برای $x, y = 0, 1, 2, \dots, N - 1$ بدست می آید. این روال معمولاً با استفاده از الگوریتم FFT اجرا می شود.

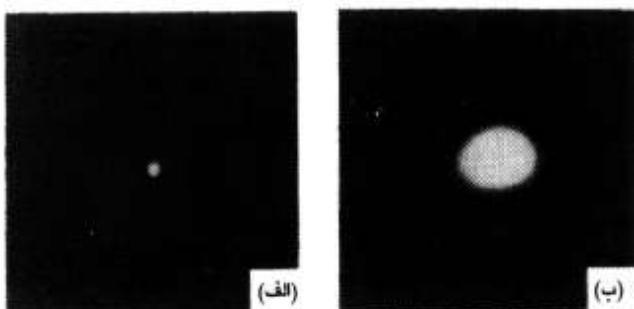
در معادله (۵. ۴-۴) به این نکته توجه کنید که اگر $H(u, v)$ در یکی از نواحی مورد علاقه در صفحه uv صفر یا خیلی کوچک باشد، در فرآیند بازیابی با مشکلات محاسباتی رویرو خواهیم شد. اگر صفرهای آنها صرفنظر کرد بدون این که اثر قابل توجهی بر نتیجه بازیابی بگذارند.

در حضور نویز مشکل جدی تری رخ می دهد. با جایگذاری معادله (۵. ۴-۲) در معادله (۵. ۴-۳)

داریم:

$$\hat{F}(u, v) = F(u, v) + \frac{N(u, v)}{H(u, v)}. \quad (۵-۴. ۶)$$

این عبارت به روشنی بیان می کند که اگر $H(u, v)$ صفر یا خیلی کوچک باشد، ممکن است اثر جزء $N(u, v)/H(u, v)$ در حاصل بازیابی $\hat{F}(u, v)$ غالب باشد. در عمل غالباً $H(u, v)$ با سرعت به صورت تابعی از فاصله تا مبدأ صفحه uv کاهش می یابد؛ با این حال معمولاً جزء نویز با سرعت کمتری

شکل ۲.۵ مات کردن یک منبع نقطه ای جهت بدست آوردن $H(u, v)$

کاهش می یابد. در چنین وضعیت هایی برای پرهیز از مقادیر کوچک $H(u, v)$ ، اغلب می توان با انجام بازیابی در یک همسایگی محدود حول مبدأ، به نتایج قابل قبولی دست یافت.

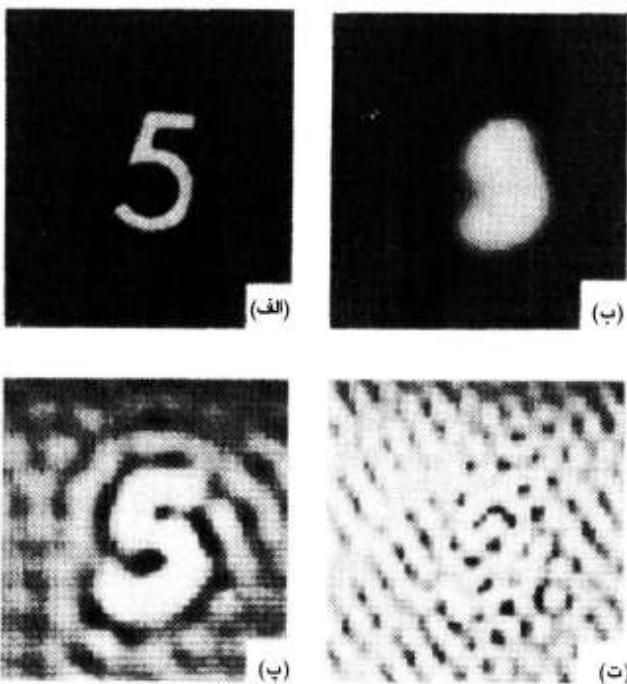
مثال: شکل ۲.۵ (الف) تصویر نقطه (x, y) و شکل ۲.۵ (ب) تصویر تخریب شده (x, y) حاصل مات شدگی (x, y) را نشان می دهد. با در نظر گرفتن این که منبع نقطه ای، تقریبی از تابع ضربه واحد باشد، نتیجه می شود

$$\begin{aligned} G(u, v) &= H(u, v)F(u, v) \\ &= H(u, v) \end{aligned}$$

زیرا $1 = [\delta(x, y)]^2$. این عبارت بیان می کند که تابع انتقال $H(u, v)$ را می توان با تبدیل فوریه تصویر تخریب شده تقریب زد. مات کردن یک تابع معلوم برای به دست آوردن تقریبی از $H(u, v)$ ، در عمل مفید است زیرا می توان از آن در رهیافت سعی و خطای جهت بازیابی تصاویری که در ابتدا مات کننده $H(u, v)$ معلوم نیست، استفاده کرد.

در شکل ۲.۵ (ب) حاصل اجرای تابع مات کننده فوق بر تصویر ایده آل شکل ۲.۵ (الف) دیده می شود. تصویر بازیابی شده شکل ۲.۵ (پ) با استفاده از معادله ۵.۴-۴ برای مقادیری از v, u که به اندازه کافی نزدیک مبدأ هستند، برای اجتناب از مقادیر بسیار کوچک $H(u, v)$ ، بدست آمد. در شکل ۲.۵ (ت) حاصل بازیابی برای یک همسایگی بزرگتر، نشان داده می شود. این نتایج به روشنی مشکلات بروز کرده بوسیله تابع میرای $H(u, v)$ را نشان می دهند.

اگر $N(u, v)$ ، $G(u, v)$ و $H(u, v)$ همه معلوم باشند، می توان از معادله ۵.۴-۲ مستقیماً یک عبارت دقیق برای فیلتر معکوس بدست آورد؛ یعنی



شکل ۳.۵ مثالی از بازیابی تصویر با فیلتر کردن معکوس: (الف) تصویر اولیه $f(x, y)$; (ب) تصویر تخریب شده (مات شده) $g(x, y)$; (پ) نتیجه بازیابی با استفاده از یک همسایگی حول مبدأ صفحه H که این همسایگی مقادیر بیش از حد کوچک $(u, v)H(u, v)$ را شامل نمی شود؛ (ت) حاصل استفاده از همسایگی بزرگتر که در آن این شرط برقرار نیست. (از [McGlamery 1967]).

$$F(u, v) = \frac{G(u, v)}{H(u, v)} - \frac{N(u, v)}{H(u, v)}. \quad (6-4.5)$$

علاوه بر مشکلات بالقوه $H(u, v)$ که درمثال قبلی تشریح شدند، مشکل دیگری که در مورد این رابطه وجود دارد این است که بذررت نویز کاملاً شناخته شده است تا این که بتوان $N(u, v)$ را محاسبه کرد.

۲۰.۴.۵ حذف ماتی حاصل از حرکت خطی یکنواخت

کاربردهای عملی بسیاری وجود دارد که در آنها می توان $H(u, v)R(u, v)$ را به طور تحلیلی به دست آورد، اما جواب مقادیر صفری در محدوده فرکانسی مورد علاقه دارد. در بخش ۴.۱، مثالی از مشکلات

۴.۵ فیلتر کردن معکوس ۳۳۳

ناشی از $H(u, v)$ میرا را ارائه کردیم. در توضیح بعدی، مسأله بازیابی تصویری را که بوسیله حرکت خطی یکنواخت مات شده است، بررسی می کنیم. این مسأله را بدلیل تأثیرات کاربردی آن و نیز بدلیل اینکه بخوبی با روش های تحلیلی توصیف می گردد، مورد بررسی خاص قرار می دهیم. همچنین جواب مات سازی یکنواخت روشی می سازد که چگونه می توان با صفحه های $H(u, v)$ در روند محاسباتی برخورد نمود. این ملاحظات بدلیل آن که در اغلب مباحث بازیابی تصویر بوسیله فیلتر کردن معکوس رخ می دهد، مهم و قابل توجه هستند.

فرض کنید که تصویر $f(x, y)$ حرکت صفحه ای دارد و $(t) x_0$ و $(t) y_0$ مؤلفه های متغیر با زمان حرکت به ترتیب درجهات x و y هستند. در این حالت نورخوردگی کلی در هر نقطه رسانه ضبط (مثلاً فیلم) با انتگرال گیری از نورخوردگی لحظه ای در دوره زمانی که دریچه^۱ باز است، بدست می آید. با فرض این که باز و بسته شدن دریچه، لحظه ای است و نیز این که فرآیند تصویربرداری نوری، کامل است، اثر حرکت تصویر جدا می شود. آنگاه اگر T طول دوره نورخوردگی باشد، تصویر مات شده (x, y) از رابطه^۲ زیر بدست می آید:

$$g(x, y) = \int_0^T f[x - x_0(t), y - y_0(t)] dt \quad (7-4.5)$$

از معادله (۷-۱.۹)، تبدیل فوریه معادله (۷-۴.۵) عبارتست از

$$\begin{aligned} G(u, v) &= \iint_{-\infty}^{\infty} g(x, y) \exp[-j2\pi(ux + vy)] dx dy \\ &= \iint_{-\infty}^{\infty} \left[\int_0^T f[x - x_0(t), y - y_0(t)] dt \right] \exp[-j2\pi(ux + vy)] dx dy. \end{aligned} \quad (8-4.5)$$

با معکوس کردن ترتیب انتگرال گیری می توان معادله (۸-۴.۵) را به شکل

$$G(u, v) = \int_0^T \left[\iint_{-\infty}^{\infty} f[x - x_0(t), y - y_0(t)] \exp[-j2\pi(ux + vy)] dx dy \right] dt. \quad (9-4.5)$$

یافته شده است. جزو داخل قلابهای^۳ بیرونی، تبدیل فوریه تابع جابجا شده $f[x - x_0(t), y - y_0(t)]$ است. آن گاه با استفاده از معادله (۳.۷-۳b) رابطه

$$\begin{aligned} G(u, v) &= \int_0^T F(u, v) \exp\{-j2\pi[ux_0(t) + vy_0(t)]\} dt \\ &= F(u, v) \int_0^T \exp\{-j2\pi[ux_0(t) + vy_0(t)]\} dt \end{aligned} \quad (10-4. 5)$$

بدست می آید که مرحله آخر از این واقعیت که $F(u, v)$ مستقل از t است، متنج گردید.
با تعریف این که

$$H(u, v) = \int_0^T \exp\{-j2\pi[ux_0(t) + vy_0(t)]\} dt \quad (11-4. 5)$$

معادله (5. 4-10) را می توان به شکل متداول

$$G(u, v) = H(u, v)F(u, v). \quad (12-4. 5)$$

بیان کرد. اگر طبیعت متغیرهای حرکت $(x_0(t))$ و $(y_0(t))$ معلوم باشد، می توان تابع انتقال $H(u, v)$ را مستقیماً از معادله (5. 4-11) بدست آورد. جهت تبیین موضوع، فرض کنید که تصویر مورد بررسی تحت حرکت خطی یکنواخت تنها درجهت محور x ها با معادله $x_0(t) = at/T$ قراردارد. در لحظه $T = t$ ، تصویر جمعبه طول a جایجا شده است. با قراردادن $0 = y_0(t)$ ، معادله (5. 4-11)

به صورت زیر در می آید:

$$\begin{aligned} H(u, v) &= \int_0^T \exp\{-j2\pi ux_0(t)\} dt \\ &= \int_0^T \exp\{-j2\pi uat/T\} dt \\ &= \frac{T}{\pi ua} \sin(\pi ua) e^{-j\pi ua}. \end{aligned} \quad (13-4. 5)$$

بوضوح H به ازای مقادیر u/a حاصل از رابطه $u = n/a$ که n یک عدد صحیح باشد، میرا و صفر می شود. وقتی $f(x, y)$ خارج از بازه $0 \leq x \leq L$ صفر (یا معلوم) باشد، می توان از مشکل حاصل از معادله (5. 4-13) دوری نمود و با اطلاع در مورد y در این بازه، تصویر کاملاً بازسازی می شود. چون y نامتغیر با زمان است، حذف این متغیر امکان می دهد که معادله (5. 4-7) را به طور موقت به صورت

$$\begin{aligned} g(x) &= \int_0^T f[x - x_0(t)] dt \\ &= \int_0^T f\left[x - \frac{at}{T}\right] dt \quad 0 \leq x \leq L. \end{aligned} \quad (14-4. 5)$$

۴.۵ فیلتر کردن معکوس ۴۳۵

بنویسیم. با جایگذاری $x = at/T$ در این عبارت و صرفنظر از یک ضریب مقیاس داریم که

$$g(x) = \int_{x-a}^x f(\tau) d\tau \quad 0 \leq x \leq L. \quad (15-4.5)$$

آنگاه با مشتق گیری نسبت به x (با استفاده از قاعده لایپنیتز^{۱)}، خواهیم داشت:

$$g'(x) = f(x) - f(x - a) \quad 0 \leq x \leq L \quad (16-4.5)$$

یا

$$f(x) = g'(x) + f(x - a) \quad 0 \leq x \leq L. \quad (17-4.5)$$

در ادامه کار مناسب است فرض کنیم که $Ka = L$ می باشد که Ka یک عدد صحیح است. آن گاه متغیر x را می توان به صورت

$$x = z + ma \quad (18-4.5)$$

بیان کرد که z مقادیری را در بازه $[0, K]$ اختیار می کند و m جزو صحیح (x/a) است. مثلاً اگر $a = 2$ و $x = 3.5$ باشد، آنگاه $1 = m$ (جزء صحیح $3.5/2$) و $z = 1.5$ خواهد بود. به روشنی همان طور که لازم است، $z + ma = 3.5$. همچنین توجه شود که برای $L = Ka$ ، اندیس m می تواند هر یک از مقادیر صحیح $1, 2, \dots, K-1$ را بگیرد. برای نمونه، اگر $x = L$ ، آن گاه $a = 1$ و $z = 0$ را بگیرید. با جایگذاری معادله (۱۸-۴.۵) در معادله (۱۷-۴.۵) داریم

$$f(z + ma) = g'(z + ma) + f[z + (m - 1)a]. \quad (19-4.5)$$

آن گاه با بیان $\phi(z)$ به عنوان بخشی از صحنه که در مدت نورخورده $a < z \leq L$ منتقل می شود، داریم:

$$\phi(z) = f(z - a) \quad 0 \leq z < a. \quad (20-4.5)$$

1) Liebnitz's rule

معادله (۵.۱۹-۴) را می توان با روش بازگشتی بر حسب $(z)\phi$ حل کرد. بنابراین برای $m = 0$

$$\begin{aligned} f(z) &= g'(z) + f(z - a) \\ &= g'(z) + \phi(z). \end{aligned} \quad (21-4. 5)$$

برای $m = 1$ ، معادله (۵.۱۹-۴) به صورت

$$f(z + a) = g'(z + a) + f(z). \quad (22-4. 5)$$

در می آید. با جایگذاری معادله (۵.۲۲-۴) در معادله (۵.۲۱-۴) داریم

$$f(z + a) = g'(z + a) + g'(z) + \phi(z). \quad (23-4. 5)$$

در مرحله بعد، با قراردادن $m = 2$ عبارت

$$f(z + 2a) = g'(z + 2a) + f(z + a) \quad (24-4. 5)$$

بدست می آید، که با جایگذاری معادله (۵.۲۳-۴) برای $f(z + a)$ در معادله فوق داریم:

$$f(z + 2a) = g'(z + 2a) + g'(z + a) + g'(z) + \phi(z). \quad (25-4. 5)$$

با ادامه این روال در نهایت داریم که

$$f(z + ma) = \sum_{k=0}^m g'(z + ka) + \phi(z). \quad (26-4. 5)$$

در هر حال وقتی $x = z + ma$ باشد، می توان معادله (۵.۲۶-۴) را به شکل

$$f(x) = \sum_{k=0}^m g'(x - ka) + \phi(x - ma) \quad 0 \leq x \leq L. \quad (27-4. 5)$$

بیان کرد. چون $(x)g$ معلوم است، مسأله به تخمین $(x)\phi$ ساده می شود.

یک روش برای تخمین مستقیم اینتابع از تصویر مات شده به صورت زیر است. ابتدا توجه کنید که وقتی x از 0 تا L تغییر می کند، m نیز از صفر تا $1 - K$ تغییر می کند. آرگومان ϕ ، $(x - ma)$ است که همواره در محدوده $a < x - ma \leq 0$ قرار دارد، بنابراین در مدت ارزیابی $f(x)$ برای $0 \leq x \leq L$

۴.۵ فیلتر کردن معکوس ۳۳۷

بار ϕ تکرار می شود. آنگاه با تعریف K

$$\tilde{f}(x) = \sum_{j=0}^m g'(x - ja) \quad (28-4.5)$$

می توان معادله (27-4.5) را به صورت

$$\phi(x - ma) = f(x) - \tilde{f}(x). \quad (29-4.5)$$

بازنویسی کرد. با ارزیابی دو طرف معادله (29-4.5) برای $a \leq x < (k+1)a$ و جمع نتایج برای $k = 0, 1, \dots, K-1$ داریم:

$$K\phi(x) = \sum_{k=0}^{K-1} f(x + ka) - \sum_{k=0}^{K-1} \tilde{f}(x + ka) \quad 0 \leq x < a \quad (30-4.5)$$

که چون $m = 0, 0 \leq x < a$ است. با تقسیم دو طرف بر K داریم

$$\phi(x) = \frac{1}{K} \sum_{k=0}^{K-1} f(x + ka) - \frac{1}{K} \sum_{k=0}^{K-1} \tilde{f}(x + ka). \quad (31-4.5)$$

البته اولین مجموع در طرف راست این عبارت، مجھول است. در هر حال برای مقادیر بزرگ K این مجموع به مقدار متوسط f نزدیک می شود. بنابراین می توان این مجموع را به عنوان ثابت A در نظر گرفت، که تقریب زیر را نتیجه می دهد:

$$\phi(x) \approx A - \frac{1}{K} \sum_{k=0}^{K-1} \tilde{f}(x + ka) \quad 0 \leq x < a \quad (32-4.5)$$

یا

$$\phi(x - ma) = A - \frac{1}{K} \sum_{k=0}^{K-1} \tilde{f}(x + ka - ma) \quad 0 \leq x \leq L. \quad (33-4.5)$$

با جایگذاری معادله (28-4.5) برای \tilde{f} داریم

[†] توجه کنید که حد بالایی مجموع درونی به جای m ، k است. اگر از معادله (18-4.5) با قرار دادن $x + ka - ma$ به جای x شروع کرد و بودیم، حد بالایی مجموع معادله (28-4.5)، k می بود زیرا از معادله (18-4.5) داریم $x + (ka - ma) = z + ma + (ka - ma) = z + ka$.

$$\begin{aligned}\phi(x - ma) &\approx A - \frac{1}{K} \sum_{k=0}^{K-1} \sum_{j=0}^k g'(x + ka - ma - ja) \\ &\approx A - \frac{1}{K} \sum_{k=0}^{K-1} \sum_{j=0}^k g'[x - ma + (k - j)a].\end{aligned}\quad (34-4. 5)$$

از معادلات (۵) و (۴-۲۸) و (۴-۲۹)، نتیجه نهایی

$$f(x) \approx A - \frac{1}{K} \sum_{k=0}^{K-1} \sum_{j=0}^k g'[x - ma + (k - j)a] + \sum_{j=0}^m g'(x - ja) \quad (35-4. 5)$$

را برای $0 \leq x \leq L$ خواهیم داشت. استفاده 'مجدد از متغیر محدود' y ، رابطه

$$f(x, y) = A - \frac{1}{K} \sum_{k=0}^{K-1} \sum_{j=0}^k g'[x - ma + (k - j)a, y] + \sum_{j=0}^m g'(x - ja, y) \quad (36-4. 5)$$

ربازای $L \leq x, y \leq 0$ می دهد. مانند قبل، فرض می شود که $f(x, y)$ تصویر مربعی است. با تعویض x و y در طرف راست معادله (۴-۳۶)، ساختار تصویری بدست می آید که در مدت نورخوردگی در جهت y حرکت می کند. از مقاهم ارائه شده می توان برای به دست آوردن یک عبارت مات زدایی که حرکت یکنواخت همزمان در دو جهت را نیز در نظر گیرد، استفاده کرد.

مثال: تصویر شکل ۵. ۴ (الف) در مدت نورخوردگی بوسیله حرکت خطی یکنواخت در یک جهت، مات شد، که مجموع فاصله پیموده شده تقریباً $1/8$ عرض عکس بود. شکل ۵. ۴ (ب) حاصل مات زدایی



(الف)



(ب)

شکل ۵. ۴ (الف) تصویر مات شده با حرکت خطی یکنواخت؛ (ب) تصویر بازیابی شده با استفاده از معادله (Sondhi [1972]). (از ۴-۳۶)

۵.۵ فیلتر کمینه میانگین مربعات (وینر)

۳۳۹

آن با استفاده از معادله (۵.۴-۳۶) است که چون حرکت در جهت y است، x و z تعویض شده اند. خطای تقریب داده شده با این معادله آزاردهنده نیست.

□

۵.۵ فیلتر کمینه میانگین مربعات (وینر)

اجازه دهد R_f و R_n ماتریس های همبستگی f و n باشند که به ترتیب با معادلات

$$R_f = E\{ff'\} \quad (1-5.5)$$

و

$$R_n = E\{nn'\} \quad (2-5.5)$$

تعريف شده اند. $E\{\cdot\}$ نشانه عملگر امید آماری است، و f و n به صورت تعریف شده در بخش ۱.۱.۵ هستند. عنصر j_{ij} از $R_f = E\{ff'\}$ می باشد، که همبستگی بین دو عنصر i_{ij} و j_{ij} است. به طریق مشابه عنصر j_{ij} از $R_n = E\{nn'\}$ همبستگی بین دو عنصر i_{ij} و j_{ij} را می دهد. چون عناصر f و n حقیقی هستند، $E\{ff'\} = E\{f,f\}$ ، $E\{n,n\} = E\{n,n\}$ است، و نتیجه می شود که $R_f = R_n$ ماتریس های متقارن حقیقی هستند. برای اغلب توابع تصویر، همبستگی بین پیکسل ها (یعنی عناصر f یا n) به خارج از فاصله ۲۰ تا ۳۰ پیکسل در تصویر گسترش نمی باید، بنابراین یک ماتریس همبستگی نامی، نواری از عناصر غیر صفر در حول قطر اصلی و صفرهای در نواحی گوشه راست بالا و چپ پایین دارد. براساس این فرض که همبستگی بین هر دو پیکسل تابعی از فاصله بین آن پیکسل ها است ته تابعی از مکان آنها، می توان $R_f = R_n$ را با ماتریس های بلوک-چرخشی تقریب زد و بنابراین می توان آنها را بوسیله ماتریس W و با فرآیند مذکور در بخش ۲.۵ قطعی کرد (Andrews [1977] و Hunt [1977]). با استفاده از A و B برای دلالت بر ماتریس ها داریم که

$$R_f = WAW^{-1} \quad (3-5.5)$$

و

$$R_n = WBW^{-1}. \quad (4-5.5)$$

دقیقاً وقتی عناصر ماتریس قطری D در رابطه $H = WDW^{-1}$ متناظر با تبدیل فوریه عناصر بلوکی H باشند، عناصر A و B ترتیب تبدیلهای عناصر همبستگی موجود در R_f و R_n خواهند بود. همان طور که

در مسئله ۴. ذکر شد، تبدیل فوریه این همبستگی ها به ترتیب طیف قدرت (یا چگالی طیفی) نامیده می شوند که در ادامه بحث آنها را با نامادهای $S_f(u, v)$ و $S_n(u, v)$ نشان

می دهیم.

با تعریف رابطه

$$\mathbf{Q}^T \mathbf{Q} = \mathbf{R}_f^{-1} \mathbf{R}_n \quad (5-5. 5)$$

و جایگذاری آن در معادله (۵. ۹-۳) داریم

$$\hat{\mathbf{f}} = (\mathbf{H}^T \mathbf{H} + \gamma \mathbf{R}_f^{-1} \mathbf{R}_n)^{-1} \mathbf{H}^T \mathbf{g}. \quad (6-5. 5)$$

با استفاده از معادلات (۵. ۲۱-۲)، (۵. ۲۳-۲)، (۵. ۳-۵) و (۵. ۵) داریم

$$\hat{\mathbf{f}} = (\mathbf{W} \mathbf{D}^* \mathbf{D} \mathbf{W}^{-1} + \gamma \mathbf{A}^{-1} \mathbf{B} \mathbf{W}^{-1})^{-1} \mathbf{W} \mathbf{D}^* \mathbf{W}^{-1} \mathbf{g}. \quad (7-5. 5)$$

با ضرب دوطرف در \mathbf{W}^{-1} و انجام چند تغییر ماتریسی، معادله (۷-۵. ۵) به شکل

$$\mathbf{W}^{-1} \hat{\mathbf{f}} = (\mathbf{D}^* \mathbf{D} + \gamma \mathbf{A}^{-1} \mathbf{B})^{-1} \mathbf{D}^* \mathbf{W}^{-1} \mathbf{g} \quad (8-5. 5)$$

ساده می شود. با در خاطرداشتن معنای عناصر \mathbf{A} و \mathbf{B} ، دانستن این که ماتریس های درون پرانتزها قطری هستند، و نیز استفاده از مفاهیم بخش ۵. ۲، ۳، می توان معادله (۸-۵. ۵) را برای

به شکل $u, v = 0, 1, 2, \dots, N - 1$

$$\begin{aligned} \hat{F}(u, v) &= \left[\frac{H^*(u, v)}{|H(u, v)|^2 + \gamma [S_n(u, v)/S_f(u, v)]} \right] G(u, v) \\ &= \left[\frac{1}{|H(u, v)|^2 + \gamma [S_n(u, v)/S_f(u, v)]} \right] G(u, v) \end{aligned} \quad (9-5. 5)$$

نوشت که $|H(u, v)|^2 = H^*(u, v)H(u, v)$ و فرض می شود که $M = N$ است.

وقتی $\gamma = 1$ ، جزء درون قلاب های بیرونی در معادله (۹-۵. ۵) به فیلتر وینر^۱ معروف ساده می شود.

اگر γ متغیر باشد، این عبارت فیلتر وینر پارامتری^۲ خوانده می شود. در غیاب نویز، $0 = S_n(u, v)$

1) Wiener filter

2) Parametric wiener filter

۵.۵ فیلتر کمینه میانگین مربعات (وینر) ۳۴۱

است و فیلتر وینر به فیلتر معکوس ایده آل تشریح شده در بخش ۵.۴ ساده می شود. در هر حال وقتی $\gamma = 1$ است، استفاده از معادله (۵.۹-۵)، دیگر جواب بهینه در مفهوم تعریف شده در بخش ۵.۲، ۳.۵ را نمی دهد، زیرا همان طور که در آن بخش بیان شد، لزماً باید برای برآورده شدن قید $\|\mathbf{g} - \mathbf{Hf}\|^2 = \|\mathbf{n}\|^2$ تنظیم شود. در هر حال می توان نشان داد که جواب به دست آمده با $\gamma = 1$ در این مفهوم که کمیت $E\{[f(x, y) - \hat{f}(x, y)]^2\}$ را حداقل می سازد، بهینه است. بروشی این روش یک معیار آماری است که با f و \hat{f} به عنوان متغیرهای تصادفی برخورد می کند.

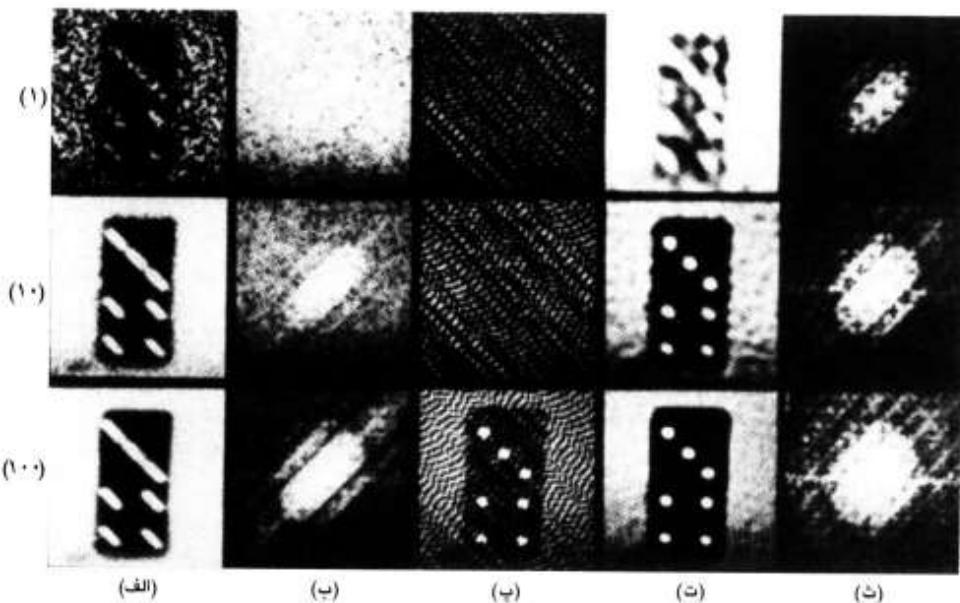
وقتی $S_n(u, v)$ و $S_f(u, v)$ مجهول باشند (مشکلی که غالباً در عمل با آن برخورد می کنیم)، گاهی اوقات مفید است که معادله (۵.۹-۵) را با رابطه'

$$\hat{F}(u, v) = \left[\frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + K} \right] G(u, v) \quad (10-5.5)$$

تقریب بزنیم که K مقدار ثابتی می باشد. در ادامه مثالی از نتایج بدست آمده با (۵.۱۰-۵) آورده می شود. در بخش ۵.۶ مسئله انتخاب ۷/۱ بهینه برای بازیابی تصویر را مقداری تشریح می کنیم.

مثال: ستون اول شکل ۵.۵ سه عکس از یک تاس^۱ را نشان می دهد که بوسیله حرکت خطی (در زاویه ۴۵°-نسبت به افق) و نویزی که واریانس آن در هر نقطه تصویر متناسب با روشناهی آن نقطه است، خراب شده است. همان طور که در سمت چپ شکل ۵.۵ دیده می شود، این سه تصویر با تغییر ضریب تناسب تولید شدند طوری که نسبت های حداکثر روشناهی به دامنه نویز در آنها بترتیب ۱، ۱۰ و ۱۰۰ است. در شکل ۵.۵ (ب) طیف فوریه تصاویر تخریب شده دیده می شوند.

از آنجا که اثرات حرکت خطی یکنواخت را می توان به طور تحلیلی بیان کرد، همان طور که در بخش ۵.۴ نشان داده شد، به آسانی می توان معادله ای را بدست آورد که $H(u, v) = R$ را توصیف نماید. شکل ۵.۵ (ب) بوسیله فیلتر کردن معکوس مستقیم^۲ طبق روال مشروح در بخش ۵.۴.۱ بدست آمد. نتایج مغلوب نویز هستند، اما همان طور که تصویر سوم نشان می دهد، فیلتر معکوس به خوبی تخریب (ماتی) حاصل از حرکت را حذف کرد. بر عکس، شکل ۵.۵ (ت) نتایج حاصل از کاربرد معادله $(5.5-10) \text{ با}^{20} = K$ را نشان می دهد که^۳ واریانس نویز است. بهبود نسبت به فیلتر کردن معکوس مستقیم بویژه برای تصویر سوم واضح است. شکل ۵.۵ (ث) طیف فوریه تصاویر بازیابی شده را نشان



شکل ۵. ۵ مثال بازیابی تصویر با فیلترهای معکوس و وینر: (الف) تصاویر تخریب شده و (ب) طیف فوریه آنها؛ (ب) تصاویر بازیابی شده با فیلتر معکوس؛ (ت) تصاویر بازیابی شده با فیلتر وینر؛ (ث) طیف فوریه تصاویر ستون (ت). (از / ۱۹۶۸/ Harris).

می دهد.

۵. ۶ بازیابی حداقل مربعات مقید

روش حداقل میانگین مربعات بدست آمده در بخش ۵. ۵، یک روش آماری است زیرا معیار بهینگی آن بر اساس ماتریس های همبستگی توابع تصویر و نویز می باشد. این بدان معناست که نتایج بدست آمده با استفاده از فیلتر وینر در مفهوم متوسط، بهینه هستند. در هر حال روش بازیابی ارائه شده در این بخش برای هر تصویر دلخواه بهینه است و تنها به دانستن میانگین و واریانس نویز نیاز دارد. مورد دیگری که بررسی می شود مسأله تنظیم γ است طوری که قید منجر به معادله ۵. ۳-۹) برآورده شود.

همان طور که در بخش ۵. ۳. ۲ ذکر شد، نتیجه بازیابی حاصل از کاربرد معادله ۵. ۳-۹) به انتخاب ماتریس **Q** بستگی دارد. در شرایط بد، گاهی معادله مذکور جواب هایی می دهد که بوسیله 'نوسانات بزرگی' مبهم شده اند. بنابراین تحقیق پذیری انتخاب **Q** طوری که این اثرات مضر حداقل شوند، مورد علاقه است. یک راه حل که بوسیله [Phillips 1962] پیشنهاد شد، بیان یک معیار بهینگی براساس

۵.۵ بازیابی حداقل مربعات مقید ۳۴۳

میزان آرامی^۱ تصویر است طوری که، به عنوان مثال، تابعی از مشتق دوم را حداقل سازد. برای این که بینیم چگونه می‌توان این معیار را به شکلی سازگار با معادله (۵.۶-۳) بیان کرد، اجازه دهید که در ابتدا حالت یک بعدی را در نظر بگیریم.

برای تابع گسته $f(x)$ ، $x = 0, 1, 2, \dots$ مشتق دوم در نقطه x را می‌توان با عبارت

$$\frac{\partial^2 f(x)}{\partial x^2} = f(x+1) - 2f(x) + f(x-1). \quad (5.6.5)$$

تقریب زد. آنگاه یک معیار مبتنی بر این عبارت، ممکن است کمینه سازی^۲ ($\partial^2 f / \partial x^2$) نسبت به x باشد؛ یعنی،

$$\left\{ \sum_i [f(x+1) - 2f(x) + f(x-1)]^2 \right\} \quad (5.6.6)$$

یا به بیان ماتریسی،

$$\{ \mathbf{f}' \mathbf{C}' \mathbf{C} \mathbf{f} \} \quad \text{کمینه سازی} \quad (5.6.7)$$

$$\mathbf{C} = \begin{bmatrix} 1 & & & & & \\ -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & -2 & 1 \\ & & & & 1 & -2 \\ & & & & & 1 \end{bmatrix} \quad (5.6.8)$$

یک ماتریس "آرام کننده"^۲ است، و برداری است که عناصرش نمونه های $f(x)$ هستند. در حالت دو بعدی گسترش مستقیم معادله (۵.۶-۱) را در نظر می گیریم. در این حالت معیار

$$\left[\frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \right]^2 \quad \text{کمینه سازی} \quad (5.6.9)$$

است که تابع مشتقی با عبارت

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} &= [2f(x, y) - f(x+1, y) - f(x-1, y)] \\ &\quad + [2f(x, y) - f(x, y+1) - f(x, y-1)] \\ &= 4f(x, y) - [f(x+1, y) + f(x-1, y) + f(x, y+1) \\ &\quad + f(x, y-1)]. \end{aligned} \quad (6-6. 5)$$

تقریب زده می شود. تابع مشتقی داده شده در معادله (5. 5-۶) همان عملگر لاپلاسین بیان شده در بخش ۷. ۳ است.

معادله (5. 6-۶) را می توان مستقیماً در یک رایانه پیاده سازی کرد. در هر حال همان عمل را می توان با کانولوکردن $f(x, y)$ و عملگر

$$p(x, y) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (7-6. 5)$$

انجام داد. همان طور که در بخش ۱. ۳ بیان شد، با گسترش $f(x, y)$ و $p(x, y)$ از خطای پیچشی^۱ کانولوشن گستته دوری می شود. پس از تشکیل $f_e(x, y)$ ، $f_e(x, y) p_e(x, y)$ را نیز به همان روش تشکیل می دهیم:

$$p_e(x, y) = \begin{cases} p(x, y) & 0 \leq x \leq 2 \quad \text{و} \quad 0 \leq y \leq 2 \\ 0 & 3 \leq x \leq M-1 \quad \text{یا} \quad 3 \leq y \leq N-1. \end{cases}$$

اگر $f(x, y)$ با ابعاد $A \times B$ باشد، چون $p(x, y)$ با ابعاد 3×3 است، M و N را با شرایط $N \geq B+3-1$ و $M \geq A+3-1$ انتخاب می کنیم.

آنگاه کانولوشن نوعی گسترش یافته به صورت

$$g_e(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_e(m, n) p_e(x-m, y-n) \quad (8-6. 5)$$

است که با معادله (5. 23-۱) هماهنگ است.

با استدلالی مشابه آنچه در بخش ۱. ۵ آمد، بیان معیار آرامی^۲ به شکل ماتریسی ممکن است. در

1) wraparound error

2) Smoothness criterion

۳۴۵ ۶.۵ بازیابی حداقل مربعات مقید

ابتدا یک ماتریس بلوک-چرخشی به شکل

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_0 & \mathbf{C}_{M-1} & \mathbf{C}_{M-2} & \cdots & \mathbf{C}_1 \\ \mathbf{C}_1 & \mathbf{C}_0 & \mathbf{C}_{M-1} & \cdots & \mathbf{C}_2 \\ \mathbf{C}_2 & \mathbf{C}_1 & \mathbf{C}_0 & \cdots & \mathbf{C}_3 \\ \vdots & & & & \\ \mathbf{C}_{M-1} & \mathbf{C}_{M-2} & \mathbf{C}_{M-3} & \cdots & \mathbf{C}_0 \end{bmatrix} \quad (9-6.5)$$

تشکیل می دهیم که هر زیرماتریس \mathbf{C}_j نیز یک ماتریس چرخشی $N \times N$ است که از عناصر سطر j آمده است؛ یعنی،
 (y) تشکیل شده است؛ یعنی،

$$\mathbf{C}_j = \begin{bmatrix} p_e(j, 0) & p_e(j, N-1) & \cdots & p_e(j, 1) \\ p_e(j, 1) & p_e(j, 0) & \cdots & p_e(j, 2) \\ \vdots & & & \\ p_e(j, N-1) & p_e(j, N-2) & \cdots & p_e(j, 0) \end{bmatrix} \quad (10-6.5)$$

از آنجا که \mathbf{C} بلوک-چرخشی است، با استفاده از ماتریس \mathbf{W} تعریف شده در بخش ۲.۲، قطری می شود. به عبارت دیگر،

$$\mathbf{E} = \mathbf{W}^{-1} \mathbf{C} \mathbf{W} \quad (11-6.5)$$

که \mathbf{E} یک ماتریس قطری است که عناصرش با رابطه ای همانند معادله (۳۹-۲.۵) به صورت زیر تعیین می شوند:

$$E(k, i) = \begin{cases} P\left(\left[\frac{k}{N}\right], k \bmod N\right) & i = k \quad \text{اگر} \\ 0 & i \neq k \quad \text{اگر} \end{cases} \quad (12-6.5)$$

در این مورد $P(u, v)$ تبدیل فوریه دو بعدی $p_e(x, y)$ است. مانند معادلات (۵.۳۷-۲) و (۵.۳۹-۲) با ضریب MN تغییر مقیاس داده شده است.

عمل کانولوشن مشروح در بالا معادل پیاده سازی معادله (۵.۶-۶) است، بنابراین معیار آرامی معادله

(۵-۶) به شکل معادله (۵-۳) خواهد بود:

$$\{\mathbf{f}^T \mathbf{C}^T \mathbf{C} \mathbf{f}\} \quad \text{کمینه سازی} \quad (13-6. 5)$$

که \mathbf{f} یک بردار MN بعدی است و ماتریس \mathbf{C} به ابعاد $MN \times MN$ می باشد. با قراردادن $\mathbf{C} = \mathbf{Q}$ ، و
یادآوری این که $\|\mathbf{Qf}\|^2 = (\mathbf{Qf})^T (\mathbf{Qf}) = \mathbf{f}^T \mathbf{Q}^T \mathbf{Qf}$ می توان این معیار را به صورت

$$\|\mathbf{Qf}\|^2 \quad \text{کمینه سازی} \quad (14-6. 5)$$

بیان کرد که همان شکل مورد استفاده در بخش ۲. ۳. ۵ است. در واقع، اگر لازم باشد که قید
برآورده شود، جواب بهینه با معادله (۵-۹) و بازای $\mathbf{C} = \mathbf{Q}$ بدست می آید:

$$\hat{\mathbf{f}} = (\mathbf{H}^T \mathbf{H} + \gamma \mathbf{C}^T \mathbf{C})^{-1} \mathbf{H}^T \mathbf{g}. \quad (15-6. 5)$$

استفاده از معادلات (۵-۲۱)، (۵-۲۲)، (۵-۲۳) و (۵-۱۱)، امکان می دهد که معادله (۵-۶) به صورت

$$\hat{\mathbf{f}} = (\mathbf{W}^* \mathbf{D} \mathbf{W}^{-1} + \gamma \mathbf{E}^* \mathbf{E} \mathbf{W}^{-1})^{-1} \mathbf{W}^* \mathbf{D}^* \mathbf{W}^{-1} \mathbf{g}. \quad (16-6. 5)$$

بیان شود. با ضرب دوطرف در \mathbf{W}^{-1} و انجام تعدادی عملیات ماتریسی معادله (۵-۶) به شکل

$$\mathbf{W}^{-1} \hat{\mathbf{f}} = (\mathbf{D}^* \mathbf{D} + \gamma \mathbf{E}^* \mathbf{E})^{-1} \mathbf{D}^* \mathbf{W}^{-1} \mathbf{g}. \quad (17-6. 5)$$

ساده می شود. با به خاطرداشتن این که ماتریسهای درون پرانتز قطری هستند و با استفاده از مفاهیم بیان
شده در بخش ۲. ۳. ۵، می توان عناصر معادله (۵-۶) را به شکل

$$\hat{F}(u, v) = \left[\frac{H^*(u, v)}{|H(u, v)|^2 + \gamma |P(u, v)|^2} \right] G(u, v) \quad (18-6. 5)$$

برای $1 \leq u, v \leq N$ بیان کرد که $|H(u, v)|^2 = H^*(u, v)H(u, v)$ و فرض
کرده ایم که $N = M$. توجه کنید که معادله (۱۸-۶) شبیه فیلتر وینر پارامتری بدست آمده در بخش ۵. ۵
است. تفاوت اصلی بین معادلات (۵-۹) و (۵-۱۸) این است که معادله دوم به اطلاعات صریحی

۶.۵ بازیابی حداقل مربعات مقید ۳۴۷

در مورد پارامترهای آماری غیر از تخمین میانگین و واریانس نویز، نیاز ندارد.

در رابطه کلی داده شده در معادله (۵.۹-۳) باید γ جهت برآوردن قید $\|\mathbf{n}\|^2 = \|\mathbf{H}\hat{\mathbf{f}}\|^2$ تنظیم شود. بنابراین جواب داده شده در معادله (۵.۱۸-۶) تنها وقتی می‌تواند بهینه باشد که γ این شرط را برآورده نماید. در ادامه یک فرآیند گام به گام^۱ برای تخمین این عامل ارائه می‌گردد. بردار مانده^۲ را به صورت

$$\mathbf{r} = \mathbf{g} - \mathbf{H}\hat{\mathbf{f}} \quad (19-6.5)$$

تعريف نمایید. با جایگذاری معادله (۵.۱۵-۶) برای $\hat{\mathbf{f}}$ در این معادله داریم:

$$\mathbf{r} = \mathbf{g} - \mathbf{H}(\mathbf{H}^T\mathbf{H} + \gamma\mathbf{C}^T\mathbf{C})^{-1}\mathbf{H}^T\mathbf{g} \quad (20-6.5)$$

که این معادله بیان می‌کند که \mathbf{r} تابعی از γ است. در واقع، می‌توان نشان داد (Hunt[1973]) که تابع

$$\phi(\gamma) = \frac{\mathbf{r}^T\mathbf{r}}{\|\mathbf{r}\|^2} \quad (21-6.5)$$

یک تابع یکنواخت صعودی از γ است. هدف تنظیم γ است طوری که

$$\|\mathbf{r}\|^2 = \|\mathbf{n}\|^2 \pm a \quad (22-6.5)$$

که a ثابت دقت می‌باشد. بروشی اگر $\|\mathbf{r}\|^2 = \|\mathbf{n}\|^2$ ، از دید معادله (۵.۱۹-۶) قید $\|\mathbf{n}\|^2 = \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2$ کاملاً برآورده می‌شود.

چون $\phi(\gamma)$ یکنواست، پیدا کردن γ که معادله (۵.۱۷-۶) را برآورده نماید، مشکل نیست. یک روش ساده بدین صورت است که:

(۱) یک مقدار اولیه برای γ تعیین نماید؛

(۲) $\phi(\gamma)$ را محاسبه نماید؛ و

(۳) اگر معادله (۵.۲۲-۶) برآورده شود، توقف نماید؛ در غیر این صورت اگر $\|\mathbf{r}\|^2 < \|\mathbf{n}\|^2$ ،

عرا افزایش و اگر $\|\mathbf{r}\|^2 > \|\mathbf{n}\|^2$ ، γ را کاهش دهید و به مرحله ۲ بازگردید.

می توان از سایر روش ها نظریه الگوریتم نیوتون-رافسون، برای افزایش سرعت همگرایی استفاده کرد.
پیاده سازی این مفاهیم نیازمند مقداری اطلاعات در مورد $\|\mathbf{n}\|^2$ است. واریانس (y, η_r) با رابطه

$$\begin{aligned}\sigma_{\eta_r}^2 &= E\{\eta_r(x, y) - \bar{\eta}_r\}^2 \\ &= E[\eta_r^2(x, y)] - \bar{\eta}_r^2.\end{aligned}\quad (23-6.5)$$

بدست می آید که

$$\bar{\eta}_r = \frac{1}{(M-1)(N-1)} \sum_i \sum_y \eta_r(x, y) \quad (24-6.5)$$

مقدار میانگین (x, y, η_r) می باشد. اگر یک متوسط نمونه ای برای تقریب زدن مقدار امید (y, η_r) بکار رود، معادله (23-6.5) به صورت

$$\sigma_{\eta_r}^2 = \frac{1}{(M-1)(N-1)} \sum_i \sum_y \eta_r^2(x, y) - \bar{\eta}_r^2. \quad (25-6.5)$$

می شود. جزو مجموع بیانگر جمع مجددورات عناصر آرایه (y, η_r) بازای $y = 0, 1, 2, \dots, N-1$ و $x = 0, 1, 2, \dots, M-1$ همان حاصل ضرب $\mathbf{n}^T \mathbf{n}$ است که طبق تعریف برابر $\|\mathbf{n}\|^2$ است. پس معادله (25-6.5) به شکل

$$\sigma_{\eta_r}^2 = \frac{\|\mathbf{n}\|^2}{(M-1)(N-1)} - \bar{\eta}_r^2 \quad (26-6.5)$$

یا

$$\|\mathbf{n}\|^2 = (M-1)(N-1)[\sigma_{\eta_r}^2 + \bar{\eta}_r^2]. \quad (27-6.5)$$

ساده می شود. این معادله از این جهت مهم است که تعیین یک مقدار را برای قید بر حسب کمیات میانگین و واریانس نویز ممکن می سازد؛ کمیاتی که اگر معلوم نباشند، در اغلب موارد عملی می توان آنها را تقریب زد یا اندازه گرفت.

روال بازیابی حداقل مربعات مقید را می توان به صورت زیر خلاصه کرد.
مرحله ۱. یک مقدار اولیه برای $\boldsymbol{\gamma}$ انتخاب کنید و با استفاده از معادله (27-6.5) تخمینی از $\|\mathbf{n}\|^2$ بدست آورید.

۷.۵ بازیابی تعاملی

۳۴۹

مرحله ۲. $\hat{F}(u, v)$ را با استفاده از معادله (۵.۱۸-۶) محاسبه نمایید. با گرفتن تبدیل فوریه معکوس از (u, v) ، $\hat{f}(u, v)$ را بدست آورید.

مرحله ۳. طبق معادله (۵.۱۹-۶) بردار مانده \mathbf{z} را تشکیل دهید و $\|\mathbf{z}\|^2 = (\gamma)\phi$ را محاسبه نمایید.
مرحله ۴. \mathbf{z} را افزایش یا کاهش دهید.

(الف) اگر $a - \|\mathbf{n}\|^2 < (\gamma)\phi$ باشد، \mathbf{z} را طبق الگوریتم فوق یا روش مناسب دیگری (نظیر فرآیند تیوتن-رافسون) افزایش دهید.

(ب) اگر $a + \|\mathbf{n}\|^2 > (\gamma)\phi$ ، \mathbf{z} را طبق الگوریتم مناسبی کاهش دهید.

مرحله ۵. به مرحله ۲ بازگردید و ادامه دهید مگر اینکه مرحله ۶ درست باشد.

مرحله ۶. اگر $a \pm \|\mathbf{n}\|^2 = (\gamma)\phi$ که a دقت برآورده شدن قید است، فرآیند تخمین را متوقف نمایید. \mathbf{f} مربوط به مقدار فعلی \mathbf{z} تصویر بازیابی شده می باشد.

مثال: شکل ۵.۶ (ب) با کاتوالوکردن تابع پخش نقطه‌ای گوسی با رابطه^۱

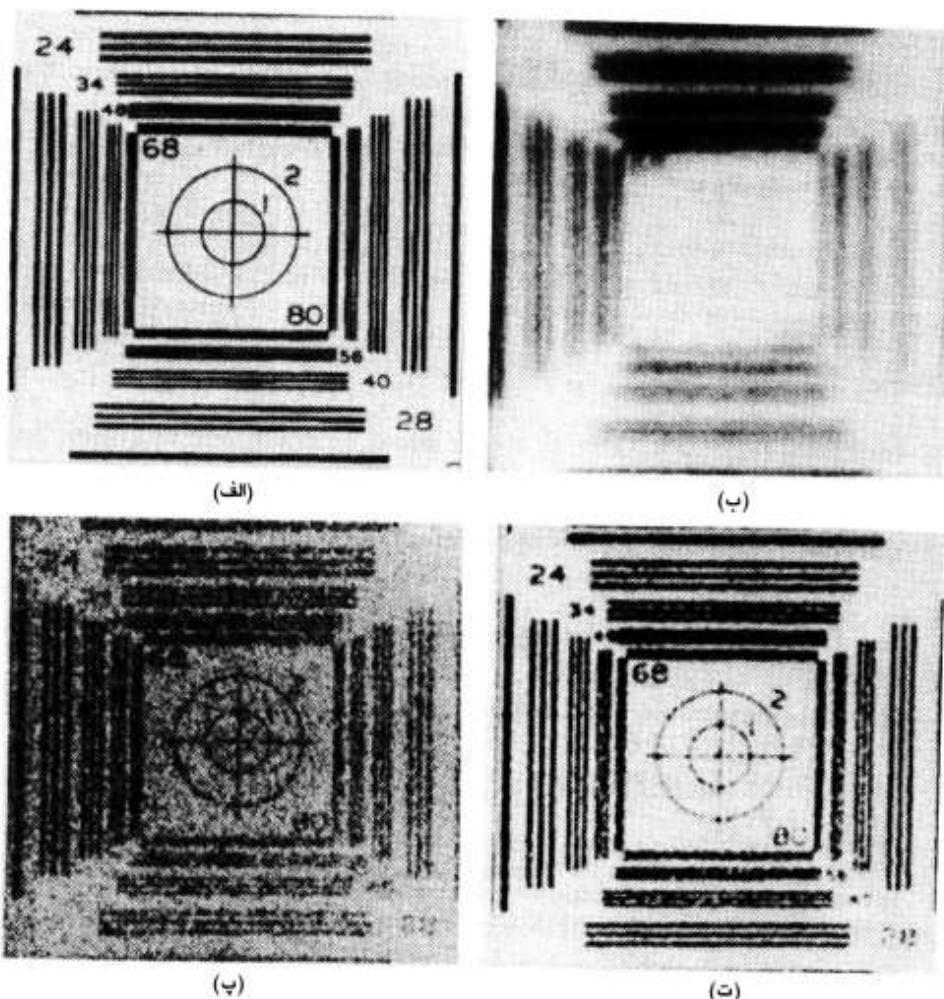
$$h(x, y) = \exp\left(-\frac{x^2 + y^2}{2400}\right)$$

و تصویر اولیه شکل ۵.۶ (الف) و اضافه کردن نویزی با توزیع یکنواخت در بازه $[0, 0.5]$ بدست آمد.

شکل ۵.۶ (ب) حاصل استفاده از الگوریتم بازیابی $\mathbf{z} = \gamma$ (فیلتر معکوس) را نشان می دهد. طبیعت نامطلوب جواب، بوسیله غلبه نویز بر تصویر بازیابی شده آشکار است. شکل ۵.۶ (ت) با استفاده از الگوریتم قبلی برای جستجوی یک \mathbf{z} که قید را برآورده نماید، به دست آمد. واریانس و میانگین چگالی یکنواخت در بازه $[0, 0.5]$ برای تخمین $\|\mathbf{n}\|^2$ استفاده شدند و عامل دقت « طوری انتخاب شد که $a = 0.025\|\mathbf{n}\|^2$ بیهود در جواب مقید نسبت به جواب فیلتر کردن معکوس مستقیم آشکار است. □

۷.۵ بازیابی تعاملی

ناکنون روی یک روش بازیابی کاملاً تحلیلی تمرکز کرده ایم. در بسیاری کاربردها، روش عملی، بازیابی تعاملی تصاویر با بهره گیری از حس ششم انسان و قابلیت های گوناگون رایانه رقمی است. در این مورد ناظر، فرآیند بازیابی را کنترل می نماید، و با تنظیم عامل های موجود، نتیجه نهایی بدست می آورد که احتمالاً برای یک کاربرد خاص کافی است.



شکل ۵.۶ (الف) تصویر اولیه؛ (ب) تصویر مات و خراب شده با نویز جمع شونده؛ (پ) تصویر بازیابی شده با فیلتر کردن معکوس؛ (ت) تصویر بازیابی شده با روش حداقل مربعات مقید. (از [Hunt 1973].)

یکی از ساده‌ترین موارد خرابی تصویر^۱ که مناسب بازیابی تعاملی می‌باشد، وقوع الگوی تداخل سینوسی دو بعدی (که غالباً نویز همدوس^۲ خوانده می‌شود) روی تصویر است. فرض کنید (x, y) پیانگر یک الگوی تداخل سینوسی با دامنه A و مؤلفه‌های فرکانسی (u_0, v_0) باشد؛ یعنی،

$$\eta(x, y) = A \sin(u_0 x + v_0 y). \quad (1-7.5)$$

1) *Image corruption*

2) *Coherent noise*

۷.۵ بازیابی تعاملی ۳۵۱

با جایگذاری مستقیم معادله (۵.۱-۹) در معادله (۳.۱-۱)، تبدیل فوریه (y, x) به صورت زیر بدست می آید:

$$N(u, v) = \frac{-jA}{2} \left[\delta\left(u - \frac{u_0}{2\pi}, v - \frac{v_0}{2\pi}\right) - \delta\left(u + \frac{u_0}{2\pi}, v + \frac{v_0}{2\pi}\right) \right]. \quad (2-7. 5)$$

به عبارت دیگر، تبدیل فوریه یکتابع سینوسی دو بعدی یک زوج ضربه با قدرت $A/2$ و $A/2$ است که به ترتیب در مختصات $(u_0/2\pi, v_0/2\pi)$ و $(-u_0/2\pi, -v_0/2\pi)$ از صفحه فرکانسی قرار گرفته اند. در این حالت تبدیل، تنها مؤلفه های موهومی را دارد.

اگر تنها عامل خرابی نویز جمع شونده باشد، از معادله (۵.۲-۴۰) نتیجه می شود که

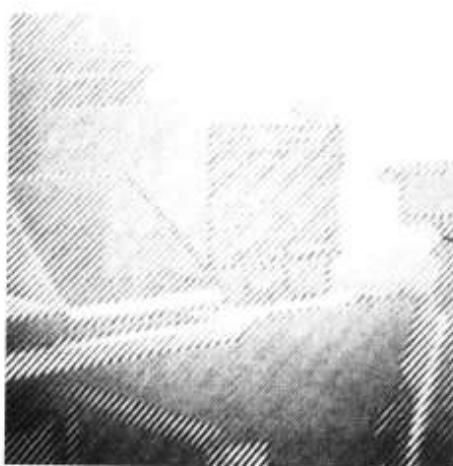
$$G(u, v) = F(u, v) + N(u, v). \quad (3-7. 5)$$

اندازه $G(u, v)$ مجموع اندازه های $F(u, v)$ و $N(u, v)$ است. اگر A به اندازه کافی بزرگ باشد، معمولاً دو ضربه موجود در $N(u, v)$ به صورت نقاط روشنی روی صفحه ظاهر می شوند، بویژه اگر دو ضربه نسبتاً دور از مبدأ باشند طوری که سهم دخالت مؤلفه های $F(u, v)$ کوچک باشد. اگر $y(x)$ کاملاً معلوم باشد، می توان با تفیری تداخل از $y(x)$ تصویر اولیه را بازیابی نمود. چون این حالت بnderت وجود دارد، یک روش مفید، شناسایی چشمی محل مؤلفه های ضربه در میدان فرکانس و استفاده از فیلتر حذف باند^۱ (بخش ۴.۶.۳ را بینید) در این محلها است.

مثال: تصویر شکل ۵.۷ (الف) بوسیله یک الگوی سینوسی به شکل معادله (۱-۷.۵)، خراب شده است. در طیف فوریه این تصویر، که در شکل ۵.۷ (ب) نمایش داده شده است، به وضوح یک زوج ضربه متقارن حاصل از تداخل سینوسی دیده می شود. شکل ۵.۷ (پ) با قراردادن دو فیلتر حذف باند به شعاع واحد در محل ضربه ها (از طریق یک کنسول رایانه ای) و سپس تبدیل فوریه معکوس گرفتن از حاصل بدست آمد. تصویر بازیابی شده با این روش عاری از تداخل است. □

در عمل بnderت رخ می دهد که تنها یک الگوی تداخل کاملاً تعریف شده، مانند آنچه گفته شد، در تصویر موجود باشد. تصاویر به دست آمده با پویشگرهای الکترو-نوری نظیر آهایی که عموماً در مأموریت های فضایی استفاده می شوند، مثال های قابل ذکری هستند. یک مشکل مشداول در مورد این

1) Bandreject filter



(الف)



(ب)



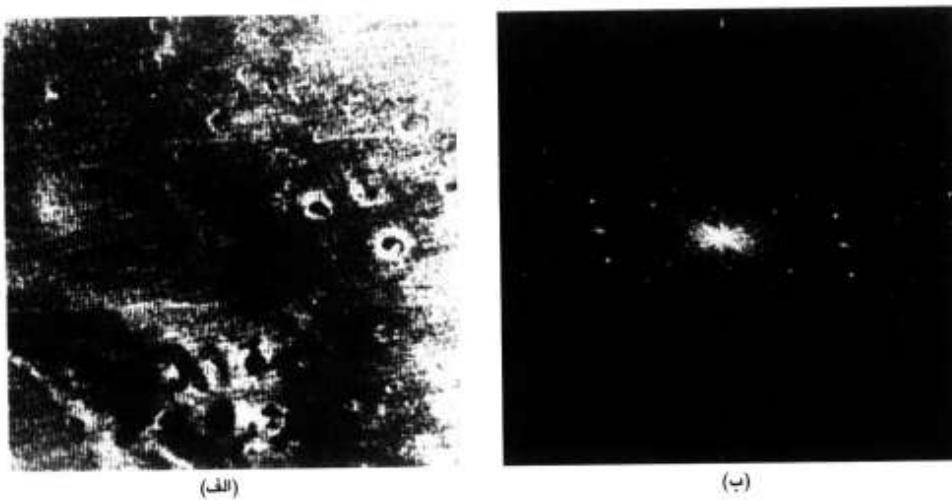
(پ)

شکل ۵. ۷ مثالی از حذف تداخل سینوسی؛ (الف) تصویر خراب شده؛ (ب) طیف فوریه که خصیه های مربوط به الگوری سینوسی را نشان می دهد؛ (پ) تصویر بازیابی شده با استفاده از فیلتر حذف باند با شاعع واحد

حسگرهای داخل ناشی از تزویج^۱ و تقویت سیگنالهای ضعیف در مدارات الکترونیکی است. در نتیجه تصاویر بازسازی شده از خروجی پویشگر حاوی یک ساختار متناوب دو بعدی تقویت شده که روی صحنۀ افتاده است، می باشدند.

شکل ۵. ۸ (الف)، که مثالی از وجود این نوع خرابی متناوب در تصویر است، یک تصویر رقمه

1) Coupling



شکل ۵.۸ (الف) عکسی از سطح مریخ که به وسیله Mariner 6 گرفته شده است؛ (ب) طیف فوریه. آن به تداخل متناوب موجود در تصویر و بر جستگی های متناظر در طیف توجه کنید. (با مجوز از NASA Jet Propulsion Laboratory)

برداشته شده از سطح کره' مریخ بوسیله فضایمای 6 Mariner می باشد. الگوی تداخل کاملاً شبیه شکل ۵.۷ (الف) است، اما الگوی تداخل شکل ۵.۸(الف) به طور قابل ملاحظه ای ریز و نامحسوس تر^۱ است. در نتیجه تشخیص آن در صفحه' فرکانس سخت تر می باشد.

شکل ۵.۸ (ب) طیف فوریه تصویر مورد نظر را نشان می دهد. مؤلفه های ستاره مانند^۲ ناشی از تداخل هستند و وجود چند زوج مؤلفه نشان می دهد که الگوی ترکیبی از چند مؤلفه سینوسی است نه فقط یک مؤلفه سینوسی . وقتی چند مؤلفه تداخل موجود باشد، روش فوق همواره مقبول تیست، زیرا ممکن است در فرآیند فیلتر کردن، اطلاعات تصویری زیادی حذف گردد. بعلاوه این مؤلفه ها عموماً دسته های تک فرکانس نیستند، بلکه دارای طیف نسبتاً پهنه هستند که حاوی اطلاعاتی درباره' الگوی تداخل می باشد. گاهی اوقات تشخیص این مؤلفه ها از زمینه تبدیل به سختی انجام می شود.

یک روش که در پردازش صحنه های مربوط به فضا مورد اقبال است، شامل تفکیک قسمت های اصلی الگوی تداخل و سپس تفریق یک بخش متغیر و وزن دار الگو از تصویر خراب شده می باشد. گرچه ما این روال را در مورد یک کاربرد خاص اجرا می کنیم، روش پایه ای کاملاً کلی است و می توان آن را

1) Subtle

1) Starlike components

برای سایر کارهای بازیابی مانند وقتی که چند تداخل متناوب وجود دارد، به کاربرد.

مرحله اول استخراج مؤلفه های فرکانسی اصلی از الگوی تداخل است. این استخراج را می توان با قرار دادن یک فیلتر میان گذر $H(u, v)$ در محل هر ضربه انجام داد (بخش ۴.۶.۳. را ببینید). اگر طوری باشد که تنها مؤلفه های مربوط به الگوی تداخل را عبور دهد، نتیجه می شود که تبدیل $H(u, v)$ فوریه الگو با رابطه

$$P(u, v) = H(u, v)G(u, v) \quad (4-7.5)$$

بدست می آید که $G(u, v)$ تبدیل فوریه تصویر خراب شده $g(x, y)$ است و اگر رقمه سازی $N \times N$ باشد، " w مقادیری در بازه $1 - N, 0, 1, 2, \dots, N$ می گیرند.

برای تشکیل $H(u, v)$ باید دقیقاً تعیین شود که چه چیزی ضربه تداخل است و چه چیزی نیست. به این دلیل، فیلتر میان گذر عموماً به طور تعاملی و همزمان با مشاهده طیف $G(u, v)$ روی نمایشگر تشکیل می شود. پس از این که $H(u, v)$ انتخاب شد، الگوی متناظر در حوزه مکان از عبارت زیر بدست می آید:

$$p(x, y) = \mathcal{F}^{-1}\{H(u, v)G(u, v)\}. \quad (5-7.5)$$

چون تصویر خراب شده با جمع $f(x, y)$ و تداخل تشکیل می شود، اگر $p(x, y)$ کاملاً معلوم باشد، تفریق الگو از $p(x, y)$ برای بدست آوردن $f(x, y)$ کار ساده ای خواهد بود. البته مسئله این است که این روش فیلتر کردن معمولاً تنها تقریبی از الگوی واقعی را می دهد. اثرات مؤلفه های غیر موجود در تخمین $p(x, y)$ را می توان با تفریق یک بخش وزن دار $w(x, y)$ از $p(x, y)$ حداقل کرد تا این که تخمینی از $f(x, y)$ بدست آید:

$$\hat{f}(x, y) = g(x, y) - w(x, y)p(x, y) \quad (6-7.5)$$

در رابطه فوق باید $w(x, y)$ معین شود.تابع $w(x, y)$ تابع وزن دهنده یا مدولاسیون^۱ نامیده می شود و هدف این روال انتخاب این تابع است طوری که حاصل در معیار معناداری بهینه شود. یک رهیافت این است که $w(x, y)$ را طوری انتخاب کنیم که واریانس $(\hat{f}(x, y) - f(x, y))^2$ روی یک همسایگی مشخص از هر نقطه

1) Weighting or modulation function

حداقل شود.

یک همسایگی با ابعاد $(1 + 2Y + 2X)$ در نقطه (x, y) در نظر بگیرید. واریانس محلی $\hat{f}(x, y)$ در مختصات (x, y) عبارتست از

$$\sigma^2(x, y) = \frac{1}{(2X + 1)(2Y + 1)} \sum_{m=-X}^X \sum_{n=-Y}^Y [\hat{f}(x + m, y + n) - \bar{\hat{f}}(x, y)]^2 \quad (7-7.5)$$

که $\bar{\hat{f}}(x, y)$ مقدار متوسط $\hat{f}(x, y)$ در همسایگی است؛ یعنی،

$$\bar{\hat{f}}(x, y) = \frac{1}{(2X + 1)(2Y + 1)} \sum_{m=-X}^X \sum_{n=-Y}^Y \hat{f}(x + m, y + n). \quad (8-7.5)$$

با استفاده از همسایگی های جزئی می توان نقاط لبه یا نزدیک لبه تصویر را نیز پردازش کرد. با جایگذاری معادله (5.7-7) در معادله (7-7.5) داریم که

$$\begin{aligned} \sigma^2(x, y) &= \frac{1}{(2X + 1)(2Y + 1)} \sum_{m=-X}^X \sum_{n=-Y}^Y \{[g(x + m, y + n) \\ &\quad - w(x + m, y + n)p(x + m, y + n)] - [\bar{g}(x, y) - \overline{w(x, y)p(x, y)}]\}^2. \end{aligned} \quad (9-7.5)$$

با فرض این که $w(x, y)$ روی همسایگی تقریباً ثابت می ماند، تقریب

$$w(x + m, y + n) = w(x, y) \quad (10-7.5)$$

برای $X \leq m \leq X + Y$ و $-Y \leq n \leq Y$ بدست می آید؛ همچنین با تقریب در این همسایگی داریم که

$$\overline{w(x, y)p(x, y)} = w(x, y)\bar{p}(x, y) \quad (11-7.5)$$

با این تقریب ها، معادله (9-7.5) به صورت

$$\begin{aligned} \sigma^2(x, y) &= \frac{1}{(2X + 1)(2Y + 1)} \sum_{m=-X}^X \sum_{n=-Y}^Y \{[g(x + m, y + n) \\ &\quad - w(x, y)p(x + m, y + n)] - [\bar{g}(x, y) - w(x, y)\bar{p}(x, y)]\}^2. \end{aligned} \quad (12-7.5)$$

در می آید. برای کمینه سازی $\sigma^2(x, y)$ ، معادله'

$$\frac{\partial \sigma^2(x, y)}{\partial w(x, y)} = 0 \quad (13-7.5)$$

را برای $w(x, y)$ حل می کنیم. حاصل عبارتست از:

$$w(x, y) = \frac{g(x, y)p(x, y) - \bar{g}(x, y)\bar{p}(x, y)}{p^2(x, y) - \bar{p}^2(x, y)}. \quad (14-7. 5)$$

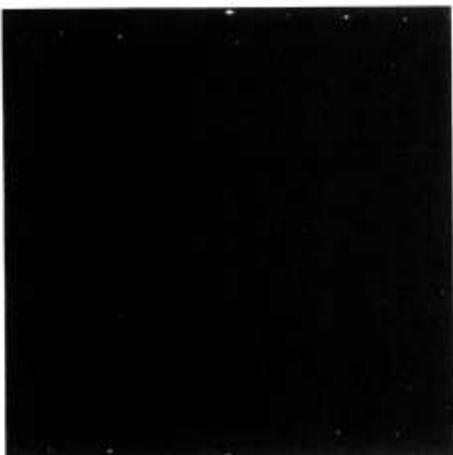
برای به دست آوردن تصویر بازیابی شده $\hat{f}(x, y)$ ، $w(x, y)$ را از معادله (14-7. 5) محاسبه کرده و سپس از معادله (5. 6) استفاده می کنیم. از آنجا که فرض می شود $w(x, y)$ در هر همسایگی ثابت است، محاسبه این تابع برای هر مقدار x و y در تصویر الزامی نیست. در عوض، $w(x, y)$ برای یک نقطه در هر همسایگی ناهمبواشانده¹ (ترجیحاً نقطه مرکزی) محاسبه می شود و سپس برای پردازش تمام نقاط تصویر درون آن همسایگی استفاده می شود.

مثال: اشکال ۹. ۱۱ تا ۹. ۱۵ نتیجه کاربرد روش فوق بر تصویر شکل ۸. ۵ (الف) رانشان می دهدند. در این مورد $N = 512$ و یک همسایگی با $Y = 15$ که $X = 5$ انتخاب شد. شکل ۹. ۵ طیف فوریه تصویر خراب شده رانشان می دهد، اما مبدأ به مرکز صفحه، فرکانس انتقال نیافته است. شکل ۹. ۱۰ (الف) طیف

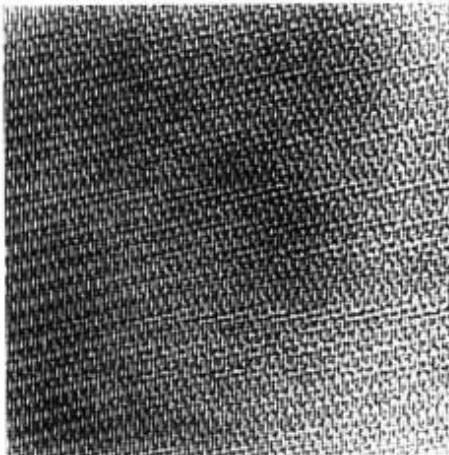


شکل ۹. ۵ طیف فوریه تصویر شکل ۸. (الف) (بدون انتقال فرکانسی). (با مجوز از Jet Propulsion Laboratory, NASA)

1) Nonoverlapping



(الف)



(ب)

شکل ۵.۱۰. (الف) طیف فوریه $p(x, y)$; (ب) الگوی تداخل متناظر با آن. (با مجوز از NASA, Jet Propulsion Laboratory).



شکل ۵.۱۱. تصویر پردازش شده (با مجوز از NASA, Jet Propulsion Laboratory).

p را نشان می دهد، که تنها ضربه های نویز موجود هستند. شکل ۵.۱۰ (ب) الگوی تداخل (u, v) حاصل از گرفتن تبدیل فوریه معکوس از $v(u, p)$ را نشان می دهد. به شباهت بین این الگو و ساختار نویز موجود در شکل ۵.۸ (الف) توجه کنید. در نهایت شکل ۵.۱۱ تصویر پردازش شده حاصل از کاربرد معادله (۵.۶-۷) را نشان می دهد. تداخل متناوب حذف می شود و تنها نویز نقطه ای که متناوب نیست، باقی میماند. این نویز را میتوان با سایر روش های نظری فیلتر کردن میانه ای^۱ پردازش کرد.

۸.۵ بازیابی در حوزه مکان

پس از این که یک فیلتر مناسب بازیابی در حوزه 'فرکانس با استفاده از روش های قبلی بدست آمد، غالباً به منظور تسريع پردازش (بخش ۴.۱ را بینید) مطلوب است که جواب را از طریق نقاب کانولوشن در حوزه 'مکان پیدا های سازی کنیم. همان طور که در بخش ۴.۵ بیان شد، میتوان مستقیماً از تابع فیلتر داده شده و از طریق معادله (۴.۱۲-۵.۱۲) ضرایب این نقاب کانولوشن را بدست آورد. گرچه بخش ۴.۵ درباره 'ارتقا است، مفاهیمی که در آنجا بیان شدند، همان کاربردها را در بازیابی دارند؛ تفاوت صرفاً در طبیعت فیلتر می باشد.

مثال: شکل ۵.۱۲ (الف) تصویر مادون قرمز مجموعه ای از اهداف نظامی در یک میدان را نشان می دهد. تصویر بوسیله تداخل تقریباً متناوب پویشگر که به صورت ریپل^۲ در جهت عمودی دیده می شود، خراب شده است. همان طور که در شکل ۵.۱۳ (الف) دیده می شود، این تداخل به علت طبیعت متناوبیش دسته هایی از انرژی مرکز را در محور عمودی طیف فوریه تصویر تولید می کند.

یک روش ساده برای کاهش اثر تداخل استفاده از فیلتر شیاری^۳ ($H(u, v)$) است که مقادیر تبدیل فوریه در محور عمودی را تضعیف می کند و سایر مقادیر تبدیل را شبیه روال مشروح در بخش ۷.۵، تنها در عدد یک ضرب می کند. شکل ۵.۱۳ (ب) چنین فیلتری را که بر روی طیف تصویر قرار گرفته است، نشان می دهد. نوارهای تاریک نواحی تضعیف کننده فیلتر هستند.

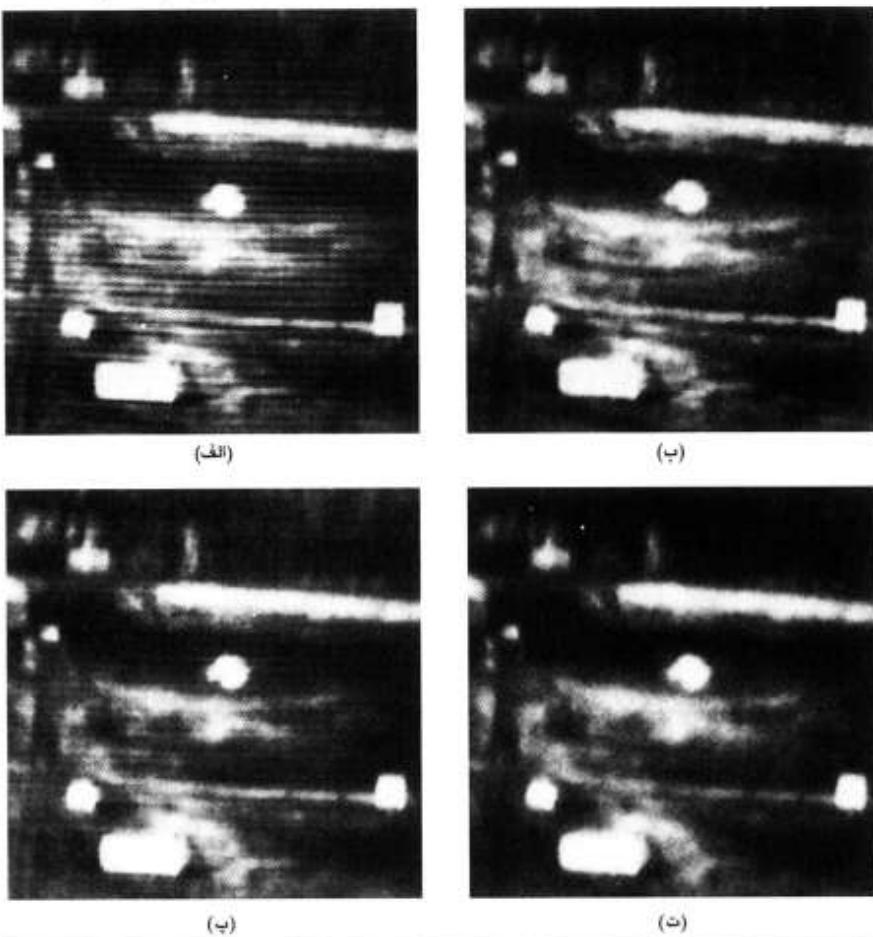
شکل ۵.۱۲ (ب) نتیجه استفاده از فیلتر شیاری و گرفتن تبدیل فوریه معکوس را نشان می دهد. توجه کنید که کاملاً تداخل از تصویر حذف شده است. تصویر شکل ۵.۱۲ (ب) با اعمال یک نقاب کانولوشن

1) Median filtering

2) Ripple effect

3) Notch filter (فیلتر باند نگلر خیلی نیز و یا شیار گرچک)

٩.٥ تبدیل های هندسی ٣٥٩

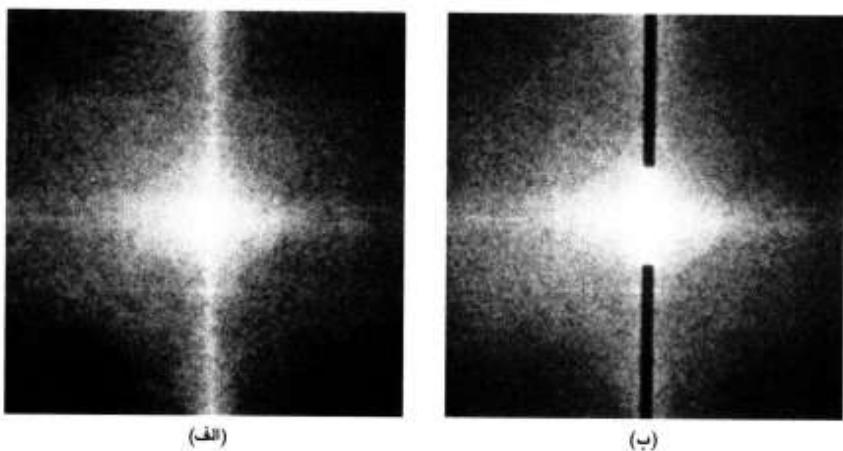


شکل ١٢.٥ (الف) تصویر مادون قرمزی که تداخل رانشان می دهد؛ (ب) تصویر بازیابی شده با استفاده از یک فیلتر شیاری در میدان فرکانس؛ (پ) تصویر بازیابی شده با استفاده از ناقاب کاتنولوشن 9×9 ؛ (ت) حاصل دو بار اعمال ناقاب . (Gonzalez و Meyer 1983)

۹ (بخش ٤ . ١ را بینید) بر تصویر خراب شده اولیه بدست آمد. ضرایب این ناقاب از فیلتر شیاری و با استفاده از معادله^٤ (۱۲-۵) تولید شدند. این ناقاب کوچک تنها تقریبی از فرآیند فیلتر کردن فوریه است، بنابراین هنوز چند خط عمودی در تصویر پردازش شده رویت می شوند. همان طور که شکل ۵ (ت) نشان می دهد، با عبور دوم ناقاب، تداخل (بازای مقدار قابل توجهی ماتی) باز هم کاهش یافت.

٩.٥ تبدیل های هندسی

این فصل را با شرحی مقدماتی در مورد کاربرد تبدیل های هندسی برای بازیابی تصویر پایان می دهیم.



شکل ۱۲.۵ (الف) طیف فوریه تصویر شکل ۱۲.۵ (الف)؛ صافی شیاری اعمال شده بر طیف.
(. Gonzalez و Meyer از ۱۹۸۳)

برخلاف روش هایی که تاکنون تشریح شدند، تبدیلات هندسی عموماً روابط مکانی بین پیکسل های تصویر را تغییر می دهند. تبدیلات هندسی اغلب تبدیلات صفحه لاستیکی^۱ نامیده می شوند، زیرا می توان آنها را مانند فرآیند چاپ تصویر روی صفحه لاستیکی و سپس کشش این صفحه طبق مجموعه ای از قانونهای پیش تعریف شده در نظر گرفت.

از نظر پردازش تصویر رقمه، تبدیل هندسی از دو عمل پایه ای تشکیل می شود: (۱) یک تبدیل مکانی^۲ که باز آرایی پیکسل ها روی صفحه تصویر را تعریف می کند؛ و (۲) یک درون یابی سطح خاکستری^۳ که به انتساب سطوح خاکستری به پیکسل های تصویر حاصل از تبدیل مکانی می پردازد. در بخش های بعدی نظریه های اساسی این مفاهیم و فایده آنها در مبحث بازیابی تصویر را شرح می دهیم.

۱.۹.۵ تبدیل های مکانی

فرض کنید که تصویر f با مختصات پیکسلی (x, y) تحت اعوجاج هندسی قرار می گیرد تا تصویر g با مختصات (\hat{x}, \hat{y}) را تولید شود. این تبدیل را می توان به صورت

$$\hat{x} = r(x, y) \quad (1.9.5)$$

1) Rubber-sheet transformations

2) Spatial transformation

3) Gray-level interpolation

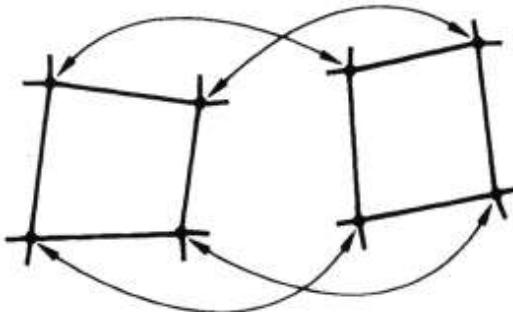
$$\hat{y} = s(x, y) \quad (2-9. 5)$$

بیان کرد که $r(x, y)$ و $s(x, y)$ تبدیلات مکانی هستند که تصویر اعوجاج هندسی یافته (\hat{x}, \hat{y}) را تولید کرده اند. برای مثال، اگر $s(x, y) = y/2$ و $r(x, y) = x/2$ باشد، اعوجاج، صرفاً کوچک کردن ابعاد $f(x, y)$ به یک دوم در هر دو جهت مکانی است.

اگر $r(x, y)$ و $s(x, y)$ به طور تحلیلی معلوم باشند، ممکن است به طور نظری بازیابی (x, y) از تصویر اعوجاج یافته (\hat{x}, \hat{y}) با اعمال تبدیلات به ترتیب عکس مقدور باشد. اما در عمل عموماً بیان تحلیلی یک مجموعه توابع $(y, r(x, y))$ که فرآیند اعوجاج هندسی را روی تمام صفحه تصویر توصیف نمایند، ممکن نیست. روشی که به طور فراوان برای غلبه بر این مشکل استفاده می شود، بیان تغییر مکان پیکسل ها با استفاده از نقاط گره¹ است، که نقاط گره زیر مجموعه ای از پیکسل هایی هستند که محل آنها در تصاویر ورودی (اعوجاج یافته) و خروجی (تصحیح شده) به دقت معلوم است.

شکل ۱۴.۵ نواحی چهار ضلعی در تصویر اعوجاج یافته و متناظر تصحیح شده آن را نشان می دهد. گوشه های چهار ضلعی ها نقاط گره متناظر هستند. فرض کنید که فرآیند اعوجاج هندسی درون نواحی چهار ضلعی با یک زوج معادلات دو خطی مدل می شود طوری که

$$r(x, y) = c_1x + c_2y + c_3xy + c_4 \quad (3-9. 5)$$



شکل ۱۴.۵ نقاط گره متناظر در دو قطعه تصویر

1) Tiepoint

و

$$s(x, y) = c_5x + c_6y + c_7xy + c_8. \quad (4-9. 5)$$

آنگاه از معادلات (۵. ۱-۹) و (۵. ۲-۹) داریم که

$$\dot{x} = c_1x + c_2y + c_3xy + c_4 \quad (5-9. 5)$$

و

$$\dot{y} = c_5x + c_6y + c_7xy + c_8. \quad (6-9. 5)$$

از آنجاکه جمیعاً ۸ نقطه گره معلوم وجود دارد، بسادگی می‌توان معادلات را برای c_1, c_2, \dots, c_8 حل کرد. این ضرایب مدلی رامی‌سازند که برای تبدیل تمام پیکسل‌های درون چهار ضلعی مشخص شده با گره‌ها به کار می‌رود. توجه شود که ضرایب از همین گره‌ها بدست آمدند. به طور کلی، برای تولید مجموعه‌ای از چهار ضلعی‌ها که تمام تصویر را پوشانند، تعداد کافی گره مورد نیاز است که هر یک از این چهار ضلعی‌ها مجموعه ضرایب خود را دارد.

روال مورد استفاده برای تولید تصویر تصحیح شده ساده است. مثلاً، برای تولید $f(0, 0)$ ، $f(x, y)$ را در معادلات (۵. ۵-۹) و (۵. ۶-۹) قرار دهید و یک زوج مختصات (\hat{x}, \hat{y}) را به دست آورید. آنگاه قرار دهید $\hat{y} = g(\hat{x}, \hat{y})$ و $\hat{y} = f(0, 0)$ که \hat{x} و \hat{y} مقادیر مختصات که اخیراً به دست آمد، هستند. آنگاه $f(0, 1) = g(x, y)$ را در معادلات (۵. ۵-۹) و (۵. ۶-۹) قرار دهید، زوج دیگری از مقادیر (\hat{x}, \hat{y}) را بدست آورید، و برای آن مقادیر مختصات، از تساوی $\hat{y} = g(\hat{x}, \hat{y})$ استفاده نمایید. این روال، پیکسل به پیکسل و سطر به سطر ادامه می‌یابد تا این که آرایه‌ای که ابعادش از ابعاد تصویر g تجاوز نمی‌کند، بدست آید. پویش ستونی (به جای سطري) همان نتایج را می‌دهد. همچنین، به منظور استفاده از ضرایب مناسب یک روال ثبت اطلاعات^۱ برای تعیین این که کدام چهار ضلعی در یک نقطه معین تصویر بکار رود، الزامی است.

1) Bookkeeping procedure

۲.۹.۵ درونیابی سطح خاکستری

روشی که اخیراً تشریح شد، با مجموعه مقادیر صحیح مختصات (y, x) پیش می رود تا تصویر تصحیح شده (x, y) بدست آورد. در هر حال بسته به ضرایب c ، معادلات $(5-9)$ و $(5-10)$ ممکن است مقادیر غیر صحیحی برای \hat{x} و \hat{y} نتیجه دهنند. چون تصویر اعوجاج یافته^۱ ورقی است، مقدار پیکسل های آن تنها در مختصات صحیح تعریف می شوند. بنابراین استفاده از مقادیر غیر صحیح برای \hat{x} و \hat{y} نگاشتی را به نقاطی از g که هیچ سطح خاکستری برایشان تعریف نشده است، ایجاد می نماید. تحت این شرایط اطلاع از مقدار سطوح خاکستری در محلهای با مختصات غیر صحیح، تها با استفاده از مقدار پیکسل های محل های با مختصات صحیح، الزامی می شود. روش مورد استفاده برای انجام این کار، درونیابی سطح خاکستری^۲ نامیده می شود.

ساده ترین روش برای درونیابی سطح خاکستری براساس روش نزدیک ترین همسایه است. این روش که درونیابی مرتبه صفر^۳ نیز نامیده می شود، در شکل ۱۵.۵ نشان داده می شود. این شکل سه مرحله را نشان می دهد: (۱) نگاشت مختصات صحیح (y, x) به مختصات کسری (\hat{y}, \hat{x}) با استفاده از معادلات $(5-9)$ و $(5-10)$ ؛ (۲) انتخاب نزدیکترین همسایه (\hat{y}, \hat{x}) با مختصه صحیح؛ و (۳) انتساب سطح خاکستری این همسایه به پیکسل قرار گرفته در (y, x) .

گرچه روش درونیابی نزدیک ترین همسایه، پیاده سازی ساده ای دارد، اما اغلب با تولید اثرات نامطلوبی، نظیر اعوجاج دادن لبه های مستقیم الخط در تصاویر با تفکیک بالا، مشکل ایجاد می کند. می توان با استفاده از روش های کارآمدتر نظیر درونیابی کانولوشن مکعبی^۴ [Bernstein 1976] نتایج بهتری^۵ بدست آورد. این روش درونیابی، سطحی از نوع $x/\sin(x)$ را بر تعداد بسیار بیشتری از نقاط همسایه (مثلاً ۱۶ نقطه) منطبق می کند تا بدين وسیله تخمین خوبی از سطح خاکستری هر نقطه^۶ مطلوب از تصویر بدست آورد. در هر حال از نقطه نظر محاسباتی، این روش پرهزینه است و به عنوان مصالحه ای منطقی بین دوروش مذکور می توان از روش درونیابی دوخطی^۷ که از سطوح خاکستری نزدیک ترین چهار همسایه استفاده می کند، نام برد. به عبارت دیگر، ایده این است که برای زوج مختصه غیر صحیح (\hat{y}, \hat{x}) ، سطح خاکستری هر یک از نزدیک ترین چهار همسایه با مختصات صحیح معلوم شود. مقدار سطح

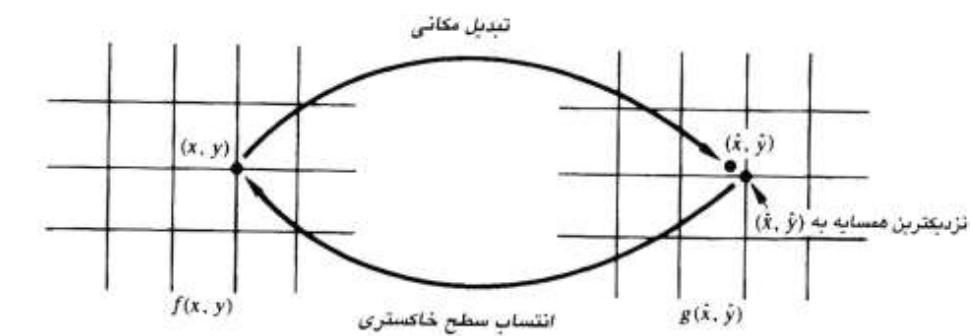
1) Gray-level interpolation

2) Zero-order interpolation

3) Cubic Convolution interpolation

4) Smoother result

5) Bilinear interpolation



شکل ۵.۱۵ درونیابی سطح خاکستری براساس مفهوم نزدیکترین همسایه

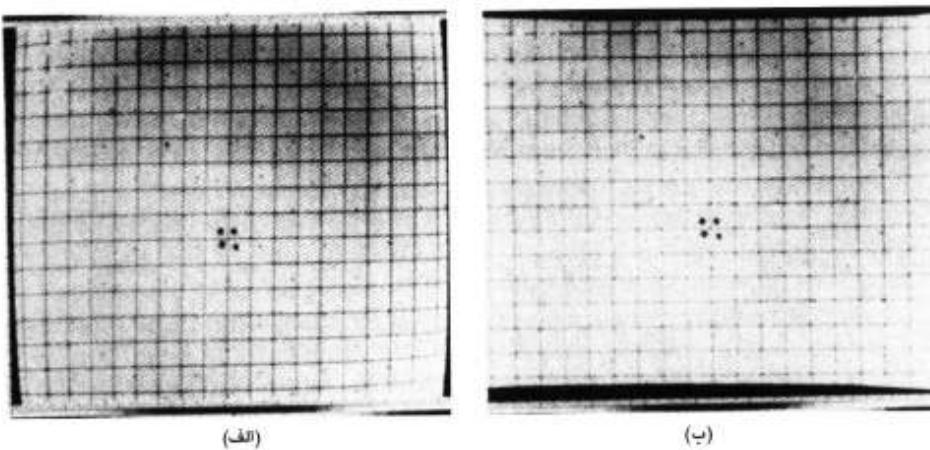
خاکستری (\hat{x}, \hat{y}) که با (\hat{x}, \hat{y}) نشان داده می شود، را می توان با استفاده از رابطه

$$v(\hat{x}, \hat{y}) = a\hat{x} + b\hat{y} + c\hat{x}\hat{y} + d \quad (7-9.5)$$

از مقادیر همسایه اش درون یابی کرد که چهار ضریب رابطه فوق بسادگی از چهار معادله چهار مجھول تعیین می شوند که این چهار معادله چهار مجھول با استفاده از چهار همسایه معلوم (\hat{x}, \hat{y}) نوشته می شوند. پس از تعیین ضرایب، مقدار (\hat{x}, \hat{y}) محاسبه می شود و این مقدار به نقطه ای در (x, y) $f(x)$ که نگاشت مکانی آن نقطه (\hat{x}, \hat{y}) بود، متنسب می شود. به تصویر کشاندن این روال با کمک شکل ۵.۱۵ آسان است. استثناء این است که به جای استفاده از مقدار سطح خاکستری نزدیکترین همسایه به (\hat{x}, \hat{y}) ، عملاً مقداری را در نقطه (\hat{x}, \hat{y}) درون یابی کرده، آن را به سطح خاکستری نقطه (x, y) متنسب می کنیم.

مثال: روش هایی که در این بخش و بخش ۹.۱ مطرح شدند، رامی توان با کاربرد در مسأله تصحیح تصویری که اعوجاج هندسی یافته است، مصور کرد. تصویر مورد پردازش در شکل ۵.۱۶ (الف) دیده می شود. این تصویر اعوجاج "بارل"^{۱)} را که در بسیاری از دوربینهای تصویربرداری مبتنی بر ویدیکون وجود دارد، به نمایش می گذارد. شبکه مستطیلی که در شکل ۵.۱۶ (الف) دیده می شود، بویژه نزدیک لبه های تصویر به شدت اعوجاج یافته است. همچنین توجه شود که اعوجاج، یکنواخت نیست و درجه اعوجاج به طور غیر خطی به صورت تابعی از فاصله تا مرکز تصویر افزایش می یابد. همان طور که در بخش ۹.۱ بیان شد، استفاده از معادلات (۵.۹-۵) و (۶-۹) نیازمند اطلاع

1) Barrel distortion



شکل ۱۶.۵ (الف) تصویر اعوجاج یافته؛ (ب) تصویر پس از تصحیح هندسی. (از O'Handley و Green[1972]).

از نقاط گره در هر دو تصویر اعوجاج یافته و تصحیح شده می باشد. در این حالت ویژه، نقاط گره، علامت های رزیو^۱ هستند که در شکل ۱۶.۵(الف) به صورت نقاط تیره کوچکی در سراسر تصویر پخش شده اند، دیده می شوند. (علامت های رزیو مربع های فلزی کوچکی هستند که مستقیماً روی سطح لامپ دوربین نهاده شده اند). چون محل این علامت دقيقاً معلوم است، از آنها به عنوان نقاط گره مطلوب استفاده می شود. شکل ۱۶.۵(ب) نتیجه استفاده از معادلات (۵.۵-۹) و (۵.۹-۶) برای نگاشت های مکانی و نتیجه کاربرد معادله (۵.۷-۹) برای درونیابی سطح خاکستری را نشان می دهد. به زیادی تصحیح هندسی حاصل از این معادلات توجه کنید.

مثال قبل تنها یکی از کاربردهای فراوان تبدیلات هندسی برای بازیابی تصویر را بیان می کنند. کاربرد مهم دیگر آنها، ثبت تصویر^۲ یا پیدا کردن تناظر بین دو تصویر است. فرآیند ثبت تصویر همانند روشی است که اخیراً برای تصحیح هندسی نشان داده شد. با این حال، تبدیل های تصویر که حاصل آنها متناظر با تصویر همان صحنه اما از زاویه دیگر باشد، مورد تأکید است. کاربردهای دیگری از روش های مشروح در این بخش، شامل تصحیح اعوجاجات نمایشگر، نگاشت های نقشه^۳، و نگاشت های کارتوگرافیک^۴ می باشد. کتابهای Castleman[1979] و Green[1983] حاوی مثالهای متعددی از این

1) Reseau mark

2) Image registration

3) Map projection

4) Cartographic projection

کاربردها است.

در بسیاری موارد تعیین نقاط گره' متناظر در دو تصویر، ممکن است کار نسبتاً مشکلی باشد. هر وضعیتی را نمی توان با وجود اثرات کنترل شده نظیر علامت های رزیو، مشخص کرد. وقتی علامت ها از پیش معلوم نباشند، معمولاً نقاط گره' با استفاده از روش های همبستگی (فصل ۹ را ببینید) تعیین می شوند تا ویژگی های متناظر در دو تصویر را بیابیم. با این حال معیارهای همبستگی تحت تأثیر عاملهای نظیر نویز و چرخش تصویر هستند و بنابراین معمولاً تأثیر مکانی کم دقت تری را بین نقاط گره' ارائه می دهند.

۱۰۰.۵ نکات پایانی

مفاهیم اصلی مطرح شده در این فصل بیانی از مسأله بازیابی تصویر در چهار چوب جبرخطی و ساده سازی جوابهای جبری براساس خواص ماتریسها ی چرخشی و بلوك- چرخشی هستند.

بیشتر روش های بازیابی بدست آمده در بخش های قبلی براساس معیار بهینگی حداقل مربعات هستند. کلمه بهینه^۱ در این مبحث تنها به یک مفهوم ریاضی و نه به پاسخ بهینه سامانه بینایی انسان اشاره دارد. درواقع کمبود فعلی دانش درباره' دریافت بصری، باعث می شود که توانیم رابطه ای کلی برای مسأله بازیابی تصویر که ترجیحات و توانایی های مشاهده گر را به حساب آورد، داشته باشیم. از دید این محدودیت ها، مزیت روال های این فصل این است که روشی پایه ای تولید می کنند که از آن می توان مجموعه ای از نتایج قیلاً معلوم (اما نامتحد^۲) را به دست آورد. بنابراین قدرت روش جبری در این است که می توان با شروع از اصول پایه ای یکسان سادگی روش های نظیر فیلترهای وینر و حداقل مربعات مقید را به دست آورد.

نتایج به دست آمده در ۸ بخش اول این فصل، براساس فرض خطی و مستقل از مکان بودن تخریب هستند. این فرض فوراً انتگرال کالولوشن را نتیجه می دهد که می توان آن را بر حسب مدل تخریب پایه ای داده شده با معادله $5 = 1(24)$ به شکل گسته بیان کرد. فرض متناسب بودن توابع ورودی با تولید ماتریس های چرخشی و بلوك- چرخشی مسأله را بیشتر ساده کرد. این ماتریس ها امکان می دهند که تمام روش های بازیابی بدست آمده در میدان فرکانس با استفاده از الگوریتم FFT دو بعدی اجرا شوند، بنابراین پیچیدگی محاسباتی ناشی از بیان ماتریسی اولیه فرآیند تخریب را بسیار کاهش می دهند.

مطلوب بخش ۵.۸ راه مناسبی را برای پیاده سازی تقریبی تابع بخش های ۵.۲-۵.۵ در حوزه مکان فراهم می آورد. در نهایت بخش ۵.۹، مسأله بازیابی تصاویری را که اعوجاج هندسی یافته اند، مطرح می نماید.

مراجع

تعریف های ارائه شده در بخش ۵.۱ از Friedland & Schwarz [1965] گرفته شده اند. زمینه مورد نیاز برای اغلب عملیات ماتریسی پایه ای مورد استفاده در این فصل در Deutsch [1965] ، Noble [1969] و Bellman [1970] موجود است. بیان مدل تخریب گسته بر حسب ماتریس های چرخشی و بلوك-چرخشی بر اساس دو مقاله از Hunt [1971] ، ۱۹۷۳ می باشد. همچنین در این مقالات و کتاب Bellman [1970] خواص قطری سازی مشروح در بخش ۵.۲ بررسی شده اند. برای اطلاعات اضافی در مورد مطالب بخش ۵.۳ و نیز استنتاج جبری روش های بازیابی گوناگون مورد استفاده در این فصل، Bellman & Andrews [1977] را ببینید. کتاب اخیر، که همه آن به مبحث بازیابی تصویر اختصاص یافته است، به تفصیل به سایر روش های بازیابی و نیز آنهایی که در اینجا بیان شدند، می پردازد.

محققان متعددی روش فیلترینگ معکوس را بررسی کرده اند. مراجع بخش ۵.۴ McGlamery [1967] ، Sondhi [1972] ، Hall & Cutrona [1968] ، Slepian [1967] هستند. مراجع دیگر در مورد روش بازیابی حداقل مربعات ارائه شده در بخش ۵.۵، Slepian [1967] ، Helstrom [1967] ، Kak & Rosenfeld [1982] ، Horner [1969] ، Rino [1969] ، Harris [1968] مقایسه روش های کلاسیک ارائه شده در این مراجع با روش جبری بخش ۵.۵ جالب است. مطالب بخش ۵.۶-۵.۷ بر اساس مقاله Hunt [1973] می باشد. سایر مراجع مرتبط با عنایون مشروح در بخش های Paris & Shack [1964] ، Twomey [1963] ، Phillips [1962] ، Pollak & Slepian [1961] ، Blackman [1968] ، Reynolds & Muellen [1967] ، Harris [1966] ، Lohman [1965] Som ، Falconer [1970] ، MacAdam [1970] ، Harris & Rushforth [1968] ، Huang [1968] Huang و Robbins ، Sawchuck [1972] ، Habibi [1972] ، Frieden [1972، ۱۹۷۴] ، [1971] Netravali [1976] ، Anderson [1974] ، Angel [1974] و Jain ، Andrews [1974] ، [1972] Gonzalez & Meyer [1983] است. شاید مطالب دیگری در مورد مباحث بخش ۵.۸ از

۹. ۵ رابطان در Green و Castleman[1979]، Bernstein[1976]، Green[1972] و O'Handley [1983] پیدا نمود.

برای مطالعه بیشتر در مورد مبحث کلی بازیابی تصویر به Stark، McDonnell & Bates[1986]، Slaney[1988]، Jain[1989]، Kak و [1987] تومورنگاری رایانه ای^۱ می پردازد. این موضوع که گرچه خارج از بحث ما است، به طور قابل ملاحظه ای در تصویربرداری پزشکی مورد علاقه است.

مسائل

۱. ۵ یک سامانه تحریب تصویر مستقل از مکان خطی با پاسخ ضربه تصویری است که دارای خطی با پهنای بسی نهایت کوچک در $a = x$ می باشد، و با $f(x, y) = \delta(x - a)$ مدل شده است. با فرض عدم وجود نویز، تصویر خروجی $g(x, y)$ را معین کنید.

درستی معادله (۵. ۲-۸) را نشان دهید.

۲. ۵ یک دانشمند باستان شناس که درباره 'پول رایج در زمان امپراطوری روم تحقیق می کند، آگاه شد که چهار سکه رومی مربوط به تحقیقش در فهرست موجودی موزه بریتانیا در لندن قرار دارند. متأسفانه پس از رسیدن به آنجا متوجه شد که اخیراً سکه ها ربوده شده اند. به او گفته شد که از تمام اقلام موجودی موزه عکس های تهیه شده است. اما عکس های سکه های مورد نظر آنقدر مات هستند که تاریخ و سایر علامت های کوچک آنها خواندنی نیست. ماتی عکس ها ناشی از متمرکز نبودن دوربین هنگام تصویربرداری آنها می باشد. شما به عنوان مشاور استخدام شده اید تا تعیین کنید که آیا می توان با پردازش رایانه ای تصاویر را بازیابی کرد طوری که دانشمند بتواند علائم آن را بخواند. به شما گفته می شود که دوربین اولیه مورد استفاده برای گرفتن عکس ها هنوز هم در دسترس است، و سکه های دیگری از همان دوره نیز موجود هستند. یک راه حل مرحله ای برای این مسأله پیشنهاد نمایید.

1) Computerized tomographic imaging

- ۴.۵** معادله ای شبیه به معادله (۵.۴-۱۳)، اما برای سرعت یکتواخت دلخواه در هر دو جهت x و z بدست آورید.
- ۵.۵** مسأله مات شدن تصویر، ناشی از شتاب یکتواخت در جهت x را در نظر بگیرید. اگر تصویر در لحظه $t=0$ در حالت سکون باشد و به مدت T ثانیه حرکت شتابدار یکتواخت با معادله $\ddot{x}_0(t) = at^2/2$ را داشته باشد، تابع انتقال $H(u, v)$ را باید.
- ۶.۵** یک کاوشگر فضایی به این منظور طراحی می شود که وقتی برای فرود در سیاره ای به آن نزدیک می شود، تصاویر آن را به زمین ارسال کند. در خلال آخرین مراحل فرود یکی از بخش های کنترل خراب می شود که باعث چرخش سریع سفینه حول محور عمودی می گردد. بر اثر این حرکت چرخشی تصاویر ارسالی در مدت دو ثانیه آخر قبل از فرود، مات می شوند. دوربین در انتهای کاوشگر در طول محور عمودی آن و رو به پایین قرار دارد. خوشبختانه، چرخش سفینه نیز حول محور عمودی آن است، بنابراین تصاویر بر اثر حرکت چرخشی یکتواخت مات می شوند. به علاوه، چرخش سفینه در مدت برداشتن هر تصویر به $8\pi/\text{رادیان}$ محدود بود. فرآیند تصویربرداری را می توان به صورت دریجه ایده آگی مدل کرد که تنها در مدتی که سفینه $8\pi/\text{رادیان}$ بچرخد، باز است. می توانید فرض کنید که حرکت عمودی در خلال تصویربرداری قابل چشم پوشی است. چگونه از مقاهیمی که در این فصل یاد گرفته اید، برای بازیابی تصاویر استفاده می کنید؟ از شما ارائه راه حل ویژه ای خواسته نمی شود، بلکه ارائه یک راه حل پایه ای خواسته می شود.
- ۷.۵** جواب مشخصی (به صورت یک یا چند معادله) برای مسأله ۶.۵ ارائه نمایید و تمام آنچه برای رسیدن به آن جواب فرض نموده اید، فهرست نمایید.
- ۸.۵** (الف) نشان دهید که چگونه معادله (۵.۸-۵) از معادله (۵.۷-۵) نتیجه می شود.
(b) نشان دهید که چگونه معادله (۵.۹-۵) از معادله (۵.۸-۵) نتیجه می شود.
- ۹.۵** ماتی تصویر حاصل از نورخوردگی طولانی در اغتشاشات جوی را می توان با تابع انتقال $H(u, v) = \exp[-(u^2 + v^2)/2\sigma^2]$ مدل کرد. نویز راناچیز فرض کنید. معادله فیلتر وینری که شما برای بازیابی تصویر مات شده با این نوع تخریب استفاده خواهید کرد، چیست؟
- ۱۰.۵** فرض کنید که مدل شکل ۱.۵، خطی و مستقل از مکان است و نشان دهید که طبیف توان

خروجی با رابطه $|G(u, v)|^2 = |H(u, v)|^2|F(u, v)|^2 + |N(u, v)|^2$ داده می شود، به معادله (۵.۴۰-۲) رجوع نمایید.

۱۱.۵ Cannon [1974] فیلتر بازیابی $R(u, v)$ را پیشنهاد کرد که شرط

$|\hat{F}(u, v)|^2 = |R(u, v)|^2|G(u, v)|^2$ را برابر آورده می نماید و مبتنی بر این فرض است که طیف توان تصویر بازیابی شده، $|\hat{F}(u, v)|^2$ ، برابر طیف توان تصویر اولیه، $|F(u, v)|^2$ است.

(الف) $R(u, v)$ را بر حسب $|F(u, v)|^2$, $|H(u, v)|^2$ و $|N(u, v)|^2$ پیدا کنید. (راهنمایی: به شکل

۵.۱، معادله (۵.۴-۲) و مسئله ۵.۱۰ مراجعه نمایید)

(ب) از نتیجه قسمت (الف) برای بیان نتیجه ای به شکل معادله (۵.۹-۵) استفاده کنید.

۱۲.۵ فرض کنید که هر عنصر تصویر به محدوده $[0, 1]$ تراز شود. آنگاه می توان هر یک از چنین عناصری را به عنوان احتمال برخورد تعداد معینی فوتون به محل معینی در تصویر تعبیر کرد.

آنتروپی^۱ به صورت $-p \ln p = E$ تعریف می شود که p احتمال و \ln لگاریتم طبیعی است (فصل ۶ را ببینید). آنتروپی یک تصویر تراز شده که به شکل برداری بیان شده است، را به صورت $-\ln \prod p_i = E$ تعریف می کنیم. در این طرز بیان، بردار \mathbf{p} با گرفتن لگاریتم طبیعی از هر جزو Ω ایجاد می شود. یک فیلتر مفید برای پرداختن به تحریبها براساس مدل دانه ای تصادفی^۲ (شبیه به مدل سازی حالت دانه ای فیلم^۳)، با کمینه سازی مقید حداقل مربعات آنتروپی منفی بدست می آید. نشان دهید که تصویر بازیابی شده با معادله غیر جبری

$$\mathbf{g} = \exp[-\mathbf{H}^T(\mathbf{H}\mathbf{f} - \mathbf{H}\mathbf{p})]$$

می توان با سطح جمله^۴ نمایی به صورت سری تیلور و نگه داشتن قسمت خطی آن، تقریبی خطی برای جواب دارای آنتروپی بیشینه در مسئله ۱۲.۵ بدست آورد. نشان دهید که این روش رابطه حداقل مربعات مقید داده شده با معادله (۵.۹-۳) را، اما با $\mathbf{Q} = \mathbf{I}$ ، نتیجه می دهد.

۱۴.۵ یک آرایش هندسی معین برای تصویر برداری اشعه ایکس، تخریب مات کننده ای را ایجاد می کند که می توان آن را به صورت کانولوشن تصویر حسن شده با تابع مکانی متقاضان چرخشی $h(r) = [(r^2 - 2\sigma^2)/\sigma^4] \exp[-r^2/2\sigma^2]$ مدل کرد که در آن $r^2 = x^2 + y^2$ است. تابع انتقال

مسایل ۳۲۱

یک فیلتر حداقل مربعات مقید را که با استفاده از آن می توانید تصاویر تولید شده با این سامانه اشue ایکس را مات زدایی کنید، به دست آورید. می توانید فرض کنید که تصاویر مربعی هستند.

- با معادله (۵ . ۷-۱۲) شروع نماید و معادله (۵ . ۷-۱۴) را به دست آورید.
- ۱۵.۵ فرض کنید که در بخش ۵ . ۹ برای تعیین تبدیل مکانی و درون یابی سطح خاکستری به جای استفاده از چهار ضلعی، از نواحی مثلثی استفاده کردید. برای نواحی نواحی مثلثی معادلات مشابه با معادلات (۵ . ۷-۵)، (۵ . ۷-۶) و (۵ . ۹-۷) چیست؟
- ۱۶.۵

فصل ۶

فسرده سازی تصویر

از آنجاکه زندگی کوتاه و اطلاعات بین بایان است...
مختصرسازی، یک قن لازم است و کار مختصرساز
کرجه ذاتی بد است، اما از نبودش بهتر است «
الدوس هاکسلی

وقتی برای تولید یک تصویر رقمی، تابع شدت نور دو بعدی متناظر، نمونه برداری و چندی^۱ می شود، مقادیر بی شماری «داده» تولید می گردد. در واقع مقدار داده های تولید شده ممکن است آنقدر بزرگ باشد که ذخیره سازی، پردازش و مخابره آن غیر عملی باشد. در چنین حالاتی، به نمایش هایی غیر از نمونه برداری دو بعدی و چندی سازی ساده که در بخش ۲.۳ توضیح داده شد، نیاز است. به عنوان مثال، برای نمایش دایرة المعارف *Encyclopaedia Britannica* در شکل رقمی، بیش از 25×10^9 گیگابایت (25 بایت) داده موردنیاز است.[†]

فسرده سازی تصویر به موضوع کاهش مقدار داده های موردنیاز برای نمایش تصویر رقمی می پردازد. اساس فرآیند کاهش، حذف داده های زاید است. از نقطه نظر ریاضی، این فرآیند معادل تبدیل یک آرایه پیکسلی دو بعدی به یک مجموعه داده ناهمبسته آماری است که معمولاً این تبدیل، قبل از ذخیره سازی یا ارسال تصویر انجام می شود. تصویر فشرده در زمانی دیگر یا در مقصد، وافسرده می شود تا تصویر اولیه یا تقریبی از آن، بازسازی شود.

توجه به فشرده سازی تصویر بیش از ۲۵ سال سابقه دارد. در ابتدا تحقیقات در این موضوع، روی توسعه روش های آنالوگ برای کاهش عرض باند ارسال ویدیو که فشرده سازی عرض باند^۲ خوانده

[†] این دایرة المعارف حدود ۲۵۰۰۰ صفحه دارد. اگر هر صفحه با نرخ ۳۰۰ نقطه در اینچ نمونه برداری و به دو سطح چندی شود، بیش از ۸۰۰۰۰۰۰ بایت (۸۰۰۰۰۰۰ بایت) داده تولید می کند.

می شود، متمن کری بود؛ گرچه اختراع رایانه های رقمی و بدنبال آن توسعه مدارهای مجتمع پیشرفت به باعث شد که توجه از روش های فشرده سازی آنالوگ به روش های رقمی معطوف شود، اخیرآبا پذیرش جهانی چند استاندارد فشرده سازی تصویر، زمینه رشد این موضوع از طریق کاربرد عملی کارهای نظری شروع شده از دهه ۱۹۴۰ فراهم گردید. شانون و همکارانش در این دهه برای اولین بار روابط احتمالاتی را برای میزان اطلاعات و نمایش، ارسال و فشرده سازی داده ها بیان کردند.

نیاز به فشرده سازی تصویر در طول سال های اخیر بتدریج رشد کرده است به طوری که اکنون این مبحث به عنوان یک «فناوری فعال ساز^۱» شناخته می شود. به عنوان مثال همواره فشرده سازی تصویر در افزایش محاسبات چند رسانه ای (یعنی استفاده از رایانه های رقمی در چاپ و نشر و تولید و توزیع ویدیو) بسیار مؤثر بوده و هست. بعلاوه، فشرده سازی تصویر، فناوری کار با تفکیک های مکانی افزایش یافته در حسگرهای تصویربردار امروزی و استانداردهای در حال تکامل پخش تلویزیون است. گذشته از این، فشرده سازی تصویر نقش خیلی مهمی در بسیاری از کاربردهای مهم و گوناگون شامل گردهمایی تصویری از راه دور، سنجش از راه دور (استفاده از تصویربرداری ماهواره ای در کاربردهای هوایی و سایر کاربردهای زمینی)، تصویربرداری استناد، تصویربرداری پزشکی، ارسال دورنگار^۲ و کترل وسایل هدایت از راه دور در کاربردهای نظامی، فضایی و مواد زايد خطرناک بازی می کند. خلاصه این که تعداد روش به افزایشی از کاربردها به پردازش، ذخیره سازی و ارسال پر بازده تصاویر دودویی، خاکستری یا رنگی نیاز دارد.

در این فصل به هر دو جنبه نظری و عملی فشرده سازی تصویر می پردازیم. بخش های ۱.۶ تا ۳.۶ مبانی فرآیند فشرده سازی تصویر را شرح می دهند. بخش ۱.۶ افزونگیهای داده ها^۳ را، که می توان از آنها در الگوریتم های فشرده سازی تصویر بهره برداری کرد، توصیف می کند، بخش ۲.۶ یک طرح مبتنی بر مدل^۴ از فرآیند کلی فشرده سازی- و افسرده سازی^۵ ارائه می کند و بخش ۳.۶ تعدادی از مفاهیم پایه ای نظریه اطلاعات و نقش آنها در تعیین محدودیت های اساسی نمایش اطلاعات را مقداری شرح می دهد.

بخش های ۴-۶.۶ جنبه های عملی فشرده سازی تصویر شامل روش های اصلی مورد استفاده و

1) Enabling technology

2) Facsimile(FAX)

3) Data redundancy

4) Model-based

5) Compression- decompression

استانداردهایی که عاملی برای افزایش محدوده و پذیرش این فرآیند بوده‌اند، را بیان می‌کنند. روش‌های فشرده‌سازی به دو گروه وسیع تقسیم می‌شوند: نگهدار اطلاعات^۱ (بدون اتلاف) و با اتلاف^۲. بخش ۶.۴ به روش‌هایی در گروه اول، که بسویه در بایگانی تصویر-مانند ذخیره‌سازی مدارک پزشکی یا قانونی-مفید هستند، می‌پردازد. این روش‌ها امکان فشرده و وافشده‌سازی تصویر را بدون ازین رفتن اطلاعات فراهم می‌کنند. بخش ۶.۵ روش‌هایی از گروه دوم را شرح می‌دهد. این روش‌ها سطوح بالاتری از کاهش داده را فراهم می‌آورند اما موجب بازسازی غیرکامل تصویر اولیه می‌شوند. فشرده‌سازی با اتلاف تصویر در کاربردهایی نظیر پخش تلویزیونی، گردشمالی تصویری^۳، و انتقال دورنگار مفید است که در آنها تحمل مقدار معینی خطای افزایش بهره^۴ فشرده‌سازی منطقی است. در نهایت بخش ۶.۶ به استانداردهای فشرده‌سازی تصویر که تاکنون ارائه شده‌اند، می‌پردازد.

۱۰۶ مبانی

عبارت فشرده‌سازی داده^۵ به فرآیند کاهش مقدار داده‌های مورد نیاز برای بیان یک مقدار معین از اطلاعات اشاره می‌کند. تفاوت آشکار بین داده و اطلاعات باید مطرح شود؛ آنها مترادف نیستند. در واقع داده‌ها، وسایل حمل اطلاعات هستند. از مقادیر متفاوتی داده می‌توان برای نمایش یک مقدار ثابت اطلاعات استفاده کرد. یک مثال آن حالتی است که یک فرد «پر حرف» و یک فرد «کم حرف» هر دو بخواهند یک داستان را تعریف کنند. در این مورد اطلاعات مورد علاقه، داستان است و کلمات، داده‌های مورد استفاده برای بیان اطلاعات هستند. اگر این دو نفر تعداد متفاوتی از کلمات را برای تعریف داستان به کاربرند، دونسخه متفاوت از داستان ایجاد می‌شود و حداقل یکی از دو نقل، دارای داده‌های غیرضروری است. یعنی دارای داده‌ها (یا کلماتی) است که یا هیچ اطلاعات مرتبط را ارائه نمی‌کنند یا فقط چیزهایی را که قبلاً معلوم بوده‌اند، دوباره بیان می‌کنند. بنابراین گفته می‌شود که این نقل، افزونگی داده دارد.

افزونگی داده مهمترین مورد در فشرده‌سازی تصویر رقمی است. افزونگی داده مفهوم مجردی^۶ نیست بلکه یک موجودیت^۷ کمیت پذیر ریاضی است. اگر n_1 و n_2 بیانگر تعداد واحدهای حامل اطلاعات در دو

1) Information preserving

2) Lossy

3) Video conferencing

4) Data compression

5) Abstract

6) Entity

مجموعه داده باشد که هر دو یک مقدار اطلاعات را نمایش دهند، می توان افزونگی نسبی داده، n_3 ، مربوط به اولین مجموعه داده (که با n_1 مشخص شد) را به صورت

$$R_D = 1 - \frac{1}{C_R} \quad (1-1.6)$$

تعریف کرد که نسبت فشرده سازی C_R ، عبارتست از:

$$C_R = \frac{n_1}{n_2} \quad (2-1.6)$$

اگر $n_1 = n_2$ باشد، $R_D = 1$ و $C_R = 0$ می باشد که نشان می دهد نمایش اول اطلاعات (نسبت به دومین مجموعه داده) حاوی هیچ داده زایدی نیست. اگر $n_1 \ll n_2$ باشد، $C_R \rightarrow \infty$ و $R_D \rightarrow 1$ میل می کنند که نمایانگر فشرده سازی زیاد و داده های زاید فراوان می باشد. در حالت آخر اگر $n_1 \gg n_2$ باشد، $C_R \rightarrow 0$ و $R_D \rightarrow \infty$ میل می کنند، که نشان می دهد دومین مجموعه داده دارای داده های بسیار بیشتری نسبت به نمایش اولیه است. البته قاعده ایین حالت، حالت غیر مطلوبی از گسترش داده ها است. به طور کلی C_R و R_D بترتیب در بازه های $(0, \infty)$ و $(1, \infty)$ قرار دارند. یک نسبت فشرده سازی عملی نظیر ۱۰ بدين معناست که تعداد واحد های حامل اطلاعات (مثلاً بیتها) در مجموعه اول ۱۰ برابر مجموعه دوم یا فشرده شده است. افزونگی متاظر ۹۰ بیان می کند که ۹۰ درصد داده های مجموعه اول زاید هستند.

در فشرده سازی تصویر رقمه سه نوع افزونگی داده را می توان شناسایی و از آنها بهره برداری کرد که عبارتند از: افزونگی رمزنگاری^۱، افزونگی بین پیکسلی^۲، و افزونگی روان بصری^۳. وقتی فشرده سازی داده حاصل می شود که دست کم یکی از این افزونگی ها کاهاش یابند یا حذف شوند.

۱.۱.۶ افزونگی رمزنگاری

در فصل ۴ با فرض اینکه سطوح خاکستری تصویر، کمیت های تصادفی هستند، با کمک روابط، ارتقای تصویر با تغییر بافت نگار را بیان کردیم. نشان دادیم که مقدار زیادی از اطلاعات ظاهری تصویر را می توان از بافت نگار و سطوح خاکستری آن بدست آورد. در این بخش از بیان مشابهی استفاده می کنیم تا

1) Compression ratio 2) Coding redundancy 3) Interpixel redundancy 4) Psychovisual redundancy

نشان دهنم چگونه بافت نگار سطح خاکستری تصویرمی تواند بصیرت فراوانی در مورد درون ساختار رمزها^۱ فراهم نماید تا مقدار داده های مورد نیاز برای نمایش آن را کاهش دهد.

یکبار دیگر اجازه دهد فرض کنیم که یک متغیر تصادفی گسته r دریازه [0, 1] سطح خاکستری یک تصویر را نمایش می دهد و هر r با احتمال $p(r)$ رخ می دهد. مانند فصل ۴،

$$p(r_k) = \frac{n_k}{n} \quad k = 0, 1, 2, \dots, L - 1 \quad (3-1.6)$$

که L تعداد سطوح خاکستری، n تعداد دفعاتی که سطح خاکستری k ام در تصویر رخ می دهد و n تعداد کل پیکسل های تصویر است. اگر تعداد بیت های به کار رفته در نمایش هر مقدار r_k باشد، تعداد متوسط بیت های مورد نیاز برای نمایش هر پیکسل عبارتست از:

$$L_{avg} = \sum_{k=0}^{L-1} l(r_k) p(r_k). \quad (4-1.6)$$

یعنی طول متوسط کلمات رمز متناسب شده به هر پیکسل تصویر با جمع حاصل ضرب های تعداد بیت بکار رفته در نمایش هر سطح خاکستری و احتمال رخداد آن سطح خاکستری پیدا می شود. بنابراین تعداد کل بیت های مورد نیاز برای رمز نمودن یک تصویر $N \times M$ ، MNL_{avg} بیت می باشد.

نمایش سطوح خاکستری یک تصویر با یک رمز دودویی m بینی طبیعی[‡]، طرف راست معادله ۴-۱.۶ را به m بیت کاهش می دهد. یعنی وقتی که m جایگزین (r_k) بشود، $L_{avg} = m$ می شود. حال می توان ثابت m را از مجموع خارج نمود و تنها مجموع $p(r_k)$ ها برای $1 - \sum_{k=0}^{L-1} p(r_k)$ باقی می ماند که البته برابر واحد است.

مثال: یک تصویر ۸ سطحی، توزیع سطوح خاکستری نشان داده شده در جدول ۶.۱ را دارد. اگر یک رمز دودویی ۳ بینی طبیعی [رمز شماره ۱ و (r_1)] از جدول ۶.۱ را بینید^۱ برای نمایش ۸ سطح خاکستری

^۱ یک رمز، سامانه ای از نمادهای (حروف، اعداد، بیتها و ...) به کار رفته در نمایش حجمی از اطلاعات یا مجموعه ای از رخدادها می باشد. به هر قسمت از اطلاعات یا رخداد، دنباله ای از نمادهای رمز که کلمه رمز خوانده می شود، متناسب می شود. تعداد نمادها در هر کلمه رمز طول آن کلمه است. یکی از معروف ترین رمزها در آوریل مال ۱۹۷۷ Paul Revere استفاده شد. عبارت "یک اگر با خشکی، دو اگر با دریا" غالباً برای توصیف آن رمز استفاده می شود که در آن یک یا دو لامپ برای نشان دادن این که آیا انگلیسی ها در خشکی یا در دریا مسافت می کردند، روشن می شدند.

[‡] یک رمز دودویی طبیعی (یاساده) رمزی است که در آن به هر رخداد یا قطعه اطلاعات که باید رمز شود (نظیر سطح خاکستری) یکی از ۲ کلمه رمز دودویی m بینی متناسب می شود.

ممکن استفاده شود، L_{avg} برابر ۳ بیت است، زیرا برای تمام r_i ها، $3 = l_i(r_i)$ می باشد. اگر از رمز شماره ۲ از جدول ۶.۱ استفاده کنیم، تعداد متوسط بیت های مورد نیاز برای رمزگذرن تصویر به

$$\begin{aligned} L_{avg} &= \sum_{k=0}^7 l_2(r_k) p_r(r_k) \\ &= 2(0.19) + 2(0.25) + 2(0.21) + 3(0.16) + 4(0.08) \\ &\quad + 5(0.06) + 6(0.03) + 6(0.02) \\ &= 2.7 \text{ bits.} \end{aligned}$$

کاهش می یابد. از معادله (۶.۱-۲) نسبت فشرده سازی حاصل، C_R ، برابر $3/2.7$ یا 1.11 می باشد. بنابراین تقریباً ده درصد داده های حاصل از کاربرد رمز شماره ۱ زاید است. میزان دقیق افزونگی را می توان از معادله (۶.۱-۱) تعیین کرد:

$$R_D = 1 - \frac{1}{1.11} = 0.099.$$

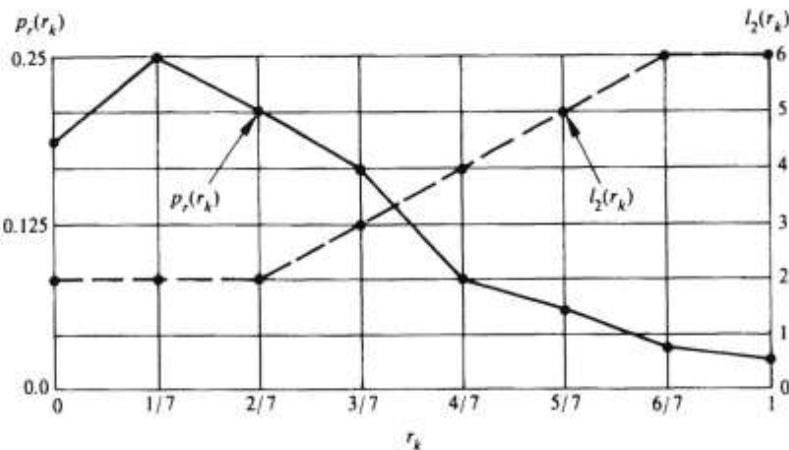
شکل ۶.۱ مبنای فشرده سازی بدست آمده با رمز شماره ۲ را نشان می دهد. این شکل هم بافت نگار تصویر [رسمی از r_i] بر حسب $p_r(r_i)$ و هم بافت نگار $[r_i]$ را نشان می دهد. چون این دوتابع تناسب معکوس با هم دارند، یعنی با کاهش (r_i) افزایش $p_r(r_i)$ اتفاق می یابد، کوتاه ترین کلمات رمز شماره ۲ به تکراری ترین سطوح خاکستری تصویر متنسب می شوند. □

در مثال قبلی انتساب بیت های کمتر به سطوح خاکستری که احتمال بیشتری دارند، باعث فشرده سازی داده ها شد. به این فرآیند، معمولاً با عنوان رمزنگاری طول متغیر^۱ اشاره می شود. اگر سطوح

جدول ۶.۶ مثالی از رمزنگاری طول متغیر

r_i	$p_r(r_i)$	Code 1	$l_1(r_i)$	Code 2	$l_2(r_i)$
$r_0 = 0$	0.19	000	3	11	2
$r_1 = 1/7$	0.25	001	3	01	2
$r_2 = 2/7$	0.21	010	3	10	2
$r_3 = 3/7$	0.16	011	3	001	3
$r_4 = 4/7$	0.08	100	3	0001	4
$r_5 = 5/7$	0.06	101	3	00001	5
$r_6 = 6/7$	0.03	110	3	000001	6
$r_7 = 1$	0.02	111	3	000000	6

1) Variable-length coding



شکل ۱۰۶ نمایش ترسیمی مبنای فشرده سازی داده ها با رمزگاری طول متغیر

خاکستری تصویر با روشی که بیش از نیاز از نمادهای رمز استفاده می کند، رمز شوند (یعنی رمز معادله $6 \cdot 4 - 1$ را حداقل نکند)، گفته می شود که رمز حاصل دارای افزونگی رمزگاری است. به طور کلی وقتی افزونگی رمزگاری وجود دارد که رمز های متناسب شده به یک مجموعه از رخدادها (نظیر مقادیر سطح خاکستری) طوری انتخاب نشده باشند که کاملاً از احتمالات رخدادها بهره بگیرند. وقتی سطوح خاکستری تصویر با یک رمز دودویی طبیعی یا ساده نمایش داده شوند، تقریباً همواره افزونگی رمزگاری وجود دارد. منشأ اصلی افزونگی رمزگاری این است که تصاویر نوعاً از اشیای تشکیل می شوند که شکل و انعکاس تقریباً قابل پیش بینی و منظمی دارند و معمولاً طوری نمونه برداری می شوند که اشیای درون تصویر بسیار بزرگتر از پیکسل ها (عناصر عکس) هستند. نتیجه طبیعی این است که در غالب تصاویر، سطوح خاکستری معینی، محتمل تر از بقیه هستند (یعنی بافت نگارهای اغلب تصاویر یکنواخت نیستند). در یک رمزگاری دودویی طبیعی تعداد بیت های یکسانی به محتمل ترین و نامحتمل ترین مقادیر متناسب می شود. بنابراین معادله $(6 \cdot 4 - 1)$ کمینه نمی شود و بنابراین افزونگی رمزگاری وجود دارد.

۲۰۱۶ افزونگی بین پیکسلی

تصاویر نشان داده شده در اشکال ۲۰.۶ (الف) و (ب) را در نظر بگیرید. همان طور که اشکال ۲۰.۶

(ب) و (ت) نشان می دهند، این تصاویر بافت نگارهای تقریباً یکسانی دارند. همچنین توجه شود که هر دو بافت نگار سه بعدی^۱ هستند، که نشان دهنده وجود سه محدوده غالب از مقادیر خاکستری می باشد. چون سطوح خاکستری این تصاویر هم احتمال نیستند، می توان از رمزنگاری طول متغیر برای کاهش افزونگی رمزنگاری حاصل از رمزگذاری دودویی طبیعی پیکسل های آنها استفاده کرد. در هر حال فرآیند رمزنگاری، میزان همبستگی بین پیکسل های تصویر را تغییر نمی دهد. به عبارت دیگر رمزهای استفاده شده برای نمایش سطوح خاکستری تصویر به همبستگی بین پیکسل ها کاری ندارند. این همبستگی ها از روابط ساختاری یا هندسی بین اشیای تصویر حاصل می شوند.

اشکال ۶ . ۲ (ث) و (ج) ضرایب خود همبستگی محاسبه شده برای یک خط از هر تصویر را نشان می دهند. این ضرایب با استفاده از یک نسخه تراز شده از معادله (۳ . ۳۷-۳) که در آن

$$\gamma(\Delta n) = \frac{A(\Delta n)}{A(0)} \quad (5-1 . 6)$$

و

$$A(\Delta n) = \frac{1}{N - \Delta n} \sum_{y=0}^{N-1-\Delta n} f(x, y) f(x, y + \Delta n). \quad (6-1 . 6)$$

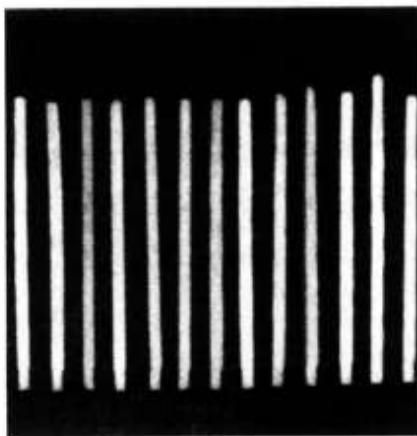
محاسبه شدند. ضریب مقیاس در معادله (۶ . ۱) تعداد متغیر عبارات جمع بازی مقادیر صحیح Δn را به حساب می آورد. البته Δn باید اکیداً کوچک تر از N ، تعداد پیکسل ها روی یک خط تصویر، باشد. متغیر x مختصه^۲ استفاده شده در محاسبه است. به تفاوت خیلی زیاد بین توابع اشکال ۶ . ۲ (ث) و (ج) توجه نمایید. از نظر کیفی می توان شکل این توابع را با تصاویر اشکال ۶ . ۲ (الف) و (ب) مرتبط نمود. این ارتباط بویژه در شکل ۶ . 2 (ج) قابل توجه است که در آن همبستگی بالای بین پیکسل های که ۴۵ و ۹۰ نمونه از هم فاصله دارند، را می توان به طور مستقیم به فاصله بین نوارهای عمودی شکل ۶ . 2 (ب) مرتبط دانست. بعلاوه پیکسل های مجاور در هر دو تصویر به شدت همبسته اند. وقتی $\Delta n = 1$ باشد، ۷ برای اشکال ۶ . ۲ (الف) و (ب) بترتیب برابر ۰.۹۹۲۲ و ۰.۹۹۲۸ است. این مقادیر نماینده اغلب تصاویر تلویزیونی که به بهترین وجه نمونه برداری شده اند، می باشند.

این تصاویر شکل مهم دیگری از افزونگی داده را که به طور مستقیم با همبستگی های بین پیکسلی

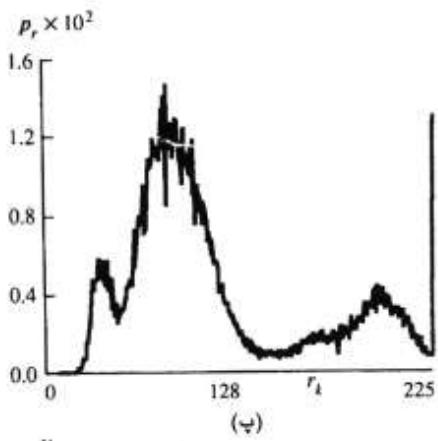
۱.۶ میانی ۳۸۱



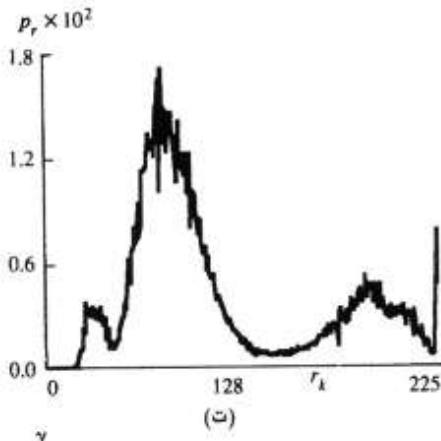
(الف)



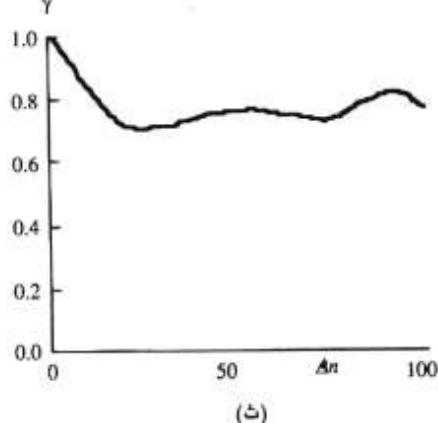
(ب)



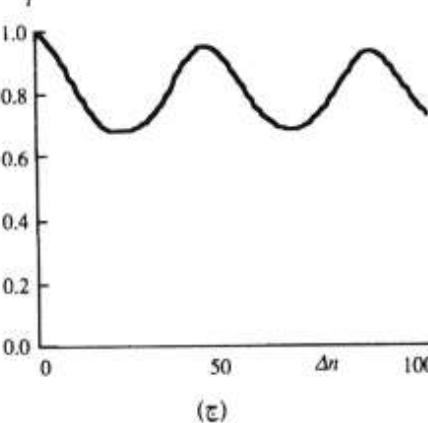
(د)



(e)



(f)



(g)

شکل ۲۰.۶ دو تصویر و بافت نگارهای سطح خاکستری و ضرایب خود همبستگی تراز شده آنها در طول یک خط افقی تصویر

درون تصویر مرتبط است، نشان می دهد. چون مقدار هر پیکسل را می توان بخوبی از مقدار همسایگانش پیش بینی کرد، مقدار اطلاعاتی که بوسیله پیکسل های منفرد حمل می شود، نسبتاً کم است. مقدار زیادی از اطلاعات بصری یک پیکسل در تصویر زاید است؛ زیرا آن را می توان براساس مقادیر همسایه اش حدس زد. اسمی گوناگونی شامل افزونگی مکانی^۱، افزونگی هندسی^۲ و افزونگی بین قابی^۳ به این وابستگی های بین پیکسلی اشاره می کند. ما از عبارت افزونگی بین پیکسلی^۴ برای اشاره به همه آنها استفاده می کنیم.

به منظور کاهش افزونگی های بین پیکسلی در یک تصویر، آرایه پیکسلی دو بعدی که معمولاً برای نظاره و تعییر بشری استفاده می شود، باید به یک قالب پربازده تر (گرچه "غیر بصری") تبدیل شود. برای مثال می توان از تفاضل های بین پیکسل های مجاور برای نمایش تصویر استفاده کرد. تبدیل هایی از این نوع (یعنی آنهایی که افزونگی بین پیکسلی را حذف می کنند) با عنوان نگاشت^۵ اشاره می شوند. اگر بتوان از داده های تبدیل شده پیکسل های اولیه را بازسازی کرد، نگاشت را معکوس پنیر^۶ می خوانند.

مثال: شکل ۶.۳. یک روال نگاشت ساده را نشان می دهد. شکل ۶.۳.(الف) یک بخش 1×3 اینچ مریع از نقشه یک مدار الکترونیکی است که با نرخی حدود ۳۳۰dpi^۷ نمونه برداری شده است. شکل ۶.۳.(ب) یک نسخه دودویی از این رسم را نشان می دهد. شکل ۶.۳.(پ) نمایش سطح خاکستری یک خط تصویر و آستانه استفاده شده برای به دست آوردن این نسخه دودویی را نشان می دهد. چون این تصویر دودویی حاوی نواحی با شدت ثابت زیادی می باشد، می توان با نگاشت پیکسل های موجود در طول هر خط پیمایش یعنی $f(x, 0), f(x, 1), \dots, f(x, N-1)$ به دنباله زوجهای $(g_1, r_1), (g_2, r_2), \dots, (g_i, r_i)$ ، نمایش پربازده تری را بدست آورد که در آن g_i دلالت بر r_i می باشد. این سطح خاکستری در طول خط و r_i دلالت بر طول دنباله (RL)^۸ دارد. به عبارت دیگر بوسیله مقدار و طول دنباله های با سطح خاکستری ثابت می توان تصویر آستانه گیری شده را به طور پربازده تری نسبت به آرایه دو بعدی پیکسل های دودویی نمایش (غیر بصری) داد.

شکل ۶.۳(ت) داده های رمز طول دنباله ای (RL) متناظر با یک خط افقی آستانه گیری شده از شکل

1) Spatial redundancy

2) Geometric redundancy

3) Interframe redundancy

4) Interpixel redundancy

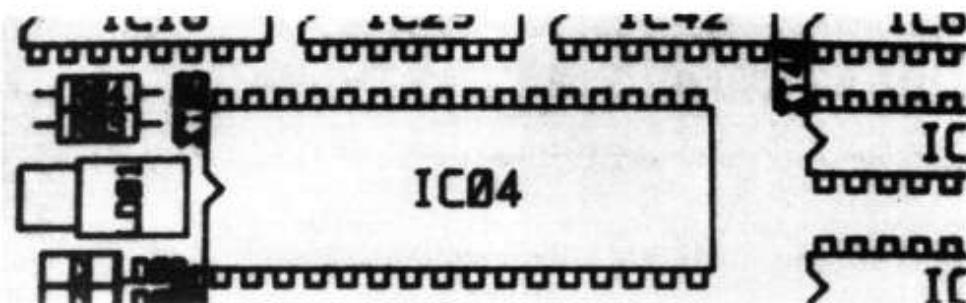
5) Mapping

6) Reversible

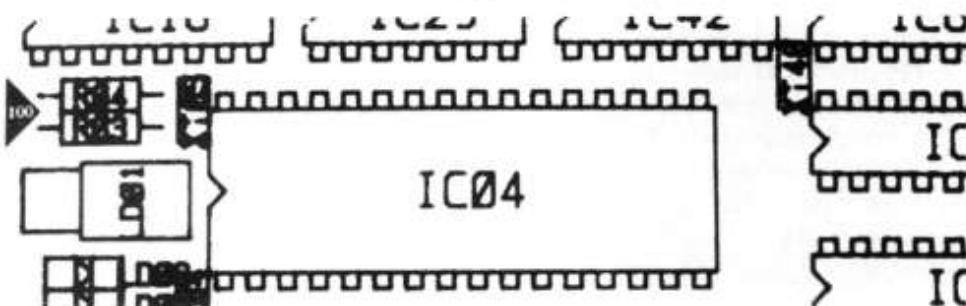
7) Dot per inch = نقطه در اینچ

8) Run length

۱۰۶ مبانی ۳۸۳

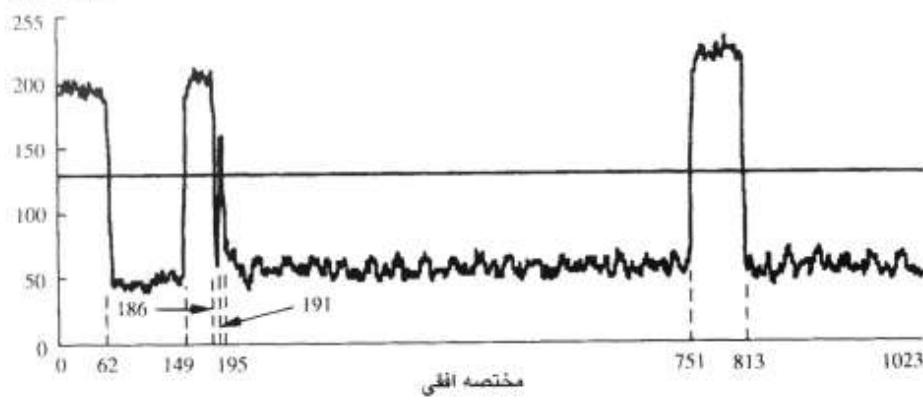


(الف)



(ب)

سطح خاکستری



(پ)

خط شماره ۱۰۰ : (۱, 63) (0, 87) (1, 37) (0, 5) (1, 4) (0, 556) (1, 62) (0, 210)

(ت)

شکل ۱۰۶ نمایش رمزگاری RL: (الف) تصویر اولیه؛ (ب) تصویر دودویی که خط افقی با شماره 100 در آن مشخص شده است؛ (پ) نمایش یک خط و آستانه دودویی سازی آن؛ (ت) رمز RL حاصل

۶. ۳. (ب) رانشان می دهد. برای نمایش 1024 بیت از داده های دودویی تنها 88 بیت مورد نیاز است. در واقع می توان تمام بخش 343×1024 پیکسلی شکل ۶. ۳. (ب) را به 12166 دنباله کاهش داد. چون برای نمایش هر زوج طول دنباله ۱۱ بیت نیاز است، نسبت فشرده سازی حاصل و افزونگی نسبی متناظر عبارتند از:

$$C_R = \frac{(1024)(343)(1)}{(12166)(11)} = 2.63$$

و

$$R_D = 1 - \frac{1}{2.63} = 0.62.$$

□

۳.۱.۶ افزونگی روان بصری

در فصل ۲ بیان کردیم که روشنایی یک ناحیه، آن طور که چشم دریافت می کند، به عواملی غیر از نور منعکس شده بوسیله آن ناحیه نیز بستگی دارد. به عنوان مثال می توان در یک ناحیه باشد نور ثابت، تغییرات شدت (نوارهای مک^۱) را احساس کرد. چنین پدیده هایی از این واقعیت نتیجه می شوند که چشم به تمام اطلاعات دیداری با حساسیت یکسانی پاسخ نمی دهد. بروشنا در پردازش بصری معمولی، اطلاعات معینی نسبت به سایر اطلاعات کمتر اهمیت دارند. به این اطلاعات افزونگی روان بصری^۲ گفته می شود و می توان بدون کاهش زیاد در کیفیت تصویر آنها را حذف نمود.

وجود افزونگی روان بصری نباید مایه تعجب باشد، زیرا معمولاً نحوه دریافت اطلاعات تصویر توسط انسان شامل تحلیل کمی هر پیکسل یا مقدار لومینانس تصویر نمی باشد. به طور کلی چشم مشاهده گر، به دنبال تشخیص ویژگی های نظری لبه ها یا نواحی بافتی می باشد و به طور ذهنی آنها را در گروه های قابل جداسازی ترکیب می کند. سپس مغز به منظور تکمیل فرآیند بینایی، این گروه هارا با دانش های قبلی مرتبط می کند.

افزونگی روان بصری به طور اساسی با افزونگی های قبل^۳ تشریح شده، تفاوت دارد. برخلاف افزونگی های بین پیکسلی و رمز نگاری، افزونگی روان بصری، مربوط به اطلاعات بصری حقیقی یا کمیت پذیر^۳ می باشد. تنها به این دلیل که معمولاً خود اطلاعات برای پردازش بصری الزامی نیست،

1) Mach bands

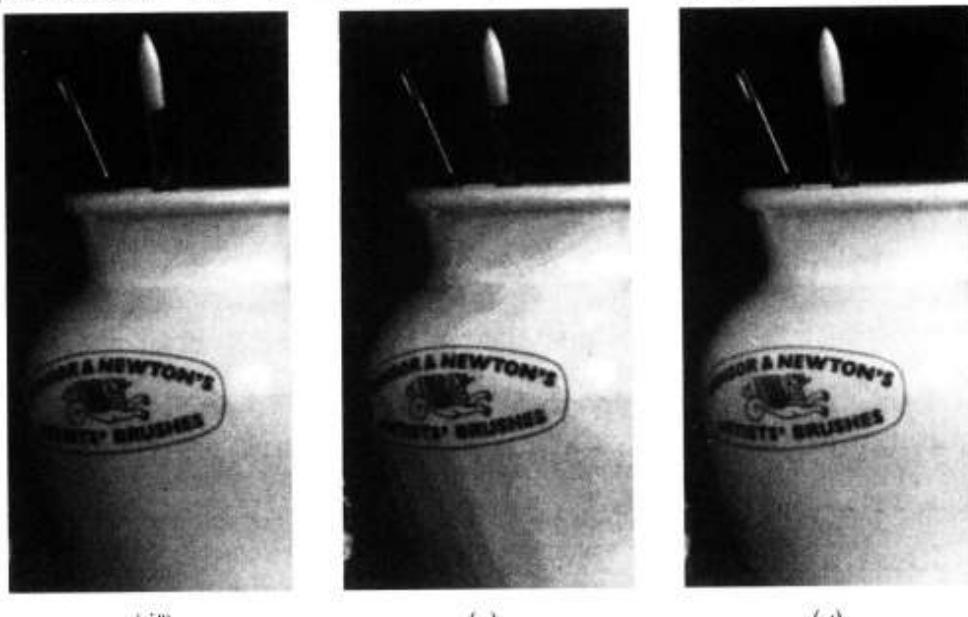
2) Psychovisually redundant

3) Quantifiable

حذف آن ممکن است. از آنجا که حذف داده های افزونگی روان بصری باعث از دست دادن اطلاعات کمی می شود، عموماً با عنوان چندی سازی^۱ اشاره می شود. این اصطلاح با کاربرد معمولی این کلمه، که معمولاً به معنی نگاشت محدوده وسیعی از مقادیر ورودی به تعداد محدودی مقادیر خروجی است، سازگار می باشد. از آنجا که چندی سازی واقعی معکوس ناپذیر است (اطلاعات بصری تلف می شود)، فشرده سازی با اتلاف داده ها را نتیجه می دهد.

مثال: تصاویر شکل ۴.۶ را در نظر بگیرید. شکل ۶.۴ (الف) یک تصویر تک رنگ دارای ۲۵۶ سطح خاکستری ممکن را نشان می دهد. شکل ۶.۴ (ب) همان تصویر را پس از چندی سازی یکنواخت به چهار بیت یا ۱۶ سطح ممکن نشان می دهد. نسبت فشرده سازی حاصل ۲ به ۱ می باشد. همان طور که در بخش ۲.۳ تشریح شد، توجه شود که در نواحی قبالاً آرام^۲ تصویر اولیه، منحنی های^۳ اشتباه وجود دارد که نتیجه طبیعی حاصل از این نمایش سطوح خاکستری تصویر است.

شکل ۶.۴ (پ) بهبود قابل ملاحظه ای را که با استفاده از نوعی چندی سازی سازگار با عجایب سامانه بینایی انسان ممکن گردیده است، نشان می دهد. گرچه نسبت فشرده سازی حاصل از روای دوم



شکل ۴.۶ (الف) تصویر اولیه؛ (ب) حاصل چندی سازی یکنواخت به ۱۶ سطح؛ (پ) حاصل چندی سازی به ۱۶ سطح IGS

حدول ۲۰۶ روال چندی سازی IGS

پیکسل	سطح خاکستری	مجموع	رمز IGS
$i - 1$	N/A	0000 0000	N/A
i	0110 1100	0110 1100	0110
$i + 1$	1000 1011	1001 0111	1001
$i + 2$	1000 0111	1000 1110	1000
$i + 3$	1111 0100	1111 0100	1111

چندی سازی نیز ۲ به ۱ است، منحنی های اشتباه به مقدار زیادی بازای افزایش دانه ای شدن که کمتر آزار دهنده است، کاهش می یابد. روش استفاده شده برای تولید این نتیجه به عنوان چندی سازی بهبود یافته سطح خاکستری^۱ (IGS) شناخته می شود. این روش حساسیت ذاتی چشم به لبه ها را تشخیص می دهد و قبل از چندی سازی با افزودن یک عدد شبه تصادفی به هر پیکسل که این عدد از بیت های پایین مرتبه پیکسل های مجاورش تولید می شود، لبه ها را می شکند. چون بیت های پایین مرتبه نسبتاً تصادفی هستند (صفحات بیتی بخش ۴ . ۱ را ببینید)، این روش معادل اضافه کردن میزانی از تصادفی بودن به لبه های ساختگی مربوط به منحنی های اشتباه می باشد، که میزان تصادفی بودن به مشخصات محلی تصویر بستگی دارد.

چندی سازی سطح خاکستری بهبود یافته نماینده یک گروه بزرگ از روالهای چندی سازی که به طور مستقیم روی سطوح خاکستری تصویر تحت فشرده سازی کار می کنند، می باشد. این روال ها عموماً کاهشی در تغییک مکانی، سطح خاکستری یا زمانی تصویر بوجود می آورند. منحنی های اشتباه یا سایر اثرات مرتبط (حرکت لرزشی، مات شدن و غیره) استفاده از روش های مکاشفه ای^۲ برای جراثم اثر بصری چندی سازی را الزامی می نمایند. برای مثال روش درهم بافته خطی^۳ به ۱ معمولی مورد استفاده در پخش تلویزیون تجاری شکلی از چندی سازی است که در آن بخش های درهم بافته قاب های متواالی با ایجاد کم کاهش در گفت تصویر دریافت شده، ترخ های پیماش ویدیویی کاهش یافته را امکان داده اند.

1) Improved gray-scale quantization

2) Heuristic

3) Line interlacing approach

۴.۱.۶ معیار مشابهت

همان طور که قبلاً ذکر شد، حذف داده های افزونگی روان بصری باعث از دست رفتن اطلاعات بصری حقیقی یا کمی می شود. چون ممکن است اطلاعات مورد علاقه از دست برود، یک وسیله تکرار پذیر با تجدید پذیر برای کمی سازی طبیعت و گستره' اتلاف اطلاعات بسیار مطلوب است. دو گروه کلی از معیارها به عنوان پایه ای برای چنین ارزیابی هایی استفاده می شوند: (۱) معیار عینی مشابهت و (۲) معیار ذهنی مشابهت.

وقتی میزان اتلاف اطلاعات به صورت تابعی از تصویر اولیه ورودی و تصویر خروجی فشرده شده و در نهایت وافشده شده بیان شود، گفته می شود که براساس معیار عینی مشابهت^۱ می باشد. یک مثال خوب، خطای ریشه میانگین مربع (rms) بین تصویر ورودی و خروجی است. اجازه دهید $(y, f(x, y))$ یک تصویر ورودی را نمایش دهد و اجازه دهید $(\hat{f}(x, y), \hat{f}(x, y))$ بیانگر تخمین یا تقریبی از $(y, f(x, y))$ باشد که از فشرده سازی و به دنبال آن وافشده سازی تصویر ورودی حاصل می شود. به ازای هر مقدار x و y ، خطای بین $(y, f(x, y))$ و $(\hat{f}(x, y), \hat{f}(x, y))$ را می توان به صورت

$$e(x, y) = \hat{f}(x, y) - f(x, y) \quad (7-1.6)$$

تعریف کرد طوری که خطای کل بین دو تصویر با ابعاد $N \times M$ به صورت زیر می باشد:

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]$$

حال e_{rms} ، خطای ریشه میانگین مربع بین $(y, f(x, y))$ و $(\hat{f}(x, y), \hat{f}(x, y))$ ، برابر ریشه دوم متوسط مجدد رات خطای روی آرایه $M \times N$ می باشد:

$$e_{\text{rms}} = \left[\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2 \right]^{1/2} \quad (8-1.6)$$

یک معیار مشابهت عینی کاملاً مرتبط، نسبت سیگنال به نویز میانگین مربعی تصویر فشرده - وافشده شده می باشد. اگر $(y, f(x, y))$ (پس از بازارایی ساده عبارات معادله (۶-۱.۷)) برابر جمع تصویر اصلی $(y, f(x, y))$ و سیگنال نویز $(y, e(x, y))$ در نظر گرفته شود، نسبت سیگنال به نویز میانگین مربعی تصویر

1) Objective fidelity criterion

خروجی که با SNR_{rms} بیان می‌شود، عبارتست از:

$$SNR_{rms} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x, y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2}, \quad (9-1.6)$$

مقدار rms نسبت سیگنال به نویز که با SNR_{rms} نشان داده می‌شود، با گرفتن ریشه دوم معادله (9-1.6) بدست می‌آید.

گرچه معیار عینی مشابهت روشی ساده و مناسب برای ارزیابی اتلاف اطلاعات عرضه می‌کند، تهایتاً اغلب تصاویر وافشده شده بوسیله انسانها نظاره می‌شوند. نتیجتاً اغلب مناسبتر است که کیفیت تصویر بوسیله ارزیابی های ذهنی بینندگان انسانی اندازه گیری شود و این کار را می‌توان با نشان دادن یک تصویر وافشده نوعی به گروه مناسبی از بینندگان و متوسط گیری ارزیابی های آنها انجام داد. ارزیابی ها را می‌توان با استفاده از یک معیار درجه بندی مطلق^۱ یا بوسیله مقایسه های پهلو به پهلوی^۲ (y) و ($f(x, y)$) انجام داد. جدول ۶.۳ یک نوع معیار درجه بندی مطلق را نشان می‌دهد. مقایسه های پهلو به پهلو را می‌توان با درجه هایی نظری^۳ {-3, -2, -1, 0, 1, 2, 3} که بترتیب نشانگر ارزیابی های ذهنی (خیلی بهتر، بهتر، کمی بهتر، همانند، کمی بدتر، بدتر، خیلی بدتر) هستند، انجام داد. در هریک از دو حالت گفته می‌شود که ارزیابی ها بر اساس معیار ذهنی مشابهت^۴ هستند.

جدول ۶.۳ معیار درجه بندی^۵ تصاویر تلویزیونی (از Behrend[1960] و Frendendall[1960])

مقدار	درجه	شرح
۱	عالی	یک تصویر با کیفیت بسیار خوب به همان خوبی که شما آرزو می‌کنید.
۲	خوب	یک تصویر با کیفیت خوب که منظره لذت‌بخشی را راه می‌کند. تداخل آزاردهنده ^۶ نیست.
۳	قابل قبول ^۷	یک تصویر با کیفیت قابل قبول. تداخل آزاردهنده نیست.
۴	مرزی ^۷	یک تصویر با کیفیت بد؛ آرزو می‌کنید که می‌توانست آن را بپرورد دهید. تداخل تقریباً آزاردهنده است.
۵	ضعیف	یک تصویر با کیفیت بسیار بد؛ امامی توانید آن را تماشا کند. فقط آن تداخل آزاردهنده است.
۶	غیرقابل استفاده	تصویری آنقدر بد که نمی‌توان آن را تماشا کرد.

1) Absolute rating scale

2) Side- by- side

3) Subjective fidelity criteria

4) Rating scale

5) Objectionable

6) Passable

7) Marginal

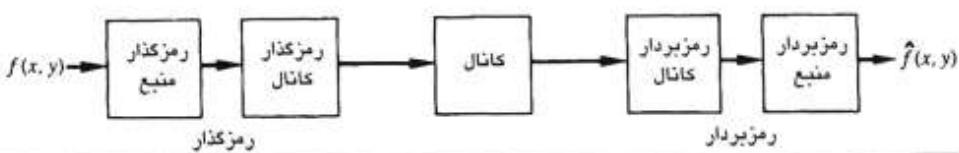
۲.۶ مدل های فشرده سازی تصویر

مثال: خطاهای rms در تصاویر چندی شده شکل ۴.۶ (ب) و (پ) بترتیب ۰.۹۳ و ۰.۷۸ و سطح خاکستری هستند. نسبت های سیگنال به نویز rms متناظر ۰.۲۵ و ۰.۳۹ هستند. گرچه این مقادیر کاملاً نزدیک هستند، ارزیابی ذهنی کیفیت دیداری دو تصویر رمز شده ممکن است برای تصویر ۴.۶ (ب) درجه مرزی ۱ و برای تصویر ۴.۶ (پ) درجه قابل قبول^۱ را نتیجه بدهد.

۲.۶ مدل های فشرده سازی تصویر

در بخش ۶.۱ به طور جداگانه سه روش کلی برای کاهش یا فشرده سازی مقدار داده های مورد نیاز برای نمایش تصویر را شرح دادیم. سامانه های عملی فشرده سازی تصویر معمولاً از ترکیب این روش ها بوجود می آیند. در این بخش مشخصات کلی چنین سامانه ای را بررسی می کنیم و یک الگوی کلی برای نمایش آن ارائه می نماییم.

همان طور که شکل ۶.۵ نشان می دهد یک سامانه فشرده سازی تصویر دارای دو بلوک ساختاری جداگانه است: یک رمزگذار^۲ و یک رمزبردار^۳. تصویر ورودی $(y, f(x))$ به درون رمزگذار که مجموعه ای از نمادهای رمز را از داده های ورودی تولید می کند، وارد می شود. پس از ارسال روی کانال^۴، خروجی کانال وارد رمزبردار شده و در آنجا تصویر $(y, \hat{f}(x))$ بازسازی می شود. در حالت کلی ممکن است $(y, \hat{f}(x))$ نسخه دوم دقیقی از $(y, f(x))$ باشد یا نباشد. در حالت مثبت، سامانه بدون خطای نگهدار اطلاعات است؛ اگر چنین نباشد مقداری اعوجاج در تصویر بازسازی شده وجود دارد. رمزگذار و رمزبردار شکل ۶.۵ هر دو دارای دوتایی یا زیربلوک نسبتاً مستقل هستند. رمزگذار از رمزگذار منبع^۵ که افزونگی های ورودی را حذف می کند و رمزگذار کانال^۶ که امنیت نویز^۷ خروجی



شکل ۶.۵ یک مدل کلی از سامانه فشرده سازی

^۱ منطقی است که انتظار داشته باشیم این بلوکها به ترتیب "فشرده ساز" و "وافرده" نامیده شوند. عبارات رمزگذار و رمزبردار تأثیر نظریه اطلاعات (که در بخش ۶.۳ تشریح می شود) بر مبحث فشرده سازی تصویر را نشان می دهند.

1) Marginal rating

2) Passable rating

3) encoder

4) decoder

5) Channel

6) Source encoder

7) Channel encoder

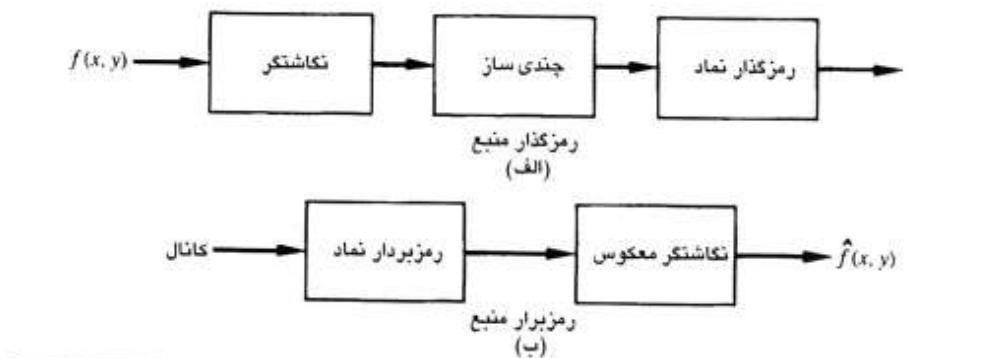
8) Noise immunity

رمزگذار منبع را افزایش می دهد، تشكیل می شود. همان طور که انتظار می رود رمزبردار نیز شامل یک رمزبردار کانال^۱ و به دنبال آن یک رمزگذار منبع^۲ است. اگر کانال بین رمزگذار و رمزبردار بدون نویز باشد (در معرض خطای نباشد)، رمزگذار و رمزبردار کانال حذف می شوند، و رمزگذار و رمزبردار کلی پتریب معادل رمزگذار و رمزبردار منبع می شوند.

۱۰.۶ رمزگذار و رمزبردار منبع

رمزگذار منبع مسئول کاهش یا حذف هر یک از افزونگی های رمزنگاری، بین پیکسلی یا روان بصری در تصویر ورودی است. کاربرد موردنظر و میزان کیفیت مورد نیاز تصویر بهترین روش رمزگذاری را تعیین می کند. معمولاً این فرآیندر امی توان با دنباله ای از سه عمل مستقل مدل نمود. همان طور که شکل ۶.۶ (الف) نشان می دهد، هریک از این اعمال برای کاهش یکی از سه نوع افزونگی مشروح در بخش ۱۰.۶ طراحی می شوند. شکل ۶.۶ (ب) رمزبردار منبع متناظر آن را نشان می دهد.

در مرحله اول فرآیند رمزگذاری منبع، نگاشت گر^۳، داده های ورودی را به یک قالب (ممولاً غیر بصری) که برای کاهش افزونگی های بین پیکسلی تصویر ورودی طراحی شده است، تبدیل می کند. این عمل در حالت کلی معکوس پذیر است و ممکن است مقدار داده موردنیاز برای نمایش تصویر را کاهش بدهد یا ندهد. رمزنگاری طول دنباله (RL)^۴ (بخش های ۱.۶ و ۲.۲.۴.۲.۲) مثالی از یک نگاشت است که در مرحله اول فرآیند کلی رمزگذاری منبع، باعث فشرده سازی داده ها می شود. نمایش



شکل ۱۰.۶ (الف) مدل رمزگذار منبع و (ب) مدل رمزبردار منبع

۴۹۱ ۲۰۶ مدل های فشرده سازی تصویر

تصویر بوسیله مجموعه ای از ضرایب تبدیل (بخش ۶.۵.۲) مثالی از حالت مخالف است. در این حالت نگاشت گر، تصویر را به آرایه ای از ضرایب تبدیل می نماید تا افزونگی های بین پیکسلی آن برای فشرده سازی در مراحل بعدی رمزگذاری قابل دسترس تر شوند.

مرحله دوم، یا بلوک چندی ساز^۱ در شکل ۶.۶(الف)، طبق یک معیار مشابه از پیش تعیین شده، دقت خروجی نگاشت گر را کاهش می دهد. در این مرحله افزونگی های روان بصری تصویر ورودی کاهش می یابند. همان طور که در بخش ۶.۳ ذکر شد، این عمل معکوس ناپذیر است. بنابراین در فشرده سازی بدون خطأ باید وجود داشته باشد.

در مرحله سوم و نهایی فرآیند رمزگذاری منبع، رمزگذار نماد^۲ برای نمایش خروجی چندی ساز، یک رمز طول ثابت یا طول متغیر ایجاد می کند و خروجی را طبق آن نگاشت می نماید. اصطلاح رمزگذار نماد، این عمل را از فرآیند کلی رمزگذاری منبع مشخص می سازد. در بیشتر حالات از یک رمز طول متغیر برای نمایش مجموعه داده های نگاشت و چندی شده استفاده می شود. رمزگذار نماد کوتاه ترین کلمات رمز را به تکراری ترین مقادیر خروجی متناسب می کند و بنابراین افزونگی رمزگاری را کاهش می دهد. البته این عمل معکوس پذیر است. با اتمام مرحله رمزگذاری نماد هر یک از سه افزونگی مشروح در بخش ۶.۱ در تصویر ورودی کاهش یافته یا حذف شده اند.

شکل ۶.۶(الف) فرآیند رمزگذاری منبع را به صورت سه مرحله متوالی نشان می دهد که همیشه هر سه مرحله در سامانه فشرده سازی به کار نمی روند. مثلاً به یاد آورید که وقتی فشرده سازی بدون خطأ مطلوب باشد، باید چندی ساز حذف شود. بعلاوه بعضی روش های فشرده سازی با ادغام بلوک هایی از شکل ۶.۶(الف) که ظاهر آجدا هستند، مدل می شوند. برای نمونه در سامانه های فشرده ساز پیشگویی در بخش ۶.۱.۵، نگاشت گر و چندی ساز اغلب با یک بلوک تکی که همزمان هر دو عمل را انجام می دهد، نمایش داده می شوند.

رمزبردار منبع شکل ۶.۶(ب) تنها دارای دو جزء است: رمزبردار نماد و نگاشت گر معکوس^۳. این دو بلوک به ترتیب عکس اعمال دو بلوک رمزگذار نماد و نگاشت گر از رمزگذار منبع را انجام می دهند. چون چندی سازی باعث اتفاق معکوس ناپذیر اطلاعات می شود، بلوک معکوس چندی ساز در مدل کلی رمزبردار منبع شکل ۶.۶(ب) وجود ندارد.

۲۰.۲۰.۶ رمزگذار و رمزبردار کانال

وقتی کانال شکل ۶.۵ نویزدار یا در معرض خطاباشد، رمزگذار و رمزبردار کانال نقش مهمی در فرآیند کلی رمزگذاری-رمزبرداری بازی می کنند. این دوبلوک برای کاهش اثر نویز کانال بوسیله افزودن یک شکل کنترل شده از افزونگی به داده های رمزشده منبع، طراحی می شوند. گرچه خروجی رمزگذار منبع مقدار کمی افزونگی دارد، اما بدون اضافه شدن "افزونگی کنترل شده"^۱ فوق، این خروجی نسبت به نویز انتقال بسیار حساس خواهد بود.

یکی از مفیدترین روش های رمزگذاری کانال بوسیله همینگ^۲ (۱۹۵۰) طراحی شد. این روش بر اساس افزودن بیت های کافی به داده هایی که رمز می شوند، می باشد طوری که اطمینان می دهد به میزان حداقل لازم، بین کلمات رمز معتبر تفاوت (فاصله^۳) وجود دارد. برای مثال همینگ نشان داد که اگر ۳ بیت افزونگی به یک کلمه ۴ بیتی اضافه شود طوری که فاصله^۴ بین هر دو کلمه رمز معتبر ۳ باشد، تمام خطاهای یک بیتی را می توان کشف و تصحیح کرد. (با افزودن بیت های افزونگی بیشتر می توان خطاهای چند بیتی را کشف و/یا تصحیح کرد). کلمه رمز ۷ بیتی همینگ (۴۰۷، ۷، $b_1 b_2 b_3 b_4 b_5 b_6 b_7$)، مربوط به عدد دودویی ۴ بیتی عبارتست از

$$\begin{array}{ll} h_1 = b_3 \oplus b_2 \oplus b_0 & h_1 = b_1 \\ h_2 = b_3 \oplus b_1 \oplus b_0 & h_3 = b_2 \\ h_4 = b_2 \oplus b_1 \oplus b_0 & h_5 = b_1 \\ & h_6 = b_0 \\ & h_7 = b_0 \end{array} \quad (1-2.6)$$

که \oplus نشانه عمل OR انحصاری است. توجه شود که بیت های h_1 , h_2 , h_3 و h_7 بترتیب بیت های توازن زوج برای میدان های بیتی $b_1 b_2 b_3 b_4 b_5 b_6$ هستند.

برای رمزبرداری از یک کلمه رمز همینگ، رمزبردار کانال باید مقدار رمزشده را برای برقراری توازن فرد روی میدان های بیتی که قبلاً توازن زوج داشته اند، وارسی نماید. یک بیت خطابا کلمه توازن غیر صفر

$$\begin{array}{l} c_1 = h_1 \oplus h_3 \oplus h_5 \oplus h_7, \\ c_2 = h_2 \oplus h_3 \oplus h_6 \oplus h_7, \\ c_4 = h_4 \oplus h_5 \oplus h_6 \oplus h_7. \end{array} \quad (2-2.6)$$

^۱ فاصله بین دو کلمه رمز دودویی به عنوان تعداد ارقامی که باید در یک کلمه تغییر نمایند تا کلمه دیگر بدست آید، تعریف می شود. مثلاً فاصله بین ۱۰۱۱۰۱ و ۰۱۱۱۰۱، ۲ است. فاصله کمینه یک رمز، کمینه فاصله بین زوج کلمات متفاوت رمز است.

۱) R. W. Hamming

2) Distance

۳.۶ عناصر نظریه اطلاعات

اگر یک مقدار غیر صفر پیدا شود، کافی است رمزبردار، تنها مکان بیتی نشان داده شده بوسیله کلمه توازن در کلمه رمز را متمم نماید. آنگاه مقدار دودویی رمزبرداری شده به صورت $h_3 h_5 h_6 h_1$ از کلمه رمز تصحیح شده استخراج می شود.

مثال: ارسال داده های IGS^۴ بیتی جدول ۶.۲ روی یک کانال مخابراتی نویزدار را در نظر بگیرید. یک خطای تک بیتی می تواند باعث شود که پیکسل واپشده شده به اندازه ۱۲۸ سطح خاکستری از مقدار صحیحش منحرف شود.^۵ یک رمزگذار کانال همینگی با اضافه کردن افزونگی کافی که کشف و تصحیح خطاهای تک بیتی را ممکن می سازند، می تواند امنیت نویز داده های IGS رمز شده این منبع را افزایش دهد. از معادله^۶ (۱-۲.۶)، خروجی رمزگذار همینگ برای اولین مقدار IGS جدول ۶.۲ برابر ۱۱۰۰۱۱۰ است. چون رمزگذار کانال همینگی تعداد بیت های مورد نیاز برای نمایش این مقدار IGS را از ۴ به ۷ افزایش می دهد، نسبت فشرده سازی مذکور در مثال IGS از ۲ به ۷/۸ یا ۱.۱۴ کاهش می یابد.

□

این کاهش در فشرده سازی هزینه افزایش امنیت نویز می باشد.

۳.۶ عناصر نظریه اطلاعات

در بخش ۶.۱ چند راه برای کاهش مقدار داده مورد نیاز برای نمایش تصویر معرفی کردیم. این پرسش به طور طبیعی مطرح می شود که: عملأً چه مقدار داده برای نمایش تصویر مورد نیاز است؟ یعنی آیا یک مقدار کمینه داده وجود دارد که برای توصیف کامل تصویر بدون از دست رفتن اطلاعات کافی باشد؟ نظریه اطلاعات چهار چوب ریاضی برای پاسخ به این سوال و سوالات مرتبط را فراهم می آورد.

۱.۳.۶ اندازه گیری اطلاعات

فرض اساسی نظریه اطلاعات این است که می توان تولید اطلاعات را به صورت فرآیندی احتمالی مدل کرد که با روشی موافق با شهود انسان اندازه گیری می شود. طبق این فرض پیشامد تصادفی E که با احتمال $P(E)$ رخ می دهد، دارای

$$I(E) = \log \frac{1}{P(E)} = -\log P(E) \quad (1-3.6)$$

^۴ یک روال ساده برای واپشده سازی داده های IGS بیتی ضرب معادل دهدۀ مقدار IGS در ۱۶ است. برای مثال اگر مقدار IGS برابر ۱۱۰۰۱۱۰ باشد، سطح خاکستری واپشده شده (۱۶)(۱۴)(۱۲) یا ۲۲۴ است. اگر با ارزش ترین این مقدار IGS به اشتباه ۰ ارسال شود، سطح خاکستری واپشده شده ۹۶ می شود. خطای حاصل ۱۲۸ سطح خاکستری می باشد.

واحد اطلاعات است. اغلب کمیت $I(E)$ اطلاعات E خوانده می شود. به طور کلی مقدار اطلاعات مربوط به پیشامد E رابطه معکوسی با احتمال E دارد. اگر $P(E) = 1$ باشد (یعنی پیشامد همواره رخ بدهد)، $I(E) = 0$ بوده و هیچ اطلاعاتی به آن پیشامد منسب نمی شود. یعنی چون هیچ عدم قطعیتی به پیشامد منسب نمی شود، هیچ اطلاعاتی با مخابره این که پیشامد رخ داده است، انتقال نمی یابد. در هر حال اگر $P(E) = 0.99$ باشد، مخابره اینکه E رخ داده است، مقدار کمی اطلاعات را منتقل می کند. مخابره اینکه E رخ نداده است، اطلاعات بیشتری را منتقل می کند زیرا این نتیجه کمتر محتمل است. پایه لگاریتم در معادله (۶.۱-۳) واحد مورد استفاده برای اندازه گیری اطلاعات را تعیین می کند.^۴ اگر لگاریتم با پایه ۲ استفاده شود، گفته می شود که اندازه گیری در واحد های ۲-گانی است. اگر پایه ۲ انتخاب شود، واحد اطلاعات بیت خوانده می شود. توجه شود که اگر $P(E) = \frac{1}{2}$ باشد، $I(E) = -\log_2 \frac{1}{2}$ یا یک بیت خواهد بود. یعنی وقتی یکی از دو پیشامد با احتمال برابر رخ می دهد، مقدار اطلاعات منتقل شده یک بیت است. یک مثال ساده از چنین وضعیتی انداختن یک سکه و مخابره نتیجه است.

۲.۳.۶ کanal اطلاعات

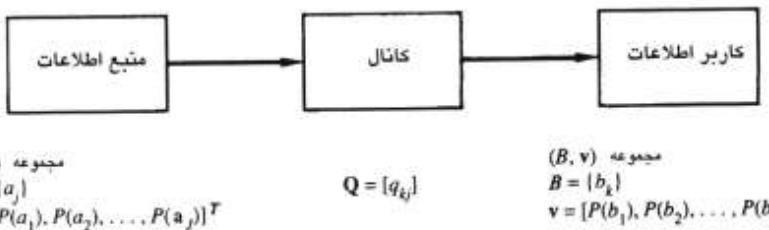
وقتی اطلاعات بین منبع اطلاعات و کاربر اطلاعات انتقال یابد، گفته می شود که منبع اطلاعات بوسیله یک کanal اطلاعات به کاربر اطلاعات متصل است. کanal اطلاعات رسانه فیزیکی است که منبع را به کاربر پیوند می دهد. رسانه فیزیکی ممکن است یک خط تلفن، یک مسیر انتشار انرژی الکترومغناطیسی یا یک سیم در یک رایانه رقمی باشد. شکل ۶.۷. مدل ریاضی ساده ای را برای سامانه اطلاعات گسته نشان می دهد. در اینجا عامل مورد توجه و وزیر ظرفیت سامانه است که به عنوان توانایی آن در انتقال اطلاعات تعریف می شود.

اجازه دهید فرض کنیم که منبع اطلاعات شکل ۶.۷، یک دنباله تصادفی از نمادهای انتخاب شده از یک مجموعه با تعداد عنصر محدود یا نامحدود را تولید می کند. یعنی خروجی منبع یک متغیر تصادفی گسته است. مجموعه نمادهای منبع $\{a_1, a_2, \dots, a_r\}$ با عنوان الفبای منبع^۱ خوانده می شود، و

^۱ وقتی پایه لگاریتم مورد استفاده در یک عبارت را صریح مشخص نکنیم، حاصل را می توان در هر پایه و واحد اطلاعات متناظر آن تغییر نمود.

1) Source alphabet

۳.۶ عناصر نظریه اطلاعات ۴۹۵



شکل ۷.۶ یک سامانه اطلاعات ساده

عناصر مجموعه، که با نشان داده می شوند، نمادها^۱ یا حروف^۲ خوانده می شوند. احتمال این رخداد که منبع نماد a_i را تولید کند، $P(a_i)$ است، و

$$\sum_{i=1}^J P(a_i) = 1. \quad (2-3.6)$$

بردار^۳ $\mathbf{z} = [P(a_1), P(a_2), \dots, P(a_J)]^T$ است که تمام احتمالات نمادهای منبع $\{P(a_1), P(a_2), \dots, P(a_J)\}$ را نمایش می دهد. مجموعه محدود^۴ (A, z) به طور کامل منبع اطلاعات را توصیف می کند.

احتمال اینکه منبع گسته، نماد a_i را گسیل نماید، $P(a_i)$ است، بنابراین طبق معادله (۱-۳.۶) اطلاعات تولید شده بوسیله این نماد منبع $I(a_i) = -\log P(a_i)$ می باشد. اگر k نماد تولید شوند، طبق قانون اعداد بزرگ برای مقدار به اندازه کافی بزرگ k ، نماد a_i (به طور متوسط) $kP(a_i)$ بار، تولید می شود. بنابراین اطلاعات بدست آمده از k نماد خروجی برابر

$$-kP(a_1) \log P(a_1) - kP(a_2) \log P(a_2) - \dots - kP(a_J) \log P(a_J)$$

و به بیان کوتاه برابر

$$-k \sum_{i=1}^J P(a_i) \log P(a_i)$$

است. متوسط اطلاعات خروجی هر منبع، که با $H(z)$ نشان داده می شود، برابر

$$H(z) = -\sum_{i=1}^J P(a_i) \log P(a_i) \quad (3-3.6)$$

است و عدم قطعیت^۱ یا آتروپی^۲ منبع خوانده می شود. آتروپی، مقدار متوسط اطلاعات (بر حسب واحد ۲-گانی بر نماد) است که با مشاهده یک خروجی منبع بدست می آید. وقتی آتروپی افزایش یابد، عدم قطعیت نیز بیشتر می شود و بنابراین اطلاعات بیشتری به منبع منتب می شود. اگر نمادهای منبع با احتمال مشابه باشند، آتروپی یا عدم قطعیت معادله (۴-۳) بیشینه^۳ می شود و منبع بزرگترین مقدار متوسط ممکن اطلاعات را در هر نماد منبع تولید می کند.

پس از مدلسازی منبع اطلاعات، نسبتاً به سادگی می توان تابع انتقال کanal اطلاعات را به دست آورد.

چون در شکل ۶.۷ ورودی کanal را به صورت یک متغیر تصادفی گسته مدل نمودیم، اطلاعات انتقال یافته به خروجی کanal نیز یک متغیر تصادفی گسته است. خروجی کanal همانند متغیر تصادفی منبع، مقادیری از یک مجموعه با تعداد محدود یا نامحدود از نمادها $\{b_1, b_2, \dots, b_K\}$ ، که الفبای کanal^۴ $P(b_i)$ خوانده می شود، به خود می گیرد. احتمال این رخداد که نماد b_i به کاربر اطلاعات برسد، است. مجموعه محدود (B, v) ، که $v = [P(b_1), P(b_2), \dots, P(b_K)]^T$ است، کاملاً خروجی

کanal و بنابراین اطلاعات دریافتی بوسیله کاربر را توصیف می کند.

احتمال یک خروجی معین از کanal، $P(b_i)$ ، و توزیع احتمال منبع، z ، با عبارت^۵

$$P(b_i) = \sum_{j=1}^J P(b_i|a_j)P(a_j) \quad (4-3.6)$$

به هم مرتبط می شوند که $P(b_i|a_j)$ احتمال شرطی دریافت نماد خروجی b_i بشرط ارسال نماد منبع^۶ است. اگر احتمالات شرطی ارجاع شده در معادله (۴-۳) در ماتریس \mathbf{Q} با ابعاد $J \times K$ به صورت زیر مرتب شوند

$$\mathbf{Q} = \begin{bmatrix} P(b_1|a_1) & P(b_1|a_2) & \dots & P(b_1|a_J) \\ P(b_2|a_1) & \ddots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ P(b_K|a_1) & P(b_K|a_2) & \dots & P(b_K|a_J) \end{bmatrix} \quad (5-3.6)$$

^۱ یکی از قوانین اساسی نظریه احتمال این است که برای هر رخداد دلخواه D و عدد رخداد دو به دو جدا از هم

^۲ $P(D) = P(D|C_1)P(C_1) + \dots + P(D|C_i)P(C_i)$ برابر است با:

^۳ ۱) Uncertainty ۲) entropy ۳) Maximize ۴) Channel alphabet ۵) Mutually exclusive

توزیع احتمال القبای کامل خروجی را می توان از رابطه^{*}

$$\mathbf{v} = \mathbf{Q}\mathbf{z}. \quad (6-3.6)$$

محاسبه کرد. ماتریس \mathbf{Q} با عناصر $P(b_k|a_i) = q_{ik}$ با عنوان ماتریس انتقال مستقیم کانال^۱ یا با اصطلاح مخفف «ماتریس کانال^۲» اشاره می شود.

برای تعیین ظرفیت یک کانال اطلاعات با ماتریس انتقال مستقیم کانال \mathbf{Q} ، در ابتدا باید با این فرض که کاربر یک خروجی ویژه b _۰ را مشاهده می کند، آنروپی منع اطلاعات را محاسبه کنیم. معادله (۶-۳.۴)

توزیعی از نمادهای منع را برای هر b _i مشاهده شده تعریف می کند، بنابراین هر b _i یک تابع آنروپی شرطی^۳ دارد. براساس مراحلی که به معادله (۶-۳.۳) متهی می شود، این تابع آنروپی شرطی که با $H(\mathbf{z}|b_i)$ نشان داده می شود، را می توان به صورت

$$H(\mathbf{z}|b_i) = -\sum_{j=1}^J P(a_j|b_i) \log P(a_j|b_i) \quad (7-3.6)$$

نوشت که $P(a_j|b_i)$ احتمال ارسال نماد a_j به شرط دریافت نماد b_i است. مقدار امید یا متوسط این عبارت روی تمام b_i ها

$$H(\mathbf{z}|\mathbf{v}) = \sum_{k=1}^K H(\mathbf{z}|b_k)P(b_k) \quad (8-3.6)$$

است که پس از جایگذاری معادله (۶-۳.۷) برای $H(\mathbf{z}|b_i)$ و چند تغییر جزئی^۴ می توان آن را به صورت

$$H(\mathbf{z}|\mathbf{v}) = -\sum_{j=1}^J \sum_{k=1}^K P(a_j, b_k) \log P(a_j|b_k) \quad (9-3.6)$$

نوشت. در اینجا $P(a_j, b_K)$ احتمال توأم a_j و b_K است. یعنی $P(a_j, b_K)$ احتمال این است که a_j فرستاده و b_K دریافت شود.

عبارت $H(\mathbf{z}|\mathbf{v})$ ابهام^۴ \mathbf{z} به شرط \mathbf{v} خوانده می شود. این عبارت نمایانگر متوسط اطلاعات یک نماد

^{*} از این واقعیت که احتمال توأم دو رخداد C و D برابر $P(C, D) = P(C)P(D|C) = P(D)P(C|D)$ است، استفاده می کنیم.

1) Forward channel transition matrix

2) Channel matrix

3) Conditional entropy function

4) Equivocation

منبع با فرض مشاهده نماد خروجی مربوط به آن می باشد. چون $H(\mathbf{z})$ متوسط اطلاعات یک نماد منبع با فرض نداشتن هیچ اطلاعی از نماد خروجی مربوطه می باشد، تفاضل بین $H(\mathbf{z})$ و $H(\mathbf{z}|\mathbf{v})$ برابر اطلاعات متوسط دریافتی از مشاهده یک نماد خروجی می باشد. این تفاضل که اطلاعات متقابل^۱ \mathbf{z} و \mathbf{v} خوانده می شود، با $I(\mathbf{z}, \mathbf{v})$ نشان داده می شود:

$$I(\mathbf{z}, \mathbf{v}) = H(\mathbf{z}) - H(\mathbf{z}|\mathbf{v}). \quad (10-3. 6)$$

با جایگذاری معادلات (۶. ۳-۳) و (۶. ۹-۳) به ترتیب بجای $H(\mathbf{z}|\mathbf{v})$ و $H(\mathbf{z})$ و یادآوری این که

$$P(a_i) = P(a_i, b_1) + P(a_i, b_2) + \dots + P(a_i, b_K)$$

$$I(\mathbf{z}, \mathbf{v}) = \sum_{j=1}^J \sum_{k=1}^K P(a_j, b_k) \log \frac{P(a_j, b_k)}{P(a_j)P(b_k)} \quad (11-3. 6)$$

که پس از تغییر بیشتر می توان آن را به صورت

$$I(\mathbf{z}, \mathbf{v}) = \sum_{j=1}^J \sum_{k=1}^K P(a_j) q_{kj} \log \frac{q_{kj}}{\sum_{i=1}^J P(a_i) q_{ki}}. \quad (12-3. 6)$$

نوشت. بنابر این اطلاعات متوسط دریافتی پس از مشاهده یک خروجی کانال اطلاعات، تابعی از توزیع احتمال نماد منبع \mathbf{z} و ماتریس کانال \mathbf{Q} است. حداقل مقدار ممکن $I(\mathbf{z}, \mathbf{v})$ صفر است و وقتی رخ می دهد که نمادهای ورودی و خروجی مستقل آماری باشند. آنگاه $P(a_i, b_k) = P(a_i)P(b_k)$ جزء لگاریتمی معادله (۶. ۱۱-۳) برای تمام j و k ها صفر است. مقدار بیشینه $I(\mathbf{z}, \mathbf{v})$ روی تمام انتخاب های ممکن از توزیع منبع \mathbf{z} ، برابر ظرفیت کانال توصیف شده با ماتریس کانال \mathbf{Q} ، یعنی C ، است. یعنی

$$C = \max [I(\mathbf{z}, \mathbf{v})] \quad (13-3. 6)$$

که بیشینه گیری روی تمام توزیع های ممکن ورودی انجام می شود. ظرفیت کانال نرخ بیشینه ای (بر حسب واحد های اطلاعات r - گانی بر نماد منبع) را تعریف می کند که با آن می توان اطلاعات را به طور قابل اطمینانی از درون کانال انتقال داد. گذشته از این ظرفیت کانال به احتمالات ورودی منبع (یعنی به چگونگی استفاده از کانال) بستگی ندارد، بلکه تابعی از احتمالات شرطی که خود کانال را تعریف می کند.

1) Mutual information

می باشد.

مثال: یک منبع اطلاعات دودویی با الفبای منبع $\{0, 1\} = A = \{a_1, a_2\}$ را در نظر بگیرید. احتمالاتی که منبع نمادهای a_1 و a_2 را تولید می کند، به ترتیب $P(a_1) = p_{bs}$ و $P(a_2) = 1 - p_{bs} = \bar{p}_{bs}$ هستند. از معادله (۳-۳) آنتروپی منبع برابر

$$H(z) = -p_{bs} \log_2 p_{bs} - \bar{p}_{bs} \log_2 \bar{p}_{bs}$$

است. چون $[1 - p_{bs}]^T = [P(a_1), P(a_2)]^T = [p_{bs}, 1 - p_{bs}]$ است، $H(z) = [p_{bs}, 1 - p_{bs}]$ تنها به عامل p_{bs} وابسته است، و طرف راست معادله، که تابع آنتروپی دودویی^۱ خوانده می شود، با $(۰) H_{bs}$ نشان داده می شود. بنابراین برای مثال (t) برابر تابع $t \log_2 t - \bar{t} \log_2 \bar{t}$ است. شکل ۶.۸. (الف) نموداری از $H_m(p_{bs})$ برای $0 \leq p_{bs} \leq 1$ نشان می دهد. توجه شود که وقتی $p_{bs} = 1/2$ مقدار بیشینه اش (یک بیت) را به دست می آورد. برای تمام مقادیر دیگر p_{bs} ، منبع کمتر از یک بیت اطلاعات تولید می کند. حال فرض کنید که اطلاعات باید روی یک کانال اطلاعات دودویی نویزدار ارسال شود و اجازه دهید که احتمال خطأ در خلال ارسال هر نماد P باشد. چنین کانالی، کانال متقاضی دودویی^۲ (BSC) خوانده می شود و با ماتریس کانال

$$\mathbf{Q} = \begin{bmatrix} 1 - p_e & p_e \\ p_e & 1 - p_e \end{bmatrix} = \begin{bmatrix} \bar{p}_e & p_e \\ p_e & \bar{p}_e \end{bmatrix}$$

تعریف می شود. برای هر ورودی یا نماد منبع، کانال BSC یک خروجی b از الفبای خروجی $B = \{b_1, b_2\} = \{0, 1\}$ را تولید می کند. احتمالات دریافت نمادهای خروجی b_1 و b_2 را می توان از معادله (۶-۳) تعیین کرد:

$$\mathbf{v} = \mathbf{Qz} = \begin{bmatrix} \bar{p}_e & p_e \\ p_e & \bar{p}_e \end{bmatrix} \begin{bmatrix} p_{bs} \\ \bar{p}_{bs} \end{bmatrix} = \begin{bmatrix} \bar{p}_e p_{bs} + p_e \bar{p}_{bs} \\ p_e p_{bs} + \bar{p}_e \bar{p}_{bs} \end{bmatrix}$$

بنابراین احتمال این که خروجی صفر باشد، برابر $p_e \bar{p}_{bs} + \bar{p}_e p_{bs}$ و احتمال این که خروجی یک باشد،

برابر $p_e p_{bs} + \bar{p}_e \bar{p}_{bs}$ می‌باشد. این احتمالات را می‌توان از عبارت قبلی برای ۷ استخراج کرد زیرا

$$\text{است. } [P(b_1), P(b_2)]^T = [P(0), P(1)]^T$$

حال می‌توان اطلاعات متقابل کانال BSC را از معادله (۶.۳-۱۲) محاسبه کرد. با بسط مجموعه‌ی
 این معادله و گردآوری اجزای مناسب داریم

$$I(z, v) = H_m(p_{bs}p_e + \bar{p}_{bs}\bar{p}_e) - H_m(p_e)$$

که (۰.۸) H_m تابع آنتروپی دودویی شکل ۶.۸ (الف) است. برای یک مقدار ثابت p ، وقتی p_{bs} صفر یا یک
 باشد، $I(z, v)$ صفر است. گذشته از این وقتی نمادهای متبع دودویی هم احتمال باشند، $I(z, v)$ به مقدار
 بیشینه اش می‌رسد. شکل ۶.۸ (ب) برای تمام مقادیر p_{bs} و یک احتمال خطای p_e معلوم، $I(z, v)$ را
 نشان می‌دهد.

طبق معادله (۶.۱۲-۳) با بیشینه گیری از اطلاعات متقابل روی تمام توزیع‌های ممکن منبع، ظرفیت
 کانال BSC بدست می‌آید. از شکل ۶.۸ (ب) که برای تمام توزیع‌های ممکن منبع دودویی (یعنی برای
 $z = [0, 1]^T$ تا $p_{bs} \leq 0$ یا برای $z = [1, 0]^T$ تا $p_{bs} \leq 1$) را در می‌کند، می‌بینیم که وقتی
 $p_{bs} = \frac{1}{2}$ باشد، $I(z, v)$ (برای هر p_e) بیشینه است. این مقدار p_{bs} متناظر با توزیع منبع $z = [\frac{1}{2}, \frac{1}{2}]^T$
 بوده و مقدار متناظر $I(z, v)$ برابر $-H_m(p_e)$ است. بنابراین ظرفیت کانال BSC برابر

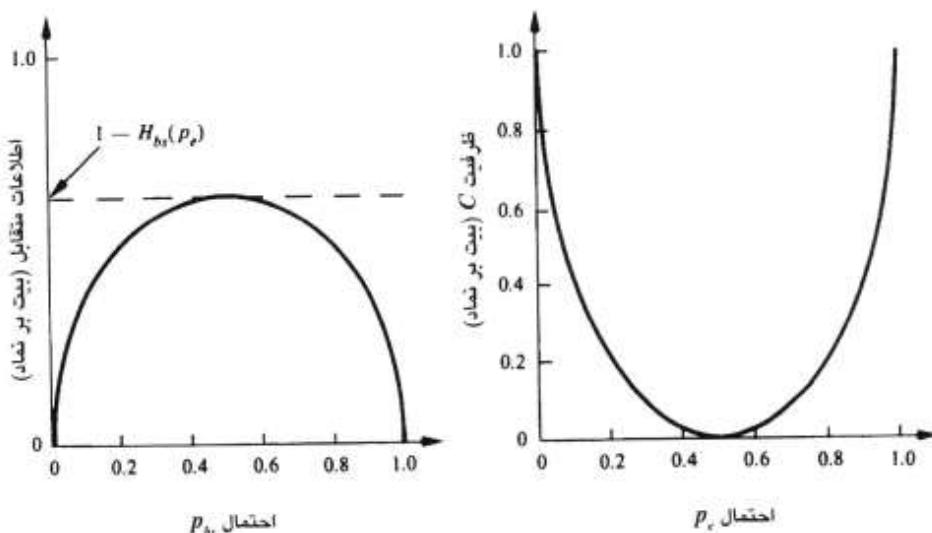
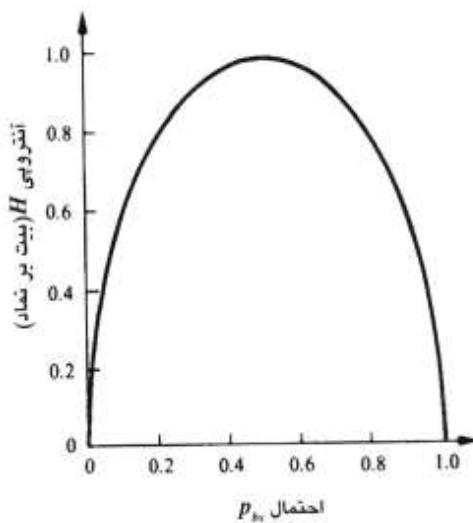
$$C = 1 - H_m(p_e).$$

می‌باشد که در شکل ۶.۸ (ب) رسم شده است. توجه شود که وقتی احتمال خطای کانال صفر باشد
 $p_e = 0$ -همچنین وقتی خطای کانال قطعی باشد ($p_e = 1$) - ظرفیت کانال به مقدار بیشینه یک بیت
 بر نماد می‌رسد. در هر دو حالت، انتقال بیشینه اطلاعات ممکن است، زیرا خروجی کانال کاملاً قابل
 پیش‌بینی است. بر عکس وقتی $p_{bs} = \frac{1}{2}$ باشد، خروجی کانال کاملاً غیرقابل پیش‌بینی می‌گردد و هیچ
 اطلاعاتی را نمی‌توان از درون کانال انتقال داد.

۳.۳.۶ قضایای اساسی رمزنگاری

چهار چوب ریاضی کلی معرفی شده در بخش ۳.۲ براساس مدل شکل ۶.۷ است که حاوی یک منبع

۳.۶ عناصر نظریه اطلاعات ۴۰۱



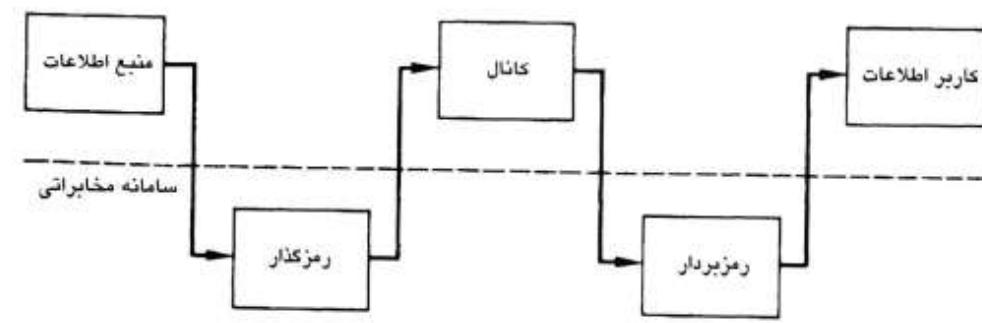
شکل ۳.۶ سه تابع اطلاعات دودویی: (الف) تابع آنترپویی دودویی؛ (ب) اطلاعات متقابل کانال متقاضی دودویی (BSC)؛ ظرفیت کانال BSC

اطلاعات، کانال و کاربر می باشد. در این بخش یک سامانه مخابراتی به مدل می افزاییم و سه قضیه پایه ای مربوط به رمزگاری یا نمایش اطلاعات را بررسی می کنیم. همان طور که شکل ۹.۶ نشان می دهد، سامانه مخابراتی بین منبع و کاربر قرارداده می شود و مشکل از یک رمزگذار و یک رمزبردار می باشد.

قضیه رمزگاری بدون نویز

وقتی کانال اطلاعات و سامانه مخابراتی هر دو عاری از خطأ باشد، وظیفه اصلی سامانه مخابراتی نمایش تا حد ممکن فشرده‌ی منبع می باشد. قضیه رمزگاری بدون نویز^۱ که قضیه اول شاتون^۲ (Shannon[1948]) نیز خوانده می شود، با این شرایط، کمینه قابل دستیابی برای متوسط طول کلمه رمز بر نماد منبع را تعریف می کند.

یک منبع اطلاعات با مجموعه محدود (A) و نمادهای منبع مستقل آماری، یک منبع بی حافظه^۳ خوانده می شود. اگر در نظر بگیریم که خروجی آن یک نماد n -تالی^۴ از الفبای منبع (n تا از نمادهای منبع بجای یک نماد تکی) باشد، خروجی منبع یک متغیر تصادفی بلوکی^۵ است. این خروجی که با α نشان داده می شود، یکی از A^n مقدار ممکن از مجموعه 'تمام دنباله های' n عنصری ممکن، $A^n = \{\alpha_1, \alpha_2, \dots, \alpha_{A^n}\}$ را می گیرد که هر α_i مشکل از n نماد از مجموعه A است. (نشانه ' A^n ')



شکل ۹.۶ مدل یک سامانه مخابراتی

1) Noiseless coding theorem
4) an n -tuple of symbols

2) Shannon's first theorem
5) Block random variable

3) Zero- memory source

مجموعه نمادهای بلوکی را از A ، مجموعه نمادهای تکی، مشخص می‌سازد. احتمال یک معین، $P(\alpha_i)$ است و با احتمالات تک نمادی $P(\alpha_i)$ رابطه

$$P(\alpha_i) = P(a_{i1}) P(a_{i2}) \cdots P(a_{in}) \quad (14-3.6)$$

را دارد که زیرنویس دوم برای شاخص گذاری n نماد از A که α_i مورد نظر را تشکیل می‌دهند، استفاده می‌شود. مانند گذشته بردار \mathbf{z}' (علامت پرایم^۱ برای دلالت بر استفاده از متغیر تصادفی بلوکی افزوده شد) نمایانگر مجموعه تمام احتمالات منبع $\{P(\alpha_1), P(\alpha_2), \dots, P(\alpha_n)\}$ می‌باشد و آنرопی منبع عبارتست از:

$$H(\mathbf{z}') = - \sum_{i=1}^n P(\alpha_i) \log P(\alpha_i).$$

با جایگذاری معادله (14-3.6) بجای $P(\alpha_i)$ و ساده سازی داریم

$$H(\mathbf{z}') = n H(\mathbf{z}). \quad (15-3.6)$$

بنابراین آنرопی این منبع اطلاعات بی حافظه (که متغیر تصادفی بلوکی را تولید می‌کند) n برابر آنرопی منبع تک نمادی متناظر است. چنین منبعی با عنوان گسترش n ام منبع تک نمادی (گسترش نیافته) اشاره می‌شود. توجه شود که گسترش اول هر منبع، همان منبع گسترش نیافته است. چون میزان اطلاعات خروجی α_i منبع، برابر $[-\log(1/P(\alpha_i))]$ است، منطقی به نظر می‌رسد که α_i را با یک کلمه رمز با طول صحیح $l(\alpha_i)$ که در رابطه زیر صدق کند، رمز نماییم:

$$\log \frac{1}{P(\alpha_i)} \leq l(\alpha_i) < \log \frac{1}{P(\alpha_i)} + 1. \quad (16-3.6)$$

به نظر می‌رسد که بهتر است خروجی α_i منبع را با یک کلمه رمز که طولش کوچکترین عدد صحیح بزرگتر از مقدار اطلاعات α_i باشد، نمایش دهیم.[†] با ضرب این نتیجه در $P(\alpha_i)$ و جمع روی تمام آنها داریم:

$$\sum_{i=1}^n P(\alpha_i) \log \frac{1}{P(\alpha_i)} \leq \sum_{i=1}^n P(\alpha_i) l(\alpha_i) < \sum_{i=1}^n P(\alpha_i) \log \frac{1}{P(\alpha_i)} + 1$$

[†] با این قید می‌توان رمزی ساخت که قابل رمزبرداری به طور یکتا باشد.

که می توان آن را به صورت زیر نیز نوشت

$$H(\mathbf{z}') \leq L'_{avg} < H(\mathbf{z}') + 1 \quad (17-3.6)$$

L'_{avg} نمایانگر طول متوسط کلمه رمز متناظر با گسترش n ام منبع اولیه است:

$$L'_{avg} = \sum_{i=1}^n P(a_i) l(a_i). \quad (18-3.6)$$

با تقسیم معادله (17-3.6) بر n و استفاده از معادله (15-3.6) که $H(\mathbf{z}')/n$ برابر $H(\mathbf{z})$ است،

داریم

$$H(\mathbf{z}) \leq \frac{L'_{avg}}{n} < H(\mathbf{z}) + \frac{1}{n} \quad (19-3.6)$$

که در حالت حدی به صورت

$$\lim_{n \rightarrow \infty} \left[\frac{L'_{avg}}{n} \right] = H(\mathbf{z}). \quad (20-3.6)$$

می شود. معادله (20-3.6) قضیه اول شانون برای یک منبع بی حافظه را بیان می کند. این معادله آشکار می سازد که می توان با رمزگاری گسترش های بی نهایت طولانی منبع، L'_{avg}/n را تا هر اندازه دلخواهی به $H(\mathbf{z})$ نزدیک کرد. گرچه این رابطه با فرض استقلال آماری نمادهای منبع بدست آمد، اما به سادگی به منابع کلی تر، مانند منبع مارکوف مرتبه m^1 که در آن وقوع نماد منبع α به m نماد قبلی (m محدود) بستگی دارد، گسترش می یابد. منابع مارکوف معمولاً برای مدل سازی همبستگی های بین پیکسلی در تصویر به کار می روند. چون $H(\mathbf{z})$ یک کران پایین برای L'_{avg}/n است [یعنی در معادله (20-3.6) وقتی n بزرگ می شود، حد L'_{avg}/n برابر $H(\mathbf{z})$ است]، بازده η برای هر روش رمزگذاری را می توان به صورت زیر تعریف کرد:

$$\eta = n \frac{H(\mathbf{z})}{L'_{avg}} \quad (21-3.6)$$

مثال: یک منبع اطلاعات بی حافظه با الفبای منبع $\{a_1, a_2\}$ ، احتمالات نماد $= \frac{2}{3}$ و

1) m -th order Markov source

جدول ۴.۶ مثالی از رمزگاری گسترشی

منبع تعدادها α_i	$P(\alpha_i)$ Eq. (6.3-14)	$I(\alpha_i)$ Eq. (6.3-1)	$l(\alpha_i)$ Eq. (6.3-16)	کلمه رمز	طول رمز
گسترش اول					
α_1 a_1	2/3	0.59	1	0	1
α_2 a_2	1/3	1.58	2	1	1
گسترش دوم					
α_1 a_1a_1	4/9	1.17	2	0	1
α_2 a_1a_2	2/9	2.17	3	10	2
α_3 a_2a_1	2/9	2.17	3	110	3
α_4 a_2a_2	1/9	3.17	4	111	3

اگر $P(a_2) = 1/3$ را دارد. آنروپی این منبع (با استفاده از معادله ۶.۳-۳. $0.918 = 1 - \sum P(a_i)$) بیت بر نماد است. اگر نمادهای a_1 و a_2 به ترتیب با کلمات رمز دودویی ۰ و ۱ بیان شوند، $L'_{avg} = 1.83$ بیت برناماد بوده و بازده رمز حاصل $1/(0.918) = 1.89$ است.

جدول ۶.۶ رمزی را که اخیراً توصیف شد و یک روش رمزگذاری دیگر براساس گسترش دوم منبع را نشان می دهد. بخش پایینی جدول ۶.۶. چهار نماد بلوکی (α_1 و α_2 , α_3 , α_4) در گسترش دوم منبع را فهرست می کند. با استفاده از معادله ۶.۳-۱۴) احتمالات آنها تابع با چهارتایی ($1/4, 1/9, 2/9, 2/9$) است. طبق معادله ۶.۳-۱۸) متوسط طول کلمه در رمزگذاری دوم $1.89 = 1/4 + 2/9 + 2/9 + 2/9$ بیت بر نماد است. آنروپی گسترش دوم دوبرابر آنروپی منبع گسترش نیافته یعنی $1.83 = 1/3 + 2/9 + 2/9 + 2/9$ بیت بر نماد است. بنابراین بازده رمزگذاری دوم $0.97 = 1.83/1.89 = 0.97$ است. این بازده کمی بهتر از بازده رمزگذاری گسترش نیافته یعنی 0.92 است. رمزگذاری گسترش دوم منبع، متوسط تعداد بیت‌های رمز بر نماد منبع را از یک بیت برنامد به $1.89/2 = 0.94$ بیت بر نماد کاهش می دهد. □

قضیه رمزگاری با نویز

اگر کanal شکل ۶.۹ نویزدار یا در معرض خطا باشد، به جای نمایش تاحد ممکن فشرده اطلاعات، رمزگذاری آن به طوری که مخابره مطمئن ممکن باشد بیشتر مورد توجه قرارمی گیرد. پرسشی که به طور طبیعی مطرح می شود، این است که: چقدر می توان خطای مخابره را کوچک کرد؟ مثال: یک کanal متقارن دودویی (BSC) احتمال خطای $p = 0.01$ را دارد (یعنی 99% در صد نمادهای منبع بدروستی از کanal انتقال می یابند). یک روش ساده برای افزایش قابلیت اطمینان چند بار تکرار مخابره هر

۴۰۶ پردازش تصویر رقمه

پیغام یا نماد دودویی است. برای مثال فرض کنید که به جای ۰ یا ۱، بترتیب پیغام های رمز شده ۰۰۰ یا ۱۱۱ ارسال شوند. احتمال این که در خلال ارسال یک پیغام سه نمادی هیچ خطای رخ ندهد، $(1-p)^3$ است. احتمال یک خطای p^3 ، احتمال دو خطای $3p^2\bar{p}$ و احتمال سه خطای p^3 است. چون احتمال خطای در ارسال یک نماد کمتر از ۵۰ درصد است، پیغام های دریافتی را می توان با استفاده از رأی اکثریت^۱ در هر سه نماد دریافتی، رمزبرداری نمود. بنابراین احتمال رمزبرداری غیر صحیح یک کلمه رمز سه نمادی جمع احتمالات خطاهای دونمادی و سه نمادی، یعنی $3p^2\bar{p} + p^3$ می باشد. وقتی هیچ خطای رخ ندهد یا فقط یک خطای رخ بدهد، قاعده رأی اکثریت پیام را به طور صحیح رمزبرداری می نماید. در این حالت به ازای $0.01 = p$ ، احتمال خطای کل به ۰.۰۰۰۳ کاهش می یابد.

با گسترش روش رمزنگاری تکراری که اخیراً توصیف شد، می توانیم خطای کلی مخابره را تا حد مطلوب کوچک نماییم. در حالت کلی برای انجام این کار، گسترش n ممکن است با استفاده از یک دنباله نماد K گانی با طول n نماد، رمز می کنیم که باید $\varphi \geq K$ باشد. این روش براساس انتخاب φ تا از K دنباله نماد ممکن به عنوان کلمه های رمز معتبر و طرح یک قاعده تصمیم که احتمال رمزبرداری صحیح را بیشینه کند، می باشد. در مثال قبل، سه بار تکرار هر نماد منبع معادل با رمزگذاری بلوکی ممکن است دودویی گسترش نیافرته با استفاده از دو تا از ۲ یا ۸ کلمه رمز دودویی ممکن می باشد. دو کلمه رمز معتبر ۰۰۰ یا ۱۱۱ هستند. اگر یک کلمه رمز نامعتبر به رمزبردار وارد شود، قاعده اکثریت از میان سه بیت رمز، خروجی را تعیین می کند.

یک منبع اطلاعات بی حافظه، اطلاعات را در نرخی (بر حسب واحد اطلاعات برناماد) برابر آتروپی آن منبع، $H(z)$ ، تولید می کند. گسترش n ممکن است اطلاعات را در نرخی برابر $nH(z)$ واحد اطلاعات بر نماد فراهم می آورد. مانند مثال قبل اگر این اطلاعات رمز شوند، نرخ بیشینه اطلاعات رمز شده $r/\log(\varphi)$ است و وقتی رخ می دهد که φ تا کلمه رمز معتبر که برای رمزگرد منبع استفاده می شوند، هم احتمال باشند. بنابراین یک رمز با اندازه φ و طول بلوکی n نماد، دارای نرخ

$$R = \log \frac{\varphi}{r}. \quad (22-3.6)$$

گفته می شود. قضیه دوم شانون^۲ (Shannon[1948]) که قضیه رمزنگاری بدون نویز نیز خوانده می شود،

1) Majority vote

2) Shannon's second theorem

به ما می گوید که برای هر C که ظرفیت کانال بی حافظه باماتریس \mathbf{Q} باشد،^۱ لاقل یک عدد صحیح r و یک رمز بلوکی با طول r نماد و نرخ $(R < C)$ وجود دارد طوری که احتمال هر خطای رمزبرداری بلوکی برای هر $0 < r$ دلخواه، کمتر یا مساوی ϵ باشد. بنابراین احتمال خطای رامی توان تا آنچه که نرخ ارسال پیام های رمز شده کمتر از ظرفیت کانال باشد، به اندازه دلخواه کوچک کرد.

قضیه رمزگذاری منبع

قضایای توصیف شده تاکنون، محدودیتهای اساسی مخابره بدون خطای روی هر یک از کانال های قابل اعتماد و غیرقابل اعتماد را توصیف می کنند. در این بخش حالتی را بررسی می کنیم که در آن کانال بدون خطای است اما فرآیند مخابره، با اتلاف همراه است. با این شرایط وظیفه اصلی سامانه مخابراتی "فسرده سازی اطلاعات" است. در اغلب حالات متوسط خطای حاصل از فشرده سازی به یک سطح بیشینه مجاز D محدود است. می خواهیم با معیار مشابه داده شده، کوچکترین میزانی را که می توان اطلاعات منبع را به کاربر انتقال داد، تعیین کنیم. در شاخه ای از نظریه اطلاعات با نام «نظریه اعوجاج نرخ»^۲ به طور ویژه به این مسئله پرداخته می شود.

اجازه دهید که خروجی های منبع اطلاعات و رمزبردار شکل ۹.۶ به ترتیب با مجموعه های محدود (A, z) و (B, v) تعریف شوند. اکنون فرض بر این است که کانال شکل ۶.۹ بدون خطای است، بنابراین ماتریس کانال \mathbf{Q} ، که طبق معادله $(6-3) z = 6-7$ مرتب می کند، را به تنهایی می توان به عنوان مدل فرآیندرمزگذاری-رمزبرداری در نظر گرفت. چون فرآیندرمزگذاری-رمزبرداری قطعی^۳ (غیرتصادفی) است، \mathbf{Q} یک کانال بی حافظه مصنوعی را که اثر فشرده سازی و افشرده سازی اطلاعات را مدل می نماید، توصیف می کند. هر بار که منبع، نماد a را تولید کند، آن نماد با یک نماد رمز نمایش داده می شود و در مقصد آن نمادرمزبرداری می شود تا به احتمال p_a نماد خروجی b را بدهد (بخش ۶.۳.۶ را بینید). برای پرداختن به مسئله رمزگذاری منبع طوری که اعوجاج متوسط کمتر از D باشد، باید رابطه ای نوشته شود که به هر تقریب ممکن در خروجی منبع، یک مقدار کمی اعوجاج را نسبت دهد. در مورد منبع گسترش نیافته، برای تعریف خطای^۴ مربوط به تولید مجدد خروجی^۵ منبع با داشتن خروجی b رمزبردار،

^۱ کانال بی حافظه کانالی است که در آن پاسخ کانال به نماد ورودی فعلی، مستقل از پاسخ آن به نمادهای ورودی قبلی است.

1) Rate distortion theory

2) Deterministic

3) Penalty

می توان از یکتابع هزینه غیر منفی $\rho(a_i, b_k)$ که معیار اعوجاج^۱ خوانده می شود، استفاده کرد. خروجی منبع تصادفی است، بنابراین اعوجاج نیز یک متغیر تصادفی است و مقدار متوسطش که با $d(\mathbf{Q})$ نشان داده می شود، عبارتست از:

$$\begin{aligned} d(\mathbf{Q}) &= \sum_{j=1}^J \sum_{k=1}^K \rho(a_j, b_k) P(a_j, b_k) \\ &= \sum_{j=1}^J \sum_{k=1}^K \rho(a_j, b_k) P(a_j) q_{kj}. \end{aligned} \quad (23-3.6)$$

نشانه $(\mathbf{Q})d$ تأکید می کند که متوسط اعوجاج، تابعی از روای رمزگذاری-رمزبرداری که با \mathbf{Q} مدل شد، است. یک روای رمزگذاری-رمزبرداری خاص، قابل قبول در سطح D ^۲ است اگر و تنها اگر متوسط اعوجاج مربوط به \mathbf{Q} کمتر یا مساوی D باشد. بنابراین مجموعه تمام روای های رمزگذاری-رمزبرداری قابل قبول در سطح D عبارتست از:

$$\mathbf{Q}_D = \{q_{kj} \mid d(\mathbf{Q}) \leq D\}. \quad (24-3.6)$$

چون هر روای رمزگذاری-رمزبرداری با یک ماتریس کانال مصنوعی \mathbf{Q} تعریف می شود، متوسط اطلاعات حاصل از مشاهده یک خروجی رمزبردار را می توان طبق معادله (۶.۳-۲۴) محاسبه نمود. بنابراین می توان تابع اعوجاج نرخ^۳ را به صورت

$$R(D) = \min_{\mathbf{Q} \in \mathbf{Q}_D} [I(\mathbf{z}, \mathbf{v})] \quad (25-3.6)$$

تعریف نمود که در واقع مقدار کمینه معادله (۶.۳-۲۴) روی تمام رمزهای قابل قبول در سطح D می باشد. توجه شود که می توان روای \mathbf{Q} کمینه سازی کرد، زیرا $I(\mathbf{z}, \mathbf{v})$ تابعی از احتمالات دربردار \mathbf{z} و عناصر ماتریس \mathbf{Q} است. اگر $R(D) = D$ باشد، $R(0) \leq H(\mathbf{z})$ است، یعنی $H(\mathbf{z})$ کمتر یا مساوی آنتروپی منبع است، یعنی $R(0) \leq H(\mathbf{z})$ است.

معادله (۶.۳-۲۵) نرخ کمینه ای را تعریف می کند که می توان اطلاعات منبع را با قید این که متوسط اعوجاج کمتر یا مساوی D باشد، به کاربر حمل کرد. برای محاسبه این نرخ [یعنی $R(D)$] کافیست که

1) Distortion measure

2) D -admissible

3) Rate distortion function

(معادله ۶.۳-۱۲) را با انتخاب مناسب \mathbf{Q} (یا q_{kj} ها) با قیود

$$q_{ij} \geq 0 \quad (26-3.6)$$

$$\sum_{i=1}^K q_{ij} = 1 \quad (27-3.6)$$

و

$$d(\mathbf{Q}) = D. \quad (28-3.6)$$

کمینه کنیم. معادلات (۶.۳-۲۶) و (۶.۳-۲۷) خواص اساسی ماتریس کانال \mathbf{Q} هستند. عناصر \mathbf{Q} باید مثبت باشند، و چون به ازای هر نماد ورودی به کانال، باید یک خروجی دریافت شود، جمع عناصر هر ستون \mathbf{Q} باید برابر واحد باشد. معادله (۶.۳-۲۸) نشان می دهد که نرخ کمینه اطلاعات، وقتی رخ می دهد که بیشینه اعوجاج ممکن، مجاز باشد.

مثال: یک منبع دودویی بی حافظه با نمادهای هم احتمال $\{0, 1\}$ و معیار اعوجاج ساده

$$\rho(a_i, b_k) = 1 - \delta_{ik}$$

در نظر بگیرید که δ_{ik} تابع ضربه یا دلتا است. اگر $a_i \neq b_k$ باشد، $\rho(a_i, b_k)$ برابر یک و در غیر این صورت برابر صفر است، بنابراین هر خطای رمزگذاری-رمزبرداری به عنوان یک واحد اعوجاج محاسبه می شود. با استفاده از حساب تغییرات می توان $R(D)$ را محاسبه کرد. با فرض اینکه $\mu_1, \mu_2, \dots, \mu_{M+1}$ ضرایب لagger از باشند، تابع معیار بسط یافته^۱

$$J(\mathbf{Q}) = I(\mathbf{z}, \mathbf{v}) - \sum_{j=1}^J \mu_j \sum_{k=1}^K q_{kj} - \mu_{M+1} d(\mathbf{Q})$$

را تشکیل می دهیم، مشتقات J/K ام آن نسبت به q_{kj} را برابر صفر قرار می دهیم (یعنی $dJ/dq_{kj} = 0$) و معادلات حاصل همراه با $I + J$ معادله مربوط به معادلات (۶.۳-۲۷) را برای مجھولات q_{kj} ، $\mu_1, \mu_2, \dots, \mu_{M+1}$ حل کنیم. اگر q_{kj} های حاصل غیر منفی باشند (یا معادله ۶.۳-۲۶ را برآورده کنند)، یک جواب معتبر پیدا می شود. برای زوج منبع و اعوجاج تعریف شده

1) Augmented criterion function

در فوق به ۷ معادله ۷ مجهول زیر می رسمیم:

$$\begin{aligned} 2q_{11} &= (q_{11} + q_{12}) \exp [2\mu_1] & 2q_{22} &= (q_{21} + q_{22}) \exp [2\mu_2] \\ 2q_{12} &= (q_{11} + q_{12}) \exp [2\mu_1 + \mu_3] & 2q_{21} &= (q_{21} + q_{22}) \exp [2\mu_2 + \mu_3] \\ q_{11} + q_{21} &= 1 & q_{12} + q_{22} &= 1 \\ q_{21} + q_{12} &= 2D. \end{aligned}$$

آنگاه با انجام دنباله‌ای از مراحل ساده جبری داریم:

$$\begin{aligned} q_{12} &= q_{21} = D \\ q_{11} &= q_{22} = 1 - D \\ \mu_1 &= \mu_2 = \log \sqrt{2(1 - D)} \\ \mu_3 &= \log \frac{D}{1 - D} \end{aligned}$$

چنان‌که

$$\mathbf{Q} = \begin{bmatrix} 1 - D & D \\ D & 1 - D \end{bmatrix}$$

چون نمادهای منبع هم احتمال هستند، حداقل اعوجاج ممکن $\frac{1}{2}$ است. بنابراین $0 \leq D \leq \frac{1}{2}$ و برای تمام D ها معادله (۶. ۳-۱۲) را برآورده می‌کند. اطلاعات متقابل مربوط به \mathbf{Q} و منبع دودویی قبل تعریف شده، با استفاده از معادله (۶. ۳-۱۲) محاسبه می‌شود. در هر حال با توجه به شباهت بین \mathbf{Q} و ماتریس کانال دودویی متقابله، مستقیماً می‌توان نوشت که

$$I(\mathbf{z}, \mathbf{v}) = 1 - H_{bv}(D).$$

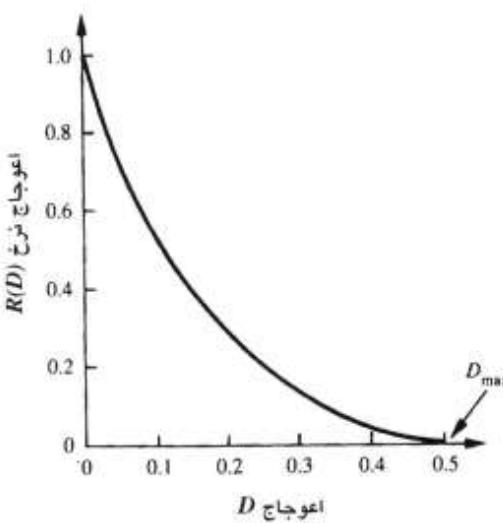
این نتیجه از مثال بخش ۳. ۲. ۶ با جایگذاری $\frac{1}{2}$ در $p_c = D$ و $p_{bc} = p_e$ می‌شود.

بدست می‌آید.تابع اعوجاج نرخ مستقیماً از معادله (۶. ۳-۲۵) نتیجه گرفته می‌شود:

$$R(D) = \min_{\mathbf{q} \in \mathbf{Q}_D} [1 - H_{bv}(D)] = 1 - H_{bv}(D).$$

آخرین ساده‌سازی براساس این واقعیت است که برای هر D معین، $H_{bv}(D) = 1$ تنها یک مقدار می‌گیرد که خواه ناخواه^۱ مقدار کمینه است. تابع حاصل در شکل ۱۰. ۶ رسم شده است. شکل آن نماینده نوعی

1) By default



شکل ۱۰۰. تابع اعوجاج نرخ برای یک منبع متقاضن دودویی

اغلب توابع اعوجاج نرخ است. توجه شود که برای هر $D \geq D_{\max}$ که $R(D) = 0$ باشد، D_{\max} است که $R(D) = 0$ است که D_{\max} مقدار بیشینه می باشد. بعلاوه $R(D)$ همواره مثبت، کاهشی یکنوا و دریازه $(0, D_{\max})$ محدب است. \square

به طور تحلیلی می توان توابع اعوجاج نرخ را برای منابع و معیارهای اعوجاج ساده، مانند مثال قبل محاسبه کرد. بعلاوه وقتی روش های تحلیلی، نامناسب یا غیرعملی باشند، می توان از الگوریتم های تکراری همگرا که برای پیاده سازی در رایانه های رقمی مناسب هستند، استفاده نمود. پس از این که $R(D)$ برای هر منبع بی حافظه و هر معیار اعوجاج نک حرفي[†] محاسبه شود، طبق قضیه رمزگاری منبع برای هر $0 < \epsilon$ ، یک r و یک رمز بلوکی به طول r و نرخ $\epsilon < R(D) + d(\mathbf{Q})$ وجود دارد طوری که $d(\mathbf{Q})$ ، متوسط اعوجاج در حرف، در رابطه $\epsilon + d(\mathbf{Q}) \leq D + R(D)$ صدق می کند. یک نتیجه عملی و مهم از این قضیه و قضیه رمزگاری با نویز این است که با شرط این که ظرفیت کانال $\epsilon + R(D) > C$ باشد، در رمزبردار می توان خروجی منبع را با هر احتمال خطای به اندازه دلخواه کوچک بازیابی کرد. این نتیجه به عنوان قضیه انتقال اطلاعات^۱ شناخته می شود.

[†] یک معیار اعوجاج نک حرفي معیاری است که در آن اعوجاج مربوط به یک بلوک از حرفها (یا نمادها) جمع اعوجاج های هر بک از حرف های (یا نمادهای) آن بلوک باشد.

1) *Information transmission theorem*

۴.۳.۶ استفاده از نظریه اطلاعات

«نظریه اطلاعات» ابزارهای اصلی مورد نیاز برای نمایش و کار با اطلاعات را به صورت مستقیم و کمی فراهم می‌آورد. در این بخش در جستجوی کاربرد این ابزارها در مسأله خاص فشرده سازی تصویر هستیم. چون فرض اساسی نظریه اطلاعات این است که تولید اطلاعات را می‌توان به صورت یک فرآیند احتمالی مدل کرد، در ابتدا یک مدل آماری برای فرآیند تولید تصویر ایجاد می‌کنیم.

مثال: مسأله تخمین محتوای اطلاعات (یا آنتروپی) تصویر ۸ بیتی زیر را در نظر بگیرید:

21	21	21	95	169	243	243	243
21	21	21	95	169	243	243	243
21	21	21	95	169	243	243	243
21	21	21	95	169	243	243	243

یک روش نسبتاً ساده این است که مدل ویژه‌ای را برای منبع فرض نموده و آنtronپی تصویر را براساس آن مدل محاسبه کنیم. برای مثال می‌توان فرض کرد که تصویر بوسیله یک «منبع خاکستری ۸ بیتی» فرضی که به طور متوالی پیکسل‌های مستقل آماری را طبق یک قانون احتمالی از پیش تعریف شده گسیل می‌کند، تولید شده است. در این حالت نمادهای منبع، سطوح خاکستری هستند و الفبای منبع مرکب از ۲۵۶ نماد می‌باشد. اگر فرض شود که احتمالات نمادها شکل استانداردی (مثل آگوستی) دارند، محتوا یا آنtronپی اطلاعات هر پیکسل تصویر را می‌توان با استفاده از معادله $(6.3-3)$ محاسبه کرد. برای نمونه در حالت توزیع یکنواخت، نمادهای منبع هم احتمال هستند و منبع با آنtronپی ۸ بیت بر پیکسل توصیف می‌شود. پس آنtronپی کل تصویر 8×4 پیکسلی، ۲۵۶ بیت است. این تصویر ویژه تنها یکی از $2^{8 \times 4} = 2^{256}$ یا (10^7) تصویر 8×4 پیکسلی هم احتمال که می‌توان بوسیله منبع تولید کرد، می‌باشد.

یک روش دیگر برای تخمین محتوای اطلاعات، مدل سازی منبع براساس فراوانی نسبی سطوح خاکستری در تصویر مورد بررسی است. یعنی تصویر مشاهده شده را می‌توان به عنوان نمونه‌ای از رفتار منبع مولدش تعبیر کرد. چون تصویر مشاهده شده تنها نمونه موجود از رفتار منبع می‌باشد، مدل سازی احتمالات نمادهای منبع با استفاده از بافت نگار سطح خاکستری تصویر نمونه، امری منطقی است: یک تخمین از آنtronپی منبع، که تخمین مرتبه اول^۱ خوانده می‌شود، را می‌توان با معادله $(6.3-3)$ ایجاد کرد.

۱) First-order estimate

۳.۶ عناصر نظریه اطلاعات ۴۱۳

سطح خاکستری	تعداد	احتمال
21	12	3/8
95	4	1/8
169	4	1/8
243	12	3/8

محاسبه کرد. در این مثال تخمین مرتبه اول ۱.۸۱ بیت بر پیکسل است. بنابراین آنتروپی منبع و/یا تصویر تقریباً ۱.۸۱ بیت بر پیکسل یا جمماً ۵۸ بیت است.

می‌توان با بررسی فراوانی نسبی بلوک‌های پیکسلی در تصویر نمونه، تخمین‌های بهتری از آنتروپی منبع مولد سطح خاکستری تصویر نمونه را محاسبه کرد که در اینجا متنظر از بلوک، گروهی از پیکسل‌های مجاور هم است. وقتی ابعاد بلوک به بی‌نهایت میل کند، تخمین به آنتروپی واقعی منبع میل می‌کند. (این نتیجه را می‌توان با روایی که در بخش ۳.۳.۶ برای اثبات درستی قضیه رمزنگاری بدون نویز استفاده شد، بدست آورد) بنابراین با فرض این که تصویر نمونه خط به خط و انتهای به ابتداء متصل است، می‌توان فراوانی نسبی زوج پیکسل‌ها (یعنی گسترش دوم منبع) را محاسبه کرد:

زوج سطوح خاکستری	تعداد	احتمال
(21, 21)	8	1/4
(21, 95)	4	1/8
(95, 169)	4	1/8
(169, 243)	4	1/8
(243, 243)	8	1/4
(243, 21)	4	1/8

تخمین آنتروپی (در اینجا نیز با استفاده از معادله $2.5/2 = 1.25$ بیت بر پیکسل است، که عمل تقسیم بر دو به خاطر لحاظ دو پیکسل در هر نماد می‌باشد. چون این تخمین با استفاده از فراوانی بلوک‌های دو پیکسلی بدست آمد، تخمین مرتبه دوم^۱ آنتروپی منبع خوانده می‌شود. گرچه تخمین‌های مرتبه ۲، ۴ و بالاتر بازهم تقریب‌های بهتری از آنتروپی منبع را فراهم می‌آورند، با این حال همگرایی آنها به آنتروپی واقعی منبع کُند است و به محاسبات زیادی نیاز دارند. برای نمونه، تصویر ۸ بیتی،^۲ (۲^۸) یعنی ۶۵,۵۳۶ زوج نماد ممکن دارد که باید فراوانی نسبی آنها محاسبه شوند. اگر بلوک‌ها ۵ پیکسلی در نظر گرفته شوند، تعداد ۵ تابی‌های ممکن^۳ (۲^{۱۲}) یعنی تقریباً ۱۰^{۳۶} است.

۱) Second-order estimate

۴۱۲ پردازش تصویر رنگی

گرچه محاسبه آتروپی واقعی یک تصویر مشکل است، تخمین های نظری تخمین های مثال قبل، اطلاعی در مورد قابلیت فشرده سازی تصویر فراهم می آورند. مثلاً تخمین مرتبه اول آتروپی یک کران پایین برای فشرده سازی است که با استفاده از رمزنگاری طول متغیر بدست می آید. (از بخش ۱.۱.۶ بیاد آورید که رمزنگاری طول متغیر برای کاهش افزونگی های رمزنگاری استفاده می شود.) بعلاوه تفاضل بین تخمین های بالا مرتبه تر آتروپی و تخمین مرتبه اول، نشانگر وجود یا عدم وجود افزونگی های بین پیکسلی است. یعنی آشکار می سازد که آیا پیکسل های یک تصویر، مستقل آماری هستند یا نیستند. اگر پیکسل ها مستقل آماری باشند (یعنی هیچ افزونگی بین پیکسلی وجود نداشته باشد)، تخمین های بالا مرتبه تر معادل تخمین مرتبه اول هستند، و رمزنگاری طول متغیر، فشرده سازی بهینه را فراهم می آورد.

برای تصویر مثال قبل، با تفاضل عددی بین تخمین های مرتبه اول و دوم می توان نگاشتی ایجاد کرد که $0.56 = 1.25 - 1.81$ بیت بر پیکسل از افزونگی موجود در تصویر را حذف کند.

مثال: در نظر بگیرید که پیکسل های تصویر مثال قبل را نگاشت نموده ایم تا نمایش زیر ایجاد شود:

21	0	0	74	74	74	0	0
21	0	0	74	74	74	0	0
21	0	0	74	74	74	0	0
21	0	0	74	74	74	0	0

این تصویر با تکرار ستون اول تصویر اولیه و تفاضل حسابی بین ستون های مجاور برای سایر عناصر ایجاد شده است. برای مثال عنصر سطر اول ستون دوم در نمایش جدید (21-21) یعنی صفر است.

توزيع تفاضل حاصل عبارتست از:

تفاضل یا سطح خاکستری	تعداد	احتمال
0	16	1/2
21	4	1/8
74	12	3/8

حال اگر در نظر بگیریم که آرایه نگاشت شده بوسیله یک "منبع تفاضلی" تولید شود، دوباره می توانیم از معادله (۳-۳.۶) برای محاسبه تخمین مرتبه اول آتروپی آرایه استفاده نماییم که برابر ۰.۴۱ بیت بر پیکسل می شود. بنابراین با رمزنگاری طول متغیر تصویر تفاضلی، می توان تصویر اولیه را تنها با ۰.۴۱ بیت

۴.۶ فشرده سازی بدون خطأ

۲۱۵

بر پیکسل یعنی جمعاً حدود 46 بیت نمایش داد. این مقدار بزرگتر از تخمین مرتبه دوم آنتروپی، یعنی ۲.۵ بیت بر پیکسل که در مثال قبل محاسبه شد، می باشد. بنابراین دانستیم که می توان نگاشت بهتری نیز پیدا نمود.

مثال های قبل روش می سازند که تخمین مرتبه اول آنتروپی، حتماً نرخ رمز کمینه تصویر نیست. علت این است که در حالت کلی پیکسل های تصویر مستقل آماری نیستند. همان طور که در بخش ۲.۶ ذکر شد، به فرآیند کمینه سازی آنتروپی تصویر، رمز نگاری منبع می گویند. در حالت بدون خطأ این فرآیند شامل دو عمل نگاشت و رمزگذاری است. اگر اتلاف اطلاعات قابل تحمل باشد، این فرآیند شامل مرحله چندی سازی نیز خواهد بود.

با استفاده از ابزارهای نظریه اطلاعات می توان به مسأله کمی پیچیده تر فشرده سازی با اتلاف تصویر نیز پرداخت. در هر حال در این مورد، نتیجه اصلی، قضیه رمز نگاری منبع است. همان طور که در بخش ۶.۳. نشان داده شد، این قضیه آشکار می کند که هر منبع بی حافظه را می توان با نرخ $R < R(D)$ که باشد، رمز کرد چنان که متوسط اعوجاج برناماد کوچکتر از D باشد. کاربرد صحیح این نتیجه در فشرده سازی با اتلاف تصویر نیازمند شناسایی یک مدل منبع مناسب، طرح یک معیار اعوجاج معنادار و محاسبه تابع اعوجاج نرخ $R(D)$ نهایی می باشد. مرحله اول این فرآیند قبل از بررسی شده است. با استفاده از یک معیار عینی مشابه از بخش ۶.۱.۰، می توان از مرحله دوم این فرآیند نیز بر احتی عبور کرد. مرحله نهایی، شامل پیدا کردن ماتریس **Q** است طوری که معادله $(6.6-3)$ را با قیود اعمال شده با معادلات $(6.6-3)-(6.6-24)$ کمینه سازد. متأسفانه این کار مشکلات ویژه ای دارد - و تنها حالات کمی از هر مورد عملی، حل شده اند. یک حالت وقتی است که تصاویر، میدان های تصادفی گوسی هستند و معیار اعوجاج یک تابع خطای مربع وزندار^۱ است. در این حالت رمزگذار بهینه، باید تصویر را به مؤلفه های کارهانز - لزو آن بسط دهد و تمام مؤلفه ها را با خطای میانگین مربع یکسانی نمایش دهد (Davisson[1972]).

۴.۶ فشرده سازی بدون خطأ

در کاربردهای متعددی فشرده سازی بدون خطأ^۲ تنها وسیله مناسب برای کاهش داده ها است. یکی از این

1) Weighted square error function

2) error-free compression

کاربردها بایگانی استناد پزشکی یا تجاری می باشد، که در آنها معمولاً به دلایل قانونی، فشرده سازی با اتلاف ممنوع است. کاربرد دیگر، پردازش تصاویر دریافتی از LANDSAT است، که در این مورد نوع استفاده و هزینه جمع آوری داده ها هر اتلافی را نامطلوب می سازد. کاربرد دیگر پرتونگاری رقمی^۱ است، که اتلاف اطلاعات می تواند دقت تشخیص رابه مخاطره اندازد. در این موارد، و موارد دیگر، کاربرد مورد نظر و طبیعت تصاویر مورد بررسی نیاز به فشرده سازی بدون خطا را مطرح می سازند.

در این بخش بر روی روش های اصلی فشرده سازی بدون خطا که اکنون استفاده می شوند، مرکز می شویم. این روش ها معمولاً نسبت های فشرده سازی بین ۲ تا ۱۰ را فراهم می آورند. گذشته از این کاربرد یکسانی در تصاویر دودویی و خاکستری دارند. همان طور که در بخش ۶.۲ نشان داده شد، روش های فشرده سازی بدون خطا معمولاً ترکیبی از دو عمل نسبتاً مستقل هستند: (۱) طرح یک نمایش دیگر از تصویری که افزونگی های بین پیکسلی آن باید کاهش یابند؛ و (۲) رمزگاری این نمایش طوری که افزونگی های رمزگاری حذف شوند. این مراحل متناظر با مراحل نگاشت و رمزگذاری نماد در مدل رمزگذاری منبع شکل ۶.۶ هستند.

۱۰.۴.۶ رمزگاری طول متغیر

ساده ترین روش فشرده سازی بدون خطای تصویر این است که فقط افزونگی رمزگاری را کاهش دهیم. معمولاً در رمزگاری دودویی طبیعی سطوح خاکستری تصویر، افزونگی رمزگاری وجود دارد. همان طور که در بخش ۱۰.۱ ذکر شد، می توان با رمزگاری سطوح خاکستری طوری که معادله (۶-۱.۶) حداقل شود، افزونگی های رمزگاری را حذف نمود. انجام این کار نیازمند استفاده از یک رمز طول متغیر که کوتاه ترین کلمات رمز ممکن را به محتمل ترین سطوح خاکستری نسبت دهد، می باشد. در اینجا چند روش بهینه و نزدیک بهینه برای ایجاد چنین رمزی را بررسی می کنیم. روابط این روش ها به زبان نظریه اطلاعات نوشته می شوند. نمادهای منبع در عمل ممکن است خود سطوح خاکستری تصویر یا خروجی یک نگاشت سطح خاکستری (مثلاً تفاضل پیکسل ها، طول دنباله های ثابت و غیره) باشند.

1) Digital radiography

رمزگاری هافمن

متداول ترین روش برای حذف افزونگی رمزگاری روش هافمن است (Huffman[1951]). هنگام رمزگاری تک به تک نمادهای یک منبع اطلاعات، رمزگاری هافمن^۱ کوچکترین تعداد ممکن نماد رمز بر نماد منبع را نتیجه می دهد. از نظر قضیه رمزگاری بدون نویز (بخش ۳.۶ را ببینید)، رمز حاصل برای یک مقدار ثابت^۲ با این قید که نمادهای منبع تک به تک رمز شوند، بهینه است.

مرحله اول در روش هافمن اجرای دنباله ای از کاهش های منبع^۳ است که هر مرحله 'کاهش منبع' با مرتب کردن صعودی یا نزولی احتمال نمادهای منبع و ترکیب دو یا چند نماد با احتمال کمتر در یک نماد که احتمال آن مجموع احتمال های نمادهای مذکور است و در کاهش بعدی منبع جایگزین آن نمادها می گردد، انجام می شود. شکل ۱۱.۶ این فرآیند را در رمزگاری دودویی نمایش می دهد (رمزهای هافمن K گانی را نیز می توان تشکیل داد). در چه ترین قسمت شکل، یک مجموعه فرضی از نمادهای منبع بر حسب مقادیر نزولی احتمال مرتب شده اند. برای انجام اولین کاهش منبع، دواحتمال پایینی یعنی ۰.۰۶ و ۰.۰۴ ترکیب می شوند تا یک "نماد مركب"^۴ با احتمال ۰.۱ تشکیل شود. این نماد مركب و احتمالش در اولین ستون کاهش منبع قرار می گیرند طوری که نمادهای منبع کاهش یافته نیز به ترتیب نزولی احتمال مرتب شوند. سپس این فرآیند تکرار می شود تا آن که به یک منبع کاهش یافته با دو نماد (در راست ترین قسمت) برسیم.

مرحله دوم روال هافمن این است که با شروع از کوچکترین منبع و بازگشت به منبع اولیه، هر منبع کاهش یافته را رمز کنیم. البته کوتاه ترین رمز دودویی برای یک منبع دونمادی، نمادهای ۰ و ۱

نماد	احتمال	کاهش منبع			
		۱	۲	۳	۴
a_2	0.4	0.4	0.4	0.4	0.6
a_6	0.3	0.3	0.3	0.3	0.4
a_1	0.1	0.1	0.2	0.3	
a_4	0.1	0.1	0.1		
a_3	0.06	0.1			
a_5	0.04				

شکل ۱۱.۶ کاهش های منبع هافمن

1) Huffman coding

2) Source reduction

3) Compound symbol

هستند. همان طور که شکل ۶.۱۲ نشان می دهد، این نمادها به دو نماد سمت راست منتب می شوند (انتساب دلخواه است؛ عکس ترتیب انتساب ۰ و ۱ نیز دقیقاً به همان خوبی کار می کند). چون نماد با احتمال 6×10^{-4} ، با ترکیب دو نماد از منبع کاهش یافته سمت چپ تولید شد، صفر استفاده شده برای رمزگردان آن به هر دونماد مولدهش نیز منتب می شود، و سپس ۰ و ۱ به طور دلخواه به هریک از آن دو افزوده می شوند تا آن دورا ز هم مشخص سازند. آنگاه این عمل برای هر منبع کاهش یافته تکرار می شود تا آن که به منبع اولیه برسیم. رمزهای هافمن نهایی در سمت چپ شکل ۶.۱۲ ظاهر می شوند. طول متوسط این رمز برابر

$$L_{avg} = (0.4)(1) + (0.3)(2) + (0.1)(3) + (0.1)(4) + (0.06)(5) + (0.04)(5) \\ = 2.2 \text{ بیت بر نماد}$$

و آنتروپی منبع برابر ۰.۹۷۳ است. طبق معادله (۶.۳-۲۰) بازده این رمز هافمن 2^{14} بیت بر نماد است.

روال هافمن برای یک مجموعه از نمادها و احتمالات با این قيد که نمادها تک تک رمز شوند، یک رمز بهینه ایجاد می کند. پس از این که رمز هافمن طراحی شد، رمزگذاری و / یا رمزبرداری به روش ساده جدول جستجو انجام می شود. خود رمز یک رمز بلوکی آنی و قابل رمزبرداری یکتا است. این رمز یک رمز بلوکی^۱ خوانده می شود، زیرا هر نماد منبع به یک دنباله معین از نمادهای رمز نگاشته می شود. این رمز آنی^۲ است، زیرا هر کلمه رمز در یک رشته از نمادهای رمز را می توان بدون مراجعه به نمادهای بعدی،

منبع اولیه			کاهش منبع			
احتمال	نماد	رمز	۱	۲	۳	۴
a_2	0.4	1	0.4 1	0.4 1	0.4 1	0.6 0
a_6	0.3	00	0.3 00	0.3 00	0.3 00	0.4 1
a_1	0.1	011	0.1 011	0.2 010	0.3 01	
a_4	0.1	0100	0.1 0100	0.1 011		
a_3	0.06	01010	0.1 0101			
a_5	0.04	01011				

شکل ۶.۱۲ روال انتساب رمز هافمن

۴.۶ فشرده سازی بدون خطای ۴۱۹

رمزبرداری کرد. این رمز قابل رمزبرداری یکتا^۱ است، زیرا هر رشته از نمادهای رمز را تنها به یک صورت می توان رمزبرداری کرد. بنابراین هر رشته از نمادهای رمز هافمن رامی توان با پیدا کردن یکی یکی نمادهای رمز از چپ به راست رمزبرداری کرد. اگر برای رمز دودویی شکل ۶.۲ رشته رمز ۰۱۰۱۰۰۱۱۱۱۰۰ از چپ به راست پیمایش شود، اولین کلمه رمز معتبر به صورت ۰۱۰۱۰ شناخته و رمزبرداری می شود که این رمز مربوط به نماد a_3 می باشد. رمز معتبر بعدی ۰۱۱ است که متاظر با نماد a_1 می باشد. ادامه این روش آشکار می سازد که کل پیام رمزبرداری شده $a_3a_1a_2a_2a_6$ می باشد.

سایر رمزهای طول متغیر نزدیک بهینه

وقتی تعداد زیادی از نمادها باید رمز شوند، ایجاد یک رمز هافمن دودویی بهینه کار سنگینی است. در حالت کلی با J نماد منبع، باید J - کاهش منبع (شکل ۶.۱۱ را ببینید) و J - انتساب رمز انجام شود (شکل ۶.۱۲ را ببینید). بنابراین ایجاد رمز هافمن بهینه برای یک تصویر با ۲۵۶ سطح خاکستری نیازمند ۲۵۴ کاهش منبع و ۲۵۴ انتساب رمز است. برای کاهش پیچیدگی محاسباتی گاهی فداکردن بهره رمزنگاری در ازای سادگی ساختمان رمز الزامی است.

جدول ۶.۵. چهار رمز طول متغیر که چنین مصالحه ای را فراهم می آورند، نشان می دهد. توجه شود که متوسط طول رمزهافمن - آخرین سطر جدول - کمتر از سایر رمزهای جدول است. رمز دودویی طبیعی، بزرگترین طول متوسط را دارد. بعلاوه نرخ رمز $40.5/4$ بیت بر نماد بدست آمده با روش هافمن به آتروپی $4/4$ بیت بر نماد منبع نزدیک است. آتروپی با استفاده از معادله $(6.3.3)$ محاسبه شده و در پایین جدول قرار دارد. گرچه سایر رمزهای جدول ۶.۵ به بازده رمزنگاری هافمن نمی رستند، اما ساختمان ساده تری دارند. همه آنها مانند روش هافمن کوتاه ترین کلمات رمز را به محتمل ترین نمادهای منبع منسب می کنند.

ستون پنجم جدول ۶.۵ یک گونه تغییریافته روش رمزنگاری هافمن موسوم به رمزنگاری هافمن ناقص^۱ را نشان می دهد. یک رمز هافمن ناقص با رمزنگاری هافمن $\#$ تا از محتمل ترین نمادهای منبع بازای یک عدد صحیح مثبت J ، $J < \#$ ، تولید می شود. یک رمز پیشوند^۲ که به دنبال آن یک رمز طول ثابت مناسب قرار دارد، برای نمایش سایر نمادهای منبع استفاده می شود. در جدول ۶.۵، به طور

جدول ۶.۵ رمزهای طول متغیر

نام منبع	احتمال	رمز دودویی	هافمن ناقص	هافمن	B_2 رمز	انتقال دودویی	انتقال هافمن
<i>Block 1</i>							
a_1	0.2	00000	10	11	C00	000	10
a_2	0.1	00001	110	011	C01	001	11
a_3	0.1	00010	111	0000	C10	010	110
a_4	0.06	00011	0101	0101	C11	011	100
a_5	0.05	00100	00000	00010	C00C00	100	101
a_6	0.05	00101	00001	00011	C00C01	101	1110
a_7	0.05	00110	00010	00100	C00C10	110	1111
<i>Block 2</i>							
a_8	0.04	00111	00011	00101	C00C11	111 000	00 10
a_9	0.04	01000	00110	00110	C01C00	111 001	00 11
a_{10}	0.04	01001	00111	00111	C01C01	111 010	00 110
a_{11}	0.04	01010	00100	01000	C01C10	111 011	00 100
a_{12}	0.03	01011	01001	01001	C01C11	111 100	00 101
a_{13}	0.03	01100	01110	10 0000	C10C00	111 101	00 1110
a_{14}	0.03	01101	01111	10 0001	C10C01	111 110	00 1111
<i>Block 3</i>							
a_{15}	0.03	01110	01100	10 0010	C10C10	111 111 000	00 00 10
a_{16}	0.02	01111	010000	10 0011	C10C11	111 111 001	00 00 11
a_{17}	0.02	10000	010001	10 0100	C11C00	111 111 010	00 00 110
a_{18}	0.02	10001	001010	10 0101	C11C01	111 111 011	00 00 100
a_{19}	0.02	10010	001011	10 0110	C11C10	111 111 100	00 00 101
a_{20}	0.02	10011	011010	10 0111	C11C11	111 111 101	00 00 1110
a_{21}	0.01	10100	011011	10 1000	C00C00C00	111 111 110	00 00 1111
انتروپی	4.0						
طول متوسط	5.0	4.05	4.24	4.65	4.59	4.13	

دلخواه ۱۲ برابر ۱۲ انتخاب شد و رمز پیشوند به عنوان سیزدهمین کلمه رمزهای هافمن تولید شد. یعنی یک نماد پیشوند^۱ که احتمالش مجموع احتمالات نمادهای a_{21} تا a_1 بود، به عنوان نماد سیزدهم در رمزگاری هافمن محتمل ترین ۱۲ نماد منبع، به کار رفت. آنگاه ۹ نماد با قیمانده با ترکیب رمز پیشوند، که ۱۰ تعیین شد، و یک مقدار دودویی ۴ بیتی که برابر زیرنویس نماد منهاهای ۱۳ بود، رمز شدند.

ستون ششم جدول ۶.۵ دومین رمز طول متغیر نزدیک بهینه موسوم به رمز B را نشان می‌دهد. وقتی احتمال نمادهای منبع از یک قانون نمایی به شکل

$$P(a_i) = c j^{-\beta} \quad (1-4.6)$$

1) Prefix symbol

تبعت نمایند، این رمز نزدیک بهینه است که β یک ثابت مثبت و ثابت تراز سازی برابر $j = 1/\sum_{j=0}^l \beta_j$ می باشد. برای مثال توزیع طول دنباله ها در نمایش دودویی یک متن نوعی تقریباً نمایی است. همان طور که جدول ۶.۵ نشان می دهد، هر کلمه رمز از بیت های ادامه^۱، که با C نشان داده می شوند، و بیت های اطلاعات^۲ که اعداد دودویی طبیعی هستند، تشکیل می شود. تنها هدف از کاربرد بیت های ادامه جداسازی کلمات رمز از یکدیگر است. بنابراین این بیت ها برای کلمات رمز متوالی در یک رشته به تناوب ۰ و ۱ می شوند. رمز B_2 در جدول ۶.۵ رمز B_2 ^۳ خوانده می شود، زیرا به ازای هر بیت ادامه از دو بیت اطلاعات استفاده می شود. دنباله رمزهای B_2 متناظر با رشته نماد منبع a_1, a_2, a_3 بسته به این که اولین بیت ادامه ۰ یا ۱ فرض شود، ۰۱۰ ۰۱۰ ۰۱۰ ۰۰۱ ۰۱۰ ۱۰۱ ۰۰۰ ۰۱۰ ۰۱۰ ۰۰۱ ۱۱۰ ۰۰۱ ۱۱۰ ۰۰۱ ۱۰۱ می باشد. دو رمز طول متغیر دیگر در جدول ۶.۵ با عنوان رمزهای انتقالی^۴ اشاره می شوند. یک رمز انتقالی با انجام چهار مرحله زیر ایجاد می شود: (۱) مرتب کردن نمادهای منبع طوری که مقادیر احتمالشان یکنواخت باشند، (۲) تقسیم بندی کل نمادها به بلوک های هم اندازه، (۳) رمز نگاری نظیر به نظری یکسان عناصر تمام بلوک ها، و (۴) افزودن نمادهای خاص انتقال به بالا^۵ / یا انتقال به پایین^۶ به هر بلوک برای شناسایی آن. هر بار که یک نماد انتقال به بالا (یا انتقال به پایین) در رمزبردار تشخیص داده شود، این نماد نسبت به بلوک مرجع پیش تعریف شده، یک بلوک به بالا (یا پایین) جایجا می شود.

برای تولید رمز انتقالی دودویی ۳ بیتی در ستون هفتم جدول ۶.۵، ابتدا ۲۱ نماد منبع طبق احتمالات وقوع آنها مرتب شده و سپس به سه بلوک هفت نمادی تقسیم می شوند. آنگاه تک تک نمادهای بلوک بالایی (a_1 تا a_3) - که بلوک مرجع در نظر گرفته می شود - با رمزهای دودویی ۰۰۰ تا ۱۱۱ رمز می شوند. رمز دودویی هشتم (۱۱۱) در بلوک مرجع قرار نمی گیرد، در عوض به عنوان کلمه کنترل انتقال به بالا که شناسه سایر بلوک ها است، استفاده می شود (در این مورد نماد انتقال به پایین استفاده نمی شود). آنگاه نمادهای دو بلوک دیگر بوسیله ترکیب یک یا دو نماد انتقال به بالا و رمزهای دودویی مورد استفاده در رمزگذاری بلوک مرجع، رمز می شوند. مثلاً نماد منبع a_1 به صورت ۱۱۱ ۱۱۱ ۰۰۰ رمز می شود.

رمز انتقالی هافمن در ستون هشتم جدول ۶.۵ به روش مشابهی تولید می شود. تفاوت اصلی در انتساب یک احتمال به نماد انتقال قبل از رمز نگاری هافمن بلوک مرجع می باشد. معمولاً این انتساب با

1) Continuation bit

2) Information bit

3) B_2 -code

4) Shift code

5) Shift -up

6) Shift -down

جمع احتمالات تمام نمادهای غیر از بلوک مرجع، یعنی با استفاده از همان مفهوم استفاده شده در تعریف نماد پیشوند در رمز هافمن ناقص، انجام می شود. در اینجا جمع روی نمادهای a_1, a_2, \dots, a_n گرفته می شود و برابر $3^9 = 1024$ است. بنابراین نماد انتقال محتمل ترین نماد است و یکی از کوتاه ترین کلمات رمز هافمن (00) به آن متنسب می شود.

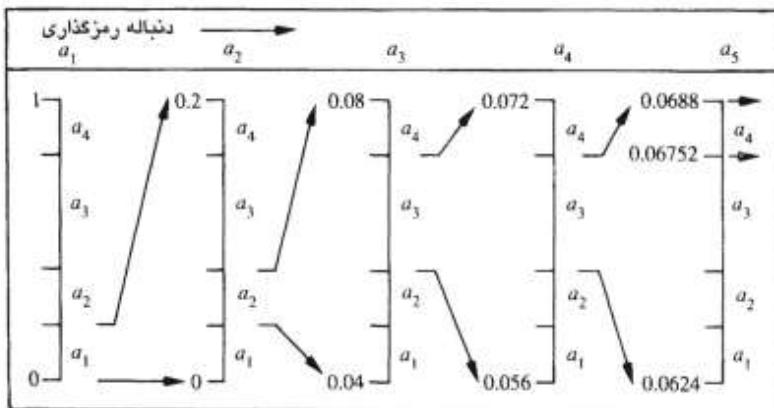
رمزنگاری حسابی

برخلاف رمزهای طول متغیر مشروح در فوق، رمزنگاری حسابی¹ رمزهای غیربلوکی تولید می کند. در رمزنگاری حسابی، که می توان سرمنش آن را تا کار Elias [1963] را بینید، کوچکتر می شود و تعداد واحدهای اطلاعات (مثل آبیت های) مورد نیاز برای نمایش آن ۰ و ۱ را تعریف می کند. وقتی تعداد نمادها در پیام افزایش می یابد، بازه ای از اعداد حقیقی بین ۰ و ۱ را تعیین می کند. چون این روش برخلاف روش هافمن نیاز ندارد که هر نماد منبع به یک تعداد صحیح از نمادهای رمز ترجمه شود (یعنی این که نمادها تک تک رمز شوند)، لذا به کران تعیین شده بوسیله قضیه رمزنگاری بدون نیاز در بخش ۶.۳.۳. دست می یابد (البته فقط در جنبه نظری). شکل ۶.۱۳. فرآیند پایه رمزنگاری حسابی را نشان می دهد. در اینجا دنباله یا پیام ۵ نمادی a_1, a_2, a_3, a_4, a_5 از یک منبع چهار نمادی، رمز می شود. در شروع فرآیند رمزنگاری فرض می شود که پیام تمام بازه $[0, 1]$ را اشغال نماید. همان طور که جدول ۶.۶ نشان می دهد، در ابتدا این بازه براساس احتمالات نمادهای منبع به چهار ناحیه تقسیم می شود. مثلاً نماد a_1 به زیر بازه $[0, 0.2]$ مربوط می باشد. چون اولین نماد پیام است که رمز می شود، بازه پیام در ابتدا به زیر بازه $[0, 0.2]$ کوچک می شود. بنابراین در شکل ۶.۱۳.، زیر بازه $(0, 0.2)$ به اندازه ارتفاع کامل نمودار، اتبساط می یابد و نقاط پایانی ارتفاع نمودار با مقادیر پایانی این زیر بازه برحسب زده می شوند. آنگاه این زیر بازه طبق احتمالات نمادهای منبع اولیه تقسیم می شود و فرآیند با نماد پیام بعدی ادامه می یابد. در این روش نماد a_2 زیر بازه را به $(0.04, 0.08)$ کوچک می کند، آن را به بازه $[0.056, 0.072]$

1) Arithmetic coding

۴۴۳

۴.۶ فشرده سازی بدون خطأ



شکل ۱۳.۶ روال رمزگاری حسابی

کوچکتر می‌کند و همین طور تا آخر، نماد پیام آخری، که باید به عنوان شاخص ویژه^۱ پایان پیام^۱ ذخیره شود، محدوده را به بازه $[0.06752, 0.0688]$ کوچک می‌کند. البته از هر عدد درون این زیربازه مثلاً ۰.۰۶۸ می‌توان برای نمایش پیام استفاده کرد.

در پیام رمز حسابی شده شکل ۱۳.۶، سه رقم دهدھی برای نمایش پیام پنج نمادی استفاده می‌شود که معادل $\frac{3}{5}$ یا ۰.۶ رقم دهدھی بر نماد منبع است و به خوبی با آنتروپی منبع، که طبق معادله $(3-3.6)$ برابر ۰.۵۸ رقم دهدھی با واحد دهگانی برناماد می‌باشد، قابل مقایسه است. وقتی طول رشته‌ای که باید رمز شود، افزایش می‌یابد، رمز حسابی حاصل به کران تعیین شده بوسیله قضیه رمزگاری بدون تویز نزدیک می‌شود. در عمل دو عامل باعث می‌شوند که کارآئی رمزگاری تواند به آن کران برسد: (۱) افزودن شاخص پیام که برای جداسازی یک پیام از دیگری لازم است؛ و (۲) استفاده از محاسبات با دقیق محدود. در پاده سازیهای عملی رمزگاری حسابی با کاربرد روش‌های مقیاس کردن و گردکردن

جدول ۶.۶ مثالی از رمزگاری حسابی

نماد منبع	احتمال	زیربازه ابتدایی
a_1	0.2	$[0.0, 0.2)$
a_2	0.2	$[0.2, 0.4)$
a_3	0.4	$[0.4, 0.8)$
a_4	0.2	$[0.8, 1.0)$

1) End-of- message indicator

(Langdon, Rissanen[1981]) عامل دوم اجرا می شود. عمل مقیاس کردن قبل از تقسیم هر زیربازه طبق احتمالات نماد، آن را به زیربازه^۱ [0, 1] تراز می کند. عمل گردکردن تضمین می کند که قطع کردنها مربوط به محاسبات با دقت محدود مانع نمایش دقیق زیربازه های رمزگاری نمی شود.

۴.۴.۶ رمزگاری بیت-صفحه ای

پس از بررسی روش های اصلی حذف افزونگی رمزگاری، اکنون یکی از چند روش فشرده سازی بدون خط را که افزونگی های بین پیکسلی تصویر را نیز حذف می کند، بررسی می نماییم. این روش که رمزگاری بیت-صفحه ای^۱ خوانده می شود، بر اساس تجزیه یک تصویر چندسطوحی (تکریگ یا رنگ) به مجموعه ای از تصاویر دودویی و سپس فشرده سازی هر یک از آن تصاویر دودویی با یکی از چند روش فشرده سازی دودویی می باشد. در این بخش متدولوژی روش های تجزیه را شرح می دهیم و مروری نیز بر تعدادی از روش های فشرده سازی متدوال خواهیم داشت.

تجزیه بیت-صفحه ای

سطح خاکستری یک تصویر خاکستری m بیتی رامی توان به شکل چند جمله ای مبنای دوی

$$a_{m-1} 2^{m-1} + a_{m-2} 2^{m-2} + \dots + a_1 2^1 + a_0 2^0. \quad (2-4.6)$$

نمایش داد. بر اساس این خاصیت یک روش ساده برای تجزیه تصویر به مجموعه ای از تصاویر دودویی جداسازی m ضریب چند جمله ای در m بیت-صفحه می باشد. همان طور که در فصل ۴ ذکر شد، بیت-صفحه مرتبه صفر از بیتها یا ضرایب a هر پیکسل تشکیل می شود و بیت-صفحه مرتبه $(m-1)$ ام از بیتها یا ضرایب a تشکیل می شود. به طور کلی هر بیت-صفحه یکی از شماره های ۰ تا $1-m$ را دارد و هر پیکسل آن برابر ضریب متناظر از چند جمله ای دودویی مربوط به پیکسل متناظرش در تصویر اصلی می باشد. عیب ذاتی این روش این است که در بعضی موارد تغییرات کوچک در سطح خاکستری اثر زیادی بر تغییرات بیت-صفحه ها دارد. مثلاً اگر یک پیکسل با شدت 127 (01111111) مجاور پیکسلی با شدت 128 (10000000) باشد، تمام بیت-صفحه ها در آن محل یک گذر از ۰ به ۱ (یا به ۰)

¹) Bit-Plane coding

خواهد داشت.

یک روش تجزیه دیگر (که اثر تغییرات کوچک سطح خاکستری را کاهش می دهد) این است که در ابتدا تصویر را با رمزگری m^1 بیتی نمایش دهیم. رمزگری m بیتی $g_0, g_1, g_2, \dots, g_{m-1}$ متناظر با چند جمله ای (۴-۲) را می توان از روابط

$$\begin{aligned} g_i &= a_i \oplus a_{i+1} & 0 \leq i \leq m-2 \\ g_{m-1} &= a_{m-1} \end{aligned} \quad (4-4)$$

محاسبه کرد. در اینجا علامت \oplus نشانه عمل یا انحصاری است. این رمز خاصیت یکتاپی دارد که کلمه های رمز متوالی آن تنها در یک بیت تفاوت دارند. بنابراین کمتر احتمال دارد که تغییرات کوچک سطح خاکستری، تمام بیت-صفحه ها را متأثر سازند. به عنوان نمونه وقتی سطوح خاکستری ۱۲۷ و ۱۲۸ مجاور هم باشند، تنها صفحه بیتی هفت گذر از ۰ به ۱ خواهد داشت، زیرا رمزهای گری متناظر با ۱۲۷ و ۱۲۸ بترتیب ۱۱۰۰۰۰۰۰ و ۰۱۰۰۰۰۰۰ هستند.

مثال: تصاویر 1024×1024 اشکال ۱۴. (الف) و (ب) برای نمایش عملکرد روش های فشرده سازی مشروح در ادامه این بخش استفاده می شوند. تصویر تک رنگ ۸ بیتی طفل با یک دوربین CCD نفیکی



(الف)

This indenture made this ninth day of our Lord one thousand nine hundred and six between John, son of Andrew Jackson of the County of Franklin of the other part and Stockley Donelson for a sum of two thousand and four hundred and eight dollars hath and by these presents sold alien enfeoffed and confirmed unto his heirs and a certain tract or parcels of land acre or a thousand acre more or less in the County of Franklin bounded as follows.

(ب)

شکل ۱۴.۶ (الف) یک تصویر تک رنگ ۸ بیتی 1024×1024 و (ب) تصویر دودویی 1024×1024

1) Gray code

1) exclusive- OR

President Andrew Jackson بالا تولید شد. تصویر دودویی ضمانت نامه ای که در سال ۱۷۹۶ بوسیله تهیه شده بود، با یک پیمایشگر تخت رقمنی شد. اشکال ۶. ۱۵. ۶ و ۶. ۱۶ هشت بیت-صفحه از تصویر طفل را که رمز دودویی و گری شده اند، نشان می دهند. توجه شود که بیت-صفحه های بالا مرتبه بسیار ساده تر از بیت-صفحه های پایین مرتبه هستند؛ یعنی دارای نواحی یکنواخت بزرگ با جزئیات، تغییرات و تصادفی بودن بسیار کمتری هستند. همچنین بیت-صفحه های رمز گری شده ساده تر از بیت-صفحه های دودویی متناظر شان هستند.

رمزنگاری ناحیه ثابت

یک روش ساده اما مؤثر برای فشرده سازی تصویر دودویی یا بیت-صفحه استفاده از کلمات رمز مخصوص برای شناساندن نواحی بزرگ یک ها یا صفرهای پوسته می باشد. در چنین روشی که رمزنگاری ناحیه ثابت^۱ خوانده می شود، تصویر به بلوک هایی با ابعاد $n \times m$ پیکسل تقسیم می شود که این بلوک ها به عنوان تمام سفید، تمام سیاه یا مختلط گروه بندی می شوند. آنگاه به محتمل ترین یا مکررترین گروه، کلمه رمز یک بیتی ۰ و به دو گروه دیگر رمزهای دو بیتی ۱۰ و ۱۱ منتسب می شود. با این روش تصویر فشرده می شود، زیرا III بیت که معمولاً برای نمایش یک ناحیه ثابت استفاده می گردند، با یک کلمه رمز دو بیتی یا یک بیتی جایگزین می شوند. البته رمز منتسب شده به گروه باشد مختلط به عنوان یک پیشوند استفاده می شود، که به دنبال آن الگوی III بیتی بلوک قرار می گیرد.

هنگام فشرده سازی استاد متی غالباً سفید، یک روش ساده تر این است که نواحی تمام سفید را با "۰" و تمام بلوک های دیگر (از جمله بلوک های تمام سیاه) را با "۱" که به دنبال آن الگوی بیتی بلوک قرار می گیرد، رمز کنیم. این روش پرش از بلوک سفید^۲ (WBS) خوانده می شود. چون تعداد کمی نواحی تمام سیاه مورد انتظار هستند، این نواحی نیز در گروه نواحی باشد مختلط قرار می گیرند؛ در نتیجه این کار، استفاده از کلمه رمز یک بیتی برای نواحی سفید بسیار محتمل ممکن می شود. یک بهبود در این روال (با بلوک های $n \times 1$)، رمز گذاری خطوط تمام سفید با "۰" و رمز گذاری سایر خطوط با "۱" که به دنبال آن دنباله رمز WBS معمولی آن خط قرار دارد، می باشد. بهبود دیگر به کار گیری یک روش تکراری است که تصویر دودویی یا بیت-صفحه را به زیر بلوک های متوالیاً کوچکتر و کوچکتر تجزیه

1) Constant area coding (CAC)

2) White block skipping (WBS)

کند. در بلوکهای دو بعدی، هر بلوک تمام سفید با "۰" رمز می شود و سایر بلوک ها به زیر بلوک هایی تقسیم می شوند که پیشوند "۱" به آنها متناسب گردیده و به روش مشابهی رمز می شوند. یعنی اگر یک زیر بلوک تمام سفید باشد، پیشوند "۱" به رمز آن داده می شود که دلالت بر این دارد که این زیر بلوک حاصل اولین تقسیم بلوک می باشد و به دنبال آن "۰" قرار می گیرد که نشانه تمام سفید بودن آن می باشد. اگر زیر بلوک تمام سفید نباشد، فرآیند تجزیه آنقدر تکرار می شود تا به زیر بلوک هایی با ابعاد بسته کوچک (از پیش گزیده) برسیم و در این مرحله هر زیر بلوک اگر تمام سفید باشد با "۰" و گرنه با "۱" همراه با الگوی بیتی بلوک رمز می شود.

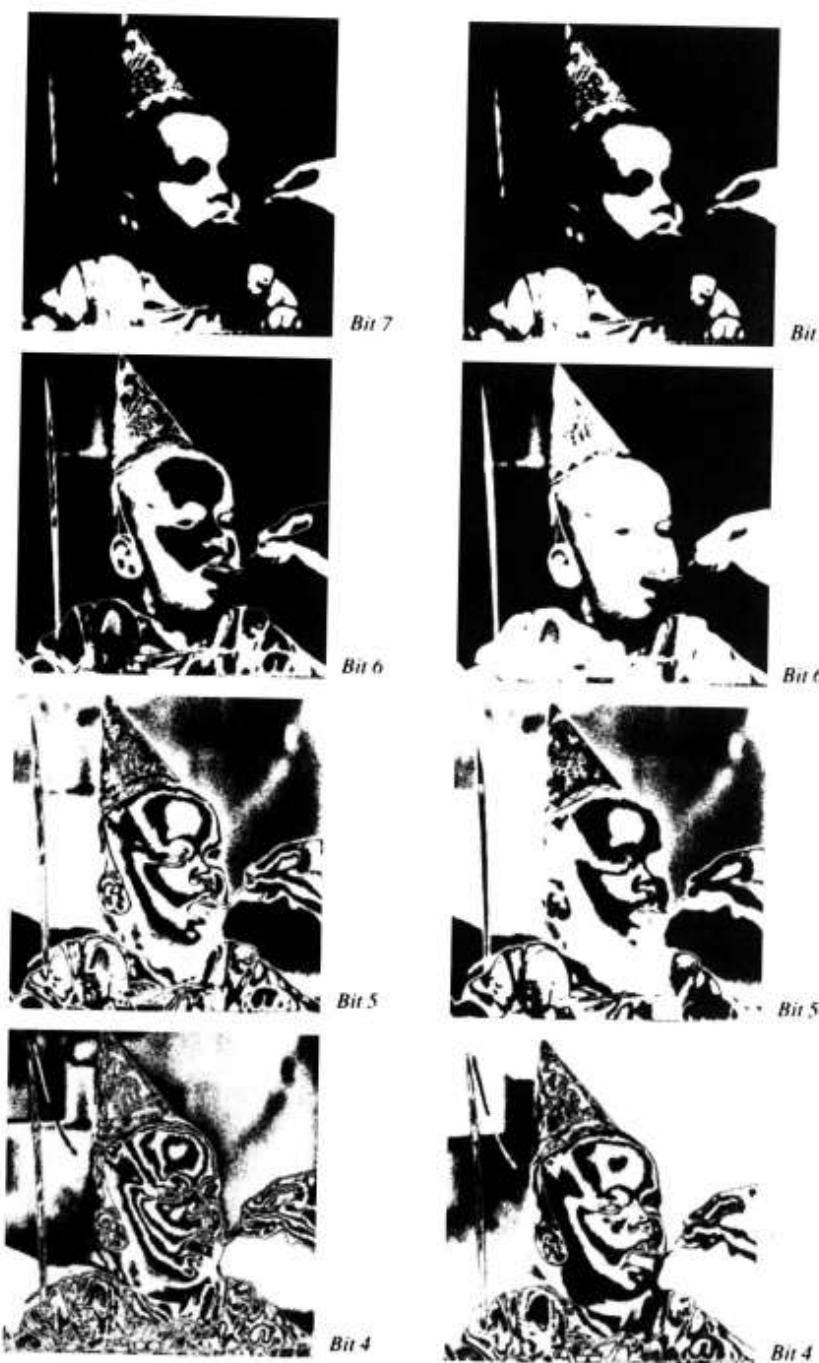
رمزنگاری RL یک بعدی

یک جایگزین مؤثر برای رمزنگاری ناحیه ثابت این است که هر سطر تصویر یا بیت-صفحه را با دنباله ای از طولها نمایش دهیم که این طولها، طول دنباله های پیکسل های سیاه یا سفید متواالی هستند. این روش که با عنوان رمزنگاری طول دنباله^۱ (RL) اشاره می شود، در دهه ۱۹۵۰ ابداع شد و یکی از تعمیم های آن به حالت دو بعدی، روش فشرده سازی استاندارد در رمزنگاری دورنگار^۲ می باشد. اساس روش این است که طول دنباله های پیوسته از "۰" یا "۱" های متواالی را که در پیمایش چپ به راست یک سطر تصویر با آنها مواجه می شویم، رمز کنیم و قراردادی نیز برای تعیین مقدار آن دنباله تعیین نماییم. متدائل ترین روش های تعیین مقدار دنباله عبارتند از: (۱) مشخص کردن مقدار اولین دنباله در هر سطر، یا (۲) فرض این که هر سطر، با یک دنباله سفید که در عمل ممکن است طول آن صفر باشد، شروع می شود.

گرچه رمزنگاری RL روشنی مؤثر برای فشرده سازی تصویر است (مثال بخش ۶.۱.۲ رایبیند)، اما معمولاً با رمزنگاری طول متغیر RL ها، می توان به فشرده سازی بیشتری دست یافت. در واقع می توان RL های سفید و سیاه را جدای از هم و با استفاده از رمزهای طول متغیری که مخصوص آمارگان هر یک از آنها طراحی شده اند، رمزنگاری کرد. برای مثال، با فرض اینکه نماد a نشانگر یک دنباله سیاه به طول ز است، با تقسیم تعداد دنباله های سیاه به طول زیر تعداد کل دنباله های سیاه تصویر، می توان احتمال تولید نماد a بوسیله منبع فرضی RL سیاه را تخمین زد. از جایگذاری این احتمالات در معادله (۶.۳-۳) تخمینی از H_0 ، آتروپی این منبع RL سیاه، به دست می آید. توضیح مشابهی نیز برای H_1 ، آتروپی

1) Run- Length

2) Facsimile (FAX)



شکل ۱۵.۶ بالا مرتبه ترین چهار بیت صفحه از تصویر شکل ۱۴. (الف) که رمز دودویی (ستون چپ) و رمز گری (سمت راست) شده اند.

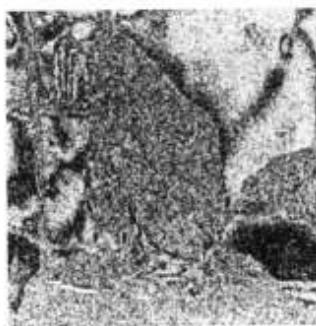
۴۰۶ فشرده سازی بدون خطای ۲۲۹



Bit 3



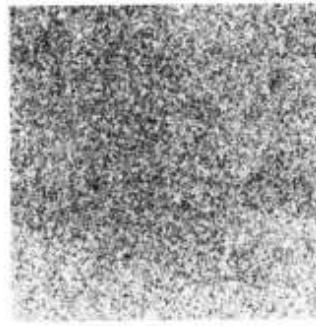
Bit 3



Bit 2



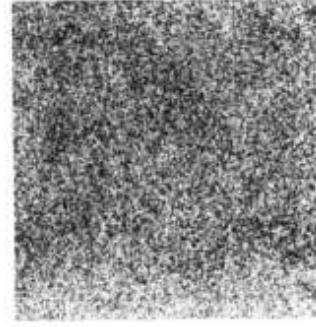
Bit 2



Bit 1



Bit 1



Bit 0



Bit 0

شکل ۱۶.۶ پایین مرتبه ترین چهار بیت صفحه از تصویر شکل ۱۴.۶ (الف) که رمز دودویی (ستون چپ) و رمزگری (ستون راست) شده‌اند.

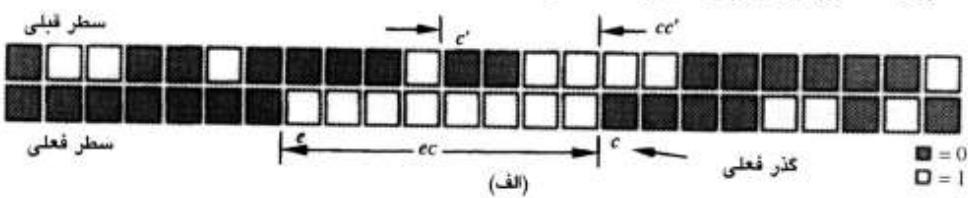
دنباله های سفید، صدق می کند. آنرودی تقریبی RL ها در تصویر عبارتست از:

$$H_{RL} = \frac{H_0 + H_1}{L_0 + L_1} \quad (4-4.6)$$

که متغیرهای L_0 و L_1 بترتیب بیانگر مقادیر متوسط راه های سیاه و سفید هستند. معادله^۱ (۴-۴.۶) تخمینی از متوسط بیت بر پیکسل مورد نیاز برای رمزنگاری RL تصویر دودویی با استفاده از رمز طول متغیر را بدست می دهد.

رمزنگاری RL دو بعدی

با تعمیم مفاهیم رمزنگاری RL یک بعدی به سادگی می توان انواع گوناگونی از روال های رمزنگاری دو بعدی را ایجاد کرد. یکی از روال های معروف تر رمزنگاری آدرس نسبی^۱ (RAC) است، که برایه ۱۷.۶ اصل ردگیری گذر های دودویی موجود در ابتدا و انتهای هر دنباله سفید یا سیاه می باشد. شکل ۱۷.۶ (الف) یک مورد از پیاده سازی این روش را نشان می دهد. توجه شود که ec فاصله c ، گذر فعلی، ta ، گذر قبلی، در خط فعلی می باشد. cc فاصله c تا اولین گذر مشابه بعد از c (در همان جهت پیمایش) روی خط قبلی، که با c' نشان داده می شود، است. اگر $ec \leq cc$ باشد، فاصله رمز آدرس نسبی شده d برابر قرار داده می شود و برای نمایش گذر فعلی در c استفاده می شود. اما اگر $cc < ec$ باشد، برابر cc سطر قبلی



فاصله اندازه کیری شده	فاصله	رمز	محدوده فاصله	رمز $h(d)$
cc'	0	0	1 - 4	0xx
ec or cc' (left)	1	100	5 - 20	10xxxx
cc' (right)	1	101	21 - 84	110xxxxxxx
ec	$d (d>1)$	$111 h(d)$	85 - 340	1110xxxxxxxx
$cc' (c' to left)$	$d (d>1)$	$1100 h(d)$	341 - 1364	11110xxxxxxxx
$cc' (c' to right)$	$d (d>1)$	$1101 h(d)$	1365 - 5460	111110xxxxxxxx

(ب)

شکل ۱۷.۶ نمایشی از رمزنگاری آدرس نسبی

1) Relative address coding

قرار داده می شود.

مشابه رمزنگاری RL، رمزنگاری آدرس نسبی نیز نیازمند اتخاذ یک قرارداد برای تعیین مقادیر دنباله می باشد. بعلاوه باید گذرهای غیرواقعی در ابتدا و انتهای هر خط و نیز یک خط شروع غیرواقعی (مثلًاً یک خط تمام سفید) در ابتدای هر صفحه فرض شوند تا این که مرزهای تصویر به سهولت قابل پردازش باشند. در نهایت چون در اغلب تصاویر واقعی، توزیع های احتمال فواصل RAC یکنواخت نیستند (بخش ۱.۱.۱ را ببینید)، مرحله نهایی فرآیند RAC، این است که فاصله RAC انتخاب شده (کوتاه ترین) و d، فاصله اش، را با استفاده از یک رمز طول متغیر مناسب رمز کنیم. همان طور که شکل ۶.(ب) نشان می دهد، از رمزی مشابه B می توان بهره برد. به کوچک ترین فاصله ها، کوتاه ترین کلمات رمز منتب می شود و سایر فاصله ها با استفاده از دو پیشوند که اولی نشان دهنده کوتاه ترین فاصله RAC است و دومی d را به یک محدوده معین از فاصله ها متنسب می کند، و نیز نمایش دودویی d (که در شکل ۶.۱۷ ب با $\dots \times \times \dots \times$ نشان داده می شود) منهای فاصله پایه آن محدوده، رمز می شوند. مانند شکل ۶.۱۷. الف، اگر ec و cc بترتیب برابر $+8$ و $+4$ باشند، کلمه رمز RAC صحیح 1100011 می باشد. در نهایت اگر $d = 0$ باشد، c مستقیماً زیر C قرار دارد، در حالی که اگر $d = 1$ باشد، ممکن است لازم شود که رمزبردار نزدیک ترین نقطه گذر را تعیین نماید، زیرا رمز 100 مشخص نمی کند که آیا اندازه گیری نسبت به سطر فعلی است یا قبلی.

ردگیری و رمزنگاری منحنی

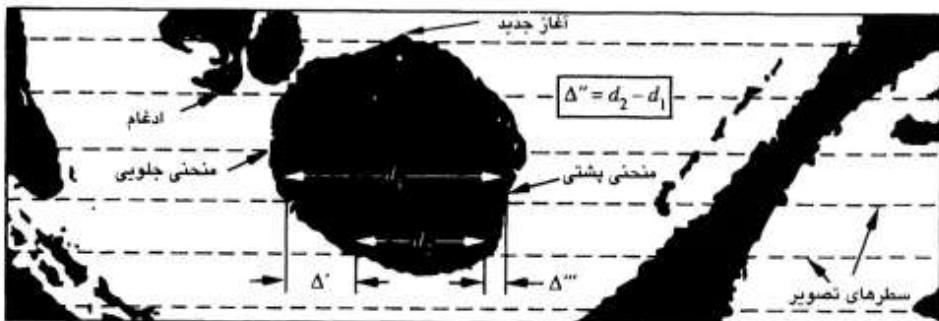
رمزنگاری آدرس نسبی، روشی برای نمایش گذرهای شدت که منحنی های یک تصویر دودویی را می سازند، است. روش دیگر، نمایش هر منحنی با یک مجموعه نقاط مرزی یا با یک نقطه مرزی و مجموعه ای از پیکانها^۱ است. روش دوم، گاهی با عنوان ردگیری مستقیم منحنی^۲ اشاره می شود. در این بخش روش دیگری را که چندی سازی تفاضلی پیشگویی^۳ خوانده می شود و مشخصات اساسی هر دو روش را در خود دارد، معرفی می کنیم. این روش یک روال ردگیری منحنی با پیمایش خطی^۴ می باشد. در چندی سازی تفاضلی پیشگویی، هم‌مان منحنی های جلو و عقب (شکل ۶.۱۸) هر شیء داخل

1) *Directionals*

2) *Direct contour tracing*

3) *Predictive differential quantizing (PDQ)*

4) *Scan line-oriented contour tracing*



شکل ۱۸.۶ عوامل الگوریتم PDQ

تصویر، ردگیری می شوند تا دنباله ای از زوج های (Δ', Δ) بدست آیند. عبارت Δ' تفاصل بین مختصات شروع منحنی های جلویی دو خط متواالی است، و Δ'' تفاصل فاصله های جلو تا عقب منحنی در دو خط متواالی است. این تفاصل ها همراه با پامهای خاصی که بر شروع منحنی های جدید (پیام شروع مجدد^۱) و پایان منحنی های قبلی (پیام ادغام^۲) دلالت دارند، هر شیوه را نمایش می دهند. اگر Δ' با تفاصل مختصات منحنی های عقبی دو خط متواالی، که با Δ''' نشان داده می شود، جایگزین شود، به روش حاصل رمزنگاری دلتای مضاعف^۳ گویند.

پیام های شروع مجدد و ادغام امکان می دهند که زوج های (Δ', Δ) یا (Δ'', Δ) تولید شده برپایه خط پیمایش، به طور مناسبی با زوج های متناظر در سطرهای قبلی و بعدی مرتبط شوند. بدون این پامها رمزبردار قادر نخواهد بود که یک زوج تفاصل را به دیگری ارجاع دهد یا منحنی های تصویر را بدرستی تعیین محل کند. برای اجتناب از رمزگذاری هر دو مختصه سطري و ستونی هر پیام شروع مجدد یا ادغام، اغلب رمزی یکتا برای شناسایی خطوط پیمایشی که حاوی پیکسل های شیشه نیستند، استفاده می شود. مرحله نهایی در هر دو رمزنگاری PDQ و DDC، نمایش Δ' یا Δ''' ، و مختصات شروع مجدد و ادغام با یک رمز طول متغیر مناسب می باشد.

مثال: این بخش را با مقایسه روش های فشرده سازی دودویی توصیف شده پایان می دهیم. هر یک از روش ها برای فشرده سازی تصاویر شکل ۱۴.۶ استفاده شدند. نزد رمز و نسبت های فشرده سازی حاصل در دو جدول ۶.۷ و ۶.۸ گردآوری شده اند. هنگام تفسیر این نتایج توجه کنید که ابتدا تخمین های مرتبه اول (بخش ۶.۳.۴ را بینید) آتروپی طول دنباله های رمزنگاری RL و فواصل PDQ و DDC

1) New start message

2) Merge message

3) Double delta coding (DDC)

۴.۶ فشرده سازی بدون خطای ۴۳۴

جدول ۶ نتایج رمزگاری بیت صفحه ای بدون خطای شکل ۶.۱۴ (الف) : بیت بر پیکسل $H \approx 6.82$

مت	نرخ رمز بیت صفحه (بیت بر پیکسل)								نرخ رمز فشرده سازی	نسبت فشرده سازی
	7	6	5	4	3	2	1	0		
رمزگاری بیت صفحه ای دودویی										
CBC (4 × 4)	0.14	0.24	0.60	0.79	0.99	—	—	—	5.75	1.4: 1
RLC	0.09	0.19	0.51	0.68	0.87	1.00	1.00	1.00	5.33	1.5: 1
PDO	0.07	0.18	0.79	—	—	—	—	—	6.04	1.3: 1
DDC	0.07	0.18	0.79	—	—	—	—	—	6.03	1.3: 1
RAC	0.06	0.15	0.62	0.91	—	—	—	—	5.17	1.4: 1
رمزگاری بیت صفحه ای کوئی										
CBC (4 × 4)	0.14	0.18	0.48	0.40	0.61	0.98	—	—	4.80	1.7: 1
RLC	0.09	0.13	0.40	0.33	0.51	0.85	1.00	1.00	4.29	1.9: 1
PDQ	0.07	0.12	0.61	0.40	0.82	—	—	—	5.02	1.6: 1
DDC	0.07	0.11	0.61	0.40	0.81	—	—	—	5.00	1.6: 1
RAC	0.06	0.10	0.49	0.31	0.62	—	—	—	4.05	1.8: 1

محاسبه و به عنوان نتیری از میزان فشرده سازی قابل دستیابی با روش های رمزگاری طول متغیر بخش ۶.۱ استفاده شدند.

نتایج موجود در جداول ۶ و ۷.۸ نشان می دهند که هر یک از روش ها قادر به حذف مقداری از افزونگی بین پیکسلی است. یعنی نرخ رمزهای حاصل، کمتر از تخمین مرتبه اول آنtronپی تصویر متناظر است. ثابت شد که رمزگاری RL بهترین روش برای تصاویر بیت-صفحه ای است، در حالی که برای فشرده سازی تصویر دودویی روش های دو بعدی (نظیر PDQ، DDC و RAC) کارآیی بهتری دارند. بعلاوه روال نسبتاً ساده رمزگاری گری شکل ۶.۱۴ (الف) کارآیی قابل حصول با رمزگاری را در حدود یک بیت بر پیکسل بهبود داد. درنهایت توجه شود که این پنج روش فشرده سازی تنها قادرند که تصویر تک رنگ را بین ۱ تا ۲ برابر فشرده سازند، در حالی که تصویر دودویی شکل ۶.۱۴ (ب) را بین ۲ تا ۵ برابر فشرده می سازند. همان طور که جدول ۶ نشان می دهد، علت این تفاوت در کارآیی این است که

جدول ۸ نتایج فشرده سازی بدون خطای تصویر دودویی : بیت بر پیکسل $H \approx 0.55$

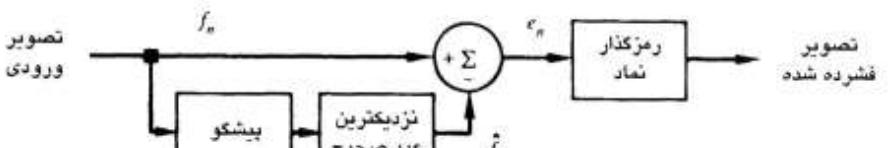
	WBS (1 × 8)	WBS (4 × 4)	RLC	PDQ	DDC	RAC
نرخ رمز (بیت بر پیکسل)	0.48	0.39	0.32	0.23	0.22	0.23
نسبت فشرده سازی	2.1: 1	2.6: 1	3.1: 1	4.4: 1	4.7: 1	4.4: 1

این الگوریتم‌ها قادر نیستند بیت-صفحه‌های پایین مرتبه تصاویر بیت-صفحه‌ای را به خوبی فشرده کنند. در واقع خانه‌های خالی جدول مربوط به مواردی هستند که الگوریتم‌های فشرده سازی باعث افزایش داده‌ها می‌شوند. در این موارد برای نمایش بیت-صفحه از داده‌های خام استفاده شد و لذا نرخ رمز کل بازای هر یک از آنها به اندازه یک بیت بر پیکسل افزایش یافت.

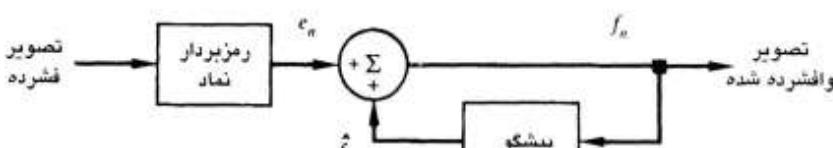
۳.۴.۶ رمزنگاری پیشگویی بی اتلاف

اکنون به یک روش فشرده سازی بدون خطای نیازمند تجزیه تصویر به بیت-صفحه‌ها نمی‌باشد، می‌پردازیم. این روش که معمولاً با عنوان رمزنگاری پیشگویی بی اتلاف^۱ اشاره می‌شود، براساس حذف افزونگی‌های بین پیکسلی از پیکسل‌های نزدیک هم و استخراج رمزنگاری اطلاعات جدید هر پیکسل می‌باشد. اطلاعات جدید پیکسل به عنوان تفاصل بین مقدار واقعی و مقدار پیشگویی شده آن نقطه تعريف می‌شود.

شکل ۱۹.۶ اجزای اصلی سامانه رمزنگاری پیشگویی بی اتلاف را نشان می‌دهد. این سامانه از یک رمزگذار و یک رمزبردار که هر دو پیشگوهای^۲ یکسانی دارند، تشکیل می‌شود. همان‌طور که پیکسل‌های تصویر ورودی، f_n ‌ها، به رمزگذار وارد می‌شوند، پیشگو مقدار هر پیکسل بعدی را براساس



(الف)



(ب)

شکل ۱۹.۶ مدلی از رمزنگاری پیشگویی بی اتلاف: (الف) رمزگذار؛ (ب) رمزبردار

1) Lossless predictive coding

2) Predictor

۴.۶ فشرده سازی بدون خطای ۴۳۵

چند تا از پیکسل های قبلی پیشگویی می کند. آنگاه خروجی پیشگوی به نزدیک ترین عدد صحیح، که با \hat{f}_n نشان داده می شود، گرد شده و تفاضل یا خطای پیشگویی^۱ از رابطه

$$e_n = f_n - \hat{f}_n \quad (5-4.6)$$

بدست می آید که این خطای جهت تولید عنصر بعدی جریان داده فشرده شده و با استفاده از یک رمز طول متغیر (بوسیله رمز گذار نماد) رمز می شود. رمزبردار شکل ۶.۱۹ (ب) "e" را از کلمات رمز طول متغیر دریافت شده، بازسازی کرده و عمل معکوس

$$f_n = e_n + \hat{f}_n \quad (6-4.6)$$

را انجام می دهد.

برای تولید \hat{f}_n می توان از روش های گوناگون محلی، سراسری و تطبیقی (بخش ۵.۱-۶ را بینید) استفاده نمود. در هر حال دربیشتر حالات پیشگویی بوسیله ترکیب خطی m پیکسل قبلی انجام می شود. یعنی،

$$\hat{f}_n = \text{round} \left[\sum_{i=1}^m \alpha_i f_{n-i} \right] \quad (7-4.6)$$

که m مرتبه پیشگوی خطی است، round تابعی است که برای نشان دادن عمل گردکردن به نزدیکترین عدد صحیح استفاده می شود، و α_i ها که $i = 1, 2, \dots, m$ هستند. در کاربردهای پیماش خطی، زیرنویس "خروچی های پیشگوی را بر حسب زمان وقوع آن ها نمایه گذاری می کند"^۲ یعنی f_n, \hat{f}_n و "e" در معادلات $(5-4.6)-(7-4.6)$ را می توان با $f(t), \hat{f}(t)$ و "f" جایگزین کرد که انمایانگر زمان است. در سایر حالات، n به عنوان نمایه ای از مختصات مکانی و / یا شماره قاب تصویر (دریک دنباله زمانی از تصاویر) استفاده می شود. برای مثال در رمزگاری پیشگویی خطی یک بعدی، معادله $(7-4.6)$ را می توان به صورت

$$\hat{f}_n(x, y) = \text{round} \left[\sum_{i=1}^m \alpha_i f(x, y - i) \right] \quad (8-4.6)$$

1) Prediction error

2) Indexes (verb)

نوشت که در اینجا هر متغیر نمایه دار به صورت تابعی صریح از مختصات مکانی^۱ و^۲ (یا بیان می شود. توجه شود که طبق معادله (۶.۴-۸)، پیشگویی خطی یک بعدی ($y, f(x)$) تنها تابعی از پیکسل های قبلی روی خط فعلی است. در رمزنگاری پیشگویی دو بعدی، پیشگویی به صورت تابعی از پیکسل های قبلی در پیمایش چپ به راست و بالا به پایین تصویر است. در حالت سه بعدی، پیشگویی براساس پیکسل های قبلی قاب فعلی و پیکسل های قاب های قبلی می باشد. چون برای m پیکسل ابتدای هر خط نمی توان معادله (۶.۴-۸) را ارزیابی کرد، بنابراین باید این پیکسل ها با استفاده از سایر روش ها (نظیر رمز هافمن) رمز شوند و به عنوان سرباری^۳ از فرآیند رمزنگاری پیشگویی در نظر گرفته شوند. توضیح مشابهی نیز برای حالات با ابعاد بیشتر وجود دارد.

مثال: رمزگذاری تصویر تک رنگ شکل ۱۴.۶ (الف) با استفاده از پیشگوی خطی مرتبه اول ساده

$$f(x, y) = \text{round} [\alpha f(x, y - 1)]. \quad (9-4.6)$$

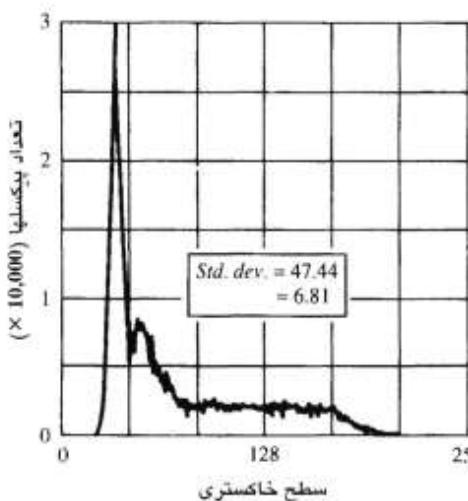
را در نظر بگیرید. یک پیشگوی با رابطه کلی فوق، پیشگوی پیکسل قبلی^۴ خوانده می شود و روال رمزنگاری پیشگویی متناظر با آن، با عنوان رمزنگاری تفاضلی^۵ یا رمزنگاری پیکسل قبلی^۶ اشاره می شود. شکل ۲۰.۶ (الف) تصویر خطای پیشگویی حاصل از معادله (۹-۴.۶) با $\alpha = 1$ را نشان می دهد. در این تصویر سطح خاکستری ۱۲۸، نمایانگر خطای پیشگویی صفر است، و تمام خطاهای پیشگویی مثبت و منفی (مریبوط به تخمین های پایین و بالا) در عدد ۸ ضرب و به ترتیب به صورت سایه های خاکستری روشن تر و تاریک تر نمایش داده می شوند. مقدار میانگین تصویر پیشگویی ۱۲۸.۰۲ است که بیانگر ۰.۰۲ بیت خطای متوسط پیشگویی است.

اشکال ۶.۲۰ (ب) و (پ) هیستوگرام سطح خاکستری تصویر شکل ۱۴.۶ (الف) و بافت نگار خطای پیشگویی حاصل از معادله (۹-۴.۶) را نشان می دهند. توجه شود که واریانس خطای پیشگویی در شکل ۶.۲۰ (پ) بسیار کوچک تر از واریانس سطوح خاکستری تصویر اولیه است. همچنین تخمین مرتبه اول آنتروپی تصویر خطای پیشگویی، بسیار کمتر از تخمین مرتبه اول آنتروپی تصویر اولیه است (۳.۹۶ بیت بر پیکسل در مقابل ۶.۸۱ بیت بر پیکسل). برخلاف این واقعیت که در تصاویر m بیتی برای نمایش دقیق دنباله^۷ خطای حاصل از معادله (۶.۴-۵) به اعداد $(1 + m)$ بیتی نیاز است، این کاهش در

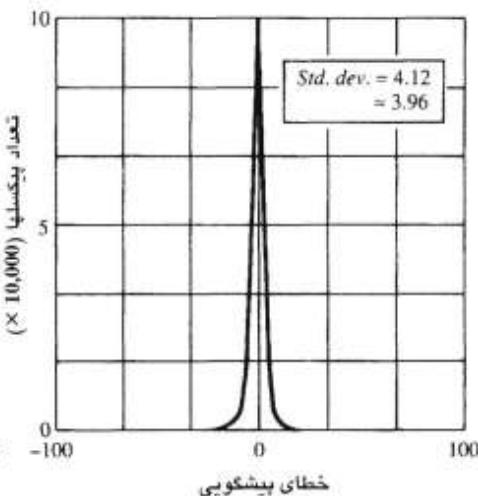
۴.۶ فشرده سازی بدون خطای ۲۳۷



(الف)



(ب)



(ب)

شکل ۲۰.۶ (الف) تصویر خطای پیشگویی حاصل از معادله (۶.۹-۴)؛ (ب) بافت نگار سطوح خاکستری تصویر اولیه؛ (پ) بافت نگار خطای پیشگویی

آتروپی نشانه حذف مقدار بزرگی از افزونگی بوسیله فرآیند رمزنگاری پیشگویی می‌باشد. گرچه هر یک از روال‌های رمزنگاری طول متغیر بخش ۴.۶ را می‌توان برای رمزگردان این دنباله خطای کار گرفت، اما نسبت فشرده سازی حاصل تقریباً به $3.96/8$ یا حدود ۰.۴۸ برابر محدود می‌شود. به طور کلی می‌توان با

تقسیم متوسط تعداد بیت های به کارفته در نمایش هریک از پیکسل های تصویر اولیه بر تخمین مرتبه اول آنتروپی داده های خطای پیشگویی، تخمینی از فشرده سازی بیشینه هر روش رمزنگاری پیشگویی بی اتفاف بدست آورد.

مثال قبلی تأکید می کند که مقدار فشرده سازی به دست آمده در رمزنگاری پیشگویی بی اتفاف مستقیماً با کاهش آنتروپی حاصل از نگاشت تصویر ورودی به دنباله خطای پیشگویی مرتبط است. چون مقدار زیادی از افزونگی بین پیکسلی بوسیله فرآیند پیشگویی و نفاذ حذف می شود، در حالت کلیتابع چگالی احتمال خطای پیشگویی قله^۱ بلندی در صفر داشته و (در مقایسه با توزیع سطح خاکستری ورودی) واریانس نسبتاً کوچکی دارد. در واقع، تابع چگالی خطای پیشگویی اغلب با pdf لابلایی ناهمبسته با میانگین صفر یعنی به صورت رابطه'

$$p_r(e) = \frac{1}{\sqrt{2}\sigma_e} \exp\left(-\frac{\sqrt{2}|e|}{\sigma_e}\right) \quad (10-4.6)$$

مدل می شود که σ_e انحراف معیار e است.

۵.۶ فشرده سازی با اتفاف

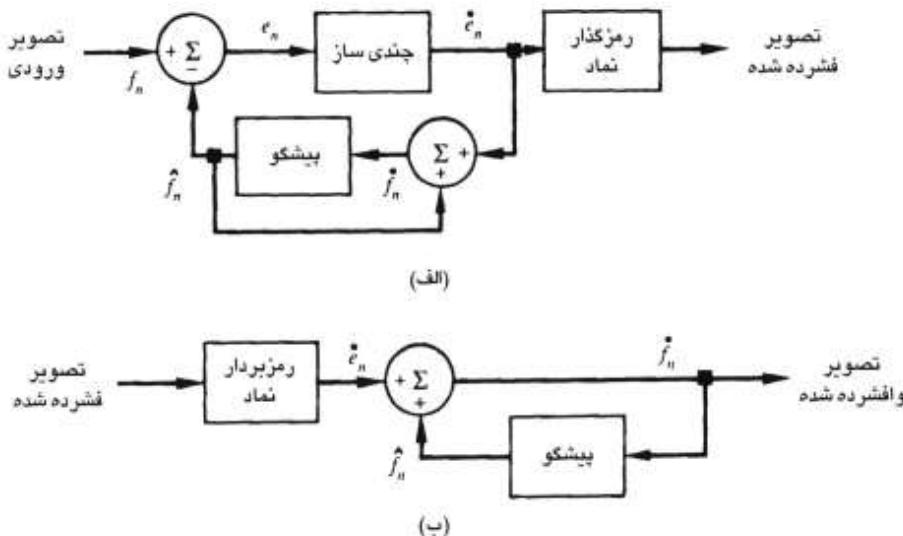
برخلاف روش های بدون خطای مشروح در بخش قبلی، رمزگذاری با اتفاف براساس مصالحه بین دقت تصویر بازسازی شده و میزان فشرده سازی می باشد. اگر اعوجاج حاصل را (که ممکن است از نظر بصری آشکار یا نااشکار باشد) بتوان تحمل کرد، می توان به افزایش قابل توجهی در میزان فشرده سازی دست یافت. در واقع بسیاری از روش های رمزنگاری با اتفاف قادرند از داده های تصاویری که بیش از ۳۰ برابر فشرده شده اند، مجدد تصاویر تک رنگ قابل تشخیص را تولید نمایند، و نیز قادرند از تصاویری که ۱۰ تا ۲۰ برابر فشرده شده اند، مجدد تصاویری تولید کنند که از تصاویر اولیه متناظر شان قابل تشخیص نیستند. در هر حال رمزنگاری بدون خطای تصاویر تک رنگ بندرت کاهش بیش از ۳ برابر داده ها را نتیجه می دهد. همان طور که در بخش ۶.۲ بیان شد، تفاوت اصلی بین این دوروش بود یا تبدیل بلوك چندی ساز شکل ۶.۶ می باشد.

۱۰۵.۶ رمزنگاری پیشگویی با اتلاف

در این بخش به مدل معرفی شده در بخش ۳.۴.۶ یک چندی ساز اضافه کرده و مصالحه حاصل بین دقت بازسازی و کارآیی فشرده سازی را بررسی می کنیم. همان طور که شکل ۲۱.۶ نشان می دهد، چندی ساز که تابع نزدیک ترین عدد صحیح از رمزگذار بدون خطای رانیز در خود دارد، بین رمزگذار تمام و نقطه ای که خطای پیشگویی شکل می گیرد، قرارداده می شود. چندی ساز، خطای پیشگویی را به بازه محدودی از خروجی ها، که با \hat{f}_n نشان داده می شوند، و مقدار فشرده سازی و اعوجاج مربوط به رمزنگاری پیشگویی با اتلاف را تعیین می کند، می نگارد.

به منظور زمینه سازی اضافه شدن مرحله 'چندی سازی'، باید رمزگذار بی خطای شکل ۱۹.۶(الف) طوری تغییر یابد که پیشگویی های تولید شده بوسیله رمزگذار و رمزبردار معادل باشند. همان طور که شکل ۲۱.۶(الف) نشان می دهد، اینکار با قرار دادن پیشگویی رمزگذار با اتلاف در حلقه بازخورد^۱ انجام می شود، که \hat{f}_n ، ورودی پیشگو، تابعی از پیشگویی های قبلی و خطاهای چندی شده متناظر آنها می باشد. یعنی،

$$\hat{f}_n = \hat{e}_n + \tilde{f}_n \quad (105.6)$$



شکل ۲۱.۶ مدلی از رمزنگاری پیشگویی با اتلاف: (الف) رمزگذار؛ (ب) رمزبردار

۴۴۰ پردازش تصویر رقمه

که \hat{f}_n به صورت تعریف شده در بخش ۶.۳.۴ می باشد. این ساختار حلقه بسته از افزایش تدریجی خطای خروجی رمزبردار جلوگیری می کند. در شکل ۶.۲۱ (ب) توجه شود که خروجی رمزبردار نیز با معادله $(6.5-1)$ تعیین می شود.

مثال: مدولاسیون دلتا^۱ یک نوع ساده اما مشهور از رمزگذاری پیشگویی با اتلاف است که در آن پیشگوی و چندی ساز به صورت های

$$\hat{f}_n = \alpha f_{n-1} \quad (2-5.6)$$

$$\dot{e}_n = \begin{cases} +\zeta & e_n > 0 \\ -\zeta & \text{در غیر این صورت} \end{cases} \quad (3-5.6)$$

تعریف می شوند که α یک ضریب پیشگویی (عموماً کوچک تر از یک) و ζ یک ثابت مثبت می باشد. خروجی چندی ساز، \hat{f}_n ، را می توان بوسیله یک بیت (شکل ۶.۲۲-الف) نمایش داد، بنابراین رمزگذار نماد شکل ۶.۲۱ (الف) می تواند از یک رمز طول ثابت یک بیتی استفاده نماید.

شکل ۶.۲۲ (ب) فرآیند مدولاسیون دلتارانمایش می دهد، که محاسبات مورد نیاز برای فشرده سازی و بازسازی دنباله ورودی

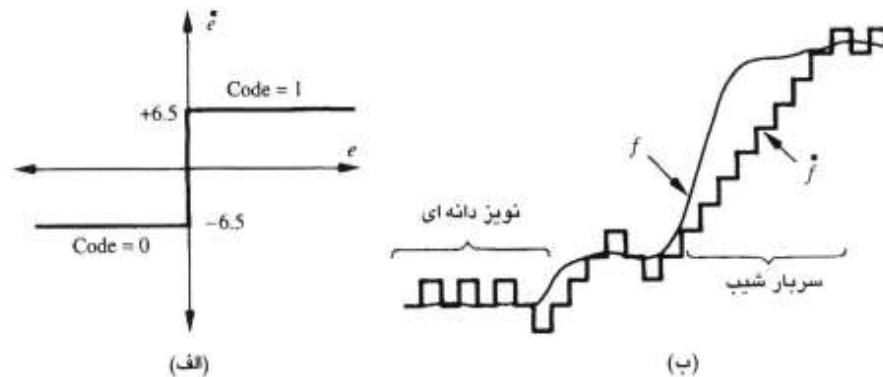
$\{14, 15, 14, 15, 13, 15, 15, 14, 20, 26, 27, 28, 27, 27, 29, 37, 47, 62, 75, 77, 78, 79, 80, 81, 81, 82, 82\}$

با زای $1 = \zeta$ در جدول آمده اند. این فرآیند با انتقال بدون خطای اولین پیکسل ورودی به رمزبردار شروع می شود. با شرط ابتدایی $14 = \hat{f}_0 = f_0$ در زوج رمزگذار و رمزبردار و با تکرار ارزیابی معادلات $(2-5.6), (3-5.6), (4-5.6)$ و $(5-5.6)$ می توان سایر خروجی ها را محاسبه نمود. پس برای مثال وقتی $n = 1$ است، $\hat{f}_1 = 14 = (1)(14) + 6.5$ ، $e_1 = 15 - 14 = 1$ ، $f_1 = 14 = 6.5 + 14 = 20.5$ ، $e_2 = 6.5 + 20.5 = 27$ و خطای بازسازی نهایی برابر $20.5 - 15 = 5.5$ سطح خاکستری است.

شکل ۶.۲۲ (ب) به طور گرافیکی داده های جدول شکل ۶.۲۲ (ب) را رسم می کند. هم ورودی و هم خروجی کاملاً رمزبرداری شده (f_n و \hat{f}_n) رسم شده اند. توجه شود که در ناحیه با تغییر سریع از $n = 14$ تا $n = 19$ که ζ آنقدر کوچک است که نمی تواند بزرگ ترین تغییرات ورودی را نمایش دهد،

1) Delta modulation (DM)

۵.۶ فشرده سازی با اتلاف



n	f	\hat{f}	e	\dot{e}	\ddot{f}	\hat{f}	\ddot{f}	$[f - \hat{f}]$
0	14	—	—	—	14.0	—	14.0	0.0
1	15	14.0	1.0	6.5	20.5	14.0	20.5	-5.5
2	14	20.5	-6.5	-6.5	14.0	20.5	14.0	0.0
3	15	14.0	1.0	6.5	20.5	14.0	20.5	-5.5
•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•
14	29	20.5	8.5	6.5	27.0	20.5	27.0	2.0
15	37	27.0	10.0	6.5	33.5	27.0	33.5	3.5
16	47	33.5	13.5	6.5	40.0	33.5	40.0	7.0
17	62	40.0	22.0	6.5	46.5	40.0	46.5	15.5
18	75	46.5	28.5	6.5	53.0	46.5	53.0	22.0
19	77	53.0	24.0	6.5	59.6	53.0	59.6	17.5
•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•

(ب)

شکل ۲۰.۶ مثالی از مدولاسیون دلتا (DM)

اعوجاجی معروف به سربار شیب^۱ رخ می دهد. بعلاوه وقتی \dot{e} بسیار بزرگتر از آن باشد که بتواند کوچک ترین تغییرات ورودی را نمایش دهد، مانند ناحیه نسبتاً هموار $0 \leq n \leq 7$ ، نویز دانه ای^۲ ظاهر می شود. در اغلب تصاویر این دو پدیده باعث لبه های مات شده و سطوح دانه ای یا نویزی (یعنی نواحی هموار اعوجاج یافته) می شوند.

اعوجاجات مذکور در مثال قبلی در تمام انواع رمزگاری پیشگویی با اتلاف متداول هستند. شدت این اعوجاجات به مجموعه پیچیده ای از تعاملات بین روش های چندی سازی و پیشگویی به کار رفته بستگی دارد. با وجود این تعاملات، معمولاً پیشگو با فرض عدم وجود نویز چندی سازی طراحی می شود.

1) Slope overload

2) Granular noise

۴۴۴ پردازش تصویر رقومی

و چندی ساز نیز برای کمینه سازی خطای خودش طراحی می شود. یعنی پیشگوی و چندی ساز مستقل از یکدیگر طراحی می شوند.

پیشگوهای بهینه

پیشگویی بهینه به کار رفته در اغلب کاربردهای رمزنگاری پیشگویی، میانگین مربع خطای پیشگویی رمزگذار^۱ یعنی

$$E\{e_n^2\} = E\{[f_n - \hat{f}_n]^2\} \quad (4-5.6)$$

را با قیود

$$\hat{f}_n = e_n + \hat{f}_{n-1} \approx e_n + f_{n-1} = f_n \quad (5-5.6)$$

و

$$\hat{f}_n = \sum_{i=1}^m \alpha_i f_{n-i}. \quad (6-5.6)$$

را کمینه می سازد. یعنی کمینه سازی میانگین مربع خطای پیشگویی به عنوان معیار بهینه سازی انتخاب گردیده، خطای چندی سازی ناچیز فرض می شود ($e_n = 0$) و مقدار پیشگویی ترکیب خطی m پیکسل قبلی می باشد.^۲ این قیود الزامی نیستند، اما تا حد زیادی تحلیل را آسان می کنند و هم زمان پیجیدگی محاسباتی پیشگورا کاهش می دهند. روش رمزنگاری پیشگویی حاصل با عنوان مدولاسیون رمز تکانه ای تفاضلی^۳ اشاره می شود.

با این شرایط مسأله طراحی پیشگویی بهینه به عمل نسبتاً آسان انتخاب m ضریب پیشگویی که عبارت

$$E\{e_n^2\} = E\left\{\left[f_n - \sum_{i=1}^m \alpha_i f_{n-i}\right]^2\right\}. \quad (7-5.6)$$

^۱ نماد (E) نشانه عملکر امید آماری است.^۲ به طور کلی پیشگویی بهینه برای یک تصویر غیر گوسی تابعی غیر خطی از پیکسل های مورد استفاده در به دست آوردن تخمین می باشد.

1) Differential pulse code modulation (DPCM)

۵.۶ فشرده سازی با ائتلاف ۴۴۳

را کمینه سازند، ساده می شود. با مشتق گیری از معادله (۶.۵-۷) نسبت به هریک از ضرایب، سپس برابر صفر قرار دادن مشتقات، و آنگاه حل معادلات همزمان حاصل با این فرض که میانگین f صفر و واریانس آن σ^2 است، معادله ماتریسی

$$\alpha = \mathbf{R}^{-1}\mathbf{r} \quad (8-5.6)$$

به دست می آید که در آن \mathbf{R}^{-1} معکوس ماتریس خود همبستگی $m \times m$

$$\mathbf{R} = \begin{bmatrix} E\{f_{n-1} f_{n-1}\} & E\{f_{n-1} f_{n-2}\} & \cdots & E\{f_{n-1} f_{n-m}\} \\ E\{f_{n-2} f_{n-1}\} & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ E\{f_{n-m} f_{n-1}\} & E\{f_{n-m} f_{n-2}\} & \cdots & E\{f_{n-m} f_{n-m}\} \end{bmatrix} \quad (9-5.6)$$

است و \mathbf{r} و α بردارهای m عنصری

$$\mathbf{r} = \begin{bmatrix} E\{f_n f_{n-1}\} \\ E\{f_n f_{n-2}\} \\ \vdots \\ E\{f_n f_{n-m}\} \end{bmatrix} \text{ و } \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} \quad (10-5.6)$$

هستند. بنابراین برای هر تصویر ورودی، ضرایبی که معادله (۶.۵-۷) را کمینه می سازند، را می توان از طریق مجموعه ای از عملیات ماتریسی مقدماتی تعیین کرد. بعلاوه این ضرایب تنها به خود همبستگی های پیکسل های تصویر اولیه بستگی دارند. واریانس خطای پیشگویی حاصل از کاربرد این ضرایب بهینه عبارتست از:

$$\sigma_e^2 = \sigma^2 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} = \sigma^2 - \sum_{i=1}^m E\{f_n f_{n-i}\} \alpha_i \quad (11-5.6)$$

گرچه ارزیابی معادله (۸.۵-۸) بسیار ساده است، اما عملاً محاسبه خود همبستگی های مورد نیاز برای تشکیل \mathbf{R} و \mathbf{r} آنقدر مشکل است که تقریباً هیچ گاه از پیشگویی های محلی^۱ (آنها که ضرایب

1) Local prediction

پیشگویی شان برای تک نک تصاویر به طور مجزا محاسبه می شوند. در اغلب حالات با فرض یک مدل ساده تصویر و جایگذاری خود همبستگی های متناظر آن در معادلات (۶.۵-۹) و (۶.۵-۱۰) مجموعه ای از ضرایب سراسری^۱ محاسبه می شوند. برای نمونه وقتی یک منبع مارکوف دو بعدی (بخش ۶.۳.۲ را بینید) دارای تابع خود همبستگی جدا پذیر

$$E\{f(x, y) f(x - i, y - j)\} = \sigma^2 \rho_i^{|} \rho_j^{|} \quad (12-5.6)$$

و نیز دارای پیشگوی خطی مرتبه چهارم تعمیم یافته

$$\begin{aligned} \hat{f}(x, y) = & \alpha_1 f(x, y - 1) + \alpha_2 f(x - 1, y - 1) \\ & + \alpha_3 f(x - 1, y) + \alpha_4 f(x - 1, y + 1) \end{aligned} \quad (13-5.6)$$

فرض شود، ضرایب بهینه حاصل (Jain[1991])

$$\alpha_1 = \rho_b \quad \alpha_2 = -\rho_c \rho_b \quad \alpha_3 = \rho_c \quad \alpha_4 = 0 \quad (14-5.6)$$

می باشد که ρ_b و ρ_c بترتیب ضرایب همبستگی افقی و عمودی تصویر مورد نظر هستند. در نهایت لازم است که جمع ضرایب پیشگویی در معادله (۶.۵-۶)، کوچک تر یا مساوی یک باشد.

یعنی،

$$\sum_{i=1}^4 \alpha_i \leq 1. \quad (15-5.6)$$

معمولًاً این قید برای کسب اطمینان از قرار گرفتن خروجی پیشگو درون محدوده مجاز سطوح خاکستری و نیز کاهش اثر نویز انتقال که معمولًاً به صورت خطوط افقی در تصویر بازسازی شده دیده می شود، اعمال می گردد. کاهش اثربندهای رمزبردار DPCM از نویز ورودی مهم است، زیرا یک خطای تکی (با شرایط واقعی) ممکن است در تمام خروجی های بعدی متشر شود. یعنی ممکن است که خروجی رمزبردار ناپایدار شود. مقید نمودن معادله (۶.۱۹-۵) به این که به طور مؤکد کوچک تر از یک باشد، اثر هر خطای ورودی را به تعداد کمی از خروجی ها محدود می کند.

مثال: خطای پیشگویی حاصل از رمزنگاری DPCM تصویر تک رنگ شکل ۶.۲۳ را با فرض خطای

1) Global coefficient

شکل ۲۳.۶ یک تصویر تک رنگ ۸ بیتی 512×512

چندی سازی صفر و با استفاده از هریک از چهار پیشگوی زیر در نظر بگیرید:

$$\hat{f}(x, y) = 0.97f(x, y - 1) \quad (16-5.6)$$

$$\hat{f}(x, y) = 0.5f(x, y - 1) + 0.5f(x - 1, y) \quad (17-5.6)$$

$$\hat{f}(x, y) = 0.75f(x, y - 1) + 0.75f(x - 1, y) - 0.5f(x - 1, y - 1) \quad (18-5.6)$$

$$\hat{f}(x, y) = \begin{cases} 0.97f(x, y - 1) & \Delta h \leq \Delta v \\ 0.97f(x - 1, y) & \end{cases} \quad (19-5.6)$$

در نظر داشته باشید که

$$\Delta v = |f(x, y - 1) - f(x - 1, y - 1)| \quad \text{و} \quad \Delta h = |f(x - 1, y) - f(x - 1, y - 1)|$$

بیانگر گرادیان های افقی و عمودی در نقطه (x, y) هستند. معادلات (۱۶-۵.۶)-(۱۹-۵.۶)

مجموعه نسبتاً پایداری از α ها را که کارآئی رضایت بخشی روی محدوده وسیعی از تصاویر فراهم می آورند، معین می کنند. طراحی پیشگوی تطبیقی معادله (۱۹-۵.۶) برای بهبود شکل ظاهری لبه ها با استفاده از اندازه محلی خواص جهتدار تصویر مانند (Δh و Δv) می باشد.

اشکال ۶ ۲۴ (الف)- (ت) تصاویر خطای پیشگویی حاصل از کاربرد پیشگوهای معادلات (۱۶-۵.۶)-(۱۹-۵.۶) را نشان می دهند. توجه شود که با افزایش مرتبه پیشگو، خطای قابل ادراک

1) edge rendition



(الف)



(ب)



(ج)



(د)

شکل ۲۴.۶ مقایسه ای از چهار روش پیشگویی خطی

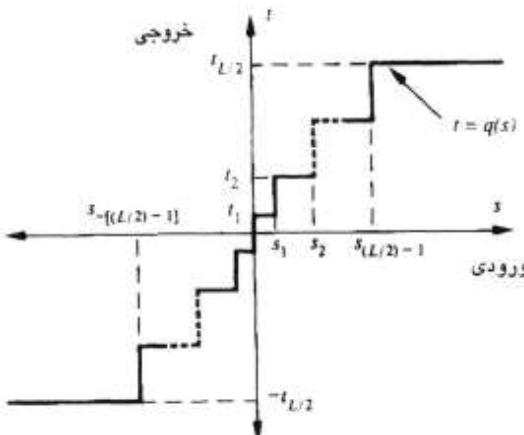
۵.۶ فشرده سازی با اثلاف ۴۳۷

بصری کاهش می یابد.[†] انحراف معیار توزیع خطاهای پیشگویی نیز از الگوی مشابهی تعیت می کنند. آنها به ترتیب ۴.۹، ۳.۷، ۴.۱ و ۳.۳ سطح خاکستری هستند.

چندی سازی بهینه

تابع چندی سازی پلکانی ($t = q(s)$) در شکل ۲۵.۶ یک تابع فرد از s است [یعنی $-q(s) = q(-s)$] که می توان آن را بوسیله $L/2$ مقدار s و $L/2$ از ربع اول نمودار کاملاً توصیف کرد. این مقادیر نقاط شکست یا ناپیوستگی های تابع را تعریف می کنند و بترتیب سطوح تصمیم^۱ و بازسازی^۲ چندی ساز خوانده می شوند. قرارداد این است که اگر s در فاصله نیمه باز (s_0, s_{0+1}) باشد، به مقدار t نگاشته شود.

مسئله طراحی چندی ساز، انتخاب بهترین s_i و t_i برای یک معیار بهینه سازی ویژه و تابع چگالی احتمال ورودی ($p(s)$) است. اگر معیار بهینه سازی، که می تواند یک معیار آماری یا روان بصری^۳ باشد، کمینه سازی میانگین مربعات خطای چندی سازی (یعنی $E\{(s - t)^2\}$) باشد و $p(s)$ نیز یک تابع زوج باشد، شرایط برای تحقق خطای کمینه ([1960]) عبارتند از



شکل ۲۵.۶ یک تابع چندی سازی نوعی

^۱ پیشگویانی که از بیش از سه یا چهار پیکسل قبلی استفاده می کنند، نسبت به پیچیدگی که به پیشگر می افزایند، مقدار کمی بهره فشرده سازی را افزایش می دهند.

برای اطلاع بیشتر در مورد معیارهای روان بصری به [Netravali 1977] و [Rubinstein 1978] و [Limb 1974] مراجعه تمایلید.

$$\int_{s_{i-1}}^{s_i} (s - t_i)p(s) ds = 0 \quad i = 1, 2, \dots, \frac{L}{2} \quad (20-5.6)$$

$$s_i = \begin{cases} 0 & i = 0 \\ \frac{t_i + t_{i+1}}{2} & i = 1, 2, \dots, \frac{L}{2} - 1 \\ \infty & i = \frac{L}{2} \end{cases} \quad (21-5.6)$$

و

$$s_{-i} = -s_i, \quad t_{-i} = -t_i \quad (22-5.6)$$

معادله (۶.۲۰-۵) بیان می کند که سطوح بازسازی، مرکز جرم نواحی زیر (s) در بازه های تصمیم هستند، در حالی که معادله (۶.۲۱-۵) دلالت بر این دارد که سطوح تصمیم، نقاط وسط فاصله بین سطوح بازسازی هستند. معادله (۶.۲۲-۵) نتیجه فرد بودن تابع η است. بازای هر L دلخواه، s ها و t هایی که معادلات (۶.۲۰-۵)-(۶.۲۲-۵) را بآورده می کنند، در مفهوم میانگین مریع خطأ، بهینه هستند. چندی ساز متناظر چندی ساز L سطحی لوید-ماکس^۱ خوانده می شود

جدول ۶.۹ سطوح تصمیم و بازسازی لویدماکس ۲، ۴ و ۸ سطحی را برای تابع چگالی احتمال لaplaci با واریانس واحد فهرست می کند (معادله ۶.۱۰-۴) را بینید. چون برای اغلب (s) های پیچیده بدست آوردن جواب صریح یا بسته برای معادلات (۶.۲۰-۵)-(۶.۲۲-۵) مشکل است، این مقادیر به روش عددی تولید شدند (Paez and Glisson[1972]). سه چندی ساز نشان داده شده به ترتیب نرخهای خروجی ثابت ۱، ۲ و ۳ بسته بر پیکسل را فراهم می آورند. چون جدول ۶.۹ برای توزیع دارای واریانس واحد تشکیل شده است، بنابراین سطوح بازسازی و تصمیم برای حالت $1 \neq \sigma$ با ضرب مقادیر جدول در انحراف معیار تابع چگالی احتمال مورد نظر بدست می آیند. سطر آخر جدول، مقادیری از اندازه پله، θ ، که معادلات (۶.۲۰-۵)-(۶.۲۲-۵) و نیز قید اضافی

$$t_i - t_{i+1} = s_i - s_{i+1} = \theta \quad (23-5.6)$$

را بآورده می کنند، فهرست می کند. اگر در رمزگذار پیشگویی با اتلاف شکل ۶.۲۱ (الف) یک رمزگذار

1) Lloyd-Max

۴۳۹ ۵.۶ فشرده سازی با اتلاف

جدول ۴.۶ چندی سازهای لویدماکس برای یکتابع چگالی احتمال لاپلاسی با واریانس واحد

سطوح	۲	۴	۸			
i	s_i	t_i	s_i	t_i	s_i	t_i
۱	∞	0.707	1.102	0.395	0.504	0.222
۲			∞	1.810	1.181	0.785
۳					2.285	1.576
۴					∞	2.994
۰		1.414		1.087		0.731

با رمز طول متغیر بکار رود، برای pdf لاپلاسی چندی ساز یکنواخت بهینه^۱ با اندازه پله ۰ نسبت به چندی ساز لوید-ماکس با رمز طول ثابت و همان مشابهت خروجی نرخ رمز پاییتری را فراهم می آورد. (O'Neil[1971])

گرچه چندی سازهای لوید ماکس و یکنواخت بهینه، تطبیقی نیستند، اما می توان با تنظیم سطوح چندی سازی بر اساس رفتار محلی تصویر به نرخ بهره بیشتری دست یافت. از دیدگاه نظری می توان نواحی با تغییر آرام را چندی سازی ریز و نواحی با تغییر سریع را بسیار درشت تر چندی نمود. با این روش، هم زمان نویز دانه ای و سربار شیب کاهش می یابد، در حالی که نرخ رمز به اندازه کمترین مقدار ممکن افزایش می یابد. هزینه این بهبود کیفیت، افزایش پیچیدگی چندی ساز است.

مثال: اشکال ۶.۲۶ (الف)، (پ) و (ث) تصاویر بازسازی شده DPCM را که به ترتیب از ترکیب چندی سازهای لویدماکس ۲، ۴ و ۸ سطحی جدول ۴.۶ با پیشگویی صفحه ای معادله (۶.۱۸-۵) به دست آمدند، نشان می دهند. چندی سازها با ضرب سطوح سطوح تصمیم و بازسازی لویدماکس این جدول در انحراف معیار خطای پیشگویی صفحه ای چندی نشده مثال قبل (یعنی $3/3$ سطح خاکستری) تولید شدند. توجه شود که لبه های تصاویر رمزبرداری شده برای سربار شیب، مات شده اند. این پدیده بویزه در شکل ۶.۲۶ (الف) که با استفاده از چندی ساز دو سطحی تولید شد، قابل توجه است، اما در اشکال ۶.۲۶ (پ) و (ث) که چهار و هشت سطح چندی سازی به کار رفته اند، کمتر آشکار است. اشکال ۶.۲۷.۶ (الف)، (پ) و (ث) تفاضل های مقیاس شده بین تصاویر رمزبرداری شده و تصویر اولیه شکل ۶.۲۳ را نشان می دهند.

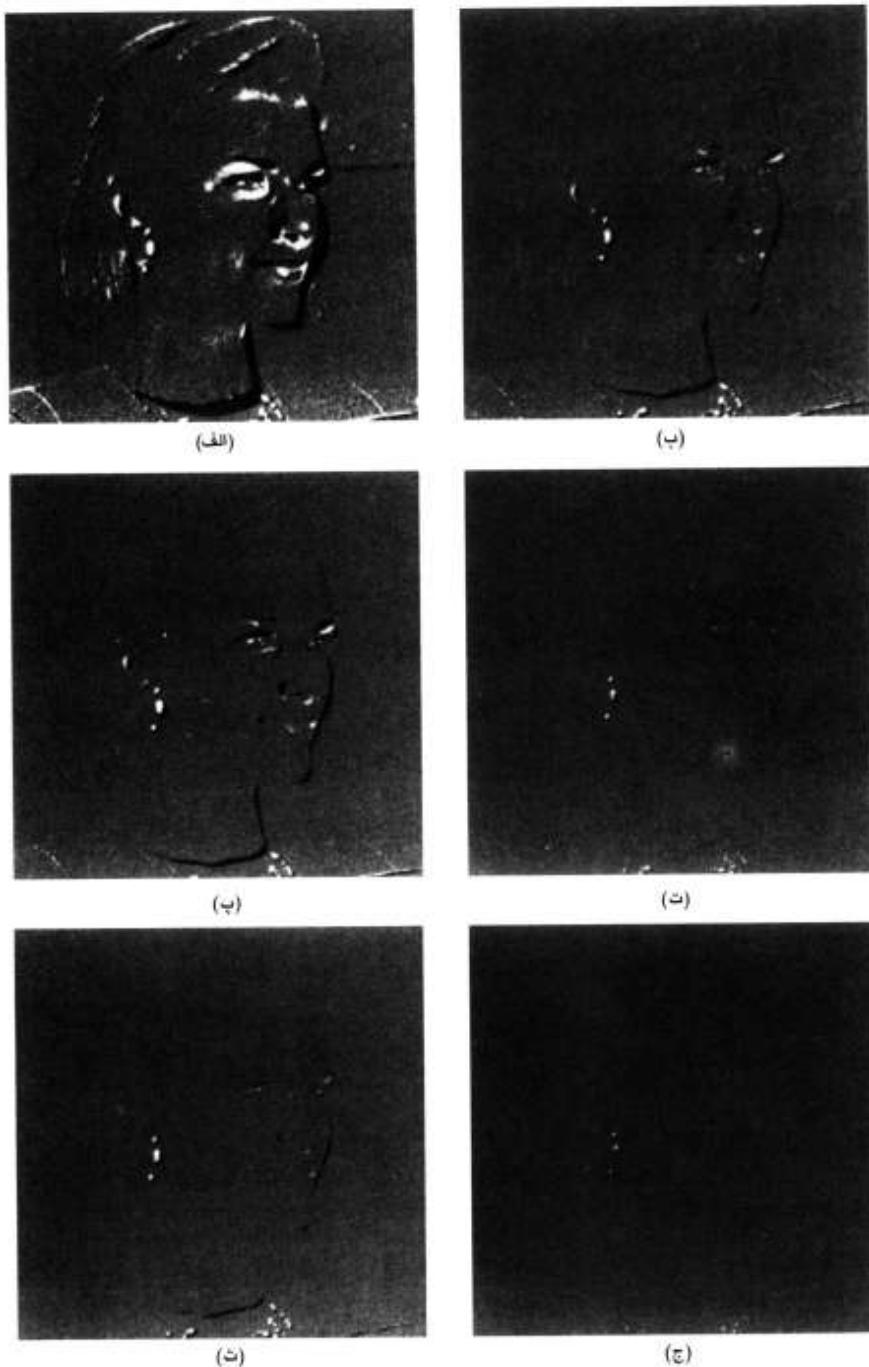
به منظور تولید تصاویر رمزبرداری شده موجود در اشکال ۶.۲۶ (ب)، (ت) و (ج) و تصاویر خطای

1) Optimum uniform quantizer



شکل ۲۶.۶ تصاویر حاصل از DPCM

۲۵۱ ۵.۶ فشرده سازی با اتلاف



شکل ۲۷.۶ تصاویر خطی DPCM مقیاس شده ($8 \times$) که متناظر با اشکال ۶.۲۶-(الف)-(ج) هستند

متناظر موجود در اشکال ۶. ۲۷. (ب)، (ت) و (ج)، از یک روش چندی سازی تطبیقی استفاده کردیم که در آن برای هر بلوک ۱۶ پیکسلی بهترین چندی ساز (در مفهوم میانگین مریع خط) از بین چهار چندی ساز ممکن انتخاب شد. این چهار چندی ساز نسخه های مقیاس شده چندی سازهای لویدماکس بهینه که قبل از توصیف شد، هستند. ضرایب مقیاس ۰.۵، ۰.۷۵، ۱.۰ و ۱.۲۵ هستند. برای مشخص کردن چندی ساز انتخاب شده به هر بلوک یک رمز دویتی افزوده شد، بنابراین سریار مربوط به انتخاب چندی ساز ۱۶/۰ بیت بر پیکسل می باشد. به کاهش زیاد خطای اثر این افزایش نسبتاً کوچک در نرخ رمز توجه کنید.

جدول ۶. ۱۰. خطاهای rms برای تصاویر تفاضل در اشکال ۶. ۲۷. (الف)- (ج)، و نیز برای چند ترکیب دیگر پیشگو- چندی ساز را فهرست می کند. توجه شود که در مفهوم میانگین مریع خط، چندی سازهای تطبیقی دو سطحی تقریباً بخوبی گونه های غیر تطبیقی چهار سطحی کار می کنند. همین طور، چندی سازهای تطبیقی چهار سطحی بهتر از روش های غیر تطبیقی هشت سطحی کار می کنند. به طور کلی نتایج عددی بیان می کنند که پیشگو های معادلات (۶. ۱۵-۵) و (۶. ۱۷-۵) و (۶. ۱۹-۵) همان مشخصات کلی پیشگوی معادله (۶. ۱۸-۵) را دارند. فشرده سازی بدست آمده با هر یک از روش های چندی سازی در سطر آخر جدول ۶. ۱۰. فهرست می شوند. توجه شود که کاهش زیاد در خطای rms معادله (۶. ۱۸-۱) که حاصل روش های تطبیقی است، اثر زیادی بر کارآیی فشرده سازی ندارد.

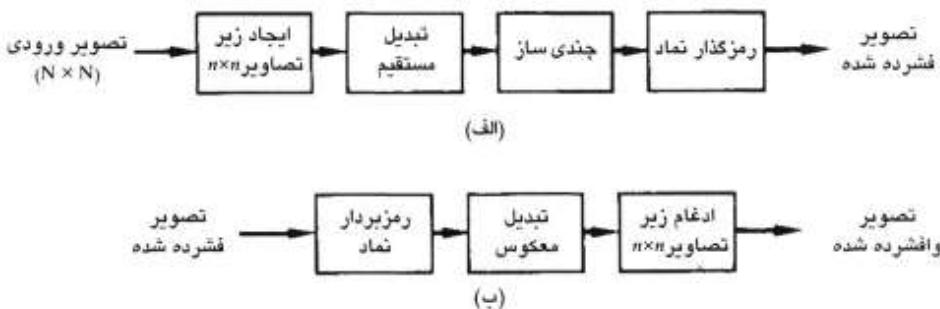
۴.۵.۶ رمزنگاری تبدیلی

روش های رمزنگاری پیشگویی مشروح در بخش ۶. ۱. ۵. مستقیماً روی پیکسل های تصویر

جدول ۶. ۱۱. جدول مقادیر خطای rms در مدلولا سیون DPCM با انتلاف بازی ای پیشگوها و چندی سازهای گوناگون

پیشگو	چندی ساز تطبیقی					
	۱ سطحی	۲ سطحی	۴ سطحی	۸ سطحی	۱۶ سطحی	۳۲ سطحی
Eq. (6.5-16)	30.88	6.86	4.08	7.49	3.22	1.55
Eq. (6.5-17)	14.59	6.94	4.09	7.53	2.49	1.12
Eq. (6.5-18)	9.90	4.30	2.31	4.61	1.70	0.76
Eq. (6.5-19)	38.18	9.25	3.36	11.46	2.56	1.14
فشرده سازی	8.00:1	4.00:1	2.70:1	7.11:1	3.77:1	2.56:1

۵.۶. فشرده سازی با اتصال ۴۵۳



شکل ۲۸.۶ یک سامانه رمزگاری تبدیلی: (الف) رمزگذار؛ (ب) رمزبردار

کار می کنند و بنابراین روش های میدان مکان^۱ خوانده می شوند. در این بخش، روش های فشرده سازی که بر اساس تبدیل تصویر هستند، را بررسی می کنیم. در رمزگاری تبدیلی^۲، ابتدا با استفاده از یک تبدیل خطی معکوس پذیر (نظیر تبدیل فوریه) تصویر به مجموعه ای از ضرایب تبدیل نگاشت می شود و سپس این ضرایب چندی و رمز می شوند. برای اغلب تصاویر طبیعی، تعداد زیادی از این ضرایب اندازه های کوچکی دارند و می توان با قبول کمی اعوجاج، آنها را بدرشتی^۳ چندی کرد (یا این که کاملاً از آنها صرف نظر کرد). هر یک از تبدیلات فصل ۳ را می توان برای تبدیل داده های تصویر پکار برد.

شکل ۲۸.۶ یک سامانه نوعی رمزگاری تبدیلی را نشان می دهد. رمزگذار چهار مرحله تجزیه به زیر تصاویر، تبدیل، چندی سازی و رمزگاری را انجام می دهد. رمزبردار معکوس مراحل فوق (به استثنای مرحله چندی سازی) را اجرا می کند. یک تصویر ورودی $N \times N$ ابتدا به زیر تصاویر $n \times n$ تقسیم می شود، و سپس با تبدیل این زیر تصاویر^۴ (N/n) آرایه تبدیل $n \times n$ بدست می آید. هدف فرآیند تبدیل، ناهمبسته کردن پیکسل های درون هر زیر تصویر، یا فشردن اطلاعات در کمترین تعداد ضرایب تبدیل است. آنگاه چندی ساز ضرایب را که کمترین اطلاعات را حمل می کنند، حذف یا بدرشتی چندی می کند. این ضرایب کمترین اثر را بر کیفیت زیر تصویر باز سازی شده دارند. در مرحله رمزگذاری، ضرایب چندی شده (معمولًا با استفاده از یک رمز طول متغیر) رمز می شوند. می توان هر یک از مراحل تمام مرحله رمزگذاری تبدیلی را با محتوای محلی تصویر تطبیق نمود که به رمزگاری تبدیلی تطبیقی^۵ موسوم است، یا این که برای تمام زیر تصاویر از مراحل کاملاً یکسانی استفاده نمود که به رمزگاری تبدیلی

1) Spatial domain method

2) Transform coding

3) Coarse

4) Adaptive transform coding

غیر تطبیقی^۱ موسوم است.

انتخاب تبدیل

سامانه های رمز نگاری تبدیلی مبتنی بر تبدیلهای کارهانن-لوو (KLT)، فوریه گسته (DFT)، کسینوسی گسته (DCT)، والش-هادامارد (WHT)، وغیره به طور گسترده ای ساختار بندی و/یا مطالعه شده اند. انتخاب یک تبدیل ویژه در یک کاربرد معین، به مقدار خطای قابل تحمل و منابع محاسباتی موجود بستگی دارد. فشرده سازی در خلال چندی سازی ضرایب تبدیل به دست می آید (نه در خلال مرحله تبدیل).

مثال: اشکال ۶.۲۹ (الف)، (ب) و (ث) سه تقریب از تصویر تک رنگ 512×512 شکل ۶.۲۳ را نشان می دهند. این تصاویر با تقسیم تصویر اولیه به زیر تصاویری با ابعاد 8×8 ، نمایش هر زیر تصویر با تبدیل DCT، WHT یا DFT آن، حذف ۵۰ درصد ضرایب تبدیل و گرفتن تبدیل معکوس از آرایه های حاصل بدست آمدند. در هر یک از حالات، ۳۲ ضریب بر جا مانده بر اساس بزرگی بیشینه انتخاب شدند. وقتی از چندی سازی یا رمز نگاری چشم پوشی کنیم، این فرآیند معادل فشرده سازی تصویر اولیه با ضریب ۲ است. توجه شود که در تمام موارد ۳۲ ضریب کنار گذاشته شده اثر بصری کمی در کیفیت تصویر باز سازی شده داشتند. در هر حال حذف آنها مقداری خطای میانگین مرربع ایجاد کرد که می توان آن خطرا را در تصاویر خطای مقیاس شده اشکال ۶.۲۹ (ب)، (ت) و (ج) مشاهده کرد. خطاهای rms بترتیب $1/28$ ، $1/86$ و $1/68$ سطح خاکستری بودند.

تفاوت های کوچک در میانگین مرربع خطاهای باز سازی مذکور در مثال قبل، مستقیماً با خواص فشردن انرژی یا اطلاعات در تبدیل به کار رفته مرتبط هستند. بیاد آورید که طبق معادله (۳.۴-۵) می توان یک تصویر $f(x, y)$ با ابعاد $n \times n$ برای $x, y = 0, 1, \dots, n-1$ را به صورت تابعی از تبدیل دو بعدی آن، $T(u, v)$ ، بیان کرد:

$$f(x, y) = \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} T(u, v) h(x, y, u, v) \quad (24-5.6)$$

توجه شود که صرفاً N جایگزین n (در معادله ۳.۴-۵) شده است و اکنون فرض می کنیم که (x, y)

¹⁾ Nonadaptive transform coding

۵.۶ فشرده سازی با اتلاف ۳۵۵



(الف)



(ب)



(پ)



(ت)



(ث)



(ج)

شکل ۶ تقریبهایی از شکل ۶. ۲۳ با استفاده از تبدیلهای (الف) فوریه، (پ) هادامارد، و (ث) کسینوسی، همراه با تصاویر خطای متناظر مقیاس شده

زیر تصویری از تصویری که باید فشرده شود، رانمایش دهد. چون پنجره معکوس $h(x, y, u, v)$ در معادله (۲۴-۵) تنها به اندیس های x, y, u و v بستگی دارد - نه به مقادیر $(y, f(x, y))$ یا $T(u, v)$ - می توان آن را به عنوان تعریف مجموعه 'تصاویر پایه'^۱ برای آرایه تعریف شده با معادله (۲۴-۵) در نظر گرفت. اگر طرز بیان معادله (۲۴-۵) را به صورت

$$\mathbf{F} = \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} T(u, v) \mathbf{H}_{uv} \quad (24-5.6)$$

تغییر دهیم، این تعبیر روشن تر می شود، که در رابطه فوق \mathbf{F} ماتریس حاوی پیکسل های یک زیر تصویر $f(x, y)$ از $n \times n$ است و

$$\mathbf{H}_{uv} = \begin{bmatrix} h(0, 0, u, v) & h(0, 1, u, v) & \cdots & h(0, n-1, u, v) \\ h(1, 0, u, v) & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ h(n-1, 0, u, v) & h(n-1, 1, u, v) & \cdots & h(n-1, n-1, u, v) \end{bmatrix} \quad (24-5.6)$$

پس \mathbf{F} به صراحت به عنوان ترکیبی خطی از n^2 ماتریس $n \times n$ - یعنی H_{uv} برای $v = 0, 1, \dots, n-1$ در معادله (۲۴-۵) - تعریف می شود. در واقع این ماتریس ها تصاویر پایه تبدیل (بخش ۳.۵) را بینند) به کار رفته در محاسبه ضرایب بسط^۲، $T(u, v)$ ها، هستند.

اگر اکنون یک تابع نقاب زنی^۳ ضرایب تبدیل را به صورت زیر

$$m(u, v) = \begin{cases} 0 & \text{اگر } T(u, v) \text{ یک معیار حذف مشخص را برآورده نماید} \\ 1 & \text{در غیر این صورت} \end{cases} \quad (24-5.6)$$

برای $u, v = 0, 1, \dots, n-1$ تعریف کنیم، از بسط

$$\hat{\mathbf{F}} = \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} T(u, v) m(u, v) \mathbf{H}_{uv} \quad (24-5.6)$$

می توان تقریبی از \mathbf{F} را به دست آورد که $m(u, v)$ برای حذف آن تصاویر پایه که کوچکترین سهم را در جمع کل معادله (۲۴-۵) دارند، تشکیل می شود. خطای میانگین مربع بین زیر تصویر \mathbf{F} و تقریب $\hat{\mathbf{F}}$ برابر

1) Basis image

2) Series expansion weighting coefficient

3) Masking function

۵.۶ فشرده سازی با اثلاف ۲۵۷

$$\begin{aligned}
 e_{ms} &= E[\|\mathbf{F} - \hat{\mathbf{F}}\|^2] \\
 &= E\left[\left\|\sum_{u=0}^{n-1} \sum_{v=0}^{n-1} T(u, v) \mathbf{H}_{uv} - \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} T(u, v) m(u, v) \mathbf{H}_{uv}\right\|^2\right] \quad (29-5.6) \\
 &= E\left[\left\|\sum_{u=0}^{n-1} \sum_{v=0}^{n-1} T(u, v) \mathbf{H}_{uv} [1 - m(u, v)]\right\|^2\right] \\
 &= \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} \sigma_{T(u,v)}^2 [1 - m(u, v)]
 \end{aligned}$$

است که $\|\mathbf{F} - \hat{\mathbf{F}}\|$ نرم ماتریسی ($\mathbf{F} - \hat{\mathbf{F}}$) است و $\sigma_{T(u,v)}^2$ ، واریانس ضریب تبدیل در محل (u, v) می باشد. ساده سازی نهایی براساس طبیعت یکه متعامد^۱ تصاویر پایه و این فرض که پیکسل های \mathbf{F} با یک فرآیند تصادفی با میانگین صفر و کوواریانس معلوم تولید شده اند، می باشد. بنابراین کل خطای میانگین مربع تقریب برابر جمع واریانس های ضرایب تبدیل محدود [یعنی ضرایبی که برای آنها $m(u, v) = 0$] می باشد. تبدیل هایی که بیشترین اطلاعات را در کمترین تعداد ضرایب توزیع نموده یا فشرده می کنند، بهترین تقریب های زیر تصویر و در نتیجه کوچک ترین خطاهای بازسازی را فراهم می آورند. در نهایت تحت فرایضی که به معادله (۲۹-۵.۶) انجامیدند، خطای میانگین مربع هر یک از (N/n) زیر تصویر یک تصویر $N \times N$ با هم برابرند. بنابراین خطای میانگین مربع $N \times N$ (که معیاری از خطای متوسط آن است)، برابر خطای میانگین مربع یک زیر تصویر می باشد.

مثال قبلی نشان داد که توانایی DCT در فشردن اطلاعات از DFT و WHT بیشتر است. با این که معمولاً برای اغلب تصاویر طبیعی این وضعیت وجود دارد، اما KLT، نه DCT، تبدیل بهینه در مفهوم فشردن اطلاعات است. یعنی KLT برای هر تصویر ورودی و هر تعداد دلخواه از ضرایب برجامانده، خطای میانگین معادله (۲۹-۵.۶) را کمیته می سازد (Mathews[1956] و Kramer[†]).

با این حال چون KLT وابسته به داده است، در حالت کلی بدست آوردن تصاویر پایه KLT برای هر یک از زیر تصویرها از لحاظ محاسباتی سنگین است. به این دلیل بندرت در عمل از KLT استفاده می شود. معمولاً بجای آن تبدیلی نظریه DCT که تصاویر پایه اش ثابت (مستقل از ورودی) هستند، انتخاب می شود. از بین تبدیل های مستقل از ورودی ممکن، تبدیل های غیر سینوسی (نظریه WHT یا تبدیل هاار) اجرای ساده تری دارند. توانایی فشردن اطلاعات در تبدیل های سینوسی (نظریه DCT یا DFT) تقریب

[†] یک شرط دیگر برای بهینگی این است که تابع پوشاننده معادله (۲۷-۵.۶) ضرایب KLT یا واریانس بیشینه را انتخاب نماید.
1) Orthonormal

نزدیک تری به همین توانایی در KLT بھیته می باشد.

بنابراین اغلب سامانه های عملی رمزنگاری براساس DCT، که مصالحه خوبی بین قابلیت فشردن اطلاعات و پیچیدگی محاسباتی فراهم می آورد، هستند. در واقع خواص DCT چنان ارزش عملی دارند که اکنون این تبدیل به صورت استاندارد بین المللی برای سامانه های رمزنگاری تبدیلی در آمده است (بخش ۶ . ۶ را ببینید). این تبدیل در مقایسه با سایر تبدیل های مستقل از ورودی چند مزیت دارد: اولاً در یک مدار مجتمع پیاده سازی شده است، ثانیاً بیشترین اطلاعات را در کمترین تعداد ضرایب فشرده می کند^۱ (برای اغلب تصاویر طبیعی)، ثالثاً ظاهر جعبه ای^۲ تصویر (مشخص شدن مرزهای بین زیرتصاویر در تصویر کامل) که اثر جعبه ای^۳ نیز خوانده می شود، را کمینه می سازد. مزیت آخر بویژه در مقایسه با سایر تبدیل های سینوسی مهم است. همان طور که شکل ۶ . ۳۰ (الف) نشان می دهد، تناوب^۴ نقطه ای ضمنی DFT باعث ناپیوستگی های مرزی می شود که این ناپیوستگی ها محتوای فرکانس بالای را برای تبدیل بوجود می آورند. بنابراین وقتی ضرایب تبدیل DFT، قطع یا چندی شوند، پدیده گیبس^۵ بوجود می آید که باعث می شود نقاط مرزی مقادیر خطداری بگیرند و این مقادیر خطدار باعث اثر جعبه ای را ایجاد می کنند. یعنی مرزهای بین زیرتصاویر مجاور مرئی می شوند، زیرا پیکسل های مرزی زیرتصاویر، مقادیر میانگین ناپیوستگی های ایجاد شده در نقاط مرزی را می گیرند (شکل ۶ . ۳۰ الف را ببینید).

DCT موجود در شکل ۶ . ۳۰ (ب) کاهش این اثر را نشان می دهد، زیرا تناوب^۶ نقطه ای ضمنی آن، ناپیوستگی های مرزی را تولید نمی کند.

انتخاب ابعاد زیرتصویر

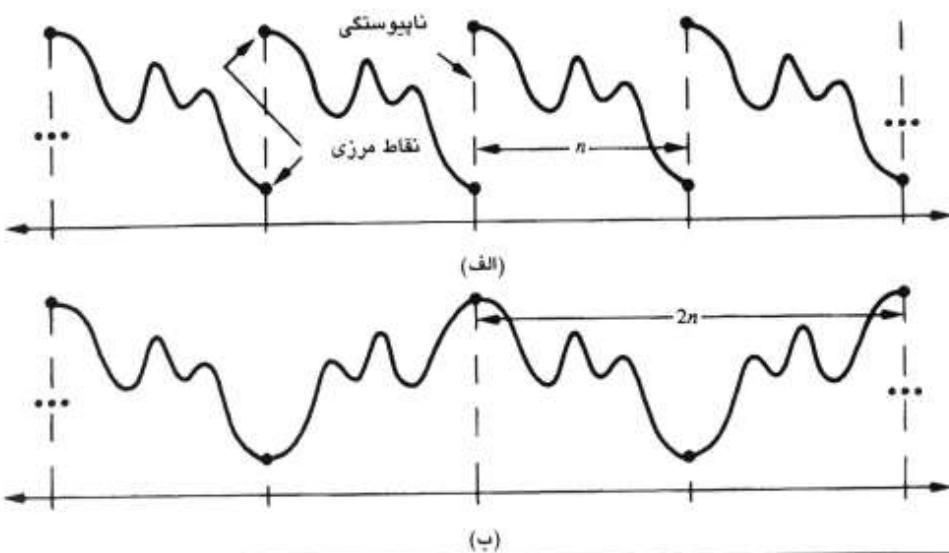
عامل مهم دیگری که برخطای رمزنگاری تبدیلی و پیچیدگی محاسباتی اثر می گذارد، ابعاد زیرتصویر است. در اغلب کاربردها، تصاویر طوری تقسیم می شوند که همبستگی (افزونگی) بین زیرتصاویر مجاور به سطح قابل قبولی کاهش یابد و^۷ نیز توان صحیحی از ۲ باشد (مانند گذشته^۸ بعد زیرتصویر است).

Ahmed و همکارانش [1974] اوین کسانی بودند که مدعی شدند تصاویر پایه KLT مربوط به یک منبع تصویری مارکوف مرتبه اول شیاهت نزدیکی با تصاویر پایه DCT دارند. در واقع وقتی ضریب همبستگی بین پیکسل های مجاور به یک نزدیک می شود، تصاویر پایه وابسته به ورودی KLT با تصاویر پایه مستقل از ورودی DCT برابر می شوند ([Clarke 1985]). Gibbs phenomenon^۹ : این پدیده که در اغلب متون مهندسی بر قریب راجع به تحلیل مدار (مثلًا [Guillemin 1949]) توصیف شده است، به این علت رخ می دهد که تبدیل فوریه تمعی تواند در ناپیوستگیها به طور یکنواخت همگراشود. بسطهای فوریه در ناپیوستگی ها مقادیر میانگین را به خود می گیرند.

1) Blocklike appearance

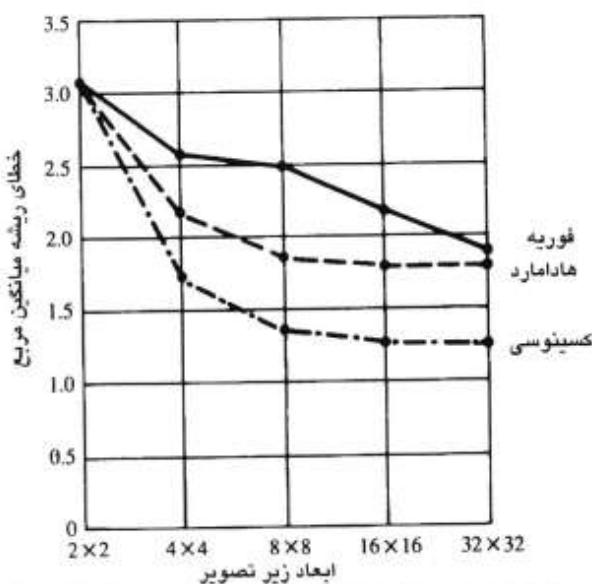
2) Blocking artifact

۳۵۹ ۵.۶ فشرده سازی با اتلاف

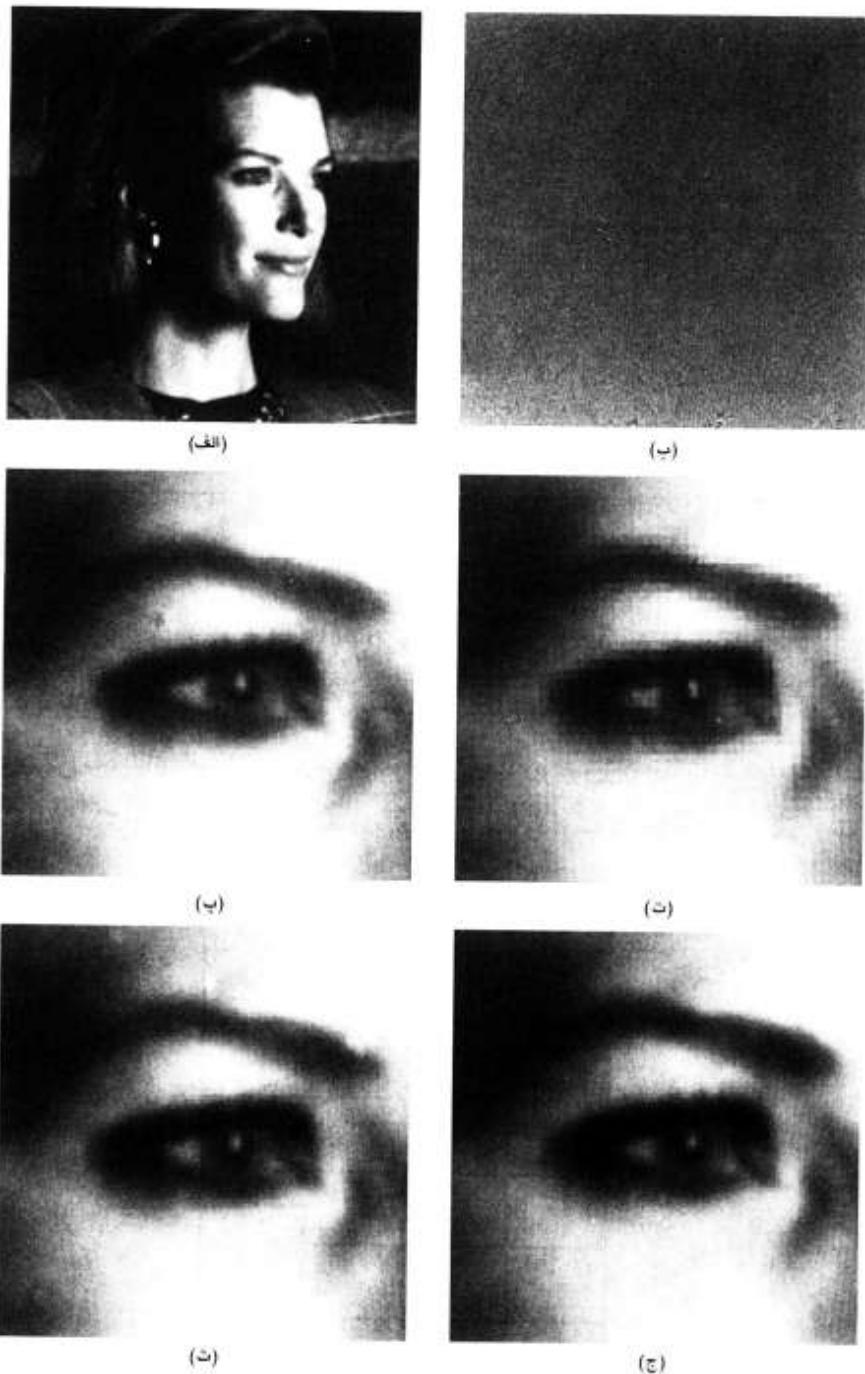


شکل ۶. ۳۰ تناوب موجود در (الف) DFT و (ب) DCT یک بعدی

شرط آخری، محاسبه تبدیل های زیرتصویر را آسان می سازد (روش های دو برابر سازی متوالی مبنای ۲ مشروح در بخش های ۴.۳ و ۵.۵ را بینید). به طور کلی با افزایش ابعاد زیرتصویر، هم میزان فشرده سازی و هم پیچیدگی محاسباتی افزایش می یابند. متداول ترین ابعاد زیرتصویر 8×8 و 16×16 هستند. مثال: شکل ۶. ۳۱ به طور گرافیکی اثر ابعاد زیرتصویر بر خطای بازسازی رمزگاری تبدیلی را نمایش



شکل ۶. ۳۱ خطای بازسازی بر حسب ابعاد زیر تصویر



شکل ۳۲.۶ تقریب هایی از شکل ۲۳.۶ با استفاده از ۲۵ درصد ضرایب DCT : (الف) و (ب) نتایج با زیرتصویر 8×8 ; (پ) تصویر اولیه بزرگنمایی شده؛ (ت) نتیجه با زیرتصویر 2×2 ; (ث) نتیجه با زیرتصویر 4×4 ; و (ج) نتیجه با زیرتصویر 8×8

۶.۵. فشرده سازی با اتلاف



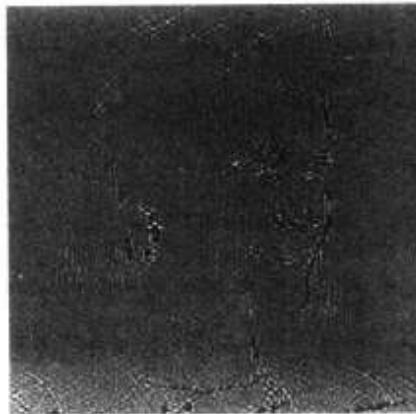
(الف)



(ب)



(پ)



(ت)



(ث)



(ج)

شکل ۳۳.۶ تقریب های از شکل ۶.۲۳ با استفاده از ۱۲/۵ درصد ضرایب $DCT \times 8 \times 8$: (الف)، (ب) و (ث)
نتایج رمزگاری آستانه ای؛ (ب)، (ت) و (ج) نتایج رمزگاری ناحیه ای

می دهد. داده های ترسیم شده با انجام مراحل زیر به دست آمدند: تقسیم تصویر تک رنگ شکل ۶.۲۳ به زیر تصاویری با ابعاد $n \times n$ بازی ۳۲، ۱۶، ۸، ۴، ۲، $n = 2, 4, 8, 16, 32$ ، محاسبه تبدیل هر زیر تصویر، حذف ۷۵ درصد ضرایب حاصل، و تبدیل معکوس آرایه های قطع شده. توجه شود که وقتی ابعاد زیر تصویر از 8×8 بزرگتر می شوند، منحنی های هادامارد و کسینوسی تخت می شوند، در حالی که در این ناحیه خطای بازسازی فوریه حتی سریع تر کاهش می یابد. برونایی این منحنی ها در مقادیر بزرگتر ۱۱ نشان می دهد که خطای بازسازی فوریه با منحنی های هادامارد تقاطع کرده و سپس به حاصل کسینوسی همگرا می شود. در واقع این نتیجه با یافته های نظری و نتایج آزمایش های گزارش شده بوسیله Limb[1980] و Netravali و Pratt[1978] نیز برای منبع تصویری مارکوف دو بعدی سازگار است.

وقتی زیر تصویر های 2×2 استفاده شوند، هر سه منحنی با هم تقاطع می کنند. در این حالت از هر چهار ضریب تنها یک ضریب (۲۵ درصد) هر آرایه تبدیل شده بر جا می ماند که این ضریب مؤلفه DC است، بنابراین در تبدیل معکوس کافیست که چهار پیکسل هر زیر تصویر با مقدار DC یا متوسط آن جایگزین شوند (برای مثال معادله ۳.۲۰-۳.۲۱ را ببینید). این وضعیت در شکل ۶.۳۲ (ت) که یک بخش بزرگ نمایی شده از نتیجه DCT 2×2 را نشان می دهد، آشکار است. توجه شود که اثر جعبه ای که در این شکل آشکار است، با افزایش ابعاد زیر تصویر به 4×4 و 8×8 (اشکال ۶.۳۲. (ث) و (ج)) کاهش می یابد. شکل ۶.۳۲ (پ) یک بخش بزرگ نمایی شده از تصویر اولیه را به عنوان مرجع نشان می دهد. بعلاوه اشکال ۶.۳۲. (الف) و (پ) مقایسه این نتایج با نتایج مثال قبلی را تسهیل می کنند.

تخصیص بیت

خطای بازسازی مربوط به بسط دباله ای قطع شده معادله (۶.۵-۲۸) تابعی از تعداد و اهمیت نسبی ضرایب تبدیل محدود، همچنین دقیقی که برای نمایش ضرایب بر جا مانده استفاده می شود، می باشد. در اغلب سامانه های رمزنگاری تبدیلی، ضرایب بر جا مانده براساس واریانس بیشینه، که رمزنگاری ناحیه ای^۱ خوانده می شود، یا براساس بزرگی بیشینه، که رمزنگاری آستانه ای^۲ خوانده می شود، انتخاب می شوند [یعنی تابع نمونه برداری معادله (۶.۵-۲۷) ایجاد می شود]. عموماً به فرآیند کلی قطع کردن، چندی سازی و رمزنگاری ضرایب زیر تصویر تبدیل شده، تخصیص بیت^۳ می گویند.

1) Zonal coding

2) Threshold coding

3) Bit allocation

مثال: اشکال ۶. ۳۳ (الف) و (ب) دو تقریب از شکل ۶. ۲۳ را نشان می دهند که در آنها ۸۷/۵ درصد ضرایب DCT هر زیر تصویر 8×8 حذف شده اند. اولین تقریب از طریق رمز نگاری آستانه ای و نگه داشتن بزرگترین هشت ضریب تبدیل بدست آمد، و تقریب دوم با روش رمز نگاری ناحیه ای تولید شد. در روش رمز نگاری ناحیه ای هر ضریب DCT یک متغیر تصادفی در نظر گرفته شد که می توان توزیع آن را روی مجموعه زیر تصاویر تبدیل شده محاسبه کرد. ۸ ضریب دارای بزرگترین واریانس ها (۱۲/۵ درصد از ۶۴ ضریب تبدیل هر زیر تصویر 8×8) تعیین محل شدن و در تمام زیر تصاویر تبدیل شده برای بازسازی زیر تصاویر اصلی فقط از همین ۸ ضریب استفاده شد. توجه شود که خطای تفاضلی رمز نگاری آستانه ای (شکل ۶. ۲۳-پ) بسیار کمتر از خطای تفاضلی رمز نگاری ناحیه ای (شکل ۶. ۳۳-ت) است. در اشکال ۶. ۳۴ (ت) و (ج) بخش کوچکی از تصاویر بازسازی شده در (الف) و (ب) بزرگنمایی شده اند.

□

رمز نگاری ناحیه ای: رمز نگاری ناحیه ای بر اساس مفهوم تعبیر اطلاعات به عنوان عدم قطعیت می باشد. بنابر این ضرایب تبدیل که دارای واریانس بیشینه هستند، بیشترین اطلاعات تصویر را حمل می کنند و باید در فرآیند فشرده سازی حذف شوند. خود واریانس ها را می توان مانند مثال قبل از مجموعه² (N/n) آرایه زیر تصویر تبدیل شده، یا بر اساس مدل تصویری مفروض (مثلًا تابع خود همبستگی مارکوف) محاسبه کرد. در هر دو حالت طبق معادله (۶. ۲۸-۵) فرآیند نمونه برداری ناحیه ای را می توان به عنوان ضرب داخلی^۱ ($T_{(l),l}$) در یک نقاط ناحیه ای^۲ که در محل متناظر با ضرایب دارای واریانس بیشینه مقدار "۱" و در سایر نقاط مقدار "۰" دارد، در نظر گرفت. معمولاً ضرایب با واریانس بیشینه حول مبدأ آرایه تبدیل تصویر قرار دارند، که نقاط ناحیه ای نوعی شکل ۶. ۳۴ (الف) را نتیجه می دهند.

ضرایب بر جا مانده در خلال فرآیند نمونه برداری ناحیه ای باید چندی و رمز شوند، بنابر این گاهی نقاط ناحیه ای که تعداد بیت های مورد استفاده برای رمز کردن هر ضریب را نشان می دهند، رسم می شوند (شکل ۶. ۳۴-ب). غالباً تعداد بیت یکسانی به هر یک از ضرایب اختصاص می یابد، یا این که یک تعداد ثابت از بیت های طور نامساوی بین آنها توزیع می شود. در حالت اول معمولاً ضرایب بوسیله انحراف معیارشان تراز و به طور یکتا خخت چندی می شوند. در حالت دوم برای هر ضریب یک چندی ساز، نظیر چندی ساز بهینه لوید-ماکس، طراحی می شود. برای ساخت چندی سازهای مورد

۱- ضرب عنصر در عنصر متناظر و جمع حاصل ضربها

2) Zonal mask

(الف)	(ب)																																																																																																																																																	
<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <tr><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> </table>	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <tr><td>8</td><td>7</td><td>6</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td></tr> <tr><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td></tr> <tr><td>6</td><td>5</td><td>4</td><td>3</td><td>3</td><td>1</td><td>1</td><td>0</td></tr> <tr><td>4</td><td>4</td><td>3</td><td>3</td><td>2</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>3</td><td>3</td><td>3</td><td>2</td><td>1</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>2</td><td>2</td><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	8	7	6	4	3	2	1	0	7	6	5	4	3	2	1	0	6	5	4	3	3	1	1	0	4	4	3	3	2	1	0	0	3	3	3	2	1	1	0	0	2	2	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0																																																																																																																																										
0	1	0	0	0	0	0	0	0																																																																																																																																										
0	0	1	0	0	0	0	0	0																																																																																																																																										
0	0	0	1	0	0	0	0	0																																																																																																																																										
0	0	0	0	1	0	0	0	0																																																																																																																																										
0	0	0	0	0	1	0	0	0																																																																																																																																										
0	0	0	0	0	0	1	0	0																																																																																																																																										
0	0	0	0	0	0	0	1	0																																																																																																																																										
0	0	0	0	0	0	0	0	1																																																																																																																																										
8	7	6	4	3	2	1	0																																																																																																																																											
7	6	5	4	3	2	1	0																																																																																																																																											
6	5	4	3	3	1	1	0																																																																																																																																											
4	4	3	3	2	1	0	0																																																																																																																																											
3	3	3	2	1	1	0	0																																																																																																																																											
2	2	1	1	1	0	0	0																																																																																																																																											
1	1	1	0	0	0	0	0																																																																																																																																											
0	0	0	0	0	0	0	0																																																																																																																																											
(ب)	(ت)																																																																																																																																																	
<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <tr><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> </table>	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <tr><td>0</td><td>1</td><td>5</td><td>6</td><td>14</td><td>15</td><td>27</td><td>28</td></tr> <tr><td>2</td><td>4</td><td>7</td><td>13</td><td>16</td><td>26</td><td>29</td><td>42</td></tr> <tr><td>3</td><td>8</td><td>12</td><td>17</td><td>25</td><td>30</td><td>41</td><td>43</td></tr> <tr><td>9</td><td>11</td><td>18</td><td>24</td><td>31</td><td>40</td><td>44</td><td>53</td></tr> <tr><td>10</td><td>19</td><td>23</td><td>32</td><td>39</td><td>45</td><td>52</td><td>54</td></tr> <tr><td>20</td><td>22</td><td>33</td><td>38</td><td>46</td><td>51</td><td>55</td><td>60</td></tr> <tr><td>21</td><td>34</td><td>37</td><td>47</td><td>50</td><td>56</td><td>59</td><td>61</td></tr> <tr><td>35</td><td>36</td><td>48</td><td>49</td><td>57</td><td>58</td><td>62</td><td>63</td></tr> </table>	0	1	5	6	14	15	27	28	2	4	7	13	16	26	29	42	3	8	12	17	25	30	41	43	9	11	18	24	31	40	44	53	10	19	23	32	39	45	52	54	20	22	33	38	46	51	55	60	21	34	37	47	50	56	59	61	35	36	48	49	57	58	62	63
1	0	0	0	0	0	0	0	0																																																																																																																																										
0	1	0	0	0	0	0	0	0																																																																																																																																										
0	0	1	0	0	0	0	0	0																																																																																																																																										
0	0	0	1	0	0	0	0	0																																																																																																																																										
0	0	0	0	1	0	0	0	0																																																																																																																																										
0	0	0	0	0	1	0	0	0																																																																																																																																										
0	0	0	0	0	0	1	0	0																																																																																																																																										
0	0	0	0	0	0	0	1	0																																																																																																																																										
0	0	0	0	0	0	0	0	1																																																																																																																																										
0	1	5	6	14	15	27	28																																																																																																																																											
2	4	7	13	16	26	29	42																																																																																																																																											
3	8	12	17	25	30	41	43																																																																																																																																											
9	11	18	24	31	40	44	53																																																																																																																																											
10	19	23	32	39	45	52	54																																																																																																																																											
20	22	33	38	46	51	55	60																																																																																																																																											
21	34	37	47	50	56	59	61																																																																																																																																											
35	36	48	49	57	58	62	63																																																																																																																																											
(ب)	(ت)																																																																																																																																																	

شکل ۳۴.۶ (الف) یک نقاب ناحیه ای، (ب) یک تخصیص بیت ناحیه ای نوعی. (ب) یک نقاب آستانه نوعی، و (ت) یک دنباله نوعی برای مرتب سازی ضرایب آستانه گیری شده. سایه ضرایبی که بر جا می مانند، را مشخص می سازد

نیاز، معمولاً ضریب مرتبه صفر با DC باتابع چگالی رالی^۱ مدل می شود، در حالی که سایر ضرایب با چگالی لاپلاسی یا گوسی مدل می شوند.^۲ تعداد سطوح چندی سازی (و بنابراین تعداد بیتها) در نظر گرفته شده برای هر چندی ساز مناسب با $\log_2 \sigma^2 / \sigma_{T(u,v)}^2$ است. این تخصیص بیت با نظریه اعوجاج نرخ سازگار است. طبق نظریه اعوجاج نرخ یک متغیر گوسی با واریانس^۲ σ^2 رانمی توان با کمتر از $\frac{1}{2} \log_2 (\sigma^2 / D)$ بیت نمایش داده و سپس آن را با خطای میانگین مربع کمتر از D بازسازی نمود (مسئله ۶.۱۱ را ببینید). نتیجه شهودی این است که محتوای اطلاعات یک متغیر تصادفی گوسی مناسب با $\log_2 (\sigma^2 / D)$ است. بنابراین به هر یک از ضرایب بر جا مانده در معادله (۶.۵-۲۸)-که در اینجا بر اساس

^۱ چون هر ضریب توکیبی خطی از پیکسل های زیر تصویر می باشد، توزیع های ضرایب تبدیل آن به سمت گوسی میل می کنند. با این حال این نتیجه در مورد ضریب DC صدق نمی کند، زیرا تصاویر غیرمنفی همواره ضرایب DC مثبت دارند.

^۲ Rayleigh

واریانس بیشینه انتخاب می شوند- باید متناسب با لگاریتم واریانس آن، بیت اختصاص داد.

رمزنگاری آستانه‌ای: معمولاً رمزنگاری ناحیه‌ای تنها با استفاده از یک نقاب ثابت برای تمام زیر تصاویر اجرا می‌شود. اما رمزنگاری آستانه‌ای از این جهت که محل ضرایب تبدیل بر جا مانده از هر زیر تصویر با زیر تصویر دیگر متفاوت است، به طور ذاتی تطبیقی است. در واقع رمزنگاری آستانه‌ای روش رمزنگاری تبدیلی تطبیقی است که به علت سادگی محاسبات آن غالباً در عمل استفاده می‌شود. همان‌طور که در مثال قبل بیان شد، مفهوم اساسی این است که برای هر زیر تصویر، بزرگترین ضرایب تبدیل از یک زیر تصویر به را در کیفیت زیر تصویر بازسازی شده دارند. چون محل بزرگترین ضرایب تبدیل از یک زیر تصویر به زیر تصویر دیگر تغییر می‌کند، معمولاً $m(u, v)$ عناصر $T(u, v)$ (به روش از پیش گزیده‌ای) باز مرتب می‌شوند تا یک دنباله یک بعدی رمز RL را تشکیل دهند. شکل ۶.۳۴ (پ) یک نقاب آستانه^۱ نوعی برای یک زیر تصویر از یک تصویر قرضی را نشان می‌دهد. این نقاب راه مناسبی را برای به تصویر در آوردن فرآیند رمزنگاری آستانه‌ای زیر تصویر مربوطه، همچنین توصیف ریاضی این فرآیند با استفاده از معادله ۶.۲۸-۵ فراهم می‌آورد. وقتی این نقاب (از طریق معادله ۶.۲۸-۵) به زیر تصویری که برای آن بدست آمده است، اعمال شود و آرایه $n \times n$ حاصل طبق الگوی مرتب سازی زیگزاگی شکل ۶.۳۴ (ت) و به صورت یک دنباله^۲ عنصری از ضرایب مرتب شود، دنباله یک بعدی حاصل حاوی دنباله‌های طولانی صفر است. معمولاً این دنباله‌ها، رمز RL می‌شوند. ضرایب غیر صفر یا بر جامانده، که متناظر با محل "۱"‌های نقاب هستند، با استفاده از یکی از رمزهای RL بخش ۶.۴ نمایش داده می‌شوند.

سه روش اصلی برای آستانه‌گیری از یک زیر تصویر تبدیل شده‌یا، به عبارت دیگر، برای ایجاد تابع نقاب زنی آستانه‌ای زیر تصویر^۳ به شکل معادله ۶.۲۷-۵ وجود دارد: (۱) روش اول این است که یک آستانه سراسری در تمام زیر تصاویر بکار رود؛ (۲) روش دوم این که برای هر زیر تصویر از آستانه مناسب آن زیر تصویر استفاده شود؛ و (۳) روش سوم این که مقدار آستانه برای هر ضریب تابعی از محل آن ضریب در زیر تصویر باشد. در روش اول میزان فشرده سازی بسته به تعداد ضرایبی که از آستانه سراسری تجاوز می‌کنند، از یک زیر تصویر به زیر تصویر دیگر تغییر می‌کند. در روش دوم که رمزنگاری N-بزرگترین ها^۳ خوانده می‌شود، در هر یک از زیر تصویرها، تعداد یکسانی از ضرایب صرف نظر می‌شوند. نتیجتاً نرخ رمز

1) Threshold mask

2) Subimage threshold masking function

3) N- largest coding

علوم واژ قبل ثابت است. روش سوم مشابه روش اول، نرخ رمز متغیری را نتیجه می دهد، اما این مزیت را دارد که در معادله (۳۰-۵. ۶) می توان با جایگزینی

$$\hat{T}(u, v) = \text{round} \left[\frac{T(u, v)}{Z(u, v)} \right] \quad (30-5. 6)$$

بحای (۳۰-۵. ۶) آستانه گیری و چندی سازی را با هم ترکیب کرد، که $\hat{T}(u, v)$ یک تقریب آستانه گیری و چندی شده از $T(u, v)$ و $Z(u, v)$ ، آرایه تراز کننده تبدیل، است:

$$Z = [Z(u, v)] = \begin{bmatrix} Z(0, 0) & Z(0, 1) & \cdots & Z(0, n-1) \\ Z(1, 0) & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ Z(n-1, 0) & Z(n-1, 1) & \cdots & Z(n-1, n-1) \end{bmatrix} \quad (31-5. 6)$$

قبل از این که بتوان از تبدیل تراز شده (آستانه گیری و چندی شده) زیرتصویر، $\hat{T}(u, v)$ ، تبدیل معکوس گرفت تا تقریبی از $F(u, v)$ بدست آید، باید آن را در $Z(u, v)$ ضرب نمود. آرایه واتراز شده^۱ حاصل، که با $\hat{T}(u, v)$ نشان داده می شود، تقریبی از $\hat{T}(u, v)$ است:

$$\hat{T}(u, v) = \hat{T}(u, v) Z(u, v). \quad (32-5. 6)$$

تبدیل معکوس $\hat{T}(u, v)$ تقریبی از زیرتصویر وافسرده شده^۲ است. شکل ۳۵. ۶ نمودار معادله (۳۰-۵. ۶) را برای حالتی که یک مقدار ویژه c به $Z(u, v)$ متنسب شده است، رسم می کند. توجه کنید که $\hat{T}(u, v)$ مقدار صحیح k را می گیرد اگر و تنها اگر

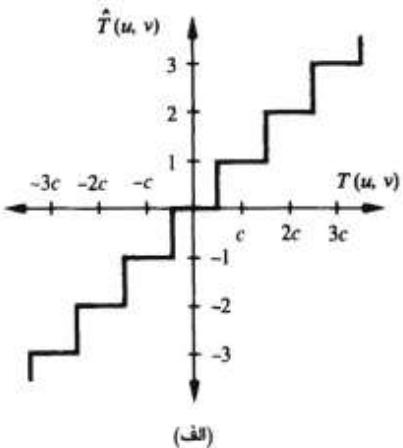
$$kc - \frac{c}{2} \leq T(u, v) < kc + \frac{c}{2}.$$

اگر $Z(u, v) > 2T(u, v)$ باشد، آنگاه $\hat{T}(u, v) = 0$ بوده و ضریب تبدیل کاملاً حذف یا کنار گذاشته می شود. وقتی $\hat{T}(u, v)$ با یک رمز طول متغیر که با افزایش بزرگی k طولش زیاد می شود، نمایش داده شود، تعداد بیت های استفاده شده برای نمایش $\hat{T}(u, v)$ با مقدار c کنترل می شود. بنابراین می توان عناصر Z را مقیاس نمود تا به سطوح فشرده سازی گوناگونی دست یابیم. شکل ۳۵. ۶(ب) آرایه

۱) Denormalized

2) Decompressed

۴۶۷ فشرده سازی با اتفاف ۵.۶



16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(ب)

شکل ۳۵.۶ (الف) یک منحنی چندی سازی رمزگاری آستانه ای (معادله ۶.۵-۳۰ را بینید)؛ (ب) یک ماتریس ترازسازی نوعی.

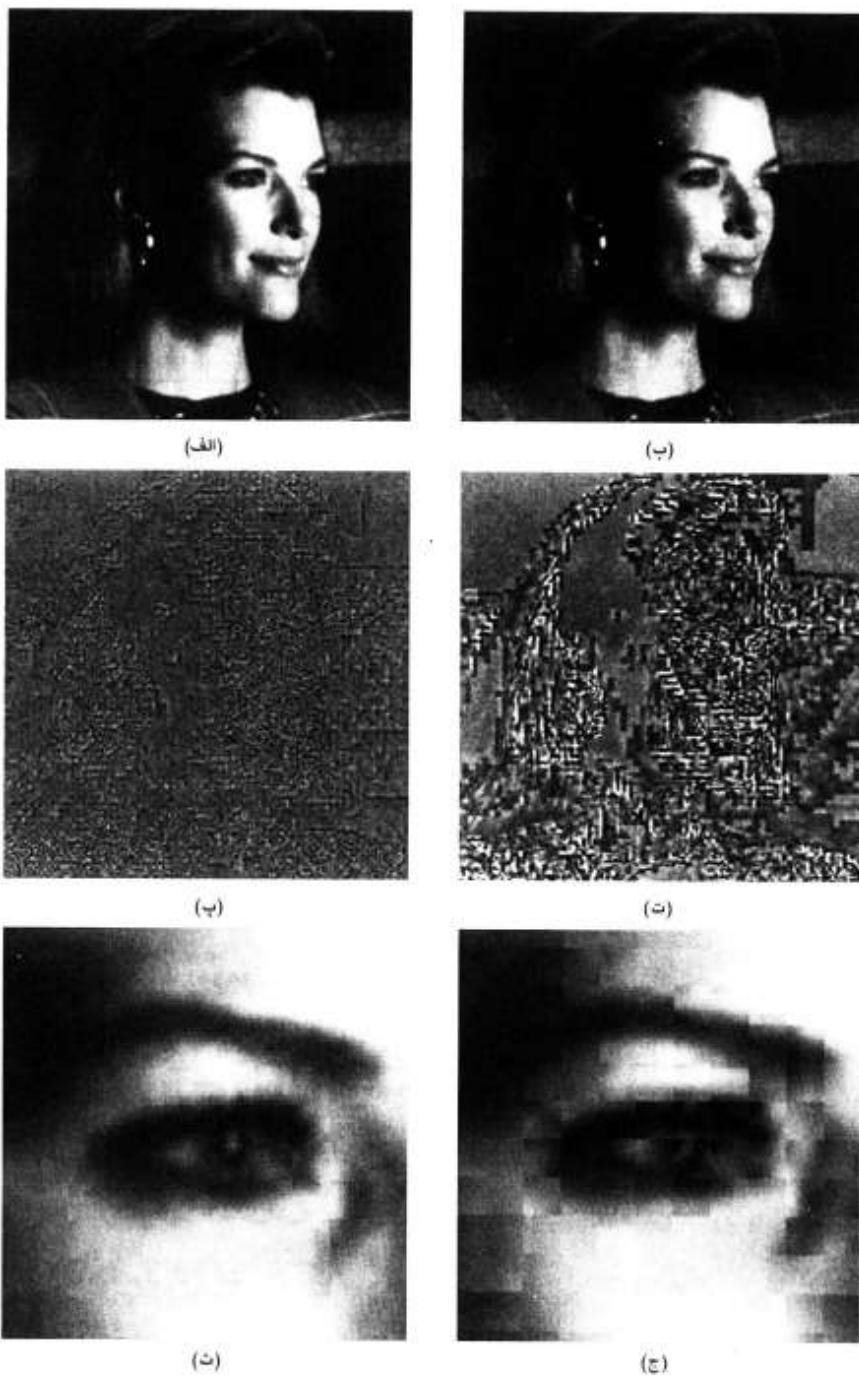
ترازسازی^۱ نوعی رانشان می دهد. این آرایه که به طور گسترده ای در استاندارد JPEG[†] استفاده شده است (بخش ۶.۶.۲ را بینید)، هر ضریب زیر تصویر را بر حسب اهمیت ادرآکی یا روان بصری حاصل از مشاهده می سنجد.

مثال: اشکال ۶.۳۶ (الف) و (ب) دو تقریب رمز آستانه ای شده از تصویر نکرنگ شکل ۶.۲۳. رانشان می دهند. هر دو تصویر با استفاده از DCT 8×8 و آرایه ترازسازی شکل ۶.۳۵ (ب) تولید شدند. نتیجه اول که یک نسبت فشرده سازی حدود ۳۴ برابر را فراهم می آورد، با اعمال مستقیم این آرایه ترازسازی بدست آمد. نتیجه دوم که تصویر اولیه را ۶۷ برابر فشرده می سازد، پس از ضرب (مقایسه کردن) آرایه ترازسازی در ۴ تولید شد. در مقایسه، متوسط نسبت فشرده سازی حاصل از کاربرد روش های بدون خطای مشروح در بخش ۶.۴، تنها ۲/۶۲ برابر است.

تفاضل های بین تصویر اولیه شکل ۶.۲۳ و تصاویر بازسازی شده اشکال ۶.۳۶ (الف) و (ب) برتری در اشکال ۶.۳۶ (ب) و (ت) نشان داده شده اند. خطاهای rms متناظر (معادله ۶.۸-۱) بینید، ۴۲، ۳۳ و ۶/۳۳ سطح خاکستری هستند. طبیعت دقیق خطاهای در تصاویر بزرگ شده اشکال ۶.۳۶ (ت) و (ج) مرئی تر است. هر یک از این تصاویر به ترتیب یک بخش بزرگ شده از اشکال ۶.۳۶ (الف) و (ب) را نشان می دهد و آنها ارزیابی بهتری از تفاوت های جزئی بین تصاویر بازسازی شده را

[†] مخفف عبارت JPEG Joint Photographic Experts Group است.

1) Normalization



شکل ۳۶.۶ (الف)، (ب)، (ث) تقریبهای شکل ۲۳.۶ با استفاده از DCT و آرایه ترازسازی از شکل ۳۵.۶ (ب)؛ (ب)، (ث)، (ج) نتایج مشابه برای ۴Z

ممکن می سازند.

۶.۶ استانداردهای فشرده سازی تصویر

بسیاری از روش های فشرده سازی بدون خطأ و با اتلاف که تاکنون توصیف شدند، نقش مهمی در توسعه و پذیرش استانداردهای اصلی فعلی برای فشرده سازی تصویر بازی کرده اند. در این بخش این استانداردها را، که در بیشتر موارد با همکاری مشترک سازمان استاندارد جهانی (ISO) و کمیته مشورتی تلفن و تلگراف جهانی (CCITT) وضع شده اند، بررسی می کنیم. این استانداردها به فشرده سازی تصویر پیوسته (نک رنگ و رنگی)، همچنین تصاویر تک قاب یا دنباله قاب^۱ (متحرک) می پردازند.

۱.۶.۶ استانداردهای فشرده سازی تصویر دو سطحی (دو دوی)

متداول ترین استانداردهای فشرده سازی این نوع تصاویر، استانداردهای CCITT گروه ۳ و ۴ برای فشرده سازی تصویر دو سطحی هستند. گرچه اکنون از آنها در انواع زیادی از کاربردهای رایانه ای استفاده می شود، اما در ابتداء فقط به عنوان روش های رمز نگاری دور نگار (FAX)^۲ برای ارسال اسناد از روی شبکه های تلفن طراحی شدند. استاندارد گروه ۳ از روش رمز نگاری RL یک بعدی غیر تطبیقی استفاده می کند که در آن آخرین $1 - K$ خط از هر گروه K خطی (برای ۴ یا ۲) به طور دلخواه با یکی از روش های دو بعدی رمز می شوند. استاندارد گروه ۴ گونه ساده شده یا خط جریانی^۳ از استاندارد گروه ۳ است که در آن تنها رمز نگاری دو بعدی مجاز است. هر دو استاندارد از روش رمز نگاری دو بعدی غیر تطبیقی یکسانی استفاده می کنند. این روش کاملاً مشابه روش رمز نگاری آدرس نسبی (RAC) که در بخش ۶.۶ توصیف شد، می باشد.

در خلال تولید استانداردهای CCITT، هشت متن نماینده برای آزمون انتخاب و به عنوان مبنای ارزیابی روش های گوناگون فشرده سازی دو دوی استفاده شدند. استانداردهای گروه ۳ و ۴ فعلی این متن را، که شامل هر دو نوع متن چاپی و دستی (به چند زبان مختلف) و حاوی تعداد کمی رسم نیز می باشند، با نسبت حدود ۱۵ برابر فشرده می سازند. چون استانداردهای گروه ۳ و ۴ براساس روش های غیر تطبیقی هستند و در بعضی حالات موجب بسط داده ها می شوند، گروه مشترک تصاویر دو سطحی^۴

(JBIG) - یک کمیته مشترک بین CCITT و ISO- موظف شده است تا یک الگوریتم تطبیقی را که بهتر از استانداردهای موجود کار کند، شناسایی نماید. سایر اهداف این کمیته مواد زیر را شامل می شوند: گسترش روش فشرده سازی گزینش شده به سایر انواع تصاویر (یعنی غیر از تصاویر آزمون)؛ استفاده از روش های ارسال و بازسازی پیش روی تصویر (مسئله ۶. ۲۲ را بینید)؛ و پردازش تصاویر نیم تن^۱ رقمی، که اغلب با استانداردهای گروه ۳ و ۴ فعلی بسط می یابند (به جای این که فشرده شوند).

فسرده سازی یک بعدی

در روش فشرده سازی یک بعدی گروه ۳ CCITT هر خط تصویر^۲ به صورت دنباله ای از کلمات رمز با طول متغیر که این کلمات رمز بیانگر طول دنباله های سفید و سیاه در پیمایش چپ به راست آن خط هستند، رمز می شود. خود کلمات رمز، دونوع هستند. اگر طول دنباله کمتر از ۶۳ باشد، بایک رمز پایان دهنده از رمز هافمن تغییر یافته جدول ۶. ۱۱ رمز می شود. اگر طول دنباله بزرگتر از ۶۳ باشد، بزرگترین رمز ترکیبی ممکن (که از طول دنباله تجاوز نمی کند) از جدول ۶. ۱۲ همراه با یک رمز پایان دهنده که بیانگر تفاضل بین رمز ترکیبی و طول دنباله واقعی است، استفاده می شود. در این استاندارد لازم است که هر خط با یک کلمه رمز RL سفید که ممکن است ۰۰۱۱۰۱۰۱ یعنی رمز دنباله سفید به طول صفر باشد، شروع شود. نهایتاً یک کلمه رمز پایان خط^۳ (EOL) ۰۰۰۰۰۰۰۰۰۰۰۰۱ برای پایان دادن هر خط و نیز اعلام اولین خط هر تصویر جدید استفاده می شود. انتهای یک دنباله از تصاویر باشش EOL متوالی نشان داده می شود.

فسرده سازی دو بعدی

روش فشرده سازی دو بعدی اتخاذ شده برای هر دو استاندارد گروه ۳ و ۴ CCITT یک روش خط به خط^۴ است که در آن محل هر گذر^۵ از دنباله سیاه به سفید یا برعکس بر حسب موقعیت عنصر مرجع^۶ که روی

^۱ در این استاندارد به تصاویر با عنوان صفحه ها و به دنباله ای از تصاویر با عنوان مستندها اشاره می شود.

1) Half-tone- documents

2) End -of- line

3) Line -by- line

4) Transition

5) Reference element

۶.۶ استانداردهای فشرده سازی تصویر ۴۷۱

جدول ۱۱.۶ رمزهای پایان دهنده CCITT

کلمه رمزسیاه	کلمه رمزنگاری	کلمه رمزنگاری	کلمه رمزسیاه	کلمه رمزنگاری	کلمه رمزنگاری
طول دبیاله	طول دبیاله	طول دبیاله	طول دبیاله	طول دبیاله	طول دبیاله
0	00110101	0000110111	32	00011011	000001101010
1	000111	010	33	00010010	000001101011
2	0111	11	34	00010011	000011010010
3	1000	10	35	00010100	000011010011
4	1011	011	36	00010101	000011010100
5	1100	0011	37	00010110	000011010101
6	1110	0010	38	00010111	000011010110
7	1111	00011	39	00101000	000011010111
8	10011	000101	40	00101001	000001101100
9	10100	000100	41	00101010	000001101101
10	00111	0000100	42	00101011	000011011010
11	01000	0000101	43	00101100	000011011011
12	001000	0000111	44	00101101	000001010100
13	000011	00000100	45	00000100	000001010101
14	110100	00000111	46	00000101	000001010110
15	110101	000011000	47	000001010	000001010111
16	101010	0000010111	48	000001011	000001100100
17	101011	0000011000	49	01010010	000001100101
18	0100111	0000001000	50	01010011	000001010010
19	0001100	0000110011	51	01010100	000001010011
20	0001000	00001101000	52	01010101	000000100100
21	0010111	00001101100	53	00100100	000000110111
22	0000011	00000110111	54	00100101	000000111000
23	0000100	00000101000	55	01011000	000000100111
24	0101000	00000010111	56	01011001	000000101000
25	0101011	00000011000	57	01011010	000001011000
26	0010011	000011001010	58	01011011	000001011001
27	0100100	000011001011	59	01001010	000000101011
28	0011000	000011001100	60	01001011	000000101100
29	00000010	000011001101	61	00110010	000001011010
30	00000011	000001101000	62	00110011	000001100110
31	00011010	000001101001	63	00110100	000001100111

خط در حال رمزنگاری^۱ قرار دارد، رمز می شود. خط قبل از رمز شده خط مرجع^۲ خوانده می شود؛ خط مرجع برای اولین خط هر تصویر جدید یک خط سفید فرضی است.

شکل ۶.۳۷. فرآیند رمزنگاری پایه ای برای یک خط پیمایش را نشان می دهد. توجه شود که مراحل اولیه فرآیند، تعیین محل چند گذر مهم یا عناصر تغییر^۱ a_0, a_1, a_2, b_1, b_2 می باشند. عنصر تغییر به عنوان پیکسلی تعریف می شود که مقدارش متفاوت از مقدار پیکسل قبلی روی همان خط باشد. مهمترین عنصر تغییر، a_0 (عنصر مرجع) است که یا در محل یک عنصر تغییر سفید فرضی در سمت چپ اولین پیکسل از

1) Coding line

2) Reference line

3) Changing element

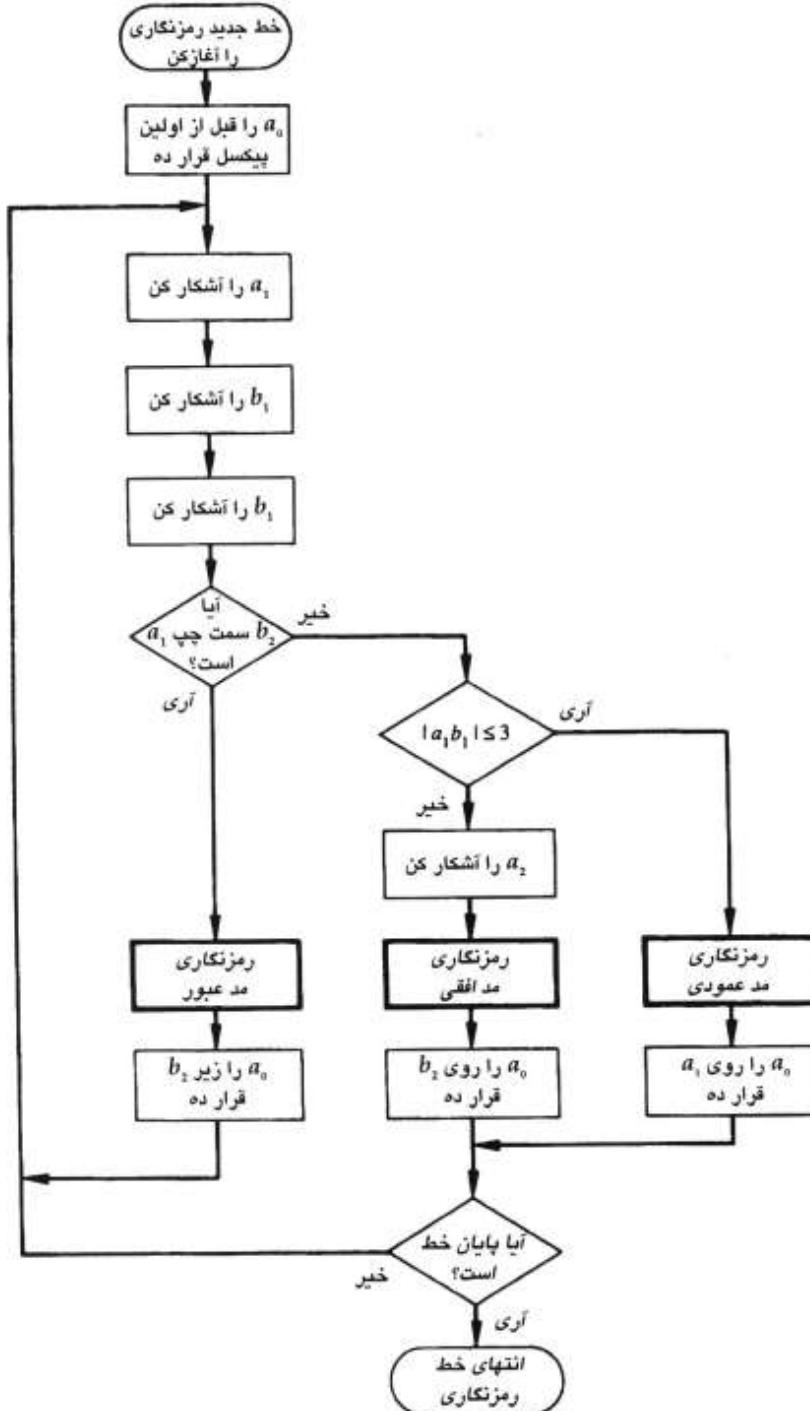
جدول ۱۲۰۶ رمز ترکیبی CCITT

طول دنباله	کلمه رمز سفید	کلمه رمز سیاه	طول دنباله	کلمه رمز سفید	کلمه رمز سیاه
64	11011	0000001111	960	011010100	00000001110011
128	10010	000011001000	1024	011010101	0000001110100
192	010111	000011001001	1088	011010110	0000001110101
256	0110111	000001011011	1152	011010111	0000001110110
320	00110110	000000110011	1216	011011000	0000001110111
384	00110111	000000110100	1280	011011001	0000001010010
448	01100100	000000110101	1344	011011010	0000001010011
512	01100101	0000001101100	1408	011011011	0000001010100
576	01101000	0000001101101	1472	010011000	0000001010101
640	01100111	0000001001010	1536	010011001	0000001011010
704	011001100	0000001001011	1600	010011010	0000001011011
768	011001101	0000001001100	1664	0110000	0000001100100
832	011010010	0000001001101	1728	010011011	0000001100101
896	011010011	0000001110010			
کلمه رمز			کلمه رمز		
1792	000000001000		2240	0000000010110	
1856	000000001100		2304	0000000010111	
1920	000000001101		2368	0000000011100	
1984	0000000010010		2432	0000000011101	
2048	0000000010011		2496	0000000011110	
2112	0000000010100		2560	0000000011111	
2176	0000000010101				

هر خط رمز نگاری جدید یا از مد رمز نگاری قبلی تعیین می شود. پس از این که محل a_0 تعیین شد، به عنوان عنصر تغییر بعدی در سمت راست a_1 روی خط رمز نگاری فعلی، a_2 به عنوان عنصر تغییر با مقدار مخالف (b_1) در سمت راست a_1 روی خط رمز نگاری فعلی، b_2 به عنوان عنصر تغییر بعدی در سمت راست b_1 روی خط مرجع شناسایی می شود. اگر یکی از این عناصر تغییر، آشکار نشود، محل آن نسبت به یک پیکسل فرضی در سمت راست آخرین پیکسل، روی خط مربوطه تعیین می شود. شکل ۶.۳۸. دو نمایش از روابط کلی بین انواع عناصر تغییر را ارائه می نماید.

پس از شناسایی عنصر مرجع فعلی و عناصر تغییر مربوطه، دو آزمون ساده برای انتخاب یکی از سه مد ممکن رمز نگاری یعنی مد عبور^۱، مد عمودی^۲ مد افقی^۳ انجام می شوند. آزمون ابتدایی که متناظر با اولین نقطه انشعاب در نمودار جریان شکل ۶.۳۷. می باشد، محل b_1 را با محل a_1 مقایسه می کند.

۶۰. استانداردهای فشرده سازی تصویر ۴۷۳

شکل ۳۷.۶ روال رمزگاری دو بعدی CCITT . عبارت $|a_i b_j|$ نمایانگر قدر مطلق فاصله بین عناصر تغییر a_i و b_j است .

آزمون دوم که متناظر با دو مین نقطه انشعاب در شکل ۶.۳۷ است، فاصله (بر حسب پیکسل) بین محل های a_1 و b_1 را محاسبه می کند و آن را با ۳ مقایسه می نماید. بسته به نتیجه این آزمونها، رمزگذار به یکی از سه بلوک رمزگاری ترسیم شده در شکل ۶.۳۷ وارد و رمزگاری مناسب اجرا می شود. آنگاه طبق نمودار جریان، جهت آمادگی برای گام بعدی رمزگاری، یک عنصر مرجع جدید تعیین می شود.

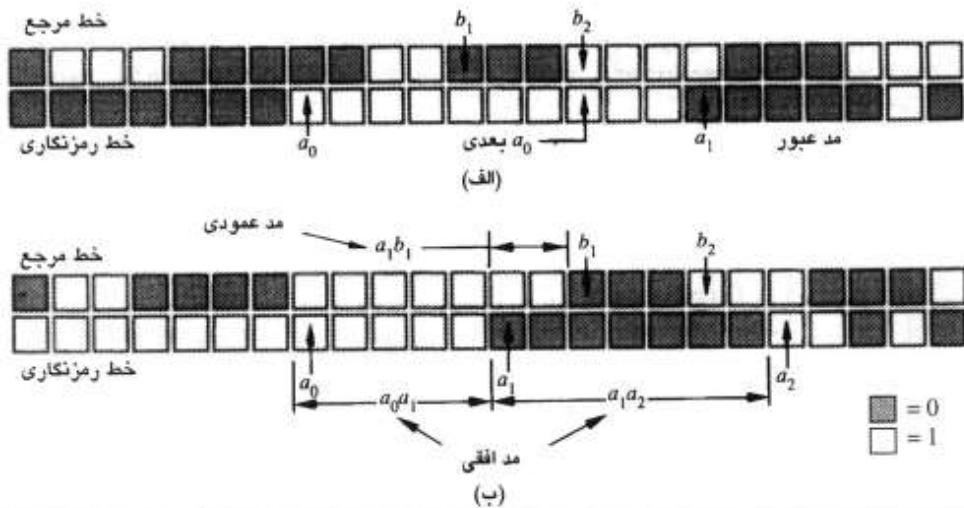
جدول ۶.۱۳. رمزهای خاص مورد استفاده برای هریک از سه مد ممکن رمزگاری را تعریف می کند. در مد عبور، که به طور مشخص حالتی را که b_1 مستقیماً بالای a_1 است، شامل نمی شود، تنها یک کلمه رمز مد عبور یعنی 0001 مورد نیاز است. همان طور که شکل ۶.۳۸. (الف) نشان می دهد، این مد دنباله های سفید یا سیاه خط مرجع را که با دنباله های سفید یا سیاه خط رمزگاری فعلی همپوشانی ندارند، شناسایی می کند. در مد رمزگاری افقی، فواصل a_1 و a_2 باید طبق رمزهای پایان دهنده^۱ و ترکیبی^۲ جداول ۶.۱۱ و ۶.۱۲ رمز شوند و سپس به کلمه رمز مدققی یعنی 001 افزوده شوند. این عمل در جدول ۶.۱۳. با عبارت $M(a_0a_1) + M(a_1a_2) + M(a_0a_2)$ نشان داده می شود، که $a_0a_1a_2$ به غیر فشرده استفاده می شود.

مثال: گرچه شکل ۶.۳۸. (ب) با عوامل به کار رفته در هر دو مد رمزگاری افقی و عمودی حاشیه تویسی شده است، در واقع وضعیت ترسیم شده حالتی از رمزگاری مد عمودی است. یعنی چون b_1 در سمت راست a_1 است، در اولین آزمون (یا مد عبور) موجود در شکل ۶.۳۷. رد می شود. آزمون دوم که تعیین می کند آیا به مد رمزگاری عمودی یا افقی وارد شویم، نشان می دهد که باید رمزگاری مد عمودی بکار رود، زیرا فاصله a_1 تا b_1 کمتر از ۳ است. طبق جدول ۶.۱۳. کلمه رمز مناسب 000010 می باشد که اشاره دارد بر این که a_1 به فاصله^۳ دوپیکسل در سمت چپ b_1 قرار دارد. برای آمادگی جهت تکرار بعدی رمزگاری، a_0 به محل a_1 انتقال می یابد.



۶.۶ استانداردهای فشرده سازی تصویر

۳۷۵



شکل ۶.۶۸ عامل های رمزگاری در (الف) مُد گذر و (ب) مدهای افقی و عمودی CCITT

۶.۶.۶ استانداردهای فشرده سازی تصویر پیوسته

CCITT و ISO چند استاندارد فشرده سازی تصویر پیوسته (در مقابل دو سطحی) را تعریف کرده اند. این استانداردها که در مراحل گوناگونی از فرآیند پذیرش هستند، به فشرده سازی هر دو نوع تصویر تکرنگ و رنگی، و نیز تصاویر ثابت و دنباله ای می پردازند. برخلاف استانداردهای فشرده سازی تصویر دو سطحی توصیف شده در بخش ۶.۶.۱، تمام استانداردهای تصویر پیوسته براساس روش های رمزگاری تبدیلی با اتلاف بخش ۶.۵.۲ هستند. برای توسعه استانداردها، کمیته های CCITT و ISO الگوریتم های پشنهدایی را از تعداد زیادی شرکت، دانشگاه و آزمایشگاه تحقیقاتی جمع آوری کردند.

جدول ۶.۱۳ جدول رمز دو بعدی CCITT

مد	کلمه رمز
0001	عبور
001 + $M(a_0 a_1)$ + $M(a_1 a_2)$	افقی
	عمودی
1	a_1 زیر b_1 است
011	a_1 یک پیکسل در سمت راست b_1 قرار دارد.
000011	دو پیکسل در سمت راست b_1 قرار دارد.
0000011	سه پیکسل در سمت راست b_1 قرار دارد.
010	a_1 یک پیکسل در سمت چپ b_1 قرار دارد.
000010	دو پیکسل در سمت چپ b_1 قرار دارد.
0000010	سه پیکسل در سمت چپ b_1 قرار دارد.
0000001xxx	...

پس بهترین آنها، براساس کیفیت تصویر و کارایی فشرده سازی، انتخاب شدند. بنابراین استانداردهای حاصل، نشانگر حداعلای رشد در فشرده سازی تصویر پیوسته هستند. این استانداردها تصویرهای ثابت و دنباله‌ای سازگار با VHS[†] را با نسبتها فشرده سازی به ترتیب حدود ۲۵ برابر و ۱۰۰ برابر با کیفیت عرضه می‌کنند.

فشرده سازی تصویر ثابت تکرینک و رنگی

CCITT و ISO با همکاری هم متدالول ترین و جامع ترین استاندارد فشرده سازی تصویر ثابت و پیوسته را که استاندارد JPEG خوانده می‌شود، ایجاد کردند. استاندارد JPEG سه سیستم رمزگاری متفاوت را تعریف می‌کند: (۱) سامانه رمزگاری مبنای^۱ با اتلاف که براساس DCT است و برای بیشتر کاربردهای فشرده سازی کافی است؛ (۲) سامانه رمزگاری گسترش یافته^۲ برای فشرده سازی بیشتر، دقت بالاتر و کاربردهای بازسازی پیشرو^۳؛ و (۳) سامانه رمزگاری مستقل^۴ بی اتلاف برای فشرده سازی معکوس پذیر. هرمحصول یا سامانه برای سازگاری با JPEG، باید سامانه مبنای پشتیبانی کند. در این استاندارد هیچ قالب فایل، تفکیک مکانی یا مدل فضای رنگ خاصی مشخص نمی‌شود.

در سامانه مبنای که اغلب سامانه مبنای دنباله‌ای^۵ خوانده می‌شود، دقت داده‌های ورودی و خروجی به ۸ بیت محدود است، در حالی که دقت مقادیر DCT چندی شده به ۱۱ بیت محدود است. عمل فشرده سازی در سه مرحله متوالی انجام می‌شود: محاسبه DCT، چندی سازی، و انتساب رمز طول متغیر. ابتدا تصویر به بلوک‌های پیکسلی با ابعاد 8×8 تقسیم می‌شود، که این بلوک‌ها از چپ به راست و از بالا به پایین پردازش می‌شوند. در پردازش هر بلوک 8×8 یا زیرتصویر، ۶۴ پیکسل آن با تفریق مقدار^{-۱}^۶، انتقال سطح داده می‌شوند، که^۷ بیشینه تعداد سطوح خاکستری است. آنگاه تبدیل کسینوسی گستته دو بعدی محاسبه شده، طبق معادله $(6.5 - 30)$ چندی شده، و با استفاده از الگوی زیگزاگ شکل^۶. (ت) بازآرایی می‌شود تا یک دنباله یک بعدی از ضرایب چندی شده را تشکیل دهد.

[†] متدالول ترین قالب برای ضبط روی نوار ویدیویی (VCR) خانگی نیم اینچی است.

- | | | |
|------------------------------|-------------------------------|----------------|
| 1) baseline coding system | 2) Extended coding system | 3) Progressive |
| 4) Independent coding system | 5) Sequential baseline system | |

۶.۶ استانداردهای فشرده سازی تصویر ۳۷۷

جدول ۱۴.۶ دسته های رمزگاری ضرایب JPEG

محدوده	دسته تفاضل DC	دسته تفاضل AC
0	0	N/A
-1, 1	1	1
-3, -2, 2, 3	2	2
-7, ..., -4, 4, ..., 7	3	3
-15, ..., -8, 8, ..., 15	4	4
-31, ..., -16, 16, ..., 31	5	5
-63, ..., -32, 32, ..., 63	6	6
-127, ..., -64, 64, ..., 127	7	7
-255, ..., -128, 128, ..., 255	8	8
-511, ..., -256, 256, ..., 511	9	9
-1023, ..., -512, 512, ..., 1023	A	A
-2047, ..., -1024, 1024, ..., 2047	B	B
-4095, ..., -2048, 2048, ..., 4095	C	C
-8191, ..., -4096, 4096, ..., 8191	D	D
-16383, ..., -8192, 8192, ..., 16383	E	E
-32767, ..., -16384, 16384, ..., 32767	F	N/A

از آنجا که آرایه' یک بعدی حاصل که با الگوی زیگزاگ شکل ۶.۴ (ت) تولید شد، طبق فرکанс مکانی صعودی مرتب شده است، روال رمزگاری JPEG طوری طراحی می شود که از ذیباله های طویل صفر که معمولاً حاصل ترتیب جدید هستند، بهره بگیرد. بویژه ضرایب AC[†] غیر صفر با استفاده از رمز طول متغیری که مقدار هر ضریب و تعداد صفرهای قبل از آن را تعیین می کند، رمز می شوند. ضریب DC نسبت به ضریب DC زیر تصویر قبلی رمز تفاضلی می شود. جداول ۱۴.۶، ۱۵.۶، ۱۶.۶، چندی سازی روشنایی در شکل ۶.۵ (ب) آمده است و می توان آن را مقایس نمود تا سطوح فشرده سازی

جدول ۱۵.۶ رمز DC پس گزیده JPEG (لومینانس)

دسته	رمز پایه	طول	دسته	رمز پایه	طول
0	010	3	6	1110	10
1	011	4	7	11110	12
2	100	5	8	111110	14
3	00	5	9	1111110	16
4	101	7	A	11111110	18
5	110	8	B	111111110	20

[†] در این استاندارد عبارت AC نمایانگر تمام ضرایب تبدیل، به غیر از ضریب DC است.

جدول ۱۶.۶ رمز AC پیش‌گزینه JPEG (لومیننس)

دسته/دبیله	رمز پایه	طول	دسته/دبیله	رمز پایه	طول
۰/۰	1010 (= EOB)	4			
۰/۱	00	3	8/1	11111010	9
۰/۲	01	4	8/2	11111111000000	17
۰/۳	100	6	8/3	111111110110111	19
۰/۴	1011	8	8/4	111111110111000	20
۰/۵	11010	10	8/5	111111110111001	21
۰/۶	111000	12	8/6	111111110111010	22
۰/۷	1111000	14	8/7	111111110111011	23
۰/۸	111110110	18	8/8	111111110111100	24
۰/۹	111111110000010	25	8/9	111111110111101	25
۰/A	111111110000011	26	8/A	111111110111110	26
۱/۱	1100	5	9/1	111111000	10
۱/۲	111001	8	9/2	111111110111111	18
۱/۳	1111001	10	9/3	111111111000000	19
۱/۴	111110110	13	9/4	1111111111000001	20
۱/۵	1111110110	16	9/5	1111111111000010	21
۱/۶	111111110000100	22	9/6	1111111111000011	22
۱/۷	111111110000101	23	9/7	1111111111000100	23
۱/۸	111111110000110	24	9/8	1111111111000101	24
۱/۹	1111111110000111	25	9/9	11111111111000110	25
۱/A	1111111110001000	26	9/A	1111111111000111	26
۲/۱	11011	6	A/1	111111001	10
۲/۲	1111000	10	A/2	1111111111001000	18
۲/۳	1111110111	13	A/3	11111111111001001	19
۲/۴	111111110001001	20	A/4	11111111111001010	20
۲/۵	111111110001010	21	A/5	11111111111001011	21
۲/۶	1111111110001011	22	A/6	11111111111001100	22
۲/۷	11111111110001100	23	A/7	11111111111001101	23
۲/۸	11111111110001101	24	A/8	11111111111001110	24
۲/۹	11111111110001110	25	A/9	11111111111001111	25
۲/A	11111111110001111	26	A/A	11111111111010000	26
۳/۱	111010	7	B/1	111111010	10
۳/۲	111110111	11	B/2	1111111111010001	18
۳/۳	1111110111	14	B/3	11111111111010010	19
۳/۴	1111111110010000	20	B/4	11111111111010011	20
۳/۵	1111111110010001	21	B/5	11111111111010100	21
۳/۶	1111111110010010	22	B/6	11111111111010101	22
۳/۷	1111111110010011	23	B/7	11111111111010110	23
۳/۸	1111111110010100	24	B/8	11111111111010111	24
۳/۹	1111111110010101	25	B/9	11111111111011000	25
۳/A	1111111110010110	26	B/A	11111111111011001	26

۶.۶ استانداردهای فشرده سازی تصویر

جدول ۱۶.۶ ادامه

دسته/دبیله	رمز پایه	طول	دسته/دبیله	رمز پایه	طول
4/1	111011	7	C/1	1111111010	11
4/2	1111111000	12	C/2	1111111111011010	18
4/3	1111111110010111	19	C/3	11111111111011011	19
4/4	11111111110011000	20	C/4	11111111111011100	20
4/5	11111111110011001	21	C/5	11111111111011101	21
4/6	11111111110011010	22	C/6	11111111111011110	22
4/7	11111111110011011	23	C/7	11111111111011111	23
4/8	11111111110011100	24	C/8	11111111111100000	24
4/9	11111111110011101	25	C/9	11111111111100001	25
4/A	11111111110011110	26	C/A	11111111111100010	26
5/1	1111010	8	D/1	11111111010	12
5/2	1111111001	12	D/2	1111111111100011	18
5/3	1111111110011111	19	D/3	11111111111100100	19
5/4	11111111110100000	20	D/4	111111111111100101	20
5/5	11111111110100001	21	D/5	111111111111100110	21
5/6	11111111110100010	22	D/6	111111111111100111	22
5/7	11111111110100011	23	D/7	111111111111101000	23
5/8	11111111110100100	24	D/8	1111111111111101001	24
5/9	11111111110100101	25	D/9	1111111111111101010	25
5/A	11111111110100110	26	D/A	1111111111111101011	26
6/1	1111011	8	E/1	111111110110	13
6/2	111111111000	13	E/2	1111111111101100	18
6/3	1111111111010011	19	E/3	11111111111101101	19
6/4	11111111110101000	20	E/4	111111111111101110	20
6/5	11111111110101001	21	E/5	111111111111101111	21
6/6	11111111110101010	22	E/6	111111111111110000	22
6/7	11111111110101011	23	E/7	111111111111110001	23
6/8	11111111110101100	24	E/8	111111111111110010	24
6/9	11111111110101101	25	E/9	111111111111110011	25
6/A	11111111110101110	26	E/A	111111111111110100	26
7/1	11111001	9	F/0	11111111110111	12
7/2	11111111001	13	F/1	11111111111110101	17
7/3	1111111110101111	19	F/2	11111111111111010	18
7/4	11111111110110000	20	F/3	111111111111110111	19
7/5	11111111110110001	21	F/4	11111111111111000	20
7/6	11111111110110010	22	F/5	11111111111111001	21
7/7	11111111110110011	23	F/6	11111111111111010	22
7/8	11111111110110100	24	F/7	11111111111111011	23
7/9	11111111110110101	25	F/8	11111111111111100	24
7/A	11111111110110110	26	F/9	11111111111111101	25
			F/A	11111111111111110	26

متفاوتی حاصل شود. گرچه جداول رمزنگاری پیش گزیده و آرایه های چندی سازی تصویب شده برای پردازش هر دو مؤلفه روشنایی و رنگینگی^۱ فراهم هستند، کاربر آزاد است که جداول و/یا آرایه های دلخواه خود را بسازد که در عمل ممکن است این آرایه ها و جداول با مشخصات تصویری که فشرده می شود، تطبیق یابند.

مثال: فشرده سازی و بازسازی زیر تصویر 8×8 زیر با استاندارد JPEG مبنای در نظر بگیرید:

52	55	61	66	70	61	64	73
63	59	66	90	109	85	69	72
62	59	68	113	144	104	66	73
63	58	71	122	154	106	70	69
67	61	68	104	126	88	68	70
79	65	60	70	77	68	58	75
85	71	64	59	55	61	65	83
87	79	69	68	65	76	78	94

تصویر اولیه متشکل از 2^8 یا 256 سطح خاکستری ممکن می باشد، بنابراین فرآیند رمزنگاری با انتقال سطح پیکسل های زیر تصویر اولیه به اندازه 2^7 - یا 128 - سطح خاکستری شروع می شود. آرایه حاصل عبارتست از

-76	-73	-67	-62	-58	-67	-64	-55
-65	-69	-62	-38	-19	-43	-59	-56
-66	-69	-60	-15	16	-24	-62	-55
-65	-70	-57	-6	26	-22	-58	-59
-61	-67	-60	-24	-2	-40	-60	-58
-49	-63	-68	-58	-51	-65	-70	-53
-43	-57	-64	-69	-73	-67	-63	-45
-41	-49	-59	-60	-63	-52	-50	-34

وقتی زیر تصویر فوق بوسیله DCT مستقیم معادله (۳.۵-۴۸) بازی $N = 8$ تبدیل شود، به صورت ذیل در می آید:

۶.۶ استانداردهای فشرده سازی تصویر
۳۸۱

-415	-29	-62	25	55	-20	-1	3
7	-21	-62	9	11	-7	-6	6
-46	8	77	-25	-30	10	7	-5
-50	13	35	-15	-9	6	0	3
11	-8	-13	-2	-1	1	-4	1
-10	1	3	-3	-1	0	2	-1
-4	-1	2	-1	2	-3	1	-2
-1	-1	-1	-2	-1	-1	0	-1

اگر از آرایه ترازساز پیشنهادی JPEG در شکل ۶.۳۵(ب) برای چندی سازی آرایه فوق استفاده شود، ضرایب حاصل پس از مقیاس کردن و قطع کردن (یعنی ضرایب تراز شده طبق معادله ۶.۳۰-۵) عبارتند از:

-26	-3	-6	2	2	0	0	0
1	-2	-4	0	0	0	0	0
-3	1	5	-1	-1	0	0	0
-4	1	2	-1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

برای نمونه ضریب DC بدین صورت محاسبه می شود

$$\begin{aligned}\hat{T}(0, 0) &= \text{round} \left[\frac{T(0, 0)}{Z(0, 0)} \right] \\ &= \text{round} \left[\frac{-415}{16} \right] = -26.\end{aligned}$$

توجه شود که فرآیند تبدیل و ترازسازی باعث می شود که تعداد زیادی از ضرایب با مقدار صفر تولید شوند. وقتی این ضرایب طبق الگوی زیگزاگ شکل ۶.۳۴(ت) مرتب شوند، دنباله ضرایب یک بعدی حاصل عبارتست از:

[-26 -3 1 -3 -2 -6 2 -4 1 -4 1 1 5 0 2 0
0 -1 2 0 0 0 0 -1 -1 EOB]

که نماد EOB نشانه پایان بلوک است. یک کلمه رمزهافمن ویژه برای EOB تهیه شده است (دسته ۰ و

طول دنباله ۰ را در جدول ۶. ۱۶. (بیینید) تا نشان دهد که سایر ضرایب دنباله بازآرایی شده صفر هستند. تولید رمز JPEG اپنگزیده برای دنباله ضرایب بازآرایی شده، با محاسبه تفاضل بین ضریب DC فعلی و ضریب DC زیرتصویر قبلی آغاز می شود. چون در اینجا زیرتصویر از شکل ۶. ۲۳. است و ضریب DC نهایی زیرتصویر سمت چپش نیز ۱۷- بوده است، تفاضل DPCM حاصل [۱۷-۲۶-] یا ۹- است و بنابراین در دسته ۴ تفاضل DC از جدول ۶. ۱۴. فرار می گیرد. طبق رمز تفاضلی هافمن پیش گزیده از جدول ۶. ۱۵.، رمز پایه مناسب برای یک تفاضل دسته ۴، رمز ۳ بیتی ۱۰۱ می باشد، از آنجا که طول کل یک ضریب دسته ۴ کاملاً رمز گذاری شده ۷ بیت است، ۴ بیت باقیمانده باید از بیت های پایین مرتبه (LSB) مقدار تفاضل تولید شوند. برای مقدار دسته تفاضل DC دلخواه (مثلاً دسته K)، بیت اضافی مورد نیاز است که یا برابر K بیت LSB تفاضل مثبت یا برابر K بیت LSB تفاضل منفی منتهای یک تعیین می شود. برای یک تفاضل ۹-، بیت های LSB مناسب ۱- (0111) یا ۰110 است و کلمه رمز DC که رمز DPCM شده است، 1010110 می باشد.

ضرایب AC غیر صفر آرایه باز مرتب شده به طور مشابهی با استفاده از جداول ۶. ۱۴. و ۶. ۱۶. رمز می شوند. تفاوت اصلی این است که هر کلمه رمز هافمن AC پیش گزیده به تعداد ضرایب صفر قبل از ضریب غیر صفری که باید رمز شود و نیز دسته آن ضریب غیر صفر بستگی دارد. (ستون با برچسب Run/Category در جدول ۶. ۱۶. را بیینید). بنابراین اولین ضریب AC غیر صفر در آرایه باز مرتب شده، یعنی ۳-، به صورت ۰100 رمز می شود. ۲ بیت اول این رمز نشان می دهد که ضریب یاد شده در دسته ۲ قرار دارد و قبل از آن هیچ ضریب صفری نیست (جدول ۶. ۱۴. را بیینید)؛ ۲ بیت آخر این رمز با همان فرآیند مورد استفاده برای به دست آوردن بیت های LSB رمز تفاضلی DC تولید می شوند. با ادامه این روش، آرایه کاملاً رمز شده عبارتست از

1010110	0100	001	100011	001	100011	001	1010	0100	001	100011	001	1010110	001	100101	11100110	110110	0110	11110100	000	1010
---------	------	-----	--------	-----	--------	-----	------	------	-----	--------	-----	---------	-----	--------	----------	--------	------	----------	-----	------

که فاصله های خالی تنها برای خوانایی درج شده اند. گرچه در این مثال مورد نیاز نبود، رمز JPEG پیش گزیده دارای یک کلمه رمز خاص برای دنباله ۱۵ صفر که پس از آن نیز یک صفر قرار دارد، می باشد (دسته ۰ و طول دنباله F در جدول ۶. ۱۶. را بیینید). تعداد کل بیتها آرایه بازآرایی شده و رمز شده (و بنابراین تعداد بیت های مورد نیاز برای نمایش یک زیرتصویر ۸ بیتی 8×8 از این مثال) ۹۲ بیت است.

۶.۶ استانداردهای فشرده سازی تصویر ۴۸۳

نسبت فشرده سازی حاصل $\frac{512}{92}$ یا حدود ۵/۶ برابر است.

برای وافشارده سازی یک زیر تصویر فشرده JPEG، در ابتدا باید رمزبردارها ضرایب تبدیل تراز شده ای که رشته بیتی فشرده را نتیجه دادند، دوباره ایجاد نماید. چون یک دنباله دودویی رمز هافمن، آنی^۱ و قابل رمزبرداری یکتا^۲ است، این مرحله به سادگی با روش ساده جدول جستجو^۳ انجام می شود. آرایه بازسازی شده ضرایب چندی شده عبارتست از:

-26	-3	-6	2	2	0	0	0
1	-2	-4	0	0	0	0	0
-3	1	5	-1	-1	0	0	0
-4	1	2	-1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

پس از واتراز سازی آرایه فوق طبق معادله (۶.۵-۳۲) داریم:

-416	-33	-60	32	48	0	0	0
12	-24	-56	0	0	0	0	0
-42	13	80	-24	-40	0	0	0
-56	17	44	-29	0	0	0	0
18	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

که به عنوان مثال ضریب DC به صورت زیر محاسبه می شود

$$\hat{T}(0, 0) = \hat{T}(0, 0) Z(0, 0) = (-26)(16) = -416.$$

با گرفتن DCT معکوس از آرایه واتراز شده طبق معادله (۳.۵-۴۹)، آرایه زیر بدست می آید:

-70	-64	-61	-64	-69	-66	-58	-50
-72	-73	-61	-39	-30	-40	-54	-59

پردازش تصویر رقمه‌ی ۴۸۴

-68	-78	-58	-9	13	-12	-48	-64
-59	-77	-57	0	22	-13	-51	-60
-54	-75	-64	-23	-13	-44	-63	-56
-52	-71	-72	-54	-54	-71	-71	-54
-45	-59	-70	-68	-67	-67	-61	-50
-35	-47	-61	-66	-60	-48	-44	-44

در نهایت با انتقال سطح هر پیکسل آرایه بدست آمده به اندازه $2^7 \times 2^7$ (یا 128×128) سطح خاکستری زیر تصویر بازسازی شده بدین صورت بدست می‌آید:

58	64	67	64	59	62	70	78
56	55	67	89	98	88	74	69
60	50	70	119	141	116	80	64
69	51	71	128	149	115	77	68
74	53	64	105	115	84	65	72
76	57	56	74	75	57	57	74
83	69	59	60	61	61	67	78
93	81	67	62	69	80	84	84

تفاوت بین زیر تصویر اولیه و زیر تصویر بازسازی شده نتیجه وجود اتلاف فرآیند فشرده سازی و افشرده سازی JPEG است. در این مثال خطاهای در محدوده $-14 \leq \text{خطا} \leq 11$ هستند و به صورت زیر توزیع شده‌اند:

-6	-9	-6	2	11	-1	-6	-5
7	4	-1	1	11	-3	-5	3
2	9	-2	-6	-3	-12	-14	9
-6	7	0	-4	-5	-9	-7	1
-7	8	4	-1	11	4	3	-2
3	8	4	-4	2	11	1	1
2	-2	5	-1	-6	0	-2	5
-6	-2	2	6	-4	-4	-6	10

کل خطای ریشه میانگین مربع فرآیند فشرده سازی و بازسازی تقریباً $5/9$ سطح خاکستری است. \square
زیر تصویر بازسازی شده در مثال قبل از نظر مکانی تقریباً در مرکز چشم راست تصویر زن در

۶.۶ استانداردهای فشرده سازی تصویر

۴۸۵

شکل ۶.۳۶ (الف) قرار دارد. توجه شود که هر دو زیر تصویر اولیه و بازسازی شده یک قله سطح خاکستری در سطح چهارم و ستون پنجم دارد، جایی که نوری به مردمک چشم زن منعکس شده است. این قله محلی باعث شد که خطای ریشه میانگین مریع زیر تصویر بازسازی شده به مقدار زیادی از خطای کل تصویر وافسرده شده تجاوز نماید. در واقع این خطای تقریباً دو برابر بزرگی خطای مربوط به شکل ۶.۳۶ (الف) که آنهم با الگوریتم JPEG مبنا فشرده شده است، می باشد. علت این است که بسیاری از زیر تصویرهای اولیه تقریباً ثابت هستند و می توان آنها را با مقدار کمی اعوجاج نمایش داد. نتایج دیگری از JPEG در شکل ۶.۳۶ (ب) و پلیت X آمده اند. در پلیت X یک تصویر اولیه تمام رنگی و یک تصویر که با استفاده از JPEG ۶۶ برابر فشرده شده است، دیده می شوند. به اثر جعبه ای ضعیف موجود در پارچه های میخکی رنگ و به پدیده حلقه زنی در تعدادی از گلهای سفید روی زمینه سیاه توجه نماید. وقتی که میزان فشرده سازی کمتر از ۲۵ برابر باشد، اثرات یاد شده به طور مجازی نامحسوس هستند.

فشرده سازی تصاویر تکرنگ و رنگی دنباله ای

تنها استاندارد رسمی برای فشرده سازی و وافسرده سازی تصاویر دنباله ای^۱، استاندارد H.261 (که با عنوان $64 \times P$ نیز اشاره می شود) از CCITT است. دو استاندارد دیگر، که I و MPEG II در حال تهیه هستند.^۲

استاندارد H.261 برای کاربردهای گرد همایی راه دور تصویری^۳ می باشد، که در این کاربردها باید تصویر کاملاً متحرک روی خطوط T1^۴ با تأخیرهای انتقال کمتر از ۱۵۰ میلی ثانیه ارسال شود. (تأخيرهایی که متجاوز از ۱۵۰ میلی ثانیه باشند، احساس مشاهده مستقیم را برای بینندگان فراهم نمی آورند.) بر عکس، استاندارد I MPEG پیشنهاد شده یک استاندارد فشرده سازی ویدیو "با کیفیت تفریحی" برای ذخیره سازی و بازیابی تصاویر فشرده روی رسانه های رقیع نظری CD-ROM^۵ می باشد. گرچه نرخهای انتقال CD-ROM های امروزی در حدود نرخ خطوط انتقال T1 هستند، استاندارد

^۱ این مطلب به زمان تالیف ویرایش دوم نسخه انگلیسی این کتاب برمی گردد.

^۲ خط T1 به وسیله Bell System برای مخابره صوت رقیع در فواصل کوتاه از ۱۰ تا ۵۰ مایل معرفی شد. ۲۴ خط تلفن معمولی نمونه برداری و رمز PCM می شوند و میس روی یک سیگنال ۱۵۴۴ مگابیت بر ثانیه PCM مالتی پلکس زمانی می شوند تاروی خط T1 انتقال یابند.

MPEG I نرخهای بیت و کیفیت های رمزگذاری بالاتر رانیز پوشش می دهد. در هر حال این استاندارد روال رمزگاری خاصی را مشخص نمی کند؛ بلکه فقط یک رشته بیت رمز شده استاندارد و رمزبردار متناظر آن را تعریف می کند. استاندارد MPEG II از نرخهای انتقال ویدیویی بین ۱۰ تا ۱۵۰ مگابیت بر ثانیه، که برای پخش تلویزیون کابلی و پخش وسیع ماهواره ای با تند باریک مناسب هستند، پشتیبانی می کند.

هردو استاندارد MPEG و H.261 روش فشرده سازی مبتنی بر DCT توصیف شده در بخش قبلی را توسعه می دهند تا بدین وسیله کاهش افزونگی های بین قابی^۱ رانیز شامل شود. این استانداردها ابتدا یک قاب شروع یا مرجع را با استفاده از یک روش شبه JPEG که مبتنی بر DCT باشد، فشرده می سازند، سپس قاب فشرده شده را وافشده می کنند تا قاب اولیه بازسازی شود، حرکت اشیا بین قاب بازسازی شده و قاب بعدی را تخمین زده و براساس مقدار حرکت تصمیم می گیرند که آیا قاب بعدی را مستقلأ یا با استفاده از قاب قبل از فشرده شده نمایند. مرحله تخمین حرکت نوعاً شامل لغزاندن هر زیر تصویر بازسازی شده حول زیر تصویرهای مجاورش در قاب بعدی و محاسبه یک معیار همبستگی (نظری جمع مریعات تفاصلهای پیکسل به پیکسل) می باشد. در واقع اغلب این فرآیند با استفاده از نموهای جزء پیکسلی^۲ (نظری لغزاندن زیر تصویر به اندازه ۱/۴ پیکسل در هر بار) انجام می شود، که این عمل درونیابی مقادیر نقاط بین پیکسل ها قبل از محاسبه معیار همبستگی را الزامی می کند. از نظر محاسباتی ممکن است این فرآیند پرهزینه باشد. چون استاندارد MPEG برای کاربردهایی است که در آنها ممکن است تغییرات سریع بسیاری در صحنه رخ دهد، در این استاندارد لازم است هر پانزده قاب یکی بدون رجوع به قاب قبلش، یک رمزی شبه JPEG شود. این ویژگی برای ویرایش ویدیو نیز مفید است. تفاوت اصلی بین MPEG و H.261 در روش تخمین حرکت می باشد. در استاندارد H.261 هر قاب با قاب قبلش مقایسه شود، اما استاندارد MPEG پیشنهاد شده تعداد قابهای قبلی را که می توان در فرآیند تخمین حرکت از آنها استفاده نمود، تعریف نمی کند.

۷. ملاحظات پایانی

اهداف اصلی این فصل ارائه مبنای نظری فشرده سازی تصویر رقمه و توصیف متدائل ترین روش های

1) Frame-to-frame redundancy

2) Subpixel increment

۳۸۷ مراجع

فسرده سازی که هسته فناوری کنونی را تشکیل می دهد، بود. گرچه سطح ارائه مطلب مقدماتی است، عمق و وسعت مطالب بیان شده به عنوان مبنای برای مطالعه مستقل در این موضوع کفایت می کند. مراجع این فصل راه رسیدن به حجم گسترده ای از نوشه ها را که درباره 'فسرده سازی تصویر و مباحث مربوطه هستند، فراهم می آورند. علاوه بر کاربردهای گسترده ای که تصویر خاکستری را به کار می بردند، همان طور که با ظهور استانداردهای جهانی مشروح در بخش ۶.۶ روشن شد، روش های فشرده سازی نقش مهم و رو به رشدی در ذخیره و ارسال تصاویر استانداریز دارند. علاوه بر تصویربرداری پزشکی، فشرده سازی یکی از چند مبحث اندک از مباحث پردازش تصویر است که به اندازه 'کافی تقاضا دارد طوری که پذیرش وسیع استانداردهای آن را تضمین می کند.

مراجع

مباحث مقدماتی این فصل که عموماً به بخش های ۱.۶ و ۲.۶ محدود می شوند، از مباحث پایه ای فشرده سازی تصویر است و می توان آن را در اغلب کتب عمومی پردازش تصویر که در انتهای فصل ۱ ذکر شدند، به همین شکل با به شکلی دیگر پیدا نمود. موضوع بخش ۱.۶ درباره چندی سازی بهبود یافته سطح خاکستری براساس Whelan[1966] و Bisignani, Richards[1966] است. برای اطلاعات بیشتر در مورد سامانه بینایی انسان Limb[1981] و Netraveli [1981]، همچنین [1966] Huang، Schreiber و [1958] Knapp، و مراجع مذکور در انتهای فصل ۲ را بینید. رمزهای کشف و تصحیح خطای بیشتر متن مقدماتی در مورد کلیدزنی یا نظریه خودکاره های محدود، همچنین متن عمومی درباره 'نظریه اطلاعات بیان شده اند.

مطلوب بخش ۳.۳ براساس چند کتاب مشهور از نظریه اطلاعات است. مراجع قابل ذکر بیشتر توضیحات روش های رمزگذاری بدون خطای بخش ۴.۶، براساس مقالات اولیه مذکور در متن یا براساس مقالات زیر هستند. مقاله کلامیک شانون با عنوان "نظریه ریاضی مخابرات" [1948] شالوده بیشتر مطالب این بخش بوده و یک مرجع عالی دیگر است.

بیشتر توضیحات روش های رمزگذاری بدون خطای بخش ۴.۶، براساس مقالات اولیه مذکور در متن یا براساس مقالات زیر هستند. الگوریتم های بیان شده نماینده ای از کار در این زمینه هستند، اما به هیچ وجه کامل نمی باشند. از جمله روش های مرتبط که بالقوه مورد علاقه هستند، کار Ziv و Lempel [1977, 1978] و Welch[1984] می باشد، که این روش ها از این نظر که رمز، همزمان با

رمزنگاری داده ها ساخته می شود، یکتا هستند. متن بیان شده درباره رمزنگاری حسابی از روش Witten و Neal [1978] پیروی می کند. یکی از پیاده سازی های بهتر رمزنگاری حسابی را می توان در کار Penne baker و همکارانش [1988] یافت. برای اطلاعات بیشتر درباره رمزنگاری بیت صفحه ای، Schwartz و [1966] Barker ومقاله آموزشی از Rabbani و [1991] Jones را که حاوی شرح خوبی در مورد رمزنگاری پیشگویی بدون اتلاف نیز می باشد، بینید. اولین بار Huang و [1975] Hussian [1972] روش پرس از بلاک سفید را به تفصیل بیان کردند. رمزنگاری آدرس نسبی و چندی سازی تفاضلی پیشگویی اولین بار به ترتیب توسط Wakahara و Yamazaki [1976] و Teramura [1972] و Huang [1972] مطرح شدند. پیشگویی تطبیقی معادله (۶.۵-۱۹) از [1958] Graham است.

متن بخش ۶.۵ دو روش اصلی رمزگذاری با اتلاف را شرح می دهد. سایر روش های گوناگون مستقیماً براساس این دوروش هستند. از میان این روش ها، روش های مهمتر عبارتند از: رمزگذاری مختلط^۱ (Habibi [1974]) که رمزنگاری تبدیلی یک بعدی و DPCM را ترکیب می کنند تا انجام محاسبات کمتر تقریباً به همان کارآئی رمزنگاری تبدیلی دو بعدی دست یابد؛ رمزنگاری باند جزئی^۲ (O'Neil Woods [1968]) که در آن یک تصویر به مجموعه ای از تصاویر (با محدوده های فرانس مکانی متفاوت) تجزیه می شود و می توان تک تک آن تصاویر را رمز DPCM نمود؛ و رمزنگاری بین قابی^۳ (Roeise [1977]) که در آن با استفاده از رمزنگاری تبدیلی یا پیشگویی افزونگی بین قاب های متوالی در یک دنباله زمانی از تصاویر کاهش می یابد. علاوه بر این، انواع گوناگونی از روش های با اتلاف ارتباط نزدیکی با دو روش توصیف شده دارند، که چند نوع از آنها عبارتند از: رمزنگاری قطع بلاک^۴ (Delp و Mitchell [1979]) که در آن برای هر بلاک $n \times n$ از تصویر تقسیم شده یک چندی ساز یک بیتی طراحی می شود؛ چندی سازی برداری^۵ (Linde و همکارانش [1980]) که در آن تصویر به بردارهای (حاوی پیکسل ها، ضرایب تبدیل وغیره) تجزیه می شود که این بردارها با یک کتاب رمز از بردارهای ممکن تطبیق داده و براساس بهترین تطبیق رمز می شوند؛ و رمزنگاری سلسله مراتبی^۶ (Knowlton [1980]) که معمولاً شامل تولید یک مجموعه داده با ساختار هرمی و با قابلیت دسترسی پیش رو به آن جهت دستیابی به نمایش های بهتر و بهتر از تصویر اصلی می باشد. این مراجع مختصر اولیه روش را ذکر نمی کنند؛ بلکه فقط نقطه

1) Hybrid encoding

2) Subband coding

3) Interframe coding

4) Block truncation coding

5) Vector quantization

6) Hierarchical coding

آغازی را برای مطالعه بیشتر درباره این روش ها مشخص می کنند. سایر مقالات مورد علاقه شامل Wintz [1971]، Gharavi [1988] و Candy [1990]، Lim [1990]، Baylon [1988] و Tabatabai [1988] و همکارانش Udpikar [1984]، Mitchell [1981] و Lema [1981]، Jain [1981] و Healy [1981] و Jain [1971] و Tanimoto [1979]، Sezan [1989]، Equitz [1989]، Gray [1984]، Raina [1987] و Fand [1989] و Rabbani [1989] و Blume [1989] و Jones [1991] می شوند.

بخش ۶ . ۶ براساس نسخه های اولیه و نهایی استانداردهای ISO و CCITT می باشد. این مستندات در سازمان های استاندارد مذکور یا مؤسسه استانداردهای ملی آمریکا¹ (ANSI) موجود هستند. چند مرجع دیگر در مورد استانداردهای فشرده سازی Hunter [1980] و Robinson [1980] و همکارانش Ang [1991] و Fox [1991] هستند.

چند مقاله مروری به مبحث فشرده سازی تصویر اختصاص یافته اند که از میان آنها Netravali [1981]، Jain [1981]، Limb [1980]، IEEE Transactions on Communications [1981]، Proceedings of DCC '91 [1985] و Proceedings of IEEE [1985] قابل ذکر هستند. مقالات در این مبحث هستند. (J. H. Reif, J. A. Storer) نماینده جدیدترین کارها در این مبحث هستند.

مسائل

۱۰.۶ (الف) آیا می توان از روش های رمزنگاری طول متغیر برای فشرده سازی یک تصویر با بافت نگار متعادل شده و 2^n سطح خاکستری استفاده کرد؟ شرح دهید.
 (ب) آیا ممکن است چنین تصویری افزونگی های بین پیکسلی داشته باشد که بتوان از آنها برای فشرده سازی تصویر بهره برد؟

۲۰.۶ گونه ای از روش های رمزنگاری RL توصیف شده در بخش ۶ . ۱ . ۲ شامل دو مرحله ذیل می باشد:

(۱) رمزنگاری دنباله های صفر یا یک (نه هر دو) و (۲) انتساب یک رمز ویژه به آغاز هر خط برای کاهش اثر خطاهای انتقال. یک زوج رمز ممکن (x_k, r_k) است، که x_k و r_k به ترتیب

مختصه آغاز و طول دنباله k ام را مشخص می‌کند. برای اعلان آغاز خط جدید از رمز $(0, 0)$ استفاده می‌شود.

الف) یک عبارت کلی برای بیشینه تعداد متوسط دنباله‌های مورد نیاز در هر خط پیمایش از تصویر دودویی $2^n \times 2^n$ که فشرده سازی داده‌ها با رمزگاری طول دنباله‌ای را تضمین کند، به دست آورید.

ب) حداقل بیشینه مجاز بازای $1024 = n$ را محاسبه نماید.

۳.۶ یک خط ۸ پیکسلی از داده‌های خاکستری $\{54, 57, 12, 13, 10, 13, 12, 13\}$ را که به طور یکنواخت با ۶ بیت چندی شده است، در نظر بگیرید. حال رمز IGS ۳ بیتی آن را ایجاد نماید.

۴.۶ خطای rms و نسبت‌های سیگنال به نویز rms برای داده‌های IGS رمزبرداری شده' مسئله ۶ را محاسبه کنید.

۵. الف) با استفاده از رمز همینگ $(4, 7)$ داده‌های IGS چندی شده جدول ۶. ۲. را رمز نماید.
ب) در پیام‌های رمز همینگ ۱۱۰۰۱۱۱، ۱۱۰۰۱۱۰ و ۱۱۰۰۰۱۰، بیت خطارا، اگر وجود دارد،

مشخص نماید؟ مقادیر رمزبرداری شده چیستند؟

۶.۶ واحد اطلاعات در مبنای e عموماً^۱ خوانده می‌شود، و واحد اطلاعات در مبنای ۱۰ هارتلی^۲ خوانده می‌شود. نسبت بین این واحدها و واحد اطلاعات در مبنای ۲ (بیت) را محاسبه نماید.

۷.۶ ثابت کنید که برای یک منبع بی حافظه با q نماد، مقدار بیشینه آنتروپی $\log q$ است که اگر و تنها اگر تمام نمادهای منبع هم احتمال باشند، بدست می‌آید. راهنمایی: کمیت $H(z) = \log q$

را در نظر داشته باشید و به نامساوی $\ln x \leq x - 1$ نیز توجه نمایید.

۸.۶ در یک کانال اطلاعات که در آن $\{0, 1\}$ ، $A = \{0, 1\}$ ، $B = \{0.75, 0.25\}^T$ و $z = [0.75, 0.25]^T$ ، و

$$\mathbf{Q} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{10} & \frac{9}{10} \end{bmatrix}$$

1) Nat

2) Hartley

۳۹۱ مسائل

احتمالات گوناگون مربوط به آن شامل $P(b = 1)$, $P(b = 0)$, $P(a = 1)$, $P(a = 0)$,
 $P(a = 0|b=0)$, $P(b=1|a=1)$, $P(b=1|a=0)$, $P(b = 0|a = 1)$, $P(b = 0|a = 0)$,
 $P(a = 0, b = 0)$, $P(a = 1|b=1)$, $P(a = 1|b = 0)$, $P(a = 0|b = 1)$
 $P(a = 1, b = 0)$, $P(a = 1, b = 1)$, $P(a = 0, b = 1)$ را محاسبه نماید.

منبع اطلاعات دودویی و کانال BSC مثال بخش ۶.۳.۲ را در نظر بگیرید و اجازه دهید که

$$p_e = \frac{1}{3} p_{ss} = \frac{1}{4}$$

(الف) آتروپی منبع چیست؟

(ب) وقتی خروجی مشاهده می شود، چقدر عدم قطعیت در مورد ورودی کمتر می شود؟

(پ) این تفاوت در عدم قطعیت چه خوانده می شود و از نظر عددی چه نسبتی با ظرفیت دارد؟

یک کانال پاک شدگی دودویی^۱ کانالی است که در آن با احتمال محدود β ممکن است یک نماد

ارسالی اصلاً دریافت نشود. کانال سه خروجی ممکن دارد: صفر، پاک شده (نماد دریافت نشده)، و یک. این سه خروجی سه سطر ماتریس کانال پاک شدگی دودویی را ایجاد می کنند.

$$\mathbf{Q} = \begin{bmatrix} 1-\beta & 0 \\ \beta & \beta \\ 0 & 1-\beta \end{bmatrix}$$

(الف) ظرفیت کانال را باید.

(ب) آیا شما یک کانال BSC با احتمال خطای ۰.۱۲۵ را ترجیح می دهید یا یک کانال پاک شدگی با

$$\text{احتمال پاک شدگی } 0.5 = \beta \text{ را؟}$$

تابع اعوجاج نرخ یک منبع گوسی بی حافظه با میانگین دلخواه و واریانس σ^2 ، بر حسب میار خطای میانگین مربع (Berger[1971]) عبارتست از

$$R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{D} & 0 \leq D \leq \sigma^2 \\ 0 & D \geq \sigma^2. \end{cases}$$

(الف) این تابع رارسم نماید.

1) Binary erasure channel

ب) D_{\max} چیست؟

پ) اگر اعوجاج‌های کمتر از ۷۵ درصد واریانس منبع، مجاز باشد، بیشینه فشرده‌سازی که می‌توان به آن دست یافت، چیست؟

۱۲.۶) چند نوع رمزهافمن برای یک منع سه نمادی وجود دارند؟
ب) آنها را ایجاد نماید.

(ب) یک رمزهای نمادهای منبع ایجاد نماید و هرگونه تفاوت بین این رمز و رمز ۲ از جدول را توضیح دهید.

ب) بهترین رمز B را برای این توزیع ایجاد کنید.

ت) بهترین رمز انتقالی دودویی ۲ بیتی را ایجاد کنید.

ت) نمادها را به دو بلوک چهارتایی تقسیم کنید و بهترین رمز انتقالی هافمن را ایجاد نماید.
ج) طول متوسط کلمات هریک از رمزهارا محاسبه کنید و آنها را با آنتروپی قسمت (الف) مقایسه نمایند.

۱۴.۶ فرآیند رمزبرداری حسابی عکس روای رمزگذاری است. با داشتن مدل رمزنگاری ذیل، پیام را رمزبرداری کنید.

نمار	احتمال
<i>a</i>	0.2
<i>e</i>	0.3
<i>i</i>	0.1
<i>o</i>	0.2
<i>u</i>	0.1
!	0.1

۶.۱۵(الف) رمزگری کامل ۴ بیتی را ایجاد نماید.

ب) یک روال کلی برای تبدیل یک عدد رمزگری شده به معادل دودویی اش ایجاد کنید و با استفاده از آن 011101010011 را رمزگیر داری نماید.

۱۶.۶ یک تصویر دودویی 64×64 پیکسلی با استفاده از WBS یک بعدی با بلوک های چهار پیکسلی رمز شده است. رمز WBS برای یک خط تصویر 0110010000001000010010000000 است که هر 0 برای نمایش یک پیکسل سیاه استفاده می شود.

الف) این خط را رمزبرداری نماید.

ب) یک روای WBS تکراری یک بعدی ایجاد کنید که با جستجوی خطوط تمام سفید (یک بلوک ۶۴ پیکسلی) آغاز می شود و به طور متواالی بازه های غیرسفید را نصف می کند تا اینکه به بلوک های چهار پیکسلی بررسد.

پ) با استفاده از الگوریتم ایجاد شده، خطی را که قبل از رمزبرداری شده است، مجدد رمزگذاری کنید. این الگوریتم باید به بیت های رمز کمتری نیاز داشته باشد.

۱۷.۶ الف) توضیح دهد که چرا در رمزگذاری آدرس نسبی، اولین گذر مشابه بعد از π روی خط قبلی به عنوان c استفاده می شود.

ب) آیا می توانید راه دیگری را طرح نماید.

۱۸.۶ یک تصویر که تابع خود همبستگی آن به شکل معادله $(6.5-12)$ با $m = 0$ است، باید با استفاده از پیشگوی مرتبه دو رمز DPCM شود.

الف) ماتریس خود همبستگی R و بردار c را ایجاد نماید.

ب) ضرایب پیشگویی بهینه را باید.

پ) واریانس خطای پیشگویی را که با استفاده از ضرایب بهینه ایجاد می شود، محاسبه نماید.

۱۹.۶ سطوح بازسازی و تصمیم لوید-ماکس را برای $L = 4$ و تابع چگالی احتمال یکتواخت زیر بدست آورید:

$$p(s) = \begin{cases} \frac{1}{2A} & -A \leq s \leq A \\ 0 & \text{در غیر این صورت} \end{cases}$$

۲۰.۶ با الگوریتم فشرده سازی گروه ۴ CCITT^۱ خط دوم از قطعه 'دوخطی زیر را رمزگذاری نماید.

۰ ۱ ۱ ۰ ۰ ۱ ۱ ۱ ۰ ۰ ۱ ۱ ۱ ۱ ۱ ۱ ۰ ۰ ۰ ۰ ۱
۱ ۱ ۱ ۱ ۱ ۱ ۰ ۰ ۰ ۱ ۱ ۱ ۰ ۰ ۰ ۰ ۱ ۱ ۱ ۱ ۱

فرض کنید که عنصر مرجع اولیه a_0 روی اولین پیکسل از قطعه خط دوم قرار دارد.

الف) تمام اعضای دسته ۳ از ضریب تفاضلی DC در JPEG را فهرست نماید.

ب) با استفاده از جدول ۶.۱۵ رمزهای هافمن پیش گزیده آنها را محاسبه کنید.

۲۲.۶

یک پرتوشناس^۱ از یک بیمارستان تحقیقاتی مشهور اخیراً در یک کنفرانس پژوهشی شرکت کرد که در آنجا سامانه‌ای برای ارسال تصاویر پرتوایکس ۴۰۹۶ × ۴۰۹۶ رقمی شده ۱۲ بیتی بر روی خطوط تلفن T1 استاندارد به نمایش گذاشته شد. آن سامانه با استفاده از روشی پیشرو تصاویر را به شکل فشرده ارسال می‌کرد که در آن روش ابتدا تقریب خوبی از تصویر پرتوایکس در مقصد بدست می‌آمد و سپس تصویر مذکور بتدريج اصلاح می‌شد تا تصویر بدون خطایی به دست آید. ارسال داده‌های موردنیاز برای تولید اولین تقریب تصویر تقریباً ۵ یا ۶ ثانیه طول می‌کشید. اصلاح‌ها در هر ۵ یا ۶ ثانیه (به طور متوسط) برای یک دقیقه بعدی انجام می‌شدند که اولین و آخرین اصلاح به ترتیب بیشترین و کمترین اثر را بر تصویر پرتوایکس بازسازی شده داشتند. پژوهش پرتوشناس بسیار تحت تأثیر این سامانه قرار گرفت، زیرا می‌توانست با استفاده از تقریب اول تصویر پرتوایکس تشخیص بیماری را شروع کند و وقتی بازسازی بدون خطای تصویر پرتوایکس انجام شد، تشخیص بیماری را کامل نماید. وی پس از بازگشت به دفترش، درخواست خرید آن دستگاه را به سرپرست بیمارستان تحويل داد. متأسفانه بیمارستان بودجه نسبتاً محدودی داشت، که این بودجه اخیراً با استخدام یک مهندس برق جوان محدودتر نیز شده بود. سرپرست بیمارستان برای آرام نمودن پرتوشناس، وظیفه طراحی چنان سامانه‌ای را به مهندس جوان داد. (او فکر کرد که ممکن است طراحی و ساخت سامانه مشابه ارزانتر تمام شود. بیمارستان تعدادی از اجزای چنان سامانه‌ای را دارد بود، اما ارسال داده‌های پرتوایکس خام بیش از ۲ دقیقه طول می‌کشید.) سرپرست بیمارستان از مهندس خواست که تا زمان برگزاری جلسه کارکنان در بعدازظهر، یک نمودار بلوکی ابتدایی از سامانه پیشنهادی تهیه نماید. با این فرصت کم و در دست داشتن تنها یک نسخه از کتاب پردازش تصویر رقمه^۱ (البته همین نسخه) مهندس قادر خواهد بود که از نظر مفهومی سامانه‌ای را طرح کند که نیازهای ارسال و فشرده سازی مربوطه را بآورده نماید. یک نمودار بلوکی مفهومی از چنین سامانه‌ای رارسم نماید و روش‌های فشرده سازی را که توصیه می‌کنید، مشخص نماید.

فصل ۷

بخش بندی تصویر

تمام یک چیز برابر جمع قسمت های آن است.

اقلیدس

تمام یک چیز بزرگتر از جمع قسمت های آن است.

ماکس ورتیمر

در فصول ۱-۳ کتاب، مباحث زمینه ای را مطرح کردیم و در فصول ۴-۶ روش های پیش پردازش را به دقت بیان کردیم. سایر فصول این کتاب به فنون استخراج اطلاعات از تصویر می پردازد. این گروه از پردازش ها را با عنوان تحلیل تصویر^۱ اشاره می کنیم.

عمولاً مرحله اول در تحلیل تصویر، بخش بندی است. با عمل بخش بندی، تصویر به قسمت های تشکیل دهنده اش تقسیم می شود. میزان بخش بندی، به موضوع مورد نظر بستگی دارد. یعنی وقتی اشیای مورد علاقه کاربرد مورد نظر از هم جدا شدند، باید بخش بندی متوقف شود. در کاربردهای رهگیری خودکار هدف زمینی از هوا، تنها شناسایی وسایل نقلیه روی جاده مورد علاقه است. در مرحله اول، جاده از تصویر جدا می شود و سپس محتویات جاده به اجزایی تقریباً به بزرگی هدف های مورد علاقه بخش می شود. هیچ فایده ای در بخش بندی کوچکتر از این مقیاس وجود ندارد و همچنین نیازی نیست که برای بخش بندی اجزایی از تصویر که خارج از مرزهای جاده قرار دارند، تلاش کنیم.

به طور کلی بخش بندی یکی از مشکل ترین کارها در پردازش تصویر است. این مرحله است که موقفيت یا شکست احتمالی تحلیل تصویر را تعیین می کند. در واقع با بخش بندی مؤثر در بیشتر موارد، به یک جواب قابل قبول می رسیم. بنابراین باید توجه زیادی معطوف افزایش احتمال دست یابی به بخش بندی قابل اعتماد بشود. در بعضی وضعیت ها نظیر کاربردهای بازرسی صنعتی، دست کم، گاهی اوقات می توان تا حدی برمیط کنترل داشت. طراح ماهر سامانه پردازش تصویر همواره توجه زیادی به چنین

1) *Image analysis*

فرصت هایی دارد. در سایر کاربردها، نظیر رهگیری هدف، طراح سامانه هیچ کنترلی بر محیط ندارد. در این شرایط روش معمول، تمرکز بر روی انتخاب انواعی از حسگرهاست که برای ارتقای اثیابی مورد علاقه در تصویر مناسب ترین هستند و نیز میزان اجزای زاید تصویر را کاهش می دهند. یک مثال خوب، استفاده از تصویر برداری مادون قرمز برای آشکارسازی اثیابی دارای تشعشع شدید گرمایی، نظیر تانکهای متحرک، است.

الگوریتم های بخش بندی تصاویر تک رنگ، معمولاً مبتنی بر یکی از دو خاصیت اصلی سطوح خاکستری یعنی ناپیوستگی یا شباهت هستند. در دسته اول، تصویر براساس تغییرات سریع سطح خاکستری بخش بندی می شود. موضوعات اصلی مورد علاقه در این دسته آشکارسازی نقاط منفرد و آشکارسازی خطوط و لبه های تصویر است. روش های اصلی در دسته دوم مبتنی بر آستانه گیری، رشد ناحیه، و تقسیم و ادغام ناحیه هستند. مفهوم بخش بندی تصویر براساس ناپیوستگی یا شباهت مقادیر سطوح خاکستری را می توان برای هر دونوع تصویر ثابت و متحرک (متغیر با زمان) به کار برد. با این حال در مورد تصویر متحرک می توان از حرکت به عنوان عاملی قوی در بهبود کارآئی الگوریتم های بخش بندی استفاده کرد.

۱.۷ آشکارسازی ناپیوستگی ها

در این بخش چند روش برای آشکارسازی سه نوع ناپیوستگی اصلی در تصویر رقمه یعنی نقاط منفرد، خطوط و لبه ها ارائه می دهیم. در عمل، عمومی ترین راه برای جستجوی ناپیوستگی ها پیمایش کامل تصویر با یک نقاب به روش مشروح در بخش ۴.۱ است. برای نقاب 3×3 شکل ۱.۷، این روال شامل محاسبه مجموع حاصل ضرب های ضرایب نقاب در سطوح خاکستری پیکسلهای ناحیه زیرش است.

یعنی پاسخ نقاب در هر نقطه 'تصویر عبارتست از

$$R = w_1 z_1 + w_2 z_2 + \dots + w_9 z_9 = \sum_{i=1}^9 w_i z_i \quad (1-1.7)$$

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

شکل ۱.۷ یک نقاب عمومی 3×3

۱۰.۷ آشکارسازی ناپیوستگی ها ۴۹۷

-1	-1	-1
-1	8	-1
-1	-1	-1

شکل ۲۰.۷ یک نقاب که برای آشکارسازی نقاط منفرد متفاوت با یک زمینه ثابت استفاده شد.

که، سطح خاکستری پیکسل زیر ضرب نسبت w است. طبق معمول پاسخ نقاب در پیکسل مرکزش تعریف می شود. وقتی مرکز نقاب روی یک پیکسل مرزی باشد، پاسخ آن با استفاده از یک همسایگی جزئی مناسب محاسبه می شود.

۱۰.۷ آشکارسازی نقطه

آشکارسازی نقاط منفرد تصویر ساده است. با استفاده از نقاب شکل ۲۰.۷، می گوییم که یک نقطه در مرکز نقاب آشکارشده است اگر

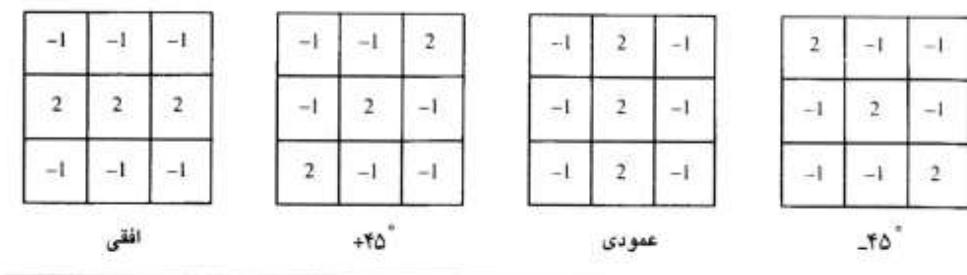
$$|R| > T \quad (20.7)$$

که T یک آستانه غیر منفی است، و R با معادله (۱۰.۱) تعیین می شود. تنها کار این رابطه اندازه گیری تفاضل های وزن داده شده بین نقطه مرکز نقاب و همسایگانش است. ایده پشتیبان این روش این است که سطح خاکستری یک نقطه منفرد کاملاً متفاوت از سطوح خاکستری همسایه هایه اش است.

نقاب شکل ۲۰.۷ همان نقاب مورد استفاده برای فیلتر کردن مکانی فرکانس بالا است (شکل ۲۴.۴ را ببینید). با این حال در اینجا فقط آشکارسازی نقاط مورد تأکید است. یعنی تنها تفاضل های به اندازه کافی بزرگ (که با T تعیین می شوند)، مورد علاقه هستند.

۲۰.۷ آشکارسازی خط

در سطح بعدی، آشکارسازی خطوط موجود در تصویر مطرح می شود. نقاب های شکل ۲۰.۳ را در نظر بگیرید. اگر اولین نقاب در تصویر حرکت نماید، به خطوط افقی (با ضخامت یک پیکسل) پاسخ



شکل ۳.۷ نقاط های خط

بسیار بزرگتری خواهد داد. با یک زمینه ثابت، وقتی پاسخ بیشینه نتیجه می شود که خط افقی از سطر وسط نقاط عبور کند. این موضوع را می توان به آسانی با رسم آرایه ساده ای از "۱" ها که دارای یک خط افقی با سطح خاکستری متفاوت (مثلث "۵" ها) باشد، بررسی کرد. آزمایش های مشابه نشان می دهد که نقاط دوم از شکل ۳.۷ به خطوط در جهت 45° ؛ نقاط سوم به خطوط عمودی؛ و نقاط چهارم به خطوط در جهت -45° بهترین پاسخ ها را می دهند. البته با توجه به این که جهت تعیین شده هر نقاط با ضرب ب بزرگتری (یعنی ۲) نسبت به سایر جهات ممکن وزن داده شده است، نیز می توان این جهت ها را تشخیص داد.

فرض کنید R_1, R_2, R_3 و R_4 بیانگر پاسخ نقاط های شکل ۳.۷ به ترتیب از چپ به راست باشد که با معادله $(1-1)$ تعیین می شوند و فرض کنید که هر یک از نقاط در تمام تصویر حرکت داده شود. اگر در یک نقطه 'معین تصویر'، برای تمام زهای که $i \neq j$ ، $|R_i| > |R_j|$ باشد، گفته می شود که به احتمال بیشتر آن نقطه مربوط به خطی در جهت نقاط است، مثلاً اگر در یک نقطه تصویر بازی $|R_1| > |R_j|$ باشد، گفته می شود که به احتمال بیشتر آن نقطه مربوط به یک خط افقی است.

۳.۱.۷ آشکارسازی لبه

گرچه به یقین در هر شرحی از بخش بندی، آشکارسازی نقطه و خط مطرح می شود، تاکنون آشکارسازی لبه متدالوی ترین راه برای آشکارسازی ناپیوستگی های معنادار در سطح خاکستری تصویر است. دلیل آن این است که در بیشتر تصاویر عملی نقاط منفرد و خطوط باریک به طور فراوان رخ نمی دهند.

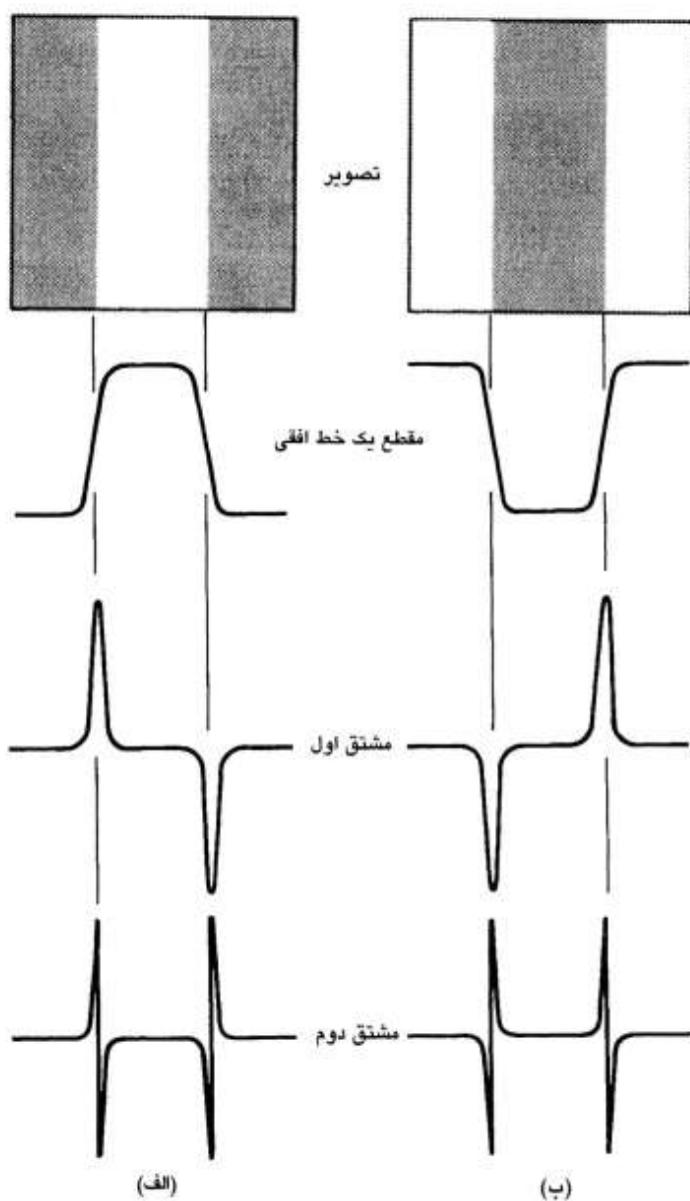
۱۰. آشکارسازی ناپیوستگی ها ۴۹۹

توصیف پایه ای

لبه مرز بین نواحی با خواص نسبتاً متفاوت سطح خاکستری است. در شرح زیر فرض می شود که نواحی مورد نظر به اندازه^۱ کافی همگن هستند که می توان تنها براساس ناپیوستگی های سطح خاکستری گذراز یک ناحیه به ناحیه دیگر را معین کرد. اگر این فرض درست نباشد، معمولاً روش های بخش بندی مشروح در بخش ۷. ۳. یا ۷. ۴ کاربردی تر از آشکارسازی لبه هستند.

نظریه^۲ پایه در بیشتر روش های آشکارسازی لبه، محاسبه یک عملگر مشتق محلی است. شکل ۷. ۴. این مفهوم را روشن می سازد. شکل ۷. ۴ (الف) تصویری از یک نوار روشن روی زمینه^۳ تیره، مقطع سطح خاکستری در طول یک خط افقی آن، و نیز مشتقات اول و دوم آن مقطع رانمایش می دهد. در این مقطع توجه شود که لبه (گذراز تاریک به روشن) به صورت یک تغییرآرام، نه سریع، سطح خاکستری مدل می شود. این مدل نشان می دهد که معمولاً لبه های تصاویر رقمی بر اثر نمونه برداری، کمی مات می شوند.

شکل ۷. ۴ (الف) نشان می دهد که مشتق اول مقطع سطح خاکستری در لبه جلویی گذرا^۱، مثبت است، در لبه عقبی آن منفی است و همان طور که مورد انتظار است، در نواحی با سطح خاکستری ثابت صفر است. مشتق دوم برای قسمتی از گذراز در طرف تیره لبه است، مثبت است، برای قسمت دیگر گذرا که در طرف روشن لبه است، منفی است، و در نواحی با سطح خاکستری ثابت، صفر است. بنابراین، از بزرگی مشتق اول می توان برای تعیین این که آیا پیکسل در روی لبه قرار دارد یا ندارد، و از علامت مشتق دوم برای تعیین این که آیا پیکسل لبه مورد نظر روی طرف تیره یا روشن لبه قرار دارد، استفاده کرد. توجه کنید که مشتق دوم در نقطه وسطی هر گذرا مقطع خاکستری یک عبور از صفر دارد. همان طور که بعداً در این بخش نشان داده می شود، عبور از صفرهاراهی قوی برای تعیین محل لبه های تصویر فراهم می آورند. گرچه تاکنون این شرح به یک برش افقی یک بعدی محدود بوده است، توضیحات مشابهی برای یک لبه در تصویر با هر جهت دلخواه نیز صدق می کند. ما در هر نقطه مطلوب تنها یک مقطع عمود بر جهت لبه تعریف کرده و نتایج را مانند شرح قبلی تعبیر می کنیم. اندازه^۱ مشتق اول تصویر در هر نقطه برابر بزرگی گرادیان است. مشتق دوم نیز با استفاده از لاپلاسین بدست می آید.



شکل ۴.۷ آشکارسازی لبه با عملگرهای مشتق: (الف) نوار روشن روی زمینه تیره؛ (ب) نوار تیره روی زمینه روشن. توجه کنید که مشتق دوم در محل هر لبه یک گذر از صفر دارد.

۵۰۱ ۱۰.۷ آشکارسازی ناپیوستگی ها

عملگرهای گرادیان

در بخش ۴.۳.۳ به اختصار مفهوم استفاده از گرادیان برای مشتق گیری تصویر را معرفی کردیم. طبق

- معادله (۴.۳-۵) گرادیان تصویر (x, y) در نقطه (x, y) برابر بردار زیر است:

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (4-1.7)$$

از مبحث تجزیه و تحلیل بردارها می دانیم که بردار گرادیان در نقطه (y, x) درجهٔ است که آهنگ تغییر ری بیشینه باشد. یک کمیت مهم در آشکارسازی لبه بزرگی این بردار است که معمولاً با عنوان گرادیان^۱ اشاره می شود و با ∇f نشان داده می شود:

$$\nabla f = \text{mag}(\nabla f) = [G_x^2 + G_y^2]^{1/2}. \quad (4-1.7)$$

این کمیت برابر آهنگ بیشینه افزایش $f(x, y)$ بر واحد طول در جهت ∇f است. مرسوم است که با استفاده از قدر مطلق، گرادیان را به صورت

$$\nabla f \approx |G_x| + |G_y| \quad (5-1.7)$$

تقریب می زند که پیاده سازی بسیار ساده تری بروزه با ساخت افزار خاص دارد.
جهت^۲ بردار گرادیان نیز یک کمیت مهم است. فرض کنید که (y, x) α بیانگر زاویه 'جهت بردار ∇f ' در نقطه (x, y) است. آنگاه از مبحث تجزیه و تحلیل بردارها داریم:

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (6-1.7)$$

که این زاویه نسبت به محور x سنجیده می شود.

طبق معادلات (۶-۱) و (۶-۲)، محاسبه 'گرادیان یک تصویر مبتنی بر بدست آوردن مشتقات جزئی $\frac{\partial f}{\partial x}$ و $\frac{\partial f}{\partial y}$ در هر پیکسل است. همان طور که در بخش ۴.۳.۳ تشریح شد (شکل ۴.۲۸)، به چند طریق می توان مشتقات را به شکل رقمه‌ی پیاده سازی کرد. در هر حال، عملگرهای سویل بینید)،

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

(الف)

-1	-2	-1
0	0	0
1	2	1

(ب)

-1	0	1
-2	0	2
-1	0	1

(پ)

شکل ۷.۵ (الف) یک ناحیه تصویری 3×3 ؛ (ب) نقاب استفاده شده برای محاسبه G_x در نقطه 'مرکز ناحیه' 3×3 ؛ (پ) نقاب استفاده شده برای محاسبه G_y در همان نقطه. اغلب به این نقاب ها با عنوان عملگرهای سوبیل اشاره می شود

این مزیت را دارند که هر دو اثر مشتق گیری و آرام سازی را با هم دارند (مسئله ۷.۳ را ببینید). چون مشتق گیری تویز را افزایش می دهد، اثر آرام سازی عملگرهای سوبیل ویژگی خیلی جذابی به شمار می آید. با توجه به شکل ۷.۵، مشتقات مبتنی بر نقاب های سوبیل عبارتند از:

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad (7-1.7)$$

و

$$G_y = (z_1 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7) \quad (8-1.7)$$

که مانند گذشته z ها سطوح خاکستری پیکسل های زیر نقاب ها در هر نقطه تصویر هستند. آنگاه مقدار گرادیان در مرکز نقاب با استفاده از معادلات (۷.۱-۴) یا (۷.۱-۵) محاسبه می شود. برای به دست آوردن مقدار بعدی گرادیان، این نقاب ها به محل پیکسل بعدی جابجا می شوند و روال قبل تکرار می شود. پس از اجرای روال فوق برای تمام نقاط تصویر، یک تصویر گرادیان با همان ابعاد تصویر اولیه بدست می آید. طبق معمول عملیات پنجه را روی مرز تصویر با استفاده از همسایگی های جزئی مناسب اجرا می شود.

۱. ۷ آشکارسازی ناپیوستگی ها ۵۰۳

مثال: شکل ۷. ۶ (الف) یک تصویر و شکل ۷. ۶ (ب) حاصل محاسبه $|G|$ آن با نقاط شکل ۷. ۵ (ب) را نشان می دهد. به یاد آورید که محور x را درجهت عمودی تعریف کردیم (شکل ۱. ۵ را بینید). بنابراین انتظار می رود که بزرگترین پاسخ های تولید شده با $|G|$ روی لبه های عمود بر محور x قرار داشته باشد. این نتیجه در شکل ۷. ۶ (ب) که پاسخ های بزرگی را در طول لبه های افقی نظری ساحل رودخانه نمایش می دهد، آشکار است. به کوچکی نسبی پاسخ در طول لبه های عمودی نیز توجه کنید. همان طور که شکل ۷. ۶ (پ) نشان می دهد، در $|G|$ حالت عکس رخ می دهد. با ترکیب این دو نتیجه از طریق معادله (۷. ۵-۱) تصویر گرادیان شکل ۷. ۶ (ت) بدست آمد.

□

لابلسین

لابلسین تابع دوبعدی $y = f(x)$ یک مشتق مرتبه دوم است که به صورت

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}. \quad (9-1. 7)$$

تعریف می شود. مانند گرادیان، معادله (۹-۱) را نیز در شکل رقمی به راه های گوناگونی می توان پیاده سازی کرد. برای یک ناحیه 3×3 شکلی که در عمل بیشتر با آن مواجه می شویم،

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8) \quad (10-1. 7)$$

است که جها قبلاً تعریف شدند. در تعریف لابلسین رقمی باید ضریب پیکسل مرکزی مثبت و ضرایب پیکسل های اطراف منفی باشد (شکل ۷. ۸. الف را بینید). چون لابلسین یک مشتق است، جمع ضرایب پنجره' آن باید صفر باشد. بنابراین هرگاه که نقطه مورد نظر و همسایگانش مقادیر پیکسانی داشته باشند، پاسخ آن صفر است. شکل ۷. ۷. یک نقاط مکانی را نشان می دهد که می توان از آن برای پیاده سازی معادله (۱۰-۱. ۱) استفاده کرد.

همان طور که قبلاً بیان شد، گرچه لابلسین به گذرهای شدت پاسخ می دهد، به چند دلیل به ندرت در عمل از آن برای آشکارسازی لبه استفاده می شود. چون لابلسین یک مشتق مرتبه دوم است، معمولاً حساسیت زیادی به نویز دارد. گذشته از این، لابلسین لبه های مضاعف تولید می کند (شکل ۴. ۷ را بینید) و قادر نیست که جهت لبه را آشکار نماید. به این دلایل لابلسین معمولاً وظیفه دوم آشکارساز را



(الف)



(ب)



(پ)



(ت)

شکل ۶.۷ (الف) یک تصویر اولیه؛ (ب) حاصل کاربرد نقاب شکل ۷.۵ (ب) برای برای به دست آوردن G_x ؛ (پ) حاصل کاربرد نقاب شکل ۷.۵ (پ) برای به دست آوردن G_y ؛ (ت) تصویر گرادیانی کامل^۱ حاصل از معادله (۶.۱-۷)

1) Complete gradient image

۱۰۵ ۱۰.۷ آشکارسازی ناپیوستگی ها

0	-1	0
-1	4	-1
0	-1	0

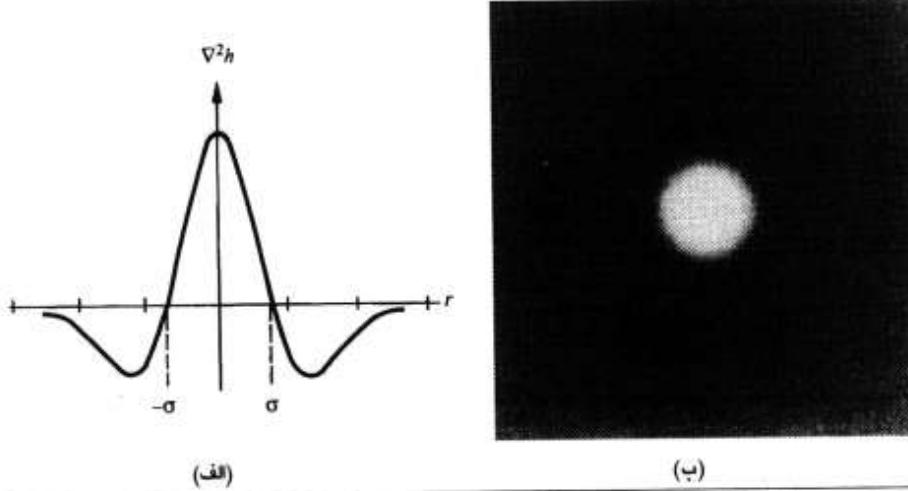
شکل ۷.۷ نقاب استفاده شده برای محاسبه لابلائین

بر عهده دارد و تعیین می کند که آیا یک پیکسل در طرف روشن لبه قرار دارد یا در طرف تاریک لبه؟ درباره سودمندی این خاصیت در بخش ۳.۵ توضیح می دهیم. یک کاربرد عمومی تر لابلائین، پیدا کردن محل^۱ لبه ها با استفاده از خاصیت عبور از صفر است (شکل ۷.۴ را ببینید). این مفهوم بر اساس کانولوشن تصویر و لابلائین تابع گوسی دو بعدی

$$h(x, y) = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (11-1.7)$$

می باشد که σ انحراف معیار است (Marr و Hildreth[1980]). فرض کنید که $y^2 = x^2 + r^2$. آنگاه از معادله (۹-۱.۷)، لابلائین h (یعنی مشتق دوم h نسبت به r) عبارتست از

$$\nabla^2 h = \left(\frac{r^2 - \sigma^2}{\sigma^4}\right) \exp\left(-\frac{r^2}{2\sigma^2}\right). \quad (12-1.7)$$

شکل ۷.۸ (الف) یک مقطع قطری از $\nabla^2 h$; (ب) $\nabla^2 h$ به صورت تابع شدت. (از Marr[1982])

1) Location

۵۰۶ پردازش تصویر رقمه

شکل ۷.۸. (الف) یک برش عرضی از این تابع متقارن چرخشی را نشان می دهد. به آرامی این تابع، نقاط عبور از صفر آن در $\sigma = \pm 2$ ، و نیز به مرکز مثبت و حوالی منفی آن توجه کنید. این شکل مدلی است که معادله ۷.۱۰) و نقاب شکل ۷.۷ برآن مبتنی هستند. شکل این معادله در پرسپکتیو سه بعدی که محور عمودی آن مربوط به شدت باشد، مانند شکل کلاه رسمی مکزیکی است. شکل ۷.۸. (ب) نیز تصویر این معادله را نمایش می دهد. می توان نشان داد (مسئله ۷.۶) که مقدار متوسط عملگر لابلسین h^2 برابر صفر است. این مطلب در مورد تصویر لابلسین حاصل از کانولوشن این عملگر با هر تصویر دلخواه نیز صحیح است.

شرح شکل ۴.۳۳. بیان می کند که کانولوشن تصویر با تابعی به شکل ۷.۸. (ب) تصویر رامات می کند که درجه مات شدگی با h^2 متناسب است. گرچه این خاصیت نیز را کاهش می دهد، مزیت اصلی معادله ۷.۱۲) در نقاط عبور از صفر می باشد.

مثال: شکل ۷.۹. (الف) را که تصویر ساده ای با تفکیک 320×320 پیکسل است، در نظر بگیرید. شکل ۷.۹. (ب) حاصل کانولوشن این تصویر با تابع h^2 را نمایش می دهد. مقدار h در این مورد برابر ۴ است. در این تصویر، سیاه بیانگر منفی ترین مقادیر و سفید بیانگر مثبت ترین مقادیر است؛ بنابراین سطوح وسط بیانگر صفر هستند. شکل ۷.۹. (پ) تصویر دودویی حاصل سیاه قرار دادن تمام مقادیر منفی شکل ۷.۹. (ب) و سفید قرار دادن تمام مقادیر مثبت آن است. در شکل ۷.۹. (ت) نیز گذر از صفرها به عنوان سرزهای بین سیاه و سفید شکل ۷.۹. (پ) تعیین شده اند. توجه کنید که محل تمام لبه های اصلی شکل ۷.۹. (الف) به دقت تعیین شده اند.

توضیح قبلی مفهوم دیگری را نیز در خود دارد: آشکارسازی لبه با عملگرهای گرادیانی در مواردی خوب کار می کند که گذرهای شدت موجود در تصویر تیز و نیز آن نیز نسبتاً کم باشد. در مواردی که لبه های مقدار نیز زیاد است، عبور از صفرها راه جایگزین هستند. عبور از صفرها با دقت بیشتری محل لبه ها را تعیین می کنند، و خواص آرام سازی h^2 نیز اثرات نیز را کاهش می دهد. بهای پرداخته شده برای این مزایا افزایش پیچیدگی و زمان محاسبات است.

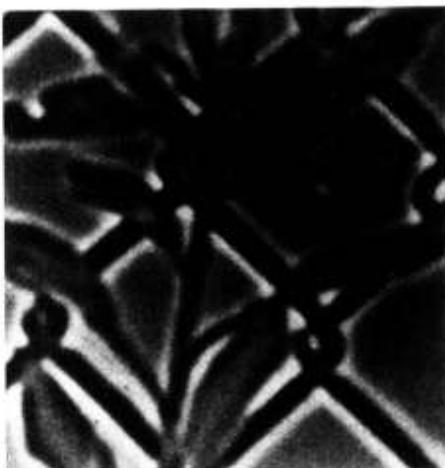
۴.۱.۷ آشکارسازی ترکیبی

با استفاده از توصیف چند نقاطی می توان روشی را برای تعیین این که آیا یک پیکسل به احتمال بیشتر

۱.۷ آشکارسازی ناپیوستگی ها ۵۰۷



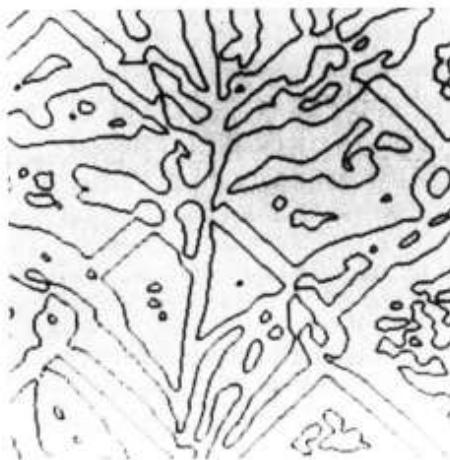
(الف)



(ب)



(ب)



(ج)

شکل ۱.۷ (الف) تصویر اولیه؛ (ب) حاصل کانولوشن (الف) با $\nabla^2 h$ ؛ (ج) حاصل دودویی سازی (ب) تا این که آشکارسازی گذر از صفرها آسان شود؛ (د) گذر از صفرها (از Marr [1982]).

یک نقطه، منفرد است، یا این که بخشی از یک خط یا لبه است، طراحی کرد. این طرح با استفاده از توصیف برداری بسیار شهودی تر می شود. فرض کنید

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_9 \end{bmatrix} \quad (13-1.7)$$

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \quad (14-1. v)$$

که همانند بخش ۱. ۳. ۷، \mathbf{w} ها ضرایب یک نقاط 3×3 (شکل ۱. ۷ را بینید)، و \mathbf{z} ها سطوح خاکستری متناظر شان هستند. با این نمادگذاری، پاسخ نقاط در معادله (۱-۱. ۷) همان ضرب داخلی بردارهای \mathbf{w} و \mathbf{z} است. یعنی،

$$R = \sum_{i=1}^9 w_i z_i = \mathbf{w}^T \mathbf{z} \quad (15-1. v)$$

که بالا نویس T بیانگر ترانهاده گیری^۱ برداری است.

اجازه دهید موقتاً فرض کنیم که نقاط \mathbf{z} به جای نه ضریب، سه ضریب دارند. در این حالت بردارها سه بعدی هستند و به آسانی می‌توان آنها را نمایش داد. همچنین فرض می‌کنیم که دو نقاط آشکارساز، مثلاً برای لبه‌ها و خط‌ها، داریم و ضرایب آنها چنان است که دو بردار متناظر شان، که با \mathbf{w}_1 و \mathbf{w}_2 نشان می‌دهیم، متعامد بوده و طول واحد دارند. در این مورد حاصل ضرب های $\mathbf{w}_1^T \mathbf{z}$ و $\mathbf{w}_2^T \mathbf{z}$ برابر تصویر \mathbf{z} به ترتیب روی بردارهای \mathbf{w}_1 و \mathbf{w}_2 هستند. زیرا برای \mathbf{w}_1 ،

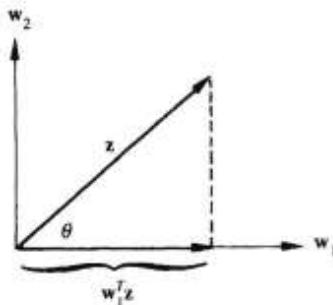
$$\mathbf{w}_1^T \mathbf{z} = \| \mathbf{w}_1 \| \| \mathbf{z} \| \cos \theta \quad (16-1. v)$$

که θ زاویه بین دو بردار است. اما چون $1 = \| \mathbf{w}_1 \|$

$$\mathbf{w}_1^T \mathbf{z} = \| \mathbf{z} \| \cos \theta \quad (17-1. v)$$

که طبق شکل ۱۰ همان تصویر \mathbf{z} روی \mathbf{w}_1 است. توضیحات مشابهی نیز برای \mathbf{w}_2 صدق می‌کند. حال فرض کنید که سه بردار متعامد $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ با بزرگی‌های واحد، متناظر با سه نقاط سه عنصری هستند. حاصل ضرب های $\mathbf{w}_1^T \mathbf{z}$ ، $\mathbf{w}_2^T \mathbf{z}$ و $\mathbf{w}_3^T \mathbf{z}$ بیانگر تصویرهای \mathbf{z} به ترتیب بر روی بردارهای $\mathbf{w}_1, \mathbf{w}_2$ و \mathbf{w}_3 می‌باشند. طبق شرح فوق این حاصل ضرب های پاسخ‌های انفرادی^۲ سه نقاط نیز هستند. به

۱۰۹ ۱۰.۷ آشکارسازی تایپوسنگی ها

شکل ۱۰.۷ تصویر z روی بردار واحد w_1

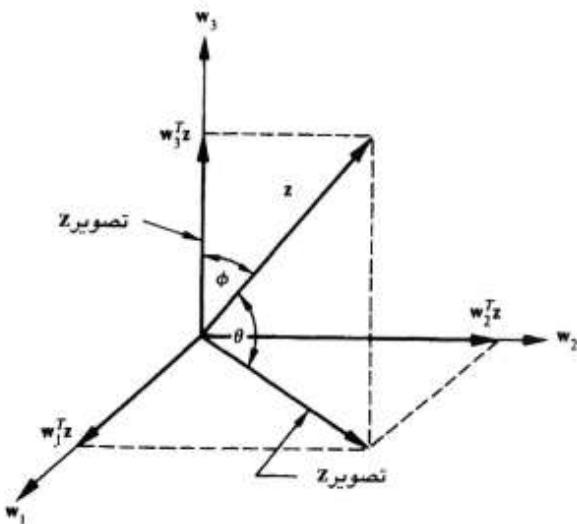
عنوان مورد بعدی فرض کنید که نقاط های ۱ و ۲ برای خطوط و نقاط ۳ برای لبه ها است. می خواهیم به این سؤال که آیا ناحیه ای که با z نشان داده شد، بیشتر به یک خط شبیه است یا یک لبه؟ پاسخ دهیم. دو نقاط اول خطوط را به ما می دهند اما ماتنها به ویژگیهای خطی z ، نه به نوع خاصی از خط، علاقه مندیم. بنابراین می توانیم با تصویر کردن z روی زیرفضای w_1 و w_2 (که در این مورد یک صفحه است) و نیز روی w_3 به سؤال فوق پاسخ دهیم. زاویه بین z و هر یک از این دو تصویر نشان می دهد که آیا z به زیرفضای خط نزدیک است یا به زیرفضای لبه. آرایش هندسی شکل ۱۱.۷ این وضعیت را نشان می دهد. بزرگی تصویر z روی صفحه w_1 و w_2 برابر کمیت $[(w_1^T z)^2 + (w_2^T z)^2]^{1/2}$ است. بزرگی (نرم^۱) z عبارتست از

$$\|z\| = [(w_1^T z)^2 + (w_2^T z)^2 + (w_3^T z)^2]^{1/2}. \quad (18-1.7)$$

پس زاویه بین z و تصویرش براین صفحه عبارتست از

$$\begin{aligned} \theta &= \cos^{-1} \left\{ \frac{[(w_1^T z)^2 + (w_2^T z)^2]^{1/2}}{[(w_1^T z)^2 + (w_2^T z)^2 + (w_3^T z)^2]^{1/2}} \right\} \\ &= \cos^{-1} \left\{ \frac{\left[\sum_{i=1}^2 (w_i^T z)^2 \right]^{1/2}}{\left[\sum_{i=1}^3 (w_i^T z)^2 \right]^{1/2}} \right\} \\ &= \cos^{-1} \left\{ \frac{1}{\|z\|} \left[\sum_{i=1}^2 (w_i^T z)^2 \right]^{1/2} \right\} \end{aligned} \quad (20-1.7)$$

1) Norm



شکل ۷.۱۱ تصویرهای z روی زیرفضای (صفحه) تعیین شده با w_1 و w_2 و روی زیرفضای w_3

که مرحله آخر روابط معادله فوق از معادله (۷.۱-۱۸) نتیجه گرفته شد. همچنین با روابط مشابهی زاویه تصویر بر زیرفضای w_3 بدست می‌آید:

$$\begin{aligned}\phi &= \cos^{-1} \left\{ \frac{1}{\|z\|} \left[\sum_{j=1}^3 (\mathbf{w}_j^T \mathbf{z})^2 \right]^{1/2} \right\} \\ &= \cos^{-1} \left\{ \frac{1}{\|z\|} |\mathbf{w}_3^T \mathbf{z}| \right\}.\end{aligned}\quad (۷.۱-۲۰)$$

بنابراین اگر $\phi < \theta$ باشد، گفته می‌شود که ناحیه z مشخصات نزدیکتری به یک خط دارد تا به برای نقاط های 3×3 مسئله نه بعدی می‌شود. هنوز هم مفاهیم معتبر است، گرچه برای تشکیل مجموعه پایه کامل بردارهای متعامد نه بعدی مورد نیاز است. نقاط های شکل ۷.۱۲ (پیشنهاد Frei و Chen [1977]) شرط تعامد را برآورده می‌کنند. چهار نقاط اول برای آشکارسازی لبه ها مناسب هستند؛ چهار نقاط دوم الگوهای^۱ مناسب برای آشکارسازی خط هستند؛ و آخرین نقاط (که برای تکمیل پایه ها اضافه شده است) متوسط پیکسل های ناحیه زیرنقطا را می‌دهند.

1) Template

۵۱۱ ۷. آشکارسازی ناپوستگی ها

$\begin{array}{ c c c } \hline 1 & \sqrt{2} & 1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -\sqrt{2} & -1 \\ \hline \end{array}$ w_1	$\begin{array}{ c c c } \hline 1 & 0 & -1 \\ \hline \sqrt{2} & 0 & -\sqrt{2} \\ \hline 1 & 0 & -1 \\ \hline \end{array}$ w_2	}
$\begin{array}{ c c c } \hline 0 & -1 & \sqrt{2} \\ \hline 1 & 0 & -1 \\ \hline -\sqrt{2} & 1 & 0 \\ \hline \end{array}$ w_3	$\begin{array}{ c c c } \hline \sqrt{2} & -1 & 0 \\ \hline -1 & 0 & 1 \\ \hline 0 & 1 & -\sqrt{2} \\ \hline \end{array}$ w_4	
$\begin{array}{ c c c } \hline 0 & 1 & 0 \\ \hline -1 & 0 & -1 \\ \hline 0 & 1 & 0 \\ \hline \end{array}$ w_5	$\begin{array}{ c c c } \hline -1 & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 0 & -1 \\ \hline \end{array}$ w_6	}
$\begin{array}{ c c c } \hline 1 & -2 & 1 \\ \hline -2 & 4 & -2 \\ \hline 1 & -2 & 1 \\ \hline \end{array}$ w_7	$\begin{array}{ c c c } \hline -2 & 1 & -2 \\ \hline 1 & 4 & 1 \\ \hline -2 & 1 & -2 \\ \hline \end{array}$ w_8	
$\begin{array}{ c c c } \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$ w_9		زیرفضای خط

شکل ۷.۲۰.۷ نقاب های معتمد (همانطور که دیده می شود، w_i ها تراز نشده اند) . (از [Frei & Chen[1977])

برای یک ناحیه 3×3 بانماد \mathbf{z} و براساس این فرض که بردارهای $\mathbf{w}_i, i = 1, 2, \dots, 9$ ، تراز شده‌اند، از توضیح قبلی نتیجه می‌شود که

$$p_e = \left[\sum_{i=1}^4 (\mathbf{w}_i^T \mathbf{z})^2 \right]^{1/2} \quad (21-1. v)$$

$$p_l = \left[\sum_{i=5}^8 (\mathbf{w}_i^T \mathbf{z})^2 \right]^{1/2} \quad (22-1. v)$$

و

$$p_a = |\mathbf{w}_9^T \mathbf{z}| \quad (23-1. v)$$

که p_e, p_l و p_a به ترتیب بزرگی تصویرهای \mathbf{z} بر روی زیرفضاهای لبه، خط و متوسط هستند. به طور مشابه:

$$\theta_e = \cos^{-1} \left\{ \frac{1}{\|\mathbf{z}\|} \left[\sum_{i=1}^4 (\mathbf{w}_i^T \mathbf{z})^2 \right]^{1/2} \right\} \quad (24-1. v)$$

$$\theta_l = \cos^{-1} \left\{ \frac{1}{\|\mathbf{z}\|} \left[\sum_{i=5}^8 (\mathbf{w}_i^T \mathbf{z})^2 \right]^{1/2} \right\} \quad (25-1. v)$$

و

$$\theta_a = \cos^{-1} \left\{ \frac{1}{\|\mathbf{z}\|} |\mathbf{w}_9^T \mathbf{z}| \right\} \quad (26-1. v)$$

که θ_e, θ_l و θ_a به ترتیب زوایای بین \mathbf{z} و تصویرهایش بر روی زیرفضاهای لبه، خط و متوسط هستند. البته این مفاهیم برای هر مجموعه 'پایه با ابعاد دلخواه' که بردارهای پایه آن متعامد باشند، تعمیم پذیر است.

مثال: تصویر موجود در شکل ۷. ۱۳. (الف) عکس هواپیک ورزشگاه فوتبال است. اشکال ۷. ۱۳. (ب)-(د) به ترتیب بزرگی تصویر برهیک از نقاب‌های پایه ۱-۹ از شکل ۷. ۱۲. را نشان می‌دهند. شکل ۷. ۱۳. (ذ) بزرگی تصویر بر زیرفضای لبه [معادله (۲۱-۱. v)]، و شکل ۷. ۱۳. (ر) بزرگی تصویر بر زیرفضای خط [معادله (۲۲-۱. v)] را نشان می‌دهد. در این مثال تصویر بر زیرفضای لبه بزرگترین پاسخ را دارد که نشانه 'محتوای لبه' بزرگی در تصویر اویله می‌باشد.



۲۰.۷ پیوند لبه و آشکارسازی مرز

فتون مشروح در بخش ۱۰.۴ ناپیوستگی های شدت را آشکار می کنند. در حالت مطلوب، این فتون تنها باید پیکسل هایی را آشکارسازند که روی مرز بین نواحی هستند. در عمل به علی مانند وجود نویز، انفصالات مرز که ناشی از روشنایی غیریکنواخت است، و سایر عواملی که ناپیوستگی های غیرطبیعی شدت را ایجاد می کنند، بندرت این پیکسل ها مرز کامل را مشخص می سازند. بنابراین معمولاً به دنبال اجرای الگوریتم های آشکارسازی لبه، الگوریتم های پیوند و الگوریتم های آشکارسازی سایر مرز های اینز اجرا می شوند تا در نهایت مرز های معناداری به دست آید. چندین روش برای این منظور مناسب هستند.

۱۰.۷ پردازش محلی

در تصویری که لبه های آن آشکار شده است، یکی از ساده ترین راهها برای پیوند نقاط لبه، تحلیل مشخصات پیکسل ها در یک همسایگی کوچک (مثالاً 3×3 یا 5×5) حول هر نقطه (y, x) است. تمام نقاط مشابه به هم پیوند می خورند، و بدین ترتیب مرزی از پیکسل هایی که چند خاصیت مشترک دارند، تشکیل می شود.

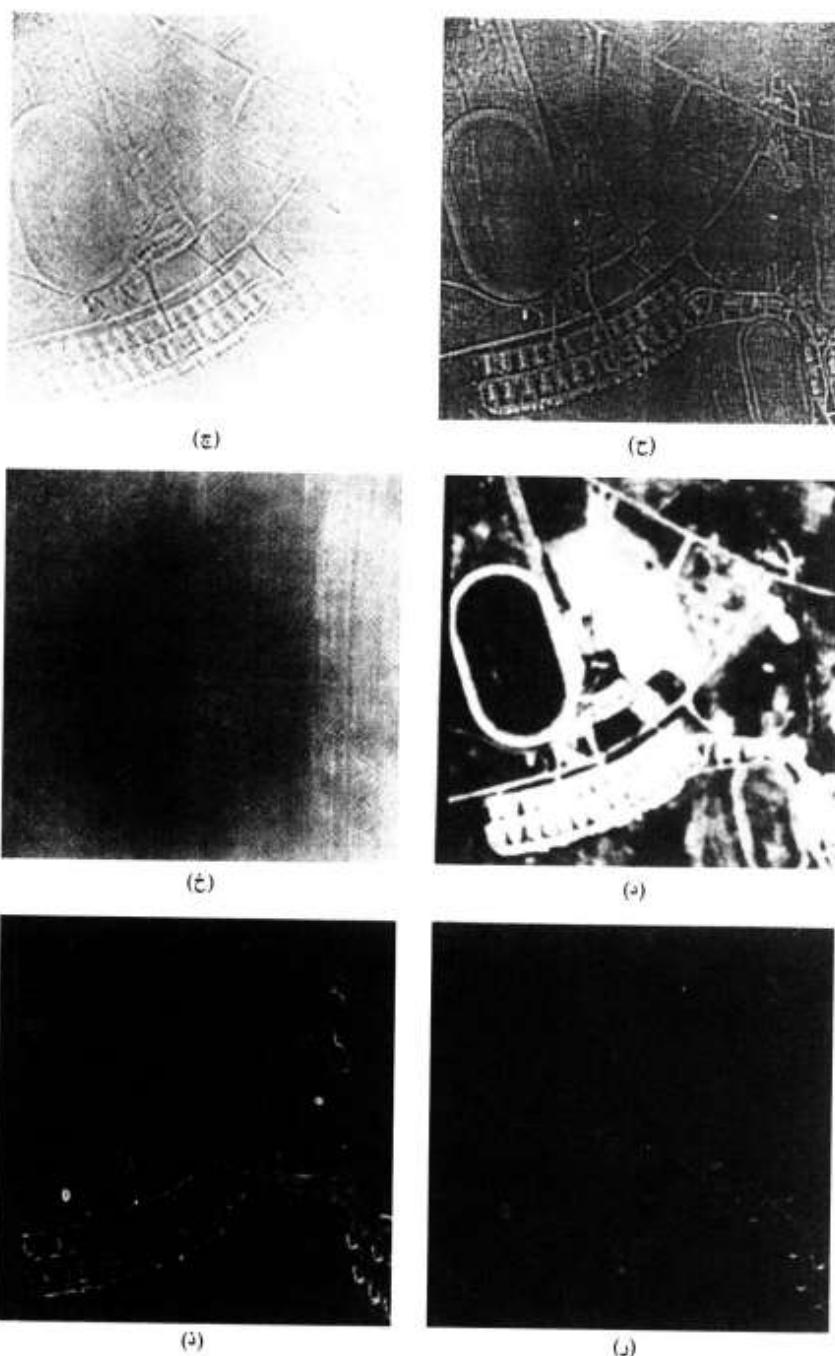
در این تحلیل دو خاصیت اصلی مورد استفاده برای تعیین شباهت پیکسل های لبه عبارتند از: (۱) بزرگی پاسخ عملگر گرادیان به کار رفته برای آشکارسازی لبه، و (۲) جهت گرادیان. همان طور که در معادله $\nabla f(x, y) = \alpha(x, y)$ یا $\nabla f(x, y) = \alpha(x, y)$ تعریف شد، اولین خاصیت با مقدار $\|\nabla f\|$ تعیین می شود. بنابراین یک پیکسل لبه با مختصات (x, y) و در همسایگی پیش تعریف شده (x', y') ، بزرگی گرادیان مشابهی با پیکسل (x', y') دارد اگر:

$$|\nabla f(x, y) - \nabla f(x', y')| \leq T \quad (1-2.7)$$

که T یک آستانه نامتفقی است.

جهت بردار گرادیان با معادله $\nabla f(x, y) = \alpha(x, y)$ تعیین می شود. بنابراین یک پیکسل لبه در (x, y) و در همسایگی پیش تعریف شده (x', y') ، زاویه مشابهی با پیکسل (x', y') دارد اگر:

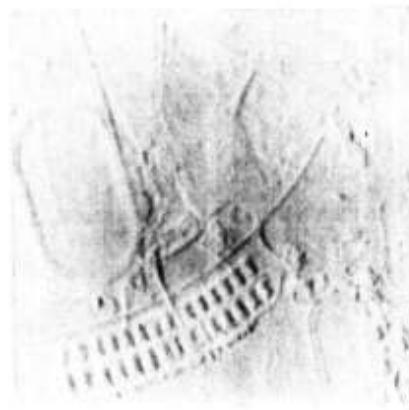
$$|\alpha(x, y) - \alpha(x', y')| < A \quad (2-2.7)$$



شکل ۱۳.۷ (الف) تصویر اولیه؛ (ب)-(ج) تصویرها به ترتیب بر زیرفضاهای w_1, w_2, w_3, w_4, w_5 ، (از *C. Hall & Frei [1976]*)



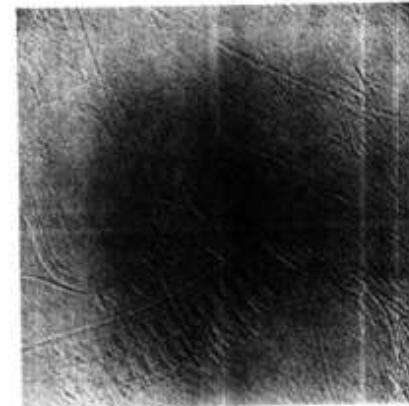
(الف)



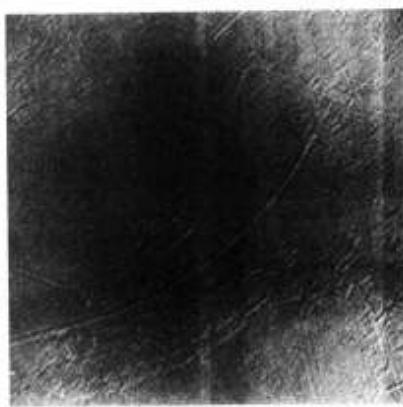
(ب)



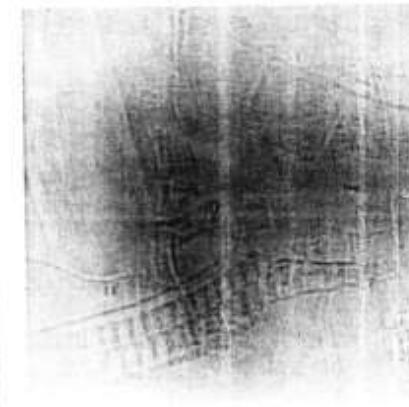
(ب)



(ج)



(د)



(ه)

شکل ۷.۱۳ (دامنه) (ج)-(د) تصویرها به ترتیب بر زیرفضاهای w_6 , w_7 , w_8 و w_9 ; (ذ) بزرگی تصویر بر زیرفضای لبه؛ (ر) بزرگی تصویر بر زیرفضای خط. (از [Hall & Frei 1976])

پردازش تصویر رقمه ۵۱۶

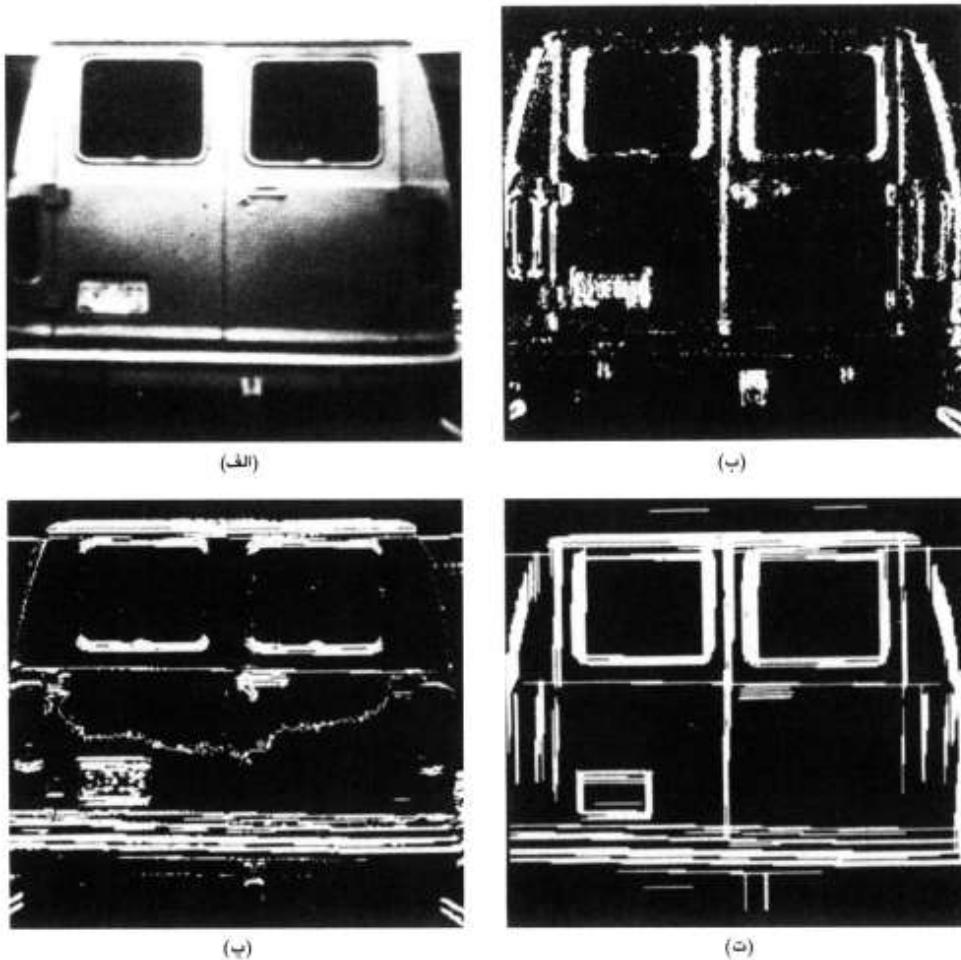
که ۱/ یک آستانه' زاویه است. توجه کنید که در واقع جهت لبه در (y, x) عمود بر جهت بردار گرادیان در آن نقطه است؛ با اینحال چون معادله' $(x-2, y-2)$ جهات را مقایسه می کند، تفاوتی ایجاد نمی شود. یک نقطه در همسایگی پیش تعریف شده' (y, x) به پیکسل (y, x) پیوند می خورد اگر هر دو معیار بزرگی و جهت برآورده شوند. این فرآیند برای تمام نقاط تصویر تکرار می شود. وقتی مرکز همسایگی پیکسل به پیکسل جابجا می شود، باید سابقه ای از نقاط پیوند خورده حفظ شود. یک روال ساده ثبت و قایع این است که به هر مجموعه از پیکسل های لبه پیوند خورده، سطح خاکستری متفاوتی با سایر مجموعه ها متنسب کیم.

مثال: برای روشن ساختن روال قبل، شکل ۷. ۱۴. (الف) را که تصویر عقب یک اتمبیل است، در نظر بگیرید. هدف پیدا کردن مستطیل هایی است که ابعادشان مناسب پلاک اتمبیل باشد. می توان با آشکار سازی لبه های مهم افقی و عمودی این مستطیل ها را تشکیل داد. اشکال ۷. ۱۴. (ب) و (پ) مؤلفه های عملگرهای سوبل را نشان می دهند. در نهایت شکل ۷. ۱۴. (ت) حاصل پیوند تمام نقاطی است که هم مقدار گرادیانشان بزرگتر از ۲۵ است و هم جهت گرادیانشان بیش از 15° اختلاف ندارند. با به کار بردن این معیارها بر سطرهای شکل ۷. ۱۴. (پ) خطوط افقی بدست آمدند، همچنین با پیمایش ترتیبی ستون های شکل ۷. ۱۴. (ب) خطوط عمودی بدست آمدند. پردازش بیشتر شامل پیوند قطعه لبه هایی که با انفصالت کوچک از هم جدا شده بودند و نیز حذف قطعه لبه های کوتاه منفرد بود. □

۴.۲.۷ پردازش سراسری با تبدیل هاف

در این بخش پیوند نقاط لبه را با تعیین این که آیا روی یک منحنی با شکل معین قرار دارند، بررسی می کنیم. در اینجا برخلاف روش تحلیلی محلی مشروح در بخش ۴. ۲. ۷، روابط سراسری بین پیکسل ها را بررسی می کنیم.

فرض کنید که برای n نقطه از تصویر می خواهیم زیر مجموعه هایی از این نقاط را که روی خطوط راست قرار دارند، پیدا کنیم. یک راه حل ممکن این است که ابتدا تمام خطوط را که با هر زوج نقطه معین می شوند، پیدا کنیم. مشکل این روال این است که شامل پیدا کردن $n^2 - n/2$ خط و سپس اجرای $n^3 - n(n-1)/2$ مقایسه هر نقطه با تمام نقاط خطوط است. این روش در تمام کاربردها جز تعدادی از کاربردهای کوچک، مانع محاسباتی دارد.



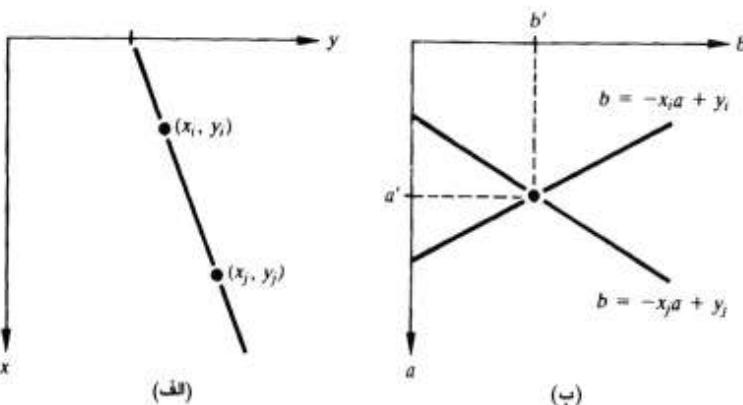
شکل ۱۴.۷ (الف) تصویر ورودی؛ (ب) مؤلفه G_y گرادیان؛ (پ) مؤلفه G_x گرادیان؛ (ت) حاصل پیوند لبه. (با مجوز از *Perceptics Corporation*)

هاف [۱۹۶۲] راه دیگری را پیشنهاد کرد که معمولاً با عنوان تبدیل هاف^۱ به آن اشاره می شود. نقطه (x_i, y_i) و معادله کلی یک خط راست در قالب شبیب-عرض از میدآ، $y_i = ax_i + b$ ، را در نظر بگیرید. خطوط بسیاری از (x_i, y_i) عبور می کنند، که هر یک از آنها بازای مقادیری از a و b معادله $y_i = ax_i + b$ را برآورده می کنند. با این حال با نوشتن این معادله به صورت $x_i a + y_i b = 1$ و در نظر گرفتن صفحه^۲ (که فضای عامل^۳ نیز خوانده می شود)، بازای یک زوج ثابت (x_i, y_i) تنها یک معادله خط در صفحه ab

1) Hough transform

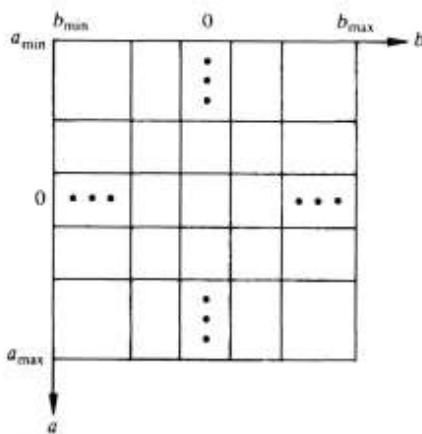
2) Parameter space

3) Single

شکل ۷.۱۵. (الف) صفحه xy ; (ب) فضای عامل

بدست می آید. گذشته از این نقطه دیگر (x_j, y_j) نیز در فضای عامل، خطی مربوط به خوددارد، و این خط با خط مربوط به (x_i, y_i) در (a', b') تلاقی می کند، که a' شیب و b' عرض از مبدأ خط گذرنده بر دو نقطه (x_i, y_i) و (x_j, y_j) در صفحه xy است. در واقع هریک از نقاط روی این خط، خطی در فضای عامل دارد که همه آنها در (a', b') تلاقی می کنند. شکل ۷.۱۵. این مفاهیم را نمایش می دهد.

همان طور که در شکل ۷.۱۶ نمایش داده شد، جذابیت محاسباتی تبدیل هاف ناشی از تقسیم فضای پارامتر به سلول های مشهور به ابانتگر^۱ است، که (a_{\min}, b_{\min}) و (a_{\max}, b_{\max}) محدوده های مورد انتظار مقادیر شیب و عرض از مبدأ هستند. سلول (i, j) ، با مقدار ابانتگر $A(i, j)$ ، متناظر با مربع مربوط به



شکل ۷.۱۶. چندی سازی صفحه عامل برای استفاده در تبدیل هاف.

مختصات فضای عامل (a, b) است. در ابتدا این سلول‌ها صفر قرار داده می‌شوند. آنگاه برای هر نقطه (x_k, y_k) در صفحه 'تصویر'، عامل a را برابر هریک از مقادیر زیر تقسیم مجاز روی محور a قرار می‌دهیم و با استفاده از معادله $-x_k a + y_k = b$ ، مقدار b متناظر را پیدا می‌کنیم. آنگاه b ‌های حاصل را به نزدیکترین مقدار مجاز روی محور b گرد می‌کنیم. اگر با انتخاب a ، جواب b به دست آید، قرار می‌دهیم xy است که روی خط $y = ax + b$ قرار دارند. دقت همراستایی^۱ این نقاط با تعداد تقسیم‌های صفحه 'ab' تعیین می‌شود.

توجه شود که تقسیم محور a به K قسمت، برای هر نقطه (x_k, y_k) ، K مقدار b متناظر با K مقدار ممکن a را نتیجه می‌دهد. با n نقطه 'تصویر'، این روش nK محاسبه دارد. بنابراین روش اخیر نسبت به n ، خطی است و حاصل ضرب nK به تعداد محاسبات مشروح در ابتدای این بخش نمی‌رسد مگر این که K به n برسد یا از آن تجاوز کند.

یک مشکل در مورد استفاده از معادله $y = ax + b$ برای توصیف خط این است که وقتی خط عمودی می‌شود، شبیه و عرض از مبدأ آن به بی‌نهایت میل می‌کند. یک راه حل برای این مشکل استفاده از توصیف نرمال^۲ خط است:

$$x \cos \theta + y \sin \theta = \rho. \quad (3-2.7)$$

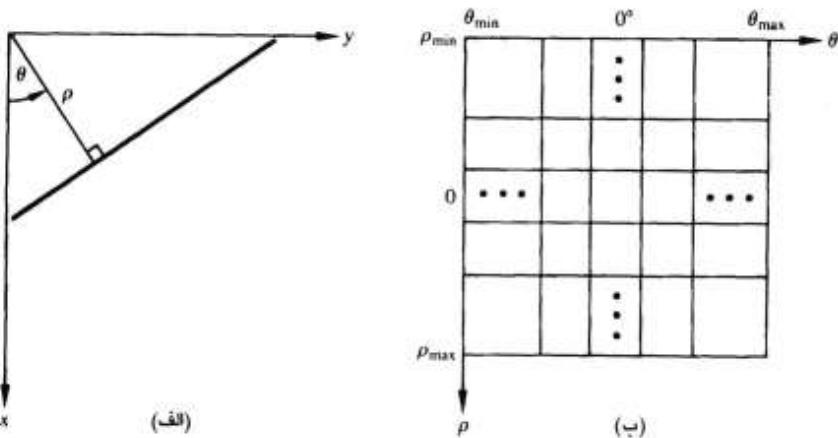
شکل ۷.۱۷ (الف) معنای عوامل مورد استفاده در معادله $(3-2.7)$ را نشان می‌دهد. از نظر تشکیل جدول ابانتگرها، این توصیف تفاوتی با توصیف شبیه-عرض از مبدأ ندارد. با این حال منحنی‌ها در صفحه 'rho theta' به جای خطوط راست سینوسی هستند. مانند گذشته M نقطه همراستا که روی خط $x \cos \theta_i + y \sin \theta_i = \rho_i$ قرار داشته باشند، در فضای عامل $\rho \theta$ ، M منحنی سینوسی دارند که در $A(i, j)$ تلاقی می‌کنند. با افزایش گام به گام θ و پیدا کردن ρ ‌های متناظر، M در ابانتگر (j) که مربوط به سلول معین شده با (ρ_i, θ_i) است، بدست می‌آید. شکل ۷.۱۷ (ب) تقسیم‌بندی فضای عامل را نمایش می‌دهد.

محدوده 'زاویه' θ ، $90^\circ \pm 90^\circ$ است که نسبت به محور x سنجیده می‌شود. بنابراین طبق شکل ۷.۱۷ (الف)، یک خط افقی یا $\theta = 0^\circ$ دارد که در آن ρ برابر طول از مبدأ مثبت است، یا $\theta = -90^\circ$.

1) Collinearity

2) Normal representation

پردازش تصویر رقمنی ۵۲۰

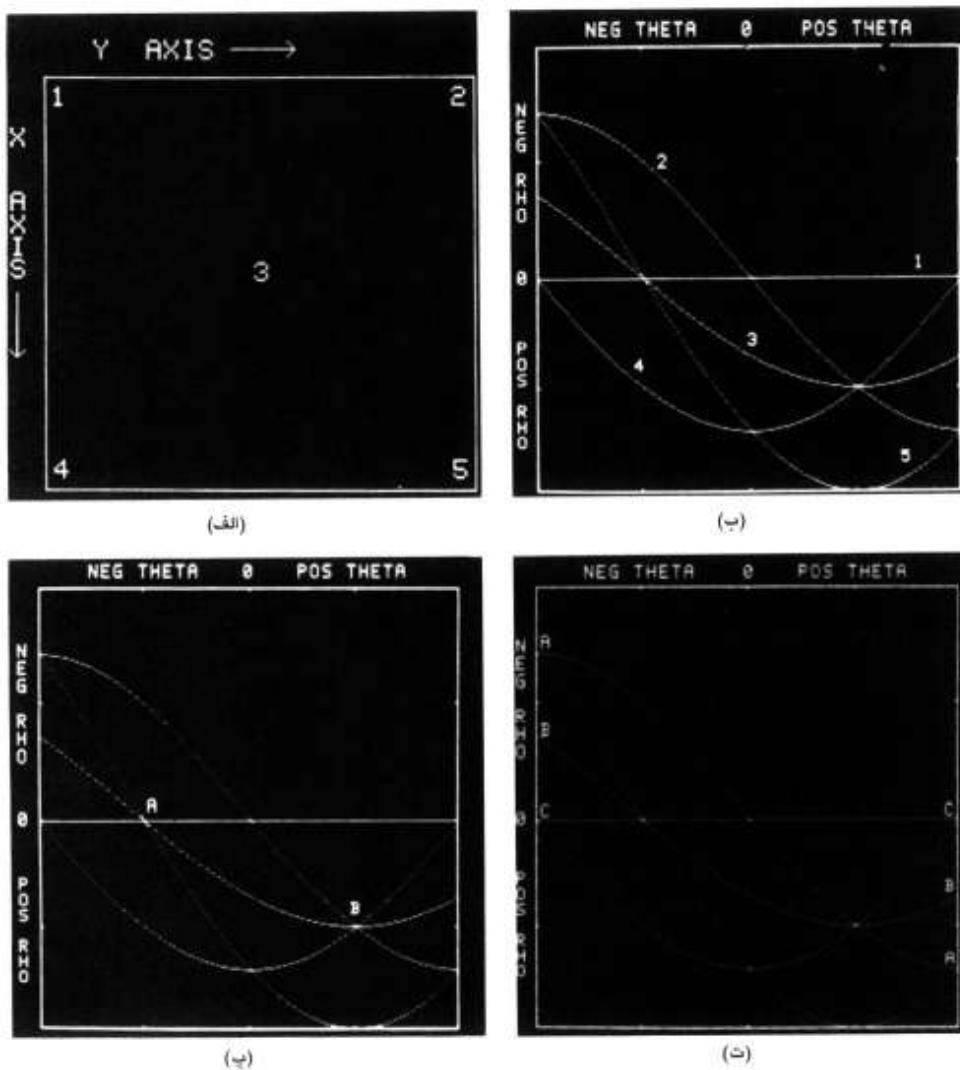


شکل ۷.۷ (الف) نمایش نرمال یک خط؛ (ب) چندی سازی صفحه $\rho\theta$ به سلولها

دارد که در آن ρ برابر عرض از مبدأ منفی است.

مثال: شکل ۷.۱۸ تبدیل هاف مبتنی بر معادله $(7.2-3)$ را نمایش می دهد. شکل ۷.۱۸.(الف) تصویری با پنج نقطه بر جسب دار را نشان می دهد. همان طور که در شکل ۷.۱۸.(ت) دیده می شود، هریک از این نقاط روی صفحه $\rho\theta$ نگاشت می شوند. مقادیر θ در محدوده $90^\circ \pm 45^\circ$ قرار دارند و مقادیر ρ در محدوده $\sqrt{2}D \pm D$ هستند که D فاصله بین گوشه های تصویر است. برخلاف تبدیل مبتنی بر استفاده از شبیه عرض از مبدأ، هریک از این منحنی ها شکل سینوسی متفاوتی با دیگران دارد. (خط افقی حاصل از نگاشت نقطه ۱ مورد خاصی از سینوس با دامنه صفر است). از هر گوشه تصویر نقطه ای نگاشت شد، بنابراین تبدیل هاف هر نقطه دیگری از تصویر بین حدود نشان داده شده در شکل ۷.۱۸.(ب) خواهد بود.

در شکل ۷.۱۸.(ب) خاصیت آشکار سازی هم راستایی تبدیل هاف نمایش داده می شود. نقطه A محل تلاقی منحنی های متناظر با نقاط ۱، ۳ و ۵ از صفحه تصویر $\rho\theta$ است. محل نقطه A نشان می دهد که این سه نقطه روی خط مستقیمی قرار دارند که از مبدأ می گذرد ($\theta = 0$) و در جهت $45^\circ - 45^\circ$ است. به طور مشابه تلاقی منحنی ها در نقطه B فضای عامل دال براین است که نقاط ۲، ۳ و ۴ روی خط مستقیمی قرار دارند که در جهت 45° است و فاصله اش از مبدأ نصف فاصله قطری مبدأ تصویر تا گوشه مخالف است.



شکل ۱۸.۷ نمایش تبدیل هاف. (با مجوز از Texas Instruments، D.R. Cate)

در نهایت شکل ۱۸.۷ (ت) دال بر این واقعیت است که تبدیل هاف در لبه های راست و چپ فضای عامل ارتباط مجاری آینه ای^۱ (نسبت به محور θ) را به نمایش می گذارد. این خاصیت که با نقاط A ، B و C از شکل ۱۸.۷ (ت) نشان داده می شود، نتیجه تغییر علامت θ و ρ در مرزهای $90^\circ \pm 90^\circ$ است. □

1) Reflective adjacency relationship

گرچه تاکنون تمرکز روی خطوط راست بوده است، تبدیل هاف را می توان برای هر تابع با شکل کلی $g(v, c) = 0$ به کاربرد، که ۷ بردار مختصات و ۳ بردار ضرایب است. مثلاً می توان با استفاده از راه اخیر، نقاطی را که روی دایره^۱

$$(x - c_1)^2 + (y - c_2)^2 = c_3^2 \quad (4-2. 7)$$

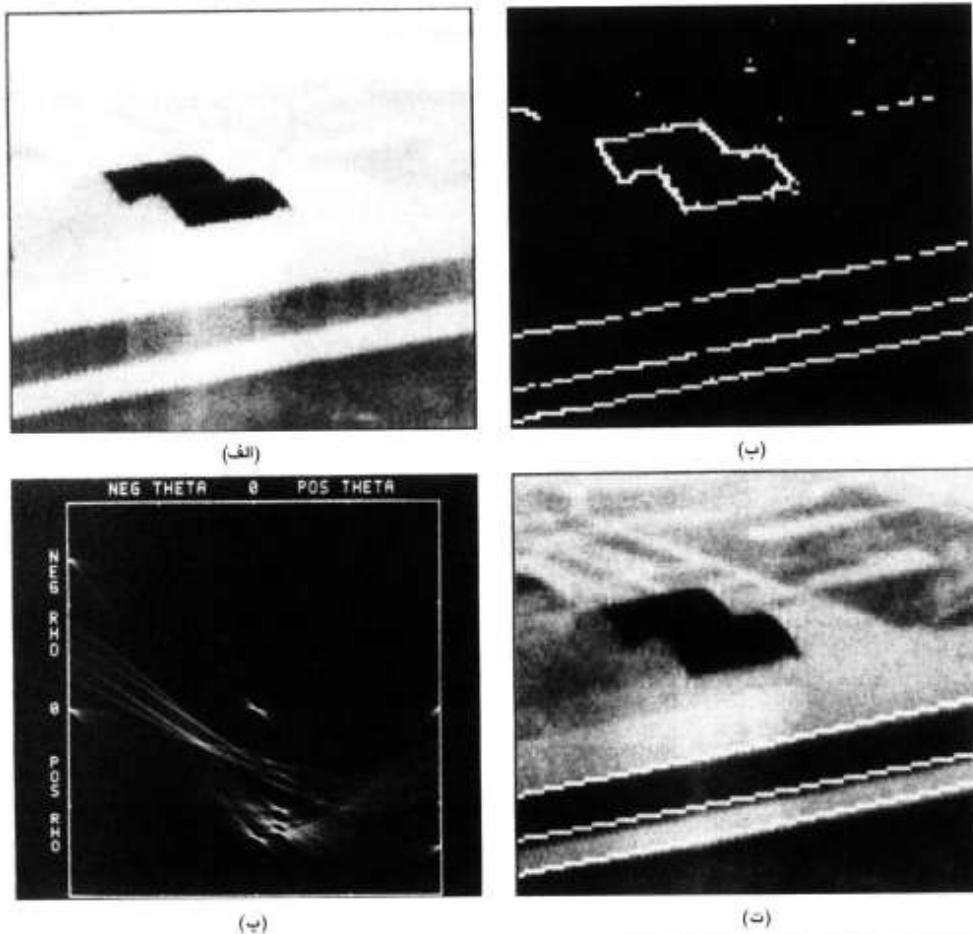
قرار دارند، آشکار کرد. تفاوت اصلی وجود سه عامل (c_1, c_2, c_3) است، که قضای عامل سه بعدی با سلول های شبه مکعبی و ابناشترگر هایی به شکل $A(i, j, k)$ را نتیجه می دهد. روال کار این است که c_1, c_2, c_3 را گام به گام افزایش دهیم و مقدار c که معادله $(4-2. 7)$ را براورد می کند، پیدا کنیم، و ابناشترگر متناظر با سلول مربوط به سه تایی (c_1, c_2, c_3) را به نگام سازیم. واضح است که پیچیدگی تبدیل هاف بستگی شدیدی به تعداد مختصات و ضرایب موجود در توصیف تابع داده شده دارد. می توان تبدیل هاف را برای آشکار سازی منحنی های بانمایش تحلیلی غیر ساده نیز تعمیم داد. [Ballard 1981] مفصلأً به این مفاهیم که گسترش های مبحث ارائه شده در اینجا هستند، پرداخته است.

حال به مسئله پیوند لبه باز می گردیم. روش پیوند لبه مبتنی بر تبدیل هاف شامل مراحل زیر است:

(۱) محاسبه 'گرادیان تصویر'، (۲) تعیین تقسیم های صفحه $\rho\theta$ ، (۳) بررسی شمارش سلول های ابناشت گر برای تعیین تمرکز های بالای پیکسل ها، و (۴) بررسی رابطه 'بین پیکسل ها در یک سلول منتخب (اصولاً برای پیوستگی^۱)'. عموماً مفهوم پیوستگی در این مورد بر اساس محاسبه 'فاصله' بین پیکسل های نامتصلی است که در حین پیمایش پیکسل های مربوط به با یک سلول ابناشترگر معین شناسایی شده اند. فاصله نقاط در صورتی اهمیت دارد که فاصله 'بین یک نقطه و نزدیکترین همسایه اش از یک آستانه' معین تجاوز کند. (برای شرحی درباره 'اتصال، همسایگی ها و معیارهای فاصله' به بخش ۴. ۲ مراجعه کنید).

مثال: شکل ۷. ۱۹. (الف) را در نظر بگیرید که یک تصویر هوایی مادون قرمز حاوی دو آشیانه 'هوایما' و یک باند پرواز را نشان می دهد. شکل ۷. ۱۹. (ب) تصویر گرادیان آستانه گیری شده حاصل از کاربرد عملگر های سوبل بخش ۱. ۳. بر تصویر اصلی است (به فاصله های کوچک در مرز های باند پرواز توجه کنید). شکل ۷. ۱۹. (پ) تبدیل هاف خطی تصویر گرادیان را نشان می دهد، و شکل ۷. ۱۹. (ت)

1) Continuity



شکل ۱۹.۷ (الف) یک تصویر مادون قرمز؛ (ب) تصویر گرادیان آن؛ (پ) تبدیل هاف آن؛ (ت) پیکسل های پیوند خورده (با مجوز D.R. Cate, Texas Instruments, Inc.)

مجموعه' پیکسل های پیوند خورده طبق دومعيار زير را (بارنگ سفید) نشان مي دهد: (۱) پیکسل ها به يكى از سه سلول انباشتگري كه بالاترین شمارش را دارند، متعلق هستند، و (۲) هيج فاصله اى طولانيتر از پنج پیکسل نیست. به ناپديد شدن فاصله ها بر اثر پیوند زدن توجه كنيد.

۳۰.۷ بوداژش سراسری با روش های نظریه گراف^۱

روش مشروح در بخش ۷.۲.۰ براساس به دست آوردن مجموعه اى از نقاط لبه با يك عملگر گرادیان

۱) Graph-theoretic techniques

است. چون گرادیان یک نوع مشتق است، به ندرت به عنوان مرحله' پیش پردازش در وضعیتهای دارای نویز زیاد مناسب است. در این بخش یک روش سراسری را شرح می دهیم که براساس نمایش قطعات لبه به شکل گراف و سپس جستجوی گراف برای مسیرهای کم هزینه' متناظر بالهای مهم می باشد. این نمایش راه قابل اعتمادی را برای عملکرد مناسب در حضور نویز ارائه می نماید. همان طور که ممکن است مورد انتظار باشد، این روال بسیار پیچیده تر است و نسبت به روش هایی که تاکنون تشریح شده اند، به زمان پردازش بیشتری نیاز دارد.

توسعه را با چند تعریف اصلی شروع می کنیم. گراف $G = (N, A)$ مشکل از مجموعه محدود و غیرتهی N از گره ها و مجموعه A از زوج های نامرتب عناصر متفاوت N است. هر زوج (n_i, n_j) از A یک کمان^۱ نامیده می شود. هر گرافی که کمان های آن جهتدار باشند، گراف جهتدار^۲ خوانده می شود. اگر کمان درجهت گره n_i به n_j باشد، به n_j دنباله^۳ گره مادر^۴ n_i گویند. فرآیند شناسایی دنباله های یک گره، گسترش^۵ آن گره نامیده می شود. در هر گراف سطوح را چنان تعریف می کنیم که سطح^۶ شامل تنها یک گره، که گره' شروع^۷ نامیده می شود، باشد؛ گره های آخرین سطح گره های هدف^۸ خوانده می شوند. به هر کمان (n_i, n_j) می توان هزینه^۹ $c(n_i, n_j)$ را متنسب کرد. دنباله ای از گره های $\dots, n_1, n_2, \dots, n_k$ که هر گره n_k جانشین گره n_i باشد، یک مسیر^{۱۰} از n_i به n_k خوانده می شود، و هزینه این مسیر

عبارتست از:

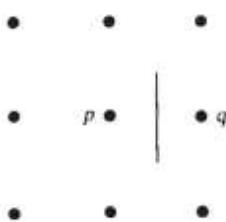
$$c = \sum_{j=2}^k c(n_{j-1}, n_j). \quad (5-2. 7)$$

در نهایت همان طور که در شکل ۷. ۲۰ دیده می شود، یک عنصر لبه^۹ مرزبین دو پیکسل p و q است که p و q همسایه های چهارگانه هستند. در این متن، یک لبه دنباله ای از عناصر لبه است. می توانیم نشان دهیم که چگونه این مفاهیم، برای آشکارسازی لبه در تصویر 3×3 شکل ۷. ۲۱ به کار می روند، که اعداد بیرون پرانتر مختصات پیکسل هستند و اعداد درون پرانترها میان شدت هستند. هر عنصر لبه تعریف شده با پیکسل های p و q یک هزینه به صورت:

$$c(p, q) = H - [f(p) - f(q)] \quad (6-2. 7)$$

- | | | | | |
|----------|-------------------|--------------|-----------------|--------------|
| 1) arc | 2) Directed graph | 3) Successor | 4) Parent | 5) Expansion |
| 6) Start | 7) Goal | 8) Path | 9) edge element | |

۵۲۵ ۲۰. ۷ پیوند لبه و آشکارسازی مرز

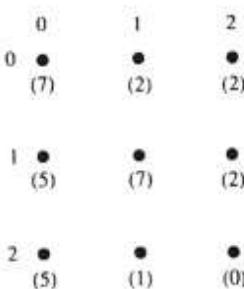


شکل ۷. ۲۰. ۷ عنصر لبه بین پکسل های p و q .

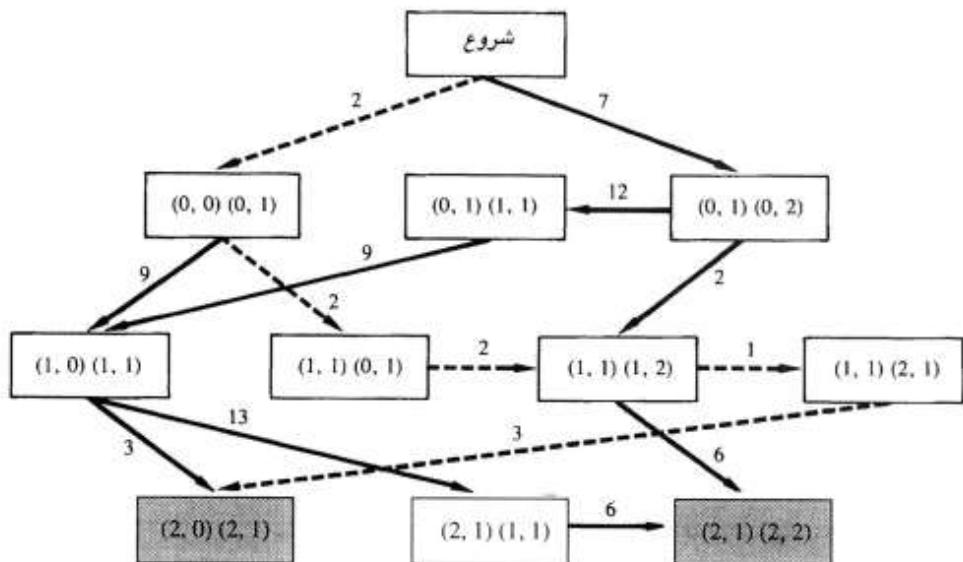
دارد که H بالاترین مقدار شدت در تصویر (در این مورد ۷) است، $f(p)$ مقدار شدت p و $f(q)$ مقدار شدت q است. همانطور که قبلاً بیان شد، p و q همسایه های چهارگانه هستند.

شکل ۷. ۲۲. گراف این مسئله را نشان می دهد. هر گره متناظر با یک عنصر لبه است، بین دو گره یک کمان وجود دارد اگر عناصر لبه متناظر که به دنبال هم در نظر گرفته می شوند، بتوانند بخشی از یک لبه باشند. هزینه هر عنصر لبه، که با استفاده از معادله (۷. ۶-۲) محاسبه می شود، در کنار کمان یا کمان های وارد شونده به آن نوشته شده است و گره های هدف به صورت مستطیل های سایه خورده نشان داده می شوند. هر مسیر بین گره 'شروع' و گره 'هدف' یک لبه ممکن است. برای سادگی فرض می شود که لبه در سطر بالا شروع می شود و در سطر پایین پایان می یابد، تا این که اولین عنصر لبه، تنها بتواند [[(2, 1), (2, 2), (2, 0), (2, 1)] یا [(0, 0), (0, 1), (0, 2)] و آخرین عنصر، تنها بتواند [(2, 1) یا [(0, 1)] باشد. خطوط مقطع بیانگر مسیر با هزینه کمیه هستند که با استفاده از معادله (۷. ۵-۲) محاسبه می شود.

شکل ۷. ۲۳. لبه متناظر آن را نشان می دهد.



شکل ۷. ۲۱. ۷ یک تصویر 3×3



شکل ۲۲.۷ گراف استفاده شده برای پیدا کردن لبه در تصویر شکل ۲۱.۷. زوج (a, b) در هر جعبه از چپ به راست به ترتیب به نقاط p و q اشاره می کنند. توجه کنید که فرض می شود در هنگام پیمایش بالا به پایین تصویر، p در سمت راست مسیر قرار دارد. خطوط مقطع یا نگر مسیر با هزینه کمینه هستند. (گرفته شده از [Martelli, 1972])

به طور کلی مسأله پیدا کردن مسیر با هزینه کمینه، محاسبات ناچیزی ندارد. نوع آرایه حل این مشکل، فداکردن بهینگی برای سرعت است، و الگوریتم زیر نماینده گروهی از روش‌هاست که به منظور کاهش جستجو از مکائسه^۱ استفاده می کنند. فرض کنید $r(n)$ تقریبی از مسیر با هزینه کمینه از گره 'شروع' به یک گره 'هدف' باشد، که این مسیر باید از n بگذرد. این هزینه را می توان به صورت تخمینی از هزینه یک مسیر با هزینه کمینه به علاوه 'تخمینی از هزینه' آن مسیر از n به گره هدف بیان کرد؛ یعنی،

$$r(n) = g(n) + h(n). \quad (6-2, 7)$$

در اینجا می توان $g(n)$ را به عنوان مسیر با هزینه کمینه از s به n که تاکنون پیدا شده است، انتخاب کرد، و $h(n)$ با استفاده از هرنوع اطلاعات شهودی موجود (نظیر گسترش گره‌های معینی براساس هزینه رسیدن به آن گره‌ها) به دست می آید. یک الگوریتم که از $r(n)$ به عنوان مبنایی برای انجام یک جستجوی گرافی استفاده می کند، به صورت زیر است:

1) Heuristics

۵۴۲ ۲.۷ پیوند به و آشکارسازی مرز

مرحله' ۱ : گره' شروع را «باز» علامت بزنید و قرار دهید $g(s) = 0$.

مرحله' ۲ : اگر هیچ گره' ای «باز» نباشد، با «شکست» خارج شوید؛ در غیر این صورت ادامه دهید.

مرحله' ۳ : گره' «باز» را که تخمین (n) آن بر حسب رابطه' $(7-2.7)$ کوچکترین است، «بسته» علامت بزنید. (گره' ها برای مقادیر کمینه' n به دلخواه، اما همواره بر حسب یک گره' هدف، انتخاب می شوند.)

مرحله' ۴ : اگر یک گره' هدف باشد، با پیمایش عکس مسیر حاصل از اشاره گرها^۱ را پایان دهید؛ در غیر این صورت ادامه دهید.

مرحله' ۵ : با تولید تمام دنباله های گره' n آن را گسترش دهید. (اگر هیچ دنباله ای وجود ندارد، به مرحله' ۲ بروید.)

مرحله' ۶ : اگر گره' دنباله n علامت ندارد، قرار دهید:

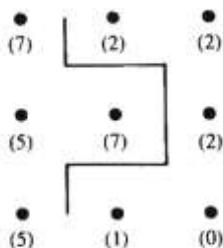
$$r(n_i) = g(n) + c(n, n_i)$$

آن را «باز» علامت بزنید و اشاره گرها را رو به n نمایید.

مرحله' ۷ : اگر گره' دنباله n «بسته» یا «باز» علامت خورده است، مقدارش را با رابطه'

$$g'(n_i) = \min[g(n_i), g(n) + c(n, n_i)].$$

بهنگام کنید. گره' های دنباله «بسته» که مقادیر g شان بدین ترتیب پایین آمدند، «باز» علامت بزنید و

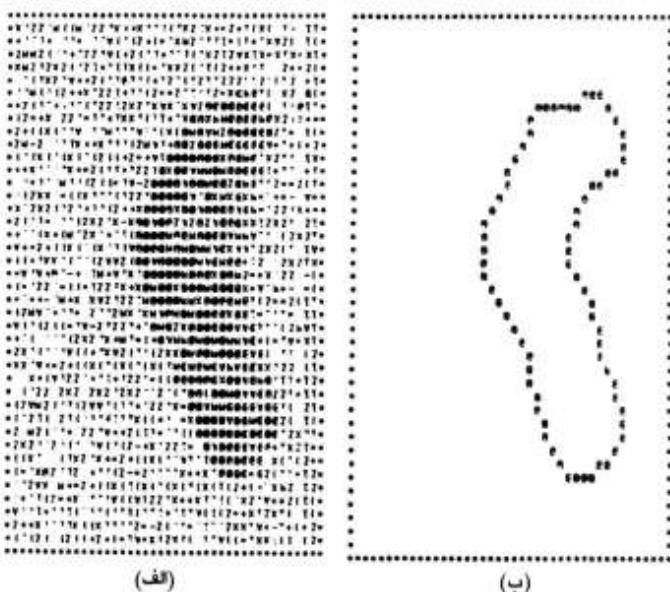


شکل ۲۳.۷ لبه متناظر با مسیر با هزینه کمینه در شکل ۷.۲۲.

اشاره گرها از تمام گره‌هایی که مقادیر g -شان پایین آمدند، به سمت h برگردانید. به مرحله^۱ بروید.

به طور کلی این الگوریتم پداکردن یک مسیر با هزینه کمینه را تضمین نمی‌کند؛ مزیت آن سرعت از طریق استفاده از شهود است. با این حال اگر (h) یک کران پایین برای هزینه کمینه از گره^۱ به گره^۰ هدف باشد، در واقع این روال مسیر بهینه‌ای به گره هدف را می‌دهد (Raphael و Nilsson [1968]). اگر هیچ اطلاعات شهودی موجود نباشد (یعنی $h \equiv 0$)، این روال به الگوریتم با هزینه یکنواخت^۱ Dijkstra[1959] ساده‌تر می‌شود.

مثال: شکل ۲۴.۷ نتیجه نوعی قابل دست یابی با این روال را نمایش می‌دهد. شکل ۲۴.۷ (الف) یک تصویر نویزدار، و شکل ۲۴.۷ (ب) حاصل بخش بندی لبه با جستجوی مسیرهای کم هزینه در گراف منتظر آن را نشان می‌دهد. با گسترش ندادن آن گره‌هایی که هزینه شان از یک آستانه معین تجاوز می‌کرد، از اطلاعات شهودی استفاده شد.



شکل ۲۴.۷ (الف) یک تصویر نویزدار؛ (ب) حاصل آشکارسازی لبه با استفاده از جستجوی مکافنه‌ای گراف. (از Martelli [1976])

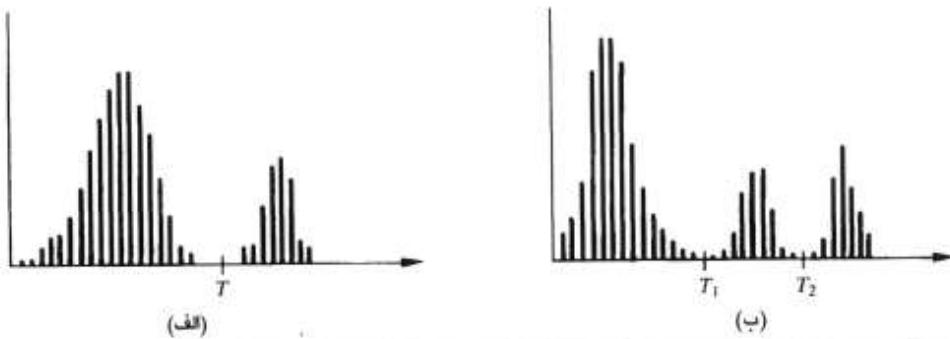
1) Uniform- cost algorithm

۳.۷ آستانه گیری

آستانه گیری یکی از مهم ترین راه های بخش بندی تصویر است. در این بخش فنون مختلفی را برای آستانه گیری ارائه می دهیم و شایستگی ها و محدودیت های هر یک از آنها را شرح می دهیم.

۱.۳.۷ معانی

فرض کنید که بافت نگار^۱ سطح خاکستری شکل ۷.۲۵ (الف) متناظر با یک تصویر $f(x, y)$ است که مشکل از اشیای روشن روی زمینه 'تیره' می باشد طوری که سطوح خاکستری پیکسل های شیشه و زمینه در دو مد^۲ غالب گروه بندی شده اند. یک راه واضح برای استخراج اشیا از زمینه، انتخاب یک مقدار آستانه^۳ T است طوری که این دو مدرا جدا کند. آنگاه هر پیکسل (y, x) که برای آن $T > f(x, y)$ یک پیکسل شیشه و در غیر این صورت یک پیکسل زمینه خوانده می شود. شکل ۷.۲۵ (ب) یک مورد کمی عمومی تر از این رهیافت را نمایش می دهد. در اینجا سه مد غالب بافت نگار، تصویر را توصیف می کنند (مثلاً دونوع شیشه روشن روی یک زمینه 'تیره'). همان رهیافت پایه^۴ پیکسل (y, x) را اگر $T_1 \leq f(x, y) \leq T_2$ در یک گروه از اشیا قرار می دهد، اگر $T_2 < f(x, y) \leq T$ در گروه دیگر قرار می دهد، و اگر $f(x, y) \leq T_1$ در گروه زمینه قرار می دهد. این نوع از آستانه گیری چند سطحی^۵ عموماً نامطمئن تر از شکل تک آستانه ای آن است. دلیل آن مشکل بودن تعیین چند آستانه است طوری که بخوبی بتوانند نواحی مورد علاقه را از هم جدا کنند، بویژه وقتی که تعداد مدهای بافت نگار عربو طه زیاد است. در حل چنین مسائلی با آستانه گیری،



شکل ۷.۲۵ بافت نگارهای سطح خاکستری که می توان آنها را (الف) با یک آستانه و (ب) با چند آستانه بخش بندی کرد.

پردازش تصویر رقمه ۵۳۰

عمولاآکار بر د آستانه' متغیر جواب بهتری می دهد.

بر اساس شرح قبلی می توان آستانه گیری را عملی در نظر گرفت که آزمون هایی را بر رویتابع T به شکل

$$T = T[x, y, p(x, y), f(x, y)] \quad (1-3. 7)$$

به کار می گیرد، که $f(x, y)$ سطح خاکستری پیکسل (x, y) و $p(x, y)$ بیانگر یک خاصیت محلی از این پیکسل - مثلاً سطح خاکستری متوسط در یک همسایگی به مرکز (x, y) - است. تصویر آستانه گیری شده به صورت $g(x, y)$ ^۱

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \leq T. \end{cases} \quad (2-3. 7)$$

تعریف می شود. بنابراین پیکسل های با برچسب ۱ (با هر سطح شدت مناسب دیگر) متناظر با اشیا هستند، در حالی که پیکسل های با برچسب ۰ متناظر با زمینه هستند.

وقتی T تنها به $f(x, y)$ بستگی داشته باشد، آستانه سراسری^۲ خوانده می شود. (شکل ۷-۲۵، الف) مثالی از چنین آستانه ای را نشان می دهد. اگر T هم به $f(x, y)$ و هم به $p(x, y)$ بستگی داشته باشد، آستانه محلی^۳ خوانده می شود. به علاوه اگر T به مختصات مکانی x و y نیز بستگی داشته باشد، آستانه پویا^۴ خوانده می شود.

۲۰۳۰۷ نقش روشنایی

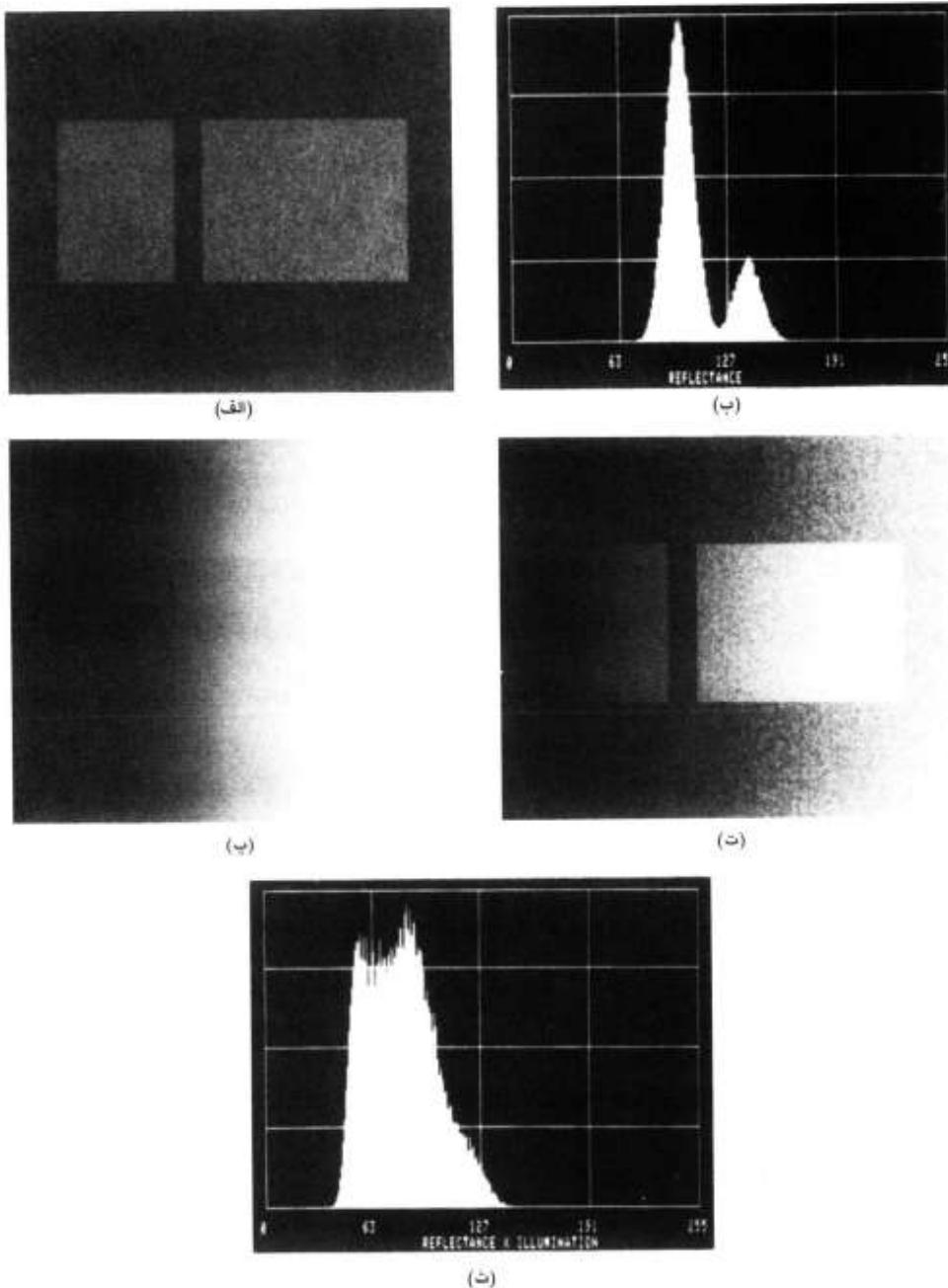
در بخش ۲. ۲. بیان کردیم که تصویر (x, y, f) را می توان به عنوان حاصل ضرب مؤلفه' انعکاس $r(x, y)$ و مؤلفه روشنایی $i(x, y)$ در نظر گرفت. هدف از این بخش شرح مختصر اثر روشنایی بر بخش بندی تصویر است.

تابع انعکاس را یانه ای شکل ۷. ۲۶ (الف) را در نظر بگیرید. بافت نگار این تابع که در شکل ۷. ۲۶ (ب) دیده می شود، به وضوح دو مدلی است و به سهولت می توان با قرار دادن یک آستانه در دره بافت نگار آن را دو قسمت کرد. ضرب تابع انعکاس شکل ۷. ۲۶ (الف) در تابع روشنایی شکل ۷. ۲۶ (پ)، تصویر (x, y, f) شکل ۷. ۲۶ (ت) را می دهد. شکل ۷. ۲۶ (ث) بافت نگار این تصویر را نمایش

1) Global

2) Local

3) Dynamic



شکل ۲۶.۷ (الف) تابع انعکاس رایانه ای؛ (ب) باند نگار تابع انعکاس؛ (پ) تابع روشنایی رایانه ای؛ (ت) تصویر حاصل از ضرب تابع روشنایی و انعکاس؛ (ث) باند نگار تصویر.

می دهد. توجه کنید که دره اولیه به طور مجازی حذف شد که در نتیجه آن بخش بندی با یک آستانه، کاری غیرممکن گردید. گرچه ما به ندرت با تابع انعکاس تنها سروکار داریم، این نمونه 'ساده نشان' می دهد که طبیعت انعکاس اشیا و زمینه ممکن است چنان باشد که آن دو به سادگی جداپذیر باشند. با این حال تصویر حاصل از روشنایی بد (در این مورد غیریکنواخت) ممکن است بخش بندی بسیار مشکلی داشته باشد.

دلیل خرابی زیاد بافت نگار شکل ۷. ۲۶(ث) را می توان با کمک شرح موجود در بخش ۴. ۴. ۳ دلیل خرابی زیاد بافت نگار شکل ۷. ۲۶(ث) را می توان با کمک شرح موجود در بخش ۴. ۴. ۳ توضیح داد. با گرفتن لگاریتم طبیعی از $f(x, y) = i(x, y) r(x, y)$ مجموع

$$z(x, y) = \ln f(x, y) = \ln i(x, y) + \ln r(x, y) = i'(x, y) + r'(x, y)$$

به دست می آید. از نظریه احتمالات (Papoulis[1965]) داریم که اگر $i'(x, y)$ و $r'(x, y)$ متغیرهای تصادفی مستقل باشند، بافت نگار (x, y) با کانولوشن بافت نگارهای $i'(x, y)$ و $r'(x, y)$ بددست می آید. اگر $i'(x, y)$ ثابت باشد، $i'(x, y)$ نیز ثابت خواهد بود، و بافت نگارش تنها یک شبه ضربه^۱ (شبیه به ضربه ایده آل^۲) خواهد بود. کانولوشن این تابع شبه ضربه با بافت نگار (y, x) ^۳، شکل اصلی این بافت نگار را بدون متغیر باقی می گذارد (شکل ۳. ۱۵ را بینید). اما اگر $i'(x, y)$ بافت نگار وسیع تری (حاصل از روشنایی غیریکنواخت) داشته باشد، فرآیند کانولوشن، بافت نگار (y, x) ^۴ را پخش می کند، که در نتیجه شکل بافت نگار (x, y) کاملاً متفاوت از شکل بافت نگار (y, x) ^۵ خواهد بود. درجه اعوجاج به وسعت بافت نگار (x, y) بستگی دارد، که آن نیز به نایکنواختی تابع روشنایی بستگی دارد.

ما به جای کار مستقیم روی تابع تصویر $f(x, y)$ ، با لگاریتم آن کار کردیم، اما اصل مسأله با کاربرد لگاریتم برای جداسازی مؤلفه های روشنایی و انعکاس به خوبی روشن می شود. این راه امکان می دهد که تشکیل بافت نگار به عنوان یک فرآیند کانولوشن در نظر گرفته شود، بنابراین روشن می سازد که چرا ممکن است یک دره بافت نگار تابع انعکاس، بر اثر روشنایی نامناسب مجازاً حذف شود.

وقتی منبع روشنایی در دسترس است، یک راه حل متداول عملی برای جبران نایکنواختی، تصویر کردن^۳ الگوی روشنایی روی یک سطح انعکاسی سفید ثابت است. این راه حل تصویر $i(x, y) g(x, y) = k$ را می دهد، که ثابتی است که به سطح بستگی دارد و $i(x, y)$ الگوی روشنایی است. پس برای هر تصویر $f(x, y) = i(x, y) r(x, y)$ حاصل از همان تابع روشنایی، تنها با تقسیم $f(x, y)$ بر $g(x, y)$ یک تابع تراز شده $h(x, y) = f(x, y) / g(x, y) = r(x, y) / k$ به دست می آید. بنابراین اگر

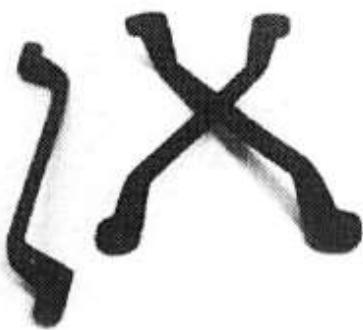
بتوان با استفاده از یک مقدار آستانه T ، (x, y) را بخش بندی کرد، $(x, y)h$ را نیز می توان با یک آستانه 'با مقدار T/K بخش بندی کرد. توجه کنید که این روش تنها در صورتی خوب کار می کند که الگوی روشنایی حاصل از $(x, y)_i$ از یک تصویر به تصویر دیگر تغییری نکند. همان طور که در بخش ۲.۴.۲ تشریح شد، نوعاً ترازسازی $(y, f(x, y))$ با $(y, g(x, y))$ با استفاده از یک واحد پردازشگر حساب و منطق (ALU) اجرا می شود.

۳.۳.۷ آستانه گیری سراسری ساده

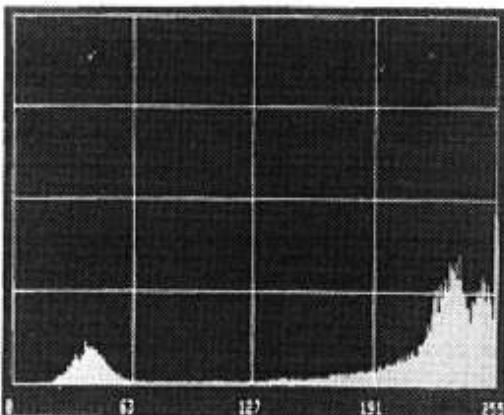
بارجou به شرح بخش ۳.۱ و همان طور که در شکل ۲۵.۷ (الف) دیده می شود، ساده ترین روش آستانه گیری تقسیم بافت نگار تصویر با استفاده از یک آستانه T است. آنگاه با پیمایش پیکسل به پیکسل و برچسب زدن هر پیکسل به عنوان شیئی یا زمینه، بسته به این که سطح خاکستری آن پیکسل بزرگتر یا کوچکتر از T باشد، تصویر بخش بندی می شود. همان طور که قبلاً بیان شد، توفیق این روش به کیفیت تقسیم بافت نگار بستگی کامل دارد.

شکل ۲۷.۷ مثالی از آستانه گیری سراسری را نشان می دهد. شکل ۲۷.۷ (الف) یک تصویر ساده و شکل ۲۷.۷ (ب) بافت نگار آن را نشان می دهد. توجه کنید که محدوده 'سطوح خاکستری' $[0, 255]$ است. شکل ۲۷.۷ (ب) حاصل بخش بندی شکل ۲۷.۷ (الف) با استفاده از آستانه $T = 90$ را نشان می دهد، که در آن با حذف سایه ها و باقی ماندن اشیا به تنها یک، به یک بخش بندی 'تمیز'^۱ دست یافته ایم. اشیای مورد علاقه در این مورد تیره تر از زمینه هستند، بنابراین هر پیکسل با سطح خاکستری کوچکتر یا مساوی T برچسب سیاه (۰) زده شد، و هر پیکسل با سطح خاکستری بزرگتر از T برچسب سفید (۲۵۵) زده شد. (هدف اصلی فقط تولید یک تصویر دودویی است، بنابراین می توان رابطه 'سیاه - سفید را برعکس کرد).

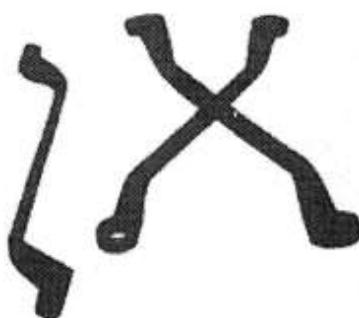
در عمل می توان امید داشت که نوع آستانه گیری سراسری که اخیراً تشریح شد، در محیط های شدیداً کنترل شده موفق باشد. یکی از زمینه هایی که در آن غالباً این کنترل شدید امکان دارد، کاربردهای بازرگانی صنعتی است. از شرح بخش ۳.۲ به خاطرآورید که روشنایی نقش مهمی در تعیین شکل بافت نگار تصویر حاصل دارد.



(الف)



(ب)



(ب)

شکل ۷۰.۷ مثالی از آستانه گیری سراسری: (الف) یک تصویر اولیه و (ب) بافت نگار آن؛ (ب) حاصل بخش بندی با $T=90$. (از [Fu, Gonzalez & Lee 1987].)

۴.۳.۷ آستانه گیری بهینه

فرض کنید که یک تصویر تنها دارای دو ناحیه اصلی روشنایی است. بافت نگار چنین تصویری را می‌توان تقریباً از تابع چگالی احتمال روشنایی ($p(z)$) در نظر گرفت. این تابع چگالی کلی، مجموع یا ترکیب دو چگالی تک مدلی است، که یکی از آنها مربوط به نواحی روشن و دیگری مربوط به نواحی تیره تصویر است. گذشته از این، ضرایب ترکیب متناسب با مساحت های هر روشنایی در عکس هستند. اگر شکل چگالی ها معلوم یا مفروض باشد، تعیین یک آستانه بهینه (بر حسب خطای کمینه) برای بخش بندی تصویر به دوناحیه روشنی ممکن است.

۵۳۵ آستانه گیری ۳. ۷

فرض کنید که یک تصویر دو مقدار دارد که با نویز گوسی جمع شونده ترکیب شده است. تابع چگالی احتمال ترکیبی عبارتست از

$$p(z) = P_1 p_1(z) + P_2 p_2(z) \quad (3-3. 7)$$

که در مورد گوسی عبارتست از:

$$p_i(z) = \frac{P_i}{\sqrt{2\pi} \sigma_i} \exp\left[-\frac{(z - \mu_i)^2}{2\sigma_i^2}\right] \quad (4-3. 7)$$

که μ_1 و μ_2 مقادیر میانگین دو سطح روشنایی هستند، σ_1 و σ_2 انحراف معیارها حول میانگین‌ها هستند، و P_1 و P_2 احتمالات پیشین دو سطح هستند. قید

$$P_1 + P_2 = 1 \quad (5-3. 7)$$

باید برآورده شود، بنابراین چگالی مرکب پنج عامل مجھول دارد. اگر تمام عوامل معلوم باشند، به سادگی آستانه بینیه تعیین می‌شود.

فرض کنید که نواحی تیره مربوط به زمینه و نواحی روشن مربوط به اشیا هستند. در این حالت $\mu_1 < \mu_2$ ، و یک آستانه T را می‌توان تعریف کرد چنان‌که تمام پیکسل‌های با سطح خاکستری کمتر از T نقاط زمینه و تمام پیکسل‌های با سطح بزرگتر از T ، نقاط شیئی در نظر گرفته شوند. احتمال گروه‌بندی (خطاطار) یک نقطه شیئی به عنوان یک نقطه زمینه عبارتست از:

$$E_i(T) = \int_{-\infty}^T p_i(z) dz. \quad (6-3. 7)$$

مشابهًاً احتمال گروه‌بندی یک نقطه زمینه به عنوان یک نقطه شیئی عبارتست از:

$$E_o(T) = \int_T^{\infty} p_o(z) dz. \quad (7-3. 7)$$

بنابراین احتمال کلی خطاطار عبارتست از:

$$E(T) = P_2 E_o(T) + P_1 E_i(T). \quad (8-3. 7)$$

برای پیدا کردن مقدار آستانه‌ای که برای آن این خطاطار کمینه است، لازم است که از $E(T)$ نسبت به T مشتق

پگیریم (با استفاده از قاعده لايب نیتز^۱) و حاصل را برابر ۰ قرار دهیم. بنابراین:

$$P_1 p_1(T) = P_2 p_2(T). \quad (9-3. 7)$$

با کاربرد این نتیجه در چگالی گوسی، گرفتن لگاریتم و ساده سازی، معادله درجه دوی

$$AT^2 + BT + C = 0 \quad (10-3. 7)$$

به دست می آید که:

$$\begin{aligned} A &= \sigma_1^2 - \sigma_2^2 \\ B &= 2(\mu_1 \sigma_2^2 - \mu_2 \sigma_1^2) \\ C &= \sigma_1^2 \mu_2^2 - \sigma_2^2 \mu_1^2 + 2\sigma_1^2 \sigma_2^2 \ln(\sigma_2 P_1 / \sigma_1 P_2). \end{aligned} \quad (11-3. 7)$$

امکان وجود در جواب بیان می کند که ممکن است برای به دست آوردن جواب بهینه دو آستانه مورد نیاز باشد.

اگر واریانسها^۲ برابر باشند، $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ، تنها یک آستانه کافی است:

$$T = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_1 - \mu_2} \ln\left(\frac{P_2}{P_1}\right). \quad (12-3. 7)$$

اگر احتمالات پیشین برابر باشند، $P_1 = P_2$ ، آستانه' بهینه متوسط میانگین هاست. همان حالت برای $\sigma = 0$ نیز صدق می کند. به طور مشابه می توان برای سایر چگالی های تک مُدی با شکل معلوم، نظریه چگالی های رالی^۳ و لوگی-نرمال^۴، آستانه' بهینه را تعیین کرد.

از روش کمینه' میانگین مریع خطای تخمین عوامل تصویر از بافت نگار استفاده کرد. مثلاً خطای میانگین مریع بین چگالی مرکب $(z)p$ و بافت نگار آزمایشی عبارتست از:

$$e_{mn} = \frac{1}{n} \sum_{i=1}^n [p(z_i) - h(z_i)]^2 \quad (13-3. 7)$$

که بافت نگار، n نقطه ای فرض شده است.

به طور کلی تعیین تحلیلی عواملی که این خطای میانگین مریع را کمینه می سازد، موضوع ساده ای

1) Liebnitz's rule

2) Variance

3) Raleigh

4) Log-normal

نیست. حتی در مورد گوسی، عمل ساده برابر ۰ قرار دادن مشتقات جزئی، به یک دستگاه معادلات غیرجبری همزمان^۱ منجر می شود که معمولاً تنها با روال های عددی حل شدنی هستند. چون گرادیان به سادگی محاسبه می شود، می توان از روش گرادیان مزدوج^۲ یا روش نیوتون برای حل معادلات غیرخطی همزمان استفاده کرد. در هر یک از این روش های تکراری مقادیر شروع باید معین باشند. فرض این که احتمالات پیشین برابر ند، ممکن است کافی باشد. می توان با آشکارسازی ۳دهای بافت نگار یا تنها با تقسیم بافت نگار به دو قسمت در طرفین میانگین، میانگین ها و واریانس های آن دو قسمت را به عنوان مقادیر شروع میانگین و واریانس تعیین کرد.

مثال: شرح زیر درباره 'راه ابداعی Chow, Kaneko[1972] برای نمایش مرزهای بطن چپ در نمودار رگهای خونی قلب (یعنی عکس های اشعه ایکس قلب در حالی که یک مایع رنگی تزریق شده است) و چگونگی انتخاب آستانه' بهینه را نمایش می دهد.

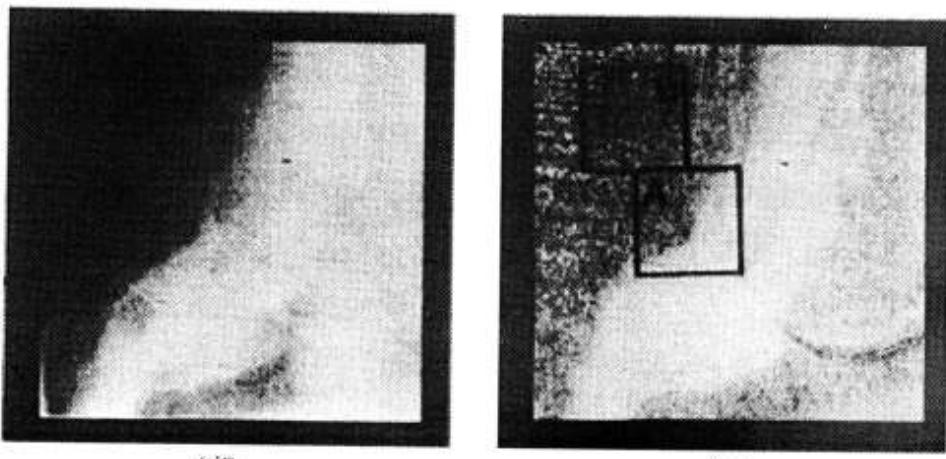
قبل از آستانه گیری، ابتدا تصاویر (با ابعاد 256×256 پیکسل) تحت مراحل پیش پردازش زیر قرار گرفتند: (۱) از مقدار هر پیکسل لگاریتم گرفته شد تا اثرات نمایی حاصل از جذب رادیواکتیو بر طرف شود، (۲) دو تصویر به دست آمده قبل و بعد از تزریق ماده رنگی از هم تفریق شدند تا این که ستون مهره های موجود در آن دو حذف بشود (بخش ۴. ۲. ۳. را بینید)، و (۳) از چند نمودار رگ های خونی متوسط گرفته شد تا نیز حذف شود (بخش ۴. ۲. ۴. را بینید). شکل ۷. ۲۸ نمودار رگ های خونی قلب را در قبل و بعد از پیش پردازش نمایش می دهد (در ادامه شرحی درباره 'ناحی A و B خواهد آمد). به منظور محاسبه آستانه های بهینه، هر تصویر پیش پردازش شده با قرار دادن یک شبکه 7×7 با 50% درصد همپوشانی بر روی آن به ۴۹ ناحیه تقسیم شد. هر یک از ۴۹ ناحیه حاصل دارای 64×64 پیکسل بود. اشکال ۷. ۲۹ (الف) و (ب) بافت نگار نواحی A و B از شکل ۷. ۲۸ (ب) هستند. توجه کنید که به وضوح بافت نگار ناحیه A دو مدل است که نشانه' وجود یک مرز در آن ناحیه است. با این حال بافت نگار ناحیه B تک مدلی است که نشانه' عدم وجود دوناچیه' به طور مشخص مجزا است.

پس از این که تمام ۴۹ بافت نگار محاسبه شدند، آزمونی دو مدل بودن انجام شد تا بافت نگار های تک مدل را حذف نماید. آنگاه بافت نگار های باقیمانده با استفاده از روش تپه نورده گرادیان مزدوج^۳ بر

1) A set of simultaneous transcendental equations

2) Conjugate gradient method

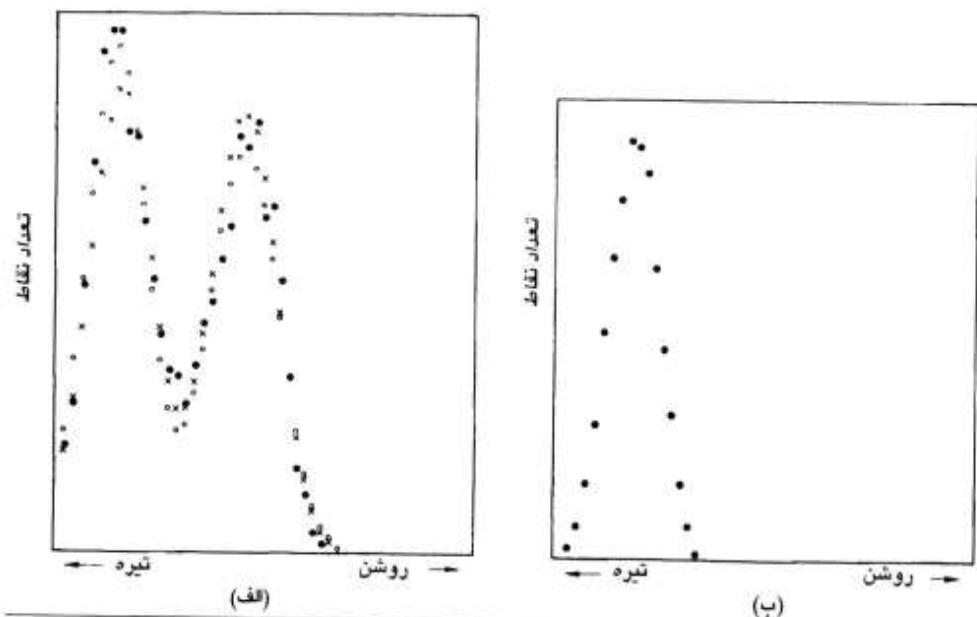
3) Conjugate gradient hill-climbing method



(الف)

(ب)

شکل ۲۸. ۷ یک تصویر از رگ های خونی قلب قبل و بعد از پردازش. (از [Chow & Kaneko[1972] منحنی های چگالی گوسی دو مدی (معادله' ۷. ۴-۳ را بینید) انطباق داده شدند تاتابع خطای ارائه شده در معادله' ۷. ۳-۱۳) کمینه شود. «Xها و Oهای شکل ۷. ۲۹ (الف)» دو انطباق بر بافت نگار نمایش داده شده با نقاط توپیر هستند. در نهایت با استفاده از معادلات (۷. ۳-۱۰) و (۷. ۳-۱۱) آستانه های بهینه بدست آمدند.

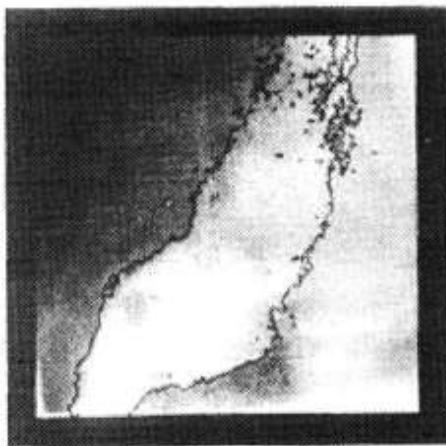


(الف)

(ب)

شکل ۲۹. ۷ بافت نگارهای (نقاط سیاه) نواحی A و B از شکل ۷. ۲۸ (ب).

(Chow, Kaneko[1972]) از



شکل ۷.۳۰ نمودار رگ های خونی قلب که مرزهای گذاشته شده روی آن را نشان می دهد.
(از Chow & Kaneko [1972])

در این مرحله از فرآیند تنها برای تواحی دارای بافت نگار دو مدل مقدار آستانه بدست آمد. آستانه های سایر تواحی با درونیابی این آستانه ها بدست آمدند. آنگاه دومین درونیابی نقطه به نقطه و با استفاده از مقادیر آستانه همسایه انجام شد تا این که در انتهای روال، به هر نقطه تصویر یک آستانه متنسب شده بود. در نهایت با استفاده از قاعده 'ا

$$f(x, y) = \begin{cases} 1 & f(x, y) \geq T_v, \\ 0 & \text{در غیر این صورت} \end{cases} \quad \text{اگر}$$

به هر پیکسل یک مقدار دودویی متنسب شد که T_v آستانه متنسب شده به پیکسل (x, y) تصویر است. [توجه کنید که این آستانه ها پویا هستند، زیرا به مختصات مکانی (x, y) بستگی دارند.] آنگاه با گرفتن گرادیان از تصویر دودویی حاصل، مرزها بدست آمدند. شکل ۷.۳۰ مرزهای گذاشته شده بر روی تصویر اصلی را نشان می دهد.

□

۵.۳.۷ انتخاب آستانه براساس مشخصات مرز

یکی از مهمترین جنبه های انتخاب آستانه، توانایی شناسایی مطمئن قله مدها در بافت نگار داده شده است. توانایی انتخاب خودکار آستانه برای وضعیت هایی که ممکن است مشخصات تصویر روی محدوده وسیعی از توزیع های شدت تغییر کند، اهمیت ویژه ای دارد. براساس شرح موجود در بخش های ۷.۳.۷-

۷. ۳. ۴. به خوبی آشکار است که وقتی قله های بافت نگار بلند، باریک و متقارن باشند و با دره های عمیق از هم جدا شوند، شانس انتخاب آستانه^۱ خوب بسیار افزایش می یابد.

یک راه برای بهبود شکل بافت نگار این است که تنها پیکسل هایی را که روی یا نزدیک مرز بین اشیا و زمینه قرار دارند، در نظر بگیریم. یک بهبود سریع و واضح این است که بافت نگارها کمتر به ابعاد نسبی اشیا و زمینه وابسته خواهد بود. برای نمونه، بافت نگار شدت تصویری که مشکل از یک ناحیه زمینه بزرگ تقریباً و ثابت و یک شیئی کوچک است، به علت تعداد زیاد پیکسل های زمینه، تحت تأثیر قله شود، بافت نگار حاصل قله هایی با ارتفاع تقریباً برابر خواهد داشت. به علاوه احتمال این که یکی از پیکسل های داده شده روی شیئی قرار داشته باشد، تقریباً برابر احتمال آن است که روی زمینه قرار داشته باشد، بنابراین تقارن قله های بافت نگار بهبود می یابد. در نهایت استفاده از پیکسل هایی که یک یا چند معیار ساده^۲ مبتنی بر عملکردهای گرادیان یا لاپلاسین را بآورده می سازند، باعث می شود که دره^۳ بین قله های بافت نگار عمیقتر شود.

مشکل اصلی روش اخیر فرض ضمنی معلوم بودن مرز بین اشیا و زمینه است. آشکار است که این اطلاعات در حین بخش بندی وجود ندارند، زیرا پیدا کردن مرز بین اشیا و زمینه، دقیقاً همان هدف نهایی بخش بندی است. با این حال با توجه به مطالب بخش ۷. ۱. ۳.، می توان با محاسبه^۴ گرادیان تصویر در هر پیکسل، نشانه ای از روی لبه بودن یا نبودن آن پیکسل بدست آورد. به علاوه استفاده از لاپلاسین می تواند اطلاعاتی راجع به این که آیا پیکسل موردنظر روی طرف تاریک (زمینه) یا روشن (شیئی) لبه قرار دارد، به مابدله. مقدار متوسط لاپلاسین در محل گذر لبه ۰ است (شکل ۷. ۴. را ببینید)، بنابراین در عمل می توان انتظار داشت که دره های بافت نگار تشکیل شده با پیکسل های انتخاب شده با معیار گرادیان / لاپلاسین کم و دور از هم^۱ باشند. این خاصیت، دره های عمیق و بسیار مطلوب را که قبل از شرح داده شد، ایجاد می کند.

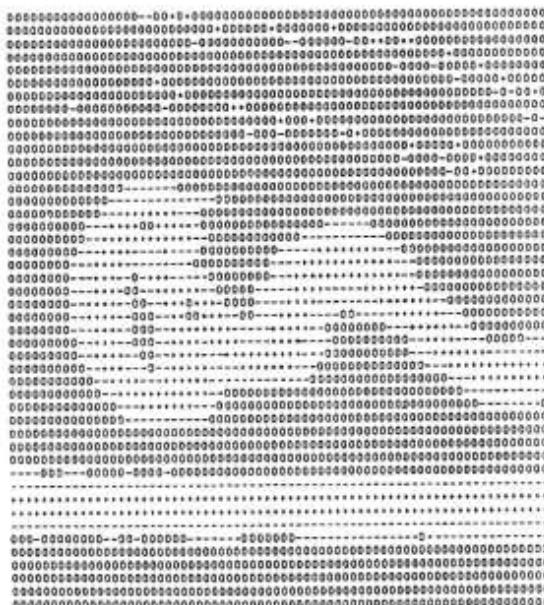
گرادیان ∇f در هر نقطه (y, x) تصویر با معادله^۵ $(7.1-4)$ یا $(7.1-5)$ به دست می آید. به طور مشابه لاپلاسین $\nabla^2 f$ با معادله^۶ $(7.1-10)$ تعیین می شود. با استفاده از این دو کمیت می توان به صورت

زیر یک تصویر سه سطحی تشکیل داد:

¹⁾ *Sparsely populated*

$$s(x, y) = \begin{cases} 0 & \nabla f < T \\ + & \nabla f \geq T \quad \nabla^2 f \geq 0 \\ - & \nabla f \geq T \quad \nabla^2 f < 0 \end{cases} \quad (14-3. v)$$

که نمادهای $+$ ، $-$ نمایانگر سطح خاکستری دلخواه اما متفاوت هستند، T یک آستانه است، و گرادیان و لاپلاسین نیز در هر نقطه (x, y) تصویر محاسبه می شوند. برای یک شیئ تیره روی یک زمینه روشن، و با مراجعه به شکل ۷. ۴. (ب)، با استفاده از معادله ۷. ۲-۱۴) تصویر $s(x, y)$ تولید می شود که در آن تمام پیکسل هایی که روی لبه نیستند (در آنها ∇f کوچکتر از T است) برچسب "۰" تمام، پیکسل های روی طرف تیره لبه برچسب "+" و تمام پیکسل های روی طرف روشن لبه برچسب "-" زده می شوند. برای یک شیئ روشن روی زمینه تیره، نمادهای $+$ و $-$ موجود در معادله ۷. ۳-۱۴) تعویض می شوند. شکل ۷. ۳۱. ۷ برچسب زنی حاصل از معادله ۷. ۳-۱۴) را برای تصویر یک علامت تیره و زیرخط دار^۱ که روی یک زمینه روشن نوشته شده بوده است، نمایش می دهد.



شکل ۷. ۳۱. ۷ تصویر یک علامت که با استفاده از معادله ۷. ۳-۱۴) رمز شده است.
(. White & . Rohrerf 1983/ از)

1) Underlined

اطلاعات حاصل از این روال را می‌توان برای تولید تصویر دودویی بخش بندی شده به کاربرد، که در این تصویر "۱" ها متناظر با اشیای موردنظر علاقه و "۰" ها متناظر با زمینه هستند. گذر از زمینه روشن به یک شیء تیره (در طول یک خط پیمایش افقی یا عمودی) باید با وقوع یک "-" در (x, y) که به دنبال آن یک "+" قرار دارد، مشخص شود. داخل شیءی مشکل از پیکسل های با برد چسب "۰" یا "+" است. در نهایت گذر از شیءی به زمینه با وقوع یک "+" که به دنبال آن یک "-" قرار دارد، مشخص می‌شود. بنابراین یک خط پیمایش افقی یا عمودی که دارای بخشی از یک شیء است، ساختار زیر را دارد:

$$(\dots) (- , +) (0 \quad \quad +) (+ , -) (\dots)$$

که (...). نمایانگر هر ترکیب دیگری از "+", "-", "۰" ها می‌باشد. داخلی ترین پرانتزها حاوی نقاط شیئی هستند و برد چسب "۱" دارند. سایر پیکسل های همان خط پیمایش برد چسب "۰" دارند، با این استثنای هر دنباله دیگر از "(۰)" یا "(+)" نیز محصور در "(+,-)" و "(-,+)" است.

مثال: شکل ۳۲. ۷. (الف) تصویری از یک چک بانکی منظره دار معمولی را نشان می‌دهد. شکل ۳۳. ۷. بافت نگار آن را به صورت تابعی از مقادیر گرادیان بزرگتر از ۵ نمایش می‌دهد. توجه کنید که این بافت نگار خواص مشروح در قبل را داراست. یعنی دو ۹۹٪ غالب دارد که متقارن و تقریباً هم ارتفاع هستند و با یک درجه مشخص از هم جدا شده‌اند. در نهایت شکل ۳۲. ۷. (ب) تصویر بخش بندی شده حاصل از معادله (۷. ۳-۱۴) با استفاده از آستانه T که برای یا نزدیک نقطه وسط درجه است، را نشان می‌دهد. این تصویر با استفاده از تحلیل دنباله‌ای که اخیراً تشریح شد، دودویی گردید. توجه کنید که گرچه T یک آستانه ثابت است، کاربرد آن محلی بود زیرا تصویر بخش بندی شده با استفاده از معادله (۷. ۳-۱۴) که مقادیر محلی گرادیان و لاپلاسین را به کار می‌برد، تولید شد.

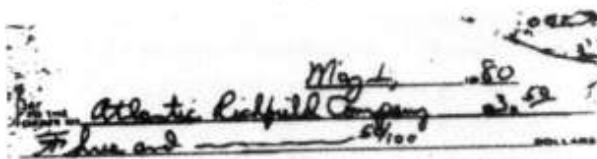
۶.۳.۷ آستانه‌های مبتنی بر چند متغیر

تاکنون به آستانه گیری از فقط یک متغیر توجه کرده‌ایم. در بعضی موارد ممکن است که حسگر از بیش از یک متغیر برای توصیف یک پیکسل تصویر، استفاده کند. یک مثال خوب تصویر برداری رنگی است که در آن از مؤلفه‌های قرمز (R)، سبز (G) و آبی (B) برای ایجاد یک تصویر رنگی مرکب استفاده می‌شود

۳۰.۷ آستانه گیری ۵۴۳



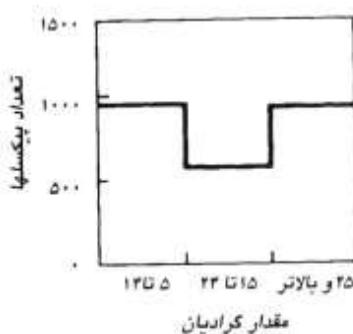
(الف)



(ب)

شکل ۷.۳۲۰ (الف) یک تصویر اولیه؛ (ب) حاصل بخش بندی آن. (از [White & Rohrer 1983])

(بخش ۷.۶.۲ را بینید). در این مورد، هر پیکسل با سه مقدار توصیف می شود و ایجاد «بافت نگار» سه بعدی ممکن می گردد. روای اصلی مانند حالت نک متغیره است. مثلاً برای سه تصویر $16 \times 16 \times 16$ سطحی منتظر با مؤلفه های RGB، یک شبکه^۱ ($16 \times 16 \times 16$) تشکیل می شود و در هر سلول مکعب تعداد پیکسل هایی که شدت مؤلفه های RGB شان برابر مختصات آن سلول است، ثبت می شود. آنگاه می توان هر درایه مکعب را بر تعداد کل پیکسل های تصویر تقسیم کرد تا بافت نگار سه بعدی تراز شده بدست آید.



شکل ۷.۳۳۰ (الف) بافت نگار پیکسلهایی که گرادیان بزرگتر از ۵ دارند. (از [White & Rohrer 1983])



(الف)



(ب)



(ب)

شکل ۷. ۳۴. بخش بندی با روش بافت نگار چند متغیره

در اینجا عمل آستانه گیری، پیدا کردن دسته^۱ نقاط در فضای سه بعدی است. به عنوان مثال فرض کنید که K دسته مهم نقاط در بافت نگار پیدا شود. می توان با انتساب یک شدت ثابت به پیکسل هایی که مؤلفه های RGB شان به یک دسته نزدیکتر است و انتساب شدت دیگری به سایر پیکسل ها، تصویر را بخش بندی کرد. این مفهوم به آسانی به مؤلفه های بیشتر و یقیناً به دسته های بیشتر گسترش پذیر است. مشکل اصلی این است که وقتی تعداد متغیرها افزایش می یابد، جستجوی دسته ها نیز پیچیده تر می شود. روش های جستجوی دسته را در منابع دیگر، مثلاً در کتاب [Tou & Gonzalez 1974] می توان پیدا کرد.

مثال: تصویر شکل ۷. ۳۴ (الف) یک تصویر تک رنگ از یک عکس رنگی است. تصویر رنگی اولیه

1) Cluster

۴.۷ بخش بندی ناحیه گرا

۵۲۵

منتقل از سه تصویر RGB ۱۶ سطحی بود. روسربی، قرمز پررنگ بود و موها و رنگ های صورت روشن بودند و مشخصات طیفی متفاوتی با پنجه و سایر ویژگی های زمینه داشتند.

شکل ۷ (۳۴. ۷) حاصل آستانه گیری حول یکی از دسته های بافت نگار است. توجه کنید که پنجه، که در تصویر تک رنگ شدتی نزدیک به موها دارد، به علت استفاده از مشخصات چند طیفی در جداسازی این دو ناحیه، در تصویر بخش بندی شده ظاهر نمی شود. شکل ۷ . (۳۴. ۷) با آستانه گیری حول یک دسته نزدیک به محور قرمز به دست آمد. در این مورد تنها روسربی و قسمتی از گل (که آنهم قرمز بود) در حاصل بخش بندی ظاهر شد. آستانه مورد استفاده برای بدست آوردن هر دو نتیجه فاصله ای به اندازه یک سلول بود. بنابراین هر پیکسلی که مؤلفه هایش خارج از سلولی بودند که مرکز دسته موردنظر را احاطه می کرد، به عنوان زمینه (سیاه) گروه بندی شد. پیکسل هایی که مؤلفه هایشان آنها را درون سلول قرار داد، با سفید رمز شدند.



بخش بندی رنگی را می توان بر مبنای هر یک از مدل های رنگی مشروح در بخش ۴. ۶. ۲ بنا نهاد. برای نمونه اصل رنگ^۱ و اشباع در بسیاری از کاربردهای تصویربرداری برای بازرگانی خودکار، خواص مهمی هستند. این خواص اهمیت ویژه ای در شبیه سازی کارهای انسان نظیر وارسی میوه ها از جهت رسیده بودن یا وارسی کالاهای تولید شده دارند. همان طور که در بخش ۴. ۶. ۲ ذکر شد، مدل HSI برای این نوع کاربردها مطلوب است زیرا ابسطه نزدیکی با روشی که مردم، احساس رنگ را توصیف می کنند، دارد. بخش بندی با استفاده از مؤلفه های اصل رنگ و اشباع یک سیگنال رنگی نیز جذابیت ویژه ای دارد، زیرا دسته های داده دو بعدی را به کار می گیرد که نسبت به، مثلاً، دسته های سه بعدی مورد نیاز برای بخش بندی RGB، تحلیل ساده تری دارند.

۴.۷ بخش بندی ناحیه گرا

هدف بخش بندی، تقسیم تصویر به چندین ناحیه است. در بخش های ۱. ۷ و ۲. ۷ با پیدا کردن مرزهای بین نواحی بر اساس ناپیوستگی های شدت، به این مسئله پرداختیم، در حالی که در بخش ۳. ۷ بخش بندی از طریق آستانه های مبتنی بر توزیع خواص پیکسلی نظیر شدت یا رنگ انجام شد. در این بخش فنون بخش بندی را که مبتنی بر پیدا کردن مستقیم نواحی هستند، شرح می دهیم.

۱.۴.۷ بیان پایه ای

فرض کنید R بیانگر کل ناحیه تصویر باشد. می توان بخش بندی را به عنوان فرآیندی در نظر گرفت که

را به n زیرناحیه، R_1, R_2, \dots, R_n قسمت می کند چنانکه

$$\bigcup_{i=1}^n R_i = R \quad (\text{الف})$$

(ب) R_i یک ناحیه متصل است، $i = 1, 2, \dots, n$.

(پ) برای تمام i و j که $i \neq j$ داشته باشند $R_i \cap R_j = \emptyset$.

(ت) به ازای i, n ، $P(R_i) = T^i$ ، $i = 1, 2, \dots, n$ ، و

(ث) به ازای $i \neq j$ $P(R_i \cup R_j) = F^i$.

که $P(R_i)$ یک گزاره منطقی روی نقاط مجموعه R_i است و ϕ مجموعه 'تهی' است.

شرط (الف) بیان می کند که بخش بندی باید کامل باشد؛ یعنی هر پیکسل باید در یک ناحیه قرار گیرد.

شرط دوم الزام می نماید که نقاط یک ناحیه باید متصل باشند (بخش ۲.۴.۲ را در مورد اتصال ببینید).

شرط (پ) می گوید که نواحی باید جدا از هم باشند. شرط (ت) به خواصی که پیکسل های درون هر یک از نواحی بخش بندی شده باید برآورده کنند، می پردازد - مثلاً $P(R_i) = T$ است اگر تمام پیکسل های ناحیه R_i شدت یکسانی داشته باشند. در نهایت شرط (ث) می گوید که نواحی R_i و R_j در مفهوم گزاره

P با هم متفاوت هستند.

۲.۴.۷ رشد ناحیه با پیوستن پیکسل

همان طور که از نامش پیداست، رشد ناحیه R روالی است که پیکسل های زیرناحیه ها را در نواحی بزرگتر گروه بندی می کند. ساده ترین این راهها پیوستن پیکسل است که با یک مجموعه از نقاط «بذر»

آغاز می شود. آنگاه با پیوستن پیکسل های همسایه هر بذر به آن براساس مشابهت خواص (نظیر سطح خاکستری، بافت، رنگ) نواحی بزرگ و بزرگتری ایجاد می شود. برای نمایش این روال شکل ۲.۵.

(الف) را در نظر می گیریم، که در آن اعداد درون سلولها بیانگر مقادیر سطوح خاکستری هستند. فرض کنید که نقاط با مختصات $(2, 3)$ و $(3, 4)$ به عنوان بذر استفاده شوند. استفاده از دو نقطه شروع، نوعی

بخش بندی را نتیجه می دهد که حداقل دارای دو ناحیه است: R_1 مربوط به بذر $(2, 3)$ و R_2 مربوط به

1) True: درست

2) False: غلط

3) Region growing

4) Pixel aggregation

۵۳۷ بخش بندی تابعه گرا ۴. ۷

	۱	۲	۳	۴	۵
۱	۰	۰	۵	۶	۷
۲	۱	۱	۵	۸	۷
۳	۰	۱	۶	۷	۷
۴	۲	۰	۷	۶	۶
۵	۰	۱	۵	۶	۵

a	a	b	b	b
a	a	b	b	b
a	a	b	b	b
a	a	b	b	b
a	a	b	b	b

a	a	a	a	a
a	a	a	a	a
a	a	a	a	a
a	a	a	a	a
a	a	a	a	a

شکل ۷ . ۳۵ مثالی از رشد تابعه با استفاده از نقاط شروع معולם: (الف) آرایه تصویری اولیه؛ (ب) حاصل بخش بندی با استفاده از تفاضل مطلق کمتر از ۳ بین سطوح شدت؛ (پ) حاصل بخش بندی با استفاده از تفاضل مطلق کمتر از ۸ بین سطوح شدت.

بذر (۴ و ۳). خاصیت P که باید برای انضمام یک پیکسل به یکی از دو ناحیه استفاده شود، این است که تفاضل مطلق بین سطوح خاکستری آن نقطه و بذر، کوچکتر از آستانه T باشد. هر پیکسلی که این خاصیت را برای هر دو بذر برآورده کند، (به دلخواه) به ناحیه R_1 منتب می شود. شکل ۷.۳۵(ب) حاصل استفاده از $T = 3$ را نشان می دهد. در این مورد، بخش بندی دارای دو ناحیه است، که نقاط درون R_1 با a و نقاط درون R_2 با b نشان داده می شوند. توجه کنید که انتخاب هر نقطه شروع دلخواه دیگر در هر یک از دو ناحیه 'حاصل، نیز همان نتیجه را خواهد داد. با این حال همان طور که شکل ۷.۳۵(ب) نشان می دهد، انتخاب $T = 8$ تنها یک ناحیه را نتیجه خواهد داد.

نمایش قبلی گرچه طبیعت ساده ای دارد، چند مشکل اساسی در رشد ناحیه را آشکار می کند. دو مشکل آشکار انتخاب بذرهای اولیه که نمایندگان خوبی برای نواحی مورد علاقه باشند و نیز انتخاب خواص مناسب برای الحق نقاط به نواحی مختلف در حین فرآیند رشد می باشد. غالباً می توان براساس طبیعت مسئله، مجموعه ای از یک یا چند نقطه' شروع را انتخاب کرد. مثلاً در کاربردهای نظامی تصویربرداری مادون قرمز اهداف مورد علاقه معمولاً داغ تر از زمینه هستند (و بنابراین روش نظری ظاهر می شوند). پس انتخاب روش ترین پیکسل ها، نقطه' شروع طبیعی برای الگوریتم رشد ناحیه است. وقتی اطلاعات پیشین موجود نباشد، روال کار این است که در هر پیکسل، مجموعه خواص پیکسانی را محاسبه کنیم که بعداً از آنها برای انتساب پیکسل ها به نواحی در حین فرآیند رشد استفاده می شود. اگر حاصل این محاسبات دسته هایی از مقادیر را نشان دهد، می توان از پیکسل هایی که خواصشان یک مرکز نقل خواص این دسته هاست، به عنوان بذر استفاده کرد. برای نمونه در شکل قبل، هیستوگرام سطح خاکستری نشان خواهد داد که نقاط با شدت های ۱ و ۷ غالبترین هستند.

انتخاب معیار شباخت نه تنها به مسئله تحت بررسی بلکه به نوع داده های تصویری موجود نیز بستگی دارد. مثلاً تحلیل تصاویر ماهواره ای گرفته شده از سطح زمین بستگی زیادی به کاربردنگ دارد. تحلیل آنها با استفاده از تصاویر تک رنگ بسیار مشکل تر خواهد بود. متأسفانه در پردازش تصویر در دسترس بودن داده های تصویری چند طیفی و سایر داده های تکمیلی همیشگی نیست، بلکه یک استثناست. نوع تحلیل ناحیه باید با مجموعه ای از توصیف گرها که مبنی بر شدت و خواص مکانی (نظیر گشتاورها یا بافت) فقط یک منبع تصویری هستند، انجام شود. توصیف گرها مفید برای مشخص کردن ناحیه را در بخش ۳.۸. شرح می دهیم.

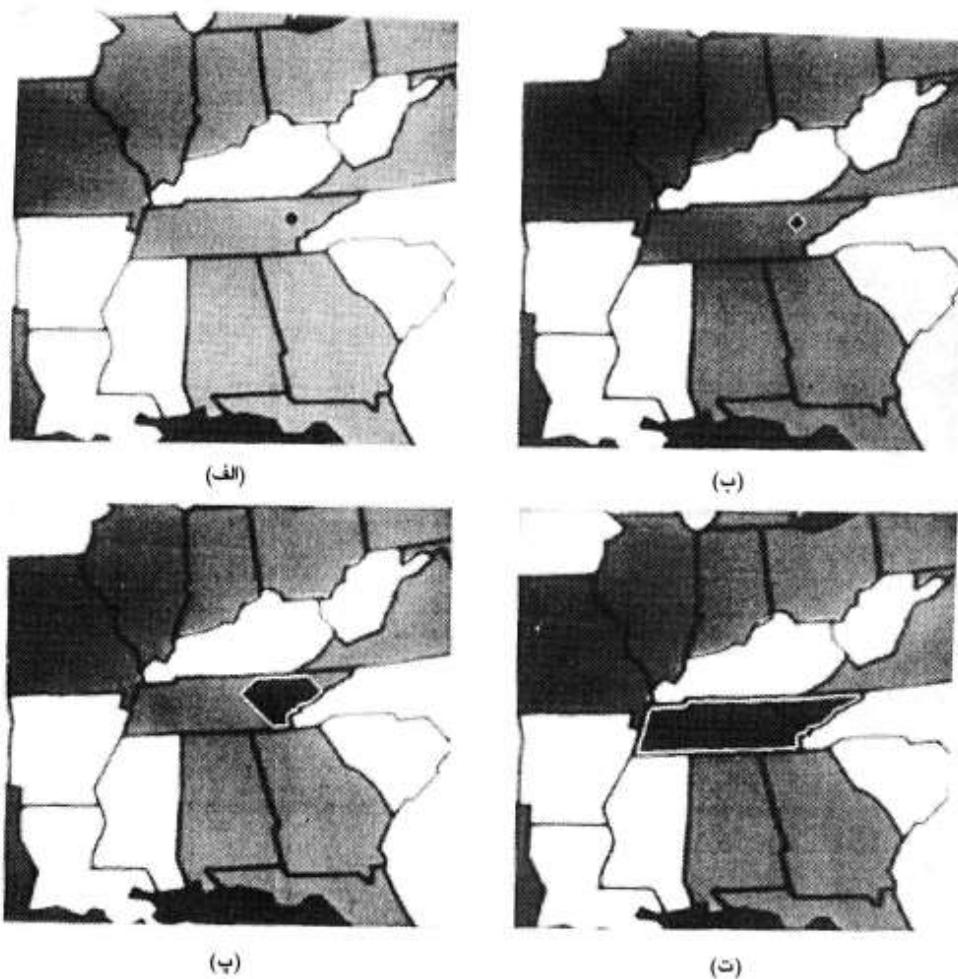
اگر از اطلاعات اتصال یا مجاورت در فرآیند رشد ناحیه استفاده نشود، استفاده از توصیف گرها به تنهایی ممکن است نتایج گمراه کننده ای بدهد. مثلاً یک آرایش تصادفی پیکسل ها را که دارای فقط سه مقدار متفاوت شدت باشد، تصور کنید. گروه بندی پیکسل های با شدت یکسان برای تشکیل یک ناحیه بدون توجه به اتصال، بخش بندی را نتیجه می دهد که در این مبحث بی معناست.

مشکل دیگر در رشد ناحیه توصیف یک قاعده توقف است. وقتی هیچ پیکسل دیگری معیارهای پیوستن به یک ناحیه را برآورده نکند، باید رشد آن ناحیه متوقف شود. قبل اذکر کردیم که معیارهایی نظری شدت، بافت، و رنگ ذاتاً محلی اند و «سابقه^۱» رشد ناحیه را در نظر نمی گیرند. معیارهای دیگری که قدرت الگوریتم رشد ناحیه را افزایش می دهند، از مفهوم اندازه ناحیه، شباهت بین یک پیکسل انتخابی و پیکسل هایی که قبل از ناحیه پیوسته اند (نظری مقایسه شدت پیکسل انتخابی و شدت متوسط ناحیه)، و نیز شکل ناحیه بهره می گیرند. استفاده از این نوع توصیف گرها مبتنی بر این فرض است که لااقل بخشی از یک مدل نتایج مورد انتظار، موجود است.

مثال: شکل ۷. (الف) بخشی از یک نقشه را که دارای یک نقطه 'بذر' تکی (یک نقطه 'سیاه') است، نشان می دهد. معیارهای مورد استفاده برای رشد ناحیه عبارتند از این که (۱) تفاضل مطلق بین سطح خاکستری بذر و نقطه انتخابی از ده درصد تفاضل بین سطوح خاکستری کمینه و بیشینه تصویر (در این مورد ۲۵۵ تجاوز نکند، و (۲) هر پیکسل افزوده شده به ناحیه لااقل به یک پیکسل قبلی ناحیه اتصال هشتگانه داشته باشد.

شکل ۷. (ب) یک ناحیه را در مراحل اولیه رشد نشان می دهد. ابتدا پیکسل های با فاصله D_4 یکسان از نقطه 'بذر'، در نظر گرفته شدند. با افزایش مقدار این فاصله برای گسترش رشد، ناحیه ای ذوزنقه ای شکل تشکیل شد (بخش ۴. ۵. ۲ را ببینید). شکل ۷. (ب) این ناحیه را در یک مرحله میانی رشد نشان می دهد. توجه کنید که چگونه شکل ذوزنقه ای براثر برخورد مرز ناحیه با پیکسل هایی که معیار سطح خاکستری را برآورده نکردن، اعوجاج یافت. شکل ۷. (ت) ناحیه کامل رشد یافته با این روش را نشان می دهد. گرچه سایر پیکسل های نواحی مجاور معیار سطح خاکستری را برآورده کردن، رشد متوقف شد زیرا این پیکسل ها به علت جدایی شان که ناشی از وجود مرز تیره حول ناحیه رشد یافته بود، معیار اتصال را برآورده نکردند.





شکل ۷.۳۶. (الف) تصویر اولیه که نقطه بذر را نشان می دهد؛ (ب) مراحل ابتدایی رشد ناحیه؛ (پ) مراحل میانی رشد ناحیه؛ (ت) ناحیه نهایی.

۳.۴.۷ تقسیم و ادغام ناحیه

روالی که اخیراً تشریح شد، نواحی را از یک مجموعه نقاط بذر رشد می دهد. روال جایگزین، تقسیم ابتدایی تصویر به یک مجموعه نواحی جدا از هم دلخواه و سپس ادغام و یا تقسیم آن نواحی در تلاش برای برآورده شدن شرایط مذکور در بخش ۷.۴.۱ است. یک الگوریتم تقسیم و ادغام برای برآوردن این قیود به صورت زیراست.

فرض کنید R نمایانگر کل ناحیه تصویر باشد و طبق شرح بخش ۷.۴.۱ گزاره 'P' را انتخاب کنید.

برای یک تصویر مربعی، یک روش برای بخش بندی R تقسیم متواالی آن به نواحی مربعی کوچکتر و کوچکتر است تا این که برای هر ناحیه R_i , $P(R_i) = T$. یعنی اگر $P(R) = F$ ، تصویر را به چهار ربع تقسیم کنید. برای هر ربع که P غلط باشد، آن ربع را به چهار ربع کوچکتر تقسیم کنید و الی آخر. همان طور که در شکل ۷. ۳۷ دیده می شود، این عمل به شکل درخت تربیعی^۱ مشهور (یعنی درختی که هر گره آن دقیقاً چهار شاخه دارد) نمایش داده می شود. توجه کنید که ریشه درخت متناظر با کل تصویر است و هر گره متناظر با یک تقسیم است. در این مورد تنها R بیشتر تقسیم شد.

اگر تنها تقسیم استفاده می شد، تقسیم نهایی به احتمال دارای نواحی مجاور با خواص یکسان بود. این مشکل را می توان با امکان اجرای همزمان ادغام و تقسیم برطرف کرد. برای اراضی قبود بخش ۴. ۷ لازم است تنها آن نواحی مجاور که اجتماع پیکسل هایشان گزاره P را برآورده می کنند، ادغام شوند؛ یعنی تنها اگر $T = P(R_j \cup R_k)$, دو ناحیه مجاور R_j و R_k ادغام می شوند.

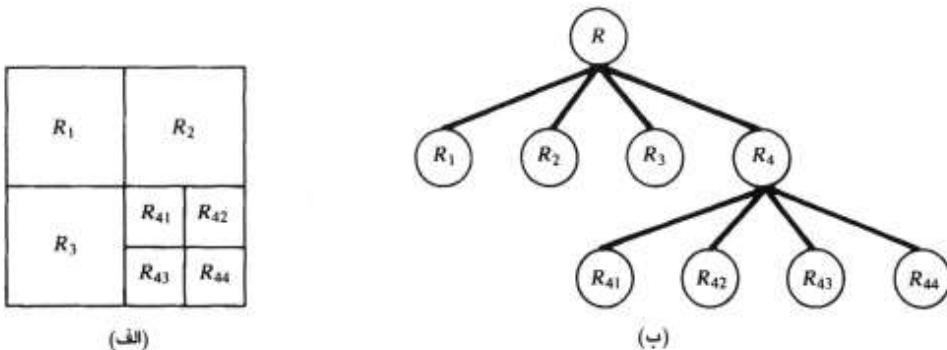
شرح قبلی را می توان با روای زیر جمع بندی کرد که در هر مرحله آن:

(۱) هر ناحیه R_i را که $P(R_i) = F$, به چهار ربع جدا از هم تقسیم می کنیم.

(۲) هر دو ناحیه مجاور R_j و R_k را که برای آنها $T = P(R_j \cup R_k)$, ادغام می کنیم؛ و

(۳) وقتی هیچ ادغام یا تقسیم بیشتری ممکن نباشد، توقف می کنیم.

چندگونه مختلف از این راه اصلی ممکن است. مثلاً یک امکان این است که در ابتدا تصویر را به مجموعه ای از بلوک های مربعی تقسیم کنیم. تقسیم بیشتر طبق شرح قبل انجام می شود، اما ادغام در ابتدا

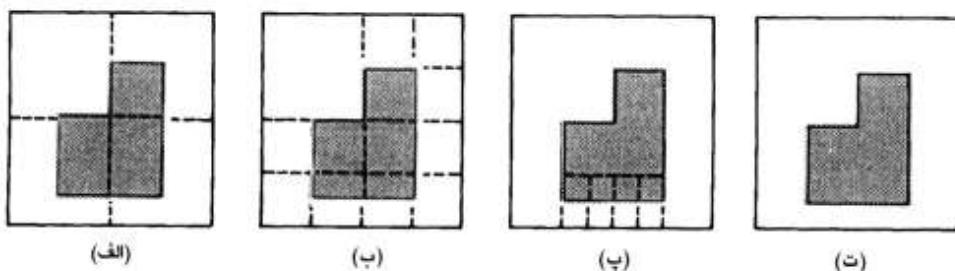


شکل ۷. ۳۷ (الف) تصویر بخش بندی شده؛ (ب) درخت تربیعی متناظر

به گروه های چهار بلوکی که زیر شاخه های یک شاخه از نمایش درخت تربیعی هستند و گزاره P را نیز برآورده می کنند، محدود می شود. وقتی ادغام های بیشتری از این نوع ممکن نباشد، روال با ادغام نهایی آن نواحی که شرط مرحله ۲ را برآورده می کنند، پایان می باید. در این مرحله، نواحی دارای ابعاد متفاوت را نیز می توان ادغام کرد. مزیت اصلی این راه این است که قبل از مرحله ادغام نهایی از درخت تربیعی یکسانی برای تقسیم و ادغام استفاده می کند.

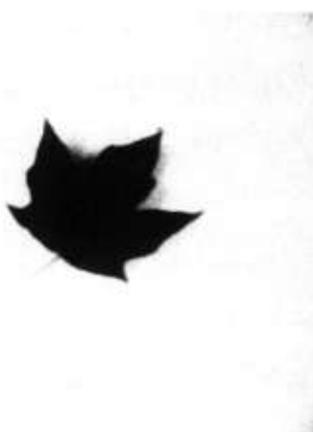
مثال ۷.۳۸. الگوریتم تقسیم و ادغام را نمایش می دهد. تصویر دارای یک شیب و زمینه است. برای سادگی هم شیب و هم زمینه سطوح خاکستری ثابتی دارند و اگر تمام پیکسل های درون R شدت یکسانی داشته باشند، آنگاه $T = P(R)$. چون برای کل ناحیه تصویر R ، $P(R) = F$ ، تصویر طبق شکل ۷.۳۸(الف) تقسیم می شود. همان طور که در شکل ۷.۳۸(ب) دیده می شود، در مرحله بعدی تنها ناحیه سمت چپ بالا، گزاره را برآورده می کند و بنابراین تغییر داده نمی شود، در حالی که سه ربع دیگر باز هم تربیع می شوند. در این مرحله می توان چند ناحیه را، به استثنای دو زیر ربعی که شامل قسمت پایین شیب هستند، ادغام کرد؛ آن دو زیر ربع^۱ گزاره را برآورده نمی کنند و باید بیشتر تقسیم شوند. شکل ۷.۳۸(پ) نتایج عمل تقسیم و ادغام را نمایش می دهد. در این مرحله تمام نواحی گزاره P را برآورده می کنند و ادغام نواحی مقتضی حاصل از آخرین عمل تقسیم نتیجه نهایی بخش بندی را می دهد که در شکل ۷.۳۸(ت) دیده می شود.

شکل ۷.۳۹. مثال عملی تری را نمایش می دهد. در این مورد اگر لاقل ۸۰ درصد پیکسل های R خاصیت $|z_i - m_i| \leq 2\sigma$ را داشته باشند، آنگاه $T = P(R)$ که سطح خاکستری پیکسل زام ناحیه R



شکل ۷.۳۸. مثالی از الگوریتم تقسیم و ادغام. (از [Fu, Gonzalez & Lee 1987])

۱) Subquadrants



(الف)



(ب)

(ب)

شکل ۷.۳۹. (الف) تصویر اولیه؛ (ب) حاصل روال تقسیم و ادغام؛ (پ) حاصل آستانه گیری از شکل ۷.۳۹. (ب)

است، و m و σ نیز به ترتیب میانگین و انحراف معیار سطوح خاکستری آن هستند. اگر $T = P(R_i)$ ، با این شرط مقدار تمام پیکسل های ناحیه R برابر m قرار داده می شدند. در شکل ۷.۳۹. (ب) حاصل کاربرد این روش بر شکل ۷.۳۹. (الف) را می بینید. به اثرات سایه روشن در بعضی گوشه های تصویر و در نزدیک برگ توجه کنید. تصویر شکل ۷.۳۹. (ب) با آستانه گیری از شکل ۷.۳۹. (ب) بدست آمد، که آستانه در وسط فاصله^۱ بین دو قله اصلی بافت نگار قرار داده شد. برای آستانه گیری، سایه روشن (وساقه برگ) حذف شدند.

□

همان طور که در مثال قبل نیز به کار گرفته شد، خواص مبنی بر میانگین و انحراف معیار می خواهند

بافت^۱ یک تاچه را کمی سازند (برای شرح مفصل درباره' بافت بخش ۳.۳.۸ را بینید). مفهوم بخش بندی بافت^۲ مبتنی بر کاربرد معیارهای بافت به عنوان گزاره های ($R_i P$) است. یعنی می توانیم با تعیین گزاره های مبتنی بر محتوای بانقی و کاربرد یکی از روش های مشروح در این بخش، بخش بندی بافت را اجرا کنیم.

۵.۷ استفاده از حرکت در بخش بندی

حرکت راهی قوی برای انسان و حیوانات برای استخراج اشیای مورد علاقه از زمینه دارای جزئیات نامرتب است. در کاربردهای تصویربرداری نظری کاربردهای روباتیک^۳، هدایت خودکار^۴ و تحلیل صحنه^۵ پویا^۶، حرکت از جایجاوی نسبی بین دستگاه تصویربردار و صحنه نشأت می گیرد. در بخش های بعدی، کاربرد حرکت در بخش بندی را هم در حوزه^۷ مکان و هم در حوزه^۸ فرکانس بررسی می کنیم.

۱۰.۷ روشهای مکانی

روش پایه ای

یکی از ساده ترین راه های آشکارسازی تغییرات بین دو قاب تصویر ($f(x, y, t_i)$ و $f(x, y, t_j)$) که به ترتیب در لحظات i و j برداشته شده اند، مقایسه پیکسل به پیکسل دو تصویر است. یک روش برای اجرای این کار تشکیل تصویر تفاضل است. فرض کنید تصویر مرجعی داریم که تنها دارای اجزای ساکن است. تفرقی این تصویر از یک تصویر بعدی که همان محیط را دارد اما دارای یک شیء متحرک نیز هست، اجزای ساکن را حذف می کند و تنها درایه های غیر صفر متناظر با اجزای غیر ساکن را باقی می گذارد. تفاضل دو تصویر که در لحظات i و j برداشته شده اند، را می توان به صورت

$$d_o(x, y) = \begin{cases} 1 & |f(x, y, t_i) - f(x, y, t_j)| > \theta \\ 0 & \text{در غیر این صورت} \end{cases} \quad (12-3.7)$$

تعریف کرد که θ یک آستانه است. توجه کنید که تنها اگر تفاضل سطح خاکستری دو تصویر در آن مختصات مقدار محسوسی، که با آستانه θ تعیین می شود، داشته باشد، $(y, x)_i^j$ واقع در مختصات

1) Texture

2) Texture segmentation

3) Robotic

4) Autonomous navigation

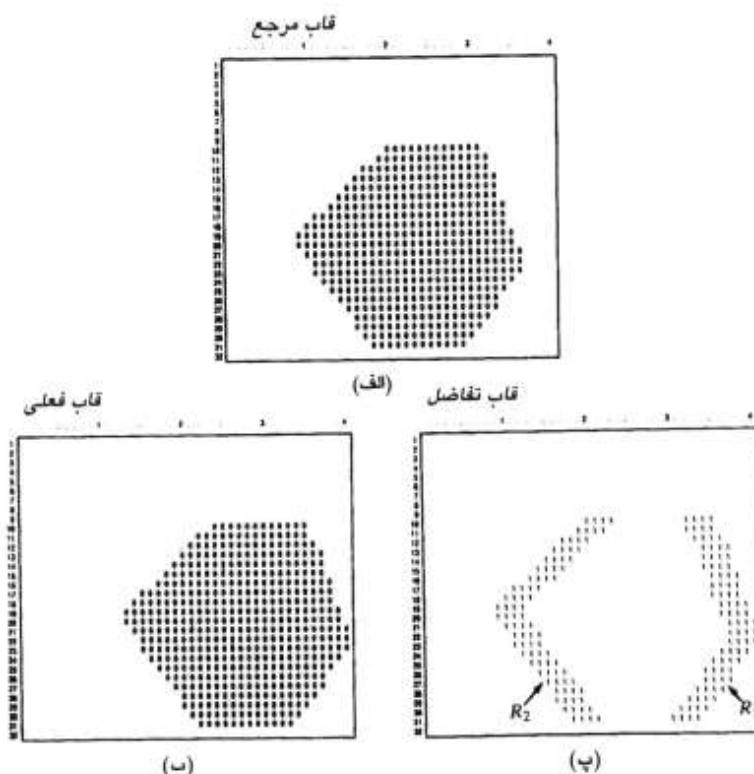
5) Dynamic scene analysis

۷. استفاده از حرکت در بخش بندی ۵۵۵

مکانی (x, y) مقدار "۱" دارد.

در تحلیل تصویر پویا تمام پیکسل های با مقدار "۱" در (x, y) حاصل حرکت شین در نظر گرفته می شوند. این روش تنها اگر دو تصویر ضبط شوند و روشنایی داخل کرانه های تعیین شده با θ نسبتاً ثابت باشد، اجرایشدنی است. در عمل غالباً درایه های با مقدار یک (x, y) ناشی از نویز هستند. نوع آن درایه ها نقاط منفردی در تصویر تفاضل هستند و یک راه ساده برای حذفشان تشکیل نواحی متصل ۴-۸ گانه از "۱" ها در (x, y) و سپس صرف نظر از هر ناحیه ای است که تعداد نقاط آن از یک تعداد پیش تعیین شده کمتر باشد. گرچه این عمل ممکن است باعث صرف نظر از اشیای کوچک و یا کم سرعت بشود، شناس این که نقاط باقیمانده در تصویر تفاضل واقعاً نتیجه حرکت باشند، را بهبود می دهد.

شکل ۷.۴۰ این مفاهیم را روشن می سازد. شکل ۷.۴۰ (الف) یک قاب تصویر مرجع را نمایش



شکل ۷.۴۰ (الف) تصویری که در لحظه t_1 برداشته شده است؛ (ب) تصویر برداشته شده در لحظه t_2 ؛ (پ) تصویر تفاضل. (از Jain [1981])

می دهد که در لحظه^۱ برش اشته شده است و دارای یک شیب باشد تثیت است که با سرعت یکنواخت روی زمینه ای با شدت ثابت حرکت می کند. شکل ۷. ۴۰ (ب) یک قاب تصویر را که در لحظه^۱ برش اشته شده است، نشان می دهد، و شکل ۷. ۴۰ (ب) تصویر تفاضل محاسبه شده با معادله^۲ (۷-۱-۵) و آستانه ای که بزرگتر از شدت ثابت زمینه است، را نمایش می دهد. توجه کنید که با فرآیند تفاضل گیری دو ناحیه^۳ جدا از هم تولید شدند؛ یکی حاصل از لبه جلویی و دیگری حاصل از لبه عقبی شیب متحرک است.

تفاضل های انباشتی^۱

همان طور که قبل^۴ بیان شد، تصویر تفاضل اغلب حاوی درایه های منفرد حاصل از نویز است. گرچه تعداد این درایه ها می توان با یک تحلیل اتصال آستانه گیری دار کاهش داد یا کاملاً حذف کرد، اما این فرآیند ممکن است اشیای کوچک یا کم سرعت را نیز حذف نماید. با بررسی تغییرات یک پیکسل در چند قاب به این مسئله می پردازیم، بنابراین یک «حافظه» را به کار می گیریم. ایده پایه ای این است که از تغییراتی که تنها گاهی روی دنباله^۵ قاب رخ می دهنده ولذا می توان آنها را با نویز تصادفی توصیف کرد، صرف نظر کنیم.

یک دنباله از قاب های تصویر $f(x, y, t_1), f(x, y, t_2), \dots, f(x, y, t_n)$ را در نظر بگیرید و را تصویر مرجع فرض نمایید. با مقایسه هر یک از تصاویر بعدی دنباله با تصویر مرجع، یک تصویر تفاضل انباشتی تشکیل می شود. به ازای هر پیکسل از تصویر انباشتی یک شمارشگر وجود دارد که هر بار بین آن پیکسل از تصویر دنباله و پیکسل متناظر از تصویر مرجع تفاوتی وجود داشته باشد، یک واحد افزایش می یابد. بنابراین وقتی قاب k^{th} با قاب مرجع مقایسه می شود، درایه^۶ یک پیکسل معین از تصویر انباشتی تعداد بارهایی که تاکنون سطح خاکستری آن محل متفاوت از مقدار پیکسل متناظر در تصویر مرجع بوده است، را نشان می دهد. تفاوت ها با استفاده از مثلاً معادله^۷ (۷-۱-۵) تعیین می شوند.

شکل ۷. ۴۱ این مفاهیم را روشن می سازد. اشکال ۷. ۴۱ (الف)-(ث) یک شیب مستطبی (توصیف شده با^۸) را نمایش می دهند که با سرعت ثابت یک پیکسل بر قاب به سمت راست حرکت می کند. تصاویر نمایش داده شده نمایانگر لحظات زمان متناظر با جابجایی های یک پیکسلی هستند. شکل ۷. ۴۱ (الف) قاب تصویر مرجع را نمایش می دهد، اشکال ۷. ۴ (ب)-(ت) قابهای ۲، ۳ و ۴ در دنباله هستند،

۱) Accumulative differences

۵۰۷ ۵. استفاده از حرکت در بخش بندی

(الف)	9 10 00000000 11 00000000 12 00000000 13 00000000 14 00000000 15 00000000 16	9 10 00000000 11 00000000 12 00000000 13 00000000 14 00000000 15 00000000 16	10 1 1 11 1 1 12 1 1 13 1 1 14 1 1 15 1 1 16	(ب)
(ب)	9 10 00000000 11 00000000 12 00000000 13 00000000 14 00000000 15 00000000 16	9 10 00000000 11 00000000 12 00000000 13 00000000 14 00000000 15 00000000 16	10 21 21 11 21 21 12 21 21 13 21 21 14 21 21 15 21 21 16	(ج)
(ج)	9 10 00000000 11 00000000 12 00000000 13 00000000 14 00000000 15 00000000 16	9 10 00000000 11 00000000 12 00000000 13 00000000 14 00000000 15 00000000 16	10 321 321 11 321 321 12 321 321 13 321 321 14 321 321 15 321 321 16	(د)
(د)	9 10 00000000 11 00000000 12 00000000 13 00000000 14 00000000 15 00000000 16	9 10 00000000 11 00000000 12 00000000 13 00000000 14 00000000 15 00000000 16	10 A98765438887654321 11 A98765438887654321 12 A98765438887654321 13 A98765438887654321 14 A98765438887654321 15 A98765438887654321 16	(ه)

شکل ۷ ۴۱. (الف) قاب تصویری مرتع؛ (ب) - (ث) قابهای ۲، ۳، ۴ و ۱۱؛ (ج) - (خ) تصاویر تفاضل انباشتی برای قابهای ۲، ۳، ۴ و ۱۱ (اعداد ۹-۱۶ نوشته شده در کتاب هر قاب فقط شماره خطوط هستند و با این شرح ارتباطی ندارند. (از Jain [1981]

و شکل ۷ ۴۱. (ث) قاب یازدهم است. اشکال ۷ ۴۱. (ج)-(خ) تصاویر انباشتی متناظر هستند، که می‌توان آنها را به صورت زیر شرح داد. در شکل ۷ ۴۱. (ج) ستون سمت چپ "۱" ها حاصل تفاضل بین شئی شکل ۷ ۴۱. (الف) و زمینه موجود در شکل ۷ ۴۱. (ب) است. ستون سمت راست "۱" ها حاصل تفاضل بین زمینه تصویر مرتع و لبه جلویی شئی متحرک است. تازمان قاب چهارم (شکل ۷ ۴۱-ت)، اولین ستون غیر صفر تصویر تفاضل انباشتی، سه شمارش را نشان می‌دهد، که بیانگر سه تفاضل بین آن ستون در تصویر مرتع و ستون متناظر در قاب‌های بعدی است. در نهایت شکل ۷ ۴۱. (خ) جمعاً ۱۰ که در مبنای ۱۶ با "A" نشان داده می‌شود) تغییر در آن محل را نشان می‌دهد. در مورد سایر درایه‌های نیز توضیح مشابهی صدق می‌کند.

غالباً در نظر گرفتن سه نوع تصویر تفاضل ابیاشتی ذیل مفید است: مطلق (PADI)، مثبت (AADI) و منفی (NADI). کمیت های مثبت و منفی با استفاده از معادله^(۷-۱) بدون قدر مطلق و با استفاده از قاب مرجع به جای $f(y, x)$ بدست می آیند. اگر سطوح خاکستری شیئی از نظر عددی بزرگتر از زمینه و تفاضل مثبت باشد، شیئی با یک آستانه['] مثبت مقایسه می شود؛ اگر تفاضل منفی باشد، شیئی با یک آستانه منفی مقایسه می شود. اگر سطوح خاکستری شیئی کوچکتر از زمینه باشند، این تعریف عکس می شود.

مثال: اشکال ۷.۴۲ (الف)-(پ) به ترتیب AADI، PADI و NADI را برای یک شیئی 20×20 پیکسلی که شدت‌ش بزرگتر از زمینه است و با سرعت ثابت درجهت جنوب شرقی حرکت می کند، نمایش می دهند. وقتی شیئی از محل اولیه اش جابجا شود، رشد مکانی PADI متوقف می شود. به عبارت دیگر وقتی شیئی که سطوح خاکستریش بزرگتر از زمینه هستند، کاملاً از محلش در تصویر مرجع جابجا شود، دیگر هیچ درایه جدیدی در تصویر تفاضل ابیاشتی مثبت تولید نمی شود. بنابراین وقتی رشدش متوقف شود، محل ابتدائی شیئی در قاب مرجع را نشان می دهد. از این خاصیت می توان در ایجاد یک مرجع از یک دنباله تصاویر پویا بهره برد. در شکل ۷.۴۲ توجه کنید که AADI هم شامل نواحی PADI و هم شامل نواحی NADI است و درایه های این تصاویر بیانگر سرعت و جهت حرکت شیئی هستند. در شکل ۷.۴۳ تصاویر موجود در شکل ۷.۴۲ باشد نور نمایش داده می شوند. □

تعیین یک تصویر مرجع

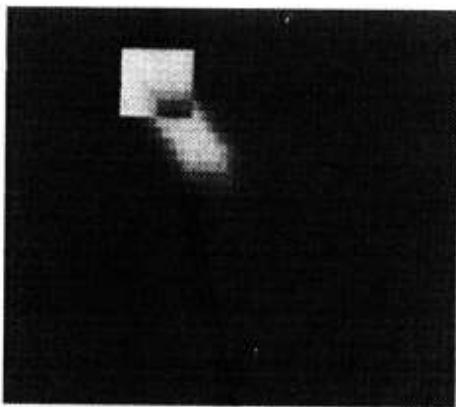
اساس توفیق روش‌های مشروح در دو بخش قبلی داشتن یک تصویر مرجع است که مقایسات بعدی را نسبت به آن انجام می دهند. همانطور که بیان شد، تفاضل بین دو تصویر در مسأله تصویربرداری پویا تا اندازه‌ای اجزای ساکن را حذف می کند، و تنها عناصری از تصویر را که متناظر با نویز و اشیای متحرك هستند، باقی می گذارد. همانطور که در بخش قبلی تشریح شد، می توان با روش فیلتر کردن مذکور در قبل یا با تشکیل یک تصویر تفاضل ابیاشتی به مسأله نویز رسیدگی کرد.

در عمل همیشه نمی توان یک تصویر مرجع به دست آورد که تنها دارای عناصر ساکن باشد، و لازم می شود که از یک مجموعه تصاویر که دارای یک یا چند شیئی متحرك هستند، تصویر مرجع بدست آید. این الزام بویژه در وضعیت هایی که صحته های شلوغ را توصیف می کنند یا در مواردی که

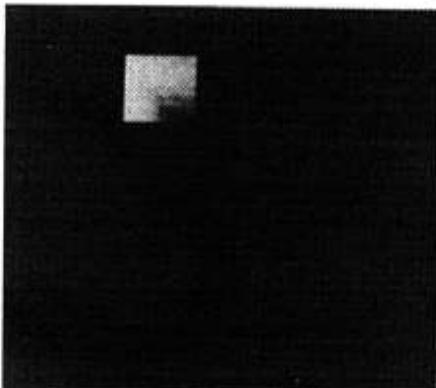
۵. استفاده از حرکت در بخش پندي

(1)

شکل ۴۲.۷ تصاویر تفاضل ایاشتی (الف) مطلق، (ب) مثبت، و (پ) منفی برای یک شیوه 20×20 پیکسلی که شدت آن بزرگتر از زمینه است و درجهٔ چوت چوب شرقی حرکت می‌کند. (از [Gain 1983])



(الف)



(ب)



(پ)

شکل ۴۳.۷ تصاویر تفاضل ابیاشی موجود در شکل ۷-۴۲ که باشد نور نمایش داده شده اند: (الف) (Jain / 1983)، (ب) (PADI)، و (پ) (AADI).

بهنگام سازی مکرر مورد نیاز است، وجود دارد. یک روال برای تولید تصویر مرجع به صورت زیر است. اولین تصویر در یک دنباله را تصویر مرجع در نظر بگیرید. وقتی یک جزء غیرساکن کاملاً از محلش در قاب مرجع جابجا شود، زمینه متناظر با آن در قاب فعلی را می توان عیناً در همان محل در تصویر مرجع قرار داد. وقتی تمام اشیای متحرک کاملاً از محلهای اولیه شان جابجا شوند، یک تصویر مرجع که تنها دارای اجزای ساکن است، ایجاد خواهد شد. جا به جایی شیئ را می توان با نظارت رشد PADI تعیین کرد.

مثال: اشکال ۷-۴۴ و ۷-۴۵ روش تشریح شده فوق را نمایش می دهند. شکل ۷-۴۴ دو قاب تصویر از یک تقاطع شهری را نمایش می دهد. تصویر اول مرجع در نظر گرفته می شود، و تصویر دوم همان

۵۶۱ ۷. استفاده از حرکت در بخش بندی



(الف)



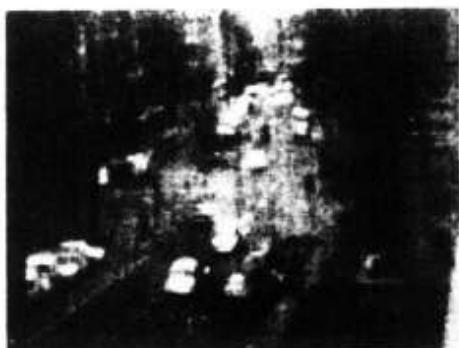
(ب)

شکل ۴۰.۷ دو قاب تصویر از یک صحنه رفت و آمد. دو شیء متحرک اصلی وجود دارند: یک اتومبیل سفید در وسط عکس و یک عابر پیاده در سمت چپ پایین عکس. (از [jain[1981]

صحنه را در لحظه‌ای بعد نمایش می‌دهد. یک اتومبیل که از چپ به راست حرکت می‌کند و یک عابر که در سمت چپ در پایین عکس از عرض خیابان عبور می‌کند، ویژگی‌های متحرک اصلی تصویر هستند. در شکل ۷. ۴۵. (الف) اتومبیل متحرک حذف شده است؛ در شکل ۷. ۴۵. (ب) عابر حذف شده است.

۴۰.۵.۷ روش‌های حوزه فرکانس

در این بخش مسئله تعیین تخمین‌های حرکت از طریق تبدیل فوریه را بررسی می‌کنیم. دنباله $f(x, y)$ در این قاب تصویر رقمنی با ابعاد $N \times M$ که با یک دوربین ساکن تولید شده



(الف)



(ب)

شکل ۴۵. (الف) تصویر حاصل از حذف اتومبیل و بازیافت زمینه؛ (ب) تصویر حاصل از حذف عابر و بازیافت زمینه تصویر (الف). از این تصویر می‌توان به عنوان مرجع استفاده کرد. (از [jain[1981]

است، را در نظر بگیرید. کار را با این فرض که تمام قاب های یک زمینه همگن باشد صفر دارند، شروع می کنیم. یک شیئی یک پیکسلی باشد واحد که با سرعت ثابت حرکت می کند، استثناست. فرض کنید که برای قاب اول ($t = 0$) صفحه تصویر روی محور x تصویر می شود^۱؛ یعنی شدت تمام پیکسل های هر ستون با هم جمع می شوند. این عمل یک آرایه^۲ یک بعدی با M درایه را می دهد که درایه های آن صفر هستند مگر در محلی که شبیه تصویر می شود. ضرب مؤلفه های این آرایه در عبارت $\exp[j2\pi k_i x \Delta t]$ ، $x = 0, 1, \dots, M - 1$ است، مجموعی برابر $\exp[j2\pi k_i x' \Delta t]$ تولید می کند. در این نمادگذاری k_i یک عدد صحیح مثبت و Δt بازه زمانی بین قابها است.

فرض کنید که در قاب دوم ($t = 1$) شبیه به مختصات (x', y') جا به جا شده است؛ یعنی یک پیکسل موازی با محور x حرکت کرده است؛ پس تکرار روال فوق مجموع $\exp[j2\pi k_i(x' + 1)\Delta t]$ را می دهد. اگر شبیه به حرکت با سرعت یک پیکسل بر قاب ادامه دهد، حاصل روال فوق در هر لحظه^۳ صحیح از زمان برابر $\exp[j2\pi k_i(x' + 1)\Delta t]$ خواهد بود، که با استفاده از رابطه^۴ اولر می توان آن را بازای $t = 0, 1, \dots, T - 1$ به صورت

$$\exp[j2\pi k_i(x' + t)\Delta t] = \cos[2\pi k_i(x' + t)\Delta t] + j \sin[2\pi k_i(x' + t)\Delta t] \quad (2-5, 7)$$

بیان کرد. به عبارت دیگر، حاصل، سینوس مختلط با فرکانس k_i است. اگر شبیه در فاصله زمانی بین قاب های ۱ پیکسل (در جهت x) حرکت می کرد، فرکانس سینوس k_i می بود. چون^۵ با گام های صحیح بین 0 تا $T - 1$ تغییر می کند، محدود کردن k_i به مقادیر صحیح باعث می شود که تبدیل فوریه گستته سینوس مختلط دو قله-یکی در فرکانس k_i و دیگری در فرکانس $-k_i$ داشته باشد. همان طور که در بخش ۳.۳. تشریح شد، قله دوم حاصل تقارن تا شده^۶ است، و می توان از آن صرف نظر کرد. بنابراین با اجرای یک فرآیند جستجوی قله در طیف فوریه، k_i به دست می آید. تقسیم این کمیت بر k_i را می دهد که مؤلفه سرعت در جهت x است، زیرا آنگ قاب، معلوم فرض می شود. به طور مشابه تصویر کردن بر محور y ، مؤلفه سرعت در جهت y ، یعنی k_y را می دهد.

دنباله ای از قاب ها که در آن ها هیچ حرکتی رخ ندهد، عبارات نمایی یکسانی تولید می کنند که تبدیل

۵۶۳ ۷. استفاده از حرکت در بخش بندی

فوریه شان دارای فقط یک قله در فرکانس صفر (فقط یک مقدار ثابت) است. بنابراین چون عملیات تشریح شده تا کنون خطی بوده است، در حالت عام که دنباله ' تصاویر شامل یک یا چند شیوه متحرک در یک زمینه ' ساکن دلخواه باشد، تبدیل فوریه دنباله ' تصاویر یک قله در فرکانس صفر متاظر با مؤلفه های تصویری ساکن و قله هایی در محل های متناسب با سرعت اشیا خواهد داشت.

این مفاهیم را می توان با روابط زیر جمع بندی کرد. برای دنباله ای از T تصویر رقمی با ابعاد $N \times M$ ، جمع تصویرهای وزن داده شده^۱ بر روی محور x در هر لحظه ' صحیح از زمان برابر است با

$$g_x(t, k_1) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y, t) e^{j2\pi k_1 x \Delta t} \quad t = 0, 1, \dots, T-1. \quad (3-5. V)$$

مشابهآ جمع تصویرها بر محور y عبارتست از

$$g_y(t, k_2) = \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} f(x, y, t) e^{j2\pi k_2 y \Delta t} \quad t = 0, 1, \dots, T-1 \quad (4-5. V)$$

که k_1 و k_2 اعداد صحیح مثبت هستند.

تبدیل فوریه های یک بعدی معادلات (۷. ۳-۵) و (۷. ۴-۵) به ترتیب عبارتند از

$$G_x(u_1, k_1) = \frac{1}{T} \sum_{t=0}^{T-1} g_x(t, k_1) e^{-j2\pi u_1 t / T} \quad u_1 = 0, 1, \dots, T-1 \quad (5-5. V)$$

و

$$G_y(u_2, k_2) = \frac{1}{T} \sum_{t=0}^{T-1} g_y(t, k_2) e^{-j2\pi u_2 t / T} \quad u_2 = 0, 1, \dots, T-1. \quad (6-5. V)$$

همان طور که در بخش ۴. ۴ بیان شد، در عمل این تبدیل ها با استفاده از الگوریتم FFT محاسبه می شوند. رابطه ' فرکانس - سرعت عبارتست از

$$u_1 = k_1 v_1 \quad (7-5. V)$$

و

$$u_2 = k_2 v_2 \quad (8-5. V)$$

در این روابط سرعت بر حسب پیکسل بر کل زمان قاب است. مثلاً $v_1 = 10$ به عنوان 10 پیکسل حرکت

1) Weighted projections

در T قاب تعبیر می شود. برای قاب هایی که با فواصل زمانی یکسان برداشته می شوند، سرعت فیزیکی واقعی به نرخ قاب و فاصله بین پیکسل ها بستگی دارد. بنابراین اگر $v_1 = 10$ و $T = 30$ ، آهنگ قاب دو تصویر در ثانیه، و فاصله بین پیکسل ها 5 cm باشد، سرعت فیزیکی واقعی در جهت x برابر است با

$$v_1 = \frac{1}{T} = (قاب 30) / (\ثانیه / قاب 2) (پیکسل / متر 5) (پیکسل 10)$$

علامت مؤلفه x سرعت با محاسبات زیر به دست می آید

$$S_{1x} = \left. \frac{d^2 \operatorname{Re}[g_i(t, k_i)]}{dt^2} \right|_{t=0} \quad (9-5.7)$$

و

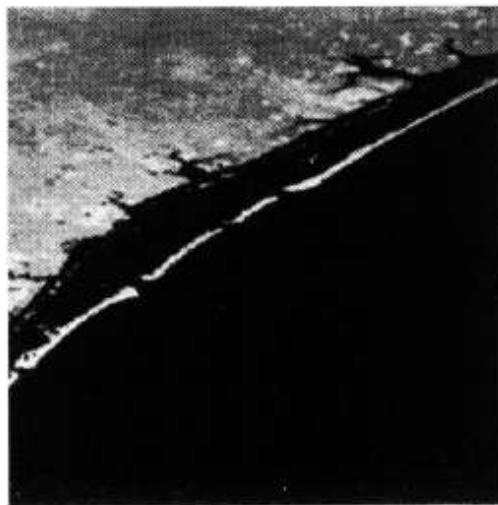
$$S_{2x} = \left. \frac{d^2 \operatorname{Im}[g_i(t, k_i)]}{dt^2} \right|_{t=0} \quad (10-5.7)$$

چون g_i سینوسی است، می توان نشان داد که اگر مؤلفه سرعت v_1 مثبت باشد، S_{1x} و S_{2x} در n ، یک لحظه دلخواه از زمان، هم علامت خواهند بود. بر عکس علامت مختلف داشتن S_{1x} و S_{2x} نشانه منفی بودن مؤلفه سرعت مذکور است. اگر یکی از S_{1x} یا S_{2x} صفر باشد، نزدیکترین لحظه بعدی، یعنی $t = n \pm \Delta t$ ، را بررسی می کنیم. در محاسبه علامت¹ نیز توضیح مشابهی صدق می کند.

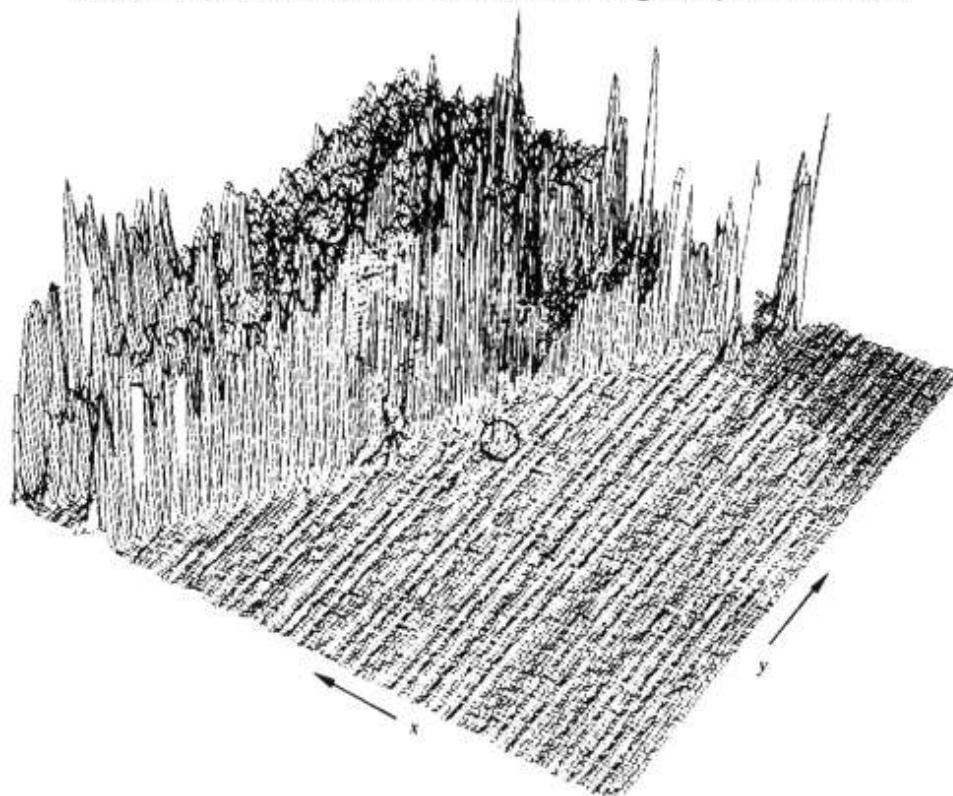
مثال: اشکال ۷.۴۶ و ۷.۴۹ مزیت راه اخیر رانمایش می دهند. شکل ۷.۴۶ یکی از قاب های دنباله ۳۲ قابی تصاویر ماهواره لندست^۱ که با افزودن نویز سفید به تصویر مرجع تولید شده است، رانشان می دهد. این دنباله دارای یک شیئ متحرک است که روی قاب اضافه شده است و با سرعت 5 m/s پیکسل بر قاب در جهت x و یک پیکسل بر قاب در جهت z حرکت می کند. شیئ مورد نظر که در شکل ۷.۴۷ مخصوص در دایره است، دارای توزیع شدت گوسی بوده و روی یک ناحیه کوچک (۹ پیکسلی) گسترده است و به سادگی نیز با چشم تشخیص داده نمی شود. نتایج محاسبه معادلات (۷.۵-۵) و (۷.۶-۵) با $k_1 = 4$ و $k_2 = 6$ به ترتیب در اشکال ۷.۴۸ و ۷.۴۹ نمایش داده می شوند. قله موجود در S_{1x} در شکل ۷.۴۸، طبق معادله (۷.۵-۵) $= 0.5$ را می دهد. به طور مشابه قله موجود در S_{2x} در شکل ۷.۴۹، طبق معادله (۷.۶-۵) $= 1.0$ را می دهد.

۱) LANDSAT

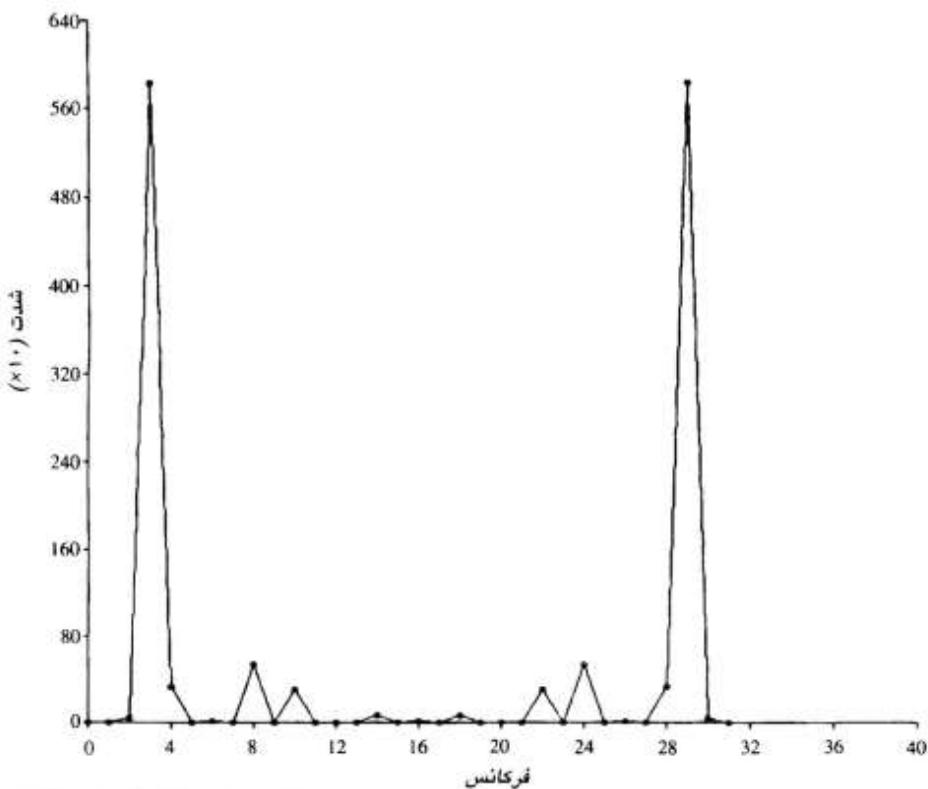
۵۶۵ ۷. استفاده از حرکت در بخش بندی



شکل ۷.۴۶ یک قاب دریافتی از LANDSAT (Cowart, Snyder & Ruedger 1983) (ا)



شکل ۷.۴۷ نمودار شدت شکل ۷.۴۶ که در آن هدف محصور در یک دایره است.
(Rajala, Raddle, & Snyder 1983)



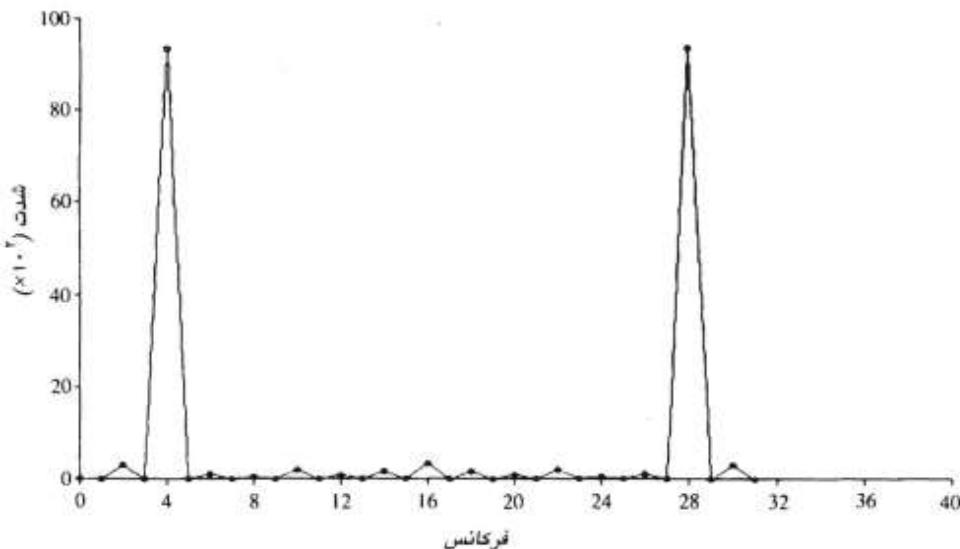
شکل ۷.۱۰.۷ طیف معادله (۷.۵-۵) که یک قله در 3 Hz را نشان می دهد.

(Rajala, Riddle & Snyder [1983])

رهنمودها برای انتخاب k_1 و k_2 رامی توان با کمک شکل های ۷.۴۸ و ۷.۴۹ شرح داد. برای نمونه، فرض کنید که به جای $4 = k_2$ از $15 = k_2$ استفاده کنیم. در آن حالت چون $1.0 = v_2$ ، قله های شکل ۷.۴۹ در $15 = k_2$ خواهند بود، که این به معنای روی هم افتادگی شدید فرکانسی است. همان طور که در بخش ۳.۹.۳ تشریح شد، رویهم افتادگی فرکانسی ناشی از کم نموده برداری (دراینجا تعداد قاب های خیلی کم، زیرا محدوده ' T با تعیین می شود) است. چون $v_1 = k_1$ ، یک امکان این است که k_1 نزدیکترین عدد صحیح به $k_1 = u_{\max} / v_{\max}$ انتخاب شود، که ' \max محدودیت فرکانس روی هم افتادگی است که با T تعیین می شود، و v_{\max} بیشینه سرعت مورد انتظار شیوه است.

۶.۷ ملاحظات پایانی

بخش بندی تصویر، یک مرحله مقدماتی و ضروری در بیشتر مسائل تشخیص خودکار الگوی تصویری



شکل ۷. ۴۹. طیف معادله (۷. ۵-۶) که یک قله در $f = 4$ را نشان می دهد.

(Rajala, Riddle & Snyder [1983])

و مسائل تحلیل صحنه است. همان طور که با محدوده مثال های ارائه شده بیان شد. رجحان یک روش بخش بندی بر دیگری به میزان زیادی به مشخصات خاص مسئله ای که بررسی می شود، بستگی دارد. روش های مذکور در این فصل، گرچه جامع نیستند، نماینده روش هایی هستند که معمولاً در عمل استفاده می شوند. از مراجع زیر می توان به عنوان مبنای برای مطالعه بیشتر در این مبحث استفاده کرد.

مراجع

کار در مورد استفاده از ماسک ها در آشکارسازی ناپیوستگی های سطح خاکستری سابقه طولانی دارد. ساختارهای نقابی که ساختارهای مذکور در بخش ۷. ۱ را تکمیل می کنند، در مقالات قدیمی Robinson [1976]، Kirsh [1971]، Prewitt [1970]، Roberts [1965] حاوی نقاب های متعدد و ارزیابی کارآئی آنهاست. شرح درباره خواص گذر از Fram & Deutsh [1975] صفر لایسنسی براساس مقاله Marr & Hildredth [1980] و کتاب Marr [1982] است. همچنین مقاله ای از Clark [1989] را درباره تقویت لبه های تولید شده با الگوریتم های گذر از صفر بینیسد. در

مقاله [Piech 1990] قسمت هایی از مقاله Clark تصحیح شده است. مقاله ای از & Malik [1990] نیز مورد علاقه است. روابط برداری مورد استفاده برای آشکارسازی همزمان ناپوستگی ها براساس مقاله Frei, Chen [1977] است؛ مقاله [Park & Choi 1990] را نیز بینید. آشکارسازی لبه گرچه سابقه طولانی دارد، اما هنوز هم مبحث تحقیقاتی فعالی می باشد، که نمونه های Petrov, Kittler [1991], Haralick, Lee [1990], Saito, Cunningham [1990] اخیر آن را در مقالات [Kittler 1991] بینید.

برای مطالعه 'عمومی' بیشتر درباره 'مبانی آشکارسازی با نقاب و کاربرد آن در پردازش تصویر و بینایی رایانه ای'، به کتب [Fu, Gonzalez & Lee 1987], Levine [1985], Kak, Rosenfeld [1982] و Jain [1989] مراجعه کنید. این کتب جزئیات بیشتری نیز درباره 'موضوع بخش ۷ . ۲ . ۱' گردآورده اند. تبدیل هاف (بخش ۷ . ۲ . ۲) را اولین بار [P.V.C. Hough 1962] در یکی از ادارات ثبت آمریکا پیشنهاد کرد که بعداً [Duda, Hart 1972] آن را منتشر کردند. [Ballard 1981] تعمیمی از تبدیل هاف را برای آشکارسازی شکل های دلخواه پیشنهاد کرد. در این باره مقالات [Hsu, Huang 1990] و Martelli [1972, 1976] را نیز بینید. موضوع بخش ۷ . ۲ . ۳ براساس دو مقاله از [Brummer 1991] است. [Ramer 1975] راه جالب دیگری را که مبتنی بر جستجوی با هزینه کمینه بود، بیان کرد. برای مطالعه بیشتر درباره 'روش های جستجوی گراف' به [Nilsson 1971, 1980] و [Umeyama 1988] مراجعه کنید. مقاله ای از [Vuylsteke, Kittler 1990] درباره 'برچسب زنی لبه ها در حضور نویز نیز مورد توجه است.

آستانه گیری یکی از اولین روش های ارائه شده برای بخش بندی تصاویر رقمه است. چند مرجع اولیه 'نوعی درباره' این مبحث، مقالات [Fornago 1963], Doyle [1962]، Narasimhan و همکاران [Rosenfeld 1965] است. مقالات مروری [Sahoo, Weszka 1978] و [Lee, Chung & Park 1988] نیز مناسب است. روش آستانه گیری بهینه ارائه شده در بخش [Perez & Gonzalez 1987] و [White & Rohrer 1983] می باشد. مقالات [Chow, Kaneko 1972] و [Allen, Muerle, Brice, Fenne 1970] ناچه گرا، [Parker 1991] و [Gonzalez 1987] را نیز بینید.

مراجع اولیه درباره 'بخش بندی ناچه گرا'، [Allen 1968] و [Muerle, Brice, Fenne 1970]

هستند. مقالات مروری [1976] Zucker، [1981] Fu، Mui ابتدا چند مفهوم متعدد از (مشترک) ^۱ را تعیین می کنند و سپس شایستگی های روش های گوناگون بخش بندی را شرح می دهند. مفهوم درخت تربیعی ^۲ که در بخش ۷. ۴. ۳. تشریح شد، در ابتدا به پیشنهاد [1972] Klinger بود، که وی این راه را تجزیه 'منظمه' نامید. مقاله [1974] Horowitz، Pavlidis را نیز بینند. پیشرفت های ارائه شده در [1982] Ballard، Brown و تایاجی نظری نتایج ارائه شده در [1983] Abel، Grosky، Jain در [1985] Haddon، Boyce نمونه هایی از پیشرفت های اخیر در این مبحث هستند. روش [1985] Mark برای ترکیب اطلاعات ناحیه ای و مرزی به منظور بخش بندی و روش بخش بندی [1990] Pavlidis، Liow که براساس مجتمع سازی رشد ناحیه و آشکار سازی لبه است، نیز جذاب هستند. نمونه هایی از روش های مبتنی بر استفاده از بافت برای بخش بندی تصویر در مقالات [1982] Haralick، Rosenfeld، Kak، [1985] Shapiro، [1991] Liu، Bouman آمده است. بخش بندی مبتنی بر مقاهم شکل شناسی Meyer، Beucher [1990] سطح خاکستری (به بخش ۸. ۴ نگاه کنید) نیز در مرحله آغازین رشد است.

مرور جامعی بر کاربرد شکل شناسی در بخش بندی تصویر داشته اند.

مطلوب بخش ۷. ۵. ۱ براساس دو مقاله از [1981] Jain، [1983] است. مطالب بخش ۷. ۵. ۲ براساس روشنی است که [1983] Rajala، Riddle & Snyder ارائه کرده اند. سایر کارهای مورد علاقه در تحلیل تصویر پریا ^۳ مقالات [1981] Aggarwal، [1983] Huang، [1985] Yachida، [1988] Adiv هستند. Koch و همکارانش [1988] یک پیاده سازی سخت افزاری جالب را برای محاسبه 'عوامل حرکت ارائه کردند. برای دست یابی به مجموعه ای از مقالات درباره 'تخمین و تحلیل حرکت به' [1989] Thompson، [1990] Cumani، Shariat، Price و همکارانش [1991] نیز مورد توجه هستند.

مسائل

- ۱.۷ یک تصویر دودویی حاوی خطوط راستی درجهات افقی، عمودی، 45° و -45° است. مجموعه ای از نقاط های 3×3 طرح کنید که بتوان از آنها برای آشکار سازی انفصالت یک پیکسلی در این خطوط استفاده کرد. فرض کنید که سطح خاکستری خطوط ۱ و سطح خاکستری زمینه ۰ است.

- ۴.۷ اگر به جای مدل لبه موجود در شکل ۷.۴ مجبور شوید از مدل شبیه لبه با شکل زیر استفاده کنید، شکل ۷.۴ چگونه می شود؟



مدل شبیه لبه

- ۴.۸ نتایج حاصل از یک بار مرور تصویر با بعضی نقاب های دو بعدی را می توان با دو بار مرور تصویر با نقاب های یک بعدی نیز به دست آورد. مثلاً حاصل استفاده از یک نقاب آرام کننده $\times 3$ با ضرایب $1/9$ را می توان با اجرای ترتیبی سه مرحله ذیل نیز به دست آورد: ۱) مرور تصویر با نقاب [۱ ۱ ۱] (۲) مرور حاصل مرحله' اول با نقاب

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

- (۳) مقایسه کردن حاصل مرحله' دوم با ضریب $1/9$. نشان دهید که نقاب های سوبل (شکل ۷.۵) را می توان با مرور یک نقاب تفاضل گیر^۱ به شکل $[1 0 -1]$ (با شکل عمودی آن) و سپس مرور یک نقاب آرام کننده به شکل $[1 2 1]$ (با شکل عمودی آن) نیز پیاده سازی کرد.

- ۴.۹ یک تصویر دودویی با ابعاد $N \times M$ پیکسل را که دارای یک مریع $n \times n$ پیکسلی از "۱" ها در مرکزش است، در نظر بگیرید. سایر پیکسل های تصویر، پیکسل های زمینه هستند که برچسب "۰" دارند.

- الف) گرادیان این تصویر را برای تقریب داده شده در معادله' (۷.۱-۵) رسم کنید. فرض کنید که G_x و G_y با استفاده از عملگرهای سوبل به دست می آیند. مقدار تمام پیکسل های تصویر گرادیان را به دست آورید.

- ب) بافت نگار جهات لبه محاسبه شده با معادله' (۷.۶-۱) را رسم کنید. در برچسب زنی ارتفاع

1) Smoothing

2) Differencing

هر قله' بافت نگار دقت کنید.

پ) لابلسین تصویر را باز ای تقریب داده شده در معادله^(۷.۱-۱۰) رسم کنید. مقدار تمام پیکسل های تصویر لابلسین را به دست آورید.

۵.۷ نشان دهید که لابلسین رقمی حاصل از معادله^(۷.۱-۱۰) با تفاضل متوسط چهار همسایه $f(x, y)$ از $f(x, y)$ تنااسب (با ضرایب $1/4$) دارد. (همان طور که در بخش ۴.۳.۲ بیان شد، فرآیند تفریق یک نسخه مات شده' $f(x, y)$ از خودش را حذف ناتیزی^۱ می نامند).

۶.۷ الف) ثابت کنید که مقدار متوسط عملگر لابلسین $h^2/7$ با معادله^(۷.۱-۱۲)، برابر صفر است.

ب) ثابت کنید که مقدار متوسط تصویر حاصل از کانوالو این عملگر با هر تصویر دلخواه نیز صفر است.

پ) آیا قسمت (ب) به طور کلی برای تقریب لابلسین داده شده با معادله^(۷.۱-۱۰)، درست است؟ شرح دهید.

۷.۷ بزرگترین پاسخ عملگرهای گرادیان سوبیل (شکل ۷.۵) برای لبه های عمودی و افقی است. عملگرهای مشهور به گرادیان گردان^۲ با ابعاد 3×3 طوری طراحی می شوند که گرادیان لبه ها در جهات E, S, SW, W, NW, N, NE است. SE را اندازه بگیرند. شکل این هشت عملگرا را با استفاده از ضرایب با مقادیر $+1$ یا -1 تعیین نمایید. با توجه به اینکه جهت گرادیان عمود بر جهت لبه است، جهت گرادیان هر نقاب را تعیین نمایید.

۸.۷ جهت خطی (خطوطی) را که هریک از نقاب های شکل ۷.۱۲ بزرگترین پاسخ را به آنها می دهد، تعیین کنید. فرض کنید که ضخامت خطوط یک پیکسل است.

۹.۷ روشی برای آشکارسازی فاصله های به طول ۱ تا L پیکسل در قطعه خطوطی تصویر گرادیان پیشنهاد کنید. فرض کنید که زمینه ثابت است، و تمام خطوط با سطح شدت یکسانی رمز شده اند و هریک به ضخامت یک پیکسل هستند. به جای تلاش برای ساخت نقاب هایی که فاصله ها را آشکارسازند، روش خود را بر اساس تحلیل اتصال همسایگی هشت گانه (بخش ۴.۲) بنا کنید.

1) Unsharp masking

2) Compass gradient

۱۰.۷) شرح دهد که چرا نگاشت هاف نقطه^۱ ۱ در شکل ۷.۱۸. (ب) یک خط راست است.

(ب) آیا این نقطه تنها نقطه‌ای است که این نتیجه را می‌دهد؟

(ب) ارتباط مجاورتی آینه‌ای نمایش داده شده در شکل ۷.۱۸. (ت) را شرح دهد.

۱۱.۷) روایی کلی برای به دست آوردن نمایش نرمال یک خط از معادله 'شیب-عرض از مبدأ

آن، $y = ax + b$ ، طرح کنید.

(ب) نمایش نرمال خط $7 - 2x = y$ را پیدا کنید.

۱۲.۷) یک زمینه^۲ کاربردی مهم برای روش‌های بخش بندی تصویر، پردازش تصاویر حاصل از

پیشامدهای مشهور به اتفاق حباب^۱ است. این تصاویر حاصل آزمایشات فیزیک

پرانرژی^۲ هستند که در آن پرتویی از ذرات با خواص معلوم روی هدفی با هسته‌های معلوم

هدایت می‌شود. یک پیشامد نوعی مشکل از مسیرهای ذرات ورودی است که هر یک از آنها

در صورت تلاقی با هم از همان نقطه تلاقی به چند مسیر دیگر تقسیم می‌شود. یک روش

بخش بندی برای آشکارسازی تمام مسیرهایی که لاقل ۱۰۰ پیکسل دارند و در یکی از شش

جهت غیرافقی $\pm 25^\circ$ ، $\pm 50^\circ$ و $\pm 75^\circ$ هستند، پیشنهاد کنید. خطای تخمین مجاز در

هر یک از این شش جهت $\pm 5^\circ$ است. یک مسیر معتبر باید لاقل به طول ۱۰۰ پیکسل باشد و

نباشد بیش از سه فاصله داشته باشد و هر یک از این فاصله‌ها نباید از ۱۰ پیکسل طولانی تر

باشد. می‌توانید فرض کنید که تصاویر جهت دودویی شدن پیش‌پردازش شده‌اند و تمام

مسیرهای نیز به جز در نقطه تلاقی که مسیرها از آن پراکنده می‌شوند، یک پیکسل عرض دارند.

روش شما باید بتواند مسیرهایی را که جهات یکسان اما مبادی متفاوت دارند، از هم تشخیص

دهد.

۱۳.۷) تمام لبه‌های ممکن داده شده با گراف شکل ۷.۲۲ را روی شکل ۷.۲۱ قراردهید.

(ب) هزینه مسیر با هزینه کمینه را محاسبه کنید.

۱۴.۷) لبه متناظر با مسیر با هزینه کمینه را در قطعه تصویر زیر پیدا کنید. توجه کنید که اعداد درون

پرانتزها بیانگر شدت هستند. فرض کنید که لبه در ستون اول شروع و در آخرین ستون ختم

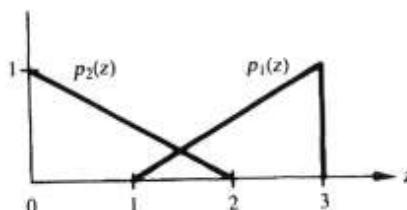
می‌شود.

	0	1	2
0	.	.	.
	(2)	(1)	(0)
1	.	.	.
	(1)	(1)	(7)
2	.	.	.
	(6)	(8)	(2)

یک تصویر بدون نویز با ابعاد $N \times N$ را در نظر بگیرید که $N/2$ ستون اول آن سطح خاکستری L_A و سایر ستون‌ها سطح خاکستری L_B دارند و $L_B > L_A$. بافت نگار این تصویر تنها دو قله هم ارتفاع دارد که یکی از آنها در L_A و دیگری در L_B قرار دارد. بخش بندی این تصویر به دو نیمه براساس محتوای سطح خاکستری آن کار آسانی است که می‌توان آن را با یک آستانه سراسری که بین L_A و L_B قرار دارد، انجام داد. با این حال حدس بزنید که اگر تصویر را در یک تابع شبیه سطح خاکستری^۱ که از مقدار ۰ در سمت چپ تا مقدار K در سمت راست تغییر می‌کند، که $L_B > K > L_A$ ، ضرب نماییم، بافت نگار تصویر حاصل به چه شبیه خواهد بود؟ قسمت‌های گوناگون این بافت نگار را برچسب بزنید.

تصویری مشکل از لکه‌های ناهمپوشان کوچک با میانگین سطح خاکستری $m_1 = 150$ و پراش $\sigma_1^2 = 400$ است که روی زمینه‌ای با میانگین $m_2 = 25$ و پراش $\sigma_2^2 = 625$ پخش شده‌اند. لکه‌ها در کل تقریباً ۲۰ درصد ناحیه تصویر را اشغال می‌کنند. براساس آستانه‌گیری روشی را برای بخش بندی لکه‌های تصویر پیشنهاد کنید.

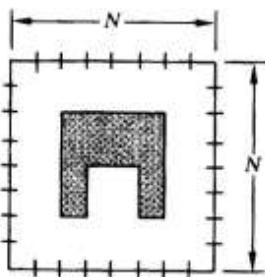
فرض کنید که تصویری توزیع‌های شدت زیر را دارد، که $p_1(z)$ متناظر با شدت اشیا و $p_2(z)$ متناظر با شدت زمینه است. با فرض این که $p_1 = p_2$ ، آستانه بهینه بین پیکسل‌های شبیه و زمینه را پیدا کنید.



1) Gray-scale wedge

پردازش تصویر رقمه ۵۷۴

- ۱۸.۷ با شروع از معادله' (۷.۹-۳) ، معادلات (۷.۱۰-۳) و (۷.۱۱-۳) را به دست آورید.
با شروع از معادلات (۷.۱۰-۳) و (۷.۱۱-۳) ، معادله' (۷.۱۲-۳) را به دست آورید.
- ۱۹.۷ تصویر مسأله ۷.۱۶ را در نظر بگیرید و یک روش بخش بندی مبتنی بر رشد ناحیه پیشنهاد کنید.
- ۲۰.۷ تصویر زیر را با استفاده از روش تقسیم و ادغام بیان شده در بخش ۷.۴.۳ بخش بندی کنید. اگر تمام پیکسل های ناحیه' R_i شدت یکسانی داشتند، قرار دهد $T = P(R_i)$. درخت تربیعی متناظر با حاصل بخش بندی را نمایش دهید.



- ۲۱.۷ باید با روش های تصویربرداری سریع، سرعت یک گلوله در هوای تعیین شود. روش انتخابی یک فلاش^۱ را که سطح تصویربرداری یک دوربین تلویزیونی را برای مدت T ثانیه روشن می کند، به کار می گیرد. گلوله ۵/۲ سانتیمتر طول و یک سانتیمتر پهنا دارد، و محدوده سرعتش $m/s \pm 250$ است. بخش نوری دوربین، تصویری تولید می کند که در آن گلوله ۱۰ درصد پهنا افقی یک تصویر رقمه 256×256 را اشغال می کند.
- الف) بیشینه مقدار T که تضمین می کند ماتی حاصل از حرکت گلوله از یک پیکسل بیشتر نمی شود، را تعیین کنید.
- ب) کمینه تعداد قابها بر ثانیه که تضمین می کند هنگام عبور گلوله از میدان دید دوربین لااقل دو تصویر کامل از آن برداشته می شود، را تعیین کنید.
- پ) یک روال بخش بندی برای استخراج خودکار گلوله از دنباله قابها پیشنهاد کنید.
- ت) روشی برای تعیین خودکار سرعت گلوله پیشنهاد نماید.

1) Flash

فصل ۸

نمایش و توصیف

«اکنون خوب فکر کن؛ آیا بارها تأکید نکرده ایم که تنها نام هایی
بدرستی انتخاب شده اند که بیانگر شباهتها و شکل چیزهایی که
می نامند، هستند.»

سقراط

پس از این که تصویر با روش های نظری روش های مشروح در فصل ۷ بخش بندی شد^۱، معمولاً پیکسل های بخش بندی شده در قالبی مناسب تر برای پردازش رایانه ای نمایش داده و توصیف می شوند. در نمایش یک ناحیه، دو انتخاب وجود دارد: (۱) می توانیم ناحیه را بر حسب مشخصه های خارجی (مرز) یا (۲) بر حسب مشخصه های داخلی (پیکسل های تشکیل دهنده) آن نمایش دهیم. با این حال انتخاب روش نمایش تنها بخشی از کار آماده سازی داده ها برای رایانه است. مرحله بعدی توصیف^۲ ناحیه بر اساس نوع نمایش انتخابی است. مثلاً می توان ناحیه را با مرزش نمایش داد، که خود مرز با ویژگی های نظری طول، جهت خط مستقیمی که نقاط انتهایی^۳ را به هم متصل می کند، و تعداد تعریف های^۴ آن، توصیف می شود. معمولاً وقتی نمایش خارجی انتخاب می شود که ویژگی های شکلی با اهمیت تر باشند، و در صورتی که خواص انعکاسی نظریرنگ و بافت اهمیت بیشتری داشته باشند، نمایش داخلی انتخاب می شود. در هر دو مورد، ویژگی های منتخب به عنوان توصیف گر، باید در حد امکان به تغییراتی نظری تغییر ابعاد، انتقال و چرخش غیرحساس باشند. غالباً توصیف گرهای مشروح در این فصل یک یا چند عدد از این خواص را دارا می باشند.

۱۰.۸ روش های نمایش

روش های بخش بندی مشروح در فصل ۷ داده های خام را به شکل پیکسل های مرزی یا پیکسل های

1) Segmented

2) Describe

3) Extreme points

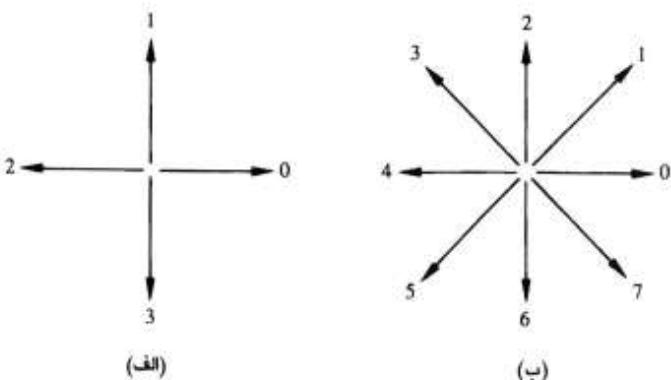
4) Concavity

درون ناحیه، نمایش می دهند. گرچه گاهی اوقات، به طور مستقیم از این داده ها برای به دست آوردن توصیف گرها (مانند تعیین بافت یک ناحیه) استفاده می شود، اما معمولاً داده ها را به نمایش هایی که در محاسبه توصیف گرها بسیار مفیدتر هستند، تبدیل می کنند. در این بخش چندین راه گوناگون برای این نوع نمایش را شرح می دهیم.

۱۰.۱.۸ رمزهای زنجیره‌ای

رمزهای زنجیره‌ای^۱ برای نمایش مرز با استفاده از یک دنباله پاره خط های متصل که طول و جهات مشخصی دارند، به کار می روند. معمولاً این نوع نمایش مبتنی بر اتصال چهارگانه یا هشت گانه پاره خط ها است. جهت هر قطعه با استفاده از روش های شماره زنی تغییر انواع موجود در شکل ۱.۸، رمز می شود.

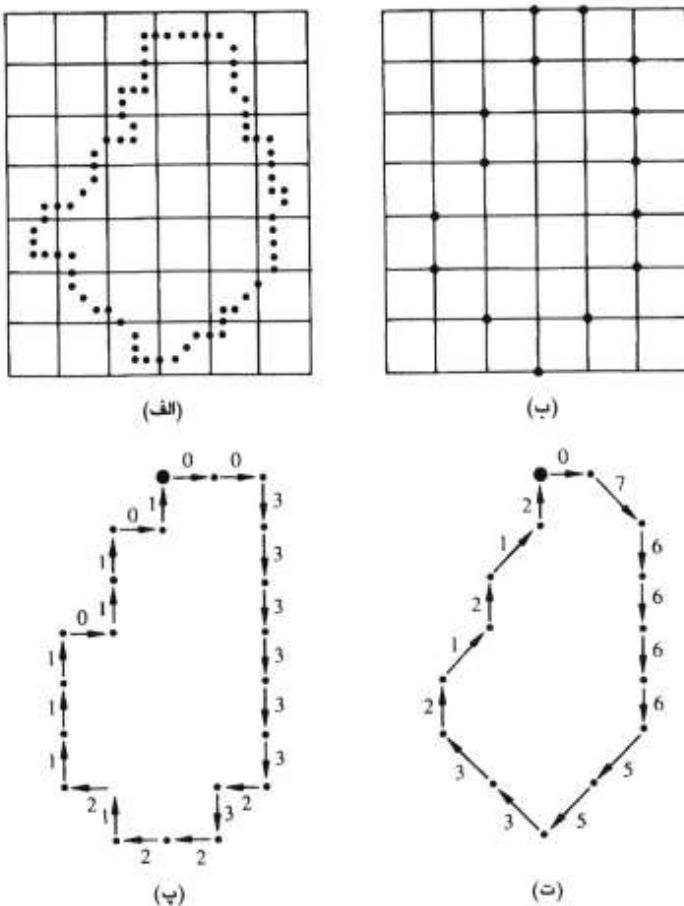
معمولآ تصاویر رقمه در قالب شبکه ای و با فواصل بکسان درجهات x و y تولید و پردازش می شوند، بنابراین می توان با دنبال کردن یک مرز، مثلاً درجهت عقربه ساعت و انتساب یک جهت به هر پاره خطی که دو پیکسل را به هم متصل می کند، یک رمز زنجیره‌ای تولید کرد. معمولاً به دو دلیل مهم، این روش به کارنمی رود: (۱) معمولاً زنجیره رمزهای حاصل بسیار طولانی است، (۲) هر اغتشاش کوچک در مرز، ناشی از نویز یا بخش بندی ناقص، باعث تغییراتی در رمز می شود که ممکن است هیچ ارتباطی با



شکل ۱.۸ جهات برای (الف) رمز زنجیره‌ای چهارجهتی و (ب) رمز زنجیره‌ای هشت‌جهتی

1) Chain code

۱۰.۸ روش های نمایش ۵۷۷



شکل ۲۰.۸ (الف) مرز رقمن که یک شبکه نمونه برداری مجدد روی آن قرار گرفته است؛ (ب) حاصل نمونه برداری مجدد؛ (پ) رمز زنجیره ای چهارجهتی؛ (ت) رمز زنجیره ای هشت جهتی

شکل مرز نداشته باشد.

یک راه متدائل برای غلبه بر این مشکلات نمونه برداری مجدد مرز با انتخاب فواصل شبکه ای بزرگتر، مانند شکل ۲.۸ (الف)، است. آنگاه در هنگام پیمایش مرز، همانند شکل ۲.۸ (ب)، به هر گره شبکه درشت تر بسته به نزدیکی مرز اولیه با آن گره، یک نقطه 'مرزی' متنسب می شود. در نهایت می توان مرز حاصل را با یک رمز چهارگانی یا هشت گانی، به ترتیب مانند اشکال ۲.۸ (پ) و (ت)، نمایش داد. در شکل ۲.۸ (پ)، نقطه 'شروع' نقطه بزرگتر بوده و مرز، کوتاه ترین مسیر خارجی چهارگانی مورد قبول در شبکه شکل ۲.۸ (ب) است. نمایش مرز شکل ۲.۸ (پ) رمز زنجیره ای "۰۱...۳۰" و

نمایش مرز شکل ۸. ۲. (ت) رمز "12...076" می باشد. همان طور که انتظار می رفت، دقت نمایش رمزی حاصل، به فواصل شبکه 'نمونه برداری' بستگی دارد.

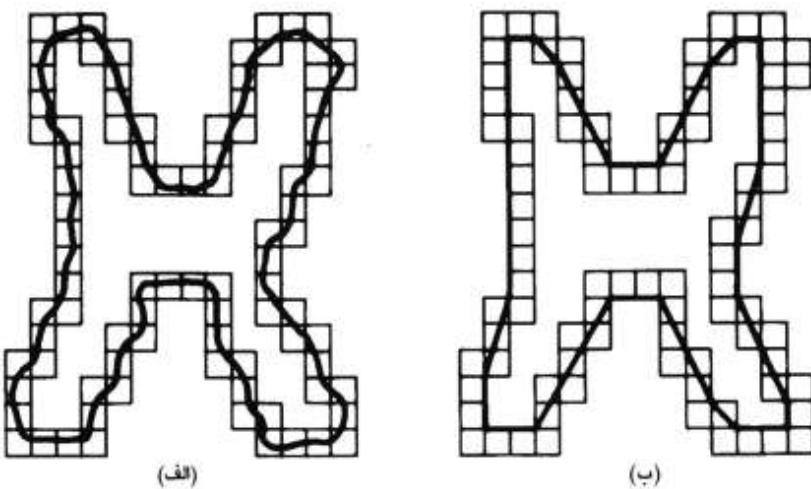
رمز زنجیره ای هر مرز به نقطه 'شروع آن' بستگی دارد. با این حال می توان با یک روال ساده، رمز را تراز^۱ کرد: با یک رمز زنجیره ای که با شروع از یک نقطه 'دلخواه' تولید شده است، به عنوان یک دنباله 'گردشی'^۲ از شماره های جهت رفتار می کنیم و نقطه 'شروع را مجددآ طوری تعریف می کنیم که دنباله 'اعداد حاصل، عدد صحیحی با بزرگی کمیته باشد. می توان با استفاده از تفاصل اول رمز زنجیره ای به جای خود رمز، آن را نسبت به گردش نیز تراز کرد. تنها با شمردن تعداد جهاتی (در خلاف حرکت عقربه ساعت) که دو عنصر مجاور رمز از هم فاصله دارند، می توان این تفاصل را به دست آورد. برای نمونه، تفاصل اول رمز زنجیره ای چهارجهتی "10103322" برابر "3133030" است. اگر بخواهیم با رمز، به عنوان یک دنباله 'گردشی' رفتار نماییم، عنصر اول تفاصل را نیز با استفاده از میزان انتقال بین اولین و آخرین جزء زنجیره محاسبه می کنیم. در مثال اخیر، حاصل "33133030" است. با تغییر ابعاد شبکه 'نمونه برداری'، می توان تراز سازی ابعاد را نیز انجام داد.

این تراز سازی ها تها وقتي دقیق هستند که مرزها با چرخش و تغییر مقیاس تصویر تغییر نکنند که در عمل به ندرت رخ می دهد. برای نمونه اگر یک شیئی در دو جهت متفاوت نمونه برداری شود، معمولاً اشکال مرزی متفاوتی به دست می آیند که درجه 'عدم تشابه آنها' متناسب با تفکیک تصویر است. طبق شرح بخش ۸. ۲. ۲.، می توان با انتخاب عناصر زنجیره ای که نسبت به فاصله 'بین پیکسل های تصویر رقemi بزرگ' هستند، یا با تغییر جهت شبکه 'نمونه برداری' مجدد در طول محورهای اصلی شیئی، این مسأله را کاهش داد.

۲۰.۱.۸ تقریب های چند ضلعی

هر مرز رقemi را می توان با یک چند ضلعی و با دقت دلخواه تقریب زد. در مورد منحنی های بسته، این تقریب وقتی دقیق است که تعداد اصلاح چند ضلعی برابر تعداد نقاط مرزی باشد، طوری که هر زوج از نقاط مجاور، یک ضلع چند ضلعی را معین نمایند. در عمل، هدف از تقریب چند ضلعی، به دست آوردن شکل کلی مرز با کمترین تعداد ممکن از پاره خط ها است، که در حالت کلی، کار ساده ای

۱۰.۸ روش های نمایش ۵۷۹



شکل ۱۰.۸ (الف) مرز شین که درون سلول ها محصور شده است؛ (ب) چند ضلعی با محیط کمینه

نیست و ممکن است به یک جستجوی تکراری طولانی تبدیل شود. با این حال چند روش تقریب چند ضلعی وجود دارند که خیلی پیچیده نبوده و نیاز به پردازش کمی دارند و برای کارهای پردازش تصویر مناسب هستند.

شرح این روش ها را با روشنی برای پیدا کردن چند ضلعی های با محیط کمینه شروع می کنیم. با یک مثال این روال به بهترین صورت شرح داده می شود. فرض کنید که مرز را در مجموعه ای از سلول های متصل، مانند شکل ۱۰.۸ (الف)، محصور کنیم. بهتر، این است که این حصار را به عنوان دو دیواره' متناظر با مرزهای بیرونی و درونی نوار سلول ها تصور کنیم و مرز شین را به عنوان یک نوار لاستیکی درون این دیواره ها در نظر بگیریم. اگر نوار لاستیکی قابل انقباض باشد، به صورت شکل ۱۰.۸ (ب) در می آید و یک چند ضلعی با محیط کمینه تولید می کند که بر شکل هندسی تعیین شده با نوار سلولی منطبق می شود. اگر هر سلول شامل تنها یک نقطه' مرز باشد، خطای بین مرز اولیه و تقریب نوار لاستیکی در هر سلول حداقل $\sqrt{2}$ خواهد بود، که 'فاصله' بین پیکسل ها است. با جایه جا کردن هر سلول طوری که مرکز آن بر پیکسل متناظر شن منطبق شود، کران بالای این خطای نصف می شود.

روش های ادغام^۱ مبتنی بر خطای یا سایر معیارها برای تقریب چند ضلعی مرز به کار رفته اند. یک

1) Merging

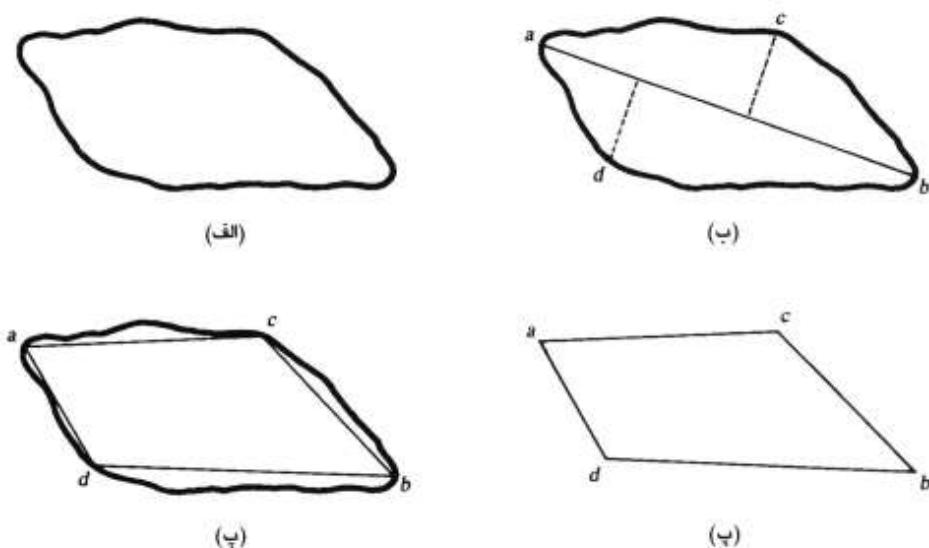
روش، ادغام نقاط مرز است تا آنجا که خطای کمینه^۱ مربع در انطباق خط برای نقاطی که تاکنون ادغام شده اند، از یک آستانه^۲ پیش تعیین شده تجاوز کند. وقتی چنین وضعیتی رخ دهد، مشخصه های خط ذخیره می شوند، خطابرا صفر قرار داده می شود و این روال تکرار می شود، بدین معنا که نقاط جدیدی از مرز ادغام می گردند تا آنجا که دوباره خطابرا، از مقدار آستانه تجاوز کند. در پایان، محل تلاقی پاره خط های مرزی مجاور هم، رؤوس چند ضلعی تعیین می شوند. یکی از مشکلات اصلی این روش این است که در حالت کلی، رؤوس، با انحنای های مرز (نظیر گوشها) متاظر نیستند، زیرا یک خط جدید، تنها وقتی شروع می شود که خطابرا از آستانه تجاوز کند. برای نمونه اگر یک خط مستقیم طویل را دنبال کنیم و آن خط به یک گوش برسد، قبل از این که خطای انطباق از مقدار آستانه تجاوز کند تعدادی از نقاط بعد از گوش (بسته به مقدار آستانه) جذب آن پاره خط خواهند شد. با این حال می توان با استفاده از عمل تقسیم همراه با فرآیند ادغام، این مشکل را کاهش داد.

یک راه برای تقسیم^۱ قطعه های مرزی، تقسیم متواالی هر قطعه به دو قسم است تا آنجا که معیار تعیین شده برآورده شود. برای نمونه، ممکن است مطلوب این باشد که فاصله^۳ عمودی بیشینه^۴ هر قطعه^۵ مرزی تا پاره خطی که دو نقطه^۶ انتهایش را متصل می کند، از یک آستانه^۷ معین تجاوز نکند. اگر از آستانه تجاوز کند، دورترین نقطه^۸ قطعه^۹ مرزی از پاره خط، یک رأس می شود، بنابراین قطعه^{۱۰} ابتدائی به دو قطعه تقسیم می شود. مزیت این روش جستجوی نقاط انحنای مهم است. برای یک مرز بسته، معمولاً بهترین نقاط شروع، دورترین دو نقطه^{۱۱} مرزی هستند. شکل ۸. ۴ (الف) مثالی از مرز شیخ را نشان می دهد، و شکل ۸. ۴ (ب) تقسیم مرز به دو قسم است با استفاده از دورترین نقاط راشن^{۱۲} را نشان می دهد. نقطه^{۱۳} دورترین نقطه^{۱۴} بالایی از خط^{۱۵} ab است. به طور مشابه نقطه^{۱۶} d دورترین نقطه^{۱۷} قطعه پایینی از آن است. شکل ۸. ۴ (پ) حاصل استفاده از روال تقسیم با آستانه ای که 25° برابر طول خط^{۱۸} ab است، راشن می دهد. چون فاصله^{۱۹} هیچ نقطه ای در قطعه^{۲۰} مرزی جدید (تا قطعه مستقیم الخط مربوط به آن) از این آستانه تجاوز نمی کند، این روال با چند ضلعی شکل ۸. ۴ (ت) پایان می یابد.

۳۰۱.۸ نمایه ها

«نمایه^{۲۱}» نمایش مرز به صورت تابع یک بعدی است و به روشهای گوناگونی می توان آن را تولید کرد.

۱۰.۸ روش های نمایش ۵۸۱

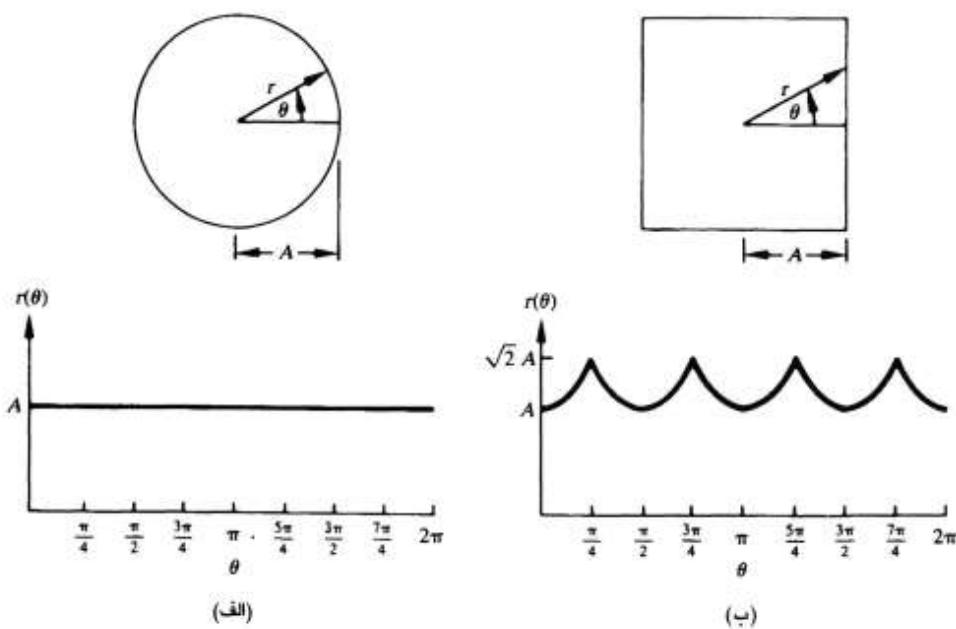


شکل ۱۰.۸ (الف) مرز اولیه؛ (ب) حاصل تقسیم مرز به چند قسمت براساس فواصل نقاط آن؛ (پ) متصل کردن رؤوس؛ (ت) چند ضلعی حاصل

یکی از ساده ترین راه رسم فاصله 'مرکز نقل تا مرز' به صورت تابعی از زاویه چرخش حول مرکز نقل، مانند شکل ۱۰.۵، است. با این حال صرف نظر از چگونگی تولید نمایه، هدف اصلی از آن، ساده کردن نمایش مرز به یک تابع یک بعدی است، که نسبت به مرز دو بعدی اولیه توصیف ساده تری دارد. نمایه های تولید شده با این روش، مستقل از انتقال هستند، اما به چرخش و مقیاس بندی بستگی دارند. با پیدا کردن راهی برای انتخاب یک نقطه 'شروع' که مستقل از جهت شکل باشد، می توان نمایه را نسبت به چرخش تراز کرد. یک راه برای انجام این کار انتخاب دورترین نقطه از مرکز نقل به عنوان نقطه 'شروع' است، به شرط آن که این نقطه یکتا و مستقل از انحرافات چرخشی موجود در شکل مورد نظر باشد. راه دیگر انتخاب دورترین نقطه 'محور ویژه اصلی' (بزرگتر) از مرکز نقل است (بخش ۳.۶ را ببینید). این روش به محاسبات بیشتری نیاز دارد اما مطمئن تر است زیرا جهت محور ویژه اصلی از روی ماتریس کوواریانس^۱ که براساس تمام نقاط منحنی^۲ تشکیل شده است، تعیین می شود. روش دیگر به دست آوردن رمز زنجیره ای مرز و سپس استفاده از روش مشروح در بخش ۱۰.۱ است، البته فرض براین است که رمز نگاری طوری انتخاب می شود که برگردشی بودن آن اثر نمی گذارد.

1) Covariance

2) Contour point



شکل ۵.۸ دو شکل مرزی ساده و نمایه های فاصله-زاویه متناظر با آنها. در (الف) $r(\theta)$ ثابت است، در حالی که در (ب) $r(\theta) = A \sec \theta$ است.

براساس این فرض ها که مقیاس بندی نسبت به هر دو محور یکنواخت است و این که نمونه برداری در بازه های مساوی از θ انجام می شود، تغییر ابعاد شکل، باعث تغییر در مقادیر دامنه 'نمایه' آن می شود. یک راه ساده 'تراز سازی' این است که تمام توابع طوری تغییر مقیاس داده شوند که همواره در محدوده 'یکسانی از مقادیر'، مثلاً $[0, 1]$ ، قرار داشته باشند. مزیت اصلی این روش سادگی آن است، اما این عیب جدی را نیز دارد که مقیاس کردن تابع تنها به دو مقدار کمینه و بیشینه آن بستگی دارد. اگر شکل ها نویز دار باشند، این نوع بستگی، ممکن است عامل انتقال خطأ از یک شیء به شیء دیگر باشد. یک راه مطمئن تر (اما پر محاسبه تر) تقسیم هر نمونه بر واریانس¹ نمایه است، البته با این فرض که واریانس صفر نباشد- مانند مورد شکل ۵.۸ (الف)- و آنقدر کوچک نیز نباشد که باعث مشکلات محاسباتی شود. واریانس یک عامل مقیاس کردن متغیر است که با تغییر ابعاد، تناسب معکوس دارد و اثری مانند کنترل خودکار بهره^۲ دارد. روش به کار رفته هر چه باشد، هدف این است که وابستگی به ابعاد از بین بروود و در همان حال چهار چوب اساسی شکل موج ها حفظ شود.

1) Variance

2) Automatic gain control

البته رسم نمودار فاصله بر حسب زاویه تنها راه تولید نمایه نیست. مثلاً می‌توان مرز را پیمود و زاویه^۱ بین خط مماس بر مرز و یک خط مرجع را به صورت تابعی از موقعیت در طول مرز رسم نمود. نمایه^۲ حاصل گرچه کاملاً متفاوت از منحنی (θ) است، حاوی اطلاعاتی درباره مشخصات پایه‌ای شکل می‌باشد. برای نمونه، قطعه‌های افقی منحنی، متناظر با خطوط مستقیم در طول مرز هستند، زیرا در آنها زاویه^۳ خط مماس ثابت است. گونه‌ای دیگر از این روش استفاده از تابع چگالی شیب^۴ به عنوان نمایه است. این تابع، بافت نگار مقادیر زاویه^۴ مماس است. چون بافت نگار معیاری از تمرکز مقادیر است، تابع چگالی شیب به قطعات مرزی دارای زوایای مماس ثابت (قطعه‌های مستقیم یا تقریباً مستقیم)، پاسخ بزرگی می‌دهد و در قطعه‌هایی که زوایای مماس تغییرات سریع دارند (گوش‌ها یا سایر خمیدگی‌های تند)، درجه‌های عمیقی دارد.

۱۰.۱.۸ قطعات مرزی

در اغلب موارد تقسیم مرز به قطعه‌ها مفید است. این تقسیم، پیچیدگی مرز را کاهش می‌دهد و بنابراین فرآیند توصیف را آسان می‌کند. این روش وقتی که مرز، دارای یک یا چند تعمیر مهم حاوی اطلاعات شکل باشد، جذایت ویژه‌ای دارد. در این مورد استفاده از بدنه^۵ محدب^۶ ناحیه^۷ محصور در مرز، ابزاری قوی برای تقسیم مرز است.

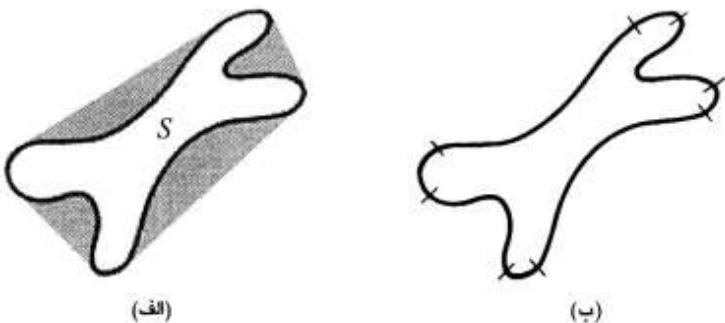
بدنه^۸ محدب H مربوط به یک مجموعه^۹ دلخواه S ، کوچکترین مجموعه^{۱۰} محدب حاوی S است. تفاضل مجموعه‌ای $S - H$ ، نقصان تحدب^{۱۱} D مربوط به مجموعه^{۱۲} S خوانده می‌شود. برای درک بهتر این نکته که چگونه می‌توان از این مفاهیم برای قسمت‌بندی مرز به قطعه‌های معنادار استفاده کرد، شکل ۱۰.۶ (الف) را در نظر بگیرید. این شکل یک شیء (مجموعه^{۱۳} S) و نقصان تحدب آن (ناحیه^{۱۴} را نشان می‌دهد. می‌توان با دنبال کردن منحنی S و علامت زدن نقاطی که در آنها انتقالی به سایه‌دار، را نشان می‌دهد. می‌توان یک چهارگانه^{۱۵} از نقصان تحدب وجود داشته است، مرز ناحیه را بخش بندی کرد. شکل ۱۰.۶ (ب) حاصل این عمل را نشان می‌دهد. توجه کنید که این روش مستقل از ابعاد و جهت ناحیه است.

در عمل به علل رقمی سازی، تویز و تغییرات ناشی از بخش بندی، مرزهای رقمی نامنظم می‌شوند. معمولاً این عوامل نقصان تحدبی را ایجاد می‌کنند که دارای اجزای کوچک بی‌معنا و به طور تصادفی

1) Slope density function

2) Convex hull

3) Convex deficiency



شکل ۶.۸ (الف) ناحیه S و نقصان تحدبش (سایه دار)؛ (ب) مرز تقسیم شده

پخش شده در سراسر مرز می باشد. به جای تلاش برای کاهش این بی نظمی ها با استفاده از پس پردازش، معمولاً تلاش می شود که قبل از بخش بندی مرز، آن را هموار^۱ کنند. چندین راه برای انجام این کار وجود دارند. یک راه، پیمایش مرز و جایگزینی مختصات هر پیکسل آن با متوسط مختصات m پیکسل همسایه اش در روی مرز می باشد. این روش بی نظمی های ریز را کاهش می دهد، اما وقت گیر بوده و مهار آن مشکل است. m های بزرگ ممکن است مرز را بیش از حد هموار نمایند، و m های کوچک نیز ممکن است در بعضی قطعه های مرزی ناکافی باشند. یک روش مطمئن تر این است که قبل از پیدا کردن نقصان تحدب یک ناحیه، تقریب چند ضلعی آن را، همانند بخش ۱.۲.۸، به دست آوریم. صرف نظر از روش به کار رفته برای هموار سازی مرز، اغلب مرزهای رقمی موردعلاقه، چند ضلعی های ساده (چند ضلعی های نامقطع با خود) هستند. [Graham & Yao 1983] الگوریتمی برای به دست آوردن بدنه' محدب چنین چند ضلعی هایی ارائه کرده اند.

مفاهیم بدنه' محدب و نقصان تحدب هر دو در توصیف ناحیه و مرز آن مفیدند. برای مثال یک ناحیه را می توان با مساحت آن، مساحت نقصان تحدبش، تعداد اجزا در نقصان تحدب، محل نسبی این اجزا و غیره توصیف کرد. در بخش ۱.۴.۴ الگوریتمی برای پیدا کردن بدنه' محدب یک ناحیه ارائه می دهیم. مراجع مذکور در انتهای این فصل حاوی سایر روش ها هستند.

۵.۱.۸ اسکلت یک ناحیه

یک راه مهم برای نمایش شکل ساختاری یک ناحیه' صفحه ای، ساده کردن آن ناحیه به یک گراف

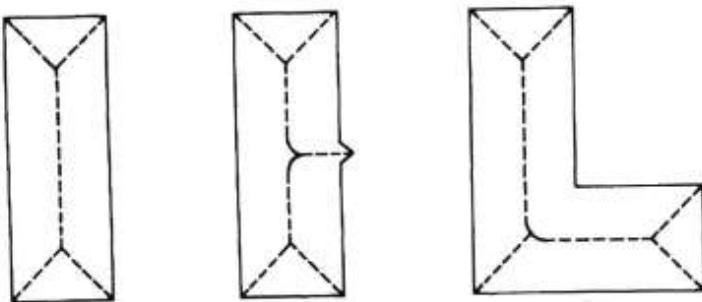
1) Smooth

۱۰۸۵ روش های نمایش

است. می توان با به دست آوردن اسکلت^۱ یک ناحیه بوسیله 'یک الگوریتم باریک سازی'^۲ (که اسکلتی سازی^۳ نیز خوانده می شود) این ساده سازی را انجام داد. روال های باریک سازی در محدوده' وسیعی از کارهای پردازش تصویر، از بازرسی خودکار صفحه های مدار چاپی تا شمارش تارهای پنبه' نسوز در صافی های هوا، نقشی محوری دارند.

اسکلت یک ناحیه را می توان با تبدیل محور میانی^۴ (MAT) پیشنهاد شده بوسیله' Blum[1967] تعریف کرد. تبدیل MAT برای ناحیه R با مرز B به صورت زیر است. برای هر نقطه' p از R ، نزدیکترین همسایه اش در B را پیدا می کنیم. اگر بیش از یک نقطه برای p پیدا شود، گفته می شود که p به محور میانی (اسکلت) R تعلق دارد. مفهوم "نزدیکترین" به تعریف فاصله بستگی دارد (بخش ۴.۵)، و بنابراین نتایج تبدیل MAT متأثر از معیار فاصله' منتخب هستند. شکل ۷.۸.۷ مثالهایی از کاربرد فاصله' اقلیدسی را نمایش می دهد.

گرچه تبدیل MAT یک ناحیه، اسکلت ظاهرآ خوبی را نتیجه می دهد، معمولاً اجرای مستقیم آن مشکلات محاسباتی دارد. تبدیل MAT شامل محاسبه' فاصله های هر نقطه' داخلی ناحیه تا تمام نقاط روی مرز آن است. الگوریتم های متعددی برای افزایش بازده' محاسباتی این تبدیل پیشنهاد شده اند. نوع آن این الگوریتم ها باریک ساز هستند که با روش تکراری، نقاط لبه' ناحیه را حذف می کنند؛ البته تحت این قیود که حذف این نقاط (۱) نقاط پایانی را حذف نکند، (۲) اتصال را قطع نکند، و (۳) باعث سایش^۵ بیش از حد نشود.



شکل ۷.۸ (الف) محورهای میانی سه ناحیه ساده

1) Skeleton 2) Thinning 3) Skeletonizing 4) Medial axis transformation 5) Erosion

در این بخش الگوریتمی برای باریک سازی نواحی دودویی ارائه می کنیم. فرض می شود که نقاط ناحیه مقدار "۱" و نقاط زمینه مقدار "۰" دارند. این روش شامل نکار متوالی دو مرحله' پایه ای زیر در نقاط منحنی است، که یک نقطه' منحنی هر پیکسل با مقدار "۱" است که حداقل یک همسایه' هشت گانه با مقدار "۰" داشته باشد. طبق تعریف همسایگی هشت گانه در شکل ۸.۸، مرحله' یک، هر نقطه' p از منحنی را که شرایط زیر در همسایگی هشتگانه اش برقرار باشد، برای حذف شدن علامت می زند:

$$(الف) \quad 2 \leq N(p_1) \leq 6;$$

$$(ب) \quad S(p_1) = 1;$$

$$(ب') \quad p_2 \cdot p_4 \cdot p_6 = 0;$$

$$(ت) \quad p_4 \cdot p_6 \cdot p_8 = 0;$$

که $N(p_1)$ تعداد همسایه های غیر صفر p_1 است؛ یعنی،

$$N(p_1) = p_2 + p_3 + \dots + p_8 + p_9 \quad (۲-۱. ۸)$$

p_9	p_2	p_3
p_8	p_1	p_4
p_7	p_6	p_5

شکل ۸.۸ همسایگی مورد استفاده در الگوریتم باریک سازی

و $S(p_1)$ تعداد تغییر حالت های "۰" به "۱" در دنباله' مرتب $p_2, p_3, \dots, p_8, p_9$ می باشد. برای مثال در شکل ۸.۹. $N(p_1) = 4$ و $S(p_1) = 3$. در مرحله' دو شرایط (الف) و (ب) مانند مرحله' یک هستند، اما شرایط (ب') و (ت) به صورت زیر تغییر می یابند.

$$(ب') \quad p_2 \cdot p_4 \cdot p_8 = 0; \quad (۸-۱. ۳)$$

$$(ت') \quad p_2 \cdot p_6 \cdot p_8 = 0.$$

۱۰.۸ روش های نمایش ۵۸۷

مرحله' یک در تمام پیکسل های مرزی ناحیه' دودویی تحت بررسی، اجرا می شود. اگر یک یا چند تا از شرایط (الف)-(ت) در همسایگی یک پیکسل نقض شوند، مقدار آن پیکسل تغییر نمی کند. اگر تمام شرایط برآورده شوند، این پیکسل برای پاک شدن علامت زده می شود، اما تا پایان پردازش تمام نقاط مرزی، پاک نمی شود. این تعویق، از تغییر ساختار داده ها در حین اجرای الگوریتم جلوگیری می کند. پس از اتمام مرحله' یک، نقاطی که علامت خورده اند، پاک می شوند (به "۰" تغییر می یابند). سپس مرحله' دو دقیقاً به همان روش مرحله' یک، اجرا می شود.

بنابراین هر بار تکرار الگوریتم باریک سازی شامل مراحل زیر است: (۱) اجرای مرحله' یک برای علامت زدن نقاط مرزی جهت پاک شدن؛ (۲) حذف نقاط علامت خورده؛ (۳) اجرای مرحله' دو برای علامت زدن نقاطی دیگر برای پاک شدن؛ و (۴) حذف نقاط علامت خورده. این روال بارها اجرا می شود تا آنجا که نتوان هیچ نقطه' دیگری را پاک کرد. در این زمان الگوریتم پایان می یابد و اسکلت ناحیه به دست آمده است.

وقتی نقطه' منحنی p تنها یک یا هفت همسایه' هشت گانه با مقدار "۱" داشته باشد، شرط (الف) نقض می شود. داشتن تنها یک همسایه' با مقدار "۱" دال براین است که p نقطه' پایان یک خط اسکلتی است و نباید پاک شود. اگر p دارای هفت همسایه با مقدار "۱" باشد، پاک کردن آن باعث سایش ناحیه به داخل می شود. شرط (ب) در نقاط روی خطوط با ضخامت یک پیکسل نقض می شود. بنابراین این شرط از انفصال قطعات اسکلت در حین اجرای باریک سازی جلوگیری می کند. دو شرط (پ) و (ت) در شرایط زیر به طور همزمان برآورده می شوند: $(p_1 = 0 \text{ یا } p_2 = 0) \text{ یا } (p_3 = 0 \text{ یا } p_4 = 0)$. بنابراین با توجه به همسایگی شکل ۸.۸، نقطه ای که شرایط فوق و نیز شرایط (الف) و (ب) را برآورده نماید، یک نقطه مرزی سمت شرق یا جنوب، یا یک نقطه' گوشه شمال غرب مرز است. در هر دو مورد نقطه' p جزء اسکلت نیست و باید حذف شود. به طور مشابه دو شرط (پ) و (ت) با این شرایط برآورده

0	0	1
1	p_1	0
1	0	1

شکل ۹.۸ نمایش شرایط (الف) و (ب) از معادله' (۸.۱-۱). در این مورد $N(p_1) = 3$ و $S(p_1) = 4$.

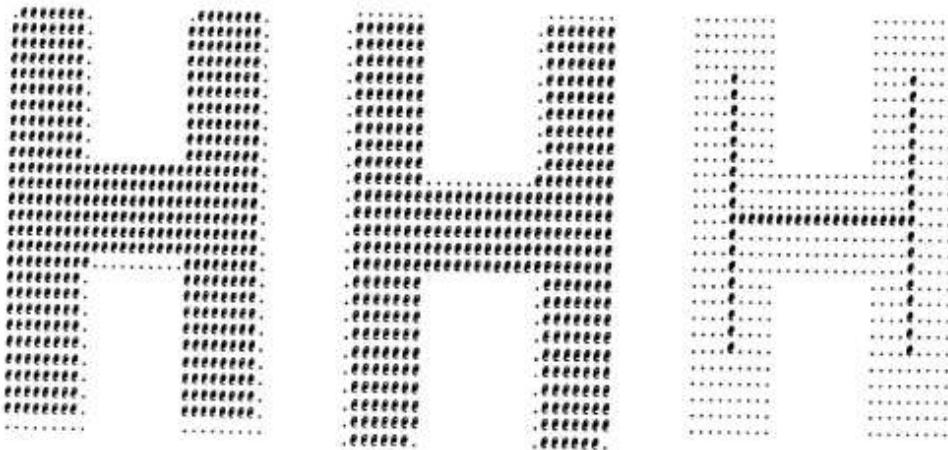
می شوند: $p_6 = 0$ یا $p_2 = 0$ (یا $p_4 = 0$) . این مقادیر متناظر با نقاط مرزی سمت شمال یا غرب، یا یک نقطه 'گوشه' جنوب شرق هستند. توجه کنید که در نقاط گوشه 'شمال شرق'، $p_2 = 0$ است و بنابراین شرایط (p) و (t) و نیز شرایط (p') و (t') برآورده می شوند.

مثال: شکل ۱۰.۸ (الف) حاصل اجرای مرحله 'یک الگوریتم باریک سازی در مرز یک ناحیه' ساده را نشان می دهد. نقطه ها نشان دهنده 'نقاط علامت خورده' هستند که در انتهای مرحله 'یک حذف شده' اند. شکل ۱۰.۸ (ب) حاصل مرحله 'دورانشان می دهد، و شکل ۱۰.۸ (ب) اسکلت به دست آمده پس از چند بار تکرار این دو مرحله 'متوالی را نشان می دهد. در شکل ۱۱.۸ اسکلت یک ناحیه دارای خواص نامنظم تر، دیده می شود.

۲.۸ توصیف گرهای مرز

۱.۲.۸ چند توصیف گر ساده

طول^۱ منحنی یکی از ساده ترین توصیف گرهای آن است. تنها با شمارش تعداد پیکسل ها در طول منحنی تقریبی سرانگشتی از طول آن به دست می آید. در رمز زنجیره ای یک منحنی که با استفاده از شبکه ای با



شکل ۱۰.۸ (الف) حاصل مرحله 'یک الگوریتم باریک سازی در اولین تکرار آن در یک ناحیه؛ (ب) نتیجه' مرحله 'دو؛ (ج) نتیجه' نهایی. (از Zhang & Suen [1984])

1) Length



شکل ۲۰.۸ مثال دیگری از باریک سازی (از Zhang & Suen [1984] از)

فاصل واحد در دو جهت متعامد به دست آمده است، جمع تعداد اجزای عمودی و افقی به اضافه $\sqrt{2}$ برابر تعداد اجزای قطری طول دقیق منحنی را می دهد.
قطر^۱ مرز B به صورت زیر تعریف می شود:

$$\text{Diam } (B) = \max_{i,j} [D(p_i, p_j)] \quad (۲۰.۸)$$

که D یک معیار فاصله است (بخش ۴.۵.۱ را بینید) و p_i و p_j دو نقطه' مرز هستند. طول قطر و جهت خطی که دو نقطه' انتهای آن را به هم متصل می کند (این خط محور اصلی^۲ مرز خوانده می شود)، توصیف گرهای مفیدی برای مرز هستند.

انحنا^۳ به عنوان آهنگ تغییر شیب تعریف می شود. در حالت کلی به دست آوردن مقادیر قابل اعتماد برای انحنا در یک نقطه از مرز رقی، مشکل است، زیرا این مرزها ناهموار هستند. با این حال استفاده از تفاضل بین شبیه های قطعات مرزی مجاور (که به صورت خطوط راست نمایش داده شده اند) به عنوان

1) Diameter

2) Major axis

3) Curvature

توصیف گر انحنا در نقطه 'تلاقي آن دو قطعه، ممکن است مفید باشد. مثلاً رئوس مرزهای اشکال ۸. ۳. (ب) و ۸. ۴. (ت) به خوبی مناسب توصیف انحنا است. وقتی این مرز در جهت حرکت عقربه' ساعت پیموده می شود، اگر تغییر شیب در آن نامنفی باشد، نقطه 'رأس ۲ جزئی از قطعه' محدب ۱ محسوب می شود؛ در غیر این صورت ۲ متعلق به یک قطعه 'مقعر' محسوب می شود. با استفاده از محدوده های تغییر شیب می توان انحنا در یک نقطه را دقیق تر توصیف کرد. برای نمونه اگر تغییر شیب در ۲ کمتر از ۱۰ درجه باشد می توان آن را بخشی از یک قطعه 'تقریباً مستقیم' محسوب کرد، و اگر تغییر شیب بیش از ۹۰ درجه باشد، آن را یک نقطه 'گوش' ۳ محسوب کرد. با این حال توجه کنید که در استفاده از این توصیف گرها باید دقت شود، زیرا تعییر آنها به طول هر قطعه نسبت به طول کلی مرز، بستگی دارد.

۲۰۲۰۸ اعداد شکلی

همان طور که در بخش ۸. ۱. ۱. توضیح داده شد، تفاضل اول رمز زنجیره ای مرز به نقطه 'شروع آن' بستگی دارد. عدد شکلی ۴ این مرز، براساس رمز چهار جهتی شکل ۸. ۱. (الف)، به صورت تفاضل اول با کوچکترین بزرگی تعریف می شود. مرتبه' یک عدد شکلی، ۱۱، به عنوان تعداد ارقام به کار رفته در نمایش آن تعریف می شود. گذشته از این، ۱۱ برای یک مرز بسته، زوج است، و مقدار آن تعداد اشکال متفاوت ممکن را محدود می کند. شکل ۸. ۱۲. تمام اشکال ممکن از مرتبه های ۴، ۶، و ۸ را همراه با رمز زنجیره ای، تفاضل اول، و اعداد شکلی متناظر شان نمایش می دهد. توجه کنید که طبق شرح بخش ۸. ۱. ۱، تفاضل های اول با درنظر گرفتن رمزهای زنجیره ای به عنوان دنباله های گردشی محاسبه شده اند. گرچه تفاضل اول یک رمز زنجیره ای مستقل از چرخش است، در حالت کلی رمز حاصل به جهت شبکه بستگی دارد. یک راه برای تراز سازی جهت شبکه به صورت زیر است.

همان طور که قبلآ بیان شد، محور اصلی^۵ مرز، پاره خطی است که دورترین دو نقطه' آن را به هم متصل می کند. محور فرعی^۶ عمود بر محور اصلی است و طول آن به اندازه ای است که بتوان با مستطیلی با اضلاع موازی با محورهای اصلی و فرعی کاملاً مرز را احاطه کرد. نسبت طول محور اصلی به محور فرعی کشیدگی^۷ مرز خوانده می شود، و مستطیلی که اخیراً ذکر شد، مستطیل پایه^۸ خوانده می شود. در

1) Convex

2) Convave

3) Corner

4) Shape number

5) Major axis

6) Minor axis

7) Eccentricity

8) Basic rectangle

۵۹۱ توصیف گرهای مرز ۲۰.۸

	مرتبه ۴	مرتبه ۶
رمز زنجیره ای	0 3 2 1	0 0 3 2 2 1
تفاضل	3 3 3 3	3 0 3 3 0 3
عدد شکلی	3 3 3 3	0 3 3 0 3 3

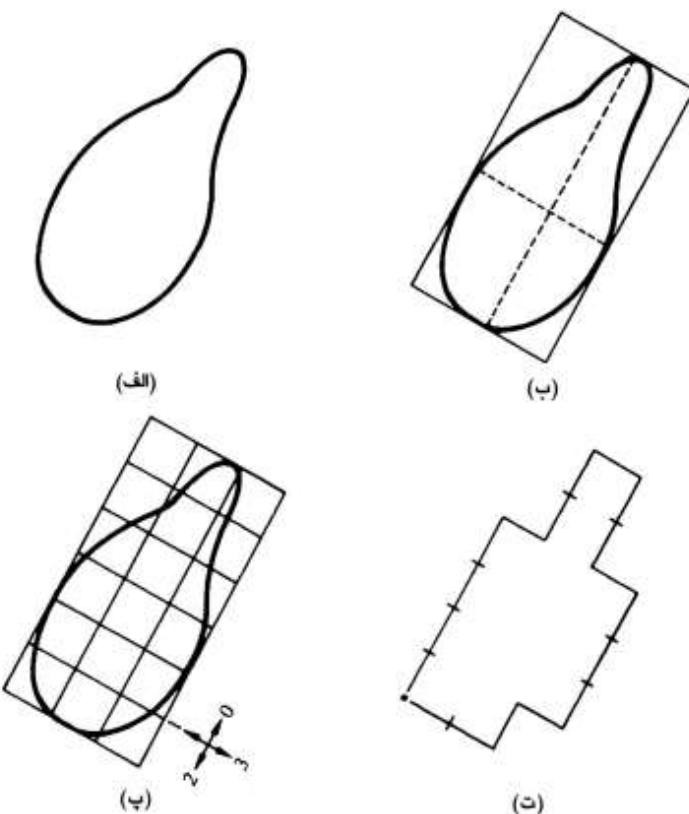
	مرتبه ۸	مرتبه ۸	مرتبه ۸
رمز زنجیره ای	0 0 3 3 2 2 1 1	0 3 0 3 2 2 1 1	0 0 0 3 2 2 2 1
تفاضل	3 0 3 0 3 0 3 0	3 3 1 3 3 0 3 0	3 0 0 3 3 0 0 3 0
عدد شکلی	0 3 0 3 0 3 0 3	0 3 0 3 3 1 3 3	0 0 3 3 0 0 3 3

شکل ۱۲۰.۸ تمام شکل های با مرتبه ۴، ۶ و ۸. در اینجا جهات شکل ۱۰.۸ (الف) استفاده می شوند، و نقطه نشانه محل شروع است.

اغلب موارد با هم راستا کردن^۱ شبکه رمز زنجیره ای با اضلاع مستطیل پایه، عدد شکلی یکتاوی بددست می آید.

در عمل برای یک مرتبه 'شکلی مطلوب مستطیل با مرتبه' n که کشیدگی آن بهترین تقریب از کشیدگی مستطیل پایه باشد، را پیدا کرده و از این مستطیل برای تعیین ابعاد شبکه استفاده می کنیم. برای مثال در حالت $n=12$ ، تمام مستطیل های با مرتبه ۱۲ (یعنی مستطیل هایی که محیط آنها ۱۲ است) عبارتند از: 4×4 ، 3×3 ، و 5×1 . اگر کشیدگی مستطیل 4×2 بهترین تقریب از کشیدگی مستطیل پایه یک مرز معین باشد، یک شبکه 4×2 روی مستطیل پایه ایجاد می کنیم و با استفاده از روال مشروح در بخش ۱۰.۸ رمز زنجیره ای مرز را به دست می آوریم. همان طور که قبلاً بیان شد، عدد شکلی از تفاضل اول این رمز به دست می آید. گرچه معمولاً مرتبه 'عدد شکلی حاصل برابر n است، اما گاهی اوقات به علت نوع فواصل شبکه ای منتخب، در مرزهای دارای تورفتگی های قابل مقایسه با این فواصل شبکه ای،

1) Aligning



0 0 0 0 3 0 0 3 2 2 3 2 2 2 1 2 1 1 : رمز زنجیره ای

3 0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 3 0 : تفاضل

0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 3 0 3 : عدد شکلی

شکل ۱۳.۸ مراحل تولید عدد شکلی

اعداد شکلی با مرتبه 'بزرگتر از n ' بدست می آیند. تحت این شرایط مستطیلی با مرتبه 'کوچکتر از n ' را مشخص کرده و این روال را تکرار می کنیم تا این که عدد شکلی حاصل، از مرتبه ' n ' گردد.

مثال: فرض کنید که برای رمز نشان داده شده در شکل ۱۳.۸ (الف)، $n = 18$ تعیین شود. برای به دست آوردن یک عدد شکلی با این مرتبه باید مراحل فوق اجرا شوند. همان طور که شکل ۱۳.۸ (ب) نشان می دهد، مرحله 'اول پیدا کردن مستطیل' پایه است. نزدیکترین مستطیل از مرتبه '۱۸' به شکل رمز، یک مستطیل 6×3 است که باید مستطیل پایه مانند شکل ۱۳.۸ (پ) تقسیم شود که در این حالت جهات رمز زنجیره ای با جهات شبکه' حاصل هم راستا هستند. مرحله 'نهایی' بدست آوردن رمز زنجیره ای و

۲۰.۸ توصیف گرهای مرز ۵۹۳

استفاده از تفاضل اول آن برای محاسبه' عدد شکلی است که در شکل ۸.۱۳(ت) نشان داده شده است. □

۳.۲.۸ توصیف گرهای فوریه

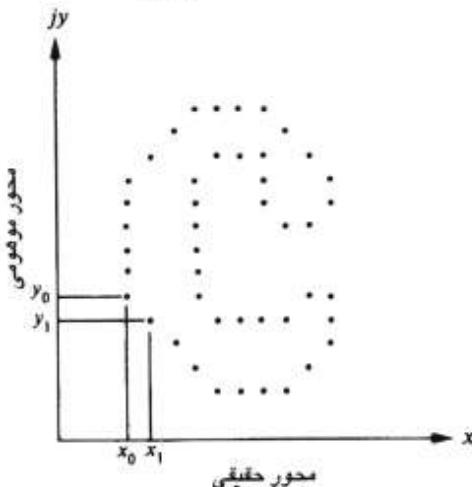
شکل ۸.۱۴ یک مرز رقمنی N نقطه ای در صفحه' xy را نمایش می دهد. با شروع از یک نقطه' دلخواه (x_0, y_0) ، در حین پیمودن مرز، مثلاً در جهت خلاف عقربه های ساعت، به زوج های مختصات $(x_1, y_1), (x_2, y_2), \dots, (x_{N-1}, y_{N-1})$ بخورد می کنیم. این مختصات را می توان به شکل ۸.۸ $x(k) = x_k$ و $y(k) = y_k$ بیان کرد. در این نمادگذاری می توان مرز را به صورت دنباله' مختصات $s(k) = [x(k), y(k)]$ به ازای $k = 0, 1, 2, \dots, N-1$ نمایش داد. گذشته از این می توان هر زوج مختصات را یک عدد مختلط در نظر گرفت طوری که به ازای $k = 0, 1, 2, \dots, N-1$

$$s(k) = x(k) + jy(k) \quad (2-2.8)$$

یعنی محور x به عنوان محور حقیقی و محور y به عنوان محور موهومی دنباله' اعداد مختلط در نظر گرفته می شود. گرچه تعبیر دنباله با قالب دیگری بیان شده است، اما طبیعت خود مرز، تغییر نیافته است. البته این نوع نمایش این مزیت بزرگ را دارد که یک مسئله' دو بعدی را به یک بعدی ساده می کند.

تبديل فوریه' گسته (DFT) $s(k)$ بازای $k = 0, 1, 2, \dots, N-1$ عبارتست از

$$a(u) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) \exp[-j2\pi uk/N] \quad (3-2.8)$$



شکل ۸.۱۴ مرز رقمنی و نمایش آن به صورت دنباله' مختلط. نقاط دلخواه (x_0, y_0) و (x_1, y_1) دو نقطه اول دنباله هستند

ضرایب مختلف $a(u)$ توصیف گرهای فوریه^۱ مرز خوانده می شوند. تبدیل فوریه^۲ معکوس (u) ها را بازیابی می کند. یعنی به ازای $k = 0, 1, 2, \dots, N-1$

$$s(k) = \sum_{u=0}^{N-1} a(u) \exp[j2\pi uk/N] \quad (4-2. A)$$

با این حال فرض کنید که تنها از M ضریب اول $a(u)$ برای بازیابی استفاده می شود. این فرض معادل این است که در معادله^۳ $s(k) = \sum_{u=0}^{M-1} a(u) \exp[j2\pi uk/N]$ قرار دهیم. پس تقریب $s(k)$ ، به ازای $k = 0, 1, 2, \dots, N-1$ عبارتست از:

$$s(k) = \sum_{u=0}^{M-1} a(u) \exp[j2\pi uk/N] \quad (5-2. A)$$

گرچه برای بدست آوردن هر جمله^۴ $s(k)$ تنها از M جمله استفاده می شود، باز هم k در محدوده^۵ ۰ تا $N-1$ تغییر می کند. یعنی همان تعداد نقاط در مرز تقریبی وجود دارند، اما نه به زیادی تعداد جمله هایی که در بازسازی یک نقطه به کار می روند. اگر تعداد نقاط مرز بزرگ باشد، معمولاً M توان صحیحی از دو انتخاب می شود تا این که بتوان با استفاده از الگوریتم FFT محاسبه^۶ توصیف گرهای را تسريع کرد. از شرح های تبدیل فوریه در فصول ۳ و ۴ بیاد آورید که مؤلفه های فرکانس بالا جزئیات ریز و مؤلفه های فرکانس پایین شکل کلی مرز را تعیین می کنند. بنابراین هر چه M کوچکتر باشد، جزئیات بیشتری از مرز از دست می روند.

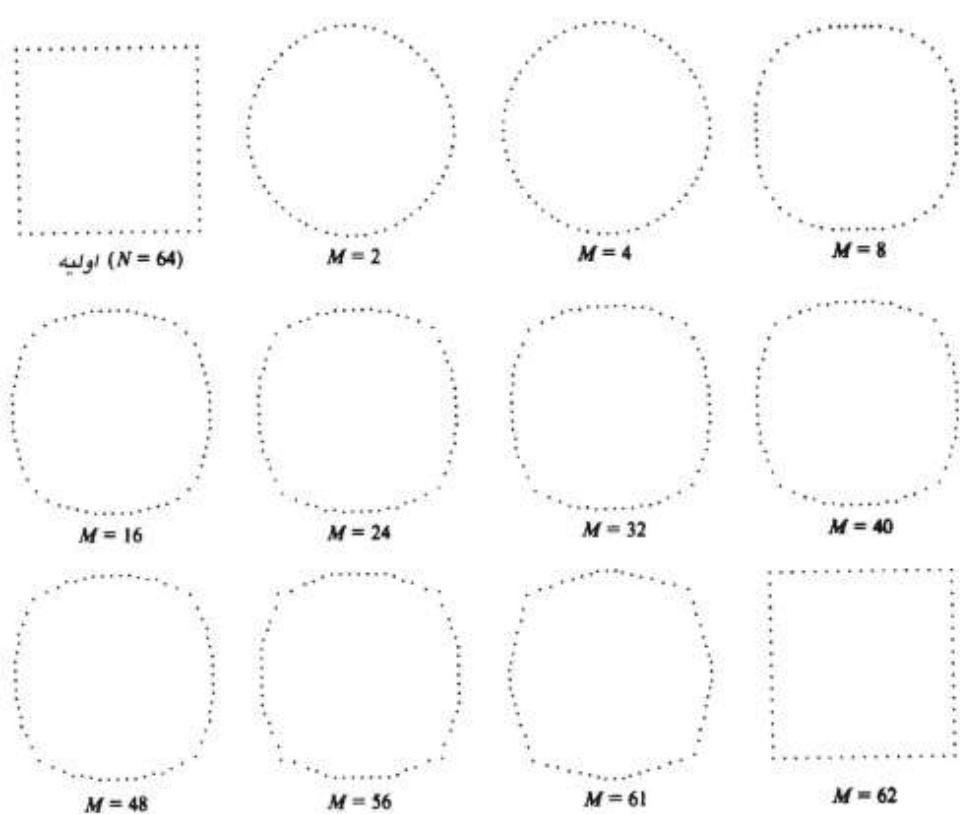
مثال: شکل ۸.۱۵ یک مرز مربعی را که مشکل از $N=64$ نقطه است و همچنین نتایج کاربرد معادله^۷ $s(k) = \sum_{u=0}^{M-1} a(u) \exp[j2\pi uk/N]$ برای بازسازی این مرز به ازای مقادیر مختلف M را نشان می دهد. توجه کنید که مقدار M باید به حدود ۸ افزایش یابد تا مرز بازسازی شده از حالت دایروی خارج و به مریع شبیه شود. اکنون به چگونگی ظهور گوشه در مرز که به ازای M در حدود ۵۶ رخ می دهد، توجه کنید، که به ازای این مقادیر M نقاط گوشه کم در مرز ظاهر می شوند. در نهایت توجه کنید که وقتی $M=61$ شود، منحنی ها به خطوط راست تبدیل می شوند، و با یک واحد افزایش M تقریباً همان شکل مرز اولیه به دست می آید. بنابراین با داشتن تعداد کمی از ضرایب پایین مرتبه، تنها می توان شکل کلی^۸ مرز را به دست آورد، اما برای تعیین دقیق ویژگی های تیز نظیر گوشه ها و خطوط راست تعداد زیادی از ضرایب بالا مرتبه مورد نیاز هستند. با

1) Fourier descriptor

2) Gross shape

۲۰.۸ توصیف گرهای مرز

۵۹۵

شکل ۲۰.۸ چند مثال بازسازی مرز از توصیف گرهای فوریه آن بازای مقادیر مختلف M .

توجه به نقشی که هریک از مؤلفه های فرکانس پایین و فرکانس بالا در تعیین شکل ناحیه ایفا می کنند، این نتیجه مورد انتظار بود.

همان طور که در مثال قبل روشن شد، می توان با تعداد کمی از توصیف گرهای فوریه شکل کلی مرز را بدست آورد. این خاصیت ارزشمند است، زیرا این ضرایب، حامل اطلاعات شکلی هستند. بنابراین می توان از آنها به عنوان معیار تشخیص بین شکل مرزهای متفاوت استفاده کرد، که در فصل ۹ با مقداری تفصیل آن را شرح می دهیم.

چند بار بیان کرده ایم که توصیف گرها باید تا حد ممکن نسبت به انتقال، چرخش و تغییر مقیاس غیر حساس باشند. در مواردی که نتایج به ترتیب پردازش نقاط، بستگی دارند، یک قید دیگر این است که

تصویف گر نسبت به نقطه 'شروع هم غیر حساس باشد. توصیف گرهای فوریه مستقیماً نسبت به این تغییرات هندسی غیر حساس نیستند، اما می توان تغییرات توصیف گرهای را با روابط ساده ای بیان کرد. مثلاً چرخش را در نظر بگیرید و از مباحث ریاضی مقدماتی به یاد آورید که چرخش یک نقطه با زاویه θ حول مبدأ صفحه 'مختلط را می توان با ضرب آن نقطه در $e^{j\theta}$ بیان کرد. با اجرای چنین کاری در هر نقطه $s(k)$ ، کل دنباله، حول مبدأ می چرخد. دنباله 'چرخش یافته' $s(k)e^{j\theta}$ است، که توصیف گرهای فوریه اش به ازای $1 - u = 0, 1, 2, \dots, N$ عبارتنداز:

$$\begin{aligned} a_s(u) &= \frac{1}{N} \sum_{k=0}^{N-1} s(k) \exp[j\theta] \exp[-j2\pi uk/N] \\ &= a_s(u) e^{j\theta} \end{aligned} \quad (6-2. 8)$$

بنابراین چرخش بر تمام ضرایب اثر یکسانی به صورت یک جمله 'ثابت ضرب شونده' دارد.

جدول ۸.۱ توصیف گرهای فوریه برای یک دنباله 'مرزی' $s(k)$ که تحت چرخش، انتقال، تغییر مقیاس، و تغییر نقطه 'شروع قرار گرفته است، را نشان می دهد. نماد Δ_x, Δ_y به صورت $\Delta_{xy} = \Delta_x + j\Delta_y$ تعریف می شود، بنابراین عبارت $s(k) = s(k) + \Delta_{xy}$ را می توان به صورت زیر نیز نوشت:

$$s(k) = [x(k) + \Delta_x] + j[y(k) + \Delta_y]. \quad (7-2. 8)$$

جدول ۱۰.۸ چند خاصیت پایه ای توصیف گرهای فوریه

تبدیل	مرز	تصویف گر فوریه
همانی	$s(k)$	$a_s(u)$
چرخش	$s_s(k) = s(k)e^{j\theta}$	$a_s(u) = a_s(u)e^{j\theta}$
انتقال	$s_s(k) = s(k) + \Delta_x$	$a_s(u) = a_s(u) + \Delta_x \delta(u)$
تغییر مقیاس	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a_s(u)$
نقطه 'شروع'	$s_p(k) = s(k - k_0)$	$a_s(u) = a_s(u) e^{-j2\pi k_0 u/N}$

به عبارت دیگر انتقال، معادل افزودن یک مقدار ثابت به مختصات تمام نقاط مرز است. توجه کنید که انتقال، هیچ اثری بر توصیف گرهاندارد، مگر برای $k = 0$ که اثر آن به صورت تابع ضریب $(k)\delta$ است.

[†] تبدیل فوریه یک ثابت یک ضریب در محل مبدأ است. به یاد آورید که تابع ضریب در هر جای دیگری صفر است.

۵۹۷ ۲.۱ توصیف گرهای مرز

در نهایت عبارت $s_p(k) = s(k - k_0)$ به معنای تعریف مجدد دنباله به صورت

$$s_p(k) = x(k - k_0) + jy(k - k_0) \quad (8-2.8)$$

است که تنها، نقطه 'شروع' دنباله را از $k = k_0$ به $k = 0$ تغییر می دهد. آخرین سطر جدول ۸.۱ نشان می دهد که تغییر در نقطه 'شروع' به طریقی متفاوت (اما معلوم) بر هر یک از توصیف گرها اثر می گذارد، زیرا عبارت ضرب شونده در $a(u)$ به u بستگی دارد.

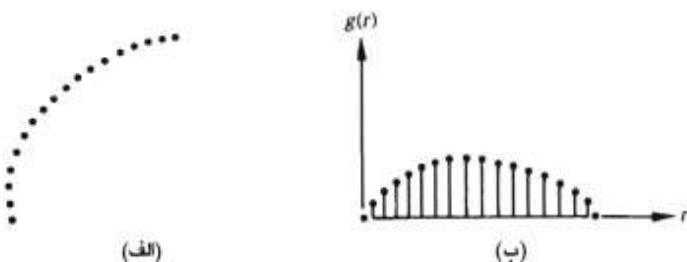
۴.۲.۸ گشتاورها

شكل قطعه های مرزی (و نمایه ها) را می توان با استفاده از گشتاورها به طور کمی توصیف کرد. برای ملاحظه چگونگی اجرای این عمل، شکل ۸.۱۶ (الف) که یک قطعه 'مرزی' را نشان می دهد، و شکل ۸.۱۶ (ب) که یک قطعه 'مرزی' را به صورت تابع یک بعدی $g(r)$ از متغیر دلخواه r نمایش می دهد، در نظر بگیرید. اجازه دهید که دامنه g را یک متغیر تصادفی v در نظر بگیریم و بافت نگار دامنه آن، (v_i) به ازای $i = 1, 2, \dots, K$ ، را تشکیل دهیم که K تعداد افزایش های دامنه گستته است. آنگاه μ_i امین گشتاور v حول میانگین آن عبارتست از

$$\mu_i(v) = \sum_{i=1}^K (v_i - m)^n p(v_i) \quad (9-2.8)$$

که در آن

$$m = \sum_{i=1}^K v_i p(v_i). \quad (10-2.8)$$



شکل ۸.۱۶ (الف) قطعه 'مرزی'؛ (ب) نمایش آن به صورت تابع یک بعدی

کمیت m به عنوان میانگین یا مقدار متوسط \bar{r} و μ_m به عنوان واریانس آن شناخته می شود. معمولاً برای تشخیص بین نمایه های با اشکال کاملاً متفاوت تنها تعداد کمی از گشتاورهای مرتبه پایین مورد نیاز است. یک راه دیگر این است که مساحت $(r)g(r)$ را به مقدار واحد تراز کنیم و سپس با آن به عنوان یک تابع بافت نگار برخورد نماییم. در این مورد، r متغیر تصادفی خواهد بود و گشتاورها عبارتند از

$$\mu_n(r) = \sum_{i=1}^L (r_i - m)^n g(r_i) \quad (11-2. 8)$$

که در آن

$$m = \sum_{i=1}^L r_i g(r_i). \quad (12-2. 8)$$

در این نمادگذاری، L تعداد نقاط مرزی است، و $\mu_n(r)g(r)$ مستقیماً با شکل $(r)g(r)$ مرتبط است. مثلاً گشتاور دوم، $(r)g(r)$ ، گستردگی منحنی حول مقدار میانگین \bar{r} و گشتاور سوم، $(r)g(r)$ ، تقارن آن نسبت به میانگین را اندازه می گیرد. از هر دو نمایش گشتاور می توان به طور همزمان در توصیف یک قطعه 'مرزی' یا نمایه استفاده کرد.

تا این مرحله تنها، توصیف مرز را به صورت توصیف توابع یک بعدی، ساده کردیم. گرچه گشتاورها متداول ترین توصیف گرهای شناخته شده تاکنون هستند، اما تنها توصیف گرهای مقید نیستند. به عنوان نمونه، یک روش دیگر محاسبه 'تبديل فوريه گسته' یک بعدی با استفاده از معادله '۲-۲. ۳'، به دست آوردن طیف آن، و استفاده از K مؤلفه، اول طبق برای توصیف $(r)g(r)$ است. مزیت گشتاورها بر سایر روش ها این است که پیاده سازی آنها ساده است و دیگر این که آنها حامل تعبیری "فیزیکی" از شکل مرز هستند. با توجه به شکل ۸. ۱۶ آشکار است که این روش نسبت به چرخش غیرحساس است. می توان با مقیاس کردن محدوده r ، ابعاد شکل را، اگر مطلوب باشد، تراز کرد.

۳.۸ توصیف گرهای ناحیه ای

۱.۳.۸ چند توصیف گر ساده

مساحت^۱ ناحیه به عنوان تعداد پکسل های درون مرز آن تعریف می شود. محیط^۲ ناحیه طول مرز آن

۳.۸ توصیف گرهای ناحیه ای ۵۹۹

است. گرچه گاهی از مساحت و محیط به عنوان توصیف گر استفاده می شود، اما این مسئله محدود به مواردی می شود که اندازه اشیای مورد علاقه نامتغیر باشد. یک کاربرد متداول تر این دو توصیف گر در سنجش فشرده‌گی^۱ یک ناحیه است، که به صورت نسبت مجدد محیط به مساحت تعریف می شود. فشرده‌گی کمیته بی بعد است (و بنابراین نسبت به تغییرات مقیاس غیرحساس است) و برای یک ناحیه^۲ دایروی کمینه است. با صرف نظر از خطای حاصل از چرخش یک ناحیه^۳ رقمه، فشرده‌گی نسبت به تغییر جهت نیز غیرحساس است.

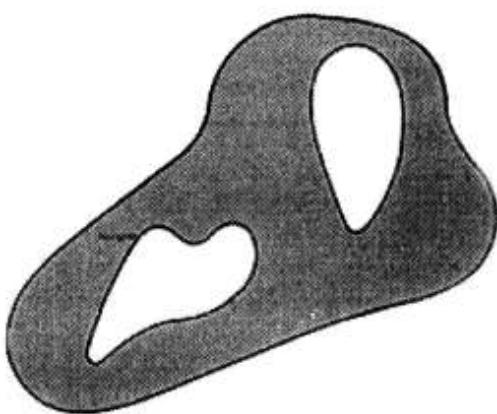
محورهای اصلی^۴ یک ناحیه، بردارهای ویژه ماتریس کوواریانس حاصل از کاربرد پیکسل های درون ناحیه به عنوان متغیرهای تصادفی، هستند (بخش ۳.۶ را ببینید). دو بردار ویژه^۵ ماتریس کوواریانس که درجهت حد اکثر گستردگی ناحیه ای هستند، متعامد می باشند. درجه^۶ گستردگی با مقدار ویژه متناظر سنجیده می شود. بنابراین گستردگی و جهت اصلی یک ناحیه را می توان با بزرگترین مقدار ویژه و بردار ویژه^۷ متناظر توصیف کرد. این نوع از توصیف نسبت به چرخش غیرحساس است اما اگر از مقادیر ویژه برای سنجش گستردگی استفاده شود، به تغییرات مقیاس بستگی دارد. یک روش که معمولاً برای جبران این مشکل به کار می رود، این است که نسبت^۸ مقدار ویژه^۹ بزرگ به کوچک را به عنوان توصیف گر در نظر بگیریم.

سایر سنجه های ساده که به عنوان توصیف گر ناحیه استفاده می شوند، عبارتند از: میانگین و میانه^{۱۰} سطوح خاکستری، مقادیر کمینه و بیشینه^{۱۱} سطح خاکستری، تعداد پیکسل های با مقادیر بالاتر و پایین تر از میانگین.

۲۰.۳.۸ توصیف گرهای توبولوژیکی

خواص توبولوژیکی برای توصیف های سراسری نواحی در صفحه^{۱۲} تصویر مفید هستند. در تعریف ساده، توبولوژی^{۱۳} مطالعه^{۱۴} خواصی از شکل است که تازمانی که هیچ بررش یا پیوندی در شکل رخ ندهد، متأثر از تغییر شکل نیستند. گاهی بررش یا پیوند، اعوجاج صفحه^{۱۵} لاستیکی^{۱۶} خوانده می شوند. مثلاً شکل ۱۷.۸ ناحیه ای با دو سوراخ را نمایش می دهد. بنابراین واضح است که اگر یک توصیف گر توبولوژیکی به عنوان تعداد سوراخ های درون یک ناحیه تعریف شود، این توصیف گر متأثر از تبدیلات کششی یا چرخشی

1) Compactness 2) Principal axes 3) Ratio 4) Topology 5) Rubber-sheet distortion



شکل ۱۷.۸ ناحیه ای با دو سوراخ

نیست. با این حال در حالت کلی اگر ناحیه بریده یا تاشود، تعداد سوراخ ها تغییر می کنند. توجه کنید که چون کشش بر فاصله اثر می گذارد، خواص توپولوژیکی به مفهوم فاصله یا هر خاصیت دیگری که مبنی بر یک سنجه^{۱)} فاصله باشد، بستگی ندارند.

خاصیت توپولوژیکی دیگری که برای توصیف ناحیه مفید است، تعداد اجزای متصل آن است. یک جزء متصل از یک مجموعه، زیر مجموعه ای با اندازه بیشینه است چنانکه هر دو نقطه اش را بتوان با یک منحنی پیوسته که تمام آن درون این زیر مجموعه قرار داشته باشد، به هم پیوندداد. شکل ۱۸.۸ ناحیه ای با سه جزء متصل را نشان می دهد. (برای شرحی درباره اجزای متصل و برچسب زنی آنها بخش های ۲۰.۴.۳ و ۲۰.۴.۴ را ببینید. در بخش ۱۷.۴.۴ یک الگوریتم شکل شناسی برای استخراج اجزای متصل ارائه می دهیم.)

برای یک شکل، می توان با استفاده از تعداد سوراخ های H و تعداد اجزای متصل C عدد اولر^{۱)} آن را تعریف کرد:

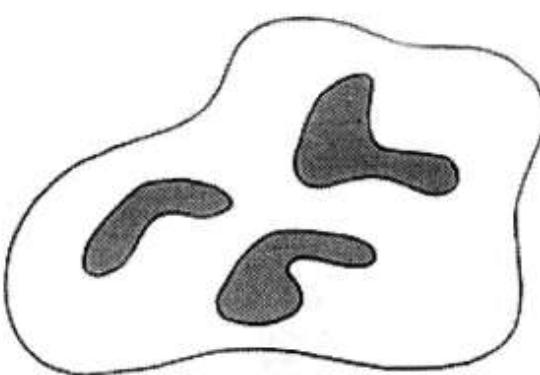
$$E = C - H \quad (1-3.8)$$

عدد اولر نیز یک خاصیت توپولوژیکی است. مثلاً نواحی نشان داده شده در شکل ۱۹.۸ به ترتیب اعداد اولر "۰" و "-۱" دارند، زیرا "A" یک جزء متصل و یک سوراخ دارد و "B" یک جزء متصل و دو

1) Euler number

۳۰.۸ توصیف گرهای ناحیه ای

۶۰۱

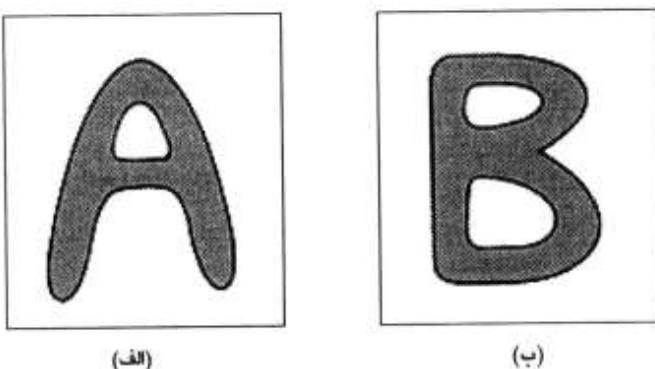


شکل ۱۸.۸ ناحیه ای با مه بخش متصل

سوراخ دارد.

نواحی نمایش داده شده با پاره خط ها (که با عنوان شبکه های چندضلعی^۱ اشاره می شوند) تعبیر بسیار ساده ای بر حسب عدد اولر دارند. شکل ۱۸.۲۰ یک شبکه چندضلعی را نشان می دهد. غالباً دسته بندي نواحی داخلی چنین شبکه ای به وجوده^۲ و سوراخ ها مهم است. نمادگذاري تعداد رؤوس با W ، تعداد بالها^۳ با Q ، و تعداد وجوه با F ، رابطه زیر را که رابطه اولر^۴ خوانده می شود، نتيجه می دهد:

$$W - Q + F = C - H \quad (۲-۳.۸)$$

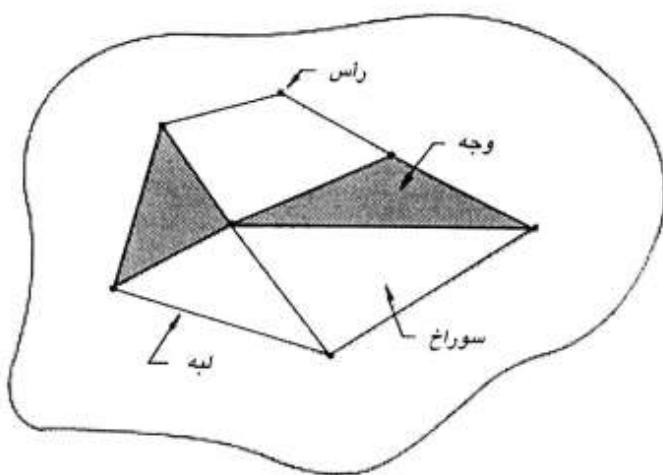
شکل ۱۹.۸ دو ناحیه که به ترتیب اعداد اولر $W=1$ - دارند

1) Polygonal network

2) Faces

3) Edge

4) Euler formula



شکل ۲۰.۸ ناحیه ای که حاوی یک شبکه چند ضلعی است

که با توجه به معادله (۳-۳.۸)، به صورت زیر با عدد اول رابطه دارد:

$$W - Q + F = C - H = E. \quad (3-3.8)$$

شبکه نشان داده شده در شکل ۲۰.۸ دارای ۷ رأس، ۱۱ لبه، دو وجه، یک جزء متصل، و سه سوراخ است؛ بنابراین

$$7 - 11 + 2 = 1 - 3 = -2.$$

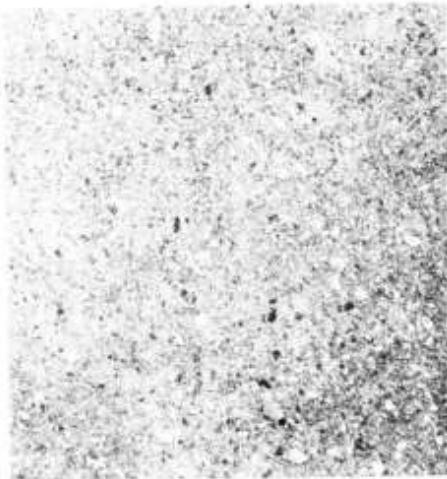
گرچه مفاهیم توبولوژیکی نسبتاً کلی هستند، ویژگی دیگری را فراهم می کنند که غالباً در بیان مشخصات نواحی درون صحنه مفید است.

۳.۳.۸ بافت

یک روش مهم برای توصیف ناحیه، کمی سازی محتوای بافتی^۱ آن است. گرچه هیچ تعریف معینی از بافت وجود ندارد، اما به نظر می رسد این توصیف گر سنجه هایی^۲ از خواصی نظیر همواری^۳، زبری^۴، و منظم بودن^۵ را فراهم می آورد (شکل ۲۱.۸ چند مثال را نشان می دهد). سه روش اصلی مورد استفاده در پردازش تصویر برای توصیف بافت ناحیه عبارتند از: آماری، ساختاری، و طیفی. روش های آماری

1) Texture 2) Measures 3) Smoothness 4) Coarseness 5) Regularity

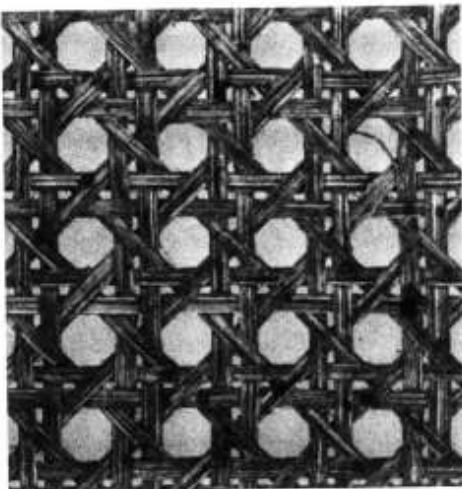
۳.۸ توصیف گرهای ناحیه ای ۶۰۳



(الف)



(ب)



(پ)

شکل ۲۱.۸ مثال هایی از باقهای (الف) هموار، (ب) زیر، و (پ) منظم. (از [Fu, Gonzalez & Lee] 1987)

خصوصیات بافت هانظیر همواری، زبری، دانه ای بودن، و غیره را تعین می کنند. روش های ساختاری با ترتیب عناصر سازنده تصویر¹، نظریه توصیف بافت مبتنی بر خطوط موازی با فواصل منظم، سر و کار دارند. روش های طیفی برآسان خواص طیف فوریه هستند و برای آشکارسازی تناوب سراسری تصویر با استفاده از شناسایی قله های پرانرژی و باریک طیف به کار می روند (بخش های ۵.۷ و ۵.۸ را ببینید).

1) *Image primitive*

روشهای آماری

یکی از ساده‌ترین روش‌ها برای توصیف بافت استفاده از گشتاورهای بافت نگار سطح خاکستری تصویر یا یک ناحیه آن است. فرض کنید z یک متغیر تصادفی و بیانگر شدت تصویر گسته باشد و فرض کنید $i = 1, 2, \dots, L$ ، بافت نگار z باشد، که در آن L ، تعداد سطوح متفاوت شدت است.

همان طور که در بخش ۸.۲.۴ بیان شد، گشتاور m حول میانگین عبارتست از

$$\mu_m(z) = \sum_{i=1}^L (z_i - m)^m p(z_i) \quad (4-3.8)$$

که m مقدار میانگین z (شدت متوسط) است:

$$m = \sum_{i=1}^L z_i p(z_i). \quad (5-3.8)$$

با توجه به معادله (4-3.8) مشخص می‌شود که $\mu_0 = 0$ ، μ_1 ، گشتاور دوم [که واریانس^۱ نیز خوانده می‌شود و با $(z)^2$ نشان داده می‌شود] اهمیت ویژه‌ای در توصیف بافت دارد. این گشتاور سنجه‌ای از وضوح^۲ سطح خاکستری است که می‌توان از آن برای تعیین توصیف گرهای تختی نسبی^۳ استفاده کرد. مثلاً سنجه^۴

$$R = 1 - \frac{1}{1 + \sigma^2(z)} \quad (6-3.8)$$

برای نواحی با شدت ثابت صفر است $[0 = (z)^2]$ اگر تمام z ‌ها برابر باشند] و به ازای مقادیر بزرگ $(z)^2$ به سمت واحد میل می‌کند. گشتاور سوم سنجه‌ای از میزان اُربیتی^۵ بافت نگار است، در حالی که گشتاور چهارم سنجه‌ای از تختی نسبی آن می‌باشد. گشتاورهای پنجم و بالاتر چنین رابطه 'ساده‌ای را با شکل بافت نگار ندارند، اما در تشخیص کمی محتوای بافت به ما کمک می‌کنند.

سنجه‌هایی از بافت که تنها با استفاده از بافت نگارها محاسبه می‌شوند، این محدودیت را دارند که هیچ اطلاعاتی راجع به محل نسبی پیکسل‌ها حمل نمی‌کنند. یک راه برای استفاده از این نوع اطلاعات در فرآیند تحلیل بافت این است که نه تنها توزیع شدت‌ها بلکه محل پیکسل‌های با مقدار شدت پیکسان یا تقریباً پیکسان را نیز مورد توجه قرار دهیم.

1) Contrast

2) Relative flatness

3) Skewness

۳.۸ توصیف گرهای ناحیه ای ۶۰۵

فرض کنید P یک عملگر مکان و A یک ماتریس $k \times k$ باشد که عنصر a_{ij} برابر تعداد بارهایی است که نقاط با سطح خاکستری z_i (در فاصله 'مکانی' مشخص شده با P) نسبت به نقاط با سطح خاکستری z_j رخ می دهد، که $0 \leq i \leq k$ و $0 \leq j \leq k$. مثلاً تصویری با سطح خاکستری $z_1 = 0$ ، $z_2 = 1$ ، $z_3 = 2$ را به صورت زیر در نظر بگیرید:

$$\begin{matrix} 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{matrix}$$

با تعریف عملگر P به صورت «یک پیکسل به راست و یک پیکسل به پایین» ماتریس A 3×3 ، به صورت زیر بدست می آید:

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

که، به عنوان مثال، a_{11} (درایه 'سمت چپ بالا') تعداد بارهایی است که پیکسلهای با سطح $0 = z_1$ در فاصله 'یک پیکسل به راست و یک پیکسل به پایین' نسبت به پیکسلی با همان سطح خاکستری ظاهر شده اند، و a_{33} (درایه 'سمت راست بالا') تعداد بارهایی است که نقاط با سطح خاکستری $0 = z_1$ در فاصله 'یک پیکسل به راست و یک پیکسل به پایین' نسبت به نقاط با سطح خاکستری $2 = z_3$ ظاهر شده اند. ابعاد A با تعداد سطوح خاکستری متفاوت در تصویر ورودی تعیین می شود. بنابراین برای استفاده از مفاهیم مشروح در این بخش، معمولاً لازم است که سطوح خاکستری تصویر مجدداً به تعداد کمی نوارهای (بازه های) سطح خاکستری چندی شوند تا این که ابعاد A بیش از حد بزرگ نباشد.

فرض کنید n تعداد کل زوج نقاط تصویر باشد که در فاصله 'P' نسبت به هم قرار دارند (در مثال قبل $n = 16$). اگر با تقسیم هر عنصر A بر n ماتریس C ایجاد شود، آنگاه C تخمینی از این احتمال توأم است که یک زوج نقطه که به فاصله 'P' از هم هستند، مقادیر (z_i, z_j) داشته باشند. ماتریس C ماتریس وقوع همزمان سطح خاکستری^۱ خوانده می شود. چون C به P بستگی دارد، می توان با انتخاب یک عملگر مکان مناسب، الگوهای بافت معینی را در تصویر آشکار کرد. مثلاً، عملگر به کار رفته در مثال قبل نسبت

1) Gray-level co-occurrence matrix

۶۰۶ پردازش تصویر رقمه

به نوارهای باشد ت ثابت درجهت ۴۵- درجه حساس است. (توجه کنید که بالاترین مقدار در $A_{11} = 4$ بود، که یک عامل آن وجود خطی از نقاط باشد "۰" درجهت ۴۵- درجه می باشد). به بیان کلی تر هدف ما تحلیل ماتریس C داده شده است تا این که بافت ناحیه‌ای را که C روی آن محاسبه شده است، دسته‌بندی کنیم. مجموعه‌ای از توصیف گرها که برای این کار مفید هستند، عبارتند از:

(۱) احتمال پیشینه،

$$\max_i(c_{ii})$$

(۲) گشتاور تفاضلی عنصری^۱ از مرتبه^k،

$$\sum_i \sum_j (i-j)^k c_{ij}$$

(۳) گشتاور تفاضلی عنصری معکوس^۲ از مرتبه^k،

$$\sum_i \sum_j c_{ij}/(i-j)^k \quad i \neq j$$

(۴) آنتروپی^۳،

$$-\sum_i \sum_j c_{ij} \log c_{ij}$$

(۵) یکنواختی^۴،

$$\sum_i \sum_j c_{ij}^2$$

هدف پایه‌ای این است که "محتوای" C را با این توصیف گرها مشخص نماییم. مثلاً اولین خاصیت، نشان دهنده بزرگترین پاسخ به عملگر P است (مانند مثال قبل). وقتی مقادیر بزرگ C تزدیک قطر اصلی باشند، توصیف گر دوم، مقدار نسبتاً کوچکتری خواهد داشت، زیرا تفاضلهای $(j-i)$ در آن نقاط کوچکتر است. توصیف گر سوم خاصیت مخالف توصیف گر دوم را دارد. توصیف گر چهارم سنجه‌ای از تصادفی بودن است که وقتی تمام عناصر C برابر باشند، بزرگترین مقدار خود را به دست می آورد. بر عکس وقتی همه c_{ij} ها برابر باشند، توصیف گر پنجم کمترین مقدار خود را به دست می آورد. یک روش برای استفاده از این توصیف گرها این است که مقادیر توصیف گرها را برای مجموعه‌ای از بافت‌های متفاوت، به سامانه "آموزش دهیم". در نتیجه بافت یک ناحیه مجھول با میزان تطبیق توصیف گرها یش با توصیف گرها ذخیره شده در حافظه سامانه تعیین می شود. در فصل ۹ مفصل تر

1) Element difference moment

2) Inverse element difference moment

3) Entropy

4) Uniformity

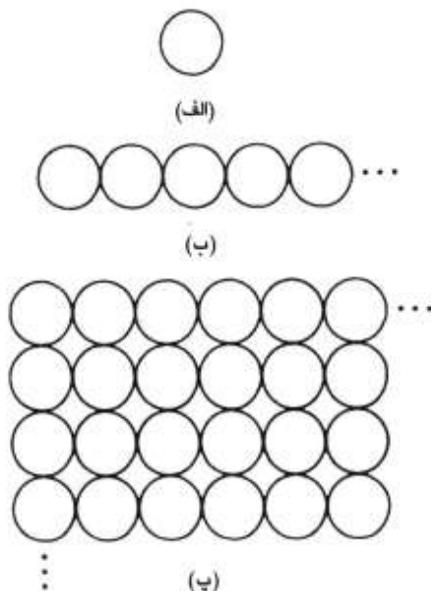
۳۰۸ توصیف گرهای ناچه ای ۶۰۲

درباره 'تطبیق شرح می دهیم .

روشهای ساختاری

همان طور که در شروع این بخش ذکر شد، دومین نوع توصیف بافت، مبتنی بر مقاہیم ساختاری است. فرض کنید که قاعده ای به شکل $S \rightarrow aS$ داریم که به معنی بازنویسی نماد S به صورت aS باشد (مثالاً سه بار اعمال این قاعده بر S رشته 'aaa' را می دهد). اگر a نشانه یک دایره باشد (شکل ۲۲.۸) و رشته ...aaa... بدین معنی باشد که "دایره ها در سمت راست هم قرار می گیرند" با قاعده $S \rightarrow aS$ می توان الگوی بافت شکل ۲۲.۸(ب) را تولید کرد.

فرض کنید که چند قاعده جدید به این روش بیفزاییم: $A \rightarrow bS$ ، $A \rightarrow c$ ، $A \rightarrow cA$ ، $S \rightarrow bA$: $S \rightarrow a$ که حضور b به این معنی است که "دایره در زیر" و حضور c به معنی این است که "دایره در سمت چپ" قرار گیرد. اکنون می توانیم رشته ای به شکل aaabccbaa تولید کنیم که متناظر با ماتریسی 3×3 از دایره ها است. به همان روش ساده می توان الگوهای بافت بزرگتر، نظیر آنچه در شکل ۲۲.۸(پ) دیده



شکل ۲۲.۸ (الف) عنصر اولیه بافت؛ (ب) الگوی تولید شده با قاعده $S \rightarrow aS$ ؛ (پ) الگوی بافت دو بعدی تولید شده با قاعده فوق و سایر قواعد

می شود، را تولید کرد. (توجه کنید که این قاعده ها می توانند ساختارهای غیرمستطیلی نیز تولید کنند).

هدف پایه ای شرح فوق این است که می توان با استفاده از یک «عنصر اولیه» بافت^۱ و قواعدی که تعداد چیدمان های^۲ عنصر (های) اولیه را محدود می کنند، الگوهای باقی پیچیده تری را تشکیل داد. این مقاومت محور توصیف های رابطه ای هستند، که در بخش ۸.۵ با تفصیل بیشتری آنها را شرح می دهیم.

روشهای طیفی

همان طور که در بخش های ۸.۵ و ۷.۵ بیان شد، طیف فوریه برای توصیف جهت مندی الگوهای دو بعدی متناوب یا تقریباً متناوب در تصویر کاملاً مناسب است. این الگوهای بافت سراسری، گرچه به آسانی به صورت تجمع دسته های پرانرژی در طیف قابل تشخیص هستند، معمولاً آشکارسازی آنها با روش های مکانی به دلیل طبیعت محلی این روشهای بسیار مشکل است.

در اینجا سه ویژگی طیف فوریه را که در توصیف بافت مفید هستند، بررسی می کنیم: (۱) قله های مهم در طیف، جهت اصلی الگوهای بافت را می دهد، (۲) محل قله ها در صفحه فرکانس، دوره^۳ تناوب اصلی الگوها را می دهد؛ و (۳) با حذف تمام مؤلفه های تناوبی توسط فیلتر کردن، عناصر تصویری نامتناوب باقی می مانند، که پس از آن می توان آنها را با روش های آماری توصیف کرد. به خاطر آورید که طیف یک تصویر واقعی حول مبدأ فرکانس متقارن است، بنابراین کافی است که تنها یک نیمه^۴ صفحه فرکانس را در نظر بگیریم. بنابراین هنگام تحلیل، هر الگوی متناوب تنها با یک قله، نه دو قله، از طیف ارتباط دارد.

با بیان طیف در مختصات قطبی، تابع $S(r, \theta)$ بدست می آید، که S تابع طبی r و θ متغیرهای این دستگاه مختصات هستند. غالباً با استفاده از تابع طیفی S ، آشکارسازی و تعبیر ویژگی های سه گانه طیف فوریه ساده می شود. برای هر یک از جهات θ ، $S(r, \theta)$ را می توان تابع یک بعدی $S(r)$ در نظر گرفت. به طور مشابه به ازای هر فرکانس ثابت و دلخواه r ، $S(r, \theta)$ تابعی یک بعدی است. با تحلیل $S(r)$ به ازای یک مقدار ثابت θ ، رفتار طیف (نظیر وجود قله ها) در طول یک مسیر شعاعی از مبدأ بدست می آید، در

(1) Texture primitive

(2) Arrangements

۳۰.۸ توصیف گرهای ناحیه ای ۶۰۹

حالی که تحلیل (θ) برای مقدار ثابت r ، رفتار طیف در طول دایره ای به مرکز مبدأ را می دهد.
با انتگرال گیری (برای متغیرهای گستته مجموع گیری) از این توابع توصیف کلی تری بدست می آید:

$$S(r) = \sum_{n=0}^{\infty} S_n(r) \quad (7-3.8)$$

و

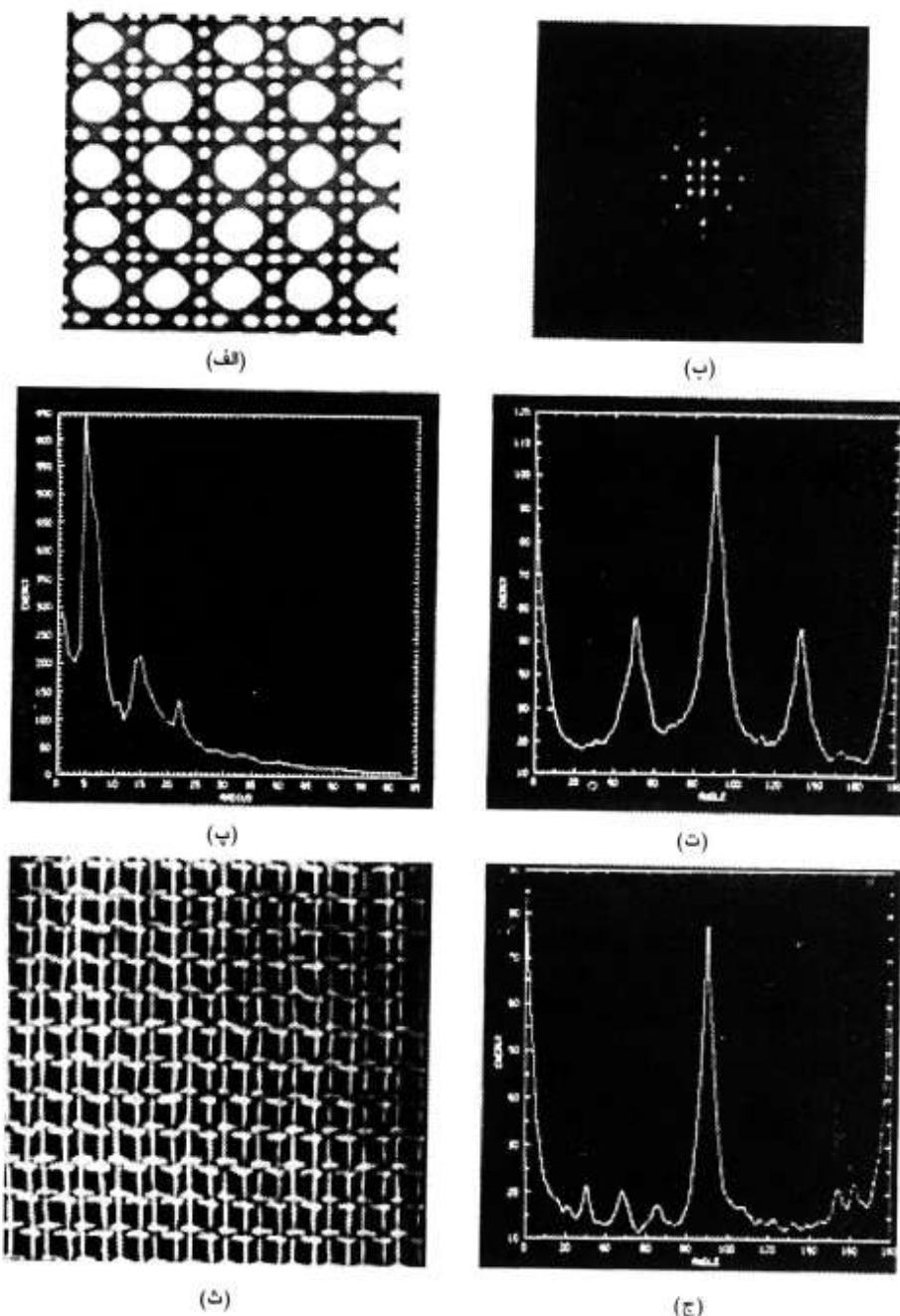
$$S(\theta) = \sum_{r=1}^{\infty} S_r(\theta) \quad (8-3.8)$$

که R شعاع دایره ای به مرکز مبدأ است. برای طیف $N \times N$ ، معمولاً $R = 2N$ انتخاب می شود.
معادلات (7-3.8) و (8-3.8) برای هر زوج مختصات (r, θ) ، زوج $[S(r), S(\theta)]$ را ایجاد می کنند. با تغییر این مختصات می توان دوتابع یک بعدی $S(r)$ و $S(\theta)$ را، که توصیف طیف انرژی بافت کل تصویر یا ناحیه تحت بررسی را تشکیل می دهد، تولید کرد. گذشته از این برای توصیف کمی رفتار این توابع می توان توصیف گرهای خود این توابع را محاسبه کرد. توصیف گرهایی که معمولاً بین منظور استفاده می شوند، عبارتند از: محل بزرگترین مقدار، میانگین و واریانس تغییرات دامنه ای و محوری (بخش ۴-۲.۴ را ببینید)، و فاصله بین میانگین و بزرگترین مقدار تابع.

مثال: شکل ۸.۲۳ کاربرد معادلات (7-3.8) و (8-3.8) برای توصیف سراسری بافت را نمایش می دهد. شکل ۸.۲۳ (الف) تصویری با بافت متناوب و شکل ۸.۲۳ (ب) نیز طیف آن را نشان می دهد. اشکال ۸.۲۳ (پ) و (ت) به ترتیب نمودارهای $S(r)$ و $S(\theta)$ را نشان می دهد. نمودار $S(r)$ ساختاری معمولی است که در نزدیک مبدأ، محتوایی پرانرژی دارد و با افزایش فرکانس، محتوای انرژی آن کاهش می یابد. نمودار $S(\theta)$ قله های بلندی با فواصل ۴۵ درجه را نشان می دهد، که به وضوح متناظر با تناوب موجود در محتوای طیفی تصویر هستند.

به عنوان نمایش چگونگی کاربرد نمودار (θ) برای تشخیص بین دوالگوی بافتی، شکل ۸.۲۳ (ث) تصویری را نشان می دهد که الگوی بافتی آن بیشتر در جهت های افقی و عمودی است. شکل ۸.۲۳ (ج) نمودار $S(\theta)$ برای طیف این تصویر را نمایش می دهد. همان طور که انتظار می رفت، این نمودار قله های بلندی با فواصل ۹۰ درجه دارد. با تحلیل شکل موج های $S(\theta)$ در این دوالگوی بافتی، تمیز دادن آن دو از هم ساده است.





شکل ۲۳.۸ (الف) تصویر یک بافت متناوب؛ (ب) طیف آن؛ (ب) نمودار $S(r)$ ؛ (ت) نمودار $S(\theta)$ ؛ (ث) تصویر یک بافت متناوب دیگر؛ (ج) نمودار $S(\theta)$. (با مجوز از D. Brzakovic University of Tennessee).

۶۱۱ ۳.۸ توصیف گرها ناحیه ای

۴.۳.۸ گشتاورها

برای تابع پیوسته دو بعدی $f(x, y)$ ، گشتاورهای مرتبه $(p+q)$ به ازای $\dots, 0, 1, 2, \dots$ به

صورت زیر تعریف می شوند:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (9-3.8)$$

قضیه یکتاپی پاپولیس^۱ [۱۹۶۵] بیان می کند که اگر تابع $f(x, y)$ پیوسته قطعه ای^۲ باشد و تنها در قسمت محدودی از صفحه xy مقادیر غیر صفر داشته باشد، گشتاورهای آن از تمام مرتبه ها وجود دارند و دنباله گشتاور (m_{pq}) نیز به طور یکتاپی^۳ معین می شود. بر عکس دنباله گشتاور (m_{pq}) نیز $f(x, y)$ را به طور یکتاپی تعیین می کند. گشتاورهای مرکزی^۴ را می توان به صورت

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (10-3.8)$$

بیان کرد که در آن

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \text{و} \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

برای تصویر رقمی معادله (۱۰-۳.۸) بدین صورت می شود:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y). \quad (11-3.8)$$

گشتاورهای مرکزی با مرتبه ناپذیرگتر از ۳ عبارتند از

$$\begin{aligned} \mu_{10} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^0 f(x, y) \\ &= m_{10} - \frac{m_{10}}{m_{00}} (m_{00}) \\ &= 0 \\ \mu_{01} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^1 f(x, y) \\ &= m_{01} - \frac{m_{10} m_{01}}{m_{00}} \end{aligned}$$

1) Papoulis uniqueness theorem

2) Piecewise continuous

3) Central moment

$$\begin{aligned}
 \mu_{20} &= \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^0 f(x, y) \\
 &= m_{20} - \frac{2m_{10}^2}{m_{00}} + \frac{m_{10}^2}{m_{00}} = m_{20} - \frac{m_{10}^2}{m_{00}} \\
 \mu_{02} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^2 f(x, y) \\
 &= m_{02} - \frac{m_{01}^2}{m_{00}} \\
 \mu_{30} &= \sum_x \sum_y (x - \bar{x})^3 (y - \bar{y})^0 f(x, y) \\
 &= m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2m_{10} \\
 \mu_{12} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^2 f(x, y) \\
 &= m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2m_{10} \\
 \mu_{21} &= \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^1 f(x, y) \\
 &= m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2m_{01} \\
 \mu_{03} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^3 f(x, y) \\
 &= m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01}.
 \end{aligned}$$

به اختصار،

$$\begin{array}{ll}
 \mu_{00} = m_{00} & \mu_{11} = m_{11} - \bar{y}m_{10} \\
 \mu_{10} = 0 & \mu_{30} = m_{30} - 3\bar{x}m_{20} + 2m_{10}\bar{x}^2 \\
 \mu_{01} = 0 & \mu_{12} = m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2m_{10} \\
 \mu_{20} = m_{20} - \bar{x}m_{10} & \mu_{21} = m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2m_{01} \\
 \mu_{02} = m_{02} - \bar{y}m_{01} & \mu_{03} = m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01}.
 \end{array}$$

گشتاورهای مرکزی تراز شده¹⁾، که با نماد η_{pq} نشان داده می شوند، به ازای

به صورت

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^q} \quad (12-3. \Delta)$$

تعریف می شوند که

$$\gamma = \frac{p + q}{2} + 1 \quad (13-3. \Delta)$$

1) Normalized central moment

۶۱۳ ۳.۸ توصیف گرهای ناحیه ای

مجموعه ای از هفت ثابت گشتاوری^۱ را می توان از گشتاورهای دوم و سوم بدست آورد:

$$\phi_1 = \eta_{20} + \eta_{02} \quad (14-3, \Delta)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (15-3, \Delta)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (16-3, \Delta)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (17-3, \Delta)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \quad (18-3, \Delta)$$

$$- 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$$

$$[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (19-3, \Delta)$$

$$+ 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \quad (20-3, \Delta)$$

$$- 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})$$

$$[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].$$

این مجموعه گشتاورها نسبت به انتقال، چرخش، و تغییر مقیاس نامغایر هستند (Hu[1962]).

مثال: تصویر موجود در شکل ۸ ۲۴. (الف) در شکل ۸ ۲۴. (ب) به نصف اندازه کوچک شد، در شکل ۸ ۲۴. (پ) وارون آینه ای شد^۲، و در اشکال ۸ ۲۴. (ت) و (ث) به ترتیب ۲ و ۴۵ درجه چرخانده شد. هفت ثابت گشتاوری داده شده در معادلات (۸. ۲۰-۳، ۸) - (۱۴-۳، ۸) برای هر یک از این تصاویر محاسبه شدند، و از مقادیر حاصل لگاریتم گرفته شد تا محدوده تغییر^۳ آنها کاهش یابد. همان طور که جدول ۸. ۲ نشان می دهد، نتایج حاصل از اشکال ۸ ۲۴. (ب) - (ث) مطابقت خوبی با ثابت های

جدول ۲.۸ ثابت های گشتاوری برای اشکال ۸ ۲۴. (الف) - (ث)

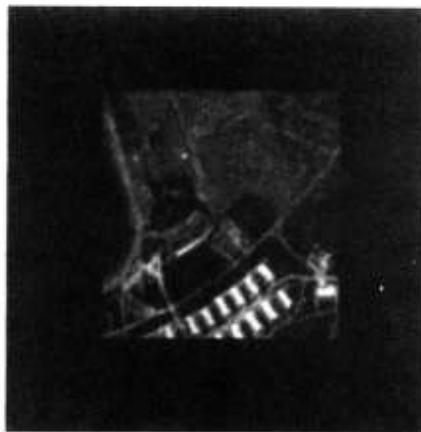
ثابت گشتاوری (لگاریتمی)	تصویر اولیه	تصویر نصف اندازه	تصویر آینه ای	تصویر ۲ درجه چرخیده	تصویر ۴۵ درجه چرخیده
ϕ_1	6.249	6.226	6.919	6.253	6.318
ϕ_2	17.180	16.954	19.955	17.270	16.803
ϕ_3	22.655	23.531	26.689	22.836	19.724
ϕ_4	22.919	24.236	26.901	23.130	20.437
ϕ_5	45.749	48.349	53.724	46.136	40.525
ϕ_6	31.830	32.916	37.134	32.068	29.315
ϕ_7	45.589	48.343	53.590	46.017	40.470

^۱ در رسیدن به این نتایج، مفاهیم خارج از محدوده این کتاب به کار گرفته می شوند. کتاب [Bell[1965] و مقاله' Hu[1962]] حاوی توضیحات مفصلی درباره' این مفاهیم هستند.

1) Invariant moment

2) Mirror- imaged

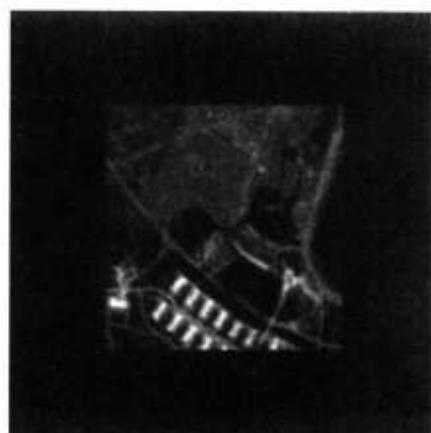
3) Dynamic range



(الف)



(ب)



(ب)



(ج)



(ج)

شکل ۲۶.۸ تصاویر مورد استفاده برای نمایش خواص ثابت های گشتاوری

www.MatlabKar.com

۴.۸ شکل‌شناسی ۶۱۵

محاسبه شده برای تصویر اولیه دارند. علت اصلی خطا را می‌توان به طبیعت رقمی داده‌ها متنسب کرد.



۴.۸ شکل‌شناسی

واژه 'شکل‌شناسی'^۱ معمولاً دلالت بر شاخه‌ای از زیست‌شناسی دارد که به شکل و ساختمان بدن حیوانات و گیاهان می‌پردازد. در اینجا در مبحث شکل‌شناسی ریاضی^۲ همان واژه را به عنوان ابزاری برای استخراج اجزای تصویر که در نمایش و توصیف شکل ناحیه مفید استند، نظریه مرزها، اسکلت‌ها، و بدن^۳ محدب، استفاده می‌کنیم. همچنین برای پیش-یا پس پردازش تصویر به روش‌های شکل‌شناسی نظریه اجرای فیلتر شکل‌شناسی^۴، پاریک‌سازی^۵، و هرس کردن^۶ علاقه مندیم.

زبان شکل‌شناسی ریاضی، نظریه 'مجموعه‌ها'^۷ است. همچنین شکل‌شناسی راهی قوی و یکتا برای تعداد زیادی از مسائل پردازش تصویر است. در شکل‌شناسی، مجموعه‌ها یا نگر شکل اثباتی تصویر استند. مثلاً 'مجموعه' تمام پیکسل‌های سیاه در تصویر دودویی، توصیف کاملی از تصویر است. در تصاویر دودویی، این مجموعه‌ها اعضای فضای دو بعدی اعداد صحیح^۸ \mathbb{Z}^2 هستند (بخش ۲.۳.۱ را ببینید). تحت این شرایط هر عنصر از مجموعه، یک چندتایی^۹ (بردار دو بعدی) است که مختصات (x, y) آن همان مختصات یک پیکسل سیاه (طبق قرارداد) در تصویر می‌باشد. تصاویر رقمی خاکستری را می‌توان به صورت زیر مجموعه‌هایی از \mathbb{Z}^2 بیان کرد. در این مورد، دو مؤلفه اول هر عنصر مجموعه، مختصات یک پیکسل و مؤلفه سوم مقدار شدت گسته آن پیکسل می‌باشد. مجموعه‌های درون فضاهای با ابعاد بیشتر می‌توانند حاوی سایر خواص تصویر، نظریه رنگ و مؤلفه‌های متغیر بازمان، باشند. در بخش‌های بعدی چند مفهوم مهم در شکل‌شناسی ریاضی را مطرح کرده و شرح می‌دهیم. گرچه قبل از این فصل مفاهیمی نظری اسکلت و مرز ناحیه را شرح دادیم، در اینجا دوباره با آنها به عنوان موارد را خاصی از مجموعه خیلی بزرگتری از اعمال شکل‌شناسی برخورد می‌کنیم. بسیاری از این اعمال را می‌توان بر حسب فضای اقلیدسی^{۱۰} بعدی E ^{۱۱} بیان کرد. با این حال ما بر تصاویر دودویی که اجزای اشان عناصری از \mathbb{Z}^2 هستند، تمرکز می‌کنیم. در بخش ۴.۸.۵ تعمیم‌های شکل‌شناسی به تصاویر خاکستری

1) Morphology

2) Mathematical morphology

3) Morphological filtering

4) Pruning

5) Set theory

6) Tuple

را شرح می دهیم.

۱۰.۴.۸ گسترش و سایش

شرح اعمال شکل شناسی را با دو عمل گسترش^۱ و سایش شروع می کنیم. این دو عمل مبنای بیشتر اعمال شکل شناسی هستند که در ادامه چند تا از آنها را نیز شرح می دهیم.

چند تعریف پایه‌ای

فرض کنید A و B مجموعه هایی در Z^2 ، به ترتیب با اجزای (a_1, a_2) و $a = (b_1, b_2)$ باشند. انتقال^۲ با $(A, x) = (x_1, x_2)$ که با (A) نشان داده می شود، به صورت زیر تعریف می گردد:

$$(A)_x = \{c | c = a + x, \quad a \in A\}. \quad (1-4. A)$$

قرینه^۳ B^\top ، که با \hat{B} نشان داده می شود، به صورت زیر تعریف می شود:

$$\hat{B} = \{x | x = -b, \quad b \in B\}. \quad (2-4. A)$$

متتم^۴ مجموعه A عبارتست از

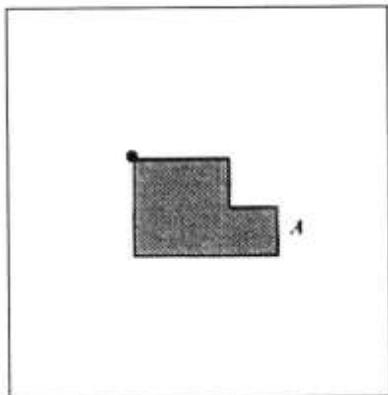
$$A^c = \{x | x \notin A\}. \quad (3-4. A)$$

در نهایت تفاضل^۵ دو مجموعه A و B ، که با $A - B$ نشان داده می شود، بدین صورت تعریف می شود

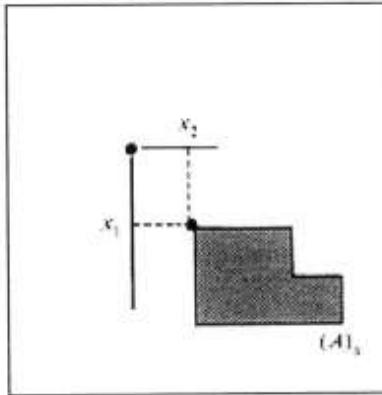
$$A - B = \{x | x \in A, x \notin B\} = A \cap B^c. \quad (4-4. A)$$

مثال: شکل ۸.۲۵. تعاریفی را که اخیراً بیان شدند، نمایش می دهد که در آنها نقطه 'سیاه نشان دهنده' مبدأ هر مجموعه است. شکل ۸.۲۵. (الف) مجموعه ای با عنوان A و شکل ۸.۲۵. (ب) انتقال آن با $x = (x_1, x_2)$ را نشان می دهد. توجه کنید که عمل انتقال با افزودن (x_1, x_2) به هر عنصر A اجرا می شود. شکل ۸.۲۵. (پ) مجموعه ای با عنوان B را نشان می دهد، و شکل ۸.۲۵. (ت) قرینه^۳ B حول

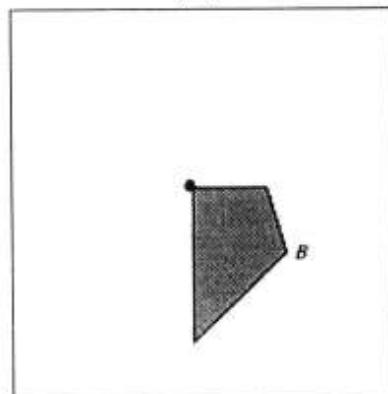
۶۱۷ شکل شناسی ۴.۸



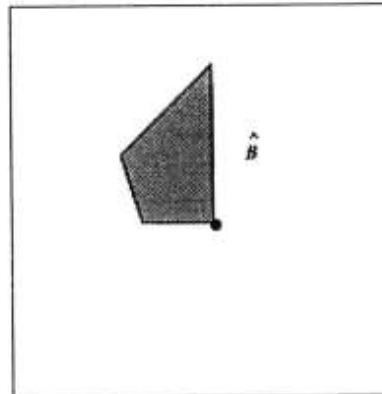
(الف)



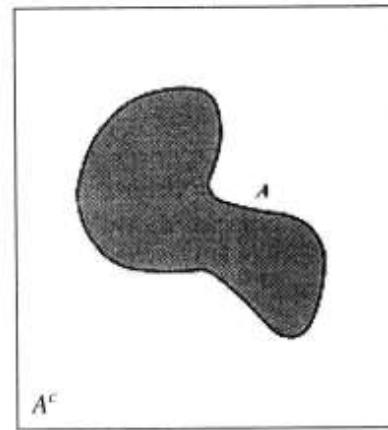
(ب)



(پ)

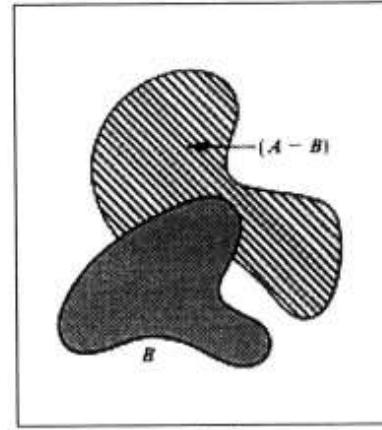


(د)



A^c

(ث)



(ج)

شکل ۲۵.۸ (الف) مجموعه 'A'؛ (ب) مجموعه 'A' که با x انتقال یافته است؛ (پ) مجموعه 'B'؛ (د) قرینه 'B'؛ (ث) مجموعه 'A' و متمم آن؛ (ج) تفاضل دو مجموعه (که خط دار نشان داده می شود). در هر یک از چهار شکل نقطه بزرگ نشانه مبدأ مجموعه است.

مبدأش را نشان می دهد. در نهایت شکل ۸. ۲۵ (ث) مجموعه 'دیگری با عنوان A و متمم آن را و شکل □ ۸. ۲۵ (ج) تفاضل بین این مجموعه 'A و مجموعه 'B از شکل ۸. ۲۵ (ج) را نشان می دهد.

گسترش

با فرض این که A و B مجموعه هایی در Z^2 باشند و \emptyset مجموعه 'نهی باشد، گسترش A با B ، که با $A \oplus B$ نشان داده می شود، به صورت زیر تعریف می شود:

$$A \oplus B = \{x | (\hat{B}), \cap A \neq \emptyset\}. \quad (5-4. ۸)$$

بنابراین فرآیند گسترش شامل بدست آوردن قرینه ' B حول مبدأش و سپس انتقال این قرینه با x است تا بدینوسیله (\hat{B}) بدست آید. آنگاه گسترش A با B مجموعه ' تمام x هایی است که، (\hat{B}) و A لااقل در یک عنصر غیر صفر همپوشانی داشته باشند. براساس این تعبیر می توان معادله ' (۸. ۴-۵) را به صورت زیر بازنویسی کرد.

$$A \oplus B = \{x | [(\hat{B}), \cap A] \subseteq A\}.$$

مجموعه ' B در گسترش و نیز در سایر اعمال شکل شناسی عموماً به عنوان عنصر ساختمانی ^۱ شناخته می شود.

معادله ' (۸. ۵-۴) تنها تعریف موجود از گسترش نیست (برای مشاهده دو تعریف دیگر ، اما معادل با این تعریف ، مسأله ۸. ۱۹ را ببینید). با این حال این تعریف مزیت بارزی بر سایر تعاریف دارد، زیرا وقتی به عنصر ساختمانی B همانند یک نقاب کانولوشن نظاره کنیم، این تعریف نسبت به سایر تعاریف شهودی تر است. گرچه گسترش ، مبتنی بر عملیات مجموعه ای و کانولوشن مبتنی بر عملیات حسابی است، فرآیند پایه ای «چرخاندن» ^۲ B حول مبدأش و سپس جابجائی متواالی آن طوری که روی تمام مجموعه ' (تصویر) A لغزانده شود، مشابه فرآیند کانولوشن مشروح در بخش های ۳. ۳ و ۴. ۱ است.

مثال: شکل ۸. ۲۶ (الف) یک مجموعه ' ساده و شکل ۸. ۲۶ (ب) یک عنصر ساختمانی و قرینه اش را

1) Structuring element 2) Flipping

۶۱۹ شکل شناسی ۴.۸

نمایش می دهد. در این مورد، عنصر ساختمانی و قرینه اش برابر هستند زیرا B نسبت به مبدأ متقاض است. خط مقطع در شکل ۸. ۲۶ (ب) مجموعه اولیه را به عنوان مرجع مقایسه نشان می دهد، و خط ممتد کرانی از x را نشان می دهد که جایگایی مبدأ \hat{B} با x های خارج آن باعث می شود که اشتراک \hat{B} و A تهی باشد. بنابراین مجموعه نقاط درون این مرز، گسترش A با B را تشکیل می دهند. شکل ۸. ۲۶ (ت) یک عنصر ساختمانی است که در جهت عمودی بیش از جهت افقی، گسترش ایجاد می کند. شکل ۸. ۲۶ (ث) حاصل گسترش با این عنصر را نشان می دهد.

□

سایش

برای مجموعه های A و B در Z^2 ، سایش A با B ، که با نماد $A \ominus B$ نشان داده می شود، به صورت زیر تعریف می شود:

$$A \ominus B = \{x | (B), \subseteq A\} \quad (6-4. 8)$$

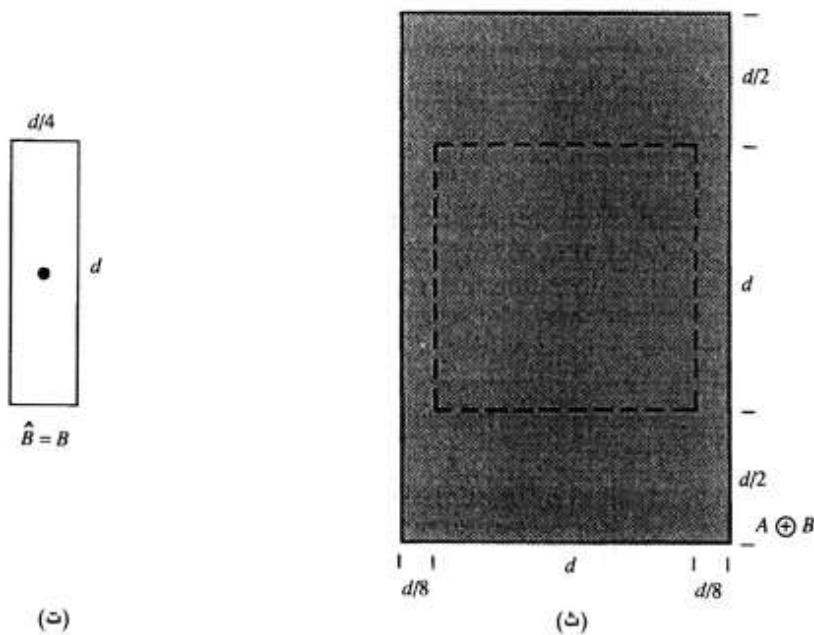
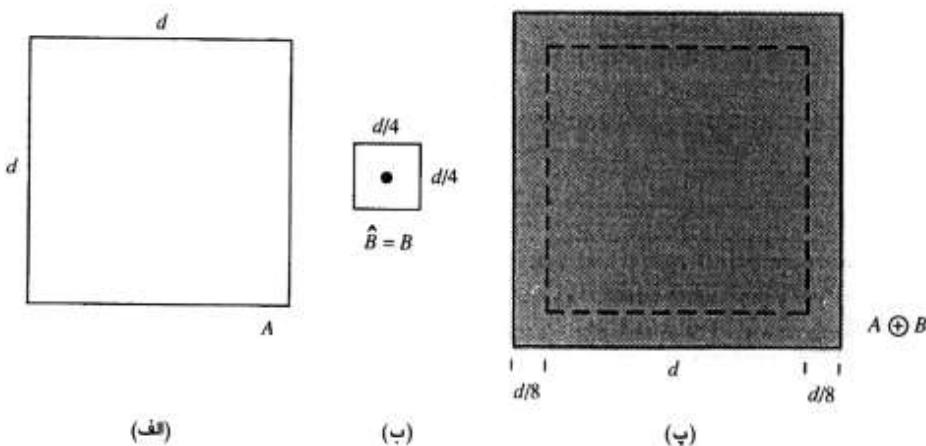
یعنی سایش A و B مجموعه تمام نقاط x است که اگر B به اندازه x جایجا شود، باز هم کاملاً درون A قرار گیرد. همانند گسترش، معادله (۶-۴. ۸) تنها تعریف سایش نیست (برای دیدن دو تعریف دیگر که معادل تعریف اخیر هستند، مسئله ۸. ۲۰ را بینید). با این حال به همان دلایلی که قبلاً در رابطه با معادله (۶-۴. ۵) ذکر شد، معمولاً در عمل معادله (۶-۴. ۸) ترجیح داده می شود.

مثال: شکل ۸. ۲۷ فرآیندی مشابه با فرآیند نمایش داده شده در شکل ۸. ۲۶ را نشان می دهد. در شکل ۸. ۲۷ (ب) نیز مانند قبل مجموعه A' به عنوان مرجع مقایسه با خط مقطع نشان داده شده است. در این شکل خط ممتد، کرانی از x هارا نشان می دهد که با جایه جایی مبدأ B با x های خارج از آن، مجموعه B به طور کامل درون A' قرار نمی گیرد. بنابراین مجموعه نقاط درون این مرز، سایش A با B را تشکیل می دهند. شکل ۸. ۲۷ (ت) یک عنصر ساختمانی طویل را نشان می دهد، و شکل ۸. ۲۷ (ث) سایش A با این عنصر را نمایش می دهد. توجه کنید که برای سایش، مجموعه اولیه به یک خط تبدیل شد. □

گسترش و سایش نسبت به دو عمل متمم گیری و قرینه گیری دوگان هم هستند. یعنی،

$$(A \ominus B)^c = A' \oplus \hat{B}. \quad (7-4. 8)$$

اکنون با هدف ارائه روشنی نوعی برای اثبات درستی عبارات شکل شناسی، این نتیجه را به طور رسمی



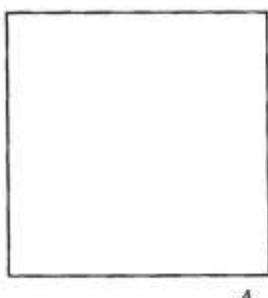
شکل ۸.۰۲ (الف) مجموعه اولیه A؛ (ب) عنصر ساختمانی مربعی و قرینه آن؛ (پ) گسترش A با B، که
سایه دار دیده می شود؛ (ت) عنصر ساختمانی طوبیان؛ (ث) گسترش A با این عنصر.

ثابت می کنیم . با شروع از تعریف سایش داریم که :

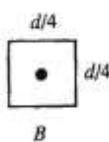
$$(A \ominus B)^c = \{x | (B)_i \subseteq A\}^c$$

اگر مجموعه (B) زیرمجموعه^{*} A باشد، آنگاه $\bigcap A' = \emptyset$ ، که در این حالت معادله[†] قبل به این

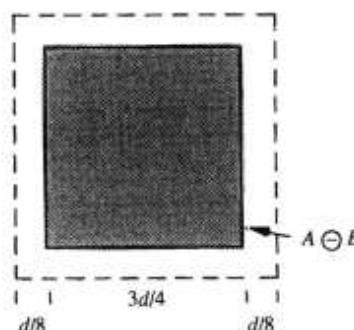
۴.۸ شکل شناسی ۶۲۱



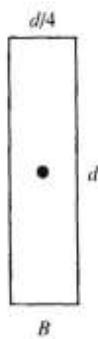
(الف)



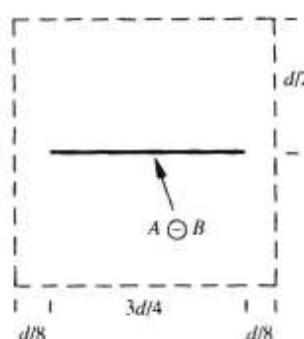
(ب)



(ج)



(د)



(ه)

شکل ۴.۸ (الف) مجموعه اولیه A ؛ (ب) عنصر ساختمانی B ؛ (ج) سایش A با B ، که سایه دار دیده می شود؛ (د) عنصر ساختمانی طویل؛ (ه) سایش A با این عنصر.

صورت می شود

$$(A \ominus B)^c = \{x | (B), \cap A^c = \emptyset\}$$

اما متمم مجموعه 'نها'ی که رابطه $(B), \cap A^c = \emptyset$ را برآورده می کنند، مجموعه ای از زها است

چنان که $(B), \cap A^c = \emptyset$. بنابراین

$$\begin{aligned} (A \ominus B)^c &= \{x | (B), \cap A^c \neq \emptyset\} \\ &= A^c \oplus \hat{B} \end{aligned}$$

که آخرین مرحله اثبات از معادله ۴.۵ نتیجه می شود. بدین ترتیب اثبات کامل شد.

۲۰۴.۸ بازگردان و بستن

همان طور که دیدیم عمل گسترش تصویر را توسعه می دهد^۱ و عمل سایش آن را کوچک می نماید^۲. در این بخش دو عمل شکل شناسی مهم دیگر را شرح می دهیم: بازگردان و بستن. معمولاً عمل بازگردان^۳ منحنی های تصویر را هموار می کند، اتصالات باریک را می شکند، و بیرون زدگی های باریک را حذف می کند. بستن^۴ نیز بخش های منحنی هارا هموار می کند، اما برخلاف عمل بازگردان معمولاً انفصلات باریک و تورفتگی های باریک و بلند را پوند می زند، سوراخ های کوچک را حذف می کند، و انفصلات منحنی ها را پر می کند.

بازگردان مجموعه 'A' با عنصر ساختمانی B، که با $B \circ A$ نشان داده می شود، بدین صورت تعریف می شود:

$$A \circ B = (A \ominus B) \oplus B \quad (8-4, A)$$

که بیان می کند بازشده A با B، با سایش A با B و سپس گسترش حاصل با B بدست می آید. بستن مجموعه 'A' با عنصر ساختمانی B، که با $B \circ A$ نشان داده می شود، بدین صورت تعریف می شود

$$A \bullet B = (A \oplus B) \ominus B \quad (9-4, A)$$

که بیان می کند بسته شده A با B، با گسترش A با B و سپس سایش حاصل با B بدست می آید. مثال: شکل ۸.۲۸. بازگردان و بستن مجموعه 'A' با یک عنصر ساختمانی دیسکی^۵ را نمایش می دهد. شکل ۸.۲۸. (الف) مجموعه 'A' را نشان می دهد، و شکل ۸.۲۲. (ب) محل های مختلف عنصر ساختمانی دیسکی را در حین فرآیند سایش نشان می دهد. وقتی این فرآیند پایان یابد شکل نامتصول ۸.۲۸. (ب) به دست می آید. در این شکل به حذف اتصال دو بخش اصلی توجه کنید. پنهانی اتصال، باریک تر از قطر عنصر ساختمانی بود؛ یعنی عنصر ساختمانی در این قسمت مجموعه جانمی گرفت، بنابراین شرایط معادله '۸.۴-۶' نقض شدند. این مطلب در مورد دو قسمت سمت راست شیخ نیز صادق بود. شکل ۸.۲۸. (ت) فرآیند گسترش مجموعه 'سایش یافته' را نشان می دهد، و شکل ۸.۲۸. (ث) نتیجه نهایی بازگردان A را نشان می دهد. به طور مشابه اشکال ۸.۲۸. (ج)-(خ)، مراحل بستن A با همان

1) Expand

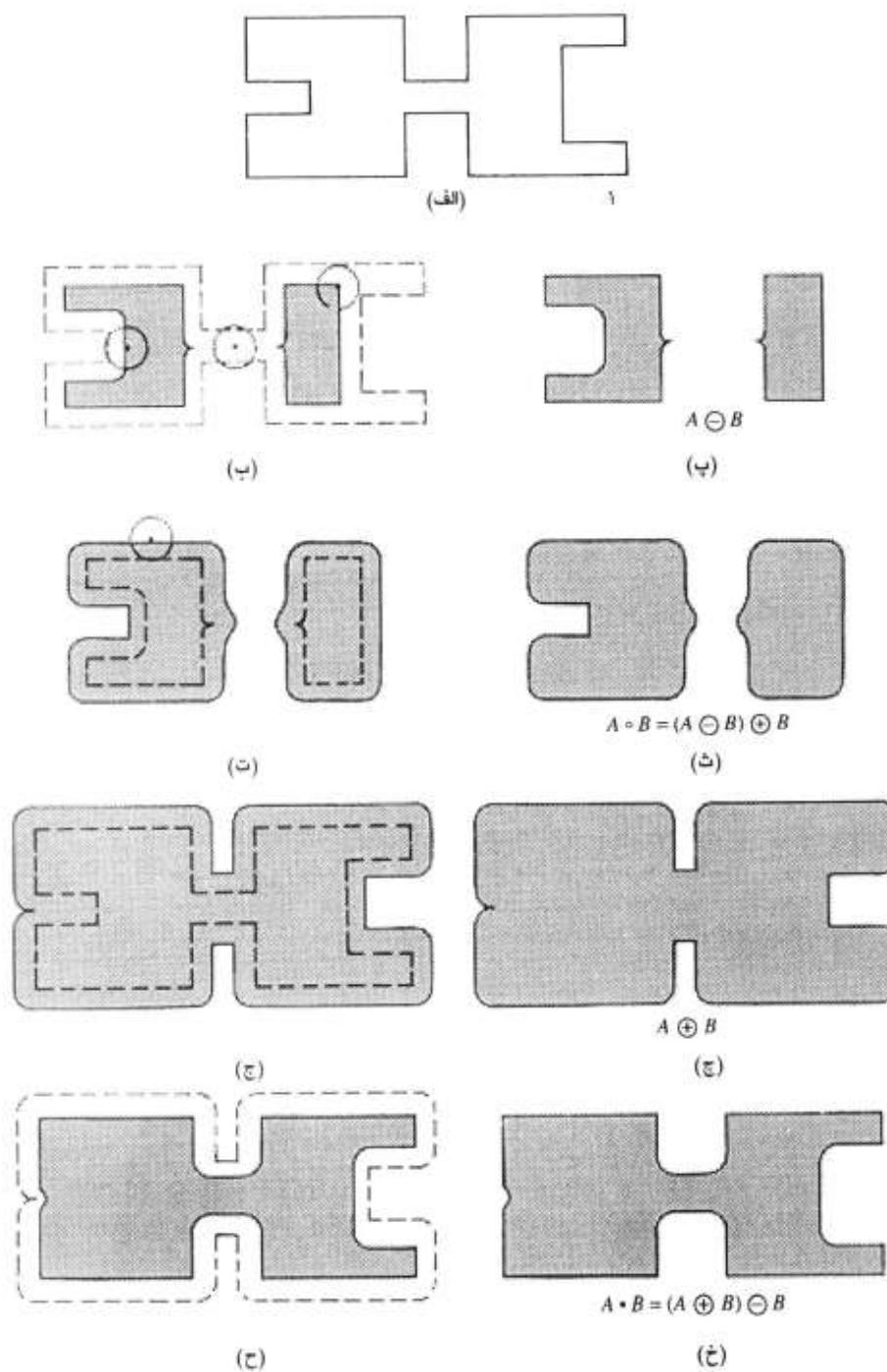
2) Shrink

3) Opening

4) Closing

5) Disk

۶۲۳ شکل شناسی ۴.۸



شکل ۴.۸ تصاویری از عملیات بازگردان و بستن

۶۴۴ پردازش تصویر رقومی

عنصر ساختمانی را نشان می دهند. در نتیجه بستن، شکاف کوچکی (نسبت به B) که در سمت چپ شیء قرار داشت، حذف شد. به هموارسازی به دست آمده در شیء بر اثر دوعمل باز کردن و بستن مجموعه A با یک عنصر ساختمانی دایروی توجه نمایید.

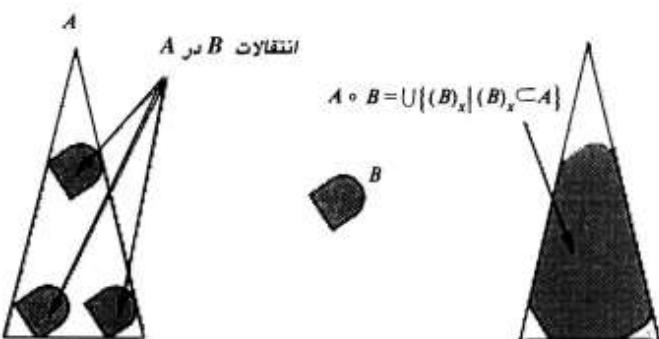
□

باز کردن و بستن تعبیر هندسی ساده ای دارند. به عنوان مثال عنصر ساختمانی دیسکی را به صورت یک "توب چرخان" (تحت) فرض نمایید. آنگاه مرز B دورترین نقطه تا مرز A است که وقتی B از داخل A روی مرز آن غلتانده می شود، مرز B به آن نقاط می رسد. با این تعبیر از گشودن، شکل ۸.۲۸. (ث) حاصل باز کردن شکل ۸.۲۸. (الف) است. توجه کنید که گوشه های رو به بیرون گرد شدند اما گوشه های رو به داخل بی تغییر باقی ماندند. عناصر بیرون زده^۱ که توب در آنها جانگرفت، حذف شدند. وجود خاصیت انطباق هندسی^۲ در عمل باز کردن به یک رابطه مجموعه ای منجر می شود که طبق آن از اجتماع تمام انتقالات B که کاملاً درون A قرار بگیرند، باز شده A با B بدست می آید. بنابراین باز کردن را می توان به صورت یک فرآیند انطباق بیان کرد چنان که

$$A \circ B = \bigcup \{(B)_x \mid (B)_x \subseteq A\}.$$

شکل ۸.۲۹. این مفهوم را نمایش می دهد؛ برای تنویر از یک عنصر ساختمانی غیر دایروی استفاده شد.

بستن نیز تعبیر هندسی مشابهی دارد، با این تفاوت که در مثال توب چرخان، B را از بیرون روی مرز A

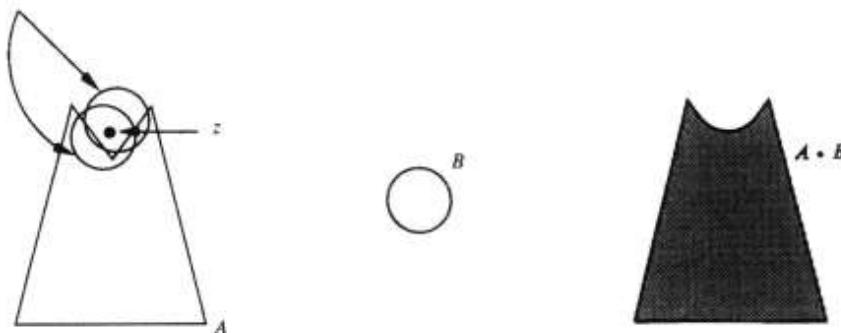


شکل ۸.۲۹. خصوصیت "انطباق" عمل باز کردن. (گرفته شده از [Giardina & Dougherty 1988]).

1) Prototyping

2) Geometric fitting property

۶۲۵ شکل شناسی ۴.۸



شکل ۴.۸۰ تعبیر هندسی عمل بستن، نقطه z ، که درون (B) قرار دارد، به حاصل عمل بستن، $A \cdot B$ ، تعلق دارد اگر و تنها اگر $\cap A \neq \emptyset$.

می غلتانیم (چون باز کردن و بستن دوگان هم هستند، لزوم غلتاندن توب از بیرون دور از انتظار نیست). با این تعبیر ذهنی، شکل ۴.۸۰ (خ) حاصل بستن شکل ۴.۸۰ (الف) است. توجه کنید که گوشه های رویه داخل گرد شدند، درحالی که گوشه های رویه بیرون بدون تغییر باقی ماندند. اندازه تورتفگی سمت چپ A به مقدار زیادی کوچک شد، زیرا توب در آن جانمی گرفت. از نظر هندسی، نقطه z عنصری از $A \cdot B$ است اگر و تنها اگر برای هر انتقال، (B) حاوی z داشته باشیم $\cap A \neq \emptyset$. شکل ۴.۸۰.

این خاصیت را نمایش می دهد.

باز کردن و بستن نیز مانند گسترش و سایش نسبت به اعمال متمم گیری و قرینه گیری مجموعه ای دوگان هم هستند. یعنی،

$$(A \cdot B)^c = (A^c \circ \hat{B}) \quad (4.10-۴)$$

اثبات این نتیجه را به عنوان تمرین باقی می گذاریم.

عمل باز کردن خواص زیر را دارد:

(الف) $B \circ A$ یک زیرمجموعه' (زیر تصویر) A است.

(ب) اگر C زیرمجموعه ای از D باشد، آنگاه $B \circ C$ زیر زیرمجموعه ای از $B \circ D$ است.

$$(A \circ B) \circ B = A \circ B \quad (4.10-۵)$$

به طور مشابه عمل بستن نیز خواص زیر را دارد:

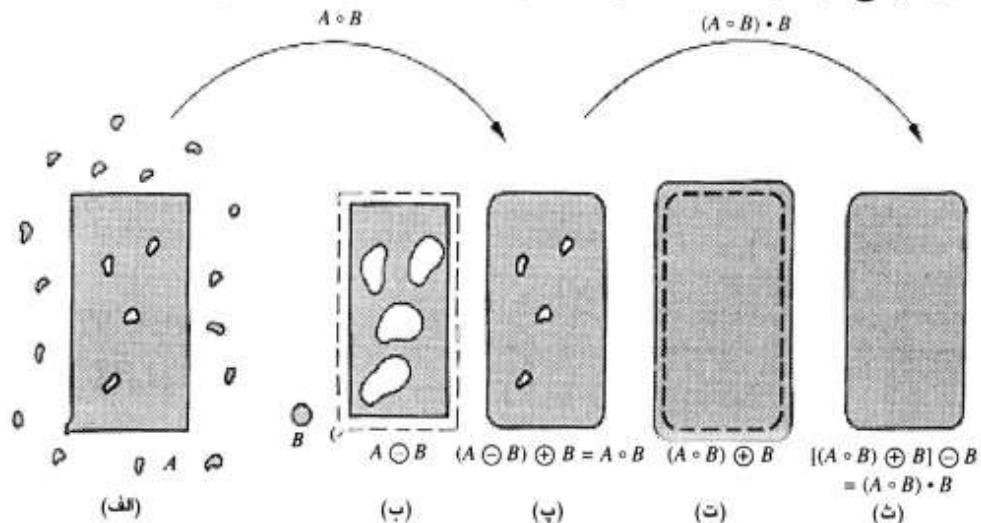
(الف) A یک زیرمجموعه' (زیر تصویر) $B \cdot A$ است.

(ب) اگر C زیرمجموعه D باشد، آنگاه $C \circ B$ نیز زیرمجموعه $D \circ B$ است.

$$(A \circ B) \circ B = A \circ B$$

این خواص به ما در فهم نتایج حاصل از فیلترهای شکل شناسی شامل اعمال باز کردن و بستن کمک می کنند. برای نمونه، ساختمند یک فیلتر مبتنی بر عمل باز کردن را در نظر بگیرید. طبق این خواص: ۱- حاصل آن زیرمجموعه ای از ورودی خواهد بود؛ ۲- یکنواخت^۱ حفظ می شود؛ و ۳- تکرار عمل باز کردن اثری بر نتیجه ندارد. خاصیت اخیر گاهی محدودیت تأثیر^۲ خوانده می شود. این نتایج در مورد عمل بستن نیز صدق می کنند.

تصویر دودویی ساده^۳ موجود در شکل ۸.۳۱(الف) را در نظر بگیرید، که این تصویر مشکل از یک شیء مستطیلی خراب شده با نویز است. در اینجا نویز به صورت عناصر تیره (سایه خورده) روی زمینه^۴ روشن با به صورت فضاهای خالی روشن روی شیء تیره ظاهر می شود. توجه کنید که مجموعه A مشکل از شیء و نویز زمینه است: البته نویز درون شیء، مرزهای داخلی را ایجاد می کند که از داخل این مرزها زمینه دیده می شود. هدف، حذف نویز و اثرات آن بر شیء است طوری که شیء نیز تاحد ممکن کمتر اعوجاج یابد. برای رسیدن به این هدف می توان از فیلتر شکل شناسی $B \circ (A \circ B)$ استفاده کرد.



شکل ۳۱.۸ فیلتر کردن شکل شناسی: (الف) تصویر اولیه نویزدار؛ (ب) حاصل سایش؛ (ب) حاصل باز کردن؛ (ت) حاصل گسترش باز شده A ؛ (ث) نتیجه نهایی که حاصل بستن باز شده A را نشان می دهد. (گرفته شده از Dougherty & Giardina[1988])

1) Monotonicity

2) Idempotence

۴.۸ شکل شناسی ۶۴۲

شکل ۴.۸(ب) حاصل باز کردن A با یک عنصر ساختمانی دیسکی که از تمام اجزای نویزی بزرگتر است، رانشان می دهد. توجه کنید که این عمل، نویز زمینه را حذف می کند اما بر مرزهای درون A اثری ندارد.

در این مثال غیرواقعی، چون تمام اجزای نویزی زمینه کوچکتر از عنصر ساختمانی بودند، در مرحله 'سایش حذف شدند. (به یاد آورید که در عمل سایش، عنصر ساختمانی باید کاملاً درون مجموعه 'تحت سایش قرار بگیرد). با این حال ابعاد اجزای نویزی درون شیئی بزرگ شدند (شکل ۴.۸-۳۱-ب)، که مورد انتظار بود. دلیل آن این است که سوراخ های درون شیئی در واقع مرزهای داخلی آن هستند و وقتی شیئی سایده می شود، باید ابعادشان افزایش باید. در نهایت شکل ۴.۸(ث) حاصل عمل بستن شکل ۴.۸(ب) نشان می دهد. همان طور که شکل ۴.۸(ت) نشان می دهد، مرزهای درون شیئی در مرحله 'گسترش از عمل بستن حذف شدند.

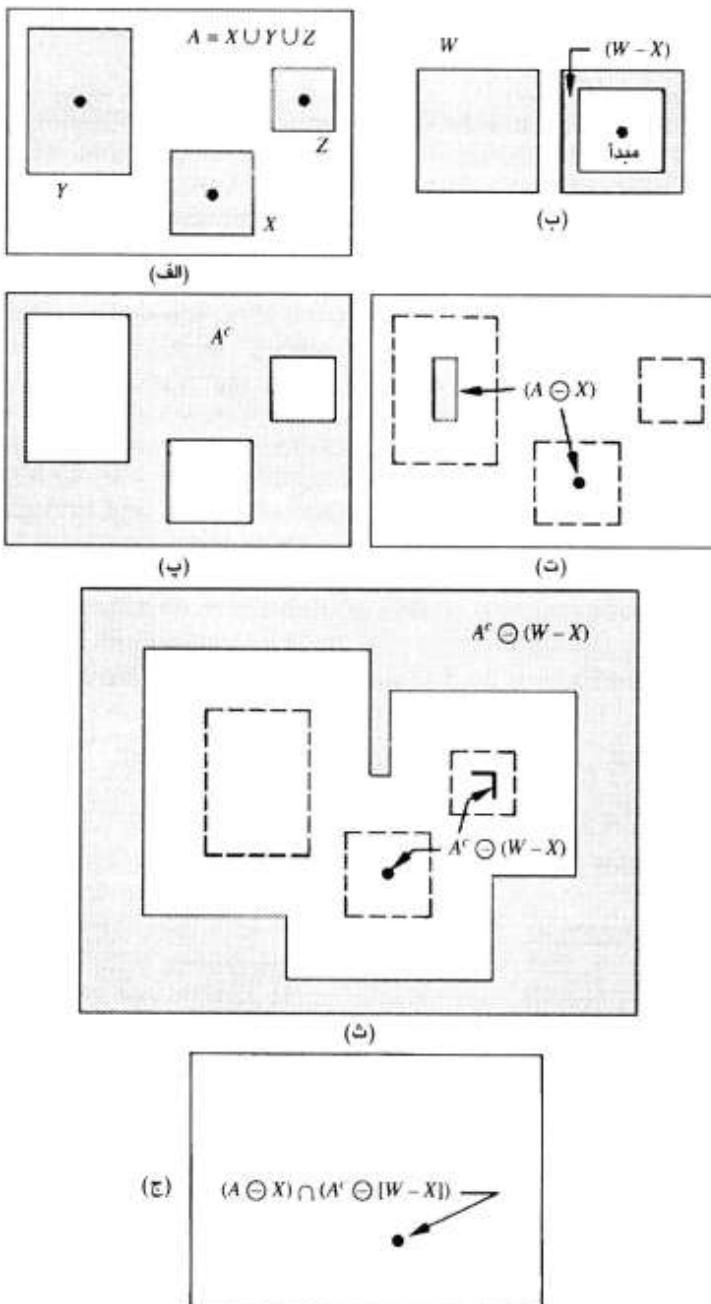
۴.۴.۸ تبدیل HOM

تبدیل شکل شناسی HOM¹ ابزاری پایه ای برای آشکارسازی شکل است. این مفهوم را با کمک شکل ۴.۸-۳۲. که متشکل از سه شکل (زیرمجموعه) بانمادهای X، Y و Z است، معرفی می کنیم. سایه در اشکال ۴.۸-۳۲(الف)-(پ) نشانگر مجموعه های اولیه است، در حالی که سایه در اشکال ۴.۸(ت) و (ث) نشانگر حاصل کاربرد عملگرهای شکل شناسی است. هدف، پیدا کردن محل یکی از شکل ها، مثلاً X، است.

فرض کنید که مبدأ هر شکل در مرکز ثقلش قرار دارد. اگر X را در پنجره 'کوچک W محصور کنیم، زمینه 'محلي² X⁷ نسبت به W، تفاضل مجموعه ای (W-X) است که در شکل ۴.۸-۳۲(ب) نیز دیده می شود. شکل ۴.۸(ب) متمم A را نشان می دهد که بعدها به آن نیاز داریم. شکل ۴.۸-۳۲(ت) سایش A با X را نشان می دهد (خطوط مقطع برای مقایسه افزوده شده اند). به یاد آورید که سایش A با X مجموعه 'مکانهای مبدأ X است چنان که X کاملاً درون A قرار گیرد. شکل ۴.۸-۳۲(ث) سایش متمم A با مجموعه 'زمینه 'محلي X-W-X، رانمایش می دهد؛ ناحیه 'سایه دار بیرونی نیز حاصل سایش است. با توجه به اشکال ۴.۸-۳۲(ت) و (ث) یادآور می شویم که مجموعه مکانهای مبدأ X کاملاً درون A

1) Hit- or- Miss transform

2) Local background



شکل ۳۲.۱ (الف) مجموعه A؛ (ب) پنجه W و زمینه محلی X نسبت به W، (c) (W - X)؛ (د) متمم A؛ (e) حاصل سایش A با X؛ (f) حاصل سایش A^c با (W - X)؛ (g) اشتراک (t) و (ث) که محل مبدأ X را، که مطلوب بود، نشان می دهد. به عنوان راهنمایی پیشنهاد می شود که X و (W - X) را روی کاغذ رسم کنید و این تپلتها را روی A و A^c حرکت دهید تا نتایج موجود در (t) و (ث) بررسی نمایید.

۶۲۹ ۴.۸ شکل‌شناسی

قرار بگیرد، اشتراک^۱ سایش A با X و سایش A^c با $(X - W)$ است که در شکل ۴.۸ (ج) دیده می‌شود. این اشتراک دقیقاً همان محل مورد جستجو است. به عبارت دیگر اگر B نشانگر مجموعه 'مرکب از X و زمینه اش باشد، تطبيق (یا مجموعه 'تطبيق های) B در A ، که با $B \odot A$ نشان داده می‌شود، عبارتست از

$$A \odot B = (A \ominus X) \cap [A^c \ominus (W - X)]. \quad (11-4. 8)$$

با قرار دادن $B = (B_1, B_2)$ ، که B_1 مجموعه 'عناصر شیئی B و B_2 مجموعه 'عناصر زمینه ای است، می‌توان مقداری این رابطه را تعمیم داد. در شرح قبلی با فرض $X = B_1$ و $B_2 = (W - X)$ داریم:

$$A \odot B = (A \ominus B_1) \cap (A^c \ominus B_2). \quad (12-4. 8)$$

با استفاده از تعریف نفاضل های مجموعه ای و رابطه 'دوگان بین سایش و گسترش می‌توان معادله^۲ (12-4. 8) را به صورت زیرنویز نوشت:

$$A \odot B = (A \ominus B_1) - (A \oplus \hat{B}_2). \quad (13-4. 8)$$

بنابراین مجموعه $B \odot A$ حاوی تمام نقاطی است که در آنها هم B_1 تطبیقی در A و هم B_2 تطبیقی در A^c بیابد.

۴.۴.۸ چند الگوریتم پایه ای شکل‌شناسی

پس از آشنایی با مطالعه پایه ای شکل‌شناسی، اکنون می‌توانیم چند کاربرد عملی شکل‌شناسی را بررسی کنیم. در برخورد با تصاویر دودویی، کاربرد اصلی شکل‌شناسی استخراج اجزایی از تصویر است که در نمایش و توصیف شکل مفید است. ما به ویژه الگوریتم های شکل‌شناسی مورد استفاده برای استخراج مرزها، اجزای متصل، بدنه' محدب، و اسکلت ناحیه را مورد بررسی قرار می‌دهیم. همچنین چند روش (برای پُر کردن ناحیه، باریک سازی، ضخیم سازی، و هرس کردن) را که اغلب همراه این الگوریتم ها به عنوان مراحل پیش-یا پس پردازش به کار می‌روند، ارائه می‌کنیم. برخلاف تصاویر قبلی، اغلب تصاویری

¹) Intersection

که در ادامه 'این بخش استفاده می شوند، "تصاویر کوچکی"^۱ هستند که برای روش ساختن طرز عمل هر فرآیند شکل شناسی در هنگام معرفی آن، طراحی شده اند. این تصاویر دودویی هستند که در آنها '۱'ها با سایه و '۰'ها با سفید نشان داده می شوند.

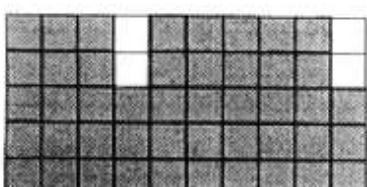
استخراج مرز

مرز مجموعه 'A'، که با $\beta(A)$ نشان داده می شود، را می توان با تفاضل مجموعه A و حاصل سایش A با B بدست آورد. یعنی،

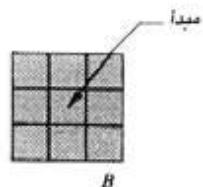
$$\beta(A) = A - (A \ominus B) \quad (14-4. ۸)$$

که B یک عنصر ساختمانی مناسب است.

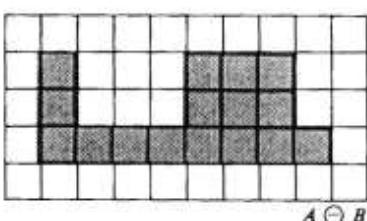
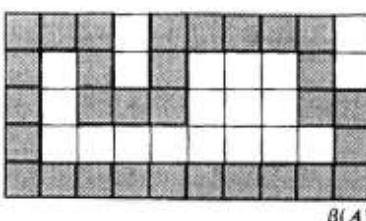
شکل ۳۳. ۸ چگونگی استخراج مرز رانمایش می دهد. این شکل، شیوه دودویی A، عنصر ساختمانی B، و حاصل کاربرد معادله '۸. (۱۴-۴.) را نشان می دهد. گرچه عنصر ساختمانی موجود در شکل ۸ (۳۳. (ب) بسیار متداول است، اما به هیچ وجه یکتا نیست. مثلاً با کاربرد عنصر ساختمانی ۵ × ۵ از ۱'ها مرزی نتیجه می شود که ۲ یا ۳ یکسل ضخامت دارد. توجه کنید که وقتی مبدأ B روی مرز



(الف) A



(ب)

(ب) $A \ominus B$ (ت) $\beta(A)$

شکل ۳۳.۸ (الف) مجموعه 'A'; (ب) عنصر ساختمانی 'B'; (پ) سایش یافته 'A' با 'B'; (ت) مرز حاصل از تفاضل مجموعه ای 'A' و سایش آن.

۶۳۱ شکل‌شناسی ۴.۸

مجموعه باشد، بخشی از عنصر ساختمانی، خارج از مجموعه قرار می‌گیرد. برخورد مناسب با این وضعیت این است که مقادیر خارج از مرز مجموعه را به طور ضمنی "۰" فرض نماییم.

برکرد ناحیه^۱

اکنون الگوریتم ساده‌ای براساس گسترش، متمم گیری، و اشتراک^۲ برای پرکردن یک ناحیه ارائه می‌کنیم. در شکل ۴.۸، مجموعه^۳ A زیرمجموعه‌ای دارد که عناصر آن نقاط مرزی متصل^۴ گانه^۵ یک ناحیه هستند. هدف این است که با شروع از یک نقطه^۶ p درون مرز، کل آن ناحیه را با "۱" برنماییم. طبق فرض چون تمام نقاط غیرمرز برحسب "۰" دارند، برای شروع، مقدار "۱" را به p متنسب می‌کنیم. آنگاه روال زیر ناحیه را با "۱" پرمی‌کنند:

$$X_k = (X_{k-1} \oplus B) \cap A^c \quad k = 1, 2, 3, \dots \quad (4.8)$$

که $p = X_0$ ، و B عنصر ساختمانی متقابن موجود در شکل ۴.۳(پ) است. اگر $X_{k-1} = X_k$ باشد، این الگوریتم در k آمین مرحله تکرار پایان می‌یابد. اجتماع^۷ X_k و A شامل مجموعه^۸ پرشده با "۱" ها و مرز آن می‌باشد.

فرآیند گسترش ارائه شده با معادله (۴.۸) اگر بدون محدودیت رها می‌شد، کل ناحیه^۹ تصویر را پر می‌کرد. لذا در این معادله یا کمک عمل اشتراک با A^c ، گسترش به درون ناحیه^{۱۰} محدود می‌گردد (گاهی اوقات به این نوع فرآیند محدود کننده، گسترش شرطی^{۱۱} گفته می‌شود). قسمت‌های بعدی شکل ۴.۸ طرز کار معادله (۴.۸) را بهتر روشن می‌سازند. گرچه در این مثال مجموعه^{۱۲} A تنها یک زیرمجموعه دارد، اما واضح است وقتی که از درون مرز هر زیرمجموعه یک نقطه داده شود، این مفهوم برای هر تعداد محدود از چنین زیرمجموعه‌هایی صادق است.

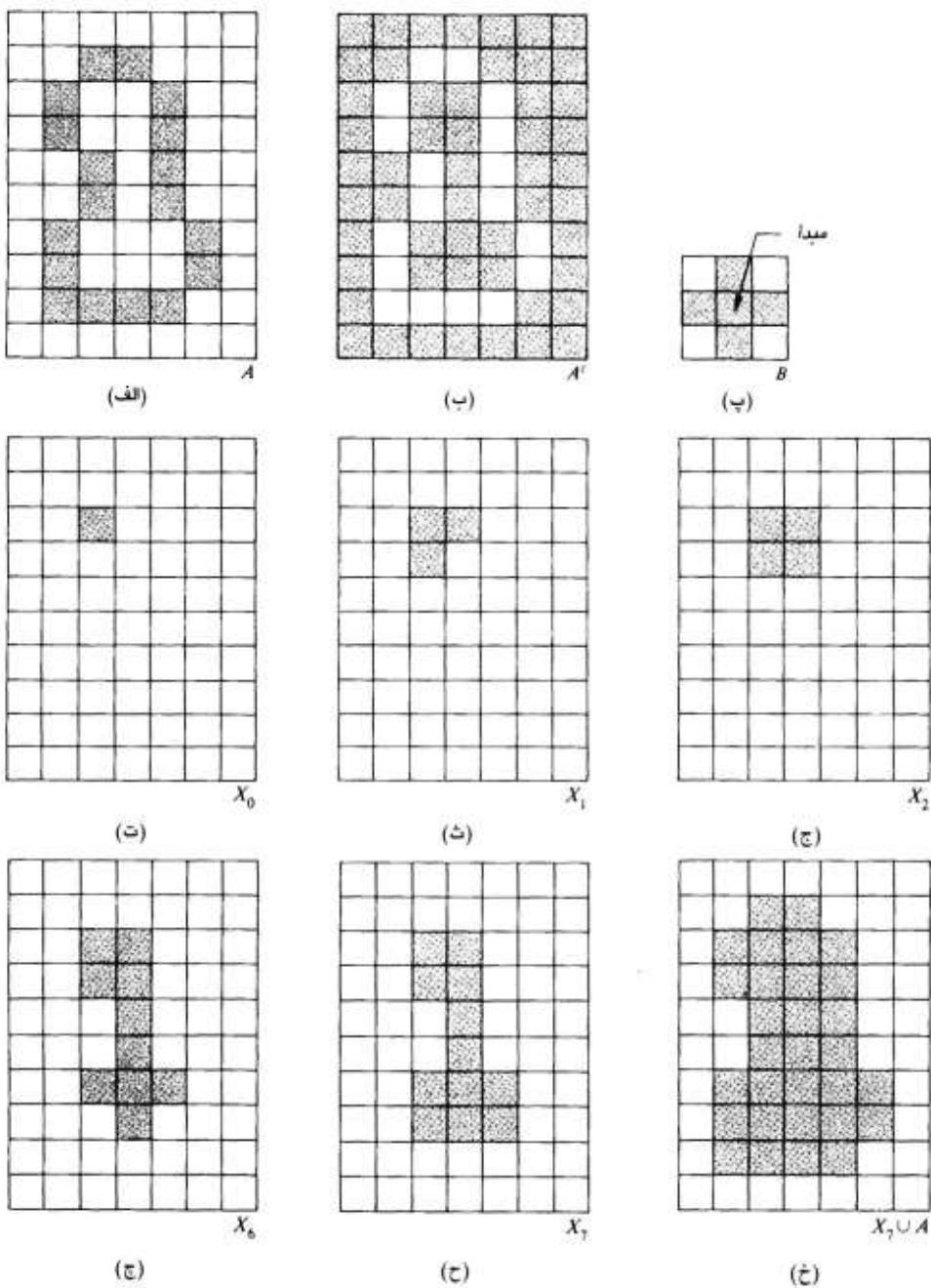
استخراج اجزای متصل

در بخش ۴.۲ مفهوم اتصال را معرفی کردیم. در عمل، استخراج اجزای متصل در تصویر دودویی اساس بسیاری از کاربردهای تحلیل تصویر خودکار است. فرض کنید^{۱۳} یانگر یک جزء متصل است که

1) Region filling

2) Intersection

3) Conditional dilation

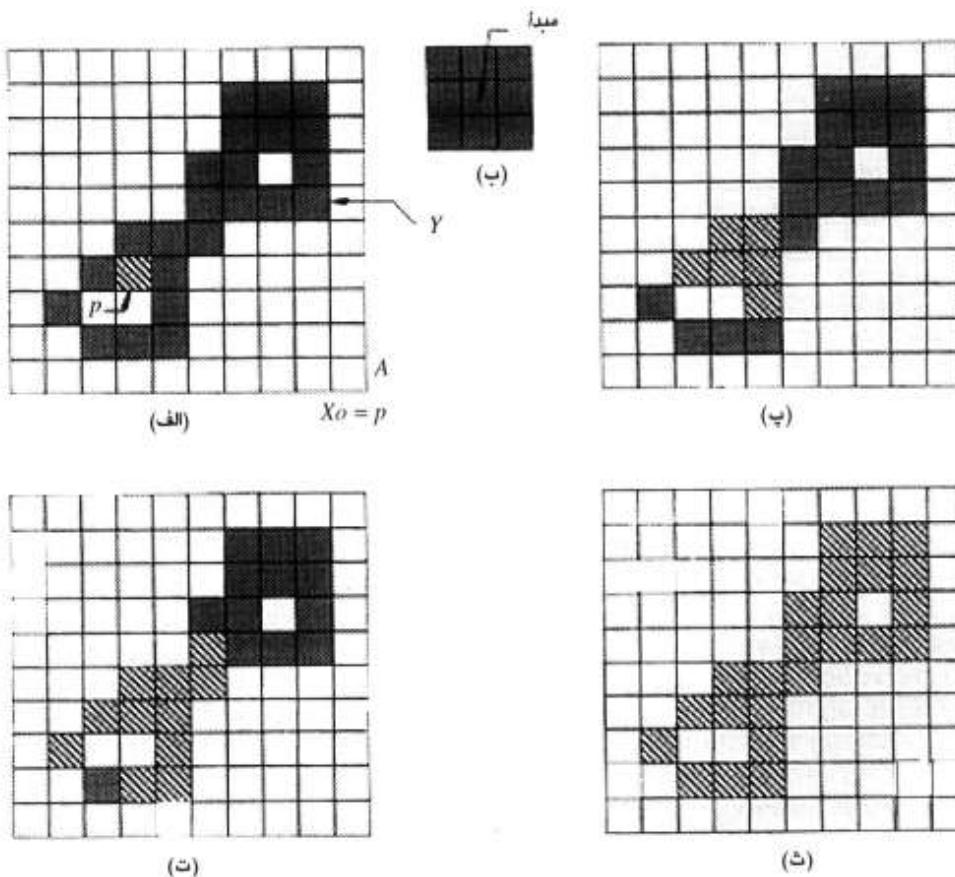


شکل ۳۴.۸ (الف) مجموعه A که حاوی یک زیرمجموعه مرزی است؛ (ب) متمم A ؛ (پ) عنصر ساختمانی B ؛ (ت) نقطه ابتدایی درون مرز؛ (ث) - (ح) مراحل مختلف معادله A . (خ) نتیجه نهایی که با اجتماع مجموعه ای (الف) و (ج) بدست آمد.

۴.۸ شکل شناسی ۶۳۳

درون مجموعه 'A' قرار دارد و فرض کنید که یک نقطه 'p' از Y معلوم است. آنگاه آن طور که در شکل ۳۵.۸ دیده می شود، عبارت تکراری زیر تمام عناصر Y را می دهد:

$$X_k = (X_{k-1} \oplus B) \cap A \quad k = 1, 2, 3, \dots \quad (16-4.8)$$



شکل ۳۵.۸ (الف) مجموعه 'A' که حاوی جزء متصلب Y و نقطه ابتدایی P است (تمام نقاط مایه دار مقدار ۱ دارند، اما متفاوت از P نمایش داده شده اند تا نشان دهنده این باشد که هنوز بوسیله الگوریتم پیدا نشده اند؛ (ب) عنصر ساختمانی؛ (پ) حاصل اولین مرحله تکراری؛ (ت) حاصل مرحله دوم؛ (ث) نتیجه نهایی.

که $X_0 = p$ ، و B یک عنصر ساختمانی مناسب است. هرگاه $X_k = X_{k-1}$ شود، الگوریتم همگرا شده است و قرار می دهیم $Y = X_k$.

معادله (۱۶-۴.۸) شکلی مشابه معادله (۱۵-۴.۸) دارد. تنها تفاوت این است که در اینجا به جای

متهم A از خود استفاده می شود، زیرا تمام عناصر مورد جستجو (یعنی، عناصر جزء متصل) برجسب "۱" خورده اند. در هر مرحله تکرار، اشتراک گیری با A ، گسترش های به مرکز عناصر با برجسب "۰" را حذف می کند. شکل ۸. ۳۵ طرز کار معادله (۸. ۱۶-۴) را نمایش می دهد. توجه کنید که در این شکل از عنصر ساختمانی، پیکسل ها متصل ۸ گانه فرض می شوند. مانند الگوریتم پرکردن ناحیه، نتایج اخیر برای هر تعداد محدودی از اجزای متصل درون A ، قابل استفاده است.

بدنه محدب

در بخش ۸. ۱. ۴ بدنه محدب یک مجموعه را به عنوان توصیف گر مقیدی از تصویر معرفی کردیم. در اینجا یک الگوریتم ساده شکل شناسی را برای به دست آوردن بدنه محدب مجموعه $C(A)$ ، ارائه می دهیم. فرض کنید که B^i ، $i = 1, 2, 3, 4$ ، نمایانگر چهار عنصر ساختمانی باشند. اجرای الگوریتم، در واقع اجرای معادله زیر است:

$$X'_i = (X \oplus B^i) \cup A \quad i = 1, 2, 3, 4 \quad \text{و} \quad k = 1, 2, 3, \dots \quad (۱۷-۴. ۸)$$

که $X'_0 = A$. اکنون فرض کنید $D' = X'_{\text{conv}}$ "conv" نشانگر همگرایی به مفهوم است. آنگاه بدنه محدب A عبارتست از

$$C(A) = \bigcup_{i=1}^4 D' \quad (۱۸-۴. ۸)$$

به عبارت دیگر، روال کار شامل اعمال تکراری تبدیل HoM بر A با B^i است؛ وقتی تغییر بیشتری در نتیجه رخ ندهد، اجتماع آن با A را به دست آورده و آن را D' می خوانیم. سپس با همین روال و با استفاده از B^2 ، B^3 را به دست می آوریم و به همین ترتیب الى آخر. اجتماع چهار D' به دست آمده بدنه محدب A می باشد.

مثال: شکل ۸. ۳۶ روال بیان شده با معادلات (۸. ۱۷-۴) و (۸. ۱۸-۴) را نمایش می دهد. شکل ۸. ۳۶ (الف) عنصر های ساختمانی به کار رفته در استخراج بدنه محدب را نشان می دهد (مبدأ هر عنصر در مرکز آن است). شکل ۸. ۳۶ (ب) مجموعه A را که می خواهیم بدنه محدبش را بدست آوریم، نشان می دهد. با شروع از $A'_0 = A$ و پس از چهار بار تکرار معادله (۸. ۱۷-۴) مجموعه نشان داده شده در شکل ۸. ۳۶ (ب) به دست آمد. آنگاه با فرض $A'_0 = A$ و استفاده از معادله (۸. ۱۷-۴)

۶۳۵ ۴.۸ شکل شناسی

مجموعه 'نشان داده شده در شکل ۸ . ۳۶(ت) بdst آمد. (توجه کنید که این مجموعه تنها با دو بار اجرای معادله ۸ . ۱۷-۴ به dst آمد). دو نتیجه' بعدی نیز به همان طریق بdst آمدند. درنهایت با اجتماع مجموعه های درون اشکال ۸ . ۳۶(پ)، (ت)، (ث)، (ج) بدنی' محدب نشان داده شده در شکل ۸ . ۳۶(ج) بdst آمد. در مجموعه' مرکب نشان داده شده در شکل ۸ . ۳۶(ح) سهم هر عنصر ساختمانی مشخص شده است.

□

باریک سازی

باریک سازی مجموعه' A با عنصر ساختمانی B ، که با $B \otimes A$ نشان داده می شود، را می توان بر حسب تبدیل HoM تعریف کرد:

$$\begin{aligned} A \otimes B &= A - (A \oplus B) \\ &= A \cap (A \oplus B)^c. \end{aligned} \quad (19-4 . \Delta)$$

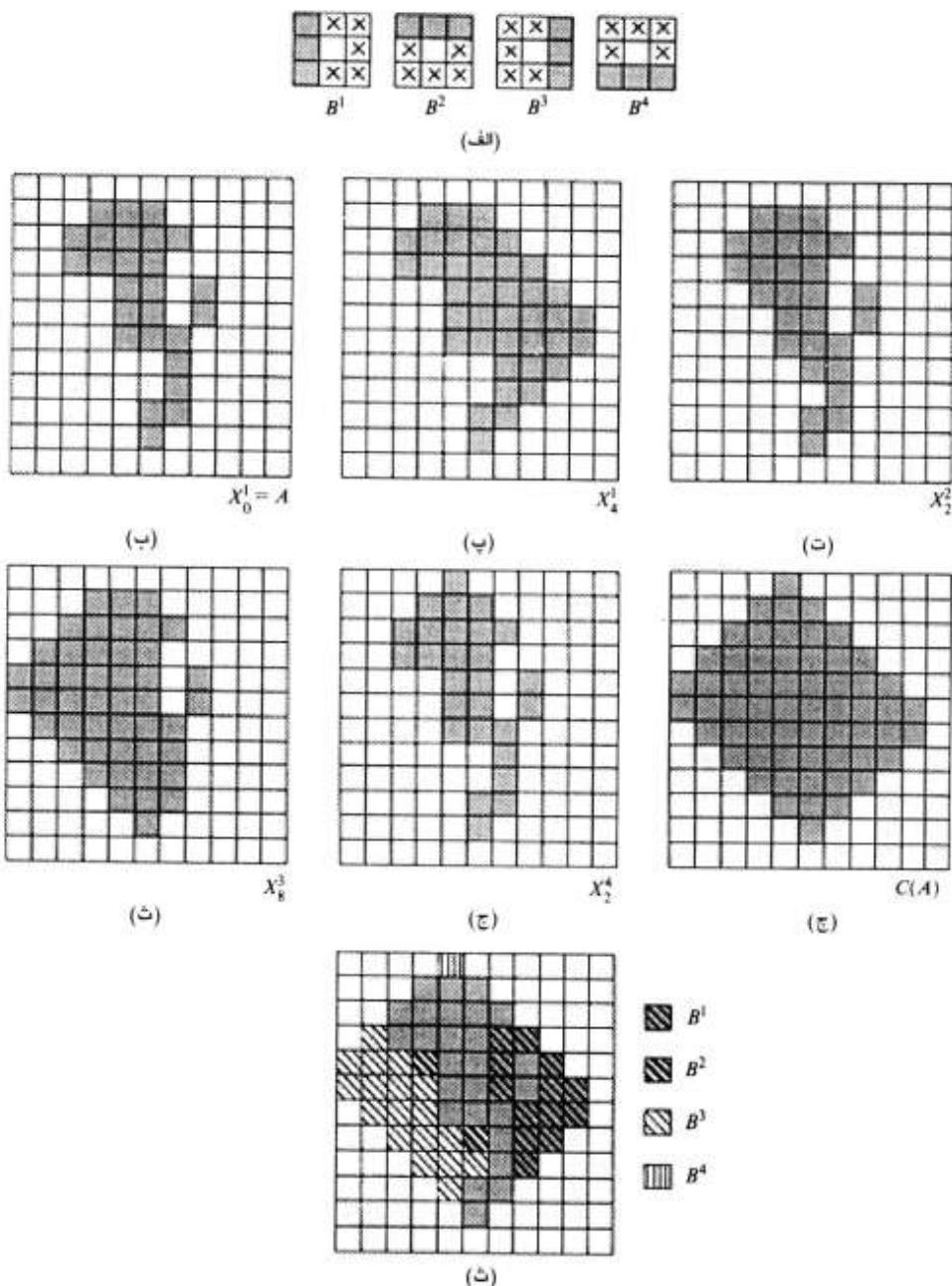
عبارتی مفیدتر برای باریک سازی متقارن A ، مبنی بر دنباله ای از عناصر ساختمانی است:

$$\{B\} = \{B^1, B^2, B^3, \dots, B^n\} \quad (20-4 . \Delta)$$

که B^i یک نسخه' چرخانده شده از B^{i-1} است. اکنون با استفاده از این مفهوم، باریک سازی با دنباله عناصر ساختمانی را به صورت زیر تعریف می کنیم

$$A \otimes \{B\} = ((\dots ((A \otimes B^1) \otimes B^2) \dots) \otimes B^n). \quad (21-4 . \Delta)$$

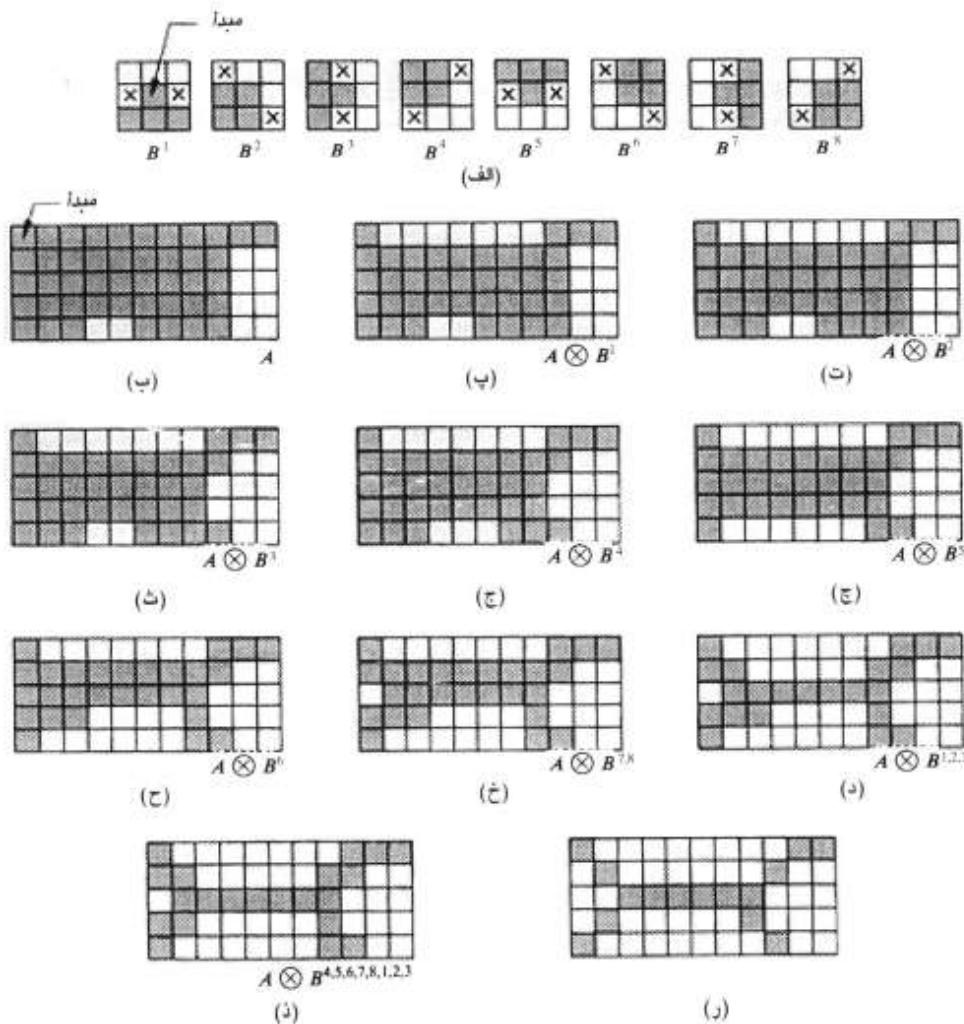
به بیان دیگر هر تکرار فرآیند بدین صورت است که ابتدا A را با یکبار مرور B^1 باریک می کنیم، سپس حاصل را با یک بار مرور B^2 باریک می کنیم، و بدین ترتیب ادامه می دهیم، تا این که در انتها حاصل را با یک بار مرور B^n باریک می کنیم. کل فرآیند تا جایی تکرار می شود که دیگر هیچ تغییری رخ ندهد.
 مثال: شکل ۸ . ۳۷ (الف) مجموعه ای از عناصر ساختمانی را که معمولاً برای باریک سازی استفاده می شود، نشان می دهد و شکل ۸ . ۳۷ (ب) مجموعه' A را که باید با این روش باریک شود، نشان می دهد. شکل ۸ . ۳۷ (پ) حاصل باریک سازی A با یک بار مرور جاروبی B^1 را نشان می دهد، و اشکال ۸ . ۳۷ (ت)-(ذ) حاصلهای مرور تصویر با سایر عناصر ساختمانی را نشان می دهند. همگرایی بعد از دو میان مرور B^4 به dst آمد. شکل ۸ . ۳۷ (ذ) نتیجه باریک شده را نشان می دهد. درنهایت شکل



شکل ۳۶.۸ (الف) عناصر ساختمانی استفاده شده در الگوریتم بدنۀ محدب (\times هانشانه، \ominus اهمیت، \oplus هستند و B^i حاصل چرخش ۹۰ درجه، B^i است)؛ (ب) مجموعه A ؛ (پ)- $(ج)$ نتایج همگرایی با چهار عنصر ساختمانی شکل (الف)؛ (ج) بدنۀ محدب؛ (ح) بدنۀ محدب که سهم هر عنصر ساختمانی را نشان می دهد.

۶۳۷ شکل شناسی ۴.۸

۴.۸ (ر) مجموعه 'باریک شده ای را نشان می دهد که برای حذف مسیرهای چندگانه به اتصال m گانه (بخش ۲. ۴. ۲ را بینید) تبدیل شده است.



شکل ۴.۸ (الف) دنباله 'عناصر ساختمانی چرخش یافته' مورد استفاده برای باریک سازی؛ (ب) مجموعه 'A' (پ) حاصل باریک سازی با اولین عنصر؛ (ت)- (خ) نتایج باریک سازی با هفت عنصر بعدی (هیچ تفاوتی بین نتایج کاربرد متواالی دو عنصر ساختمانی هفتم و هشتم وجود نداشت)؛ (د) حاصل استفاده مجدد از اولین عنصر ساختمانی (هیچ تفاوتی بین نتایج کاربرد متواالی سه عنصر ساختمانی اول وجود نداشت)؛ (ذ) حاصل همگرا شده؛ (ر) تبدیل به اتصال m گانه.

ضخیم سازی

ضخیم سازی، دو گان باریک سازی در حوزه' شکل شناسی است و با عبارت

$$A \odot B = A \cup (A \oplus B) \quad (22-4, A)$$

تعریف می شود که B عنصر ساختمانی مناسب برای ضخیم سازی است. طبق گذشته، ضخیم سازی را می توان به صورت یک عمل دنباله ای تعریف کرد:

$$A \odot \{B\} = ((\dots ((A \odot B^1) \odot B^2) \dots) \odot B^n). \quad (23-4, A)$$

عناصر ساختمانی مورد استفاده برای ضخیم سازی، هم شکل عناصر ساختمانی باریک سازی در شکل ۸. ۳۷ (الف) هستند، البته با این تفاوت که تمام "۱" ها و "۰" ها با هم تعویض شده اند. با این حال در عمل به ندرت از یک الگوریتم ضخیم سازی جداگانه استفاده می شود. در عوض روال معمول این است که ابتدا زمینه' مجموعه' مورد نظر را باریک کنیم و سپس حاصل را متمم نماییم. به بیان دیگر، برای ضخیم سازی مجموعه' A ، $C = A^c$ را تشکیل داده، C را باریک کرده و سپس C^c را بدست می آوریم. شکل ۸. ۳۸. این روال را نمایش می دهد.

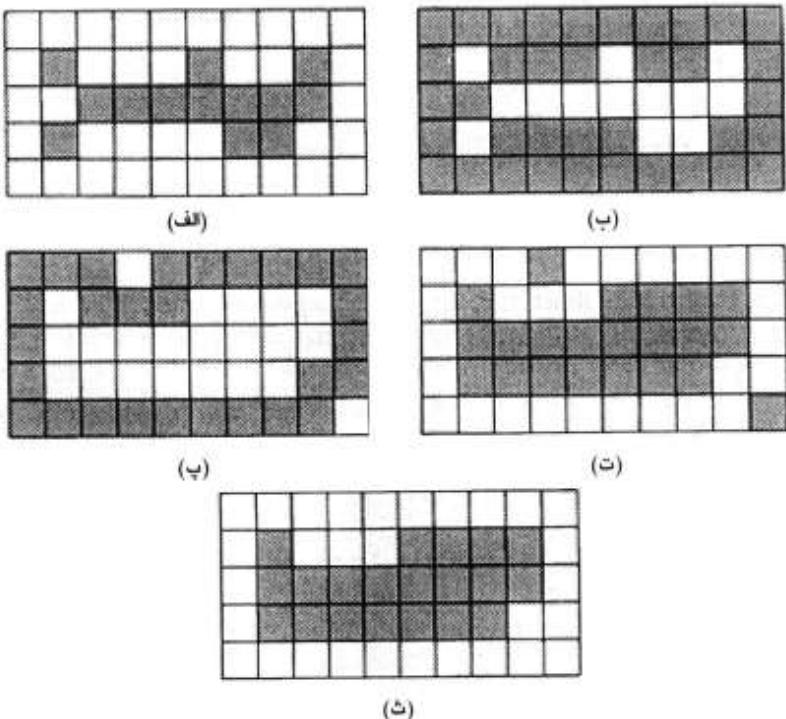
همان طور که در شکل ۸. ۳۸. (ت) دیده می شود، بسته به طبیعت A ممکن است این روال چند نقطه' منفرد را نیز نتیجه دهد. بنابراین بعد از ضخیم سازی تصویر با این روش، برای حذف نقاط منفرد معمولاً یک مرحله پس-پردازش ساده اجرا می شود. در شکل ۸. ۳۸. (پ) توجه کنید که زمینه' باریک شده، مرزی برای فرآیند ضخیم سازی ایجاد می کند. این ویژگی مفید، در پیاده سازی مستقیم ضخیم سازی با استفاده از معادله' (۲۳-۴) وجود ندارد، و یکی از دلایل اصلی استفاده از باریک سازی زمینه'، اجرای ضخیم سازی است.

اسکلت ها

مفهوم «اسکلت» و استخراج آن از ناحیه را در بخش ۱. ۸. ۵ معرفی کردیم. در اینجا از دیدگاه شکل شناسی آن را بررسی می کنیم.

اسکلت (Serra[1982] & Maragos[1987]) نشان داد که اسکلت

۶۳۹ شکل‌شناسی ۴.۸



شکل ۳۸.۸ ضخیم‌سازی از راه باریک سازی زمینه: (الف) مجموعه A ; (ب) متمم آن؛ (پ) حاصل باریک سازی متمم A ; (ت) مجموعه ضخیم شده حاصل از متمم کردن (پ)؛ (ث) نتیجه نهایی که حدف نقاط منفرد را نشان می‌دهد

مجموعه (ناحیه) A را می‌توان بر حسب سایش و باز کردن بیان کرد. یعنی، اگر $S(A)$ نشانگر اسکلت A باشد، می‌توان نشان داد که

$$S(A) = \bigcup_{k=0}^K S_k(A) \quad (24-4.8)$$

و

$$S_k(A) = \bigcup_{k=0}^K \{(A \ominus kB) - [(A \ominus kB) \circ B]\} \quad (25-4.8)$$

که B یک عنصر ساختمانی و $(A \ominus kB)$ یانگر k بار سایش متوالی A با عنصر ساختمانی B است؛ یعنی

$$(A \ominus kB) = ((\dots (A \ominus B) \ominus B) \ominus \dots) \ominus B$$

و K آخرین مرحله' تکرار است قبل از این که A به مجموعه' تهی سایش یابد. به عبارت دیگر،

$$K = \max \{k \mid (A \ominus kB) \neq \emptyset\}.$$

همانند گذشته، نماد $(^o)$ برای نشان دادن عمل باز کردن در معادله' (۲۵-۴) استفاده می شود.

روابط داده شده با معادلات (۲۴-۴) و (۲۵-۴) بیان می کنند که $S(A)$ ، اسکلت A ، رامی توان با اجتماع زیرمجموعه های اسکلتی^۱ ($S_i(A)$) بدست آورد. همچنین می توان نشان داد که با استفاده از معادله'

$$A = \bigcup_{k=0}^K (S_k(A) \oplus kB) \quad (26-4, ۸)$$

می توان A را بازسازی کرد، که در این رابطه $(S_k(A) \oplus kB)$ یانگر k بار گسترش متوالی ($S_k(A)$ با عنصر ساختمانی B است، یعنی

$$(S_k(A) \oplus kB) = ((\dots (S_k(A) \oplus B) \oplus B) \oplus \dots) \oplus B$$

و حد بالای اجتماع گیری^۲، K ، مانند قبل است.

مثال: شکل ۳۹. ۸ مفاهیم اخیر را با تصویر نشان می دهد. ستون اول، مجموعه' اولیه (در بالای ستون) و دوسایش متوالی آن با عنصر ساختمانی B را نشان می دهد. توجه کنید که با بیش از دو بار سایش A مجموعه های ستون اول با عنصر ساختمانی B را نمایش می دهد. با استفاده از خصوصیت انطباق عمل باز کردن که در رابطه با شکل ۲۹. ۸ تشریح شد، به آسانی می توان نتایج ستون دوم را توضیح داد. ستون سوم حاوی تفاضل های مجموعه ای اشکال ستون اول و دوم است.

ستون چهارم حاوی دو اسکلت ناقص و اسکلت نهایی (در پایین ستون) است. اسکلت نهایی نه تنها

1) Skeleton subset

2) Limit of the summation

۶۴۱ شکل‌شناسی ۴.۸

بیش از حد لازم ضخیم است، بلکه مهمتر این که متصل هم نیست. این نتیجه دور از انتظار نیست، زیرا در روابط قبلی اسکلتی سازی هیچ اتصال را تضمین نمی‌کند. شکل‌شناسی، بیانی قوی بر حسب سایش‌ها و بازکردن‌های مجموعه^۱ داده شده تولید می‌کند. با این حال اگر، طبق معمول، لازم باشد که اسکلت تا حد ممکن باریک، متصل و کمتر سایش یافته باشد، باید از روش‌های مکاشفه‌ای^۱ نظریه‌گوریتی که در بخش ۱.۸ مطرح شد، استفاده کرد.

ستون پنجم ($S_0(A) \oplus B, S_1(A) \oplus B, S_2(A) \oplus 2B$) = $(S_2(A) \oplus B) \oplus B$ را نمایش می‌دهد. در نهایت ستون آخر باز سازی مجموعه^۱ A را نمایش می‌دهد، که طبق معادله (۸.۴-۲۶) اجتماع زیر مجموعه‌های اسکلتی گسترش یافته^۱ موجود در ستون پنجم است. □

هرس کردن

روش‌های هرس کردن یک مکمل ضروری برای الگوریتم‌های باریک سازی و اسکلت سازی هستند، چرا که این الگوریتم‌ها، اجزایی پارازیتی در تصویر باقی می‌گذارند که باید با پس پردازش حذف شوند. شرح را با مسئله هرس کردن آغاز کرده و سپس یک راه حل شکل‌شناسی براساس مطالب بخش‌های قبلی معرفی می‌کنیم. بنابراین ما از آنچه تاکنون آموخته کرده و نشان می‌دهیم که چگونه می‌توان با ترکیب روشهای از قبیل معرفی شده یک مسئله را حل نمود.

یک راه عمومی برای تشخیص خودکار نویسه‌های دست نویس تحلیل اسکلت هر نویسه است. اسکلت‌ها غالباً با اجزای پارازیتی که در حین سایش و برای نایکنواختی خطوط تشکیل دهنده^۱ نویسه بوجود می‌آیند، همراه هستند. با این فرض که طول یک جزء پارازیتی از سه پیکسل تجاوز نمی‌کند، یک روش شکل‌شناسی برای حل این مشکل را ارائه می‌کنیم.

شکل ۸.۴۰ (الف) اسکلت حرف "a" دست نویس را نشان می‌دهد. جزء پارازیتی در سمت چپ نویسه همان چیزی است که می‌خواهیم حذف کنیم. راه حل حذف شاخه^۱ پارازیتی براساس حذف متواലی نقاط انتهایش می‌باشد. البته این کار سایر شاخه‌های نویسه را نیز کوتاه (یا حذف) می‌کند، بنابراین در غیاب سایر اطلاعات ساختاری، فرض می‌شود که هر شاخه^۱ با طول سه پیکسل یا کمتر باید حذف شود. برای هر مجموعه^۱ ورودی A ، باریک سازی A با دنباله‌ای از عناصر ساختمانی که تنها برای آشکارسازی

1) Heuristic formulation

$k \setminus A \ominus kB$	$(A \ominus kB) \circ B$	$S_k(A)$	$\bigcup_{k=0}^K S_k(A)$	$S_k(A) \oplus kB$	$\bigcup_{k=0}^K S_k(A) \oplus kB$
0					
1					
2					

A
 B

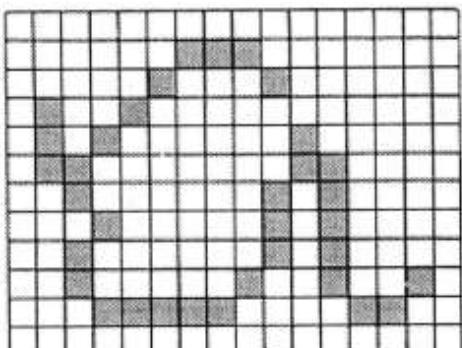
شکل ۳۹.۸ ۳۹ مثالی از پاده سازی معادلات $(A \ominus kB) \circ B = A - S_k(A)$. مجموعه اولیه در بالای ستون سمت چپ دیده می شود، و اسکلت شکل شناسی آن در پایین ستون چهارم دیده می شود. مجموعه بازسازی شده در پایین ستون ششم دیده می شود.

نقاط پایانی طرح شده است، نتیجه مطلوب را می دهد. یعنی از رابطه زیر استفاده کنید:

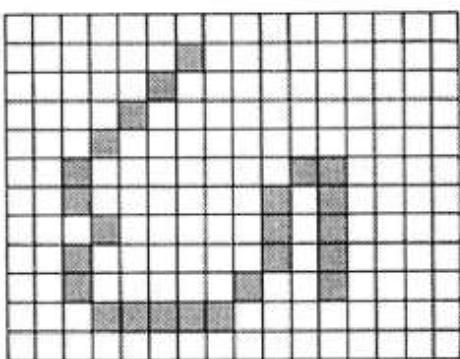
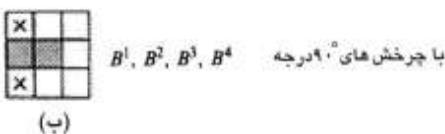
$$X_1 = A \otimes \{B\} \quad (27-4, 8)$$

که $\{B\}$ نشانگر دنباله [معادله ۲۰-۴] را بینند. نشان داده شده در اشکال ۸.۴۰ (ب) و (پ) می باشد. دنباله عناصر ساختمانی، مشکل از دو ساختار متفاوت است، که هریک از این دو ساختار

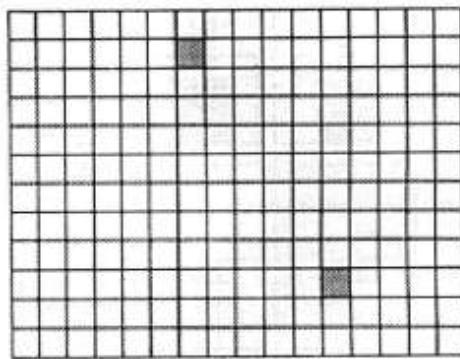
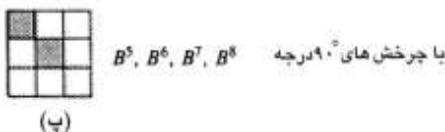
۶۴۳ شکل شناسی ۴.۸



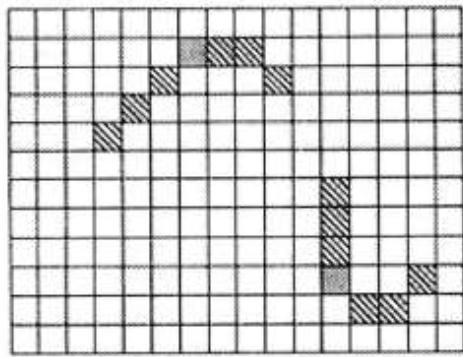
(الف)



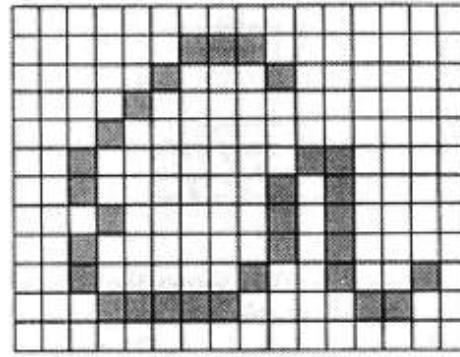
(ت)



(ث)



(ج)



(ه)

شکل ۴.۸ مثالی از هرس کردن: (الف) مجموعه اولیه؛ (ب) و (پ) عناصر ساختمانی استفاده شده برای حذف (باریک سازی) نقاط انتهایی؛ (ت) حاصل سه بار باریک سازی؛ (ث) نقاط پایانی (ت)؛ (ج) حاصل گسترش نقاط انتهایی که باید زیر مجموعه شکل (الف) باشد؛ (ج) تصویر هرس شده

چهار بار با زاویه 90° درجه چرخانده می شود تا جمعاً ۸ عنصر ساختمانی ایجاد شود. علامت \times در شکل

۴۰. ۸ (ب) نشانه 'وضعیت "بی اهمیت"' است، به این مفهوم که صفر یا یک بودن مقدار پیکسل موجود در آن محل اهمیتی ندارد. نتایج مطروحه در گزارشات متعددی که درباره 'شکل شناسی نوشته شده اند، مبتنی بر کاربرد تنها یک عنصر ساختمانی، مشابه عنصر ساختمانی شکل ۸. ۴۰ (ب)، هستند، اما در تمام طول ستون اول وضعیت بی اهمیت دارند. این نتایج غلط هستند. زیرا به عنوان مثال این عنصر ساختمانی، نقطه قرارگرفته در سطر هشتم از ستون چهارم شکل ۸. ۴۰ (الف) را به عنوان یک نقطه پایان شناسایی می کند، بنابراین آن را حذف و اتصال خطوط را قطع می کند.

سه بار اعمال معادله '۸. ۴. ۲۷-۴' بر A ، مجموعه X نمایش داده شده در شکل ۸. ۴۰ (ت) را می دهد. مرحله 'بعدی بازیابی نویسه به شکل اولش، البته بدون شاخه های پارازیتی، می باشد. برای انجام این کار در ابتدا لازم است مجموعه X (شکل ۸. ۴۰ ت) که حاوی تمام نقاط پایانی X است، تشکیل شود:

$$X_2 = \bigcup_{k=1}^8 (X_1 \odot B^k) \quad (28-4. 8)$$

که B^k ها آشکارساز های نقاط پایان هستند که قبل از نیز به کار رفتند. مرحله 'بعد سه بار گسترش نقاط پایان با استفاده از مجموعه A به عنوان محدودساز^۱ می باشد:

$$X_3 = (X_2 \oplus H) \cap A \quad (29-4. 8)$$

که H یک عنصر ساختمانی 3×3 از "۱" هاست. همان طور که در شکل ۸. ۴۰ (ج) دیده می شود، این نوع از گسترش شرطی همانند پر کردن ناحیه و استخراج اجزای متصل، از ایجاد عناصر با مقدار "۱" خارج از ناحیه مورد علاقه جلوگیری می کند. در نهایت همانطور که در شکل ۸. ۴۰ (ج) دیده می شود، از اجتماع X_3 و X نتیجه 'نهایی بدست می آید:

$$X_4 = X_3 \cup X_1 \quad (30-4. 8)$$

در موارد پیچیده تر، گاهی استفاده از معادله '۸. ۴. ۲۹-۴' "نوك های"^۲ تعدادی از شاخه های پارازیتی رانگه می دارد. این وضعیت وقتی ممکن است رخ دهد که نقاط پایانی این شاخه ها نزدیک

۶۲۵ شکل شناسی ۴.۱

اسکلت باشند. حتی اگر معادله^۸ $4x - 27 = 0$ آنها را حذف کند، دوباره ممکن است در حین گسترش اضافه شوند زیرا جزء نقاط معتبر در A هستند. آشکارسازی و حذف عناصر پارازیتی آسان است زیرا آنها نواحی متصل هستند، مگر این که تمام عنصر پارازیتی دوباره اضافه شوند (که اگر این عناصر نسبت به خطوط معتبر کوتاه باشند، مورد نادری است).

در این مرحله طبیعی است که فکر کنیم باید راه های ساده تری نیز برای حل مسئله وجود داشته باشد. مثلاً می توانستیم همه 'نقاط پاک شده را پیدا کرده و فقط نقاط مناسب را به هریک از نقاط پایانی باقی مانده از کاربرد معادله^۸ $4x - 27 = 0$ متصل نماییم. این انتخاب درست است، اما مزیت روش ارائه شده این است که استفاده از ساختارهای ساده شکل شناسی کل مشکل را حل کرد. در وضعیت های عملی وقتی مجموعه ای از چنین ابزارهایی موجود باشد، این مزیت وجود دارد که دیگر لازم نیست هیچ الگوریتم جدیدی نوشته شود و تنها توابع شکل شناسی مورد نیاز را در یک رشته عملیات متوالی ترکیب می کنیم.

جدول ۴.۳ عملگرهای شکل شناسی ارائه شده در این بخش را گردآوری می کند. شکل ۴.۱ نیز انواع پایه ای عناصر ساختمانی به کار رفته را گردآوری می کند.

جدول ۴.۳ خلاصه ای از عملگرهای شکل شناسی و خواص آنها

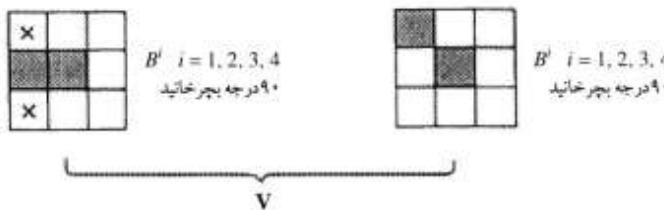
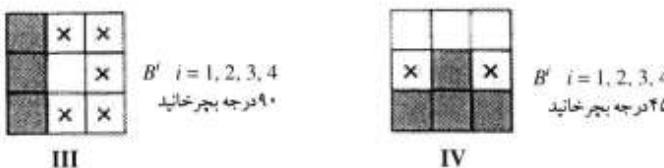
عملگر	معادله	توضیحات ^۹
انتقال	$(A)_x = \{c c = a + x, \text{ for } a \in A\}$	مبدأ A را به نقطه x انتقال می دهد.
قرینه	$\hat{B} = \{x x = -b, \text{ for } b \in B\}$	تمام عناصر B را حول مبدأ قرینه می کند.
متهم	$A^c = \{x x \notin A\}$	مجموعه 'نقاطی که در A نیستند.
تفاضل	$A - B = \{x x \in A, x \notin B\} = A \cap B^c$	مجموعه 'نقاطی که به A تعلق دارند اما به B تعلق ندارند.
گسترش	$A \oplus B = \{x (B)_x \cap A \neq \emptyset\}$	مرز A را "گسترش می دهد" (I)
سایش	$A \ominus B = \{x (B)_x \subseteq A\}$	مرزهای A را "می برد". (I)
باز کردن	$A \circ B = (A \ominus B) \oplus B$	منحنی هارا هموار می کند، اتصالات باریک را قطع می کند و تکه های کوچک و نوک های تیز را حذف می کند. (I)
بسن	$A \bullet B = (A \oplus B) \ominus B$	منحنی هارا هموار می کند، اتصالات باریک و

عملگر	معادله	توضیحات
تبديل HoM	$\begin{aligned} A \odot B &= (A \ominus B_1) \cap (A' \ominus B_2) \\ &= (A \ominus B_1) - (A \oplus B_2) \end{aligned}$	و تورفتگی های باریک بلند را پسوند می دهد، و سوراخ های کوچک را حذف می کند. (I)
استخراج مرز	$\beta(A) = A - (A \ominus B)$	مجموعه 'نقاط مرزی' A . (I)
پرکردن ناحیه	$X_i = (X_{i-1} \oplus B) \cap A'; \quad X_0 = p$ $k = 1, 2, 3, \dots$	با داشتن نقطه ای در یک ناحیه از A ، آن ناحیه را پر می کند. (II).
جزء متصل	$X_i = (X_{i-1} \oplus B) \cap A; \quad X_0 = p$ $k = 1, 2, 3, \dots$	با داشتن نقطه ای در جزء متصل Y از A ، آن جزء را پیدا می کند. (I).
بلند' محدب	$\begin{aligned} X'_i &= (X'_{i-1} \odot B') \cup A; \quad i = 1, 2, 3, 4, \\ k &= 1, 2, 3, \dots, X'_0 = A, \text{ and} \\ D' &= X'_{i_{\max}} \\ C(A) &= \bigcup_{i=1}^k D' \end{aligned}$	بدنه' محدب $C(A)$ ، A ، را می باید، که "نشانه همگرایی به مفهوم" $X'_k = X'_{k-1}$ می باشد. (III)
باریک سازی	$\begin{aligned} A \otimes B &= A - (A \odot B) \\ &= A \cap (A \odot B)' \end{aligned}$	مجموعه 'A را باریک می کند. دو معادله' اول دو تعریف پایه ای باریک سازی هستند. دو معادله' آخر باریک سازی با دنباله ای از عناصر ساختمانی را نشان می دهند. معمولاً در عمل از این روش استفاده می شود. (IV)
ضخیم سازی	$\begin{aligned} A \odot B &= A \cup (A \oplus B) \\ A \odot \{B\} &= ((\dots((A \odot B^1) \odot B^2) \dots) \\ &\quad \otimes B^n) \\ \{B\} &= \{B^1, B^2, B^3, \dots, B^n\} \end{aligned}$	مجموعه 'A را ضخیم می کند. (توضیحات قبلی درباره' عناصر ساختمانی را ببینید). از عنصر ساختمانی (IV) با "0" ها و "1" های معکوس استفاده می کند.
اسکلت	$\begin{aligned} S(A) &= \bigcup_{i=0}^K S_i(A) \\ S_i(A) &= \bigcup_{k=0}^K \{(A \ominus kB) \\ &\quad - [(A \ominus kB) \circ B]\} \\ A &= \bigcup_{i=0}^K (S_i(A) \oplus kB) \end{aligned}$	اسکلت A ، $S(A)$ ، را می باید. معادله' آخر نشان می دهد که می توان A را از زیر مجموعه های اسکلتی آن، $S_i(A)$ ها، بازسازی کرد. در هر سه معادله، K تعداد مراحل تکرار است که بعد از آنها مجموعه 'A به مجموعه 'تهی سایش می باید. نماد B نشانه 'k بار سایش متوالی با A است. (I)

۶۲۷ شکل شناسی ۴.۸

عملگر	معادله	توضیحات [†]
هرس کردن	$X_1 = A \otimes \{B\}$	X_4 حاصل هرس کردن مجموعه 'A' است . تعداد بارهایی که معادله 'B' به کار می رود تا X_1 بدست آید ، باید مشخص شود . در دو معادله 'اول عنصر ساختمانی (V) استفاده می شوند . معادله 'سوم از عنصر ساختمانی (I) استفاده می کند .
	$X_2 = \bigcup_{i=1}^n (X_1 \odot B^i)$	
	$X_3 = (X_2 \oplus H) \cap A$	
	$X_4 = X_1 \cup X_3$	

[†] اعداد روی درون پرانتزها به عنصر(های) ساختمانی مورد استفاده در فرآیند شکل شناسی اشاره دارند (شکل ۴۱.۸ را ببینید).



شکل ۴.۸ پنج نوع پایه ای عناصر ساختمانی که در این بخش مورد استفاده قرار گرفتند . مبدأ هر عنصر در مرکزش است و × هاشانه مقادیر "بی اهمیت" هستند .

مثال: این بخش را با یک مثال عملی که کاربرد شکل شناسی در مراحل اولیه پردازش در یک سامانه تشخیص نویسه را نمایش می دهد، پایان می دهیم. این سامانه قادر به خواندن شناسه های پستی^۱ دست نویس روی بسته های پستی آمریکا است.

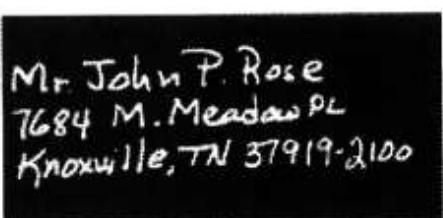
1) ZIP code

شکل ۸ . (الف) بخش آدرس یک بسته را بعد از آستانه گیری تصویرنمايش می دهد. یک مرحله' پردازش اولیه بعد از تعیین محل بخش آدرس ، استخراج تمام اجزای متصل در ناحیه' حاوی این بخش است (برای شرح مفصل درباره' اجزای متصل ، بخش ۲ . ۴ . و توضیحات قبلی در این بخش را ببینید). سپس هر جزء متصل با کوچکترین جعبه ای که کاملاً آن جزء را احاطه کند ، محصور می شود. اکنون جعبه ها و محتویات آن اساس بخشی از فرآیند هستند که به استخراج ناحیه' حاوی شناسه پستی می پردازد. شکل ۸ . (ب) نمایی نزدیک از ناحیه' حاوی شناسه پستی در این مثال خاص است. در این شکل سه ایراد دیده می شود: دورقم اول (7, 3) به هم متصل هستند، اولین "۱" در وسط قطعی دارد، و حلقه' رقم 2 نیز قطعی دارد.

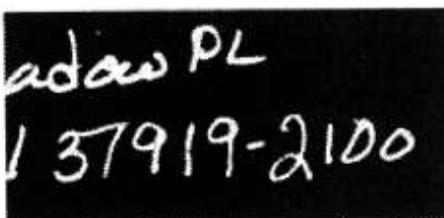
بعد از تعیین محل شناسه پستی ، به راه های گوناگونی می توان نویسه های پیوند خورده را آشکار کرد. برای نمونه اگر تحلیل جعبه های حول نویسه ها کمتر از پنج نویسه (یا کمتر از نه نویسه در مورد «شناسه پستی به اضافه' چهار رقم») را آشکار کند، جستجو برای نویسه های پیوند خورده با بررسی عرض نسبی جعبه هایی که نویسه هارا محصور می کنند، شروع می شود (در اینجا با نویسه ها به عنوان اجزای متصل برخورد می شود). یک جعبه با عرض بیش از معمول معمولاً متناظر با دو یا چند نویسه' پیوند خورده است. یک روش شکل شناسی برای جدا کردن نویسه های پیوند خورده سایش محتویات هر جعبه است تا این که نواحی "شبه نویسه ای" منفصل بدست آیند. برای نمونه مشکل تماس ۳ و ۷ با پنج بار تکرار عمل سایش در جعبه' حاوی این نویسه ها بر طرف شد. شکل ۸ . (ب) حاصل جداسازی نویسه هارانشان می دهد. توجه کنید که تنها در ناحیه ای که احتمال حضور نویسه های متصل وجود داشت، سایش اجرا شد.

ممولاً با عمل گسترش می توان مشکل نویسه های قطعی دار را حل کرد. در پیش پردازش ، به عنوان مثال ، وقتی که جعبه های محیط بر چند نویسه ، ابعادی کوچک تر از مقدار مورد انتظار دارند یا وقتی که دو یا چند جعبه چیدمان غیر معمولی دارند، مثلاً روی هم قرار گرفته اند، احتمالاً نویسه های قطعی دار حضور دارند. وضعیت اخیر در مورد ۱ موجود در شکل ۸ . (ب) صدق می کند که نشانگر قطعی دار بودن آن می باشد. با فرض این که دلیل چنین انفصالاتی نایکتواختی در عرض خط است، این انتظار ایجاد می شود که انفصال های دیگری نیز در شناسه پستی وجود داشته اند اما اجزای متصل جدا از هم (نظیر انفصال در ۲) را نتیجه نداده اند. بنابراین اجرای گسترش روی تمام نویسه ها منطقی است، البته با نظرارت بر این که

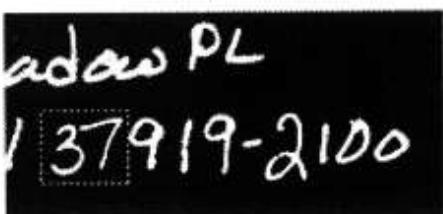
۶۲۹ شکل شناسی ۴.۸



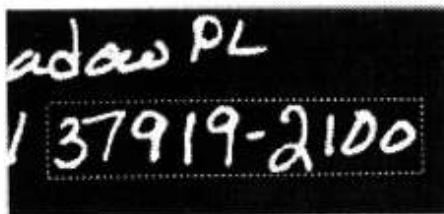
(الف)



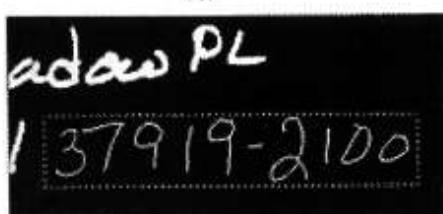
(ب)



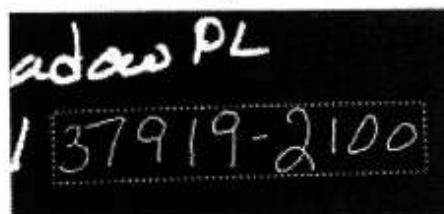
(ج)



(د)



(ه)



(ج)

شکل ۴۲.۸ (الف) میدان آدرس آستانه گیری شده؛ (ب) نمای نزدیک^۱ شناسه پستی که نویسه های متصل به هم یا منفصل را نشان می دهد؛ (پ) نتیجه بعد از پنج بار تکرار سایش در جعبه ای که نویسه های متصل به هم را محصور می کند (اکنون نویسه ها جدا از هم هستند)؛ (ت) نتیجه بعد از سه بار تکرار گسترش در کل شناسه پستی (انفصالات در چپ ترین "۱" و در "۲" پیوند خوردن)؛ (ث) اسکلت (ت) که شاخه های پارازیتی در گوشه ۷ و در یکی از "۰" ها را نشان می دهد؛ (ج) نتیجه بعد از هفت بار تکرار هرس کردن که حذف شاخه های پارازیتی را نشان می دهد. (با مجوز از *Perceptics corporation*)

نویسه های جدید- یا نویسه هایی که قبل از هم جدا شده اند- با اجرای گسترش به هم متصل نشوند. شکل ۴۲.۸ (ت) حاصل این فرآیندر اپس از سه بار اجرای گسترش روی تمام نویسه ها نشان می دهد. توجه کنید که ناپیوستگی در وسط ۱ و ناپیوستگی در حلقة ۲ هر دو پیوند خوردن.

یکی از روش های اصلی برای تشخیص ساختاری نویسه، تحلیل اسکلت هر نویسه است. شکل ۴۲.۸ (ث) که با الگوریتم ارائه شده در بخش ۸.۱.۵. به دست آمد، اسکلت نویسه های شناسه پستی

1) Close-up

شکل ۸.۴۲ (ت) را نشان می دهد. از توضیحات ارائه شده در مورد هرس کردن به ياد آوريد که يكى از مشکلات اسکلت ها ایجاد شاخه های پارازیتی در آنها است که اگر راه حل صحیحی برای آن پیدا نشود، منبع خطای بزرگی در تشخیص نویسه است. در این مثال يك شاخه 'کوچک در گوش' ۷ وجود دارد، و يك شاخه 'بزرگ نیز در بالای يكی از ۰ ها ایجاد شده است. همان طور که در شکل ۸.۴۲ (ج) دیده می شود، پس از هفت بار اجرای فرآیند هرس کردن هر دو شاخه 'پارازیتی حذف شدند. همان طور که قبل از شاد، انتخاب تعداد تکرارهای هرس کردن معمولاً به صورت مکاشفه ای انجام می گیرد. برای نمونه اگر شاخه 'روی ۰ در سمت راست پایین آن می بود، ممکن بود که نویسه Q باشد، و حذف کامل شاخه موجب خطا می شد. هیچ راه مطمئنی غیر از به کارگیری دانش زمینه ای^۱ برای حل این مشکل وجود ندارد. در این مثال می دانیم که نویسه باید عدد باشد زیرا در يك شناسه پستی آمده است. در وضعیت های پیچیده تر (مانند آدرس های خیابانی) استفاده از دانش زمینه ای شامل بررسی تناظر بین شناسه پستی و نام خیابان های معتبر برای آن شناسه می باشد.

۵.۴.۸ تعمیم به تصاویر خاکستری

در این بخش عملگرهای اساسی گسترش، سایش، باز کردن و بستن را به تصاویر خاکستری تعمیم می دهیم. آنگاه از این عملگرهای معرفی چند الگوریتم پایه ای شکل شناسی مناسب با تصاویر خاکستری استفاده می کنیم. در اینجا نیز همانند بخش ۸.۴.۴ تمرکز ما بر استفاده از شکل شناسی خاکستری برای استخراج اجزای تصویری مفید برای نمایش و توصیف شکل است. بویژه الگوریتم هایی برای استخراج مرز با گرادیان شکل شناسی، و نیز برای تقسیم بندی ناحیه براساس محتوای بافتی ارائه می کنیم. همچنین الگوریتم هایی را برای هموارسازی و ارتقا، که اغلب در مراحل پیش-یا پس پردازش مورد استفاده هستند، شرح می دهیم.

در تمام شرح های بعدی، با توابع تصویر رقمه به شکل $f(x, y)$ و $b(x, y)$ سروکار داریم که $f(x, y)$ تصویر ورودی و $b(x, y)$ یک عنصر ساختمانی (زیر تصویر) است. فرض براین است که این توابع در مفهوم معرفی شده در بخش ۳.۱.۲ گستته هستند. یعنی اگر Z نشانه 'مجموعه' اعداد صحیح حقیقی باشد، فرض براین است که (y, x) اعداد صحیحی از $Z \times Z$ هستند و f و b توابعی هستند که به هر زوج

1) Contextual knowledge

۴.۸ شکل شناسی ۶۵۱

مجزا از مختصات (y, x) یک مقدار سطح خاکستری (یک عدد حقیقی از مجموعه اعداد حقیقی R) متنسب می کنند. اگر سطوح خاکستری نیز اعداد صحیح باشند، Z جایگزین R می شود.

گسترش

گسترش تصویر خاکستری f با b ، که با $f \oplus b$ نشان داده می شود، بدین صورت تعریف می شود:

$$(f \oplus b)(s, t) = \max\{f(s-x, t-y) + b(x, y) | (s-x, t-y) \in D_f, (x, y) \in D_b\} \quad (31-4. 8)$$

که D_f و D_b به ترتیب حوزه های f و b هستند. طبق گذشته، b عنصر ساختمانی فرآیند شکل شناسی است، البته توجه کنید که اکنون b یک تابع است نه یک مجموعه.

شرط این که عامل های جابجاپی $(s-x)$ و $(t-y)$ درون حوزه f باشند، مشابه شرط موجود در تعریف دودویی گسترش است که در آنجا دو مجموعه باید لاقل در یک عنصر همپوشانی داشته باشند. همچنین توجه کنید که شکل معادله '۳۱-۴. ۸' مشابه کانولوشن دو بعدی (معادله '۳۵-۳') است، که در اینجا عمل بیشینه گیری، جایگزین جمع کانولوشن شده و عمل جمع نیز جایگزین ضرب شده است. با استفاده از توابع یک بعدی ساده، نمادها و طرز کار معادله '۳۱-۴. ۸' را نمایش می دهیم. برای توابع یک متغیره، معادله '۳۱-۴. ۸' به صورت زیر ساده می شود:

$$(f \oplus b)(s) = \max\{f(s-x) + b(x) | s-x \in D_f \text{ و } x \in D_b\}.$$

از شرح کانولوشن به یادآورید که $(-x)f$ با چرخش $(x)f$ حول مبدأ محور x به دست می آید. همانند کانولوشن، تابع $(s-x)f$ برای s های مثبت به راست و برای s های منفی به سمت چپ جایه جایی شود. این دو شرط که مقدار $(s-x)f$ باید در حوزه f و مقدار x باید در حوزه b باشد، به طور ضمنی بیان می کند که f و b همپوشانی دارند. این شرایط مشابه شرایط تعریف دودویی گسترش که دو مجموعه باید دست کم در یک عنصر همپوشانی داشته باشند، است. در نهایت برخلاف حالت دودویی، به جای عنصر ساختمانی b ، حرکت می کند. معادله '۳۱-۴. ۸' را می توان طوری نوشت که به جای f ، b حرکت کند. با این حال اگر D کوچک تر از D_f باشد (که این شرط در قریب به اتفاق موارد عملی برقرار است)،

شکل معادله' (۳۱-۴) اندیس گذاری ساده تری دارد و همان حاصل رامی دهد. از نظر مفهومی لغزش روی f در واقع تفاوتی بالغزش f روی b ندارد.

گسترش خاصیت جایه جایی دارد، بنابراین می توان با جایه جایی f و b ، و استفاده از معادله' (۳۱-۴) $f \ominus b$ را نیز محاسبه کرد. نتیجه، مشابه است، و اکنون تابع b حرکت داده می شود. اما سایش جایه جایی پذیر نیست و نمی توان از جایه جایی تابع بهره گرفت. با الزام این که در هر دو عمل گسترش و سایش b حرکت نماید، عبارتی متفاوت (و یچیده تر) نتیجه می شود. بنابراین به خاطر سادگی (و همسانی با سایر نوشته ها) معادله' (۳۱-۴) را برای بیان گسترش انتخاب کردیم، و در خاطر داریم که هر دو روش مفهوم یکسانی دارند. در شکل ۴۳.۸ یک مثال دیده می شود.

چون گسترش مبتنی بر انتخاب مقدار بیشینه' $b + f$ در یک همسایگی با شکل عنصر ساختمانی است، اثر کلی گسترش بر تصویر خاکستری دو وجه مختلف دارد: (۱) اگر تمام مقادیر عنصر ساختمانی مثبت باشند، تصویر خروجی روشن تر از تصویر ورودی خواهد بود؛ و (۲) جزئیات تیره، بسته به چگونگی ارتباط مقادیر و شکلشان با عنصر ساختمانی بکار رفته در گسترش، کاهش می یابند یا حذف می شوند.

سایش

سایش تصویر خاکستری، که با $b \ominus f$ نشان داده می شود، به صورت زیر تعریف می شود:

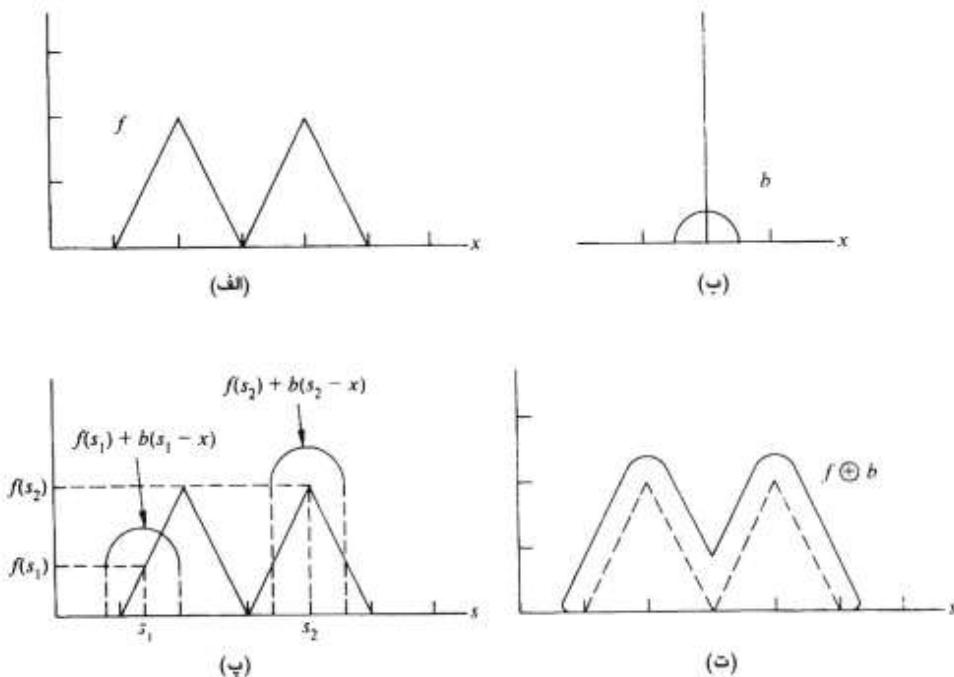
$$(f \ominus b)(s, t) = \min\{f(s + x, t + y) - b(x, y) | (s + x, t + y) \in D_f; (x, y) \in D_b\} \quad (32-4, 8)$$

که D_f و D_b به ترتیب حوزه های f و b هستند. شرط این که عوامل جایه جایی $(s + x)$ و $(t + y)$ درون حوزه' f باشند، مشابه شرط موجود در تعریف دودویی سایش است، طوری که عنصر ساختمانی همواره به طور کامل درون مجموعه' تحت سایش قرار گیرد. توجه کنید که شکل معادله' (۳۲-۴) ثبیه محاسبه همبستگی دو بعدی (معادله' ۳.۳۹-۳) است با این تفاوت که در اینجا عمل کمینه گیری جایگزین جمع های موجود در محاسبه همبستگی و عمل تفریق نیز جایگزین ضرب های آن شده است.

با سایش یک تابع بعدی ساده طرز عمل معادله' (۳۲-۴) را نشان می دهیم. برای توابع یک متغیره معادله' سایش به صورت زیر ساده می شود:

$$\text{معادله' سایش} = \min\{f(s + x) - b(x) | x \in D_b\}$$

۴۳.۸ شکل شناسی ۶۵۳

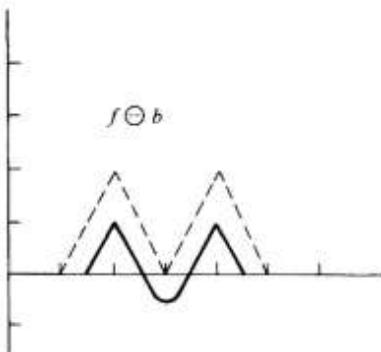


شکل ۴۳.۸ نتیجه گسترش f بالغزاندن b روی آن. از نظر ریاضی بالغزاندن f روی b ، اندیس‌گذاری ساده‌تر می‌شود. در هر حال نتیجه یکسان است. (گرفته شده از [Giardina & Dougherty 1988].)

$$(f \ominus b)(s) = \min\{f(s+x) - b(x) | (s+x) \in D_f \text{ و } x \in D_b\}.$$

مانند فرآیند همبستگی، تابع $f(s+x)$ برای s های مثبت به چپ و برای s های منفی به راست جا به جا می‌شود. دو شرط $s \in D_f$ و $x \in D_b$ الزام می‌کند که حوزه b کاملاً درون حوزه تابع f جابجا شده قرار گیرد. همان طور که در بالا ذکر شد، این شرایط مشابه شرایط تعریف دودویی سایش هستند، که در آنجا نیز عنصر ساختمنی می‌بایست کاملاً درون مجموعه 'تحت سایش' قرار گیرد.

در نهایت برخلاف تعریف دودویی سایش، f به جای عنصر ساختمنی b انتقال می‌باید. می‌توان معادله '۴.۴-۳۲' را طوری نوشت که تابع b انتقال باید، اما این روش عبارتی با اندیس‌گذاری پیچیده تر را نتیجه می‌دهد. چون لغزاندن f روی b مفهومی مانند لغزاندن b روی f دارد، به همان دلایلی که در پایان شرح گسترش ذکر شد، از معادله '۴.۴-۳۲' استفاده می‌شود. شکل ۴۴.۸ حاصل سایش تابع شکل ۴۳.۸ (الف) با عنصر ساختمنی شکل ۴۳.۸ (ب) را نشان می‌دهد.



شکل ۸.۴۳ سایش تابع شکل ۸.۴۲(الف) با عنصر ساختمانی شکل ۸.۴۳(ب). (گرفته شده از Giardina & Dougherty[1988]).

همان طور که معادله ۸.۴-۳۲ نشان می دهد، سایش براساس انتخاب مقدار کمینه ' $f - b$ ' در یک همسایگی به شکل عنصر ساختمانی، است. اجرای سایش دو اثر بر تصویر خاکستری می گذارد:

- (۱) اگر تمام عناصر عنصر ساختمانی مثبت باشد، تصویر خروجی تیره تر از تصویر ورودی می شود؛ و
- (۲) اثر جزئیات روشن تصویر ورودی که "سطحی" کوچک تر از عنصر ساختمانی دارند، کاهش می یابد، که درجه 'کاهش' به مقادیر خاکستری حول جزئیات روشن و به شکل و دامنه' مقادیر عنصر ساختمانی بستگی دارد.

طبق گذشته، گسترش و سایش نسبت به دو عمل متمم گیری و انعکاس، دو گان هم هستند. یعنی،

$$(f \ominus b)^c(x, y) = (f^c \oplus \hat{b})(x, y) \quad (8.4-33)$$

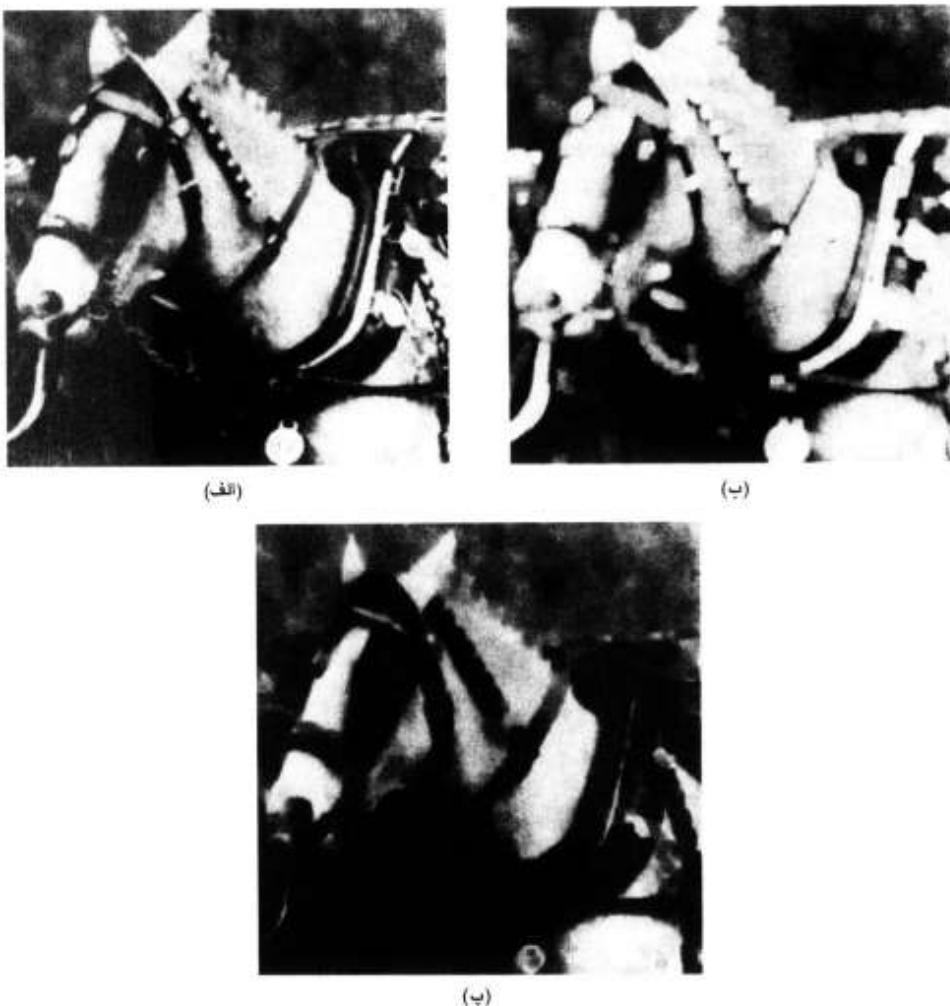
که $(f \ominus b)^c = b(-x, -y)$ در ادامه بحث بجز در مواردی که برای روشن شدن مطلب مورد نیاز است، با حذف آرگومان های توابع طرز بیان روابط را ساده می کنیم.

مثال: شکل ۸.۴۵(الف) یک تصویر خاکستری 512×512 ساده را نشان می دهد، و شکل ۸.۴۵(ب) حاصل گسترش این تصویر را با یک عنصر ساختمانی "سرتخت"^۱ که به شکل سطوح موازی^۲ با ارتفاع واحد و ابعاد 5×5 پیکسل است، نشان می دهد. براساس شرح قبلی انتظار می رود که تصویر حاصل از گسترش، روشن تر از تصویر اولیه باشد و جزئیات تیره آن کاهش یابند یا حذف شوند. این اثرات به وضوح در شکل ۸.۴۵(ب) نمایان است. نه تنها این تصویر روشن تر از تصویر اولیه است، بلکه ابعاد

1) Flat-top

2) Parallel piped

۴۰.۸ شکل شناسی ۶۵۵



شکل ۴۰.۸ (الف) تصویر اولیه؛ (ب) نتیجه گسترش؛ (پ) نتیجه سایش. (با مجوز از A. Morris, Leica Cambridge, Ltd)

ویژگی های تیره آن، نظیر سوراخ های بینی و اجزای تیره افسار اسب، نیز کاهش یافته اند.
شکل ۴۰.۸(پ) حاصل سایش تصویر اولیه را نشان می دهد. به اثر مخالف سایش نسبت به گسترش توجه کنید. تصویر سایش یافته تیره تر است، و ابعاد بخش های کوچک روشن آن (نظیر پرج های افسار^۱) کاهش یافته اند.

□

1) Studs on the rein

بازگردن و بستن

عبارات مربوط به بازگردن و بستن تصاویر خاکستری، همشکل گونه های دودویی شان هستند. بازگردن تصویر f با تصویر (عنصر ساختمانی) b ، که با $b \circ f$ نشان داده می شود، عبارتست از:

$$f \circ b = (f \ominus b) \oplus b. \quad (34-4. \Delta)$$

همانند حالت دودویی، بازگردن مشکل از سایش f با b و سپس گسترش حاصل با b است. به طور مشابه بستن f با b ، که با $b \bullet f$ نشان داده می شود، عبارتست از

$$f \bullet b = (f \oplus b) \ominus b. \quad (35-4. \Delta)$$

بازگردن و بستن تصویر خاکستری نسبت به دو عمل متمم گیری و انعکاس، دو گان هستند. یعنی،

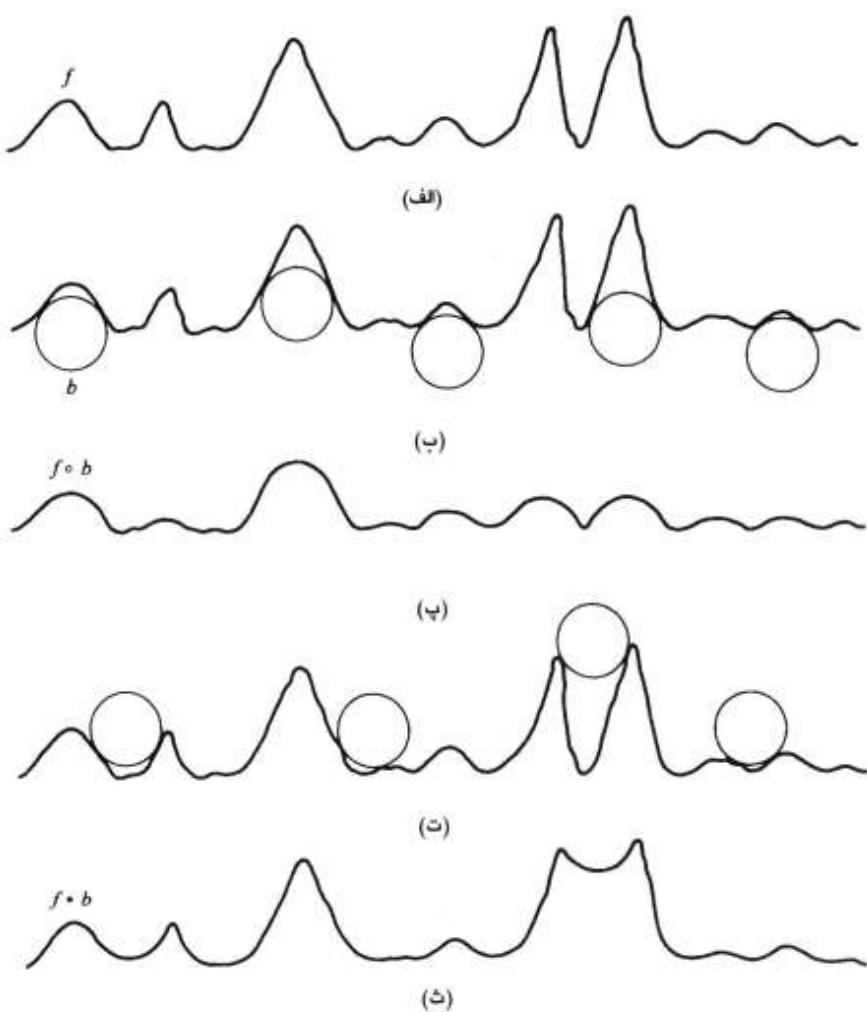
$$(f \bullet b)^c = f^c \circ \hat{b} \quad (36-4. \Delta)$$

چون $(y, f^c) = -f(x, y)$ ، می توان معادله $(36-4. \Delta)$ را به صورت $(\hat{b} \circ b) = (-f \circ f)$ نیز نوشت.

بازگردن و بستن تصاویر، تعبیر هندسی ساده ای دارد. فرض کنید که تابع تصویر $f(x, y)$ را در پرسپکتیو^۱ سه بعدی (نظیر نقشه ناهمواری ها) مشاهده کنیم، که در آن محور های x و y ، مختصات مکانی و محور سوم، روشنایی (یعنی مقادیر f) است. در این نوع نمایش، تصویر به صورت سطح گسته ای ظاهر می شود که مقدار آن در هر نقطه (x, y) همان مقدار f در آن نقطه است. حال فرض کنید که می خواهیم حاصل بازگردن f با یک عنصر ساختمانی کروی را به دست آوریم، و عنصر ساختمانی را به صورت یک "توب چرخان" در نظر می گیریم. آنگاه طرز اجرای بازگردن f با \hat{f} را می توان از نظر هندسی به صورت فرآیند هل دادن توب به سطح زیرین f همگام با غلتاندن آن طوری که تمام سطح زیرین f پیمایش شود، تعبیر کرد. آنگاه حاصل بازگردن $b \circ f$ ، سطحی است از بلندترین نقاطی که قسمتی از سطح کره هنگام غلتیدن بر سطح زیرین f به آنها می رسد. شکل ۴۶. ۸ این مفهوم را نمایش می دهد. شکل ۴۶. ۸ (الف) برای سادگی یک خط پیمایش از یک تصویر خاکستری را به صورت تابعی پیوسته

۱) Perspective

۴۵۷ شکل شناسی ۴.۸



شکل ۴.۸ (الف) یک خط پیمایش از تصویر خاکستری؛ (ب) مکانهای مختلف توب چرخان در حین باز کردن (پ) نتیجه عمل باز کردن (ت) مکانهای مختلف توب چرخان در حین بستن؛ (ث) نتیجه عمل بستن.

نشان می دهد. شکل ۴.۸ (ب) توب چرخان را در مکان های مختلف نشان می دهد، و شکل ۴.۸ (پ) حاصل باز کردن f با b در طول خط پیمایش را نشان می دهد. دامنه و تیزی^۱ تمام قله های باریکتر از قطر توب، کاهش یافت. در کاربردهای عملی معمولاً از باز کردن برای حذف جزئیات روش کوچک (نسبت به ابعاد عنصر ساختمانی) استفاده می شود، در حالی که کل سطوح خاکستری و ویژگی های روش

1) Sharpness

بزرگتر نسبتاً بدون تغییر باقی میماند. سایش تصویر اولیه، جزئیات کوچک را حذف و تصویر را تیره تر میکند. گسترش بعد از سایش مجدد روشی تصویر را افزایش میدهد بدون این که جزئیات حذف شده با سایش را بازیابی کند.

شکل های ۴۶. (ت) و (ث) بستن f با b را نمایش میدهند. در اینجا توب روی سطح بالایی f میلغزد و قله ها به همان شکل اولیه شان باقی میمانند (البته تا زمانی که فاصله 'کمبنه' دو قله بیش از قطر توب باشد). در کاربردهای عملی معمولاً از بستن برای حذف جزئیات تیره تصویر استفاده میشود، در حالی که ویژگی های روش نسبتاً بدون تغییر باقی میمانند. گسترش تصویر اولیه، جزئیات تیره را حذف و تصویر را روش نتر میکند و سایش بعد از گسترش تصویر را تیره تر مینماید بدون این که جزئیات حذف شده با گسترش را بازیابی کند.

عمل بازکردن خواص زیر را دارد:

$$(f \circ b) \leftarrow f \quad (1)$$

$$(2) \text{ اگر } f_2 \leftarrow f_1 \text{ آن گاه } (f_2 \circ b) \leftarrow (f_1 \circ b)$$

$$\cdot (f \circ b) \circ b = f \circ b \quad (3)$$

عبارت $\leftarrow u$ بیانگر این است که حوزه u زیرمجموعه ای از حوزه v است و نیز برای هر (x, y) از حوزه v است $u(x, y) \leq v(x, y)$.

به طور مشابه عمل بستن خواص زیر را دارد:

$$f \leftarrow (f \bullet b) \quad (1)$$

$$(2) \text{ اگر } f_2 \leftarrow f_1, \text{ آن گاه } (f_2 \bullet b) \leftarrow (f_1 \bullet b)$$

$$\cdot (f \bullet b) \bullet b = f \bullet b \quad (3)$$

سودمندی این عبارات مشابه گونه های دودویی شان است. همانند حالت دودویی در هر دو عمل بازکردن و بستن، غالباً خواص (۲) و (۳) به ترتیب یکنواخت صعودی^۱ و محدودیت تأثیر^۲ خوانده میشوند. □

مثال: شکل ۴۷. (الف) حاصل باز کردن شکل ۴۵. (الف) با عنصر ساختمانی مورد استفاده در همان شکل را نشان میدهد. توجه کنید که ابعاد جزئیات کوچک روش، کوچک تر شد، اما بر سطوح

1) Increasing monotonicity

2) Idempotence

۶۵۹ ۴. آنالیز شکل



(٦)



(一)

شکل ۸ نتیجه (الف) باز کردن و (ب) بستن شکل ۸. ۴۵ (الف). (با مجوز از A. Morris, Leica Ltd, Cambridge, Ltd)

خاکستری تیره تر هیچ اثر قابل توجهی گذاشته نشد. شکل ۸ (ب) حاصل بستن شکل ۸ (الف) را نشان می دهد. به کاهش ابعاد در قسمتهای تیره' کوچک و اثر نسبتاً کم آن بر بخش های روشن توجه کنند. □

چند گاربرد از شکل شناسی خاگستری

مبحث شکل‌شناسی را با بیان کاربردهای گوناگون مفاهیم ارائه شده پایان می‌دهیم. تمام تصاویر به ابعاد 512×512 هستند و با استفاده از عنصر ساختمانی به کار رفته در شکل ۸.۴۵ پردازش می‌شوند، مگر این که خلاف آن ذکر شود.

آرام‌سازی^۱ با شکل‌شناسی. یک راه برای آرام‌سازی تصویر، بازکردن و سپس بستن آن است. بر اثر اجرای این دو عمل متوالی هر دو نقایص تیره و روشن و نیز نویز حذف یا تضعیف می‌شوند. شکل ۴۸. ۸ یک نسخه آرام شده از تصویر اوله موجود در شکل ۸. ۴۵ (الف) را نشان می‌دهد.

گردایان شکل شناسی. علاوه بر عملگرهایی که قبلًاً در مورد حذف نتایج کوچک تیره و روش تشریح شدند، گسترش و سایش غالباً برای محاسبه گردایان شکل شناسی^۲ تصویر، که با پنهان داده می‌شود،



شکل ۴۸.۱ نتیجه آرام سازی تصویر شکل ۴۵.۸ (الف) با شکل شناسی. (با مجوز از A. Morris, Leica, Ltd. (Cambridge, Ltd.

نیز به کار می روند:

$$g = (f \oplus b) - (f \ominus b). \quad (37-4. ۸)$$

شکل ۴۹.۸ حاصل محاسبه گرادیان شکل شناسی تصویر شکل ۴۵.۸ (الف) را نشان می دهد. همانطور که انتظار می رفت، گرادیان شکل شناسی، تغییرات تند سطح خاکستری تصویر ورودی را



شکل ۴۹.۸ نتیجه گرادیان شکل شناسی تصویر شکل ۴۵.۸ (الف). (با مجوز از A. Morris, Leica, Ltd. (Cambridge, Ltd.

۶۶۱ شکل‌شناسی ۴.۸



شکل ۴.۸۰ حاصل اجرای تبدیل TH بر تصویر شکل ۴.۲۵ (الف). (با مجوز از A. Morris, Leica (Cambridge, Ltd

مشخص می‌کند. برخلاف گرادیان‌های حاصل از روش‌هایی نظربر عملگرسوبل، گرادیان‌های شکل‌شناسی حاصل از کاربرد عناصر ساختمانی متقاضن، کمتر به جهت لبه بستگی دارند. هریته^۱ این مزیت افزایش قابل توجه محاسبات است.

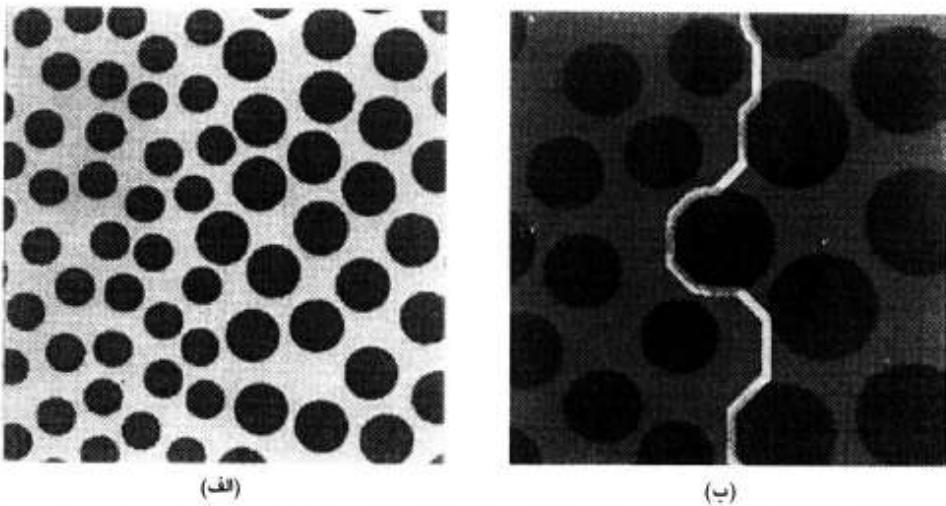
تبدیل TH^1 . تبدیل شکل‌شناسی TH یک تصویر، که با^۲ نشان داده می‌شود، به صورت زیر تعریف می‌شود:

$$h = f - (f \circ b) \quad (38-4.8)$$

که طبق گذشته f تصویر ورودی و b تابع عنصر ساختمانی است. این تبدیل -که در ابتدا به علت استفاده از عنصر ساختمانی استوانه‌ای بالوله- موازی^۳ سرتخت این نام را گرفت- برای ارتقای جزئیات تصویر در حضور سایه مفید است. شکل ۴.۸۰ حاصل اجرای تبدیل TH بر تصویر شکل ۴.۲۵ (الف) را نشان می‌دهد. به ارتقای جزئیات در ناحیه 'زمینه' زیر سر اسب توجه کنید.

بخش بندی بافتی. شکل ۴.۸۱ (الف) یک تصویر خاکستری ساده متشکل از دو ناحیه بافتی را نشان می‌دهد. ناحیه^۴ سمت راست از قطره‌های دایروی با قطر بزرگ‌تر از قطرات سمت چپ تشکیل می‌شود. هدف پیدا کردن مرز بین دو ناحیه براساس محتوای بافتی آنها است. گرچه ممکن است در ابتدا براساس

1) top-hat transformation 2) Parallel piped

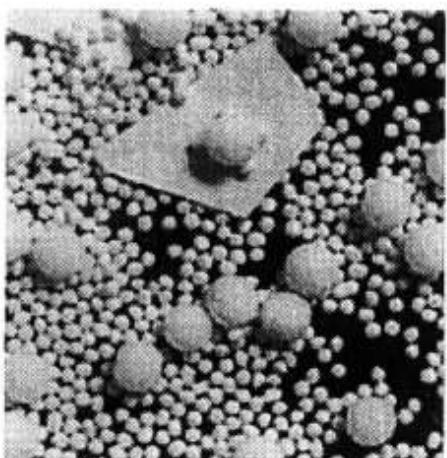


شکل ۸.۵ (الف) تصویر اولیه؛ (ب) مرز حاصل از بخش بندی. (با مجوز از A. Morris, Leica (Cambridge, Ltd

مطالب فصل ۷ راه حلهای گوناگونی به ذهن راه یابد، اما یک راه حل شکل شناسی که از عملگرهای بستن و باز کردن بهره گیرد، فواید ویژه‌ای دارد.

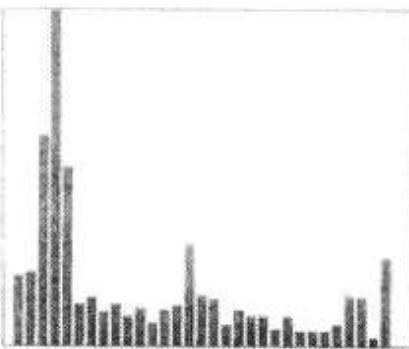
چون عمل بستن جزئیات تیره تصویر را کاهش می‌دهد، روال کار در این مورد خاص، بستن تصویر ورودی با دنباله‌ای از عناصر ساختمانی به طور متواالی بزرگتر است. وقتی ابعاد عنصر ساختمانی در حدود ابعاد قطره‌های کوچک بشود، این قطره‌ها حذف می‌شوند و در ناحیه سمت چپ شکل ۸.۵ (الف) تنها یک زمینه روشن باقی می‌ماند. در این مرحله از فرآیند، تنها قطره‌های بزرگ در سمت راست و زمینه روشن در سمت چپ و بین قطره‌های بزرگ باقی می‌ماند. سپس عمل باز کردن تصویر با یک عنصر ساختمانی که ابعادش نسبت به فاصله بین قطره‌های بزرگ باقیمانده بزرگ باشد، یکبار اجرا می‌گردد. این عمل تکه‌های روشن بین قطره‌هارا حذف می‌کند، و در سمت راست تصویر ناحیه‌ای تیره با قطره‌های تیره بزرگ و تکه‌هایی با تیرگی یکسان بین این قطره‌ها باقی می‌گذارد. در این مرحله از فرآیند یک ناحیه روشن در سمت چپ و یک ناحیه تیره در سمت راست ایجاد شده است. آنگاه یک آشکارساز آستانه‌ای ساده مرز بین دو ناحیه بافتی را مشخص می‌کند. شکل ۸.۵(ب) مرز حاصل را که روی تصویر اولیه قرار گرفته است، نشان می‌دهد. توصیه می‌شود که با استفاده از روش توب چرخان که در شکل ۸.۴۶ توصیف شد، این مثال با تفصیل بیشتری بررسی شود.

۴.۸ شکل شناسی ۶۶۳



(الف)

Size Dist'n



(ب)

شکل ۵۲.۸ (الف) تصویر اولیه ای که دارای قطعه های همپوشاننده است؛ (ب) توزیع اندازه. (با مجوز از Morris, Leica Cambridge, Ltd.

دانه سنجی^۱. دانه سنجی مبحثی است که فقط به تعیین توزیع اندازه، قطعه های تصویر می پردازد. شکل ۵۲.۸ (الف) تصویری مشکل از اشیای روشن با سه اندازه متفاوت را نشان می دهد. این اشیانه تنها همپوشانی دارند، بلکه آنقدر نامرتب هستند که نمی توان قطعه های منفرد را آشکار کرد. چون این قطعه ها نسبت به زمینه روشن تر هستند، می توان از راه شکل شناسی زیر برای تعیین توزیع ابعاد آنها استفاده کرد. ابتدا حاصل های باز کردن تصویر اولیه با چند عنصر اولیه^۲ که به طور متوالی ابعاد بزرگتری دارند، به دست می آید و پس از هر عمل باز کردن تفاصل بین تصویر اولیه و تصویر باز شده محاسبه می شود. در پایان فرآیند، این تفاصل ها تراز می شوند^۳ و با استفاده از آنها بافت نگار مربوط به توزیع اندازه قطعه ها تشکیل می شود. این روش مبتنی بر این نظریه است که عملگر باز کردن با یک اندازه^۴ خاص، بیشترین اثر را بر آن دسته از نواحی تصویر ورودی می گذارد که قطعاتی با اندازه مشابه دارند. بنابراین با محاسبه^۵ تفاصل بین دو تصویر ورودی و خروجی معیاری از تعداد نسبی چنین قطعاتی بدست می آید. شکل ۵۲.۸ (ب) توزیع اندازه^۶ بدست آمده در این مورد را نشان می دهد. بافت نگار حاصل، نشان دهنده وجود قطعاتی با سه اندازه^۷ متفاوت در تصویر ورودی می باشد. این نوع پردازش برای توصیف آن دسته از نواحی که بخش غالب آنها را نویسه تشکیل داده است، مفید است.

¹⁾ Granulometry²⁾ Normalized

۵.۸ توصیف گرهای رابطه ای

روش های مشروح در بخش های ۸-۲، ۸-۴ معمولاً به هر یک از مرزها و نواحی مورد علاقه در تصویر به صورت جداگانه، اعمال می شوند. سطح بعدی پیچیدگی در فرآیند توصیف، سازماندهی این اجزا است تا از تمام روابط ساختاری که ممکن است بین آنها وجود داشته باشد، استفاده نمایم.

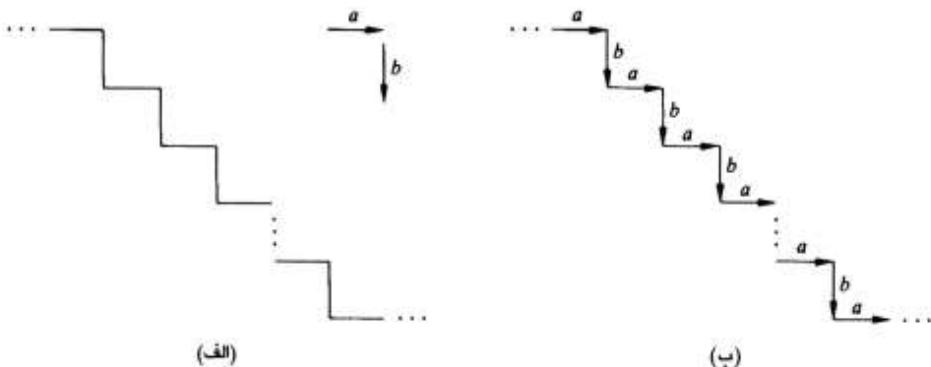
این مفهوم را با ساختار پله ای ساده موجود در شکل ۸. ۵۳ (الف) معرفی می کنیم. فرض کنید این ساختار بخشی از یک تصویر است و ما می خواهیم آن را با روشی رسمی توصیف نماییم. با تعریف دو عنصر اولیه a و b ، می توان شکل ۸. ۵۳ (الف) را به صورت شکل ۸. ۵۳ (ب) رمز نمود. آشکارترین خاصیت ساختار رمز شده، تکراری بودن عناصر a و b است. بنابراین یک روش ساده توصیف این است که با استفاده از این عناصر اولیه رابطه ای بازگشتی بنویسیم. یک امکان برای تحقق این هدف استفاده از قواعد بازنویسی^۱

$$, S \rightarrow aA \quad (1)$$

$$, A \rightarrow bS \quad (2)$$

$$, A \rightarrow b \quad (3)$$

است که S و A متغیر و a و b ثابت های متناظر با عناصر اولیه از پیش تعریف شده هستند. قاعده (۱) نشان می دهد که می توان S ، که نماد آغاز^۲ خوانده می شود، را با عنصر اولیه a و متغیر A جایگزین کرد.

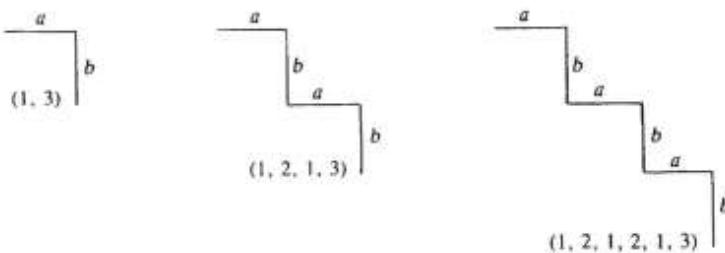


شکل ۸. ۵۳ (الف) یک ساختار پلکانی ساده؛ (ب) ساختار رمز شده

1) Rewriting rule

2) starting symbol

۶۶۵ ۵.۸ توصیف گرهای رابطه ای

شکل ۵.۸ نمونه دستاوردهای قواعد $A \rightarrow bS$, $S \rightarrow aA$ و $bA \rightarrow b$.

خود متغیر A را نیز می‌توان با b و S یا با a تنها جایگزین کرد. با جایگزینی A با bS دوباره به قاعده 'اول بازگشته و می‌توان این روال را تکرار کرد. جایگزینی A با a به روال پایان می‌دهد، زیرا هیچ متغیری در عبارت باقی نمی‌ماند. شکل ۵.۸. چندنمونه از کاربرد این قواعد را نشان می‌دهد، که اعداد زیر ساختارها، ترتیب کاربرد قواعد ۱، ۲، و ۳ را نشان می‌دهند. ارتباط بین a و b حفظ می‌شود، زیرا این قواعد طوری تنظیم شده‌اند که همواره بعد از a ، b باشد. قابل ذکر است که با استفاده از این سه قاعده 'ساده' بازنویسی، می‌توان ساختارهای "مشابه" بسیار زیادی را تولید (یا توصیف) کرد. همان‌طور که در فصل ۹ نشان می‌دهیم، مزیت دیگر این روش داشتن مبنای نظری محکم است.

چون رشته‌ها ساختارهای یک بعدی هستند، کاربرد آنها در توصیف تصویر، نیازمند تعیین روشنی مناسب برای ساده‌سازی روابط مکانی دو بعدی به شکل یک بعدی است. در بیشتر مواقع کاربرد رشته‌ها در توصیف تصویر، مبتنی بر ایده 'استخراج پاره خط‌های متصل از اشیای مورد علاقه در تصویر' است. یک روش برای تحقق این هدف، دنبال کردن منحنی شیء و رمزکردن نتیجه کار با قطعه‌های با جهت و/یا طول مشخص است. شکل ۵.۸. ۵۵ این روال را نمایش می‌دهد.

یک روش تقریباً کلی تر دیگر، این است که بخش‌های تصویر (نظیر نواحی همگن کوچک) را با پاره خط‌های جهت دار توصیف کنیم. پاره خط‌های جهت دار را می‌توان به راه‌های دیگری علاوه بر اتصالات ابتداء-به-انتها^۱ نیز به هم متصل کرد. شکل ۵.۸. ۵۶ (الف) این روش و شکل ۵.۸. ۵۶ (ب) چند عمل نوعی را که می‌توان روی عناصر اولیه تعریف کرد، نمایش می‌دهد. شکل ۵.۸. ۵۶ (پ) مجموعه‌ای از عناصر اولیه معین را نشان می‌دهد که از پاره خط‌های تعریف شده در چهار جهت تشکیل شده‌اند، و

1) Head- to- tail connection



شکل ۸.۵۵ رمزگاری مرز ناحیه با پاره خط های جهت دار

شکل ۸.۵۶ (ت) تولید مرحله به مرحله' یک شکل معین را نشان می دهد. در این شکل نماد (d) نشان دهنده 'عنصر اولیه'، با جهت معکوس است. توجه کنید که هر ساختار ترکیبی تنها یک ابتدا و یک انتها دارد. تیجه' مورد علاقه آخرین رشته از عناصر اولیه است، که ساختار کامل را توصیف می کند. این نوع رشته ها برای کاربردهای مناسب هستند که در آنها می توان اتصال عناصر اولیه را به روش ابتدا به انتها یا یک روش پیوسته' دیگر بیان کرد. در فصل ۹ لازم می شود که بتوانیم با ساختارهای منفصل نیز کار کنیم، و یکی از مفیدترین روش ها برای این کار استفاده از توصیف گرهات درختی است.

درخت T مجموعه' محدودی از یک یا چند گره است که در آن

(۱) یک گره یکتایی موسوم به ریشه' وجود دارد، و

(۲) سایر گره ها به m مجموعه' جدا از هم T_1, T_2, \dots, T_m تقسیم می شوند، که هر یک از این مجموعه ها خود یک درخت است و زیر درخت T خوانده می شود.

سرشاخه درخت T مجموعه' گره های انتهایی درخت (برگها^۳) است. این گره ها به ترتیب از چپ به راست درون این مجموعه قرار می گیرند. مثلاً درخت زیر، گره ریشه' s و سرشاخه xy را دارد.



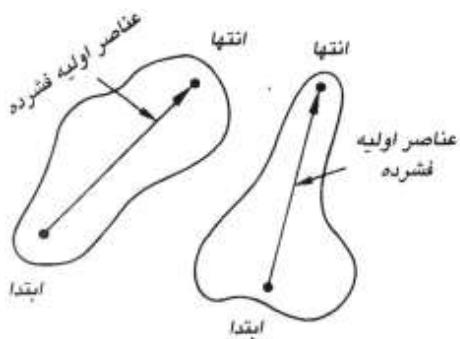
1) Root

2) Subtree

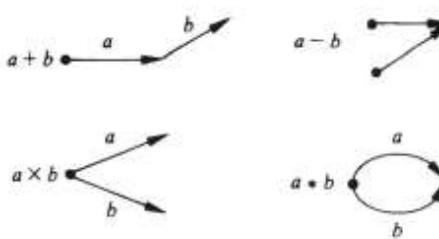
3) Tree frontier

4) Leave

۵.۸ توصیف گرهای رابطه ای ۶۶۷



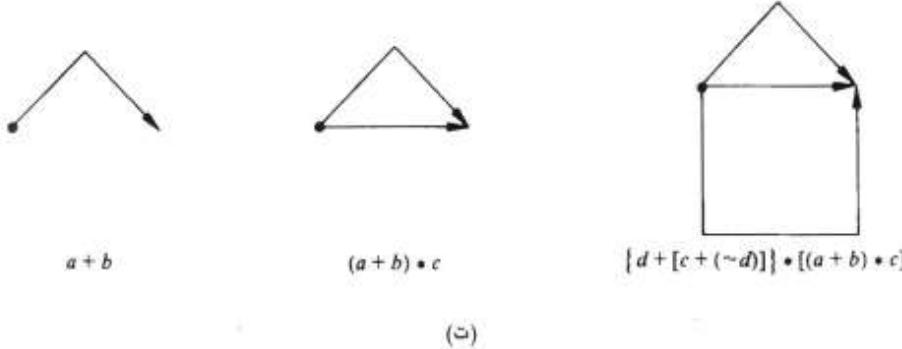
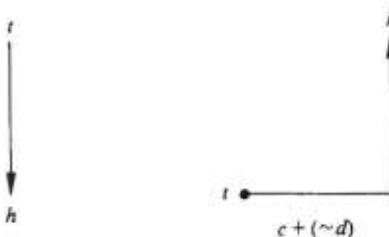
(الف)



(ب)

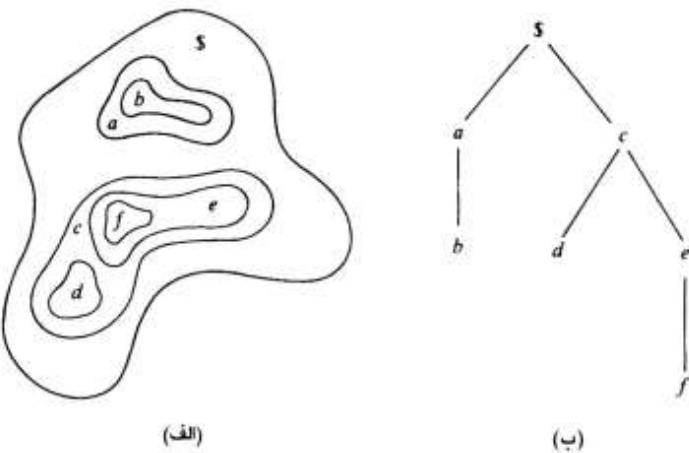


(پ)



(ت)

شکل ۵.۸ (الف) عناصر اولیه فشرده؛ (ب) عملیات بین عناصر اولیه؛ (پ) مجموعه ای از عناصر اولیه خاص؛ (ت) مرحله ایجاد یک ساختار.



شکل ۸.۵۷ (الف) پک ناحیه ترکیبی ماده؛ (ب) نمایش درختی حاصل از کاربرد رابطه "درون"^۱

معمولًا دو نوع اطلاعات از درخت مفید هستند: (۱) اطلاعات درباره 'گره که به صورت مجموعه ای از واژه ها که آن گره را توصیف می کنند، ذخیره شده است، و (۲) اطلاعاتی که یک گره را با همسایگانش مرتبط می کند که به صورت مجموعه ای از اشاره گرها به آن همسایه ها ذخیره شده است. همان طور که در توصیف تصویر نیز به کار رفت، نوع اول اطلاعات، عنصر اولیه الگو^۱ را معرفی می کند، در حالی که نوع دوم اطلاعات، ارتباط فیزیکی عنصر اولیه با سایر زیرساختهای را تعریف می کند. برای مثال شکل ۵۷. ۸ ((الف) را می توان با استفاده از رابطه "درون" یا یک درخت نمایش داد. بنابراین اگر ریشه 'درخت بانماد \$ نشان داده شود، طبق شکل ۸. ۵۷ (ب) سطح اول ترکیب شامل a و c درون \$ بوده و طبق شکل ۸. ۵۷ (ب) این سطح از ترکیب، دوشاخه 'وجود آمده از ریشه را تولید می کند. سطح بعدی شامل درون a ، و d درون c می باشد. در نهایت با اضافه کردن شاخه 'f درون c درخت کامل می شود.

۶۰۸ نکات پایانی

نمایش و توصیف اشیا یا نواحی که تصویر بخش بندی شده، جزء مراحل ابتدایی بسیاری از سامانه‌های تحلیل خودکار تصویر هستند. مفاهیم شکل‌شناسی معرفی شده، راهی قوی برای استخراج عناصر این

1) Inside of

مراجع ۶۶۹

نواحی ارائه می دهند. این مفاهیم برای نمایش و توصیف شکل، نظیر مرزها، اسکلت، و بدنه' محدب مفید هستند. همان طور که در انواع روش های توصیف ارائه شده در این فصل نشان داده شد، مسأله مورد بررسی، مارابه انتخاب یک روش از میان روش های گوناگون ملزم می کند. هدف، انتخاب توصیف گرهایی است که تفاوت های اساسی بین اشیا یا گروه های اشیا را در خود "نگه می دارند" ، و در همان حال تا حد امکان مستقل از تغییرات عواملی نظیر محل، اندازه، وجهت هستند.

مراجع

رمز زنجیره ای مشروح در بخش ۱.۱.۱ اولین بار توسط Freeman[1961, 1974] پیشنهاد شد. برای مطالعه' بیشتر درباره' تقریب های چند ضلعی مقاله' SKlansky و همکارانش[1972]، کتاب Pavlidis[1961]، و مقاله ای از Bengtsson & Eklundh[1991] را ببینید. مراجع درباره' نمایه ها عبارتند از : Ambler و همکاران[1975]، Ballard & Brown[1982]، Nahim[1975] Gupta & Srinath[1988] . اطلاعات بیشتر درباره' الگوریتم های پداکردن بدنه' محدب و نقصان تحبد یک مجموعه، در Graham & Yao[1985] و Preparata & Shamos[1985] وجود دارند. الگوریتم اسکلت سازی مشروح در بخش ۱.۱.۵ از Suen & Zhang[1984] است. چند راهنمایی مفید دیگر درباره' خواص و طرز اجرای این الگوریتم در مقاله' Lu و Wang موجود است. مقاله' Chin& Jang[1990] ارتباط جالبی بین شرح بخش ۱.۱.۵ و مفهوم باریک سازی از بخش ۱.۴.۸ ایجاد می کند. گرچه تاکنون تلاش هایی برای استفاده از اسکلت های تصاویر خاکستری صورت گرفته است با داده های دودویی مرتبط است. پداکردن نقاط انتهایی و سایر نقاط بر جسته روی مرزهای رقیع، نقش مهمی در بعضی از روش های مشروح در بخش های ۱.۱ و ۱.۲ دارد. راجع به این موضوع الگوریتمی که توسط Ten & Chin[1989] پیشنهاد شد، بیشتر مورد توجه است.

مراجع بخش ۱.۲.۱ عبارتند از : Toussaint[1982]، Fischler[1980]، Shamos[1978]، Mokhtarin & Mackworth[1986]. RosenfelTen & Kak[1982] را نیز ببینید. مبحث اعداد شکلی براساس کار Bribiesca[1981]، Bribiesca & Guzman[1980] است. الگوریتم پیشنهادی

Niz برای پیدا کردن مستطیل پایه' یک منحنی بسته از رمز زنجیره ای آن قابل توجه است. مراجع بخش ۸. ۲. ۳.، ۲. ۴. برا اساس نظریه' احتمال مقدماتی است. Schalkoff[1989] را نیز بینید. مطلب بخش ۸. ۲. ۴. بر اساس نظریه' احتمال مقدماتی است. جزئیات بیشتر درباره' مطالب بخش های ۸. ۱. ۳. و ۸. ۲. ۳. در Duda & Hart[1973] و در Ballard & Brown[1982] آمده است. اخیراً توجه زیادی به توصیف گرهای بافت مبذول شده است. Bajcsy & Haralick[1973]، Lieberman[1976] برای مطالعه' بیشتر درباره' جنبه های آماری بافت Cross & Jain[1983]، Haralick[1979]، Tomita، و همکاران[1982] را بینید. درباره' بافت ساختاری Wechsler[1980] مرور خوبی بر تحلیل بافت دارد. او لیه ای از Bajcsy[1973] است. مقاله ای از Chen & Wang[1991]، Bouman & Liu[1991]، و Hu[1962] از ۴. ۳. ۸. روش گشتاور ثابت مشروح در بخش ۸. ۳. ۴. از ۴. ۳. ۸. از ۴. ۳. ۴. از ۴. ۳. ۸. در Hall& Wang[1978]، و Bell[1965] وجود دارد.

مراجع پایه ای برای بخش ۸. ۴. کتاب های Giardina & Serra[1982، 1988]، Dougherty[1988]، و Dougherty[1992] هستند. در این مورد، فصل ۶ از کتاب Vanetsanopoulos & Pits[1990] نیز مورد توجه است. مقالات Haralick و همکاران[1987]، Maragos[1987] مروری آموزشی بر روش های شکل شناسی مورد استفاده در پردازش تصویر فراهم می آورند. گرچه از نیمه' دهه' ۱۹۸۰ به بعد فنون شکل شناسی مبحث تحقیقی فعالی در آمریکا شده است، قبلاً Golay در نیمه و پایان دهه' ۱۹۶۰ عناصر مهم این مبحث را ارائه کرده بود (برای مثال، Golay[1969] و یک مقاله' جدیدتر از Preston[1983] را بینید). از اوایل دهه' ۱۹۷۰ تاکنون در اروپا کارهای زیادی روی شکل شناسی دودویی و خاکستری انجام شده است و مراجع Serra [1982]، Meyer & Beucher[1990]، Rahnamai بسیار خوبی در مورد این کارها هستند. یک مقاله جدیدتر از Schonfeld & Goutsias[1991] تشریح شد، استفاده از فیلترهای عنوان توسعه ای از مفاهیم پایه ای شکل شناسی خاکستری که در بخش ۸. ۴. معرفی شدند، مطرح می باشد. همان طور که در شکل شناسی برای بازیابی تصاویر دودویی، مثال دیگری از سودمندی مطالب بخش ۸. ۴. می باشد.

قضیه' مشهور به قضیه همشکلی آمبرا^۱، پیوندی پایه ای بین شکل شناسی دودویی و خاکستری ایجاد می کند. گرچه طرح این مبحث خارج از محدوده' شرح بخش ۸.۴ بود، با مراجعه به کتاب های Dougherty[1992]، Giardina & Dougherty[1988]، Serra[1982] و نیز مقاله' Haralick و همکاران [1987] می توان فهم عمیق تری درباره' این مبحث به دست آورد. روشی برای تجزیه' اعمال شکل شناسی خاکستری به اعمال شکل شناسی دودویی توسط Mitchell & Shih[1989] و Fu[1982] ارائه شد. در نهایت مراجع بخش ۸.۵ Thomason & Gonzalez[1978] و Gonzalez & Thomason[1978] هستند.

مسائل

۱.۸ (الف) نشان دهید که با تعریف مجدد نقطه' شروع یک رمز زنجیره ای طوری که دنباله' اعداد حاصل عدد صحیحی با بزرگی کمینه بشود، رمز از نقطه' شروع ابتدائی در روی رمز مستقل می شود.

(ب) نقطه' شروع تراز شده رمز ۱۱۰۷۶۷۶۵۵۴۳۳۲۲ را پیدا کنید.

۲.۸ (الف) نشان دهید که، طبق توضیع بخش ۱.۸، تفاضل اول رمز زنجیره ای نسبت به چرخش، تراز شده است.

(ب) تفاضل اول رمز ۱۱۱۲۲۱۳۳۳۲۳۲۳۰۳۰۳۰۱۰۱۰۱ را محاسبه کنید.

۳.۸ (الف) نشان دهید که روش تقریب چند ضلعی توار لاستیکی مشروح در بخش ۱.۸، یک چند ضلعی با محیط کمینه را نتیجه می دهد.

(ب) نشان دهید که اگر هر سلول، متناظر با یک پیکسل مرزی باشد، خطای بیشینه' ممکن در آن سلول $d\sqrt{2}$ خواهد بود، که d فاصله' شبکه ای بین پیکسل ها است.

۴.۸ (الف) اگر در روش ادغام مشروح در بخش ۱.۸ آستانه' خطای صفر قرار داده شود، اثر آن بر چند ضلعی حاصل را شرح دهید.

(ب) اثر آن بر روش تقسیم چه خواهد بود؟

۵.۸ (الف) نمایه' یک مرز مربعی را با استفاده از روش زاویه' مماس که در بخش ۱.۸.۳ تشریح شد، رسم کنید.

(ب) تابع چگالی شبکه آن را نیز رسم کنید. فرض کنید که مریع در راستای محورهای x و y است،

1) *Umbra homomorphism theorem*

و فرض کنید که محور X خط مرجع است. از نزدیکترین گوش به مبدأ شروع کنید.

۶.۸ محور میانی (الف) دایره، (ب) مربع، و (ج) مثلث متساوی الاضلاع را باید.

۷.۰.۸ (الف) برای هریک از اشکال زیر، عملی را که در مرحله' یک از الگوریتم باریک سازی بخشن

۱.۸.۵. اجرا می شود، شرح دهید.

(ب) مرحله' دورانیز شرح دهید. در تمام موارد فرض کنید که $p = 1$.

1	1	0
1	p	0
1	1	0

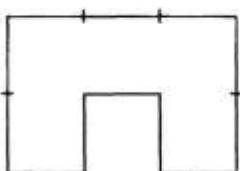
0	0	0
1	p	0
0	0	0

0	1	0
1	p	1
0	1	0

1	1	0
0	p	1
0	0	0

۸.۸.۸ (الف) مرتبه' عدد شکلی برای شکل زیر چیست؟

(ب) عدد شکلی آن را بدست آورید.



روال مشروح در بخش ۷.۰.۸. برای استفاده از توصیف گرهای فوریه، شامل بیان مختصات منحنی به صورت اعداد مختلط، محاسبه DFT این اعداد، و نگه داشتن تعداد خیلی کمی از مؤلفه های DFT به عنوان توصیف گرهای شکل مرز می باشد. آنگاه DFT معکوس، تقریبی از منحنی اولیه را نتیجه می دهد. DFT چه گروهی از شکل منحنی ها، از اعداد حقیقی تشکیل می شود و سامانه محور های شکل ۸.۱۴ چگونه باید تنظیم شود که اعداد حقیقی بدست آیند؟

۹.۰.۸ کمترین تعداد توصیف گرهای گشتاوری را که برای تفکیک نمایه های شکل ۸.۵ مورد نیاز هستند، تعیین نماید.

۱۱.۸ عدد اولر نویسه های ۰, ۱, ۸, ۹ و X را باید.

۱۲.۸ شما باید یک سامانه پردازش تصویر برای آشکارسازی نواقص درون صفحات پلاستیکی جامد معین شده طراحی کنید. صفحات بوسیله یک سامانه عکسبرداری اشعه ایکس کم توان که تصاویر $8 \times 512 \times 512$ می دهد، بررسی می شوند. در غیاب نواقص، تصاویر همواری ظاهر می شوند که میانگین سطح خاکستری آنها 100 و واریانس نویز 400 است. نواقص به صورت نواحی شبه جبابی ظاهر می شوند که پیکسل های آن نواحی به اندازه 50×20 باشند، صفحه معیوب در نظر گرفته می شود. با استفاده از تحلیل بافت راهی برای این مسئله پیشنهاد نمایید.

۱۳.۸ ماتریس وقوع همزمان سطح خاکستری یک تصویر 5×5 را که بخشی از یک صفحه شطرنجی با "۱" ها و "۰" های متناظر است، به دست آورید اگر (الف) عملگر مکان P به عنوان "یک پیکسل حرکت به راست" ، و (ب) "دو پیکسل حرکت به راست" تعریف شود. فرض کنید که پیکسل سمت چپ بالا مقدار "۰" دارد.

۱۴.۸ یک تصویر شطرنجی مشکل از مربعهای $m \times m$ سیاه و سفید متناظر در نظر بگیرید. یک عملگر مکان معین کنید که ماتریس وقوع همزمان قطری دهد.

۱۵.۸ (الف) گسترش دایره ای به شعاع 2 با یک عنصر ساختمانی دایروی با شعاع $2/4$ را رسم کنید.
 (ب) با استفاده از همین عنصر ساختمانی یک مربع 2×2 را نیز گسترش دهید.
 (ج) با استفاده از همان عنصر ساختمانی یک مثلث متساوی الاضلاع به طول ضلع 2 را نیز گسترش دهید.

(د) قسمت های (الف) - (ج) را برای سایش نیز تکرار کنید.

۱۶.۸ مثال گسترش نمایش داده شده در اشکال ۸ . ۲۶ (ت) و (ث) را تکرار کنید، اما بدأ عنصر ساختمانی را در گوشه سمت چپ بالای آن فرض کنید.

۱۷.۸ (الف) را برای سایش نیز تکرار کنید و حاصل را با اشکال ۸ . ۲۷ . (ث) مقایسه نمایید.
 با استفاده از مفاهیم معرفی شده در بخش ۸ . ۴ . یک الگوریتم شکل شناسی برای تبدیل مرز متصل هشتگانه به مرز متصل m گانه طرح نمایید (بخش ۲ . ۴ . را بینید). یک شرط مهم این است که الگوریتم شما تواند اتصال راقطع کند. می توانید فرض کنید که مرز ورودی

کاملاً متصل و به ضخامت یک پیکسل است (اما می تواند نقاط انشعاب داشته باشد).

درستی معادله' (۸. ۴-۱۰) را ثابت کنید.

یک تعریف دیگر از گسترش عبارتست از

$$A \oplus B = \{c \in Z^2 | c = a + b, \quad a \in A \quad \text{و} \quad b \in B\}$$

{برای بعضی از a و b هایی که}

(الف) نشان دهید که این تعریف معادل تعریف موجود در معادله' (۸. ۴-۵) است.

(ب) نشان دهید که این تعریف با تعریف دیگر گسترش به صورت $A \oplus B = \bigcup_{n \in N} (A)$ معادل است.

(این معادله به عنوان جمع مینکووسکی^۱ دو مجموعه نیز خوانده می شود.)

(ج) نشان دهید که $A \oplus B = \bigcup_{n \in N} (A)$ معادل تعریف ارائه شده در معادله' (۸. ۴-۵) می باشد.

یک تعریف دیگر از سایش عبارتست از

$$A \ominus B = \{c \in Z^2 | c + b \in A, \quad b \in B\}$$

{برای هر b که}

(الف) نشان دهید که این تعریف معادل تعریف ارائه شده با معادله' (۸. ۴-۶) است.

(ب) نشان دهید که این تعریف با تعریف دیگر سایش به صورت $A \ominus B = \bigcap_{n \in N} (A)$ نیز معادل است.

(اگر $b - a$ جایگزین شود، این عبارت تفریق مینکووسکی دو مجموعه نیز خوانده می شود.)

(ج) نشان دهید که $A \ominus B = \bigcap_{n \in N} (A)$ معادل تعریف ارائه شده با معادله' (۸. ۴-۵) می باشد.

فرض کنید که تصویر $f(x, y)$ و عنصر ساختمانی (y, b) در معادله' (۸. ۴-۳۱) هر دو

مستطیل شکل هستند و حوزه های D_f و D_b به ترتیب با $([B_{11}, B_{12}], [F_{11}, F_{12}])$ و $([B_{21}, B_{22}], [F_{21}, F_{22}])$

[نشان داده می شوند. مثلاً بازه های بسته $[F_{11}, F_{12}]$ و $[F_{21}, F_{22}]$ محدوده های

تغییر y و x در نقاطی از صفحه' xy هستند که تابع $f(x, y)$ در آنها تعریف شده است.

(الف) با فرض $(x, y) \in D_b$ عباراتی را برای بازه های تغییر متغیرهای y و x طوری که

معادله' (۸. ۴-۳۱) برآورده شود، به دست آورید. این بازه ها در محورهای x و y حوزه'

مستطیلی $(s, t) \oplus (s, t)$ در صفحه' st را تعریف می کنند.

1) Minkowsky addition

ب) قسمت (الف) را برای سایش با تعریف ارائه شده در معادله '۸ . ۴-۳۲' تکرار کنید.

۲۲۰.۸ تصویر خاکستری ($y(x)$) بوسیله شبه ضربه های نویزی ناهمپوشانه که می توان آنها را به صورت استوانه های کوچک با شعاع های $R_{min} \leq r \leq R_{max}$ و دامنه های $A_{min} \leq a \leq A_{max}$ مدل کرد، خراب شده است.

(الف) یک روش فیلتر کردن شکل شناسی برای حذف این نویز از تصویر طرح کنید.

(ب) قسمت (الف) را تکرار کنید، اما اکنون فرض کنید که حداقل چهار شبه ضربه نویزی همپوشاننده در تصویر وجود دارد.

۲۳۰.۸ ۲۳.۸ یک رابطه مکانی^۱ و نمایش درختی متناظر آن را برای الگوی شطرنجی مرکب از مربع های سیاه و سفید، ارائه کنید. فرض کنید که عنصر سمت چپ بالا، سیاه است، و ریشه درخت متناظر با آن عنصر است. درخت شمانی تواند در هر گره بیش از دو انشعاب داشته باشد.

فصل ۹

تشخیص و تعبیر

یکی از جالب ترین جلوه های جهان این است که می توان آن را ترکیبی از الگوها دانست. یک الگو ترتیب و آرایشی از عناصر است. الگو با ترتیب عناصر سازنده اش و نه با طبیعت ذاتی آنها، شناخته می شود.^۱

نوربرت وینر^۲

بحث پردازش تصویر رقمی را با ارائه چند روش تشخیص و تعبیر تصویر پایان می دهیم. مطالب این فصل بیشتر به کاربردهایی که نیازمند تحلیل خودکار تصویر هستند، مربوط می شود.

تحلیل تصویر شامل فرآیندهای کشف، شناسایی، و فهم الگوهای مرتبط با یک کار تصویری است. یکی از اهداف اصلی تحلیل رایانه ای تصویر این است که ماشین بتواند بعضی توانایی های انسان را تا حدودی تقریب بزند. برای مثال در سامانه ای که مخصوص خواندن نوشته های تایپ شده است، الگوهای مورد علاقه، نویسه های الفبایی- عددی هستند و هدف، رسیدن به آن چنان دقیقی در تشخیص نویسه است که تا حد ممکن به توانایی بالای انسان در انجام این کار نزدیک باشد.

بنابراین سامانه خودکار تحلیل تصویر باید بتواند درجات گوناگونی از هوشمندی را ارائه کند. مفهوم هوشمندی^۳ تا حدی، بویژه راجع به ماشین، تقریبی است. با این حال شناخت انواع مشخصه هایی که معمولاً با هوشمندی مرتبط هستند، مشکل نیست. در این مورد چند مشخصه فورآ به ذهن می آیند: (۱) توانایی استخراج اطلاعات مرتبط، از زمینه ای با جزئیات نامرتب؛ (۲) توانایی یادگیری از مثال ها و تعمیم آن طوری که در شرایط جدید و متفاوت قابل استفاده باشد و (۳) توانایی استنتاج از اطلاعات ناقص.

سامانه های تحلیل تصویر با چنین مشخصاتی را تنها برای محیط های عملیاتی محدودی می توان طراحی و پیاده سازی کرد. ماهنوز نمی دانیم که چگونه می توان کارآیی این سامانه ها را حتی به نزدیک توانایی های

انسان در اجرای کارهای عمومی تحلیل تصویر رساند. با این حال تحقیقات در مورد سامانه های زیستی و محاسباتی، به طور دائم نظریه های جدید و امیدبخشی در چگونگی تشخیص بصری انسان مطرح می کند. با این حال، قسمت اعظم تحلیل تصویر رایانه ای کنونی براساس روابط تجربی که برای حل مسائل خاصی طرح شده اند، می باشد. مثلاً بعضی ماشین ها قادرند نوشه های چاپی را که قالب مناسبی داشته باشند، با سرعت هایی که بارها سریع تر از سرعت خواندن ماهرترین انسان ها است، بخوانند. با این حال این نوع سامانه ها بسیار خاص هستند و توسعه پذیری کمی دارند یا اصلًا توسعه پذیر نیستند. بنابراین محدودیت های عملی و نظری فعلی در مبحث تحلیل تصویر، راه حل هایی را تحمیل می کنند که شدیداً وابسته به مسئله هستند.

در بخش ۱.۵ گفته شد که فصول ۱-۳ به مطالب پایه می پردازند، فصول ۴-۶ برای پیش پردازش تصویر مفیدند، و فصول ۷-۹ بیشتر به کاربردهای تحلیل تصویر مربوط می شوند. این نحوه گروه بندی، گرچه از نظر سازماندهی مطالب به طور جدی مورد نیاز است، نباید این گونه تعبیر شود که مطالب فصول ۱-۶ برای تحلیل تصویر مفید نیستند. تصویربرداری، ارتقا، بازیابی، و فشرده سازی برای کاربردهایی که داده های تصویر رقمه را تحت پردازش انسانی یا ماشینی قرار می دهند، مشترک هستند. در توضیحات بعدی روشن می شود که طراحی سامانه های تحلیل تصویر، نیازمند آگاهی از غالب مطالب این کتاب است.

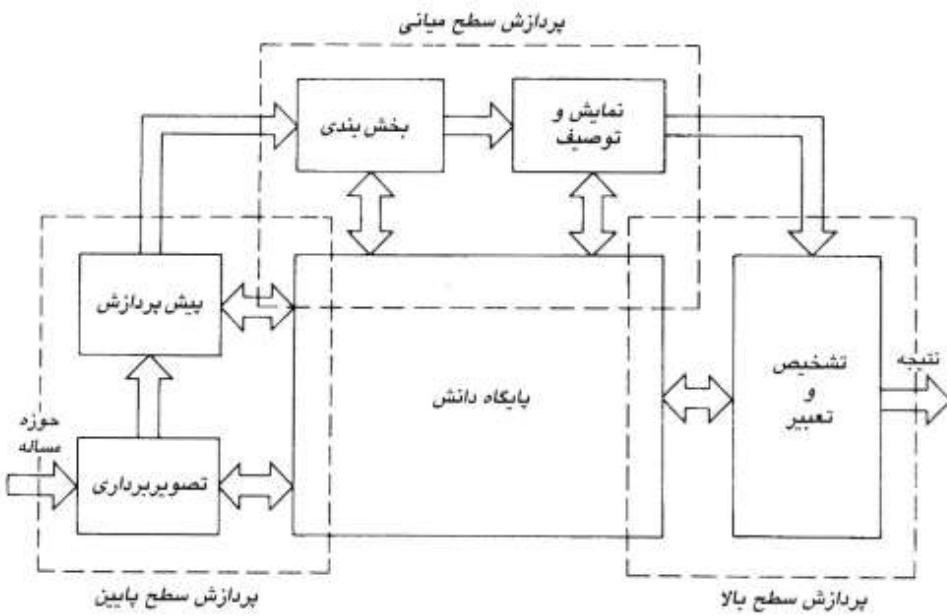
۱.۹ عناصر تحلیل تصویر

تقسیم روش های تحلیل تصویر به سه گروه اصلی مفید به نظر می رسد. این گروه ها عبارتند از: (۱) پردازش سطح پایین، (۲) پردازش سطح میانی، و (۳) پردازش سطح بالا. گرچه این تقسیم بندی ها مرز های قطعی ندارند، اما چهار چوب مفیدی برای دسته بندی فرآیندهای مختلفی که اجزای اصلی سامانه خودکار تحلیل تصویر هستند، فراهم می آورند. شکل ۱.۹ این مفاهیم را جمع بندی می کند. همپوشانی خطوط خط چین نشان دهنده عدم وجود مرز مشخص بین فرآیندها هستند. برای مثال، بسته به کاربرد می توان آستانه گیری را ابزار ارتقا (پیش پردازش) یا ابزار بخش بندی در نظر گرفت.

پردازش سطح پایین^۱ به اعمالی گفته می شود که واکنش هایی خودکار هستند و اصلًا نیازی به

¹⁾ Low-level processing

۶۷۹ ۱۰.۹ عناصر تحلیل تصویر



شکل ۱۰.۹ عناصر تحلیل تصویر

هوشمندی ندارند. ما تصویربرداری و پیشپردازش (فصل ۶-۲) را به عنوان اعمال سطح پایین در نظر می‌گیریم. این گروه شامل فعالیت‌هایی از تشکیل تصویر تا جبران سازی هایی نظیر کاهش نویز یا مات‌زادی است. توابع سطح پایین را می‌توان با فرآیندهای سنجش و تطابق که شخص هنگام ورود از هوای آفتابی به یک سالن تاریک برای جستجوی صندلی خالی انجام می‌دهد، مقایسه کرد. قبل از این که تصویر مناسبی موجود باشد، نمی‌توان فرآیند (هوشمندانه) پیدا کردن جای خالی را شروع کرد. فرآیند بعدی مغز، تطابق سامانه بینایی برای تولید تصویر مناسب است، که واکنشی خودکار و غیر هوشمندانه می‌باشد. پردازش سطح میانی^۱ به عمل استخراج و مشخص کردن اجزای (مثلاً نواحی) تصویر که حاصل فرآیند سطح پایین هستند، می‌پردازد. همان طور که شکل ۱۰.۹ نشان می‌دهد، فرآیندهای سطح میانی شامل بخش بندی و توصیف با استفاده از روش‌های نظیر روش‌های فصل ۷ و ۸ می‌باشد. برای رفتار هوشمندانه باید در روشهای بخش بندی انعطاف پذیر بعضی توانایی‌ها ایجاد شود. مثلاً برای پونزدزدن انفصالت کوچک در مرز بخش بندی شده، به عناصر کارآمدتری نسبت به واکنش‌های خودکار سطح پایین نیاز است.

1) *Intermediate-level processing*

در نهایت پردازش سطح بالا^۱ شامل تشخیص و تعبیر است که موضوعات اصلی این فصل می باشند. این دو فرآیند تشابه بیشتری به مفهوم عام درک هوشمندانه^۲ دارند. اکثریت روش های مورد استفاده در پردازش سطح پایین و میانی روابط نظری نسبتاً تعریف شده ای دارند. در هر حال وقتی به مبحث تشخیص و خصوصاً تعبیر، وارد می شویم، دانش و فهم ما از اصول پایه ای آن کم دقت تر و سطحی تر است. این کمبود نسبی فهم، در نهایت منجر به مطرح شدن محدودیت ها و مطلق گرایی هایی می شود که به منظور کاهش پیچیدگی در سطحی قابل قبول لازم هستند. محصول نهایی، سامانه ای با قابلیت های عملیاتی بسیار خاص می باشد.

بخش های بعدی به (۱) روش های تشخیص با نظریه تصمیم^۳ تشخیص، (۲) روش های ساختاری تشخیص، و (۳) روش های تعبیر تصویر می پردازند. تشخیص با نظریه تصمیم (بخش ۹ . ۳) بر اساس نمایش الگوها به شکل بردار و سپس جستجوی روش هایی برای گروه بندی و انتساب این بردارها به دسته الگوهای متفاوت است. روش های اصلی تشخیص با نظریه تصمیم، دسته بندی کننده های حداقل فاصله، همبسته سازها، دسته بندی کننده بیز^۴، و شبکه های عصبی هستند. در تشخیص ساختاری (بخش ۹ . ۴) الگوها به شکل نمادین (نظیر رشته ها یا درخت ها) نمایش داده می شوند، و روش های تشخیص بر اساس تطبیق نماد یا بر اساس مدل هایی هستند که با الگوهای نمادی به عنوان جمله هایی از یک زبان مصنوعی رفتار می کنند. تعبیر تصویر (بخش ۹ . ۵) به مجموعه ای از عناصر تصویری تشخیص داده شده، معنای نسبت می دهد. مهمترین مفهومی که زیربنای روش های تعبیر تصویر است، سازماندهی و استفاده مؤثر از دانش درباره حوزه مورد نظر است. روش های فعلی تعبیر تصویر بر اساس منطق گزاره ها، شبکه های معنایی^۵، و سامانه فرآوری^۶ (بويژه خبره^۷) هستند.

۲.۹ الگوها و دسته های الگو

همان طور که در بخش ۹ . ۱ بیان شد، قابلیت تشخیص الگو تا حدی برای تحلیل تصویر اساسی است. در اینجا الگو^۸، توصیفی کمی یا ساختاری از شئی با هر چیز مورد علاقه دیگری در تصویر است. به طور کلی الگو از یک یا چند توصیف گر، نظیر توصیف گرهای فصل ۸، تشکیل می شود. به عبارت

1) High-level processing

2) Intelligent cognition

3) Decision-theoretic method

4) Bayes

5) Semantic networks

6) Production system

7) Expert

8) Pattern

۶۸۱ ۲. الگوها و دسته های الگو

دیگر، الگو، چیدمانی^۱ از توصیف گرها است. (در نوشته های راجع به تشخیص نویسه غالب از کلمه ویژگی ها^۲ برای دلالت بر توصیف گرها استفاده می شود). یک دسته الگو^۳ خانواده ای از الگوها است که تعدادی خواص مشترک دارند. دسته های الگو بانمادهای w_1, w_2, \dots, w_M نشان داده می شوند، که M تعداد دسته ها است. تشخیص الگوی ماشینی روش هایی را به کار می برد تا الگوهارا- به طور خودکار و با کمترین حد ممکن دخالت انسان- به دسته های متناظر شان منتب نماید.

سه چیدمان اصلی الگو که در عمل استفاده می شوند، بردارها (برای توصیف ها کمی)، و رشته ها و درخت ها (برای توصیف های ساختاری) هستند. بردارهای الگو با حروف کوچک پررنگ نظیر x, y, z نمایش داده می شوند و به شکل زیر هستند:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (1-2.9)$$

که مؤلفه^۴ x_i بیانگر i امین توصیف گر است و n تعداد چنین توصیف گرهایی می باشد. بردارهای الگو به صورت ستون (یعنی ماتریس های $1 \times n$) نمایش داده می شوند. بنابراین بردار الگو را می توان به شکل معادله^۵ (۱-۲.۹) یا به شکل $(x_1, x_2, \dots, x_n)^T = \mathbf{x}$ بیان کرد، که T بیانگر عمل ترانهاده گیری^۶ است.

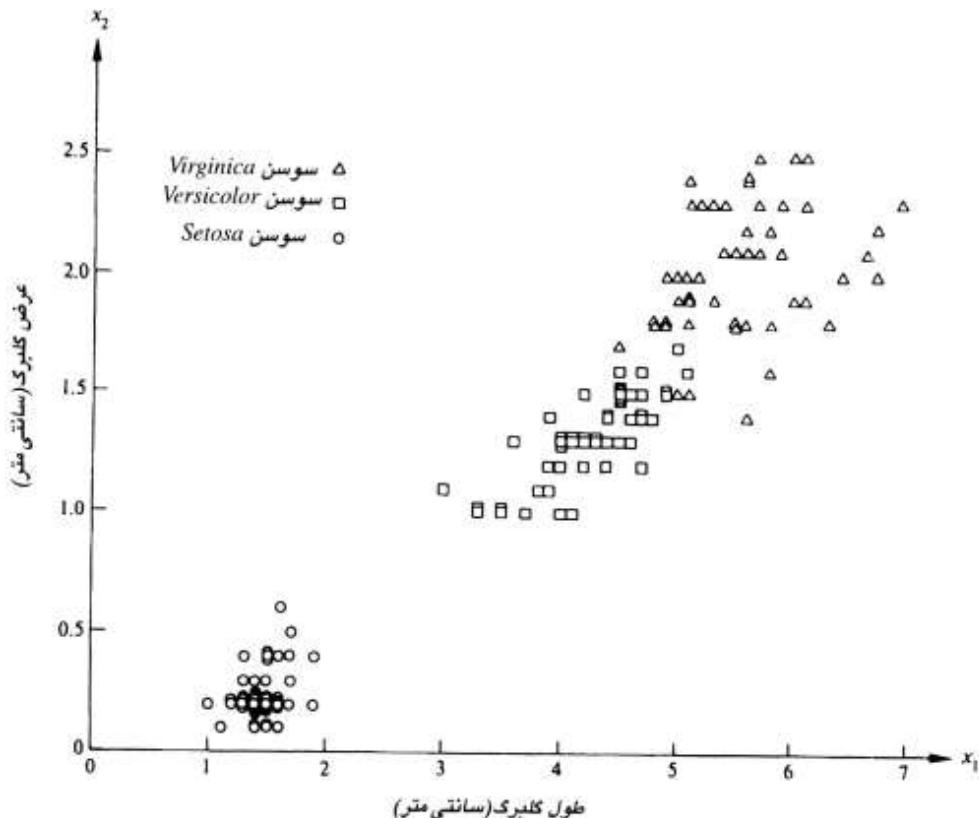
طبیعت مؤلفه های بردار الگوی \mathbf{x} به روش اندازه گیری مورد استفاده در توصیف الگو واقعی بستگی دارد. مثلاً فرض کنید مثلاً می خواهیم سه نوع از انواع گلها^۷ی سوسن (*Setosa*، *Virginica*، و *Versicolor*) را با اندازه گیری طول و عرض گلبرگ هایشان توصیف کنیم. در این مورد با بردارهای دو بعدی به شکل

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (2-2.9)$$

سروکار داریم که x_1 و x_2 به ترتیب متناظر با طول و عرض گلبرگ هستند. در این مورد سه دسته الگو که با w_1, w_2 و w_3 نشان داده می شوند، به ترتیب متناظر با گونه های *Setosa*، *Virginica*، و

Versicolor هستند.

چون طول و عرض گلبرگ های تمام گل ها تا حدی متغیرند، بردارهای الگویی که این گل ها را توصیف می کنند، نه تنها بین دسته های متفاوت، بلکه درون یک دسته نیز متغیرند. شکل ۲.۹ اندازه گیری های طول و عرض چند نمونه از هر نوع گل سوسن را نشان می دهد. پس از این که مجموعه ای از سنجه ها^۱ (در این مورد دو تا) انتخاب شدند، بردار الگو نمایش کاملی برای هر نمونه 'واقعی خواهد بود. بنابراین هر گل در این مورد نقطه ای در فضای اقلیدسی دو بعدی است. باید ذکر کنیم که در این مورد سنجه های طول و عرض دسته سوسن Setosa را به اندازه^۲ کافی از دو دسته 'دیگر جدا کرد، اما انواع Versicolor و Virginica را در حد مناسبی از هم جدا نکرد. این نتیجه، مسأله کلاسیک انتخاب ویژگی^۳



شکل ۲.۹ اندازه گیری طول و عرض گلبرگ های تعدادی گل سوسن از سه نوع مختلف (برگرفته از & Duda (Hart 1973)

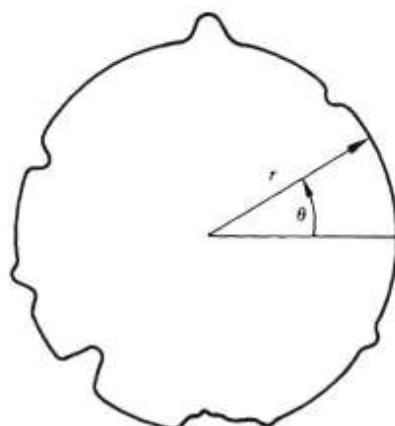
1) Measurements

2) Feature selection

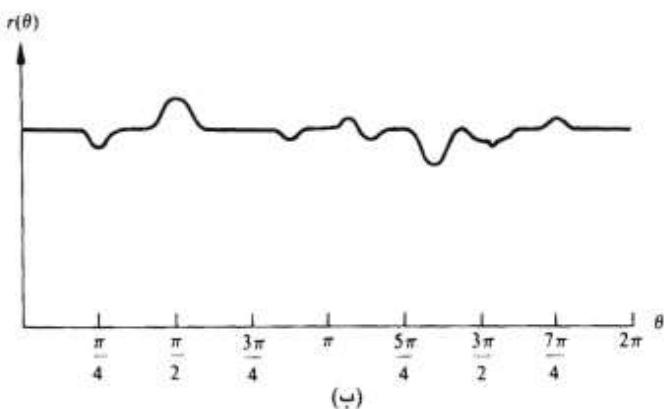
۶۸۳ ۹. الگوها و دسته های الگو

رامطراح می کند، که در آن درجه^۱ جداپذیری دسته ها بشدت به سنجه های منتخب برای آن کاربرد بستگی دارد. در بخش ۳. ۳. ۳. بیشتر در این مورد صحبت می کنیم.

شکل ۹. ۹ مثال دیگری از تولید بردار الگورانشان می دهد. در این مورد، به انواع متفاوتی از اشکال نویزدار علاقه مندیم که نمونه ای از آن در شکل ۹. ۳ (الف) دیده می شود. اگر هر شبیه را با نمایه^۱ آن (بخش ۸. ۳. ۱ را ببینید) توصیف کنیم، سیگنال های یک بعدی به شکل ۹. ۳(ب) را به دست می آوریم. با نمونه برداری از این توابع در مقادیر مشخصی از θ ، که با $\theta_1, \theta_2, \dots, \theta_n$ نشان داده می شوند، و قرار دادن $x_1 = r(\theta_1), x_2 = r(\theta_2), \dots, x_n = r(\theta_n)$ می توان بردارهای الگورا تشکیل



(الف)



شکل ۹. ۹ شبیه نویزدار و نمایه متناظر با آن

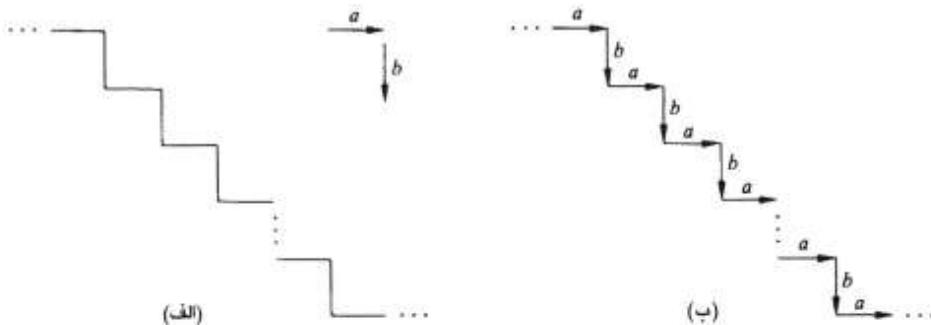
داد. این بردارها در فضای اقلیدسی // بعدی به نقطه تبدیل می شوند، و هر دسته الگو در فضای // بعدی به صورت «ابری» حول آن نقطه ظاهر می شود.

به جای استفاده مستقیم از دامنه نمایه، می توان به عنوان مثال ابتدا // گشتاور یک نمایه معین (بخش ۸. ۴.) را محاسبه و سپس از آنها به عنوان مؤلفه های بردار الگوی مربوطه استفاده کرد. در واقع همان طور که احتمالاً تاکنون روشن شده است، به راه های بیشمار دیگری نیز می توان بردارهای الگو را تولید کرد، که بعضی از آنها در این فصل ارائه می دهیم. در مورد گشتاور، مفهوم کلیدی که باید به آن توجه داشته باشیم، این است که انتخاب سنجه یا سنجه هایی که هر مؤلفه بردار الگو را براساس آن ها تعیین می کنیم، اثر عمیقی بر عملکرد نهایی سامانه تحلیل تصویر مبتنی بر روش بردار الگو دارد.

استفاده از این ترفندها برای تولید بردارهای الگو، دسته هایی از الگو که با اطلاعات کمی مشخص می شوند، را به دست می دهدند. در بعضی کاربردها، مشخصات الگوها به بهترین وجه ممکن با روابط ساختاری توصیف می شوند. مثلاً تشخیص اثر انگشت براساس روابط بین ویژگی های ریز^۱ آن است. این ویژگی ها اجزایی ابتدایی هستند که خواص خطوط اثر انگشت، نظیر قطع فوری خطوط، انشعاب، ادغام و عدم اتصال قطعه ها را همراه با اندازه و محل نسبی آنها، توصیف می کنند. این نوع از مسائل تشخیص، که در آنها علاوه بر معیارهای کمی درباره هر ویژگی، روابط مکانی بین ویژگی هاییز در تعیین عضویت در یک دسته نقش دارند، معمولاً با روش های ساختاری بهتر حل می شوند.

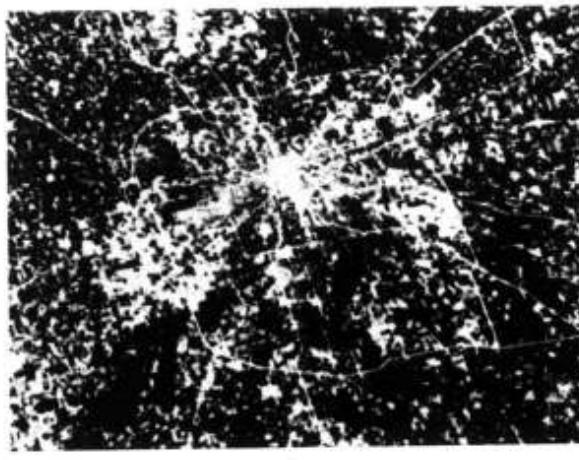
شکل ۹. ۴. (الف) الگوی پلکانی ساده ای را نشان می دهد. این الگو باید مشابه روش مورد استفاده در شکل ۹. ۳. نمونه برداری و سپس بر حسب بردار الگو بیان شود. در هر حال، در این روش توصیف، ساختار پایه که از تکرار دو عنصر ابتدایی ساده تشکیل شده است، از دست می رود. همان طور که در شکل ۹. ۴. (ب) دیده می شود، توصیف معنادارتر این است که عناصر a و b را تعریف کرده و اجازه دهیم که الگو به صورت رشته نمادهای ... $ababab = \dots$ تعریف شود. در این نحوه نمایش ساختار این دسته خاص از الگوها حفظ می گردد به شرط آن که اتصال با روش ابتدا - به - انتها تعریف شود و تنها نمادهای متناب، مجاز باشند. این روش ساختاری به الگوی پلکانی با هر طولی قابل اعمال است، اما در مورد سایر ساختارهایی که می توان با ترکیبات دیگر عناصر اولیه a و b تولید کرد، قابل استفاده نیست (به خاطر آورید که روش های پیچیده تر تولید رشته های الگو در بخش ۸. ۵ بیان شدند).

۶۸۵ ۲۰. الگوهای دسته های الگو



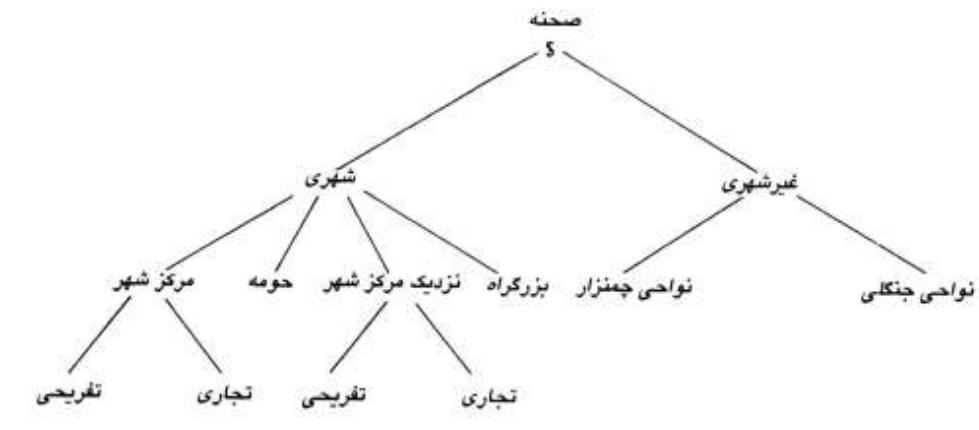
شکل ۹.۴ (الف) ساختار پله ای؛ (ب) ساختاری که بر حسب عناصر ابتدایی a و b رمز شده تا نمایش رشته ای ... $ababab$... بدمت آید

نمایش های رشته ای به میزان کافی الگوهای اشیاء و سایر چیزهایی که ساختارشان مبتنی بر اتصال نسبتاً ساده اجزای ابتدایی که معمولاً مربوط به شکل مرز هستند، تولید می کند. روشی قوی تر برای بسیاری از کاربردها استفاده از توصیف گرهای درختی است، که در بخش ۸.۵ تعریف شدند. در واقع اغلب روش های مرتب سازی مراتبی به ساختارهای درختی منتهی می شوند. مثلاً، عکس هوانی شکل ۹.۵ (الف) صحنه ای را که مرکب از نواحی شهری و روستایی است، نشان می دهد. اجازه دهید کل ناحیه، تصویر را با نماد $\$$ تعریف کنیم. نمایش درختی (وازگون) شکل ۹.۵ (ب) با استفاده از



(الف)

شکل ۹.۵ (الف) عکس هوانی نواحی شهری و غیر شهری؛ (ب) نمایش درختی آن.
. (Brayer, Swain & Fu [1977] از)



شکل ۹.۵ ادامه

رابطه 'ساختاری "مركب از"^۱' به دست آمد. بنابراین ریشه 'درخت کل صحنه را نمایش می دهد. سطح بعدی نشان می دهد که صحنه، متشكل از نواحی شهری و روستایی است. ناحیه 'روستایی متشكل از نواحی چمنزار و جنگلی است. ناحیه 'شهری را می توان به اجزای مرکز شهر، نزدیک مرکز شهر، حومه و بزرگراه ها تقسیم کرد. در نهایت، ناحیه 'مرکز شهر به اجزای تفریحی و تجاری تقسیم می شود. می توانیم این نوع تقسیم بندی را ادامه دهیم تا جایی که به سقف توانایی مان در جدادسازی مناطق مختلف تصویر بررسیم. روش های استفاده از هر یک از آرایش های مختلف الگو در مرحله 'تشخیص سامانه های تحلیل تصویر را در طی دو بخش بعدی ارائه می دهیم.

۳.۹ روش های نظریه تصمیم

رهیافت های تشخیص با نظریه تصمیم، مبتنی بر استفاده از توابع تصمیم (یا تبعیض)^۲ هستند. فرض کنید که $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ بیانگر بردار الگوی \mathbf{x} بعدی، که در بخش ۹.۲ تشریح شد، باشد. مشکل اصلی در تشخیص الگو با نظریه تصمیم برای M دسته الگوی w_1, w_2, \dots, w_M ، پیدا کردن M تابع تصمیم $d_1(\mathbf{x}), d_2(\mathbf{x}), \dots, d_M(\mathbf{x})$ است با این خاصیت که اگر الگوی \mathbf{x} متعلق به دسته w_i باشد، آنگاه

$$d_i(\mathbf{x}) > d_j(\mathbf{x}) \quad j = 1, 2, \dots, M; j \neq i. \quad (1-3.9)$$

1) Composed of

2) Decision (or discriminant) functions

۶۸۷ ۳. روش های نظریه تصمیم

به عبارت دیگر، الگوی نامعلوم \mathbf{x} به دسته الگوی ω تعلق دارد اگر، پس از جای گذاری \mathbf{x} در تمام توابع تصمیم، $d_i(\mathbf{x})$ بزرگترین مقدار عددی را بدهد. قیود به دلخواه تعیین می شوند.

مرز تصمیمی که دسته ω را از دسته ω' جدا می کند، با مقادیری از \mathbf{x} که برای آنها $d_i(\mathbf{x}) = d_{i'}(\mathbf{x})$ باشد، صورت معادل، با مقادیری از \mathbf{x} که برای آنها

$$d_i(\mathbf{x}) - d_{i'}(\mathbf{x}) = 0. \quad (2-3. 9)$$

باشد، تعیین می شود. روال معمول این است که مرز تصمیم بین دو دسته با تابع $d_i(\mathbf{x}) - d_{i'}(\mathbf{x}) = 0$ تعیین شود. بنابراین برای الگوهای دسته ω ، $d_i(\mathbf{x}) > 0$ و برای الگوهای دسته ω' ، $d_{i'}(\mathbf{x}) < 0$.

هدف اصلی مطالب این بخش، ارائه روش های گوناگون جهت پیدا کردن توابع تصمیمی که معادله '۱-۳. ۹) را بآورده می کنند، می باشد.

۱.۳.۹ تطبیق

دسته بندی کننده حداقل فاصله

فرض کنید که هر دسته الگو با یک بردار پیش نمونه (یا میانگین)¹⁾ به شکل

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{\mathbf{x} \in \omega_j} \mathbf{x} \quad j = 1, 2, \dots, M \quad (3-3. 9)$$

بیان شود که N_j تعداد بردارهای الگو در دسته ω_j است و عمل جمع روی این بردارها اجرا می شود. یک راه برای تعیین عضویت بردار الگوی مجهول \mathbf{x} در دسته این است که آن را به دسته ای منتسب نماییم که نزدیکترین بردار پیش نمونه را به این الگو دارد. استفاده از فاصله اقلیدسی برای تعیین نزدیکی، مسأله را به محاسبه 'معیار فاصله' می پرساند.

$$D_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_j\| \quad j = 1, 2, \dots, M \quad (4-3. 9)$$

ساده می کند که $D_j(\mathbf{x}) = \|\mathbf{a}\|^{1/2}$ نرم اقلیدسی است. آنگاه اگر $D_j(\mathbf{x})$ کوچکترین فاصله باشد، \mathbf{x} را به دسته ω_j منتسب می کنیم. یعنی، در این تعریف نزدیکترین فاصله به مفهوم بهترین تطبیق است. به سادگی

1) Prototype (or mean) vector

می توان نشان داد (مسئله ۹. ۲.) که این تعریف معادل ارزیابی توابع

$$d_j(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_j - \frac{1}{2} \mathbf{m}_j^T \mathbf{m}_j, \quad j = 1, 2, \dots, M \quad (5-3. 9)$$

و انتساب \mathbf{x} به دسته' ω_j است اگر $d_j(\mathbf{x})$ بزرگترین مقدار عددی را بدهد. همان طور که در معادله' ۹. ۳. ۱) تعریف شد، این بیان با مفهوم تابع تصمیم موافق است.

طبق معادلات ۹. ۲-۳) و (۹. ۵-۳)، مرز تصمیم بین دسته های ω_i و ω_j برای یک دسته بندی کننده' حداقل فاصله عبارتست از:

$$\begin{aligned} d_{ij}(\mathbf{x}) &= d_i(\mathbf{x}) - d_j(\mathbf{x}) \\ &= \mathbf{x}^T (\mathbf{m}_i - \mathbf{m}_j) - \frac{1}{2} (\mathbf{m}_i - \mathbf{m}_j)^T (\mathbf{m}_i - \mathbf{m}_j) = 0. \end{aligned} \quad (6-3. 9)$$

سطح معین شده با معادله' ۹. ۶-۳)، عمود منصف پاره خطی است که \mathbf{m}_i و \mathbf{m}_j را به هم وصل می کند (مسئله ۹. ۳. را ببینید). برای $n=2$ ، این عمود منصف یک خط، برای $n=3$ یک صفحه، و برای $n>3$ یک ابر صفحه^۱ است.

مثال: شکل ۹. ۶. دو دسته الگوراکه از نمونه های سوسن شکل ۹. ۲. استخراج شده اند، نشان می دهد. دو دسته سوسن *Setosa* و *Versicolor*، که بترتیب با ω_1 و ω_2 نشان داده می شوند، دارای بردارهای میانگین نمونه ای^۱ (۹. ۵-۳)، توابع

$$\begin{aligned} d_1(\mathbf{x}) &= \mathbf{x}^T \mathbf{m}_1 - \frac{1}{2} \mathbf{m}_1^T \mathbf{m}_1 = 4.3x_1 + 1.3x_2 - 10.1 \\ d_2(\mathbf{x}) &= \mathbf{x}^T \mathbf{m}_2 - \frac{1}{2} \mathbf{m}_2^T \mathbf{m}_2 = 1.5x_1 + 0.3x_2 - 1.17 \end{aligned} \quad \text{تصمیم}$$

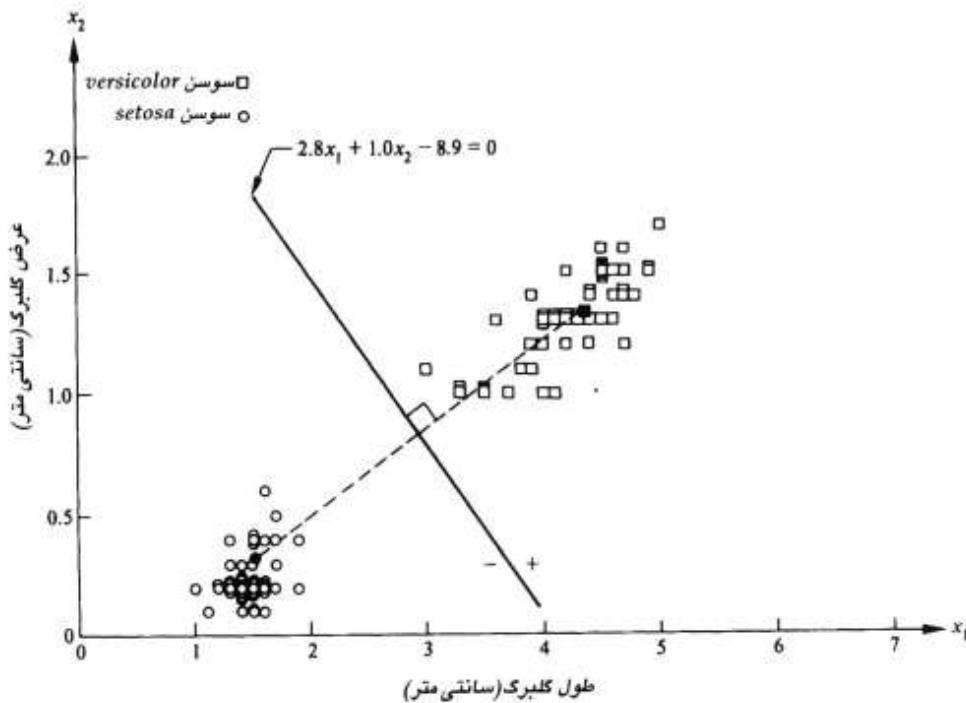
هستند. طبق معادله' ۹. ۶، معادله' مرز عبارتست از

$$\begin{aligned} d_{12}(\mathbf{x}) &= d_1(\mathbf{x}) - d_2(\mathbf{x}) \\ &= 2.8x_1 + 1.0x_2 - 8.9 = 0. \end{aligned}$$

شکل ۹. ۶ نمایشی از این مرز را نشان می دهد (توجه کنید که چون محورها مقیاس یکسانی ندارند، الگوها و مرز درجهت عمودی جای بجا شده اند). جایگذاری هر الگوی دسته' ω_1 ، $d_{12}(\mathbf{x}) > 0$ را نتیجه می دهد. بر عکس، هر الگو از دسته' ω_2 ، $d_{12}(\mathbf{x}) < 0$ را نتیجه می دهد. به عبارت دیگر، با داشتن یک

1) Hyperplane

۳۰. ۹ روش های نظریه 'تصمیم' ۶۸۹

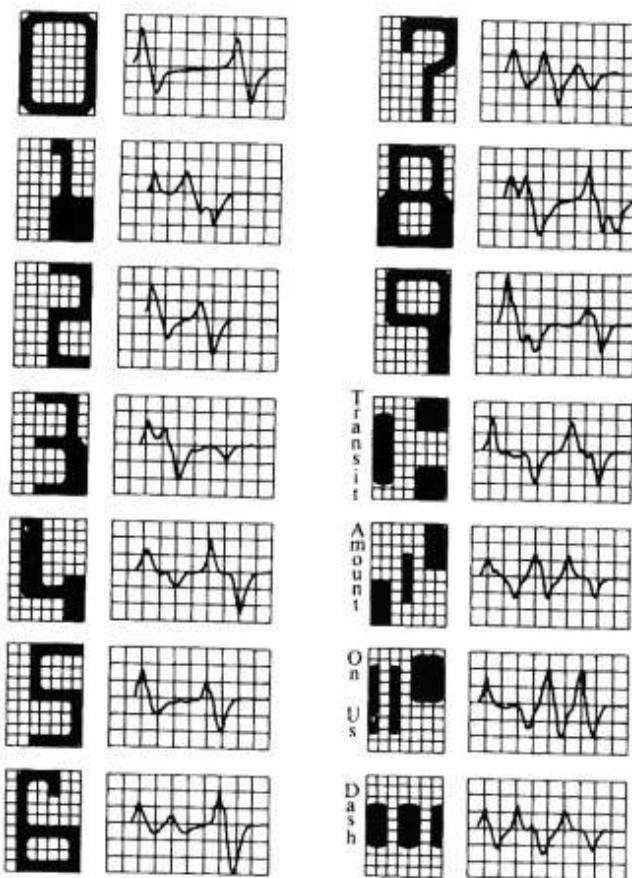


شکل ۹.۶ مرز تصمیم در دسته بندی کننده' حداقل فاصله برای دسته های سوسن versicolor و سوسن setosa

الگو که به یکی از این دو دسته تعلق دارد، علامت $(*)_{12}$ برای تعیین آن الگو کافی است. □

در عمل، دسته بندی کننده' حداقل فاصله وقتی خوب کار می کند که فاصله بین میانگین ها در مقایسه با پراکندگی^۱ یا اتفاقی بودن^۲ هر دسته حول میانگین آن، بزرگ باشد. در بخش ۹.۳.۲ نشان می دهیم که دسته بندی کننده' حداقل فاصله وقتی که توزیع هر دسته حول میانگین آن به شکل "فوق ابر"^۳ کروی در فضای الگوی^۴ بعدی باشد، کارآیی بهینه (از نظر کمینه سازی متوسط خطای ناشی از دسته بندی نادرست) را بدست می دهد.

در عمل بندرت رخ می دهد که به طور همزمان فاصله' میانگین ها بزرگ و پراکندگی دسته ها نسبتاً کوچک باشد، مگر این که طراح سامانه، طبیعت ورودی را کنترل نماید. یک مثال عالی، سامانه های طراحی شده برای خواندن قلم^۵ های تویسه ای سبک دار، نظیر مجموعه قلم تویسه ای E-13B از مؤسسه' بانکداری آمریکا، است. همان طور که شکل ۹.۷ نشان می دهد، این مجموعه' قلم ویژه شامل ۱۴



شکل ۹.۷ مجموعه نویسه های فونت E-13B انجمن بانکداران آمریکا و شکل موج های متناظر آنها

نویسه است که روی یک شبکه 7×9 طراحی شده اند تا خواندن آسان باشد. معمولاً نویسه ها با جوهری که حاوی مواد مغناطیسی کاملاً پودر شده است، چاپ می شوند. قبل از خواندن، جوهر، تحت میدان مغناطیسی قرار می گیرد تا آشکارسازی نویسه ها آسان شود. به عبارت دیگر، مسأله بخش بندی با بر جسته کردن مصنوعی نویسه ها حل می شود.

معمولانویسه ها در جهت افقی با یک هد^۱ تک شیاری و مخصوص عمل خواندن که باریکتر و بلندتر از نویسه ها است، پیمایش می شوند. وقتی که هد از روی نویسه عبور می کند، سیگنال الکترونیکی یک بعدی تولید می گردد. این سیگنال متناسب با آهنگ افزایش یا کاهش ناحیه^۲ جوهری زیر هد می باشد.

۱) Head

مثالاً، شکل موج مربوط به عدد "۰" در شکل ۹. ۷ را در نظر بگیرید. وقتی که هد از چپ به راست حرکت می کند، ناحیه زیر هد شروع به افزایش کرده و در نتیجه یک مشتق مثبت (تفییری با آهنگ مثبت) تولید می گردد. وقتی که هد شروع به ترک پایه سمت چپ "۰" می کند، ناحیه زیر هد شروع به کاهش می کند و در نتیجه یک مشتق منفی تولید می شود. وقتی که هد در ناحیه وسطی نویسه است، ناحیه زیر هد ثابت بوده و مشتق صفر تولید می گردد. وقتی که هد به پایه راست نویسه "۰" می رسد، الگوی خروجی فوق تکرار می شود. این طراحی قلم تضمین می کند که شکل موج هرنویسه متفاوت از سایر نویسه ها باشد. همچنین تضمین می کند که قله ها و صفر های هر شکل موج تقریباً در خطوط عمودی شبکه زمینه که این شکل موج ها روی آن نمایش داده می شوند، رخ دهند، که این نکته در شکل ۹. ۷ دیده می شود.

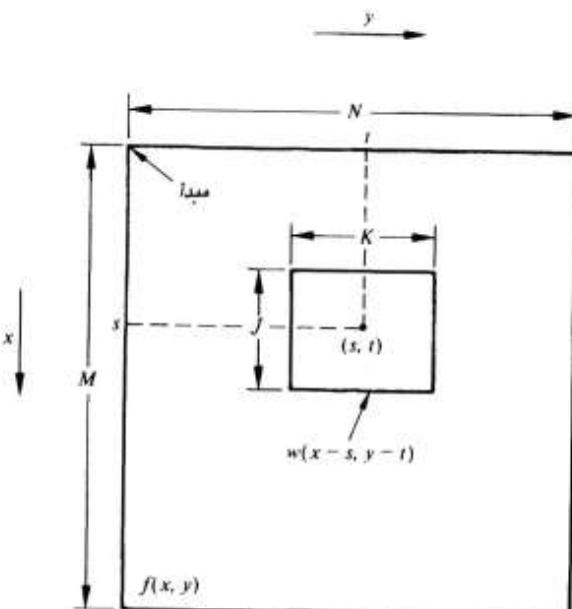
قلم E-13B این خاصیت را دارد که نمونه برداری شکل موج ها تنها در این نقاط، اطلاعات کافی برای دسته بندی صحیح آن ها را به ما می دهد. استفاده از جوهر مغناطیسی به تولید شکل موج های تمیز کمک نموده و بنابراین پخش را کمینه می کند.

طراحی دسته بندی کننده حداقل فاصله برای این کاربرد ساده است. مانند مقادیر نمونه های هر شکل موج را ذخیره کرده و هر مجموعه از نمونه ها را به صورت بردار نمونه $m, 1, 2, \dots, 14 = i$ نمایش می دهیم. هنگامی که بخواهیم نویسه مجھولی را دسته بندی کنیم، راهکار این است که آن را به روشی که قبلًاً توصیف شد، پیمایش کنیم، نمونه های شکل موج آن در نقاط شبکه ای را به صورت بردار X نمایش دهیم و در نهایت با انتخاب دسته متناظر با بردار پیش نمونه ای که بزرگترین مقدار را در معادله $(4. 5-3)$ می دهد، دسته عضویت آن بردار را تعیین کنیم. با مدارهای آنالوگی که دارای بانکهای مقاومتی باشند، می توان به سرعتهای بالای دسته بندی دست یافت (مسئله ۹. ۴ را ببینید).

تطبیق یا همبستگی

در بخش ۳. ۸. ۳ مفهوم پایه ای همبستگی تصویر را معرفی کردیم. اکنون، آن را به عنوان مبنایی برای پیدا کردن تطبیق های زیر تصویر $w(x, y)$ به ابعاد $K \times J$ درون تصویر $f(x, y)$ به ابعاد $M \times N$ و با فرض $M \leq J \leq K$ در نظر می گیریم. گرچه روش همبستگی را می توان به صورت برداری بیان کرد (مسئله ۹. ۵ را ببینید)، کار کردن مستقیم با یک قالب پنجره ای، قابل درک تر (و متداول تر) است.

در ساده ترین شکل، همبستگی بین $w(x, y)$ و $f(x, y)$ عبارتست از



شکل ۸.۹ آرایشی برای به دست آوردن همبستگی $w(x, y)$ و $f(x, y)$ در نقطه (s, t) .

$$c(s, t) = \sum_x \sum_y f(x, y)w(x - s, y - t) \quad (7-3.9)$$

که $t = 0, 1, 2, \dots, N-1$ ، $s = 0, 1, 2, \dots, M-1$ و مجموع، روی ناحیه ای از تصویر که w و f را همپوشانی دارند، گرفته می شود. شکل ۸.۹ این روال را نمایش می دهد. در این شکل مبدأ $f(x, y)$ را در گوشه چپ بالا و مبدأ $w(x, y)$ را در مرکزش فرض می کنیم. برای هر مقدار (s, t) درون w در ناحیه 'تصویر حرکت' می کند تا کاربرد معادله (7-3.9) مقدار c را می دهد. با تغییر s و t در $w(x, y)$ در ناحیه 'تصویر حرکت' می کند تا تابع $c(s, t)$ بدست آید. مقدار بیشینه $c(s, t)$ نشان دهنده ' محلی است که $w(x, y)$ بهترین تطبیق را با $f(x, y)$ دارد. توجه کنید که برای مقادیر s و t که نزدیک لبه های $f(x, y)$ باشند، دقت از دست می رود و مقدار خطای متناسب با ابعاد $w(x, y)$ می باشد.

تابع همبستگی داده شده در معادله (7-3.9) این عیب را دارد که به تغییرات در دامنه $f(x, y)$ و $w(x, y)$ حساس است. مثلاً دو برابر کردن تمام مقادیر $f(x, y)$ ، مقدار $c(s, t)$ را نیز دو برابر می کند. روشی که مکرر برای غلبه بر این مشکل استفاده می شود، اجرای تطبیق با ضریب همبستگی¹ است. این

1) Correlation coefficient

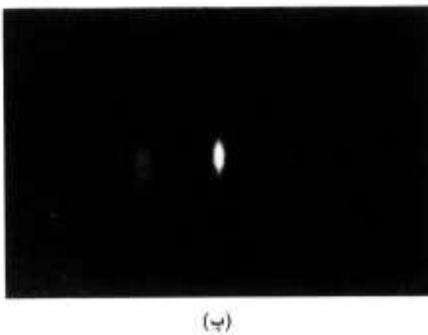
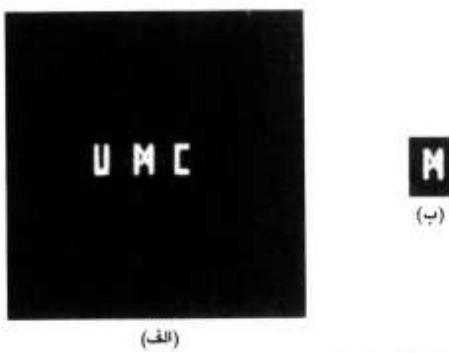
۶۹۳ ۴. ۳. روش های نظریه 'تصمیم'

ضریب به صورت

$$\gamma(s, t) = \frac{\sum_i \sum_y [f(x, y) - \bar{f}(x, y)][w(x - s, y - t) - \bar{w}]}{\left[\sum_i \sum_y [f(x, y) - \bar{f}(x, y)]^2 \sum_i \sum_y [w(x - s, y - t) - \bar{w}]^2 \right]^{1/2}} \quad (8-3, 9)$$

تعريف می شود که در آن $1 \leq t \leq M-1$ ، $0 \leq s \leq N-1$ ، $0 \leq y \leq 2$ ، \bar{w} مقدار متوسط پیکسل های $w(x, y)$ (که تنها یک بار محاسبه می شود)، و $\bar{f}(x, y)$ مقدار متوسط $f(x, y)$ در ناحیه 'مشترک' با محل فعلی w است، و مجموع های زیر ناحیه 'مشترک' بین f و w گرفته می شوند. ضریب همبستگی $\gamma(s, t)$ مستقل از محدوده 'تفاوت دامنه' $w(x, y)$ و $f(x, y)$ است و به محدوده '۱-تا ۱' مقیاس می شود (مسئله ۹.۵ را بینید).

مثال: شکل ۹.۹. ۹. مفاهیمی را که اخیراً تشریح شدند، به تصویر می کشد. شکل ۹.۹. (الف) $w(x, y)$ و شکل ۹.۹. (ب) $w(x, y)$ را نمایش می دهد. ضریب همبستگی $\gamma(s, t)$ به صورت تصویری در شکل



شکل ۹.۹. مثالی از همبستگی. به روشنایی $\gamma(s, t)$ در محلی که دو حرف تطبیق می کنند، توجه کنید.
(برگرفته از Hall و همکارانش [1971]).

۹. ۹. (ب) نشان داده می شود، به مقدار (شدت) بالاتر ($1/s^2$) در محلی که بهترین تطبیق بین $f(x, y)$ و (x, y) وجود دارد، توجه کنید.



گرچه از طریق ضربه همبستگی می توانتابع همبستگی را برای تغییرات دامنه^۱ تراز کرد، ترازسازی برای تغییرات اندازه^۲ و چرخش مشکل تر است. ترازسازی^۳ برای اندازه شامل مقیاس کردن مکانی است، اگر فرآیندی که به تهابی به محاسبات زیادی نیاز دارد. ترازسازی برای چرخش از آن هم مشکل تر است. اگر بتوان از روی (x, y) f قاعده ای برای چرخش آن استخراج کرد، آنگاه بسادگی می توان (y, x) ^۴ را طوری چرخاند که با (x, y) f هم راستا شود. در هر حال، اگر طبیعت چرخش نامعلوم باشد، جستجو برای بهترین تطبیق نیازمند چرخش های بی شمار (x, y) ^۵ است. این روال غیرعملی است، و در نتیجه، در مواردی که چرخش دلخواه یا نامقید است، بندرت از همبستگی استفاده می شود.

در بخش ۳. ۸. بیان شد که در حوزه 'فرکانس نیز می توان از طریق FFT همبستگی را انجام داد. اگر f و w هم اندازه باشند، این روش ممکن است از پیاده سازی مستقیم همبستگی در حوزه 'مکان کارآمدتر باشد. اگر معادله^۶ (۷-۳) استفاده شود، معمولاً^۷ بسیار کوچکتر از f است. تخمین مبتنی بر مصالحه انجام شده توسط Campbell[1969] بیان می کند که اگر تعداد عبارات غیر صفر در w کمتر از ۱۳۲ (زیر تصویری با ابعاد تقریباً 13×13 پیکسل) باشد، پیاده سازی مستقیم معادله^۶ (۷-۳) با FFT کارآمدتر است. البته این تعداد به ماشین و الگوریتم های استفاده شده بستگی دارد، اما روش FFT به طور محسوسی مشکل تر است و معمولاً مستقیماً از معادله^۶ (۸-۳) محاسبه حوزه 'فرکانس به طور محسوسی مشکل تر است. می شود.

۴. ۳. ۹ دسته بندی گشته های آماری بهینه

مبانی

در این بخش روشی احتمالاتی برای تشخیص ارائه می کنیم. همانند اغلب مباحثی که به اندازه گیری و تعبیر پدیده های فیزیکی می پردازند، بدلیل این که معمولاً دسته های الگو به طور اتفاقی تولید می شوند، ملاحظات احتمالاتی در تشخیص الگو مهم هستند. همان طور که در توضیحات بعدی دیده می شود،

۶۹۵ ۳. روش های نظریه 'تصمیم'

می توان راهی برای دسته بندی بدست آورده که در مفهوم پایین ترین احتمال متوسط خطاهای دسته بندی، بهینه باشد.

احتمال این که الگوی ویژه x از دسته ω_i باشد، با نماد $p(\omega_i/x)$ نشان داده می شود. اگر دسته بندی کننده الگو نتیجه بگیرد که x از دسته ω_i است، اما در واقع از دسته ω_j باشد، یک تلف^۱ ایجاد می شود که با نماد L_{ij} نشان داده می شود. چون الگوی x می تواند به یکی از M دسته' تحت بررسی متعلق باشد، متوسط تلف ناشی از انتساب x به دسته ω_i عبارتست از

$$r_i(x) = \sum_{k=1}^M L_{ik} p(\omega_k/x). \quad (9-3. 9)$$

معادله' (9-3. 9) در زبان نظریه تصمیم غالباً خطای متوسط شرطی^۲ یا تلف نامیده می شود. از نظریه احتمال پایه ای داریم که $(b/a) = [p(a)p(b/a)]/p(b)$. با استفاده از این عبارت می توان معادله' (9-3. 9) را به شکل

$$r_i(x) = \frac{1}{p(x)} \sum_{k=1}^M L_{ik} p(x/\omega_k) P(\omega_k) \quad (10-3. 9)$$

نوشت که $(x/\omega_k) p$ تابع چگالی احتمال الگوهای دسته ω_k است و $P(\omega_k)$ نیز احتمال 'خداد دسته ω_k ' است. چون $(x/\omega_k) p$ مثبت و برای تمام $(x/\omega_1), (x/\omega_2), \dots, (x/\omega_M)$ ها، $j = 1, 2, \dots, M$ مشترک است، می توان آن را از معادله' (10-3. 9) حذف کرد. بدون این که بر ترتیب نسبی این توابع از کمترین مقدار تا بزرگترین مقدار اثری بگذارد. آنگاه رابطه' تلف متوسط به صورت زیر ساده می شود:

$$r_i(x) = \sum_{k=1}^M L_{ik} p(x/\omega_k) P(\omega_k). \quad (11-3. 9)$$

دسته بندی کننده می تواند برای هر الگوی مجهول داده شده یکی از M دسته ممکن را انتخاب کند. اگر دسته بندی کننده برای هر الگوی x ، $r_1(x), r_2(x), \dots, r_M(x)$ را محاسبه و الگو را به دسته' دارای کمترین تلف متساب نماید، تلف متوسط کل نسبت به تمام تصمیم ها کمینه خواهد بود. دسته بندی کننده ای که تلف متوسط کل را کمینه می کند، دسته بندی کننده' بیز^۳ خوانده می شود. بنابراین دسته بندی کننده' بیز الگوی مجهول x را به دسته ω_i منتب می کند اگر برای $i = 1, 2, \dots, M$ که $i \neq j$

1) Loss

2) Conditional average risk

3) Bayes classifier

باشد. به عبارت دیگر، x به دسته ω متنسب می شود اگر $r_i(x) < r_j(x)$

$$\sum_{k=1}^M L_{ki} p(x/\omega_k) P(\omega_k) < \sum_{q=1}^M L_{qj} p(x/\omega_q) P(\omega_q). \quad (12-3. 9)$$

در بسیاری از مسائل تشخیص، تلف تصمیم صحیح، صفر است، و تمام تصمیم های ناصحیح، تلف غیر صفر اما یکسانی (مثلًا ۱) دارند. در این حالات، تابع تلف به صورت

$$L_{ij} = 1 - \delta_{ij} \quad (13-3. 9)$$

در می آید که اگر $j = 1, i = 1$ و $\delta_{ij} = 0$ ، $i \neq j$ است. معادله (۱۳-۳. ۹) بیانگر تلف ۱ برای تصمیم های ناصحیح و تلف صفر برای تصمیم های صحیح است. با جایگذاری معادله (۱۳-۳. ۹) در معادله (۱۱-۳. ۹) داریم:

$$\begin{aligned} r_i(x) &= \sum_{k=1}^M (1 - \delta_{ki}) p(x/\omega_k) P(\omega_k) \\ &= p(x) - p(x/\omega_i) P(\omega_i). \end{aligned} \quad (14-3. 9)$$

آنگاه دسته بندی کننده 'بیز الگوی x ' را به دسته ω متنسب می کند اگر

$$p(x) - p(x/\omega_i) P(\omega_i) < p(x) - p(x/\omega_j) P(\omega_j) \quad (15-3. 9)$$

یا به طور معادل، اگر

$$p(x/\omega_i) P(\omega_i) > p(x/\omega_j) P(\omega_j) \quad j = 1, 2, \dots, M; j \neq i. \quad (16-3. 9)$$

با مراجعه به شرح متنه به معادله (۹. ۱-۳)، می بینیم که دسته بندی کننده 'بیز برای توابع تلف ۰-۱' چیزی بیش از پیاده سازی توابع تصمیم به شکل

$$d_i(x) = p(x/\omega_i) P(\omega_i) \quad j = 1, 2, \dots, M \quad (17-3. 9)$$

نیست که در آنها بردار الگوی x به دسته ω متنسب می شود اگر برای تمام $j \neq i$ ، $d_i(x) > d_j(x)$ باشد. این نظر که متوسط تلف دسته بندی نادرست را کمینه

۴. روش های نظریه 'تصمیم' ۶۹۷

می کنند، بهینه هستند. در هر حال برای برقراری این بهینگی، باید توابع چگالی احتمال الگوهای هر دسته، همچنین احتمال رخداد هر دسته، معلوم باشد. معمولاً برآورده شدن شرط دوم مشکل نیست. برای نمونه، اگر تمام دسته ها هم احتمال باشند، آنگاه $\frac{1}{M} = P(\omega_i)$. حتی اگر این رابطه درست نباشد، معمولاً می توان با استفاده از دانش موجود در مورد هر مسئله از این احتمالات مطلع شد. تخمین توابع چگالی احتمال $(x/\omega_i)^p$ مطلب دیگری است. اگر بردارهای الگو، x ها، n بعدی باشند، $(x/\omega_i)^p$ یک تابع n متغیره است، که اگر شکلش معلوم نباشد، برای تخمین آن به روشهای نظریه 'احتمال چند متغیره' نیاز داریم. اجرای این روش ها مشکل است، بویژه اگر تعداد الگوهای نماینده هر دسته بزرگ نباشد یا اگر شکل توابع چگالی احتمال خوش رفتار^۱ نباشد. به این دلایل، استفاده از دسته بندی کننده بیز معمولاً مبتنی بر فرض وجود عبارتی تحلیلی برای توابع چگالی گوناگون و سپس تخمین عوامل آن عبارت از الگوهای نمونه هر دسته است. تاکنون شایع ترین شکل مفروش برای $(x/\omega_i)^p$ ، تابع چگالی احتمال گوسی است. هرچه این فرض به واقعیت نزدیکتر باشد، دسته بندی کننده بیز به تلف متوسط کمینه، نزدیکتر می شود.

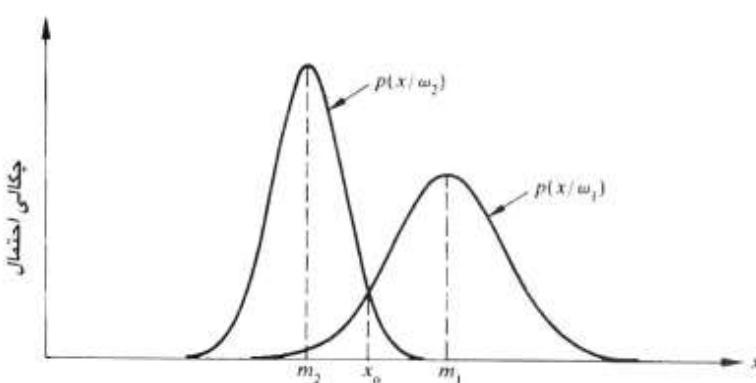
دسته بندی کننده بیز برای دسته های الگوی گوسی

برای شروع اجازه دهدید مسئله ای یک بعدی ($n = 1$) شامل دو دسته الگو ($m = 2$) که به ترتیب دارای چگالی های گوسی با میانگین های m_1 و m_2 و انحراف معیارهای σ_1 و σ_2 می باشند، را در نظر بگیریم. طبق معادله^۲ (۳-۱۷) تابع تصمیم بیز به شکل

$$d_j(x) = p(x/\omega_j)P(\omega_j) \quad (18-3.9) \\ = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{(x - m_j)^2}{2\sigma_j^2}\right]P(\omega_j) \quad j = 1, 2$$

باشند که اکنون الگوها اعدادی هستند که با نماد x نشان داده می شوند. شکل ۹. ۱۰ نموداری از تابع چگالی احتمال این دو دسته را نشان می دهد. مرز بین دو دسته، که با نماد x نشان داده می شود، نقطه ای است که در آن $d_1(x_0) = d_2(x_0)$. اگر دو دسته هم احتمال باشند، آنگاه $\frac{1}{2} = P(\omega_1) = P(\omega_2)$ ، و مرز تصمیم مقداری از x است که $p(x_0/\omega_1) = p(x_0/\omega_2) = p(x_0/\omega_1)$. همان طور که در شکل ۹. ۱۰ نیز دیده

¹⁾ Well-behaved



شکل ۹.۱۰.۹ توابع چگالی احتمال برای دو دسته الگوی یک بعدی. اگر احتمال وقوع دو دسته برابر باشد، نقطه مرز تصمیم است.

می شود، این نقطه محل تلاقی دو تابع چگالی احتمال است. هر الگو (نقطه) که در سمت راست x_0 باشد، به عنوان عضو دسته w_1 دسته بندی می شود. به طور مشابه، هر الگو که در سمت چپ x_0 باشد، به عنوان عضو دسته w_2 دسته بندی می شود. وقتی دسته ها هم احتمال نباشند، اگر احتمال رخداد دسته w_1 بیشتر باشد، x_0 به سمت چپ جابجا می شود، و بر عکس اگر احتمال رخداد دسته w_2 بیشتر باشد، x_0 به سمت راست جایه جا می شود. این نتیجه مورد انتظار است، زیرا دسته بندی کننده می خواهد تلف دسته بندی نادرست را کمینه نماید. برای نمونه، در حالت حدی که دسته w_1 اصلاً رخ ندهد، دسته بندی کننده همواره الگوهای را به دسته w_2 منتهی کند و هرگز اشتباه نمی کند (یعنی، x_0 به بی نهایت منفی منتقل می شود).

در حالت w_1 بعدی، چگالی گوسی بردارهای دسته الگوی زام به شکل

$$p(\mathbf{x}|w_i) = \frac{1}{(2\pi)^{n/2} |\mathbf{C}_i|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mathbf{m}_i)^T \mathbf{C}_i^{-1} (\mathbf{x} - \mathbf{m}_i) \right] \quad (19-3.9)$$

است و هر چگالی به طور کامل با بردار میانگین \mathbf{m}_i و ماتریس کوواریانس \mathbf{C}_i خود مشخص می شود:

$$\mathbf{m}_i = E_i[\mathbf{x}] \quad (20-3.9)$$

۶۹۹ ۳. ۹ روش های نظریه 'تصمیم

$$\mathbf{C}_i = E_i\{(\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T\} \quad (21-3. 9)$$

که $\{E_i\}$ برای نشان دادن مقدار امید آرگومان روی الگوهای دسته ω به کار می رود. در معادله ۹. ۱۹-۳، n تعداد ابعاد بردارهای الگو، و $|\mathbf{C}_i|$ دترمینان ماتریس \mathbf{C}_i است. تقریب زدن مقدار امید E_i با مقدار متوسط کمیات مورد نظر، تخمینی از بردار میانگین و ماتریس کواریانس^۱ را می دهد:

$$\mathbf{m}_i = \frac{1}{N_i} \sum_{\mathbf{x} \in \omega_i} \mathbf{x} \quad (22-3. 9)$$

و

$$\mathbf{C}_i = \frac{1}{N_i} \sum_{\mathbf{x} \in \omega_i} \mathbf{x}\mathbf{x}^T - \mathbf{m}_i\mathbf{m}_i^T \quad (23-3. 9)$$

که N_i تعداد بردارهای الگو در دسته ω_i است، و مجموع، روی این بردارها گرفته می شود. در ادامه این بخش مثالی از چگونگی استفاده از این دو عبارت ارائه می دهیم.

ماتریس کواریانس، متقارن و نیمه معین مثبت^۲ است. عنصر قطری c_{kk} ، واریانس عنصر k بردارهای الگو است. عنصر غیر قطری c_{jk} ، کواریانس x_j و x_k است. وقتی عناصر x_j و x_k مستقل آماری باشند، $c_{jk} = 0$ وقتی عناصر غیر قطری ماتریس کواریانس صفر باشند، تابع چگالی گوسی چند متغیره به حاصل ضرب تابع های چگالی گوسی یک متغیره هر یک از عناصر \mathbf{x} ساده می شود. طبق معادله ۹. ۱۷-۳، تابع تصمیم بیز دسته ω_i به شکل $p(\mathbf{x}/\omega_i)P(\omega_i) = p(\mathbf{x})d_i(\mathbf{x})$ می باشد. با این حال به دلیل شکل نمایی تابع چگالی گوسی، کار با لگاریتم طبیعی این تابع تصمیم، سهل تر است. به عبارت دیگر، می توانیم از شکل لگاریتمی

$$d_i(\mathbf{x}) = \ln[p(\mathbf{x}/\omega_i)P(\omega_i)] \quad (24-3. 9)$$

$$= \ln p(\mathbf{x}/\omega_i) + \ln P(\omega_i).$$

استفاده کنیم. از نظر کارایی دسته بندی، این عبارت معادل معادله ۹. ۱۷-۳ است زیرا لگاریتم، یک تابع یکنواهی صعودی است. به بیان دیگر، ترتیب^۳ عددی تابع تصمیم در معادلات ۹. ۱۷-۳ و ۹. ۲۴-۳ یکسان است. با جایگذاری معادله ۹. ۱۹-۳ در معادله ۹. ۲۴-۳ داریم:

$$d_i(\mathbf{x}) = \ln P(\omega_i) - \frac{n}{2} \ln 2\pi - \frac{1}{2} \ln |\mathbf{C}_i| - \frac{1}{2} [(\mathbf{x} - \mathbf{m}_i)^T \mathbf{C}_i^{-1} (\mathbf{x} - \mathbf{m}_i)]. \quad (25-3. 9)$$

1) Covariance

2) Positive semidefinite

3) Order

عبارت $\ln 2\pi n/2$ برای تمام دسته های یکسان است، بنابراین می توان آنها را از معادله^۹ (۲۵-۳) حذف کرد، که در این صورت، برای $M, \dots, j = 1, 2, \dots$ نتیجه می شود:

$$d_i(\mathbf{x}) = \ln P(\omega_i) - \frac{1}{2} \ln |\mathbf{C}_i| - \frac{1}{2}[(\mathbf{x} - \mathbf{m}_i)^T \mathbf{C}_i^{-1} (\mathbf{x} - \mathbf{m}_i)] \quad (26-3. 9)$$

معادله^۹ (۲۶-۳) توابع تصمیم بیز برای دسته های الگوی گوسی را با شرط تابع تلف ۱-۰ ارائه می دهد.

توابع تصمیم داده شده با معادله^۹ (۲۶-۳) «آبر مربعی»^۱ (توابع مربعی در فضای n بعدی) هستند، زیرا در این معادله هیچ عبارتی با درجه 'بالاتر از دو' بر حسب مؤلفه های \mathbf{x} ظاهر نمی شود. آنگاه، به روشنی، بهترین کاری که دسته بندی کننده بیز می تواند برای الگوهای گوسی انجام دهد، قرار دادن یک سطح تصمیم مرتبه دوم سراسری بین هر زوج از دسته های الگو است. اگر جامعه های الگو کاملاً گوسی باشند، هیچ سطح دیگری نمی تواند تلف متوسط کمتری را در دسته بندی نتیجه دهد.

اگر تمام ماتریسهای کوواریانس برابر باشند، برای تمام $M, \dots, j = 1, 2, \dots$ $\mathbf{C}_j = \mathbf{C}$ می شود و با حذف تمام عبارات مستقل از j معادله^۹ (۲۶-۳) به صورت

$$d_i(\mathbf{x}) = \ln P(\omega_i) + \mathbf{x}^T \mathbf{C}^{-1} \mathbf{m}_i - \frac{1}{2} \mathbf{m}_i^T \mathbf{C}^{-1} \mathbf{m}_i \quad (27-3. 9)$$

در می آید که توابع تصمیم خطی برای $M, \dots, j = 1, 2, \dots$ هستند. به علاوه اگر $\mathbf{C} = \mathbf{I}$ ، که \mathbf{I} ماتریس همانی است، و $P(\omega_i) = 1/M$ ، برای $M, \dots, j = 1, 2, \dots, M$ باشد، آنگاه

$$d_i(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_i - \frac{1}{2} \mathbf{m}_i^T \mathbf{m}_i \quad j = 1, 2, \dots, M. \quad (28-3. 9)$$

معادله^۹ (۲۸-۳) توابع تصمیم برای دسته بندی کننده حداقل فاصله، آن طور که در معادله^۹ (۵-۳) داده شد، را معین می کند. بنابراین اگر (۱) دسته های الگو گوسی باشند، (۲) تمام ماتریس های کوواریانس برابر ماتریس همانی باشند، و (۳) تمام دسته ها احتمال رخداد یکسانی داشته باشند، دسته بندی کننده حداقل فاصله در معیار بیز بهینه است. دسته های الگوی گوسی که این شرایط را برآورده

¹⁾ Hyperquadric

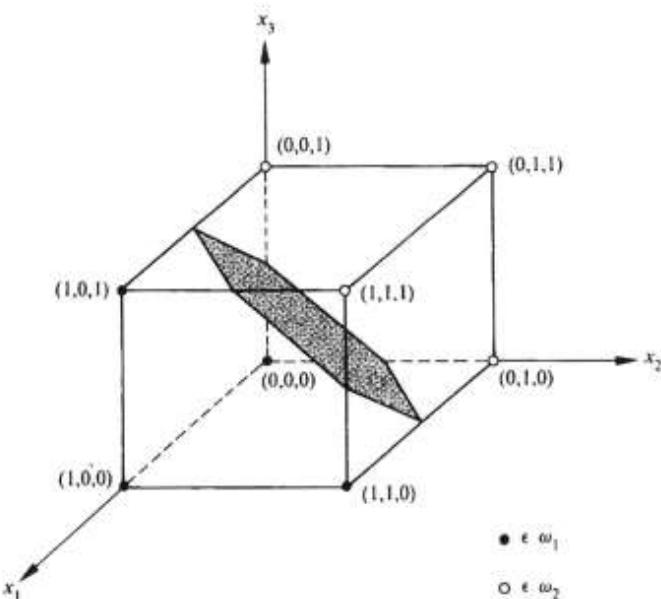
۴. ۳. روش های نظریه / تصمیم ۷۰۱

می کنند، ابرهای کروی هستند که در \mathbb{R}^n بعد شکل یکسانی دارند (و آنکه خوانده می شوند). دسته بندی کننده حداقل فاصله بین هر زوج دسته، یک ابرصفحه قرار می دهد، که این ابرصفحه، عمود منصف پاره خطی است که مراکز دو کره را به هم متصل می کند. در حالت دو بعدی، دسته ها نواحی دایروی هستند، و مرزها به صورت خطوطی می باشند که پاره خط های واصل مراکز هر یک از زوج دایره ها را دو قسمت می کنند.

مثال: شکل ۹ . ۱۱ آرایش ساده ای از دو دسته الگو در فضای سه بعدی را نشان می دهد. با فرض این که الگوهای هر دسته نمونه هایی از یک توزیع گوسی هستند، با استفاده از این الگوها نحوه اجرای دسته بندی کننده بیز را نمایش می دهیم.

با اعمال معادله (۹ . ۲۲-۳) به الگوهای شکل ۹ . ۱۱ داریم که

$$\mathbf{m}_1 = \frac{1}{4} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{m}_2 = \frac{1}{4} \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}$$



شکل ۱۱ . ۹ دو دسته الگوی ساده و مرز تصمیم بیز بین آنها .

به طور مشابه، اعمال معادله^۹ $(۲۳-۳)$ به دسته الگو، دو ماتریس کوواریانس می دهد که در این مورد با هم برابرند:

$$\mathbf{C}_1 = \mathbf{C}_2 = \frac{1}{16} \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}$$

چون ماتریس های کوواریانس برابر هستند، توابع تصمیم بیز با معادله^۹ $(۲۷-۳)$ داده می شوند. اگر فرض کنیم که $\ln P(\omega_i) = P(\omega_i) = 1/2$ توان حذف کرد، و در نتیجه داریم:

$$d_i(\mathbf{x}) = \mathbf{x}' \mathbf{C}^{-1} \mathbf{m}_i - \frac{1}{2} \mathbf{m}_i' \mathbf{C}^{-1} \mathbf{m}_i$$

که در آن

$$\mathbf{C}^{-1} = \begin{bmatrix} 8 & -4 & -4 \\ -4 & 8 & 4 \\ -4 & 4 & 8 \end{bmatrix}$$

با بسط $d_i(\mathbf{x})$ توابع تصمیم بدست می آیند:

$$d_1(\mathbf{x}) = 4x_1 - 1.5 \quad d_2(\mathbf{x}) = -4x_1 + 8x_2 + 8x_3 - 5.5.$$

آنگاه سطح تصمیمی که این دو دسته را جدا می کند، عبارتست از:

$$d_1(\mathbf{x}) - d_2(\mathbf{x}) = 8x_1 - 8x_2 - 8x_3 + 4 = 0.$$

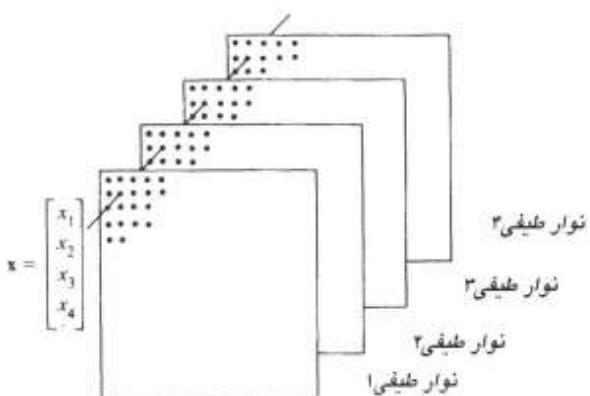
شکل ۹ . ۱۱ بخشی از این سطح را نشان می دهد، که مشاهده می شود دسته ها بخوبی از هم جدا شده اند.

یکی از موفق ترین کاربردهای روش دسته بندی بیز، دسته بندی تصاویر سنجش از راه دور گرفته شده از پیماشگرهای چند طیفی روی سفینه ها، ماهواره ها، و ایستگاه های فضایی می باشد. داده های تصویری فراوان تولید شده با این منابع، علاقه زیادی را به دسته بندی و تحلیل خودکار تصویر در مبحث سنجش از راه دور، ایجاد می کنند. کاربردهای سنجش از راه دور متفاوت هستند و شامل کاربردهای زمینی،

۴۰۴ ۳. روش های نظریه ^۱ تصمیم

صورت برداری محصولات کشاورزی، تشخیص آفات محصولات کشاورزی، جنگلداری، بررسی کیفیت آب و هوا، مطالعات زمین شناسی، پیش بینی هوا، و گروهی دیگر از کاربردها که اهمیت زیست محیطی دارند، می شود. مثال زیر یک کاربرد نوعی را نشان می دهد.

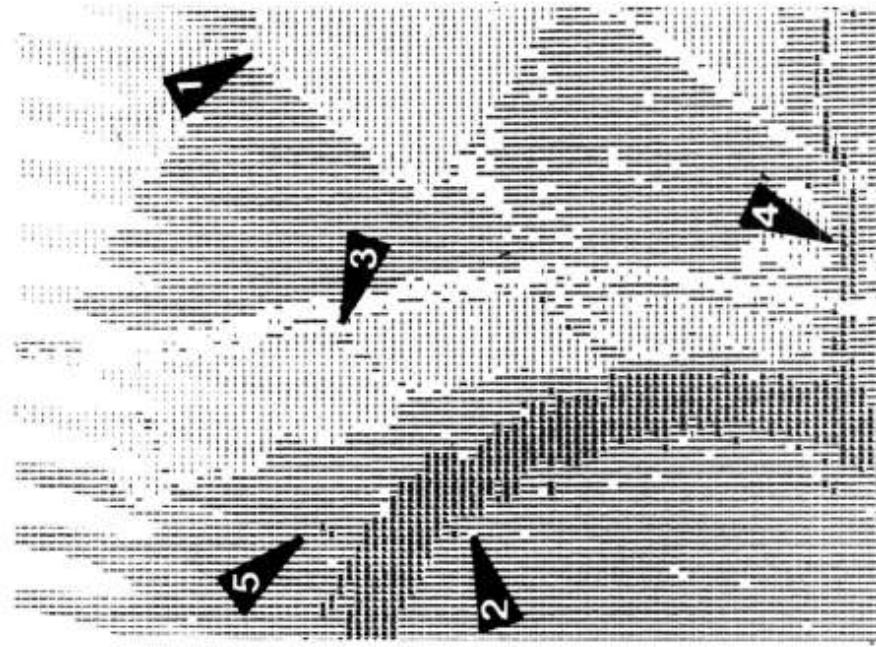
مثال: یک پیماشگر چند طبقی در چند محدوده مشخص طول موج به نور پاسخ می دهد؛ مثلاً، ۰.۴۰-۰.۴۴، ۰.۵۸-۰.۶۲، ۰.۶۶-۰.۷۲، ۰.۸۰-۱.۰۰ میکرون (m^{-6}). این محدوده ها به ترتیب در نوارهای بخش، سبز، قرمز، و مادون قرمز هستند. با پیماش هر ناحیه از زمین، پیماشگر چهار تصویر رقیم، برای هر نوار رنگی یک تصویر، تولید می کند. اگر این تصاویر کاملاً ذخیره شوند، حالتی که عموماً در عمل وجود دارد، می توان آنها را به صورت پشت سرهم، مانند شکل ۱۲.۹، تصور کرد. بنابراین می توان هر نقطه روی زمین را با یک بردار الگوی ۴ عنصری به شکل ^۱ $x = (x_1, x_2, x_3, x_4)$ نمایش داد، که x_1 نماینده 'بخش' است، x_2 نماینده 'سبز' است، و الی آخر. اگر تصاویر با تفکیک ۵۱۲ × ۵۱۲ باشند، هر پشتہ از این چهار تصویر چند طبقی را می توان با ۲۶۲۱۴۴ بردار الگو نمایش داد. در دسته بندی کننده 'بیز گوسی' نیازمند تخمین بردار میانگین و ماتریس کوواریانس هر دسته است. در کاربردهای سنجش از راه دور که بسیار شبیه مثال قبل هستند، این تخمین ها با جمع آوری داده های چند طبقی هر منطقه مورد علاقه و سپس استفاده از این نمونه ها به دست می آیند. شکل ۱۲.۹(الف) تصویری نوعی را که از راه دور و از درون یک سفینه برداشته شده است، نشان می دهد (این



شکل ۱۲.۹ تشكیل بردار الگو از پیکسل های چهار تصویر رقیم که با پیماشگر چند طبقی تولید شده اند.



(الف)



(ب)

شکل ۱۳.۹ (الف) تصویر چند خلعی؛ (ب) نتایج چاپ شده دسته بندی ماشینی با استفاده از دسته بندی کننده بیز. (با مجوز از آزمایشگاه کاربردهای سنجش از دور، دانشگاه Purdue)

۳۰.۹ روش های نظریه / تصمیم ۷۰۵

تصویر نسخه 'تک رنگی از تصویر اولیه' چند طیفی است). در این مورد خاص، هدف دسته بندی نواحی مختلف، نظیر گیاهان، آبها، و زمین های باир است. شکل ۹ . ۱۳(ب) نتایج دسته بندی رایانه ای با استفاده از دسته بندی کننده 'بیز گوسی را به شکل چاپ رایانه ای نمایش می دهد. پیکان ها به موارد مورد علاقه اشاره می کنند. پیکان ۱ به گوشه ای از یک مزرعه و پیکان ۲ به یک رودخانه اشاره می کند. پیکان ۳ به یک بیشه 'کوچک بین دو ناحیه' بایر و پیکان ۴ به یک رود فرعی اشاره می کند که بدرستی به وسیله 'سامانه شناسایی شد. پیکان ۵ به استخر کوچکی که در شکل ۹ . ۱۳(الف) تقریباً غیرقابل تشخیص است، اشاره می کند. با مقایسه 'تصویر اولیه و خروجی رایانه، نتایج تشخیص آشکار می شوند. این نتایج به نتایجی که انسان با مشاهده بدست می آورد، بسیار نزدیک هستند.



قبل از شروع بخش بعدی، جالب است ذکر کنیم که دسته بندی پیکسل به پیکسل تصویر آنطور که در مثال قبلی توصیف شد، تصویر را به دسته های گوناگونی بخش بندی می کند. این روش شباهتی با روش آستانه گیری با چند متغیر، آنطور که در بخش ۳ . ۶ به اختصار تشریح شد، دارد.

۳۰.۹ شبکه های عصبی

زمینه

روش های مسروچ در دو بخش قبلی مبتنی بر استفاده از الگوهای نمونه برای تخمین عوامل آماری معینی از هر دسته الگو هستند. دسته بندی کننده 'حداقل فاصله' به طور کامل با بردار میانگین هر دسته مشخص می شود. به طور مشابه، دسته بندی کننده 'بیز برای جامعه های گوسی' به طور کامل با بردار میانگین و ماتریس کوواریانس هر دسته مشخص می شود. الگوهایی (با دسته عضویت^۱ معلوم) که برای تخمین این عوامل استفاده شدند، معمولاً الگوهای آموزش^۲ خوانده می شوند، و مجموعه 'چنین الگوهایی از هر یک از دسته ها، مجموعه 'آموزشی^۳ خوانده می شود. فرآیندی که در آن با استفاده از مجموعه 'آموزشی، توابع تصمیم بدست می آیند، یادگیری^۴ یا آموزش^۵ خوانده می شود.

در دوره های افتخاری که اخیراً تشریح شدند، آموزش کار ساده ای است. الگوهای آموزش هر دسته به روش ساده ای برای محاسبه 'عوامل تابع تصمیم متناظر با آن دسته' به کار می روند. بعد از این که عوامل مورد نظر

1) Class membership

2) Training Pattern

3) Training set

4) Learning

5) Training

تخمین زده شدند، ساختار دسته بندی کننده ثابت می شود، و کارآئی نهایی آن به میزان برآورده شدن شرایط آماری که در به دست آوردن روش دسته بندی فرض شده اند، بستگی دارد.

خواص آماری دسته های الگو در یک مسئله غالباً نامعلوم هستند یا نمی توان آنها را تخمین زد (شرح کوتاه در بخش قبلی راجع به مشکل بودن کار با آمارگان چند متغیره را به خاطر آورید). در عمل بهترین راه پرداختن به چنین مسائل تصمیم-نظریه ای، روشهایی هستند که از طریق آموزش مستقیماً توابع تصمیم مورد نیاز را بدست می آورند. با این شرایط انجام فرض هایی راجع به شکل کلی توابع چگالی احتمال، یا سایر اطلاعات احتمالاتی درباره' دسته های الگوی مورد بررسی لازم نیست. در این بخش رهیافت های گوناگونی که این شرایط را برآورده می کنند، شرح می دهیم.

اساس مطلبی که در ادامه می آید، استفاده از تعداد زیادی عناصر محاسباتی غیرخطی (که نورون^۱ خوانده می شوند) می باشد. این عناصر به صورت شبکه هایی که مجسم کننده' روش اتصال سلول های عصبی در مغز هستند، سازمان یافته اند. مدل های حاصل بر اساس های گوناگونی شامل شبکه های عصبی^۲، رایانه های عصبی^۳، مدل های پردازش موازی توزیع شده^۴(PDP)، سامانه های نورونی^۵، شبکه های وفقی لایه ای^۶، و مدل های اتصالی^۷ مورد استفاده قرار می گیرند. در اینجا از نام شبکه های عصبی^۸ استفاده می کنیم. همچنین از این شبکه ها به عنوان وسایلی برای بهبود وفقی ضرائب توابع تصمیم از طریق عرضه' متوالی مجموعه الگوهای آموزشی به آنها استفاده خواهد شد.

توجه به شبکه های عصبی به ابتدای دهه ۱۹۴۰ میلادی باز می گردد، که یکی از نمونه های آن، کار McCulloch & Pitts[1943] است. این دو نفر مدل های نورونی را به شکل وسایل آستانه ای دودویی و الگوریتم های اتفاقی با تغییر حالت های ناگهانی ۰-۱ و ۱-۰ در نورونها، به عنوان پایه ای برای مدل سازی سامانه های عصبی پیشنهاد کردند. کار بعدی [1949] Hebb بر پایه مدل های ریاضی بود که طی آن تلاش کرد یادگیری با استفاده از خواص تقویتی یا انجمانی انجام گیرد.

در فاصله' نیمه' دهه' ۱۹۵۰ تا ابتدای دهه' ۱۹۶۰ میلادی، دسته ای از ماشین های یادگیرنده^۹ که توسط Rosenblatt[1957, 1962] ابداع شد، تحرک زیادی در محققان و کاربران نظریه' تشخیص الگو ایجاد

1) Neuron 2) Neural network 3) Neuro-computer 4) Parallel distributed processing models

5) Neuromorphic network 6) Layered self-adaptive network 7) Connectionist model

8) Neural net (or neural network) 9) Learning machines

کرد. علت علاقه زیاد به این ماشین ها، که پرسپترون^۱ خوانده می شدند، دستیابی به اثبات های ریاضی بود. که نشان می دادند وقتی پرسپترون ها با مجموعه های آموزشی جدا پذیر و خطی آموزش داده شوند، ظرف چند مرحله تکرار محدود به جواب همگرا می شوند. این جواب به شکل ضرایب ابرصفحه هایی بود که دسته های حاصل از الگوهای مجموعه آموزشی را به درستی از هم جدا می کردند.

متاسفانه بزودی تمام امیدها به کشفی که ظاهرآ یک مدل نظری قوی برای یادگیری بود، به نامیدی تبدیل گردید. پرسپترون پایه و بعضی از تعمیم های آن برای کارهای پردازش الگوی عملی، در آن زمان ناکافی بود. تلاش های بعدی برای افزایش توان ماشین های شبیه پرسپترون با استفاده از چند لایه از این ماشین ها، گرچه مفهوم جذابی دارد، اما الگوریتم های آموزشی کارآمد نظیر آنها که سبب جذابیت پرسپترونها شده بودند، را کم داشتند. [1965] Nilsson مبحث ماشین های آموزشی را در نیمه 'دهه' ۱۹۶۰ جمع بنده کرد. چند سال بعد، [1969] Minskey & Papert تحلیلی دلسرد کننده درباره 'محدودیت ماشین های شبیه پرسپترون' ارائه کردند. طبق توضیحات [1986] Simon، این دیدگاه تا نیمه 'دهه' ۱۹۸۰ میلادی باقی بود. در مقاله وی، که ابتدا در سال ۱۹۸۴ و در فرانسه چاپ شد، Simon تحت عنوان "تولد و مرگ یک افسانه" پرسپترون را به کناری نهاد.

نتایج اخیر توسط [1986] Rumelhart, Hinton & Williams در رابطه با توسعه الگوریتم های آموزشی جدید برای پرسپترون های چند لایه، تا حد زیادی مسأله را عوض کرده اند. روش آنها، که غالباً قاعده 'دلتای تعمیم یافته'^۲ برای آموزش با پس انتشار^۳ خوانده می شود، روش آموزش مؤثری برای ماشین های چند لایه است. گرچه نمی توان نشان داد که این الگوریتم آموزشی در مفهومی مشابه با پرسپترون تک لایه به جواب همگرامی شود، اما قاعده 'دلتای تعمیم یافته' در مسائل عملی مختلفی با موفقیت استفاده شده است. این توفيق، ماشین های شبیه پرسپترون چند لایه را به عنوان یکی از مدل های اصلی شبکه های عصبی فعلی تثیت کرده است.

گرچه پیشرفت هایی نظیر قواعد آموزشی جدید برای ماشین های چند لایه، معرفی مدل های جدید برای شبکه های عصبی و بعضی نتایج کاربردی دلگرم کننده، علاقه مجددی را به شاخه 'تشخیص الگو' توسط ماشین های یادگیرنده برانگیخته اند، تحقیق در این مبحث هنوز هم در مراحل ابتدایی است. معزز انسان صدها بیلیون نرون دارد، که در شبکه 'پیچیده ای سازمان یافته' اند و هر نرون ممکن است به چند

هزار نرون دیگر متصل باشد. البته این موضوع هنوز درک نشده است که چگونه این شبکه' باتوازی انبوهش، که در آن هر نرون در سرعت هایی از مرتبه' میلی ثانیه عمل می کند (در مقایسه با سرعت های نانو ثانیه ای قطعات الکترونیکی فعلی)، می تواند کارهایی نظیر جمع آوری، ذخیره، نمایش، بازیابی، و تحلیل داده های حسی^۱ بسیار پیچیده را با سرعت های شگفت آور انجام دهد. تفکر درباره' توانایی بسیار زیاد مغز و روشی که مغز تصاویر افراد، مکان ها، یا واقعی را تنها بر اساس بخش هایی از اطلاعات آنها، نظیر صدا یا بوی آشنا یا خاطره یک روز تاریخی، بازسازی می کند، بسیار بهت آور است.

با این شرایط کارهای انجام شده در شبکه های عصبی مصنوعی نسبتاً ناچیز است. بنابراین مشکلات پیش روی این مبحث تحقیقاتی زیاد است. در اینجا چند جنبه از وضعیت فعلی پیاده سازی شبکه های عصبی چند لایه را معرفی می کنیم. شرح را با معرفی پرسپترون به عنوان مدل پایه ای نرون شروع می کنیم. سپس جنبه های گوناگون آموزش پرسپترون ها را تحت شرایط جدا پذیری و جدا ناپذیری دسته ها شرح می دهیم. این مطلب به عنوان مبنای طرح، توسعه و نمایش قاعده' دلتای تعمیم یافته برای آموزش شبکه های عصبی چند لایه می باشد.

پرسپترون برای دو دسته الگو

در پایه ای ترین شکل، پرسپترون یکتابع تصمیم خطی که دو مجموعه آموزشی جدا پذیر و خطی را از هم جدا می کند، پیاده سازی می نماید. شکل ۹ . ۱۴ (الف) مدل پرسپترون برای دو دسته الگورانشان می دهد. پاسخ این وسیله بر مبنای جمع وزن دار ورودی هایش است؛ یعنی،

$$d(\mathbf{x}) = \sum_{i=1}^n w_i x_i + w_{n+1} \quad (29-۳ . ۹)$$

که یک تابع تصمیم خطی بر حسب مؤلفه های بردارهای الگو است (معادله' ۹ . ۱-۲ را ببینید). ابتدا ضرایب $w_1, w_2, \dots, w_n, w_{n+1}$ را بخواهیم شوند، در ورودی ها ضرب و سپس حاصل ضرب ها وارد عنصر آستانه گیر می شوند. از این نظر، وزنها مشابه سیناپس های^۳ سامانه عصبی انسان هستند. تابعی که خروجی جمع کننده را به خروجی نهایی می نگارد، گاهی اوقات تابع فعالیت^۴ خوانده می شود.

وقتی $d(\mathbf{x}) > 0$ عنصر آستانه گیر باعث می شود که خروجی پرسپترون $+1$ باشد، که نشان دهنده این

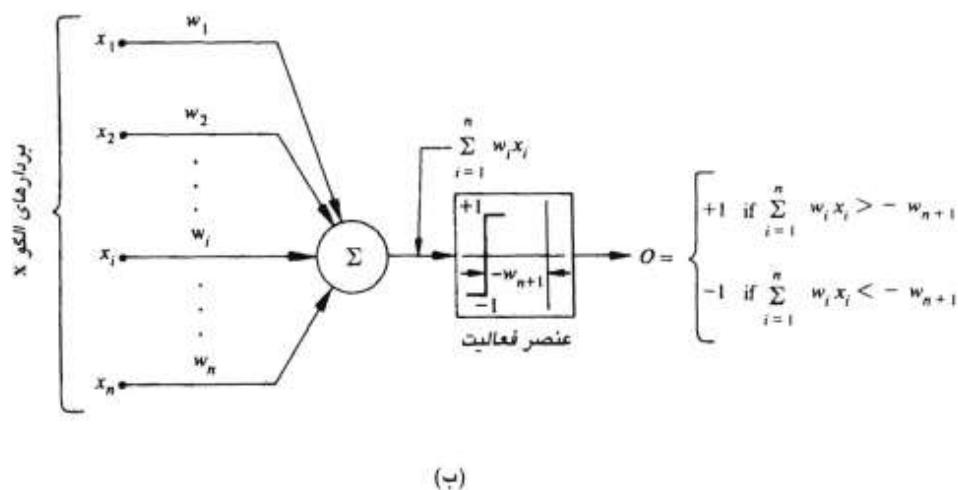
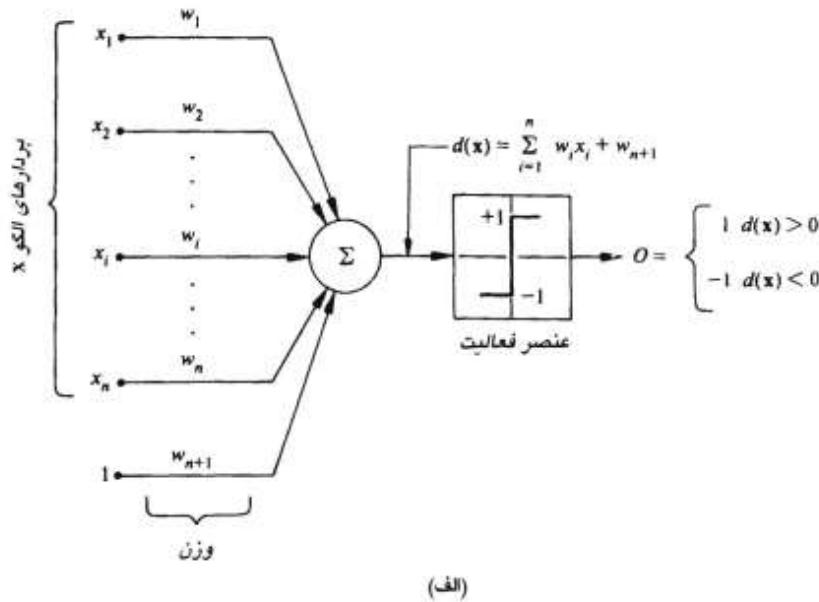
1) Sensory data

2) Weight

3) Synapse

4) Activation function

۷۰۹ ۳. روش های نظریه تصمیم



شکل ۱۴۰۹ دو نمایش معادل از مدل پرسپترون برای دسته های دو الگویی

است که الگری \mathbf{X} به عنوان عضو دسته w تشخیص داده شده است. وقتی $0 < d(\mathbf{x})$ عکس آن درست است. این نحوه کار با توضیحات قبلی راجع به معادله (۲-۳) در مورد استفاده از یک تابع تصمیم برای دو دسته الگو موافقت دارد. وقتی $0 = d(\mathbf{x})$ روی سطح تصمیمی که دو دسته الگورا جدا می کند، قرار دارد و یک وضعیت میانی را می دهد. مرز تصمیمی پرسپترون با برابر صفر قرار دادن معادله (۲۹-۳) بدست می آید:

$$d(\mathbf{x}) = \sum_{i=1}^n w_i x_i + w_{n+1} = 0 \quad (30-3, 9)$$

یا

$$w_1 x_1 + w_2 x_2 + \dots + w_n x_n + w_{n+1} = 0 \quad (31-3, 9)$$

که معادله یک ابرصفحه^۱ در فضای الگوی w بعدی است. از نظر هندسی، ضریب اول جهت ابرصفحه را تعیین می کنند، و آخرین ضریب، w_{n+1} ، متناسب با فاصله 'عمودی' مبدأ تا ابرصفحه است. یعنی اگر $w_{n+1} = 0$ ، ابرصفحه از مبدأ صفحه الگوی می گذرد. به طور مشابه اگر $w_i = 0$ ، ابرصفحه موازی محور x_i است.

خروجی عنصر آستانه گیر در شکل ۹.۱۴(الف) به علامت $d(\mathbf{x})$ بستگی دارد. به جای آزمودن کل تابع برای تعیین مثبت یا منفی بودن آن، می توانیم بخش مجموع معادله (۲۹-۳) را با w_{n+1} مقایسه کنیم، که در این مورد خروجی سامانه عبارت است از

$$O = \begin{cases} +1 & \text{اگر } \sum_{i=1}^n w_i x_i > -w_{n+1} \\ -1 & \text{اگر } \sum_{i=1}^n w_i x_i < -w_{n+1}. \end{cases} \quad (32-3, 9)$$

این پیاده سازی معادل شکل ۹.۱۴(الف) است که در شکل ۹.۱۴(ب) نیز نشان داده می شود. تنها تفاوت این هستند که تابع آستانه ای به میزان w_{n+1} - جایه جا می شود و ورودی ثابت یک نیز دیگر وجود ندارد. بعداً در این بخش وقتی پیاده سازی شبکه های عصبی چند لایه را شرح می دهیم، به این روابط

1) Hyperplane

۳.۹ روش های نظریه تصحیم ۷۱۱

معادل مراجعه می کنیم.

رابطه' دیگری که معمولاً در عمل به کار می رود، این است که بردارهای الگوریتم افزودن عنصر $(n+1)$ که بدون توجه به دسته' عضویت الگو، همواره برابر ۱ است، افزایش دهیم. یعنی، بردار الگوی افزایش یافته^۱ \mathbf{y} ، از بردار الگوی \mathbf{x} با قرار دادن $x_i = y_i$ ، برای $i = 1, 2, \dots, n$ ، و افزودن عنصر اضافی $y_{n+1} = 1$ بدست می آید. آنگاه معادله^۲ (۹.۲۹-۳) به صورت

$$\begin{aligned} d(\mathbf{y}) &= \sum_{i=1}^{n+1} w_i y_i \\ &= \mathbf{w}^T \mathbf{y} \end{aligned} \quad (33-3.9)$$

می شود که اکنون^۳ $(1, y_1, y_2, \dots, y_n)^T = \mathbf{y}$ بردار الگوی افزایش یافته است، و $\mathbf{w} = (w_1, w_2, \dots, w_n, w_{n+1})^T$ بردار وزن خوانده می شود. این طرز بیان از نظر نوشتمناسب تر است. با این حال بدون توجه به رابطه' مورد استفاده، مشکل اصلی، پیدا کردن \mathbf{w} با استفاده از یک مجموعه' آموزشی داده شده از هر دو دسته می باشد.

الگوریتم های آموزش

الگوریتم هایی که در زیر مطرح می شوند نماینده' روش های متعددی هستند که در طول سال ها برای آموزش پرسپترون ها پیشنهاد شده اند.

دسته های جداپذیر خطی . یک الگوریتم تکراری ساده برای بدست آوردن بردار وزن نهایی برای دو مجموعه' آموزشی جداپذیر خطی بدین صورت است: ابتدا برای دو مجموعه' آموزشی از بردارهای الگوی افزایش یافته که به ترتیب به دسته های الگوی \mathbf{w}_1 و \mathbf{w}_2 تعلق دارند، فرض کنید (۱) \mathbf{w}_1 بیانگر بردار وزن ابتدایی باشد، که می توان آن را به دلخواه انتخاب کرد. آنگاه در k امین مرحله' تکرار؛ اگر $\mathbf{y}(k)$ و $\mathbf{w}^T(k)\mathbf{y}(k) \leq 0$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + c\mathbf{y}(k) \quad (34-3.9)$$

جایگزین کنید، که c ضریب تصحیح مثبت^۴ است. اما اگر $\mathbf{y}(k) \in \omega_2$ و $\mathbf{w}^T(k)\mathbf{y}(k) \geq 0$ را با

1) Augmented

2) Weight vector

3) Positive correction increment

$$\mathbf{w}(k+1) = \mathbf{w}(k) - c\mathbf{y}(k). \quad (35-3 . ۹)$$

جایگزین کنید. در غیر این صورت، $\mathbf{w}(k)$ را بدون تغییر رها کنید؛ یعنی قرار دهید

$$\mathbf{w}(k+1) = \mathbf{w}(k). \quad (36-3 . ۹)$$

به بیان ساده این الگوریتم تنها وقتی \mathbf{w} را تغییر می دهد که الگوی مورد استفاده در مرحله k آموزش با استفاده از بردار وزن آن مرحله، به طور نادرست دسته بندی شود. ضریب تصحیح c مثبت است و اکنون ثابت فرض می شود. این الگوریتم گاهی با عنوان قاعده 'تصحیح با ضریب ثابت' نامیده می شود.
این روش آموزش به روشنی بر مبنای مفهوم پاداش - و - تنبیه است. "پاداش" ماشین برای دسته بندی صحیح یک الگو، در واقع عدم "تبیه" است. به عبارت دیگر، اگر ماشین الگوی را به طور صحیح دسته بندی کند، به وسیله عدم تغییر \mathbf{w} پاداش داده می شود. اما اگر الگو را به طور نادرست دسته بندی کند، با انجام تغییر در بردار وزن تنبیه می شود. وقتی الگوریتم همگرا شده است که کل چرخه آموزش برای هر دو دسته الگوها بدون تغییر وزن ها انجام شود. اگر دو مجموعه الگوی آموزشی جداپذیر و خطی باشند، قاعده 'تصحیح با ضریب ثابت' در تعداد محدودی از مراحل همگرا می شود. اثباتی از این نتیجه را، که گاهی اوقات قضیه آموزش پرسپترون^۲ خوانده می شود، می توان در کتاب های [Nilsson 1965]، Gonzalez & Tou[1974]، Hart & Duda[1973] پیدا کرد.

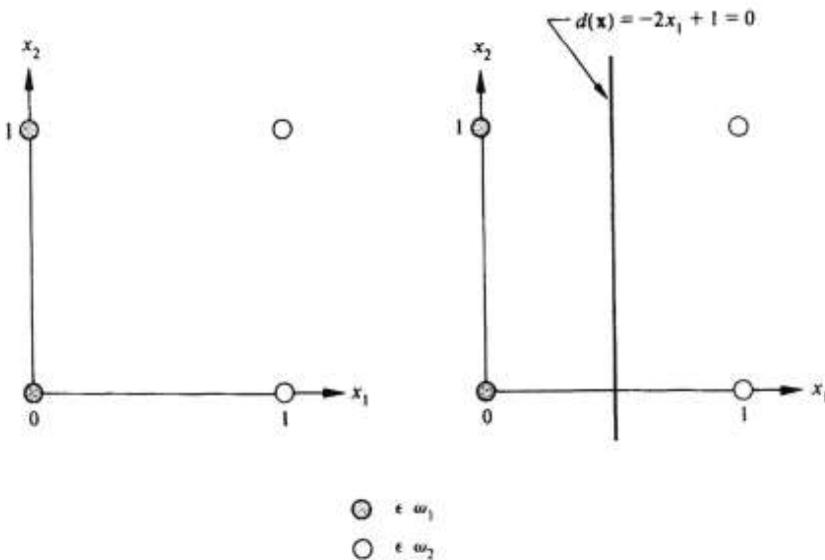
مثال: دو مجموعه آموزشی نشان داده شده در شکل ۹ . ۱۵ (الف) را، که هر یک متشکل از دو الگو هستند، در نظر بگیرید. چون دو مجموعه آموزشی جداپذیر هستند، الگوریتم آموزش موفق خواهد بود.

قبل از اجرای الگوریتم، الگوها افزایش یافته می شوند که طی آن مجموعه آموزشی \mathbf{w} به دست می آیند. با قرار دادن $1 = \mathbf{w}(1), 0 = \mathbf{w}(2)$ و دادن ترتیب الگوها به الگوریتم، دنباله مراحل

زیر نتیجه می شود:

$$\mathbf{w}^T(1)\mathbf{y}(1) = [0, 0, 0] \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0 \quad \mathbf{w}(2) = \mathbf{w}(1) + \mathbf{y}(1) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

۷۱۳ ۹. روش های نظریه تصمیم



شکل ۹.۱۵ نمایشی از الگوریتم آموزش پرسپترون: (الف) الگوهایی که به دو دسته تعلق دارند؛ (ب) مرز تصمیم تعیین شده با آموزش.

$$\begin{aligned} \mathbf{w}^T(2)\mathbf{y}(2) &= [0, 0, 1] \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 1 & \mathbf{w}(3) &= \mathbf{w}(2) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ \mathbf{w}^T(3)\mathbf{y}(3) &= [0, 0, 1] \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1 & \mathbf{w}(4) &= \mathbf{w}(3) - \mathbf{y}(3) = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \\ \mathbf{w}^T(4)\mathbf{y}(4) &= [-1, 0, 0] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1 & \mathbf{w}(5) &= \mathbf{w}(4) = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$

که در مراحل اول و سوم به علت دسته بنده ناصحیح، طبق معادلات (۹.۳۴-۳) و (۹.۳۵-۳) تغییرهای در بردار وزن انجام شد. چون جواب تنها وقتی بدست می آید که الگوریتم در تمام الگوهای آموزشی بدون خطا انجام شود، باید مجدداً مجموعه آموزشی به الگوریتم داده شود. فرآیند یادگیری

ماشین با قرار دادن $y(1) = y(8)$ ، $y(2) = y(5)$ ، $y(3) = y(6)$ و $y(4) = y(7)$ ادامه کار با همین روش دنبال می شود. در $k = 14$ به همگرایی می رسیم و بردار وزن نهایی $w(14) = (-2, 0, 1)^T$ بدست می آید. تابع تصمیم متناظر عبارت است از $r - 2y_1 + d(y) = -2y_1 + d(y)$. با قرار دادن $y_i = x_i$ به فضای الگوی اولیه باز گشته و بدین وسیله $r - 2x_1 + d(x) = -2x_1 + d(x)$ بدست می آید که وقتی برابر صفر قرار داده شود، همان معادله 'مرز تصمیم' نشان داده شده در شکل ۹.۱۵(ب)، خواهد بود.

دسته های جداپذیر. در عمل، دسته های الگوی جداپذیر و خطی (نادر) استثناء هستند نه قاعده. نتیجتاً در خلال دهه های ۱۹۶۰ و ۱۹۷۰ میلادی تحقیقات زیادی برای تولید و توسعه 'روش های کار با دسته های الگوی جداپذیر' انجام شد. با پیشرفت های اخیر در آموزش شبکه های عصبی، بسیاری از روش هایی که برای رسیدگی به جداپذیری پیشنهاد شده اند، تنها از لحاظ تاریخی مورد توجه هستند. در هر حال یکی از اولین روش هایی که مستقیماً با این موضوع مرتبط است، 'قاعده' دلتای اولیه می باشد. این روش که معروف به قاعده Widrow - Hoff، یا 'قاعده دلتای کمینه' میانگین مربع (LMS) برای آموزش پرسپترون ها است، در هر مرحله 'آموزش خطای بین پاسخ واقعی و مطلوب را کمینه می کند.

تابع معیار

$$J(w) = \frac{1}{2} (r - w^T y)^2 \quad (37-3.9)$$

را در نظر بگیرید که در آن r پاسخ مطلوب است (یعنی، اگر y ، بردار الگوی آموزشی، متعلق به دسته '۱' باشد، $r = 1$ و اگر y متعلق به دسته '۰' باشد، $r = 0$). وظیفه این است که گام به گام w را در جهت شبیه منفی تابع $J(w)$ تصحیح کنیم تا این که کمینه 'این تابع را بیابیم. این کمینه که متناظر با دسته بندی صحیح است، در $y = w^T r$ رخ می دهد. اگر (k) w بیانگر بردار وزن در k امین مرحله 'تکرار باشد، می توان الگوریتم کلی نزول گرادیان را به صورت

$$w(k+1) = w(k) - \alpha \left[\frac{\partial J(w)}{\partial w} \right]_{w=w(k)} \quad (38-3.9)$$

نوشت که $w(k+1)$ مقدار جدید w است، و $0 < \alpha$ ، بزرگی مقدار تصحیح را تعیین می کند. از معادله '۳۷-۳' داریم که

$$\frac{\partial J(w)}{\partial w} = -(r - w^T y)y. \quad (39-3.9)$$

۷۱۵ ۳. روش های نظریه' تصمیم

با جای گذاری این نتیجه در معادله (۴۰-۳) نتیجه می شود:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \alpha[r(k) - \mathbf{w}^T(k)\mathbf{y}(k)]\mathbf{y}(k) \quad (40-3, 9)$$

که بردار وزن شروع، یعنی (۱) \mathbf{w} ، دلخواه است. با انجام تصحیح تنها در وقتی که الگو به طور نادرست دسته بندی شده است، این امکان حاصل می شود که معادله (۴۰-۳) را به شکل الگوریتم آموزش پرسپترون داده شده با معادلات (۳۶-۳) - (۳۴-۳)، بیان کنیم.

با تعریف تغییر (دلتا) در بردار وزن به صورت

$$\Delta\mathbf{w} = \mathbf{w}(k+1) - \mathbf{w}(k) \quad (41-3, 9)$$

می توانیم معادله (۴۰-۳) را به شکل الگوریتم تصحیح دلتا^۱، یعنی

$$\Delta\mathbf{w} = \alpha e(k)\mathbf{y}(k) \quad (42-3, 9)$$

نویسیم که

$$e(k) = r(k) - \mathbf{w}^T(k)\mathbf{y}(k) \quad (43-3, 9)$$

بوده و خطای ناشی از بردار وزن $\mathbf{w}(k)$ برای الگوی داده شده $\mathbf{y}(k)$ می باشد.
معادله (۴۳-۳) خطابازی بردار وزن $\mathbf{w}(k)$ را بدست می دهد. اگر آن را به $\mathbf{w}(k+1)$ تغییر دهیم، اما الگو بدون تغییر باقی بماند، خطاب به صورت زیر در می آید

$$e(k) = r(k) - \mathbf{w}^T(k+1)\mathbf{y}(k). \quad (44-3, 9)$$

آنگاه تغییر در خطاب عبارتست از

$$\begin{aligned} \Delta e &= [r(k) - \mathbf{w}^T(k+1)\mathbf{y}(k)] - [r(k) - \mathbf{w}^T(k)\mathbf{y}(k)] \\ &= -[\mathbf{w}^T(k+1) - \mathbf{w}^T(k)]\mathbf{y}(k) \\ &= -\Delta\mathbf{w}^T\mathbf{y}(k). \end{aligned} \quad (45-3, 9)$$

1) *Delta correction algorithm*

۲۱۶ پردازش تصویر رقمه

$$\text{اما } \Delta\mathbf{w} = \alpha e(k)\mathbf{y}(k), \text{ بنابراین}$$

$$\begin{aligned}\Delta e &= -\alpha e(k)\mathbf{y}^T(k)\mathbf{y}(k) \\ &= -\alpha e(k)\|\mathbf{y}(k)\|^2.\end{aligned}\quad (46-3 . 9)$$

بنابراین تصحیح وزن ها، خط را با ضریب $\alpha\|\mathbf{y}(k)\|^2$ کاهش می دهد. الگوی ورودی بعدی چرخه تطابق جدیدی را شروع می کند و خطای بعدی را با ضریب $(1 + \alpha\|\mathbf{y}(k)\|^2)$ کاهش می دهد، به همین ترتیب الى آخر.

انتخاب α پایداری و سرعت همگرای را کنترل می کند (Stearns & Widrow[1985]). برای پایداری باید $0 < \alpha < 2$. یک محدوده عملی برای α عبارتست از $0.1 < \alpha < 1.0$. گرچه در اینجا نشان داده نشده است اما الگوریتم معادله^۹ (۴۰-۳) یا معادلات^۹ (۴۲-۳) و^۹ (۴۳-۳) به جوابی همگرا می شود که خطای میانگین مربع را روی الگوهای مجموعه آموزشی کمینه می کند. وقتی که دسته های الگو جداپذیر باشند، جواب حاصل از روال Widrow - Hoff ممکن است بتواند یا تواند ابر صفحه جداکننده را تولید کند؛ یعنی، جواب میانگین مربع خط، یک جواب در مفهوم قضیه آموزش پرسپترون نیست. این عدم قطعیت هزینه استفاده از الگوریتمی است که در این مورد ویژه در هر دو حالت جداپذیر و جداانپذیر همگرا می شود.

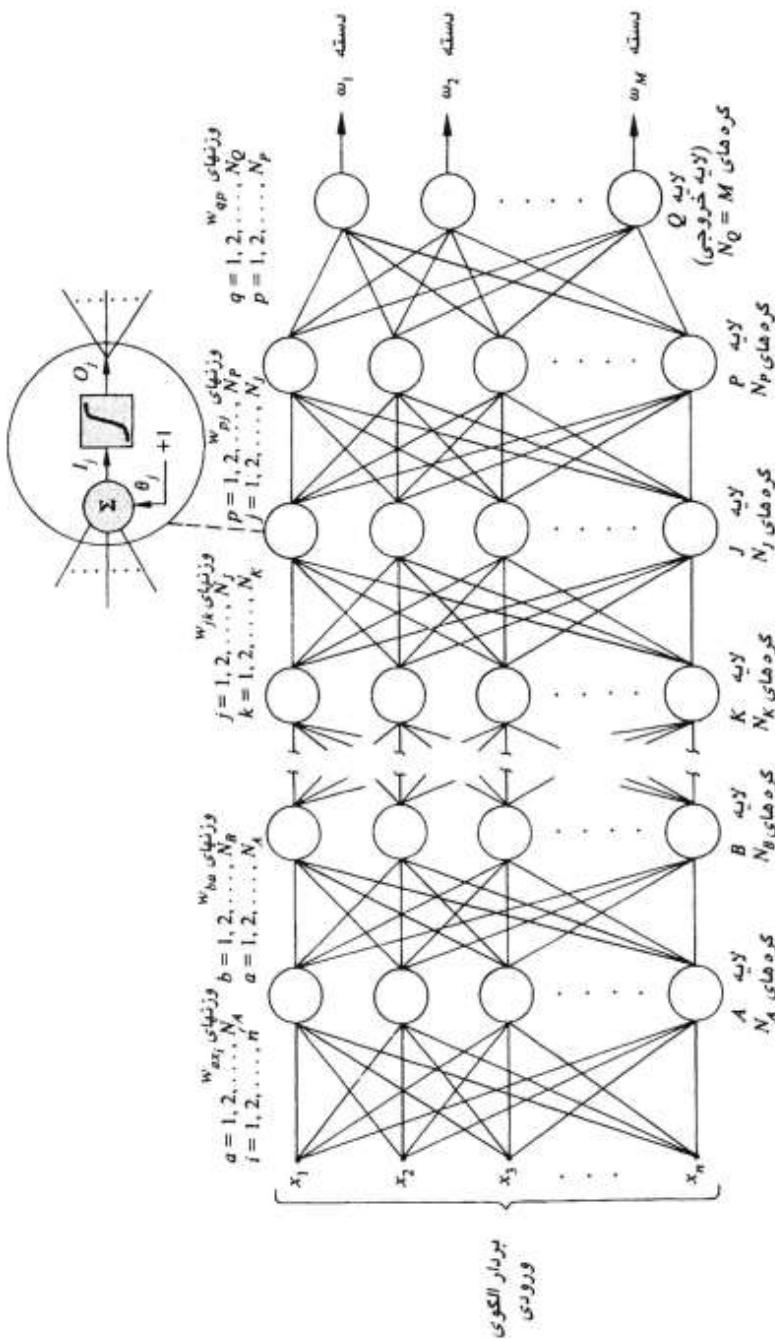
دو الگوریتم آموزش پرسپترون که تاکنون شرح داده شدند، را به آسانی می توان به مواردی که شامل بیش از دو دسته هستند، نیز توسعه داد. براساس تاریخچه ای که در شروع این بخش ارائه شد، معرفی این الگوریتم های آموزش چند دسته ای، کاربرد کمی دارد. در عوض مبحث آموزش چند دسته ای را در حوزه شبکه های عصبی پیشخور چند لایه.

شبکه های عصبی پیشخور چند لایه

اکنون می توانیم مستقل از این که دسته ها جداپذیر هستند یا نیستند، به توابع تصمیم در مسائل تشخیص الگوی چند دسته ای پرسپترون و ساختارهایی را به کار گیریم که متشکل از لایه هایی از عناصر محاسباتی پرسپترونی هستند.

ساختار پایه ای، شکل ۹ . ۱۶ ساختار سامانه ای مدل شبکه عصبی مورد بررسی را نشان می دهد. این شبکه متشکل از لایه هایی از گره های محاسباتی هم ساختار (نرون ها) می باشد. نرون ها طوری مرتب

۳. روش های نظریه' تصمیم ۷۱۷



شده اند که خروجی هر نرون از یک لایه به ورودی تمام نرون های لایه بعدی متصل است. تعداد نرون های لایه 'اول، که لایه 'A خوانده می شود، $N_A = n$ می باشد. غالباً // ابعاد بردارهای الگوی ورودی است. تعداد نرون ها در لایه 'خروجی، که لایه 'Q خوانده می شود، با نماد N_Q نشان داده می شود. شبکه برابر M است که M تعداد دسته های الگو که شبکه عصبی با آنها آموزش دیده است، می باشد. شبکه عصبی بردار الگوی X را به عنوان عضو دسته ' تشخیص می دهد اگر خروجی m شبکه "فعال"^۱ و سایر خروجی ها "غیرفعال"^۲ باشند.

همان طور که نزدیک نمای^۳ موجود در شکل ۹ . ۱۶ نشان می دهد، هر نرون همان شکل مدل پرسپترون که قبلًا تشریح شد، را دارد (شکل ۹ . ۱۴ را بینید)، البته با این تفاوت که یکتابع «سیگموید»^۴ با برش نرم^۵ جایگزین تابع فعالیت با برش سخت^۶ شده است. در پاده سازی قاعده آموزش با پس انتشار، مشتق پذیری در طول تمام مسیرهای شبکه عصبی مورد نیاز است. تابع فعالیت سیگمویدی زیر مشتق پذیری مورد نیاز را دارد است:

$$h_j(I_j) = \frac{1}{1 + \exp[-(I_j + \theta_j)/\theta_0]} \quad (47-3 . 9)$$

که $I_j, I_j = 1, 2, \dots, N_j$ ، نمایانگر ورودی به عنصر فعالیت هر گره از لایه ' J ام شبکه است، θ_j مقدار افست^۷ است، و θ_0 نیز شکل تابع سیگموید را کنترل می کند. در شکل ۹ . ۱۷ ، معادله (۹ . ۴۷-۳) همراه با حدود پاسخ های "فعال" و "غیرفعال" خروجی هر گره رسم شده است. بنابراین وقتی این تابع ویژه استفاده شود، سامانه برای هر مقدار I_j که بزرگتر از θ_0 باشد، خروجی "فعال" می دهد. به طور مشابه، سامانه برای هر مقدار I_j که کوچکتر از θ_0 باشد، خروجی "غیرفعال" می دهد. همان طور که شکل ۹ . ۱۷ نشان می دهد، تابع فعالیت سیگمویدی، همیشه مثبت است، و تنها وقتی می تواند به مقادیر حدی ۰ و ۱ برسد که ورودی عنصر فعالیت به ترتیب به مقادیر بی نهایت منفی و بی نهایت مثبت برسد. به این دلیل، مقادیر نزدیک ۰ و ۱ (مثلاً ۰.۰۵ و ۰.۹۵) مقادیر غیر فعال و فعال خروجی نرنونهای شکل ۹ . ۱۶ را تعریف می کنند. در اصل، می توان از توابع فعالیت متفاوتی برای لایه های متفاوت یا حتی گره های متفاوت در یک لایه از شبکه عصبی استفاده کرد. در

1) High

2) Low

3) Blow-up

4) Sigmoid

5) soft-limiting

6) Hard-limiting

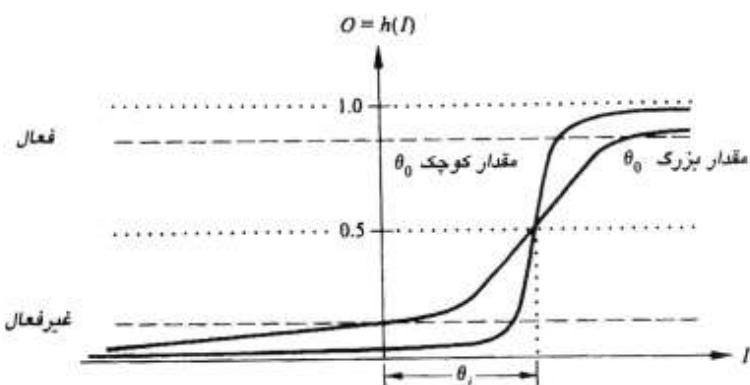
7) Offset

۳.۹ روش های نظریه تصمیم ۷۱۹

عمل، روش معمول این است که در سراسر شبکه از یک شکل تابع فعالیت استفاده شود. با مراجعه به شکل ۹.۱۴(ب)، افست θ_0 در شکل ۹.۱۷ مشابه ضربی وزنی w_{i+1} در شرح قبلی پرسپترون است. با تبدیل افست θ_0 به ضربی اضافی که ورودی واحد هر یک از گره های شبکه را تغییر می دهد، می توان تابع آستانه گیری جابجا شده را به صورت شکل ۹.۱۴(الف) پیاده سازی کرد. به منظور پیروی از طرز نگارشی که غالباً در متون این مبحث استفاده می شود، ورودی های ثابت +1 به گره های شکل ۹.۱۶ را به طور مجزا نشان نمی دهیم. در عوض، این ورودی و وزن تغییر دهنده آن، θ_0 را از اجزاء ضروری و ضمنی هر یک از گره های شبکه در نظر می گیریم. همان طور که در مورد نزدیک نمای شکل ۹.۱۶ ذکر شد، برای هر گره N_j در لایه J کی از چنین ضرایب وجود دارد. برای ساختاری که در شکل ۹.۱۶ دیده می شود، ورودی هر یک از گره ها در هر یک از لایه ها جمع وزنی خروجی های لایه J قبل است. با فرض این که لایه K ، قبل از لایه J باشد، ورودی عنصر فعالیت هر گره از لایه J ، که با I_j نشان داده می شود، برای N_j به صورت:

$$I_j = \sum_{k=1}^{N_K} w_{kj} O_k \quad (48-3.9)$$

بیان می شود، که N_j تعداد گره های لایه J ، N_K تعداد گره های لایه K ، و w_{kj} وزن هایی هستند که خروجی های O_k از گره های لایه K را قبل از این که وارد گره های لایه J بشوند، تغییر می دهند.



شکل ۹.۱۷ تابع فعالیت سیگمویدی معادله (۹.۳-۴۷)

۷۲۰ پردازش تصویر رقمه

خروجی های لایه K ، برای $k = 1, 2, \dots, N_K$ عبارتند از:

$$O_k = h_k(I_k) \quad (49-3. 9)$$

طرز زیرنویس گذاری استفاده شده در معادله (۹ . ۴۸-۳) مهم است، زیرا در باقی مانده این بخش از آن استفاده می کنیم. در ابتدا ذکر کنیم که I_1, I_2, \dots, I_{N_K} ، بیانگر ورودی به عنصر فعالیت^۱ گره 'زم' در لایه 'J' است. بنابراین I_1 بیانگر ورودی به عنصر فعالیت اولین (بالاترین) گره در لایه 'J'، I_2 بیانگر ورودی به عنصر فعالیت گره 'دوم در لایه 'J' است، و به همین ترتیب الى آخر. هر گره از لایه 'J'، N_K ورودی دارد، که به هر یک از ورودیها وزن متفاوتی داده می شود. بنابراین N_K ورودی اولین گره 'لام' با ضرایب $w_{11}, w_{12}, \dots, w_{1N_K}$ ، وزن داده می شوند؛ ورودی های گره 'دوم با ضرایب $w_{21}, w_{22}, \dots, w_{2N_K}$ وزن داده می شوند؛ و به همین ترتیب الى آخر. بنابراین جمعاً به ضریب برای مشخص کردن وزن دهنده های خروجی های لایه 'K' در هنگام ورودشان به لایه 'J' $\times N_J \times N_K$ ضریب برای مشخص کردن وزن دهنده های خروجی های لایه 'K' در معادله (۹ . ۴۷-۳) داریم:

$$h_k(I_k) = \frac{1}{1 + \exp \left[- \left(\sum_{j=1}^{N_K} w_{jk} O_j + \theta_k \right) / \theta_0 \right]} \quad (50-3. 9)$$

که همان شکل تابع فعالیت مورد استفاده در ادامه این بخش است.

در خلال آموزش، تطابق نرون های لایه 'خروجی' کار ساده ای است، زیرا خروجی مطلوب هر گره معلوم است. مشکل اصلی در آموزش شبکه 'چند لایه'، تنظیم وزن های لایه های پنهان^۲، یعنی لایه های غیر از لایه 'خروجی'، است.

آموزش با پس انتشار^۳. با تمرکز بر لایه 'خروجی' شروع می کنیم. خطای مربعی کل بین پاسخ های مطلوب، O_q ، و پاسخ های واقعی متناظر، r_q ، در گره های لایه 'خروجی' Q عبارتست از:

$$E_Q = \frac{1}{2} \sum_{q=1}^{N_Q} (r_q - O_q)^2 \quad (51-3. 9)$$

1) Activation element

2) Hidden layer

3) Back propagation

۴۰۳. ۹ روش های نظریه 'تصمیم' ۷۲۱

که N تعداد گره های لایه خروجی Q بوده و ضریب $1/2$ برای سادگی رابطه حاصل از مشتق گیری اضافه شده است.

هدف این است که یک قاعده آموزش، شبیه به قاعده دلتا، طرح کنیم که با جستجوی کمینه برای تابع خطایی به شکل معادله (۹. ۵۱-۳)، تنظیم وزن های هر یک از لایه هارا امکان می دهد. همانند گذشته، تغییر وزن ها متناسب با مشتق جزئی خطایی نسبت به وزن ها، مارا به تبعیجه می رساند. به عبارت دیگر،

$$\Delta w_{qp} = - \alpha \frac{\partial E_Q}{\partial w_{qp}} \quad (52-3. 9)$$

که P لایه قبل از Q است، Δw_{qp} به صورت تعریف شده در معادله (۹. ۴۲-۳)، و α یک ضریب صحیح مثبت است.

خطای E_Q تابعی از خروجی ها، O_q ها، بوده که خود آنها نیز توابعی از ورودی ها، I_p ها، هستند. با استفاده از قاعده زنجیری، مشتق جزئی E_Q به صورت زیر محاسبه می شود:

$$\frac{\partial E_Q}{\partial w_{qp}} = \frac{\partial E_Q}{\partial I_q} \frac{\partial I_q}{\partial w_{qp}} \quad (53-3. 9)$$

طبقه معادله (۹. ۴۸-۳)،

$$\frac{\partial I_q}{\partial w_{qp}} = \frac{\partial}{\partial w_{qp}} \sum_{p=1}^{N_p} w_{qp} O_p = O_p \quad (54-3. 9)$$

با جای گذاری معادلات (۹. ۵۲-۳) و (۹. ۵۳-۳) در معادله (۹. ۵۴-۳) داریم:

$$\begin{aligned} \Delta w_{qp} &= - \alpha \frac{\partial E_Q}{\partial I_q} O_p \\ &= \alpha \delta_q O_p \end{aligned} \quad (55-3. 9)$$

که

$$\delta_q = - \frac{\partial E_Q}{\partial I_q} \quad (56-3. 9)$$

به منظور محاسبه $\partial E_Q / \partial I_q$ ، با استفاده از قاعده زنجیری، مشتق جزئی را بر حسب آهنگ تغییر E_Q

نسبت به O_q و آهنگ تغییر O نسبت به I بیان می کنیم. یعنی،

$$\delta_q = -\frac{\partial E_Q}{\partial I_q} = -\frac{\partial E_Q}{\partial O_q} \frac{\partial O_q}{\partial I_q} \quad (57-3. 9)$$

طبق معادله (۵۱-۳)،

$$\frac{\partial E_Q}{\partial O_q} = -(r_q - O_q) \quad (58-3. 9)$$

واز معادله (۴۹-۳) داریم که:

$$\frac{\partial O_q}{\partial I_q} = \frac{\partial}{\partial I_q} h_q(I_q) = h'_q(I_q) \quad (59-3. 9)$$

با جای گذاری معادلات (۵۸-۳) و (۵۹-۳) در معادله (۵۷-۳) داریم که:

$$\delta_q = (r_q - O_q) h'_q(I_q) \quad (60-3. 9)$$

که δ_q متناسب با خطای $O_q - r_q$ است. با جای گذاری معادلات (۵۶-۳) - (۵۸-۳) در معادله (۵۵-۳) نهایتاً داریم که

$$\begin{aligned} \Delta w_{qr} &= \alpha(r_q - O_q) h'_q(I_q) O_p \\ &= \alpha \delta_q O_p. \end{aligned} \quad (61-3. 9)$$

پس از این کهتابع $(I_q) h_q$ مشخص شد، تمام اجزای معادله (۶۱-۳) معلوم هستند یا این که می توان آنها در شبکه مشاهده کرد. به عبارت دیگر، بعد از دادن هر الگوی آموزشی به ورودی شبکه، می دانیم که پاسخ مطلوب هر گره خروجی، r_p ، چه باید باشد. مقدار O_p هر گره خروجی را می توان همانند O_q که به ترتیب ورودی عناصر فعالیت لایه Q و خروجی گره های لایه P هستند، مشاهده کرد. بنابراین می دانیم که چگونه ضرایبی را که اتصالات بین لایه آخر و لایه ماقبل آخر را وزن می دهند، تنظیم کنیم.

با حرکت از لایه خروجی به سمت لایه ورودی، اکنون اجازه دهید بررسی کنیم که در لایه P چه اتفاقی می افتد. با انجام همان روال داریم که:

$$\Delta w_{pr} = \alpha(r_p - O_p) h'_p(I_p) O_i = \alpha \delta_p O_i \quad (62-3. 9)$$

۷۴۳ ۹. روش های نظریه تصمیم ۳۰.

که جزء خطاب عبارتست از

$$\delta_p = (r_p - O_p)h'_p(I_p). \quad (63-3. 9)$$

به استثنای r_p ، تمام اجزای معادلات (۹. ۶۲-۳) و (۹. ۶۳-۳) یا معلوم هستند یا می توان آنها را در شبکه مشاهده کرد. جزء r_p در لایه های داخلی هیچ مفهومی ندارد زیرا پاسخ مطلوب هر گره 'داخلی' بر حسب عضویت الگورانمی دانیم. تنها می توان پاسخ مطلوب گره های لایه 'خروجی' را، که دسته بندی نهایی الگوها در آن انجام می شود، تعیین کرد. اگر این اطلاعات را در مورد گره های داخلی داشته باشیم، دیگر به لایه های بعد نیازی نیست. بنابراین باید راهی بیابیم که δ_p را بر حسب کمیات معلوم یا کمیاتی که می توان در شبکه مشاهده کرد، بیان کنیم.

با برگشت به معادله (۹. ۵۷-۳)، جزء خطاب در لایه P را به صورت

$$\delta_p = -\frac{\partial E_p}{\partial I_p} = -\frac{\partial E_p}{\partial O_p} \frac{\partial O_p}{\partial I_p}. \quad (64-3. 9)$$

می نویسیم. همانند قبیل، بدون هیچ مشکلی جزء $O_p/\partial I_p$ را به صورت زیر بیان می شود:

$$\frac{\partial O_p}{\partial I_p} = \frac{\partial h_p(I_p)}{\partial I_p} = h'_p(I_p) \quad (65-3. 9)$$

از آنجا که می توان I_p را مشاهده کرد، عبارت فوق وقتی معلوم می شود که تابع h_p تعیین شود. جزئی که r_p را تولید کرد، $O_p/\partial I_p$ بود، بنابراین این جزء باید به طریقی که حاوی r_p نباشد، بیان شود. با استفاده از قاعده 'زنگیری'، مشتق را به صورت

$$\begin{aligned} -\frac{\partial E_p}{\partial O_p} &= -\sum_{q=1}^{N_Q} \frac{\partial E_p}{\partial I_q} \frac{\partial I_q}{\partial O_p} = \sum_{q=1}^{N_Q} \left(-\frac{\partial E_p}{\partial I_q} \right) \frac{\partial}{\partial O_p} \sum_{p=1}^{N_P} w_{qp} O_p \\ &= \sum_{q=1}^{N_Q} \left(-\frac{\partial E_p}{\partial I_q} \right) w_{qp} \\ &= \sum_{q=1}^{N_Q} \delta_q w_{qp} \end{aligned} \quad (66-3. 9)$$

می نویسیم که آخرین مرحله از معادله (۹. ۵۶-۳) نتیجه می شود. با جای گذاری معادلات (۹. ۶۵-۳) و (۹. ۶۶-۳) در معادله (۹. ۶۴-۳)، عبارت مطلوب برای δ_p بدست می آید:

$$\delta_p = h'_p(I_p) \sum_{q=1}^{NQ} \delta_q w_{qp} \quad (67-3.9)$$

اکنون می توان عامل δ را محاسبه کرد، زیرا تمام اجزای آن معلوم هستند. بنابراین معادلات (۶۷-۳.۹) و (۶۷-۳.۹) به طور کامل قاعده آموزش لایه P را تعیین می کنند. اهمیت معادله (۶۷-۳.۹) این است که δ را از روی کمیات δ_q ، که در لایه P بعد از P محاسبه شدند، محاسبه می کند. پس از این که جزو خطوط و وزن های لایه P محاسبه شدند، به طریق مشابه می توان با استفاده از این کمیات، خطوط و وزن های لایه P را نیز محاسبه کرد. به بیان دیگر، با شروع از خطای در لایه خروجی، راهی برای پس انتشار خطای سمت ورودی شبکه پیدا کرده ایم. می توانیم روال آموزش را به صورت زیر جمع بندی کرده و تعیین دهیم. برای هر دو لایه دلخواه K و J ، که K لایه P قبل از J است، با استفاده از معادله

$$\Delta w_{jk} = \alpha \delta_j O_k \quad (68-3.9)$$

وزن های w_{jk} را، که اتصالات بین این دو لایه را تغییر می دهند، محاسبه می کنیم. اگر J لایه خروجی باشد، δ عبارتست از:

$$\delta_j = (r_j - O_j) h'_j(I_j). \quad (69-3.9)$$

اگر J لایه داخلی و P لایه بعد آن (سمت راستش) باشد، آنگاه δ برای $N, 2, \dots, 1 = j$ با معادله زیر به دست می آید:

$$\delta_j = h'_j(I_j) \sum_{p=1}^{Np} \delta_p w_{jp} \quad (70-3.9)$$

با استفاده از تابع فعالیت معادله (۶۹-۳.۹) و (۶۹-۳.۵) بازای $1 = \theta_0$ خواهیم داشت:

$$h'_j(I_j) = O_j(1 - O_j) \quad (71-3.9)$$

که در این مورد معادلات (۶۹-۳.۹) و (۶۹-۳.۹) به صورت های جالب تری بیان می شوند که برای لایه خروجی به صورت:

۴. ۳. روش های نظریه ' تصمیم ۷۲۵

$$\delta_i = (r_i - O_i)O_i(1 - O_i) \quad (72-3. 9)$$

و برای لایه های داخلی به صورت:

$$\delta_i = O_i(1 - O_i) \sum_{p=1}^{N_p} \delta_p w_{ip} \quad (73-3. 9)$$

می باشد. در هر دو معادله (9. ۹)، (73-۳)، (72-۳) و (9. ۹)، (73-۳)، $N_j = j$ می باشد.

معادلات (9. ۹)، (70-۳) - (68-۳) قاعده ' دلتای تعمیم یافته برای آموزش شبکه ' عصبی پیشخور چند لایه ' شکل ۹. ۱۶ را تشکیل می دهند. این فرآیند با مجموعه ' دلخواهی از وزن های سراسر شبکه (که همه با هم برابر نیستند) شروع می شود. آنگاه اجرای قاعده ' دلتای تعمیم یافته در هر مرحله ' تکرار شامل دو مرحله پایه است. در مرحله اول، یک بردار آموزشی به شبکه داده می شود و این بردار در لایه های شبکه متشر می شود تا برای هر گره، خروجی O را تولید کند. آنگاه خروجی های O مربوط به گره های لایه خروجی با پاسخ های مطلوب، r ، مقایسه می شوند تا اجزای خطأ، δ ها، تولید شوند. مرحله دوم حرکت برگشتی از درون شبکه است که در خلال آن سیگنال خطای مناسب به هر گره داده می شود و تغییرات وزن متناظر با آن خطاهای انجام می شوند. همان طور که قبلاً با مقداری تفصیل بیان شد، این روال به وزن های بایاس θ نیز اعمال می شود. با این وزن های نیز مانند وزنی که ورودی واحد به اتصال جمع کننده ' یک گره ' شبکه را تغییر می دهد، رفتار می شود.

روال معمول این است که خطای شبکه، و نیز خطاهای مرتبط با هر یک از الگوهای ورودی را دگیری کنیم. در یک آموزش موفق، با افزایش تعداد تکرارها خطای شبکه کاهش می یابد و مسأله به مجموعه ' پایداری از وزن ها' که با آموزش بیشتر، تنها نوسانات کوچکی در آنها باقی می ماند، همگرا می شود. روش معمول برای تعیین این که آیا یک الگو در خلال آموزش بدرستی دسته بندی شده است یا نه، این است که مشاهده کنیم آیا پاسخ گره ' خروجی مربوط به دسته عضویت آن الگو 'فعال' و همزمان خروجی سایر گره ها 'غیرفعال' شده است یا نه.

بعد از این که سامانه آموزش دید، الگوهار ابا استفاده از عوامل تعیین شده در خلال آموزش، دسته بندی می کند. در حالت عادی، تمام عملیات بازخورد^۱ قطع هستند. آنگاه هر الگوی ورودی در لایه های

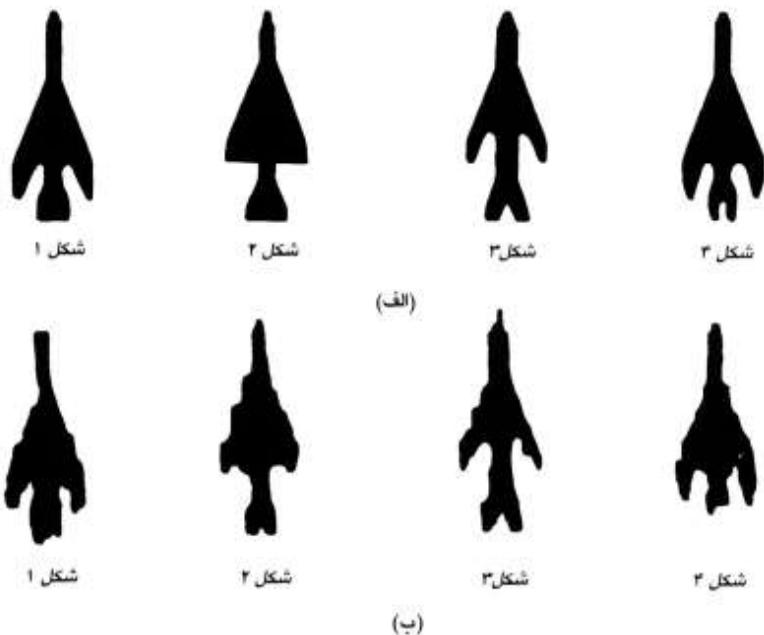
1) Bias weights

2) Feedback

شبکه پیش می رود، و الگو به عنوان عضو دسته یکتایی که گره 'خروجی' متناظر با آن فعال شود، دسته بندی می شود. اگر بیش از یک خروجی فعال شود، یا همه 'خروچی ها' غیرفعال شوند، یا می توان دسته بندی را نادرست اعلام کرد یا این که الگورابه دسته ای که گره 'خروچی' متناظر با آن بالاترین مقدار عددی را دارد، منتب کرد.

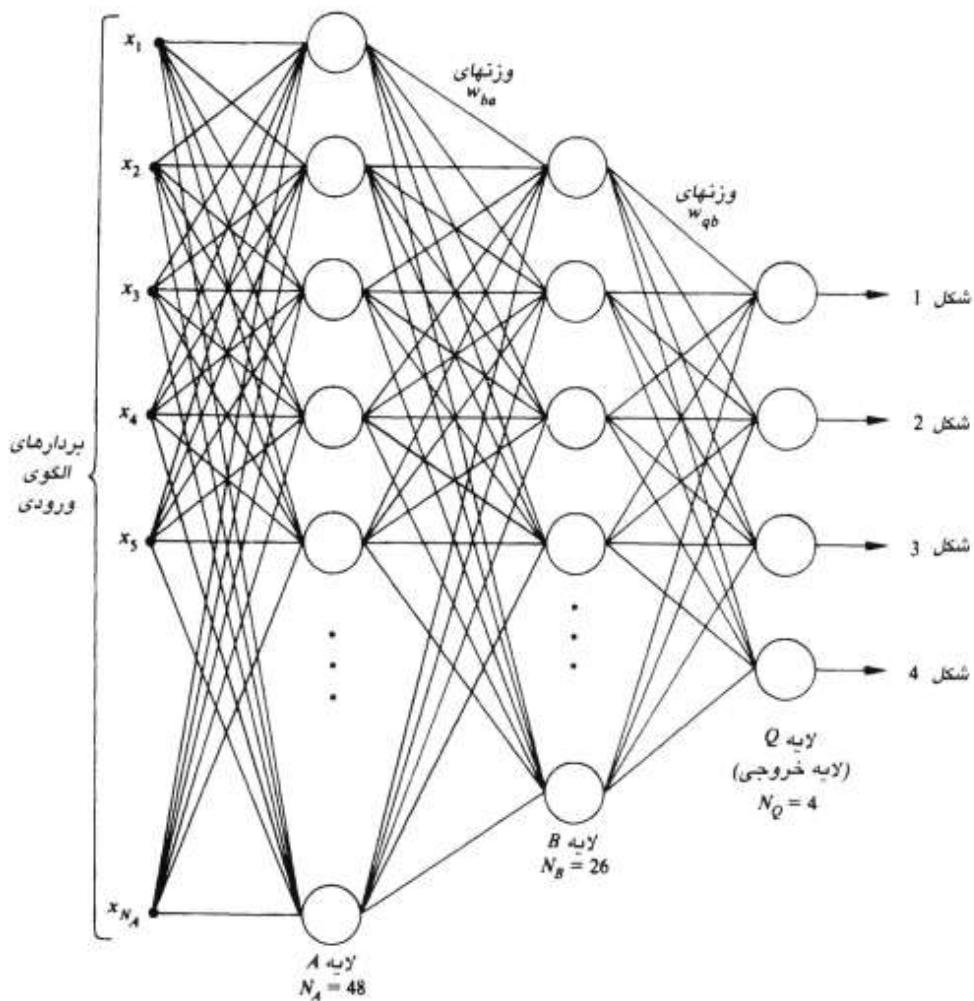
مثال: اکنون نشان می دهیم که چگونه شبکه 'عصبی' به صورت ساختار شکل ۹.۱۶، آموزش داده شد تا بتواند چهار تصویر موجود در شکل ۹.۱۸(الف)، و نیز نسخه های نویزدار آنها که در شکل ۹.۱۸(ب) دیده می شوند، را از هم تشخیص دهد.

بردارهای الگو با محاسبه 'نمایه های تراز شده' این اشکال (بخش ۸.۱.۳) را بینید) و سپس برداشتی ۴۸ نمونه هم فاصله از هر نمایه تولید شدند. بردارهای ۴۸ بعدی حاصل ورودی های شبکه 'عصبی' پیشخور سه لایه شکل ۹.۱۹ بودند. تعداد گره های نزونی لایه اول ۴۸ انتخاب شد، که متناظر با ابعاد بردارهای الگوی ورودی است. چهار نزون لایه سوم (خروچی) متناظر با تعداد دسته های الگو



شکل ۹.۱۸ (الف) شکل های مرجع و (ب) شکل های نویزدار نویعی که در آموزش شبکه 'عصبی' شکل ۹.۱۹ استفاده شدند. (از Gupta و همکارانش [1990])

۴. ۳. روش های نظریه' تصمیم ۷۲۷



شکل ۱۹.۹ شبکه عصبی سه لایه استفاده شده برای تشخیص تصاویر شکل ۹.

است، و تعداد نرون های لایه میانی به طور تجربی، ۲۶ (متوجه تعداد نرون های لایه های ورودی و خروجی) تعیین شد. هیچ قاعده مشخصی برای تعیین تعداد گره های لایه های داخلی شبکه عصبی وجود ندارد، بنابر این این تعداد معمولاً بر اساس تجربه قبلی انتخاب می شود، و یا این که در ابتدا بدله خواه انتخاب و سپس با آزمون اصلاح می گردد. در این مورد، چهار گره لایه خروجی از بالا به پایین، به ترتیب بیانگر دسته های $\omega_j = 1, 2, 3, 4$ هستند. پس از این که ساختار شبکه تعیین شد، می بایست توابع فعالیت هر واحد و لایه انتخاب شوند. تمام توابع فعالیت طوری انتخاب شدند که در معادله

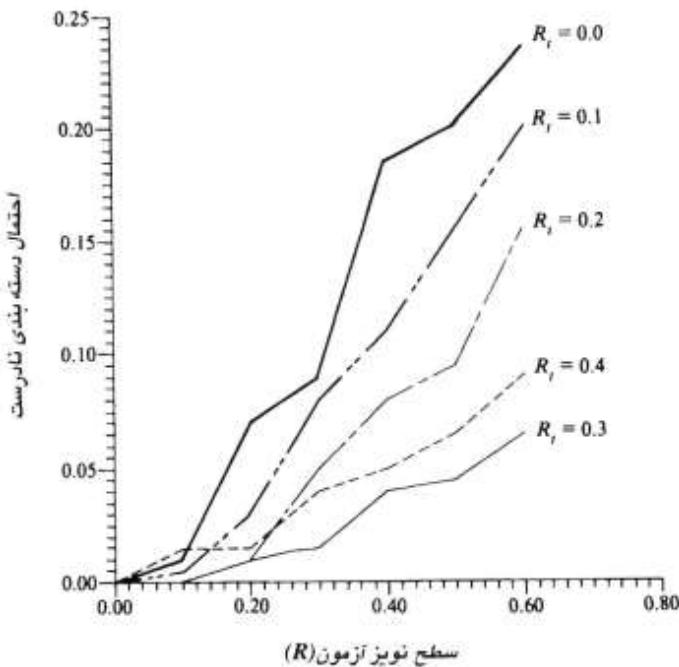
۹ . ۳-۵۰) صدق کنند. تابدین و سیله، طبق شرح قبلی، معادلات (۹ . ۳-۷۲) و (۹ . ۳-۷۳) نیز برقرار باشند.

فرآیند آموزش به دو قسمت تقسیم شد. در اولین قسمت، وزن ها با مقادیر تصادفی کوچکی بامانگین صفر مقدار اولیه داده شدند، و آنگاه شبکه با بردارهای الگوی متناظر با نمونه های بدون نویز نظیر تصاویر موجود در شکل ۹ . ۱۸(الف)، آموزش داده شد. در خلال آموزش، گره های خروجی پاییده می شدند.^۱ وقتی برای هر الگوی آموزشی از دسته^۲، عناصر لایه 'خروجی برای شبکه، تصاویر نمونه' هر چهار دسته را آموخته است. به عبارت دیگر، برای هر الگو از دسته^۳، عیوبیت گره خروجی متناظر با آن دسته فعال ($0.95 \geq O_i \neq q$) و در همان زمان خروجی سایر گره ها غیرفعال ($O_i = 1, 2, \dots, N_0$) می بودند.

قسمت دوم آموزش با نمونه های نویز داری که به صورت زیر تولید شدند، انجام شد. به هر پیکسل منحنی^۴ در یک شکل بدون نویز، دو عدد P و $R = 1 - P$ متناسب شد که P احتمال باقیماندن آن پس از اعمال نویز در مختصات اولیه و $R = 1 - P$ احتمال انتساب تصادفی آن به مختصات یکی از هشت پیکسل همسایه اش بود. با کاهش P (یعنی افزایش R) درجه 'نویز افزایش یافت. دو مجموعه داده 'نویز دار تولید شدند. اولین مجموعه دارای ۱۰۰ الگوی نویز دار از هر دسته بود که با تغییر R از ۰.۱ تا ۰.۶ تولید شده بودند. در مجموع ۴۰۰ الگو به دست آمد. این مجموعه، که مجموعه 'آزمون^۵' خوانده می شود، برای تعیین کارایی سامانه آموزش دیده، استفاده شد.

برای آموزش سامانه با داده های نویز دار، چند مجموعه 'نویز دار تولید شدند. اولین مجموعه مشکل از ۱۰ نمونه برای هر دسته بود، که با استفاده از $R = 0$ تولید شده بودند. R دلالت بر مقداری از R دارد که برای تولید داده های آموزشی استفاده شد. با شروع از بردارهای وزن بدست آمده در قسمت اول (بدون نویز) آموزش، سامانه امکان یافت که با مجموعه داده 'جدید تحت آموزش قرار گیرد. چون $R = 0$ به مفهوم عدم وجود نویز است، این آموزش مجدد، تعمیمی از آموزش بدون نویز اولیه بود. پس از جایگزینی وزنهای حاصل از روش فوق در شبکه و ارائه مجموعه داده های آزمون به آن، نتایج به صورت منحنی با برچسب $R = 0$ در شکل ۹ . ۲۰ بدست آمد. با تقسیم تعداد الگوهایی که به صورت نادرست دسته بندی

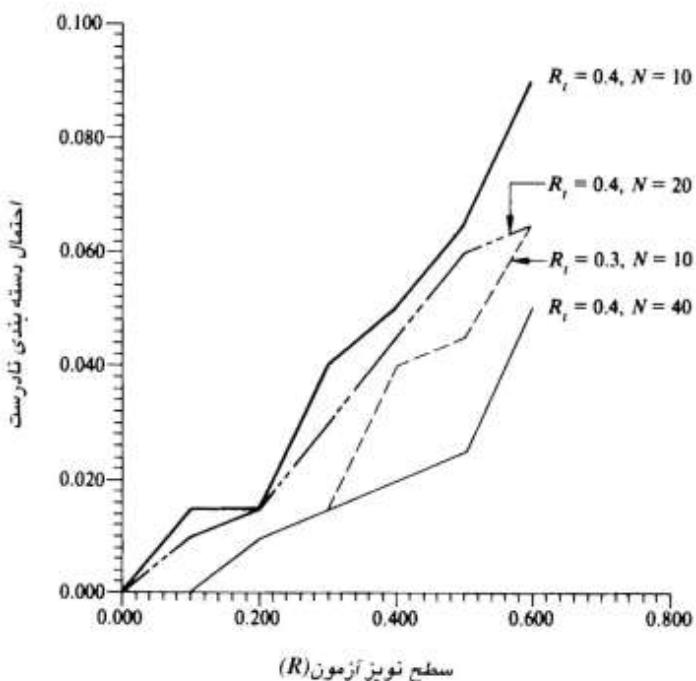
۷۲۹ ۳. ۴ روش های نظریه ' تصمیم



شکل ۹ ۲۰. ۹ کارآئی شبکه عصبی به صورت تابعی از سطح نویز. (از Gupta و همکارانش (1990)).

شده اند، بر تعداد الگوهای آزموده شده، احتمال دسته بندی نادرست بدست می آید، که معمولاً به عنوان معیاری از کارآئی شبکه استفاده می شود.

پس از آن با شروع از بردارهای وزن آموخته شده در حالت $R_i = 0$ ، سامانه با مجموعه داده های نویزدار حالت $R_i = 0.1$ مجدداً تحت آموزش قرار گرفت. سپس کارآئی تشخیص با دادن نمونه های آزمون به سامانه و استفاده از بردارهای وزن جدید تعیین شد. شکل ۹. ۲۰ نتایج حاصل از ادامه این آموزش و آزمون مجدد بازای $R_i = 0.2, 0.3, 0.4$ را نشان می دهد. همانطور که انتظار می رفت، اگر سامانه بخوبی آموزش دیده بود، احتمال دسته بندی نادرست الگوهای مجموعه آزمون با افزایش مقدار R_i کاهش می یافت، زیرا در مقادیر بالاتر R_i سامانه با داده های نویزدارتری آموزش دیده بود. در شکل ۹. ۲۰ نتیجه بازای $R_i = 0.4$ یک استثناء است. علت آن تعداد کم نمونه های استفاده شده برای آموزش سامانه است. یعنی، با تعداد نمونه های استفاده شده، شبکه قادر نبود که خود را به اندازه 'کافی با تغییرات بیشتر شکل ناشی از سطوح بالاتر نویز تطبیق دهد. درستی یا نادرستی این فرضیه با نتایج موجود در شکل ۹. ۲۱ معین می شود، که در آن با افزایش تعداد نمونه های آموزشی، احتمال دسته بندی نادرست کمتر



شکل ۲۱.۹ بهبود کارآئی بازای $R_t = 0.4$ با افزایش N ، تعداد الگوهای آموزش. به تفاوت مقیاس این نمودار با نمودار شکل ۹.۲۰ توجه کنید. (از Gupta و همکارانش (1990) با تأثیرگذاری)

شده است. در شکل ۹.۲۱، منحنی $R_t = 0.3$ از شکل ۹.۲۰ به عنوان مرجعی برای مقایسه نشان داده شده است.

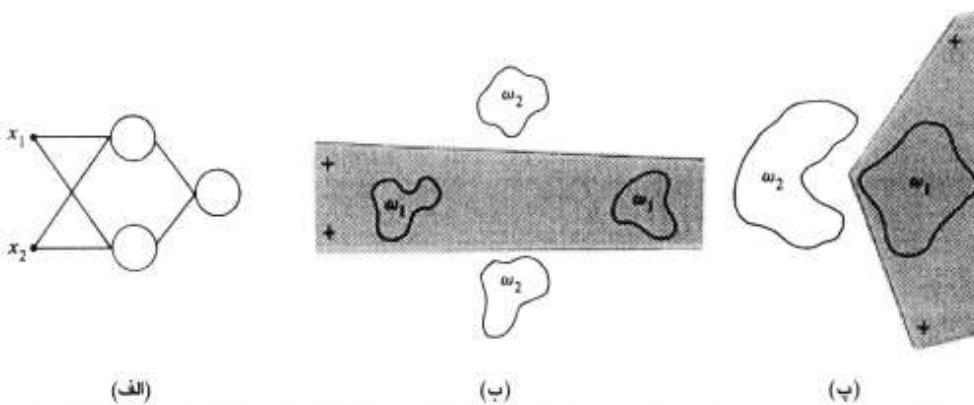
نتایج قبل نشان می دهند که شبکه عصبی سه لایه بعد از یک آموزش ساده می تواند شکل های تخریب شده با نویز را شناسایی کند. حتی وقتی با داده های بدون نویز آموزش داده شد ($R_t = 0$) در شکل ۹.۲۰، سامانه توائنت نزدیک به ۷۷ درصد داده های بسیار تخریب شده با نویز ($R_t = 0.6$) در شکل ۹.۲۰ را به درستی تشخیص دهد. وقتی سامانه با داده های نویز دارتری ($R_t = 0.3, 0.4$) آموزش دید، میزان تشخیص صحیح تا نزدیک ۹۹ درصد افزایش یافت. ذکر این نکته مهم است که روش آموزش بدین ترتیب بود که با افزایش های کوچک و منظم نویز، قدرت دسته بندی سامانه افزایش یافت. وقتی طبیعت نویز معلوم باشد، این روش برای بهبود خواص همگرایی و پایداری شبکه عصبی در خلال آموزش، ایده آل است.

پیچیدگی سطوح تصمیم. همان طور که قبلاً نشان دادیم، پرسپترون یک لایه، سطح تصمیم ابر صفحه ای دارد. اکنون وقت آن است که پرسیم: شکل سطوح تصمیم شبکه 'چند لایه ای، نظری مدل شکل

۷۳۱ ۹. ۳. روش های نظریه تصمیم

۹. ۱۶. چیست؟ در توضیحات زیر روش می شود که شبکه 'سه لایه' قادر است سطوح تصمیمی با پیچیدگی دلخواه که از ابرصفحه های متقاطع تشکیل شده اند، ایجاد کند.

به عنوان شروع، شبکه 'دو لایه' دو ورودی شکل ۹. ۲۲(الف) را در نظر می گیریم. با دو ورودی، الگوها دو بعدی هستند، و بنابراین هر گره در لایه 'اول شبکه'، یک خط در فضای دو بعدی ایجاد می کند. خروجی های فعال و غیرفعال این دو گره را به ترتیب با ۰ و ۱ نشان می دهیم. فرض می کنیم که خروجی ۱ دلالت بر این دارد که بردار ورودی متناظر با گره ای در لایه 'اول'، در طرف مثبت خط قرار دارد. آنگاه ترکیبات ممکن خروجی گره های لایه 'اول' که به تنها گره 'لایه' دوم وارد می شوند، عبارتند از: $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$. اگر دو ناحیه تعریف کنیم، یک ناحیه برای دسته ω_1 که در طرف مثبت هر دو خط قرار دارد، و دیگری برای دسته ω_2 که در سایر نواحی قرار دارد، گره 'خروچی' می تواند با یک AND منطقی ساده هر الگوی خروجی را به عنوان عضو یکی از این دو ناحیه دسته بندی کند. به عبارت دیگر، تنها وقتی که هر دو خروجی لایه 'اول' باشند، گره 'خروچی' پاسخ ۱ که دلالت بر دسته ω_1 دارد، می دهد. عمل AND را می توان با یک گره 'عصیبی' با شکل تشریح شده در قبل، انجام داد که باید مقداری در فاصله نیمه باز $[1, 2)$ انتخاب شود. بنابراین اگر پاسخ های خروجی لایه 'اول' را ۰ و ۱ فرض کنیم، تنها وقتی مجموع دو خروجی لایه 'اول' بزرگتر از ۱ باشد، پاسخ گره 'خروچی' فعال خواهد بود که دال بر دسته ω_1 می باشد. اشکال ۹. ۲۲(ب) و (پ) نشان می دهند که چگونه شبکه 'شکل ۹. ۲۲(الف)' می تواند دو دسته الگو را که نمی توان با یک سطح از هم جدا کرد، پدرستی از هم تشخیص دهد.



شکل ۹. ۲۲. (الف) شبکه عصیبی پیشخور دو لایه' دو ورودی؛ (ب) و (پ) مثال هایی از مزه های تصمیم که می توان برای این شبکه تعیین کرد

اگر تعداد گره های لایه اول به سه افزایش باید، شبکه 'شکل ۹ . ۲۲ (الف) مرز تصمیمی شامل سه خط تقاطع ایجاد می کند. برای این که دسته '۱ در طرف مثبت هر سه خط قرار داشته باشد، ناحیه 'محدبی لازم است که در میان سه خط محصور شده باشد. در واقع می توان با افزایش تعداد گره های لایه 'اول شبکه 'عصبی دو لایه، ناحیه محدب باز یا بسته 'دلخواه را ایجاد کرد.

قدم منطقی بعدی افزایش تعداد لایه های سه لایه است. همانند گذشته در این مورد نیز گره های لایه 'اول خطوط را ایجاد می کنند. آنگاه گره های لایه 'دوم عمل AND را انجام می دهند تا بین وسیله از خطوط مختلف چندین ناحیه ایجاد کنند. گره های لایه 'سوم به نواحی مختلف، دسته 'عضویت متنسب می کنند. برای نمونه فرض کنید که دسته '۱ مشکل از دو ناحیه 'جزا باشد که هر یک از آنها در مجموعه 'متقاوئی از خطوط محصور است. آنگاه دو تا از گره های لایه 'دوم برای نواحی مربوط به یک دسته الگو است. وقتی یکی از این دو گره فعال شود، باید یکی از گره های خروجی بتواند رخداد آن دسته را اعلام کند. با فرض این که در لایه 'دوم حالات فعال و غیرفعال به ترتیب با ۱ و ۰ نشان داده شوند، این توانایی با تغییر گره های خروجی شبکه طوری که بتواند عمل OR منطقی را انجام دهدن، ایجاد می شود. با استفاده از گره های عصبی با شکلی که قبل از تشریح شد، این کار را با انتخاب مقدار θ از فاصله '۰,۱] انجام می دهیم. حال هر گاه لاقل یکی از گره های لایه 'دوم مرتبط با آن گره 'خروجی، فعال شود (خروجی ۱ بدهد)، گره 'منتظر در لایه 'خروجی فعال می شود، که بیان می کند الگوی پردازش شده به دسته 'مرتبط با آن گره تعلق دارد.

شکل ۹ . ۲۳ توضیحات قبلی را جمع بندی می کند. ردیف سوم نشان می دهد که پیچیدگی نواحی تصمیم ایجاد شده با شبکه 'عصبی سه لایه دلخواه است. یک مشکل جدی که معمولاً در ساختار لایه 'دوم رخ می دهد این است که باید به ترکیبات گوناگون مرتبط با دسته های ویژه به درستی پاسخ دهد. دلیل آن این است که خطوط، دقیقاً در محل تلاقی با سایر خطوط متوقف نمی شوند بلکه معمولاً مقداری ادامه می یابند، و در نتیجه الگوهای یک دسته ممکن است در هر دو طرف خطوط فضای الگو واقع شوند. از لحظه عملی لایه دوم ممکن است در تعیین خطوطی که باید برای هر دسته 'الگوی معین، AND شوند، مشکل داشته باشد - یا حتی ممکن است اصلاً نتواند این عمل را انجام دهد. مطرح کردن OR انحصاری در ستون سوم شکل ۹ . ۲۳ ناشی از این واقعیت است که اگر الگوهای ورودی دو بعدی و دودویی

۹. روش های نظریه تصمیم ۷۳۳

ساختار شبکه	نوع نواحی تصمیم	جواب مسئله انحصاری	دسته هایی با نواحی شبکه ای شده	کلی ترین شکل های سطح تصمیم
یک لایه	یک ابر صفحه			
دو لایه	نواحی محدب باز یا بسته			
سه لایه	دلخواه (پیچیدگی به وسیله تعداد گره ها محدود نی شود)			

شکل ۲۳.۹ انواع نواحی تصمیم که می توان با شبکه های پیشخور یک و چند لایه که دارای یک و دو لایه از واحد های مخفی و دورودی هستند، تشکیل داد. (گرفته شده از [Lippman 1987]).

باشد، تنها چهار الگوی متفاوت می توان تشکیل داد. اگر الگوها طوری مرتب شوند که دسته ' w ' متشکل از الگوهای $(1, 0), (0, 1)$ و دسته ' w ' متشکل از الگوهای $(1, 1), (0, 0)$ باشد، دسته عضویت الگوها در این دو دسته با تابع منطقی OR انحصاری (XOR) به دست می آید، یعنی تنها وقتی یکی و فقط یکی از دو متغیر ۱ باشد، خروجی ۱ و در غیر این صورت ۰ است. بنابراین مقدار ۱ در خروجی XOR نشانه الگوهای دسته ' w ' و مقدار ۰ نشانه الگوهای دسته ' w ' است.

توضیحات قبلی با روش ساده ای به بعدی تعمیم داده می شود: به جای خطوط، با ابر صفحه ها کار می کنیم. شبکه 'تک لایه'، یک ابر صفحه ایجاد می کند. شبکه 'دو لایه'، نواحی محدب دلخواهی که شامل تقاطع های ابر صفحه ها هستند، ایجاد می کند. شبکه 'سه لایه'، سطوح تصمیمی با پیچیدگی دلخواه ایجاد می کند. تعداد گره های استفاده شده در هر لایه، پیچیدگی دو مورد آخر را تعیین می کند. تعداد دسته ها در مورد اول به دو محدود می شود. در دو مورد دیگر، تعداد دسته ها دلخواه است، زیرا تعداد گره های خروجی را می توان طوری انتخاب کرد که مناسب مسئله تحت

بررسی باشد.

با در نظر گرفتن توضیحات قبلی، منطقی است که بپرسیم: چرا بعضی افراد به مطالعه 'شبکه های عصبی' باشند از سه لایه علاوه دارند؟ در حالی که شبکه 'عصبی' سه لایه می تواند سطوح تصمیم با پیچیدگی دلخواه را ایجاد کند. جواب در روش آموزش شبکه سه لایه قرار دارد. قاعده 'آموزش شبکه' شکل ۹.۱۶ معیار خطابی را کمینه می کند اما چیزی درباره 'چگونگی ارتباط گروه های ابر صفحه ها با گره های مشخصی در لایه 'دوم شبکه' سه لایه از نوعی که قبل از تشریح شد، نمی گوید. در واقع، در مبحث شبکه های عصبی چگونگی انجام مصالحه بین تعداد لایه ها و تعداد گره های هر لایه بدون تصمیم گیری قطعی باقی می ماند. در عمل معمولاً این مصالحه با سعی و خطا یا با تجربه 'قبلی درباره 'حوزه' مسأله داده شده، انجام می شود.

۴.۹ روش های ساختاری

روش هایی که در بخش ۹.۳ ارائه گردیدند، به صورت کمی به الگوها می پردازند و روابط ساختاری نهفته در شکل الگورامورد توجه قرار نمی دهند، اما روش های ساختاری تشریح شده در این بخش، به دنبال دستیابی به تشخیص الگو با تمرکز دقیق بر این نوع روابط هستند.

۱۰.۹ اعداد شکلی تطبیقی

روالی تقریباً مشابه با مفهوم حداقل فاصله که در بخش ۹.۳.۱ در مورد بردارهای الگو معرفی شد، را می توان برای مقایسه 'مرزهای ناحیه' که بر حسب اعداد شکلی توصیف شده اند، بیان کرد. طبق شرح ارائه شده در بخش ۸.۲.۲، درجه 'تشابه' k ، بین دو مرز ناحیه ای (شکل) A و B ، به عنوان بزرگترین مرتبه ای که هنوز اعداد شکلی شان یکسان هستند، تعریف می شود. مثلاً در موردی که اعداد شکلی مرزهای بسته با رمزهای زنجیره ای چهار جهتی بیان شده اند، A و B درجه 'تشابه' k دارند اگر

$$s_{k+1}(A) \neq s_{k+1}(B), s_k(A) = s_k(B), \dots, s_4(A) = s_4(B)$$

$$s_{k+4}(A) \neq s_{k+4}(B)$$

فاصله^{۱)} بین دو شکل A و B به صورت معکوس درجه 'تشابه' شناسان تعریف می شود:

1) Degree of similarity

2) Distance

۷۳۵ ۴.۹ روش های ساختاری

$$D(A, B) = \frac{1}{k}. \quad (1-4.9)$$

این فاصله خواص زیر را دارد:

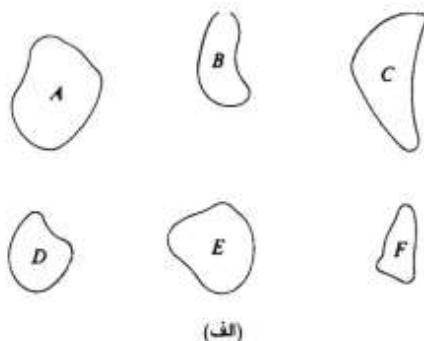
$$D(A, B) \geq 0$$

$$D(A, B) = 0 \quad \text{اگر} \quad A = B \quad (2-4.9)$$

$$D(A, C) \leq \max[D(A, B), D(B, C)].$$

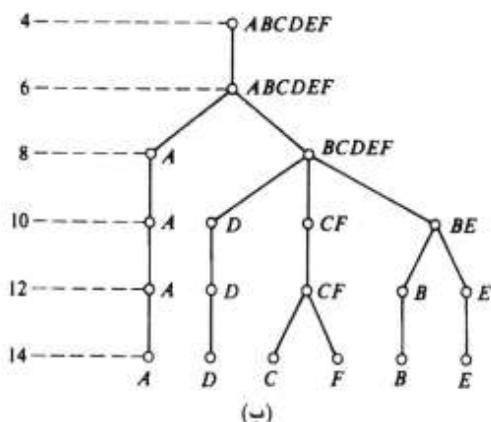
از k یا D می توان برای مقایسه' دو شکل استفاده کرد. اگر درجه' شباهت استفاده شود، هر چه k بزرگتر باشد، شکل ها شبیه تر هستند (توجه کنید که برای اشکال کاملاً یکسان، k بی نهایت است). وقتی معیار فاصله استفاده شود، عکس آن درست است.

مثال: فرض کنید که شکل F را داریم و می خواهیم نزدیکترین تطبیقش را در مجموعه ای از پنج شکل دیگر که در شکل ۲۴.۹(الف) دیده می شوند، بیابیم. این مسأله مشابه داشتن پنج شکل



(الف)

درجه



(ب)

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
<i>A</i>	∞	6	6	6	6	6
<i>B</i>		∞	8	8	10	8
<i>C</i>			∞	8	8	12
<i>D</i>				∞	8	8
<i>E</i>					∞	8
<i>F</i>						∞

(ب)

شکل ۲۴.۹ (الف) شکل ها؛ (ب) درخت شباهت؛ (ب) ماتریس شباهت. (از . (Bribisca&Guzman[1980])

نمونه و تلاش برای پیدا کردن بهترین تطبیق با یک شکل ناشناخته داده شده است. جستجو را می توان با کمک درخت شباهت رسم شده در شکل ۹ . (ب) نمایش داد. ریشه 'درخت متناظر با کمترین درجه' ممکن شباهت است، که در این مثال چهار است. به استثنای شکل A که درجه 'شباهتش با سایر اشکال ۶ است، سایر شکل ها دست کم درجه 'شباهت ۸ نسبت به هم دارند. با ادامه 'درخت به پایین، می فهمیم که شکل D نسبت به سایر شکل ها درجه 'شباهت ۸ دارد، و به همین ترتیب الی آخر. شکل های F و C بیشتر با هم تطبیق می کنند و درجه 'شباهت بزرگتری نسبت به هر دو شکل دیگر دارند. از طرف دیگر، اگر A شکل ناشناخته ای می بود، تمام آتجه ما می توانستیم با استفاده از این روش بگوییم، این بود که ۸ نسبت به چهار شکل دیگر درجه 'شباهت ۶ دارد. همان طور که در شکل ۹ . (پ) دیده می شود، این اطلاعات را می توان به صورت ماتریس شباهت^۱ جمع بندی کرد.

□

۴.۳.۹ تطبیق رشته

فرض کنید که دو مرز ناحیه ای A و B به صورت رشته هایی رمز شده (بخش ۸ . ۵ را بینید) و به ترتیب با b_1, b_2, \dots, b_m و a_1, a_2, \dots, a_n نشان داده شوند. فرض کنید M بیانگر تعداد تطبیق های بین دو رشته باشد، به صورتی که اگر $a_k = b_k$ در محل k ام یک تطبیق وجود دارد. تعداد نمادهایی که با هم تطبیق نمی کنند، عبارتست از

$$Q = \max(|A|, |B|) - M \quad (۴-۴ . ۹)$$

که $\text{arg} Q$ طول (تعداد نمادهای) نمایش رشته ای آرگومان است. می توان نشان داد که اگر و تها اگر A و B یکسان باشند، $Q = 0$ (مسئله ۹ . ۲۱ را بینید).

یک معیار شباهت ساده بین A و B، نسبت

$$R = \frac{M}{Q} = \frac{M}{\max(|A|, |B|) - M} \quad (۴-۴ . ۹)$$

است. بنابراین در تطبیق کامل، R نامحدود است و وقتی هیچ یک از نمادهای A و B تطبیق نکنند (در این مورد $Q = 0$)، R صفر است. چون عمل تطبیق، نماد به نماد انجام می شود، نقطه 'شروع روی هر یک

1) Similarity matrix

۴.۹ روش های ساختاری ۷۳۷

از مرزها از نظر کاهش محاسبات مهم است. هر روشی که به همان نقطه 'شروع یا نزدیک آن تراز شود، مفید است و نسبت به تطبیق بدون فکر^۱، مزیت محاسباتی دارد. چرا که تطبیق بدون فکر از نقاطی دلخواه روی هر رشته شروع کرده، سپس هر یک از رشته هارا (با چرخش به دور خود) انتقال داده و برای هر انتقال معادله^۲ .۴-۴ را محاسبه می کند. بزرگترین مقدار R بهترین تطبیق را می دهد.

مثال: شکل های ۹ . ۲۵ (الف) و (ب) مرزهای نمونه از هر یک از دو دسته شیئی که با انطباق چندضلعی تقریب زده شده اند، را نشان می دهند (بخش ۸ . ۱ . ۲ را ببینید). اشکال ۹ . ۲۵ (پ) و (ت) به ترتیب تقریب های چندضلعی متناظر با مرزهای نمایش داده شده در شکل های ۹ . ۲۵ (الف) و (ب) هستند. این رشته ها با پیمایش محیط چندضلعی در جهت عقریه 'ساعت و محاسبه' زاویه 'داخلی بین اضلاع آن تشکیل شدند. زوایا به یکی از هشت نماد ممکن متناظر با فواصل^۳ ۴۵°، رمز شدند؛ یعنی، $0^\circ < \theta \leq 45^\circ$ ؛ $\alpha_1: 45^\circ < \theta \leq 90^\circ$ ؛ $\alpha_2: 90^\circ < \theta \leq 135^\circ$ ؛ $\alpha_3: 135^\circ < \theta \leq 180^\circ$ ؛ $\alpha_4: 180^\circ < \theta \leq 225^\circ$ ؛ $\alpha_5: 225^\circ < \theta \leq 270^\circ$ ؛ $\alpha_6: 270^\circ < \theta \leq 315^\circ$ ؛ $\alpha_7: 315^\circ < \theta \leq 360^\circ$ ؛ $\alpha_8: 360^\circ < \theta \leq 45^\circ$.

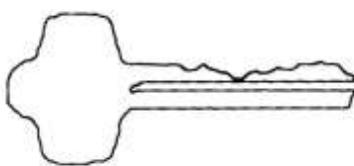
شکل ۹ . ۲۵ (ث) نتایج محاسبه 'معیار R برای پنج نمونه' شیئی ۱ (نسبت به خودشان) را نشان می دهد. درایه ها متناظر با مقادیر R هستند، و برای مثال، نماد ۱.۱ به رشته 'سوم از دسته شیئی ۱ اشاره دارد. شکل ۹ . ۲۵ (ج) نتایج مقایسه 'رشته های دسته شیئی دوم با خودشان را نشان می دهد. در نهایت شکل ۹ . ۲۵ (ج) جدول مقادیر R حاصل از مقایسه 'رشته های یک دسته با دسته' دیگر را نشان می دهد. توجه کنید که در اینجا تمام مقادیر R به طور چشمگیری کوچکتر از هر یک از درایه های دو جدول قبلی هستند. این نشان می دهد معیار R به درجه 'بالایی از تفکیک بین دو دسته شیئی دست یافته است. مثلاً اگر دسته 'عضویت رشته' a.1 نامعلوم باشد، کوچکترین مقدار R حاصل از مقایسه 'این رشته با رشته های نمونه (پیش نمونه) دسته' ۱ . ۱ خواهد بود. بر عکس بزرگترین مقدار R در مقایسه 'رشته' مذکور با رشته های دسته' ۲ . ۱ خواهد بود. این نتایج نشان می دهد که رشته 'a.1 عضوی از دسته شیئی ۱ است. این راه دسته بندی، مشابه دسته بندی کننده حداقل فاصله 'معرفی شده در بخش ۹ . ۳ . ۱ می باشد. □

۴.۹ روش های قالبی

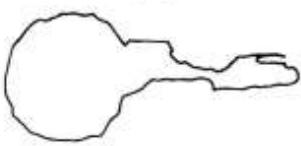
روش های قالبی^۴ از متد اول ترین روش ها برای حل مسائل تشخیص ساختاری هستند. نظریه حاکم بر تشخیص الگوی قالبی عبارت است از: تعیین مجموعه ای از الگوهای ابتدائی^۵ (بخش ۸ . ۵ را ببینید)،



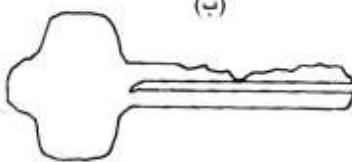
(الف)



(ب)



(ج)



(د)

<i>R</i>	1.a	1.b	1.c	1.d	1.e
1.b	16.0				
1.c	9.6	26.3			
1.d	5.07	8.1	10.3		
1.e	4.67	7.2	10.3	14.2	
1.f	4.67	7.2	10.3	8.5	23.7

(ه)

<i>R</i>	2.a	2.b	2.c	2.d	2.e
2.b	33.5				
2.c	4.75	5.8			
2.d	3.6	4.23	19.3		
2.e	2.83	3.25	9.17	18.3	
2.f	2.63	3.0	7.71	13.5	27.0

(ز)

<i>R</i>	1.a	1.b	1.c	1.d	1.e	1.f
2.a	1.24	1.50	1.32	1.47	1.55	1.48
2.b	1.18	1.43	1.32	1.47	1.55	1.48
2.c	1.02	1.18	1.19	1.32	1.39	1.48
2.d	1.02	1.18	1.19	1.32	1.39	1.40
2.e	0.93	1.07	1.08	1.19	1.24	1.25
2.f	0.89	1.02	1.02	1.14	1.11	1.18

(ج)

شکل ۲۵.۹ (الف) و (ب) مرزهای نمونه دودسته شیخ متفاوت؛ (پ) و (ت) تقریبهای چند ضلعی متناظر با آنها؛ (ه) و (ز) جداول *R*. (گرفته شده از [Sze & Yang 1981].)

تعیین مجموعه ای از قواعد به شکل گرامر^۱ که اتصالات بین الگوهای ابتدایی را قانونمند می کند، و تعیین یک تشخیص دهنده^۲ که خودکاره^۳ خوانده شده و ساختارش با مجموعه 'قواعد درون گرامر' معین می شود. در ابتدا گرامرهای رشته ای^۴ و خودکاره ها را در نظر می گیریم و سپس مبحث را به گرامرهای درختی و خودکاره های متناظر شان بسط می دهیم. یادآوری می شود که رشته ها و درختها تو صیف گرهای اصلی مورد استفاده در این کتاب برای تو صیف الگوهای ساختاری هستند.

تشخیص قالبی رشته ها

شرح زیر مبتنی بر این فرض است که نواحی یا اشیای مورد علاقه^۵ در تصویر، طبق شرح بخش ۵ . ۸، با استفاده از عناصر ابتدایی مناسب به شکل رشته بیان شوند.

گرامرهای رشته ای. فرض کنید که دو دسته ω_1 و ω_2 را داریم که الگوهایشان رشته هایی از عناصر ابتدایی هستند. هر عنصر ابتدایی را می توان به عنوان یک نماد معتبر در الفبای گرامر تغییر کرد. گرامر مجموعه ای از قواعد قالب^۶ است (بنابراین نام تشخیص قالبی را دارد) که تولید جمله های تشکیل شده از نمادهای الفبا را قانونمند می کند. مجموعه 'جملات تولید شده با گرامر G '، زبان آن گرامر خوانده می شود و با نماد $L(G)$ نشان داده می شود. بنابراین، جملات، رشته هایی از نمادها هستند (که نمایانگر الگوها می باشند)، و زبانها نیز متناظر با دسته های الگو هستند.

دو گرامر G_1 و G_2 را در نظر بگیرید که قواعد قالب آنها چنان است که G_1 تنها جملاتی را که متناظر با الگوهای دسته ω_1 هستند و G_2 تنها جملات متناظر با الگوهای دسته ω_2 را تولید می کند. پس از این که گرامرهایی با این خواص تعیین شدند، اجرای فرآیند تشخیص الگوی قالبی، ساده است. برای جمله ای که نمایانگر یک الگوی مجهول است، باید تعیین کنیم که این الگو در کدام زبان نمایانگر یک جمله 'معتبر' است. اگر جمله متعلق به $(G_1 \cup G_2)$ باشد، می گوییم که الگو از دسته ω_1 است. به طور مشابه اگر جمله در $L(G_2)$ معتبر باشد، گفته می شود که الگو از دسته ω_2 است. اگر جمله متعلق به هر دو زبان باشد، تصمیم یکتایی نمی توان گرفت. جمله ای که در هر دو زبان نامعتبر باشد، رد می شود.

وقتی بیش از دو دسته الگو موجود باشند، روش دسته بندی قالبی همانند روش فوق است، البته با این

1) Grammar

2) Recognizer

3) Automaton

4) String grammar

5) Syntax

تفاوت که گرامرهای بیشتری (لاقل یکی برای هر دسته) در فرآیند به کار می روند. برای دسته بندی، الگو به دسته Σ تعلق دارد اگر نمایانگر جمله^۱ معتبری از (G/L) باشد. همان طور که در بالا ذکر شد، اگر جمله ای به زیان های متفاوت متعلق باشد، تصمیم یکتاوی نمی توان گرفت. جمله ای که در تمام زبانها نامعتبر باشد، رد می شود.

برای کار با رشته ها، گرامری را به صورت چهار تابی

$$G = (N, \Sigma, P, S) \quad (5-4. 9)$$

تعريف می کنیم که

N = مجموعه^۲ محدودی از متغیرهای معروف به غیرپایانه ها^۱

Σ = مجموعه^۳ محدودی از ثابت های معروف به پایانه ها^۲

P = مجموعه^۴ قواعد بازنوسی که فرآوری ها^۵ خوانده می شوند

در N = نماد شروع^۶

لازم است که مجموعه های N و Σ جدا از هم^۷ باشند. در شرح بعدی، حروف بزرگ A, B, \dots, S ، حروف کوچک a, b, c, \dots ، در شروع الفانشانگر پایانه ها هستند. حروف کوچک v, w, x, y, z ، تا انتهای الفانشانگر رشته های پایانه ها هستند. حروف یونانی کوچک $\alpha, \beta, \theta, \dots$ ، نشانگر رشته های مرکب از پایانه ها و غیرپایانه ها هستند. جمله^۸ تهی^۹ (جمله^{۱۰} بدون نماد) با λ نشان داده می شود. در نهایت برای مجموعه V از نمادها، نشانه^{۱۱} V^* نشانگر مجموعه^{۱۲} تمام جمله هایی است که از عناصر V تشکیل شده اند.

گرامرهای رشته ای با شکل فرآوری هایشان مشخص می شوند. در تشخیص الگوی قالبی، گرامرهای منظم^{۱۳} و گرامرهای آزاد^{۱۴} مورد علاقه ویژه هستند. گرامرهای منظم تنها فرآوری هایی به شکل $aB \rightarrow aB$ یا $a \rightarrow A$ دارند، که A و B در Σ در^{۱۵} است. گرامرهای آزاد فرآوری هایی به شکل $\alpha \rightarrow A$ دارند، که A در N و α در مجموعه^{۱۶} Σ^* (N) است؛ یعنی α می تواند هر رشته^{۱۷} مرکب از پایانه ها و غیرپایانه ها، جز رشته تهی^{۱۸}، باشد.

1) Nonterminal

2) Terminal

3) Production

4) Starting symbol

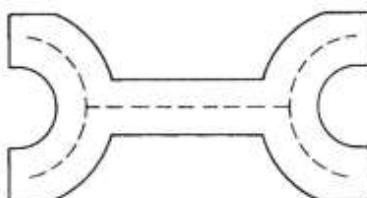
5) Disjoint

6) Empty sentence

7) Regular grammar

8) Context-free grammar

۴. ۹ روش های ساختاری ۷۲۱



(الف)



(ب)



(پ)

شکل ۴.۹ (الف) یک شیئ که با اسکلت نمایش داده شده است؛ (ب) عناصر ابتدایی (پ) ساختار تولید شده با استفاده از گرامر رشته‌ای معمولی.

مثال: فرض کنید که شیئ نشان داده شده در شکل ۹.۲۶(الف) با اسکلت نمایش بیان شود و عناصر ابتدایی نشان داده شده در شکل ۹.۲۶(ب) نیز برای توصیف ساختار این اسکلت‌ها (و مشابه آنها) تعریف شوند. گرامر $G = (N, \Sigma, P, S)$ ، $N = \{A, B, S\}$ ، $\Sigma = \{a, b, c\}$ ، $S = (N, \Sigma, P, S)$ را در نظر بگیرید که پایانه‌های a ، b ، و c متناظر با عناصر ابتدایی نشان داده شده در شکل ۹.۲۶(ب) هستند. همان‌طور که قبلاً بیان شد، S نماد شروع می‌باشد که رشته‌های $L(G)$ از آن تولید می‌شوند. برای نمونه، یک بار استفاده از اولین فرآوری و سپس دوبار کاربرد فرآوری دوم، $aa \Rightarrow abA \Rightarrow abbA$ را نتیجه می‌دهد، که (\Rightarrow) بیانگر یک اشتقاق^۱ رشته‌ای با شروع از S و با استفاده از فرآوری‌های مجموعه P است. اولین فرآوری، بازنویسی S به صورت aa و دومین فرآوری بازنویسی A به صورت ba را ممکن می‌سازد. با یک غیرپایانه در رشته $abbA$ ، می‌توانیم کار را ادامه دهیم. مثلاً دوبار دیگر کاربرد فرآوری دوم، که به دنبال آن یک بار فرآوری سوم و یک بار فرآوری چهارم به کار رود، رشته $abbbbabc$ را می‌دهد، که متناظر با ساختار نشان داده

1) Derivation

شده در شکل ۹ . ۲۶(ب) است. پس از کاربرد فرآوری چهارم هیچ غیرپایانه ای باقی نمی ماند، بنابراین وقتی این فرآوری استفاده شود، کار پایان می یابد. زبان تولید شده با قواعد این گرامر $L(G) = \{ab^n c \mid n \geq 1\}$ است، که "n شانه" n بار تکرار نماد b می باشد. به عبارت دیگر، G تنها قادر به تولید اسکلت هایی به شکل ۹ . ۲۶(ب) اما با طول دلخواه است.

کاربرد معانی. در مثال قبل فرض کردیم که اتصال بین عناصر ابتدایی تنها در نقاط نشان داده شده در شکل ۹ . ۲۶(ب) رخ می دهد. در وضعیت های پیچیده تر، قواعد اتصال و همچنین اطلاعات راجع به سایر عوامل (نظری طول و جهت عنصر ابتدایی)، و تعداد بارهایی که می توان یک فرآوری را به کار برد، باید مشخص شود. این کار را می توان با استفاده از قواعد معنایی^۱ ذخیره شده در پایگاه دانش^۲ شکل ۹ . ۱ انجام داد. قالب موجود در قواعد فرآوری، ساختار شیئ را تعیین می کند، در حالی که معنا با صحت آن سروکار دارد. مثلاً جمله 'فرترن^۳ = A' از نظر قالب صحیح است، اما نه وقتي معنای صحیحی دارد که $C \neq 0$ باشد.

فرض کنید که اطلاعات معنایی را به گرامر تشریح شده در مثال قبل پیوند بزنیم. اطلاعات را می توان به صورت بیان شده در جدول ۹ . ۱ به قواعد فرآوری پیوند زد. با استفاده از اطلاعات معنایی، قادریم با به کار گیری تعداد کمی از قواعد قالبی، دسته ای وسیع (اما همانطور که انتظار می رود، محدود) از الگوهای را توصیف کنیم. برای نمونه، با تعیین جهت^۴ در جدول ۹ . ۱، لازم نیست برای هر یک از جهات ممکن، عناصر ابتدایی را تعیین کنیم. به طور مشابه با توجه به این واقعیت که تمام عناصر ابتدایی هم جهت هستند، ساختارهای نامفهومی که از شکل های پایه ای نشان داده شده در شکل ۹ . ۲۶(الف) فاصله دارند، را در نظر نمی گیریم.

خودکاره ها به عنوان تشخیص دهنده های رشته. تاکنون گفتیم که گرامرها تولیدکننده های^۵ الگو هستند. حال این مسأله را بررسی می کنیم که آیا الگوی تولید شده با گرامر G به زبان $L(G)$ تعلق دارد یا خیر. مفاهیم پایه ای در تشخیص قالبی را می توان با ایجاد مدل های ریاضی ماشین های محاسبه گر، که خودکاره ها^۶ خوانده می شوند، نمایش داد. با داشتن یک رشته الگوی ورودی، خودکاره ها قادر است تشخیص بدهد که آیا الگو به زبانی که خودکاره ها با آن مرتبط است، تعلق دارد یا خیر. در اینجا تنها

1) Semantic rule

2) Knowledge base

3) FORTRAN

4) Generator

5) Automata

۴. ۹ روش های ساختاری ۷۲۳

جدول ۱۰. ۹ مثالی از پیوند اطلاعات معنایی با قواعد فرآوری

فرآوری	اطلاعات معنایی
$S \rightarrow aA$	اتصالات به a تنها در نقطه انجام می شوند. جهت a ، که با β نشان داده می شود، بوسیله 'جهت عمود منصف خطی' که دو نقطه انتهایی دو پاره خط بدون نقطه را به هم متصل می کند، معین می شود. هر پاره خط 3 سانتی متر است.
$A \rightarrow bA$	اتصالات به a تنها در نقاط انجام می شوند. چند اتصالی ها مجاز نیستند. جهت b باید ممکنند جهت a باشد. طول b ، 25 سانتی متر است. این فرآوری را نمی توان بیش از ده بار به کار برد.
$A \rightarrow bB$	a و b باید هم جهت باشند. اتصالات باید ساده باشند و تنها در نقاط انجام شوند.
$B \rightarrow c$	c و a باید هم جهت باشند. اتصالات باید ساده باشند و تنها در نقاط انجام شوند.

روی خودکاره های محدود^۱، که تشخیص دهنده زیان های تولید شده با گرامرهای معمولی هستند، تمرکز می کنیم.

خودکاره محدود^۲ به صورت پنج تابع

$$A_f = (Q, \Sigma, \delta, q_0, F) \quad (6-4. 9)$$

تعریف می شود که Q مجموعه 'محدود و غیرتپی از حالات'^۳، Σ الفایی^۴ ورودی محدود، δ نگاشتی^۵ از $\Sigma \times Q$ (مجموعه 'زوج های مرتب تشکیل شده از عناصر Q و Σ) به درون مجموعه 'تمام زیرمجموعه های Q ، q_0 حالت شروع^۶ و F (زیرمجموعه Q) مجموعه ای از حالات پایانی، یا پذیرفته شده^۷ است.

مثال: خودکاره داده شده با معادله^۸ (۶-۴. ۹)، رابازای $\{q_0, q_1, q_2\}$ ، $\Sigma = \{a, b\}$ ، $Q = \{q_0, q_1, q_2\}$ ، $\delta(q_1, a) = \{q_2\}$ ، $\delta(q_0, b) = \{q_1\}$ ، $\delta(q_0, a) = \{q_2\}$ ، $\delta(q_1, b) = \{q_0\}$ ، $\delta(q_2, a) = \{q_1\}$ و $\delta(q_2, b) = \{q_0\}$ در نظر بگیرید. برای مثال اگر خودکاره در حالت q_0 ورودی a باشد، حالتش به q_2 تغییر می یابد. به طور مشابه اگر ورودی بعدی b باشد، خودکاره به حالت q_1 تغییر می یابد، و بدین ترتیب الی آخر. در این مورد حالات ابتدایی و پایانی

1) Finite automata

2) Finite automaton

3) State

4) Alphabet

5) Mapping

6) Starting state

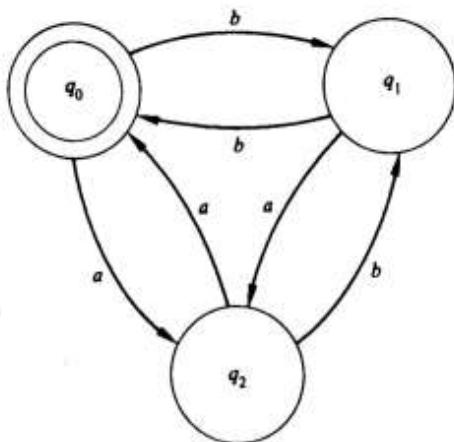
7) Final, or accepting, state

□

پیکسان هستند.

شکل ۹. ۲۷. یک نمودار حالت^۱ را برای خودکاره‌ای که اخیراً تشریح شد، نشان می‌دهد. این نمودار برای هر حالت یک گره دارد و کمان‌های جهت دار انتقالات ممکن بین حالات را نشان می‌دهند. حالت پایانی به صورت دو دایره نشان داده می‌شود، و هر کمان با نمادی که باعث انتقال بین حالات پیوند خورده با آن کمان می‌شود، برچسب زده می‌شود. در این مورد حالات ابتدی و پایانی پیکسان هستند. گفته می‌شود که رشته' w به وسیله' خودکاره پذیرفته یا تشخیص داده شده است، اگر با شروع از حالت q_0 و پیمودن نمادهای w از چپ به راست تا انتهای، در نهایت خودکاره در حالت پایانی باشد. مثلاً خودکاره' شکل ۹. ۲۷. رشته' $w = abbabb$ را تشخیص می‌دهد اما رشته' $w = ababab$ را رد می‌کند.

تนาظر یک به یکی بین گرامرها منظم و خودکاره‌های محدود وجود دارد. یعنی زبان، به وسیله' خودکاره' محدودش تشخیص داده می‌شود اگر و تنها اگر با گرامر منظم تولید شود. طراحی تشخیص دهنده' قالبی رشته‌ها براساس مفاهیمی که تاکنون تشریح شدند، کار ساده‌ای بوده و مشکل از به دست آوردن یک خودکاره' محدود از گرامر منظم داده شده می‌باشد. اجازه دهید که گرامر مورد نظر با G نشان داده شود، که $S = X_0 \cup X_1 \cup \dots \cup X_n$ و فرض کنید N متشکل از X_0, X_1, \dots, X_n غیرپایانه



شکل ۹. ۲۷. خودکاره' محدود

1) State diagram

۴. روش های ساختاری ۷۲۵

دیگر X_1, X_2, \dots, X_n باشد. مجموعه Q مربوط به خودکاره با $n+2$ حالت $\{q_0, q_1, \dots, q_n, q_{n+1}\}$ چنان که بازای $i \leq n \leq i+1$ متناظر با X_i و q_{i+1} حالت پایانی باشد، تشکیل می شود. مجموعه 'نمادهای ورودی' با مجموعه 'پایانه های G ' یکسان است. نگاشتهای درون δ با استفاده از دو قاعده که مبتنی بر فرآوری های G باشد، به دست می آیند؛ به عبارت دیگر، برای هر i و j ، که $0 \leq j \leq n+1$ و $0 \leq i \leq n$:

۱. اگر $aX_j \rightarrow aX_i$ در P باشد، آنگاه $\delta(q_i, a)$ حاوی q_j است.

۲. اگر $X_i \rightarrow a$ در P باشد، آنگاه $\delta(q_i, a)$ حاوی q_{i+1} است.

برعکس، با داشتن خودکاره 'محدود' $A_r = (Q, \Sigma, \delta, q_0, F)$ ، گرامر منظم $G = (N, \Sigma, P, X_0)$ را بدست می آوریم، که در آن N مشکل از عناصر Q است و نماد شروع X_0 نیز متناظر با q_0 می باشد، و فرآوری های G به صورت زیر بدست می آیند:

۱. اگر q_i در $(Q, \Sigma, \delta, q_0, F)$ باشد، فرآوری $aX_i \rightarrow aX_j$ در P وجود دارد.

۲. اگر حالتی از F در $(Q, \Sigma, \delta, q_0, F)$ باشد، فرآوری $a \rightarrow X_i$ در P وجود دارد.

مجموعه 'پایانه' Σ در هر دو مورد یکسان است.

مثال: خودکاره 'محدود' برای گرامر مرتبط با شکل ۹.۲۶، بانوشن فرآوری ها به صورت $X_0 \rightarrow aX_1$ ، $X_0 \rightarrow bX_2$ ، $X_1 \rightarrow bX_2$ و $X_1 \rightarrow bX_1$ به دست می آید. آنگاه، $A_r = (Q, \Sigma, \delta, q_0, F)$ ، که $\delta(q_0, a) = \{q_1\}$ ، $Q = \{q_0, q_1, q_2, q_3\}$ ، $\Sigma = \{a, b\}$ و نگاشتهای عبارتند از: $\delta(q_1, c) = \{q_3\}$ ، $\delta(q_1, b) = \{q_1, q_2\}$. برای تکمیل، می نویسیم:

$\emptyset \leftarrow \delta(q_0, b) = \delta(q_0, c) = \delta(q_1, a) = \delta(q_1, c) = \delta(q_2, a) = \delta(q_2, b) = \emptyset$ مجموعه 'تهی' است و نشان می دهد که این انتقالات برای این خودکاره تعریف نشده اند. \square

تشخیص قالبی درختها

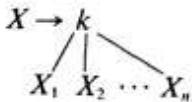
با پیروی از قالبی مشابه با شرح قبلی برای رشته ها، اکنون شرح را طوری بسط می دهیم که شامل توصیف گرهای درختی الگو نیز بشود. دوباره فرض می کنیم که نواحی یا اشیای مورد علاقه در تصویر با استفاده از عناصر اولیه مناسب به شکل درخت هایی، آن طور که در بخش ۸.۵ شرح داده شد، بیان شوند.

گرامرها درختی. گرامر درختی^۱ به صورت پنج تابع

$$G = (N, \Sigma, P, r, S) \quad (7-4. ۹)$$

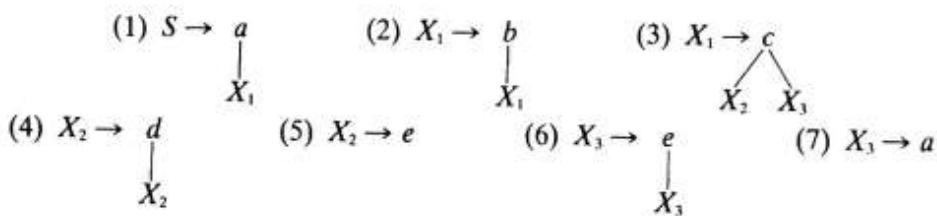
تعريف می شود که مانند گذشته، N و Σ به ترتیب مجموعه های غیرپایانه ها و پایانه ها هستند؛ S ، که درون N قرار دارد، نماد شروع است و در حالت کلی می تواند یک درخت باشد؛ P مجموعه ای از فرآوری های به شکل $T_i \rightarrow T_j$ است، که T_i و T_j درخت هستند؛ و r تابع مرتبه بندی^۲ است که بیانگر تعداد شاخه های مستقیم گره ای است که در گرامر برچسب پایانه دارد. گرامرها درختی گسترشی^۳ که

فرآوری هایی به شکل



دارند و در آن ها X_n, X_2, \dots, X_1 غیرپایانه ها و k پایانه است، ارتباط ویژه ای با بحث ما دارد.

مثال: اسکلت ساختار نشان داده شده در شکل ۹ . ۲۸ (الف) رامی توان با استفاده از یک گرامر درختی که در آن $\{a, b, c, d, e\}$ و $N = \{X_1, X_2, X_3, S\}$ و $\Sigma = \{X_1, X_2, X_3, S\}$ باشد، تولید کرد. با فرض اتصال ابتدا-به- انتهای عناصر ابتدایی خطی، و بافرض شکل ۹ . ۲۸ (ب) هستند، تولید کرد. با فرض اتصال ابتدا-به- انتهای عناصر ابتدایی خطی، و بافرض اتصالات دلخواه به محیط دایره، گرامر مورد بررسی، فرآوری هایی به شکل زیر دارد:



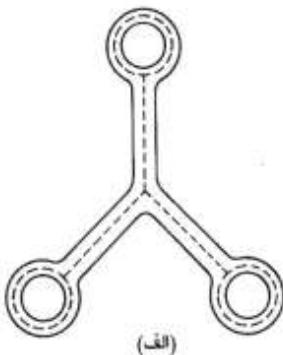
در این مورد توابع مرتبه بندی به صورت $r(b) = r(d) = r(e) = \{1\}$ ، $r(a) = \{0, 1\}$ و $r(c) = \{2\}$ هستند. با این قيد که فرآوری های ۲ ، ۴ و ۶ به تعداد بارهای یکسان به کار روند، ساختاری تولید می شود که در آن تمام سه پایه طول یکسانی دارند. به طور مشابه با این قيد که فرآوری های ۴ و ۶ به تعداد بارهای یکسانی به کار روند، ساختاری تولید می شود که حول محور عمودیش متقاضان است. این

1) Tree grammar

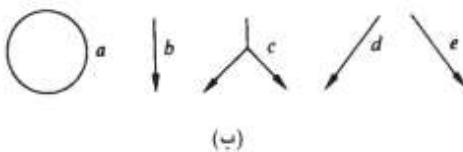
2) Ranking function

3) Expansive tree grammar

۷۴۷ ۴. روش های ساختاری



(الف)



(ب)

شکل ۲۸.۹ (الف) یک شیئ و (ب) عناصر ابتدایی مورد استفاده در تماش اسکلت آن به وسیله گرامر درختی

نوع اطلاعات معنایی مشابه شرح قبلی در رابطه با جدول ۹.۱ و پایگاه دانش شکل ۹.۱ است. \square
 خودکاره های درختی. در حالی که خودکاره محدود معمولی، رشته ورودی را نماد به نماد از چپ به راست پیماش می کند، خودکاره درختی باید همزمان از تمام گره های سرشاخه درخت ورودی (برگها به ترتیب از چپ به راست) شروع کند و در طول مسیرهای موازی به سمت ریشه پیش رود. به طور خاص خودکاره سرشاخه - به - ریشه به صورت

$$A_r = (Q, F, \{f_k \mid k \in \Sigma\}) \quad (8-4.9)$$

تعريف می شود که

Q مجموعه محدودی از حالات است،

F مجموعه حالات پایانی و زیر مجموعه ای از Q است، و

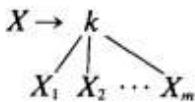
f_k رابطه ای در $Q^m \times Q^m$ است چنان که m مرتبه ای از k می باشد.

نماد Q^m نشانگر m بار ضرب کارتزین Q در خودش است: $Q^m = Q \times Q \times Q \times \dots \times Q$.

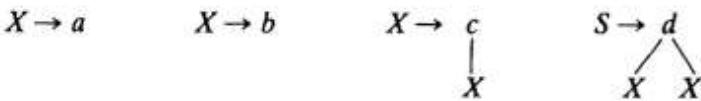
از تعريف ضریف کارتزین، می دانیم که این عبارت به معنی مجموعه تمام m تایی های مرتب از عناصر Q

است. مثلاً اگر $Q^3 = Q \times Q \times Q = \{x, y, z \mid x \in Q, y \in Q, z \in Q\}$ ، $m = 3$ توجه داشته باشید که رابطه 'R' بین مجموعه 'A' و مجموعه 'B'، زیرمجموعه‌ای از ضرب کارتزین A در B است؛ یعنی، $R \subseteq A \times B$. بنابراین هر رابطه در $Q^m \times Q^m$ ، زیرمجموعه‌ای از مجموعه ' $Q^m \times Q^m$ ' است.

برای گرامر درختی گسترشی، $G = (N, \Sigma, P, r, S)$ ، خودکاره' درختی متناظر را با قرار دادن $Q = N$ تشکیل می‌دهیم که در آن $\{S\} = F$ و برای هر نماد a در Σ ، رابطه ' f_k ' را بدين صورت تعريف می‌کنیم که f_k است اگر و تنها اگر فراوری



در G موجود باشد. مثلاً، گرامر درختی ساده' $G = (N, \Sigma, P, r, S)$ را با $\Sigma = \{a, b, c, d\}$ فراوری های



و مرتبه‌بندی های $r(d) = \{2\}$ ، $r(c) = \{1\}$ ، $r(b) = \{0\}$ ، $r(a) = \{0\}$ در نظر بگیرید. خودکاره' درختی متناظر، $A_i = (Q, F, \{f_k \mid k \in \Sigma\})$ ، با قرار دادن $Q = \{S, X\}$ ، $S = \{S\}$ مشخص می‌شود، که روابط آن به صورت زیر تعريف می‌شوند:

$f_a = \{(\emptyset, X)\}$,	$X \rightarrow a$	ناشی از فراوری
$f_b = \{(\emptyset, X)\}$,	$X \rightarrow b$	ناشی از فراوری
$f_c = \{(X, X)\}$,	$X \rightarrow c$	ناشی از فراوری
$f_d = \{(X, X, S)\}$,	$S \rightarrow d$	ناشی از فراوری

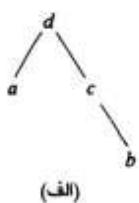
تعییر رابطه' f این است که به گرهی با برچسب a و بدون شاخه (بنابراین نماد تهی \emptyset)، حالت^۱

1) State

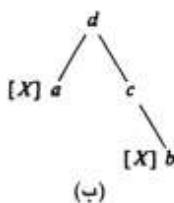
۴۰.۹ روش های ساختاری ۷۴۹

متتب می شود. تعبیر f این است که به گرهی با برچسب c و با یک شاخه که حالت X دارد، حالت X متتب می شود. تعبیر رابطه ' f ' این است که به گرهی با برچسب d و با دو شاخه که هر یک حالت X دارند، حالت S متتب می شود.

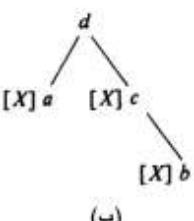
به منظور این که بینیم چگونه این خودکاره درختی، درخت تولید شده با گرامری که قبلاً تشریح شد، را تشخیص می دهد، درخت نشان داده شده در شکل ۲۹.۹ (الف) را در نظر بگیرید. خودکاره ' A ' در ابتدا از طریق روابط ' f ' و ' f_f ' به ترتیب حالاتی را به گره های سرشاخه ' a ' و ' b ' متتب می کند. همان طور که شکل ۲۹.۹ (ب) نشان می دهد، در این مورد طبق دو رابطه ' f ' و ' f_f '، به هر دو برگ، حالت X متتب می شود. اکنون خودکاره، یک سطح از سرشاخه بالاتر می رودویر مبنای ' f ' و حالت شاخه 'گره' ' c '، در شکل ۲۹.۹ (ب) نشان می دهد. همان طور که در شکل ۲۹.۹ (ب) دیده می شود، باز هم حالت حالاتی را به گره ' c ' متتب می کند. همان طور که در شکل ۲۹.۹ (ب) دیده است، باز هم حالت متتب شده بر مبنای ' f '، X است. با یک سطح بالاتر رفتن، خودکاره به گره ' d ' می رسدو چون به هر دو شاخه آن حالتی متتب شده است، رابطه ' f_f ' جهت انتساب حالت S به گره ' d ' استفاده می شود. چون d آخرین گره است و حالت S در F است، خودکاره، این درخت را به عنوان عضو معتبری از زبان گرامر درختی که قبلاً بیان شد، می پذیرد (تشخیص می دهد). شکل ۲۹.۹ نمایش نهایی دنباله 'حالات طی شده در طول مسیرهای سرشاخه - به - ریشه را نشان می دهد.



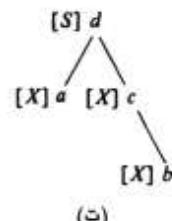
(الف)



(ب)



(پ)



(ت)

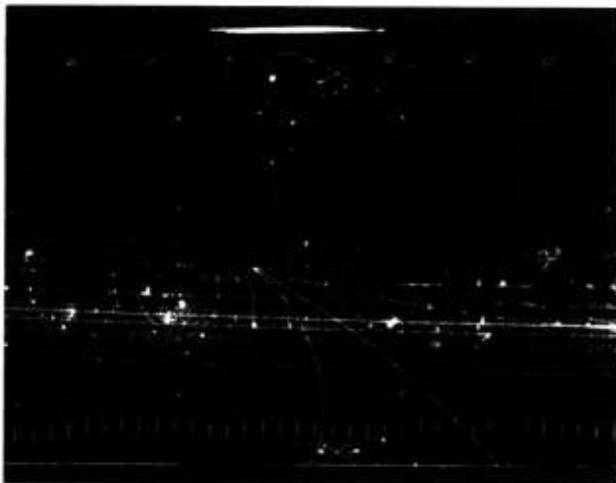
شکل ۲۹.۹ مراحل پردازش خودکاره درختی سرشاخه به ریشه: (الف) درخت ورودی؛ (ب) انتساب حالت به گره های سرشاخه؛ (پ) انتساب حالت به گره های میانی؛ (ت) انتساب حالت به گره ریشه. (از *Gonzalez&Thomason[1978]*)

مثال: یک کاربرد جالب گرامرهای درختی، تحلیل عکس های گرفته شده از رخدادهای اتاق حباب^۱ است. این عکس ها در خلال آزمایش های فیزیک پرانرژی^۲ گرفته می شوند که در آنها پرتوی از ذرات با خواص معلوم به سمت هدفی با هسته های معلوم هدایت می گردد. یک رخداد نوعی، شامل پرتوهای ثانویه ذرات است که از نقطه برخورد پراکنده شده اند و در شکل ۹. ۳۰ مثالی از آن دیده می شود. مسیرهای ورودی، خطوط موازی افقی هستند. به ساختار معمولاً درختی رخداد در نزدیک وسط عکس توجه کنید.

یک آزمایش نوعی صدها هزار عکس تولید می کند که بسیاری از آنها حاوی رخدادهای مورد نظر نیستند. بررسی و طبقه بندی این عکس ها خسته کننده و وقت گیر است، لذا نیاز به الگوریتمهای خودکار پردازش و روش های خودکار تشخیص الگو احساس می شود.

گرامر $G = (N, \Sigma, P, r, S)$ را می توان مشخص کرد که درخت های نمایانگر رخدادهایی از نوع آنچه در اتاق حباب هیدروژنی اتفاق می افتد، را تولید کند. این رخدادها حاصل جریان های ورودی ذرات باردار مثبت می باشند. در این مورد $\{S, X_1, X_2\} = \{a, b\}$ ، $N = \Sigma$ ، و عناصر ابتدایی a و b

$a:$	\cap	کمان محدب
$b:$	\cup	کمان مقعر



شکل ۹. ۳۰ عکس اتاق حباب. (از [Fu & Bhargava [1973]

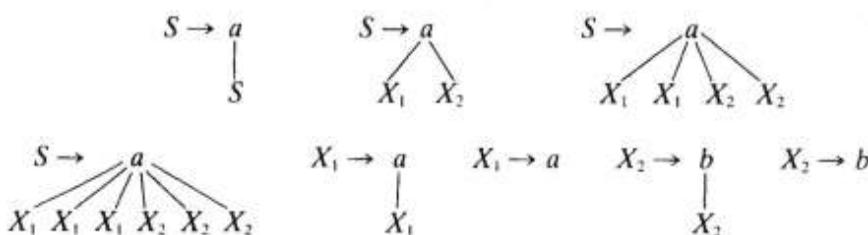
1) Bubble chamber event

2) High-energy physics

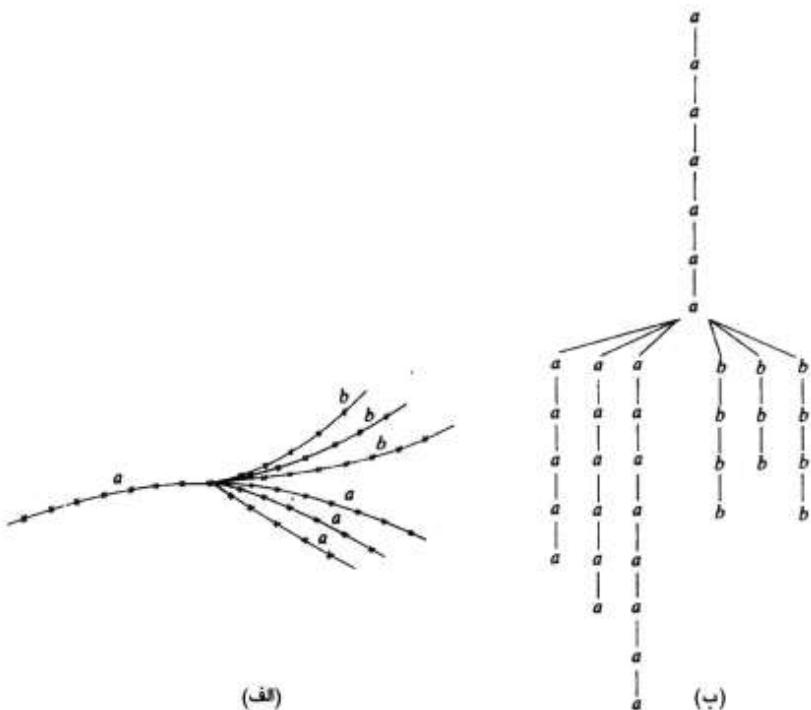
۹.۴ روش های ساختاری ۷۸۱

به صورت زیر تعبیر می شوند:

فرآوری های درون P عبارتند از:



مرتبه بندی ها عبارتند از $\{6, 4, 2, 1, 0\}$. فرآوری های شاخه دار، نمایانگر تعداد مسیرهای منشعب شده از برخورد هستند. این تعداد زوج است و معمولاً از شش تجاوز نمی کند. شکل ۹.۳۱ (الف) رخداد برخورد موجود در شکل ۹.۳۰ که به بخش های محدب و مقعر تقسیم شده است، را نشان می دهد و شکل ۹.۳۱ (ب) نمایش درختی متناظر با آن را نشان می دهد. با



شکل ۹.۳۱ (الف) رخداد رمز شده ای از شکل ۹.۳۰؛ (ب) نمایش درختی متناظر. (از *Fu & Bhargava [1973]*)

گرامر داده شده در بالا می توان این درخت و نیز گونه های آن را تولید کرد.

خودکاره^۱ درختی مورد نیاز برای تشخیص انواع درخت هایی که اخیراً تشریح شدند، با استفاده از روال مشروح در توضیحات قبلی تعریف می شود. بنابراین $(A, = (Q, F, \{f_k \mid k \in \Sigma\})$ ، با فرض این که $\{f_k \mid k \in \Sigma\} = \{f_a, f_b\}$ معین می شود، که روابط f_a و f_b به صورت $f_a = \{(S, S), (X_1, X_2, S), (X_1, X_1, X_2, X_2, S)(X_1, X_1), (\emptyset, X_1)\}$ و $f_b = \{(X_2, X_2) = (\emptyset, X_2)\}$ تعریف می شوند. به عنوان تمرین نشان دهید که این خودکاره درخت موجود در شکل ۹ . ۳۱ (ب) را می پذیرد.

پادگیری^۲

روش های تشخیص قالبی معرفی شده در قبل، نیازمند تعیین خودکاره های (تشخیص دهنده های) مناسب برای هر دسته مورد بررسی است. در وضعیت های ساده، با بررسی می توان خودکاره های موردنیاز را به دست آورد. در موارد پیچیده تر، ممکن است الگوریتمی برای یادگیری خودکاره ها از روی الگوهای ساده (نظیر رشته ها یا درختها) موردنیاز باشد. به علت تناقض یک - به - یک بین خودکاره ها و گرامرها، گاهی مسأله یادگیری به صورت یادگیری مستقیم گرامرها از روی الگوهای نمونه بیان می شود، که این فرآیند معمولاً استنتاج گرامری^۳ خوانده می شود. در این بخش روی الگوریتمی برای یادگیری خودکاره های محدود تمرکز می کنیم. مراجع انتهای فصل راهنمایی لازم برای روش های یادگیری گرامرها و خودکاره های درختی و همچنین سایر روش های تشخیص قالبی را فراهم می آورند.

فرض کنید که تمام الگوهای یک دسته به وسیله 'گرامر مجھول G تولید می شوند و مجموعه^۴ محدودی

از نمونه ها^۵ R^+ با خاصیت

$$R^+ \subseteq \{\alpha \mid \alpha \text{ in } L(G)\} \quad (9-4. 9)$$

موجود است. در واژه شناسی ارائه شده در بخش ۹ . ۲ ، R^+ که مجموعه نمونه^۶ مثبت^۷ خوانده می شود، تنها مجموعه ای از الگوهای آموزشی مرتبط با گرامر G است. مجموعه^۸ نمونه^۹ R^+ ، کامل ساختاری^{۱۰} گفته می شود اگر هر فرآوری G لاقل برای تولید یک عنصر^{۱۱} R^+ به کار رود. ما می خواهیم یک خودکاره

1) Learning

2) Grammatical inference

3) Positive sample set

4) Structurally complete

۴.۹ روش های ساختاری ۷۵۳

محدود، A_f را که رشته های R^+ و احتمالاً بعضی رشته های شبیه به رشته های R را می پذیرد، یاد بگیریم (بسازیم).

براساس تعریف خودکاره' محدود و تاظرین G و A_f ، نتیجه می شود که $\Sigma^* \subseteq R^+$ ، که مجموعه تمام رشته های متشکل از عناصر Σ است. فرض کنید که z رشته ای از Σ^* باشد چنان که بازای بعضی w های Σ^* در R^+ باشد. برای عدد صحیح مثبت k ، دنباله' z^k نسبت به R را به صورت مجموعه' $h(z, R^+, k)$ تعریف می کنیم، که

$$h(z, R^+, k) = \{w \mid zw \text{ in } R^+, |w| \leq k\}. \quad (10-4. 9)$$

به عبارت دیگر، دنباله' z^k مجموعه' رشته های w با این خواص است که (۱) zw در R^+ باشد، و (۲) طول w کوچکتر یا مساوی k باشد.

یک روال برای یادگیری خودکاره' $A_f(R^+, k) = (Q, \Sigma, \delta, q_0, F)$ از مجموعه' نمونه' R^+ و مقدار ویژه' k ، این است که فرض کنیم

$$Q = \{q \mid q = h(z, R^+, k) \text{ for } z \text{ in } \Sigma^*\} \quad (11-4. 9)$$

و برای هر a در Σ

$$\delta(q, a) = \{q' \mid q \in Q \mid q' = h(za, R^+, k), \text{ with } q = h(z, R^+, k)\}. \quad (12-4. 9)$$

همچنین، قرار می دهیم:

$$q_0 = h(\lambda, R^+, k) \quad (13-4. 9)$$

و

$$F = \{q \mid q \in Q, \lambda \in q\} \quad (14-4. 9)$$

که λ رشته' تهی (رشته ای بدون نماد) است. لازم به یادآوری است که $A_f(R^+, k)$ حالت هایی دارد زیرمجموعه ای از مجموعه' تمام k دنباله که می توان از R^+ ساخت، هستند.

مثال: فرض کنید که $\{a, ab, abb\} = R^+ = k = 1$. از تعریف قبلی نتیجه می گیریم که

1) k tail of z

$$\begin{aligned}
 z = \lambda, & \quad h(\lambda, R^+, 1) = \{w \mid \lambda w \text{ in } R^+, |w| \leq 1\} \\
 & = \{a\} \\
 & = q_0; \\
 z = a, & \quad h(a, R^+, 1) = \{w \mid aw \text{ in } R^+, |w| \leq 1\} \\
 & = \{\lambda, b\} \\
 & = q_1; \\
 z = ab, & \quad h(ab, R^+, 1) = \{\lambda, b\} \\
 & = q_1; \\
 z = abb, & \quad h(abb, R^+, 1) = \{\lambda\} \\
 & = q_2.
 \end{aligned}$$

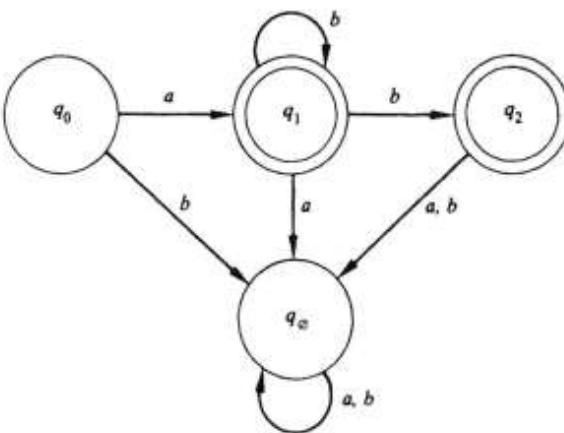
در این مورد، سایر رشته های z موجود در Σ^* ، رشته های zw که متعلق به R^+ نیستند، را می دهند و منجر به حالت چهارمی می شوند که با نماد q_\emptyset نشان داده شده و متناظر با حالتی است که 'مجموعه' تهی باشد. بنابراین حالات عبارتند از $\{\lambda\}$, $q_0 = \{a\}$, $q_1 = \{\lambda, b\}$, $q_2 = \{q_0, q_1, q_2\}$. q_\emptyset و $q_0 = \{a\}$, $q_1 = \{\lambda, b\}$, $q_2 = \{q_0, q_1, q_2\}$ ، که مجموعه 'به دست می آیند، تنها برچسب های حالت q_0, q_1, q_2 در تشکیل مجموعه Q استفاده می شوند.

مرحله بعدی، به دست آوردن توابع انتقال است. چون $q_0 = h(\lambda, R^+, 1)$, نتیجه می شود که $\delta(q_0, b) = h(\lambda b, R^+, 1) = h(b, R^+, 1) = q_\emptyset$ و $\delta(q_0, a) = h(\lambda a, R^+, 1) = h(a, R^+, 1) = q_1$ به صورت مشابه $q_1 = h(a, R^+, 1) = h(ab, R^+, 1)$ و $q_2 = h(ab, R^+, 1) = q_1$. همچنین، $\delta(q_1, a) = h(aa, R^+, 1) = h(aba, R^+, 1) = q_\emptyset$ و $\delta(q_1, b) = \{q_1, q_2\}$; یعنی، $\delta(q_1, b) \supseteq h(abb, R^+, 1) = q_2$. با ادامه روایی که در بالا تشریح شد، به دست می آوریم که: $\delta(q_2, a) = \delta(q_2, b) = \delta(q_\emptyset, a) = \delta(q_\emptyset, b) = q_\emptyset$. با ادامه روایی که در بالا تشریح مجموعه 'حالات نهایی'، حاوی حالتی است که در نمایش k دنباله شان¹، رشته 'تهی' λ را دارند. در این مورد، $F = \{q_1, q_2\}$ و $q_1 = \{\lambda\}$, $q_2 = \{\lambda, b\}$ ، بنابراین $A_f(R^+, 1) = (Q, \Sigma, \delta, q_0, F)$ داده می شود

براساس این نتایج، خودکاره 'حاصل با' $A_f(R^+, 1)$ داده می شود که $F = \{q_1, q_2\}$, $\Sigma = \{a, b\}$, $Q = \{q_0, q_1, q_2, q_\emptyset\}$

1) k - tail representation

۴.۹ روش های ساختاری ۷۵۵



شکل ۴.۹ نمودار حالت برای خودکاره' محدود $A_f(R^+, 1)$ که از مجموعه' $R^+ = \{a, ab, abb\}$ استنتاج شده است.

۹. ۳۲ نمودار حالت را نشان می دهد. خودکاره، رشته هایی به شکل "abⁿ" را را می پذیرد. به عبارت دیگر، این روال تکرار منظم نماد "b" را شناسایی کرده است. \square

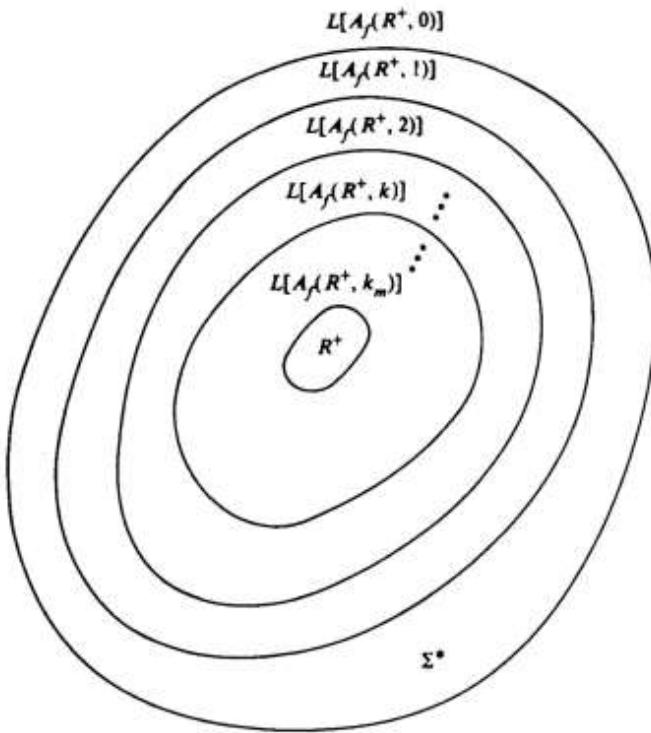
مثال قبل به روشنی نشان می دهد که مقدار k طبیعت خودکاره' حاصل را کنترل می کند. خواص زیر مثالهایی از بستگی $(A_f(R^+, k))$ به این عامل هستند.

خاصیت ۱. برای تمام k هایی که $0 \leq k \leq R^+$ ، $L[A_f(R^+, k)] \subseteq L[A_f(R^+ k)]$ زیان پذیرفته شده به وسیله $(A_f(R^+, k))$ است.

خاصیت ۲. اگر k برابر، یا بزرگتر از، طول بزرگترین رشته در R^+ باشد، $R^+ = L[A_f(R^+, k)]$ و اگر $0 < k = \Sigma^*$.

خاصیت ۳. $L[A_f(R^+, k + 1)] \subseteq L[A_f(R^+, k)]$.

خاصیت ۱ تضمین می کند که $A_f(R^+, k)$ لااقل رشته های درون مجموعه' نمونه' R^+ را می پذیرد. اگر k برابر، یا بزرگتر از، طول بلندترین رشته در R^+ باشد، آنگاه طبق خاصیت ۲، خودکاره تنها رشته های درون R^+ را می پذیرد. اگر $0 < k = \Sigma^*$ دارای یک حالت $\{q_0\} = \lambda$ خواهد بود، که به عنوان هر دو حالت ابتدایی و پایانی عمل می کند. آنگاه توابع انتقال برای a از λ به شکل $\delta(q_0, a) = q_0$ خواهد بود. بنابراین $L[A_f(R^+, 0)] = \Sigma^*$ ، و خودکاره رشته' تهی λ و تمام رشته های تشکیل شده از نمادهای Σ را می پذیرد. نهایتاً، خاصیت ۳ بیان می کند که با کاهش k



شکل ۹.۳۳ ارتباط بین $L[A_f(R^+, k)]$ و k . مقدار k چنان است که (طول بلندترین رشته در R^+) $\geq k_m$.

محدوده زیان پذیرفته شده با $A_f(R^+, k)$ کاهش می یابد.

این سه خاصیت تنها با تغییر عامل k ، کنترل طبیعت $A_f(R^+, k)$ را ممکن می سازند. اگر حدسی از زیان L_0 باشد که نمونه R^+ از آن انتخاب شدو اگر k خیلی کوچک باشد، این L_0 استنتاج آزادی را تشکیل می دهد که ممکن است شامل اغلب یا تمام رشته های درون Σ باشد. با این حال، اگر k برابر طول بلندترین رشته R^+ باشد، استنتاج در این مفهوم که $A_f(R^+, k)$ تنها رشته های درون R^+ را می پذیرد، نگهدار^۱ است. شکل ۹.۳۳ این مفاهیم را با شکل نشان می دهد.

مثال: مجموعه $R^+ = \{caaab, bbaab, caab, bbab, cab, bbb, cb\}$ را در نظر بگیرید.

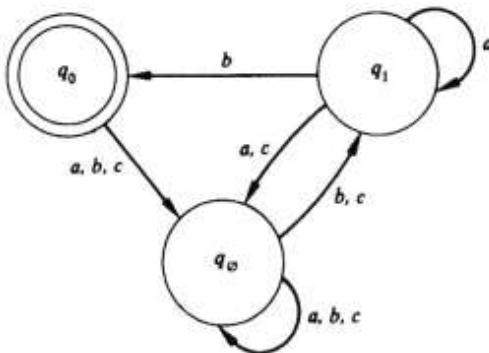
برای $k = 1$ ، با پیروی از روال بکار رفته در مثال قبل داریم:

1) Conservative

۷۵۷ روش های ساختاری ۴.۱

- | | |
|-------------------|---|
| 1. $z = \lambda,$ | $h(z, R^+, 1) = \{\emptyset\} = q_\emptyset;$ |
| 2. $z = c,$ | $h(z, R^+, 1) = \{b\} = q_1;$ |
| 3. $z = ca,$ | $h(z, R^+, 1) = \{b\} = q_1;$ |
| 4. $z = cb,$ | $h(z, R^+, 1) = \{\lambda\} = q_0;$ |
| 5. $z = caa,$ | $h(z, R^+, 1) = \{b\} = q_1;$ |
| 6. $z = cab,$ | $h(z, R^+, 1) = \{\lambda\} = q_0;$ |
| 7. $z = caaa,$ | $h(z, R^+, 1) = \{b\} = q_1;$ |
| 8. $z = caab,$ | $h(z, R^+, 1) = \{\lambda\} = q_0;$ |
| 9. $z = caaab,$ | $h(z, R^+, 1) = \{\lambda\} = q_0;$ |
| 10. $z = b,$ | $h(z, R^+, 1) = \{\emptyset\} = q_\emptyset;$ |
| 11. $z = bb,$ | $h(z, R^+, 1) = \{b\} = q_1;$ |
| 12. $z = bba,$ | $h(z, R^+, 1) = \{b\} = q_1;$ |
| 13. $z = bbb,$ | $h(z, R^+, 1) = \{\lambda\} = q_0;$ |
| 14. $z = bbaa,$ | $h(z, R^+, 1) = \{b\} = q_1;$ |
| 15. $z = bbab,$ | $h(z, R^+, 1) = \{\lambda\} = q_0;$ |
| 16. $z = bbaab,$ | $h(z, R^+, 1) = \{\lambda\} = q_0.$ |

$\Sigma = \{a, b, c\}$, $F = \{q_0\}$, $Q = \{q_0, q_1, q_\emptyset\}$ است که $A_f(R^+, 1) = (Q, \Sigma, \delta, q_0, F)$ خودکاره و انتقالات در نمودار حالت شکل ۳۴ نشان داده شده اند. برای این که خودکاره رشته را پذیرد، رشته باید با a , b یا c شروع شود و با b پایان یابد. همچنین $A_f(R^+, 1)$ رشته های بازگشتی در a , b یا c را می پذیرد.



شکل ۳۴.۹ نمودار حالت خودکاره $A_f(R^+, 1)$ که از مجموعه 'نمونه' $R^+ = \{caaab, bbaab, caab, bbab, cab, bbb, cb\}$ استنتاج شده است.

مزیت اصلی روش فوق سادگی پیاده سازی است. با کمترین تلاشی می توان روال ساخت را روی رایانه های رقمی شبیه سازی کرد. عیب اصلی آن، تصمیم گیری در مورد مقدار مناسب k است، گرچه با توجه به سه خاصیتی که قبلاً تشریح شد، این مسأله تا حدی آسان می شود.

۵.۹ تعبیر

تاکنون بیشتر روی فرآیندهای مختلف، از تصویربرداری و پیش پردازش تابخش بندی، توصیف، و تشخیص شیء، تمرکز کرده ایم. در این بخش، مسأله استفاده از کل اطلاعات حاصل از این فرآیندها در تعبیر^۱ محتویات تصویر را شرح می دهیم. به عبارت دیگر، علاقه مندیم که به تصویر معنا دهیم، فرآیندی که علاوه بر عنوان تعبیر تصویر غالباً با عنوانی فهم تصویر^۲ یا تحلیل صحنه^۳ نیز خوانده می شود. برای همگونی با توضیحات قبلی، به استفاده از اصطلاح تعبیر تصویر^۴ ادامه می دهیم.

۱۰.۹ زمینه

تعبیر رایانه ای تصویر فرآیند بسیار پیچیده ای است. مشکلات معمولاً ناشی از زیادی داده هایی که باید پردازش شوند و عدم وجود ابزارهای اساسی پردازش برای رسیدن از داده ها (آرایه، پیکسل ها) به نتیجه^۵ مطلوب (شرح دقیق محتویات تصویر) می باشد. بنابراین بدون وجود ابزارهای عمومی برای تعبیر ساختار نیافته تصویر، مجبوریم از روش هایی که احتمال موفقیت قابل قبولی دارند، استفاده کنیم. این قید مارابه یک مصالحه^۶ دو مرحله ای هدایت می کند که: (۱) کلیت مسأله را محدود کنیم، و (۲) سعی کنیم از دانش بشری نیز در فرآیند استفاده کنیم. محدود کردن مسأله، کار ساده ای است. در صورت امکان از تمام وسائل موجود برای محدود کردن شرایط ناشناخته استفاده می کنیم تا بدین وسیله مسأله مورد نظر ساده شود. وقتی متغیرهای مسأله در کنترل نباشند، مجبوریم که محدوده (و دقت) نتایج مورد انتظار را کاهش دهیم. برای به کار گیری دانش بشری در تعبیر تصویر لازم است که قالبی را برای بیان این دانش برگزینیم. سه روش اصلی مورد استفاده مبتنی بر (۱) منطق رسمی^۷؛ (۲) شبکه های معنایی^۸؛ و (۳) سامانه های فراوری هستند. غالب سامانه های منطقی بر مبنای حساب گزاره ای^۹ مرتبه اول می باشند، که زبانی با منطق بر

1) Interpret

2) Image understanding

3) Scene analysis

4) Image interpretation

5) Formal logic

6) Predicate calculus

نمادی است که با آن می توان طیفی از جملات از واقعیات ساده تا عبارات نسبتاً پچیده را بیان کرد. ابزارهای موجود در حساب گزاره ای این امکان را می دهند که دانش بر حسب قواعد منطقی بیان شود و با استفاده از آنها می توان درستی عبارات منطقی را اثبات (یارد) کرد.

روشهای معنایی، دانش را به شکل شبکه های معنایی^۱ نمایش می دهند، که این شبکه ها، گرافهای بر چسب دار و جهت دار هستند. این گرافهای صورت شهودی روابط بین عناصر تصویر را بیان می کنند. در این نوع طرز بیان، اشیاء به عنوان گره های گراف نمایش داده می شوند، و روابط بین اشیاء به صورت کمان های بر چسب داری که گره های مختلف را به هم متصل می کنند، بیان می شوند. همان طور که در بخش ۹ . ۵ . ۴ نشان می دهیم، شبکه های معنایی ابزاری قوی برای تعبیر تصویر فراهم می کنند.

روش های مبتنی بر سامانه های فرآوری^۲ (که روش های مبتنی بر قاعده^۳ نیز خوانده می شوند) بیشترین توجه را در کاربردهای تعبیر تصویر به خود جلب کرده اند. دلایل این توجه عبارتند از: در دسترس بودن ابزارهای متعدد برای توسعه این سامانه ها و این واقعیت که دانش بشری را می توان به صورت شهودی، ساده، و با روندی افزایشی در این سامانه ها به کار برد. بویژه سامانه های خبره^۴، که به حوزه های با وظیفه مشخص محدود نند، قادر به انجام محدوده^۵ وسیعی از کاربردهای پردازش تصویر هستند.

۲۰.۵.۹ انواع دانش

تقسیم توابع پردازش تصویر به سه گروه (پردازش سطح پایین، متوسط و بالا) که در بخش ۹ . ۱ تشریح شدند، برای الگوریتم ها درست است، اما بویژه برای طبقه بندی انواع دانش مورد نیاز برای کارهای پردازش تصویر صحیح نیست. یک تقسیم بندی بسیار مفیدتر عبارتست از: (۱) دانش فرآیندی^۶، (۲) دانش دیداری^۶، و (۳) دانش جهانی^۷.

دانش فرآیندی به عملیاتی شبیه انتخاب الگوریتم ها و تنظیم عوامل این الگوریتم ها (مثلاً انتخاب مقادیر آستانه) می پردازد. دانش دیداری به جنبه های مختلف تشکیل تصویر، نظری انتظار سایه داشتن از نور پردازی مایل شیئی سه بعدی، توجه می کند. دانش جهانی بیانگر دانش کلی درباره 'حوزه' یک مسئله است. از جمله مثال های این نوع دانش، روابط معلوم بین اشیای تصویر (برای نمونه، تقاطع باندهای پرواز و

1) Semantic network

2) Production

3) Rule-based

4) Expert system

5) Procedural knowledge

6) Visual knowledge

7) World knowledge

حرکت در تصویر هوایی فرودگاه) و روابط بین صحنه و محیطش (نظیر افزایش انعکاس جاده در شب بر اثر باران می دهد) هستند.

به طور کلی، دانش های فرآیندی و دیداری برای پردازش سطح پایین و میانی، و دانش جهانی برای پردازش سطح بالا استفاده می شود. بویژه، دانش جهانی مبنای کارهای تعییر تصویر را تشکیل می دهد. بدون توجه به نوع دانش، یکی از مهمترین جنبه های نمایش دانش برای سامانه پردازش تصویر استقلال تا حد ممکن آن از کاربرد است. بنابراین هدف این است که از هدر رفتن دانش در نرم افزار^۱ یا سخت افزار دوری کیم. بنابراین الگوریتم های پایه ای پردازش باید مبتنی بر دانش مستقل از حوزه باشند، و دانش لازم درباره کاربردهای مشخص باید در پایگاه دانش سامانه موجود و تا حد ممکن از این الگوریتم ها مستقل باشد.

۳.۵.۹ سامانه های منطقی (حساب گزاره ای)

منطق گزاره ای بیش از یک قرن است که به صورت متون علمی مدون و جامع تهیه شده و برای بیان پیشنهادها و برای استنتاج واقعیت های جدید از روی پایگاه دانش مفید است. یکی از قویترین عناصر این نوع دانش، حساب گزاره ای مرتبه اول^۲ است که یک سامانه منطقی است و قادر به کار با محدوده وسیعی از عبارات ریاضی، همچنین کار با جملات یک زبان طبیعی، نظیر انگلیسی، می باشد.

تعریف ها

اجزای اصلی حساب گزاره ای، نمادهای گزاره^۳، نمادهای تابع^۴، نمادهای متغیر^۵، و نمادهای ثابت^۶ هستند. نماد گزاره، رابطه ای در حوزه مورد بحث را نمایش می دهد. مثلاً جمله "پنج کوچکتر از ده است" را می توان به صورت LESS THAN (five, ten) بیان کرد، که نماد گزاره بوده و پنج و ده نمادهای ثابت هستند.

جدول ۹ . ۲ سایر مثال هارانشان می دهد. در این مثال ها، نماد گزاره (نظیر MOTHER)، گزاره ای را که دارای یک یا دو آرگومان^۷ است، معرفی می کند. آرگومانها ممکن است ثابت^۸، نظیر Mary،

1) Code

2) First- order predicate calculus

3) Predicate symbol

4) Function symbol

5) Variable symbol

6) Constant symbol

7) Argument

8) Constant

Poe، و MARRIED، باشند. آرگومانها نیز ممکن است توابعی از سایر آرگومان ها باشند. مثلاً [MARRIED father (John), mother (John)] نمایانگر عبارت "پدر John با مادر John ازدواج کرده است" می باشد. در اینجا John یک نماد ثابت، father و mother نمادهای تابع، و MARRIED یک نمادگزاره است. در مثال آخر جدول ۲.۲، BEHIND نمادگزاره بوده و X و Y نمادهای متغیر هستند. گزاره هایی نظیر گزاره های جدول ۲.۲، اتم^۱ نیز خوانده می شوند. همان طور که در جدول ۲.۳ دیده می شود، اتم ها را می توان با پیونددهنده های منطقی^۲ باهم ترکیب کرد تا جمله ها تشکیل شوند. در جدول ۲.۳، این پیونددهنده های منطقی، معناهای شناخته شده زیر را دارند: " \wedge "(AND)، " \vee "(OR)، " \sim "(NOT)، و " \Rightarrow "(نتیجه می دهد)؛ $\forall x$ ، که کمیت نمای عمومی^۳ خوانده می شود، به معنای "برای تمام x ها" است. به طور مشابه، $\exists x$ ، که کمیت نمای وجودی^۴ خوانده می شود، به معنای "لاقل یک x وجود دارد" می باشد. چهار مثال اول جدول ۲.۳ نمادهای ثابت و دو مثال آخر نمادهای متغیر را به کار می برند. عبارات منطقی حاصل از پیوند عبارات مختلف با \wedge (\vee) ترکیبات عطفی^۵ (فصلی)^۶ خوانده می شوند. عبارات رسمی (قانونی) حساب گزاره ای، روابط خوب ساختار یافته^۷ (wffs) خوانده می شوند.

عبارت منطقی در قالب غیر گزاره ای^۸ خوانده می شود اگر شامل اتم ها، پیونددهنده های منطقی، کمیت نمایی عمومی و کمیت نمایی وجودی باشد. عبارت منطقی در قالب گزاره ای^۹ خوانده می شود اگر به شکل $[A_1 \wedge A_2 \wedge \dots \wedge A_n \Rightarrow B_1 \vee B_2 \vee \dots \vee B_m]$ باشد، که A_i ها

جدول ۲.۹ مثال هایی از گزاره ها

گزاره	جمله ضربی
FEMALE(Mary)	مؤنث است Mary
MOTHER(Mary)	مادر است Mary
BIGGERTHAN(oceans, lakes)	اقیانوس ها بزر گتر از دریاچه ها هستند.
WRITE(Poe, The Raven)	Poe را Raven نوشت.
BEHIND(x, y)	X پشت y است.

1) Atom

2) Logical connective

3) Universal quantifier

4) Existential quantifier

5) Conjunction

6) Disjunction

7) Well-formed formulas

8) Nonclausal from syntax

9) Clausal from syntax

جدول ۹ مثال هایی از عبارات گزاره ای

جمله خبری	عبارت گزاره ای
Mary مؤنث و مادر است.	FEMALE(Mary) \wedge MOTHER(Mary)
Mary مذکر یا مؤنث است.	MALE(Mary) \vee FEMALE(Mary)
Mary مذکر نیست	\sim MALE(Mary)
اگر Mary مادر باشد، مؤنث است.	MOTHER(Mary) \Rightarrow FEMALE(Mary)
هر کسی یا مذکر یا مؤنث است.	($\forall x$)[MALE(x) \vee FEMALE(x)]
Jaws را نوشته شخصی وجود دارد که	($\exists x$)WRITE(x , Jaws)

و B ها، اتم هستند. قسمت های چپ و راست عبارت گزاره ای به ترتیب شرط^۱ و نتیجه^۲ خوانده می شوند. وقتی طرف شرط عبارت استنتاج^۳ خالی باشد، عبارتی به شکل $P \Rightarrow$ به معنای P تعبیر می شود. بر عکس، عبارت $\Rightarrow P$ به معنای $\sim P$ تعبیر می شود.

عبارت "برای هر x ، اگر x یک شخص و یکی از والدین باشد، آنگاه x یا مادر و یا پدر است." را در نظر بگیرید. در قالب گزاره ای، این عبارت به صورت ذیل نوشته می شود:

$$(\forall x)[\text{PERSON}(x) \wedge \text{PARENT}(x) \Rightarrow \text{MOTHER}(x) \vee \text{FATHER}(x)].$$

در قالب غیر گزاره ای عبارت مذکور به صورت زیر می شود:

$$(\forall x)[\sim \text{PERSON}(x) \vee \sim \text{PARENT}(x) \vee \text{MOTHER}(x) \vee \text{FATHER}(x)].$$

به سادگی می توان هم ارزی این دو عبارت را بررسی کرد (جدول ۹ . ۴ را ببینید). در واقع، همواره تبدیل از قالب غیر گزاره ای به گزاره ای و بالعکس ممکن است. بنابراین این دو نمایش، بیان معادلی دارند.

جدول ۹ . ۴ روابط بین پیوندهای منطقی گوناگونی که اخیراً معرفی شدند، را نشان می دهد. از منطق مقدماتی، باید محتویات پنج ستون اول کاملاً برایتان آشنا باشد. عملگر استنتاج ممکن است زیاد آشنا نباشد. طرف چپ عبارت استنتاج، مقدمه^۴ و طرف راست آن، نتیجه^۵ خوانده می شود. جدول

1) Condition

2) Conclusion

3) Implication

4) Antecedent

5) Consequent

۷۶۳ تعبیر ۴

جدول ۴.۹ جدول درستی پیوندهای منطقی

A	B	$\sim A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$
T	T	F	T	T	T
T	F	F	F	T	F
F	T	T	F	T	T
F	F	T	F	F	T

۹. نشان می دهد که عبارت استنتاج ارزش T (درست) دارد اگر نتیجه ارزش T داشته باشد (بدون توجه به ارزش مقدمه) یا اگر مقدمه ارزش F داشته باشد (بدون توجه به ارزش نتیجه)؛ در غیر این صورت عبارت استنتاج ارزش F دارد. این تعریف -که در آن یک استنتاج درست است هرگاه مقدمه نادرست باشد- منشأ بسیاری از سوءتفاهم ها است، زیرا ممکن است منجر به عبارات نامفهومی گردد. مثلاً جمله 'بی مفهوم' "اگر زمین مربعی باشد، آنگاه تمام حیوانات به رنگ آبی هستند" در حساب گزاره ای ارزش T دارد زیرا مقدمه نادرست است. در عمل، معمولاً چنین مشکلی رخ نمی دهد اما در خاطر داشته باشید که گاهی عبارت منطقی در زبان طبیعی بی مفهوم است.

مثال: عبارات حساب گزاره ای زیر مفاهیم معرفی شده در توضیحات قبلی را نشان می دهند.

۱. اگر تصویر رقمه ای باشد، آنگاه پیکسل هایش گسته هستند.

$DIGITAL(image) \Rightarrow DISCRETE(pixels)$.

۲. تمام تصاویر رقمه ای پیکسل های گسته دارند:

$(\forall x)[[IMAGE(x) \wedge DIGITAL(x)] \Rightarrow (\exists y)[PIXEL-IN(y, x) \wedge DISCRETE(y)]]$.

عبارت فوق به صورت زیر خوانده می شود: برای تمام x ها که x تصویر و رقمه باشد، دست کم یک y وجود دارد که y پیکسلی در x است و y گسته است (به کاربرد متغیرها توجه کنید).

۳. همه تصاویر رقمه نیستند:

$(\forall x)[IMAGE(x)] \Rightarrow (\exists y)[IMAGE(y) \wedge \sim DIGITAL(y)]$.

این عبارت به این صورت خوانده می شود: برای تمام x ها، اگر x تصویر باشد، آنگاه x بی وجود دارد

که u تصویر است و u رقمه نیست.

۴. تصاویر رقمه رنگی اطلاعات بیشتری از تصاویر رقمه تک رنگ دارند:

$$(\forall x)(\forall y)[[\text{IMAGE}(x) \wedge \text{DIGITAL}(x) \wedge \text{COLOR}(x)] \wedge [\text{IMAGE}(y) \wedge \text{DIGITAL}(y) \wedge \text{MONOCHROME}(y)] \Rightarrow \text{MOREINFO}(x, y)].$$

این عبارت خوانده می شود: برای تمام x ها و تمام y ها، اگر x تصویر رنگی رقمه باشد و اگر y

□ تصویر رقمه تک رنگ باشد، آنگاه x اطلاعات بیشتری از y دارد.

چند برابری مهم

درستی برابری های زیر را می توان با استفاده از جدول ۹ . ۴ و منطق مقدماتی بررسی کرد:
این برابری ها برای ساده سازی عبارات منطقی بسیار مفید هستند.

A	معادل است با	$\sim(\sim A)$	
$\sim A \Rightarrow B$	معادل است با	$A \vee B$	قانون خلف ^۱
$\sim B \Rightarrow \sim A$	معادل است با	$A \Rightarrow B$	قانون دمورگان ^۲
$\sim A \vee \sim B$	معادل است با	$\sim(A \wedge B)$	
$\sim A \wedge \sim B$	معادل است با	$\sim(A \vee B)$	قوانين توزیعی ^۳
$(A \wedge B) \vee (A \wedge C)$	معادل است با	$A \wedge (B \vee C)$	
$(A \vee B) \wedge (A \vee C)$	معادل است با	$A \vee (B \wedge C)$	قوانين جایه جایی ^۴
$B \wedge A$	معادل است با	$A \wedge B$	
$B \vee A$	معادل است با	$A \vee B$	
$A \wedge (B \wedge C)$	معادل است با	$(A \wedge B) \wedge C$	قوانين انجمنی ^۵
$A \vee (B \vee C)$	معادل است با	$(A \vee B) \vee C$	
$(\exists x)[\sim P(x)]$	معادل است با	$\sim(\forall x)P(x)$	بعلاوه،
$(\forall x)[\sim P(x)]$	معادل است با	$\sim(\exists x)P(x)$	

1) Contrapositive law

2) De Morgan's laws

3) Distributive

4) Commutative

5) Associative

استنتاج با اثبات قضیه

در منطق گزاره‌ای، قواعد استنتاج را می‌توان به $wffs$ معین و مجموعه‌های $wffs$ اعمال کرد تا بین وسیله $wffs$ جدیدی به دست آورد. در زیر مثال‌هایی از قواعد استنتاج آورده می‌شوند که W ‌ها نشان دهنده $wffs$ هستند:

مدوس پونتر^۱

از $(W1 \wedge W2) \Rightarrow W2$ استنتاج می‌شود.

مدوس تولنتر^۲

از $(\sim W2 \wedge \sim W1) \Rightarrow (\sim W1)$ استنتاج می‌شود.

پروجکشن (تصویر کردن)^۳

از $(W1 \wedge W2) \Rightarrow W1$ استنتاج می‌شود.

تخصیص فرآگیر^۴

از $(\forall x)W(x)$ استنتاج می‌شود $W(c)$.

که c نماد ثابت است، عبارت عمومی "از F استنتاج می‌شود G " بین معنی است که $G \Rightarrow F$ همواره درست است (یعنی، F همواره G را نتیجه می‌دهد)؛ این عبارت به ما اجازه می‌دهد که در عبارات منطقی، G را جایگزین F کنیم.

قواعد استنتاج از روی $wffs$ داده شده، $wffs$ مشتق شده^۵ را تولید می‌کنند. در حساب گزاره‌ای، $wffs$ مشتق شده، قضایا^۶ خوانده می‌شوند، و دنباله‌ای از کاربرد قواعد استنتاج که در فرآیند اشتقاق استفاده شدند، اثبات^۷ قضیه را تشکیل می‌دهند. این مفهوم در بحث ما اساسی است، زیرا کارهای تعبیر تصویر متعددی را می‌توان بر حسب اثبات قضیه و با استفاده از حساب گزاره‌ای بیان کرد. در این روش می‌توان از قواعد استنتاج و مجموعه‌ای از واقعیات شناخته شده برای به دست آوردن واقعیات جدید یا اثبات درستی یک فرض استفاده کرد.

در حساب گزاره‌ای می‌توان از دو روش پایه‌ای برای اثبات درستی عبارات منطقی استفاده کرد.

روش اول با کار مستقیم با قالب‌های غیر گزاره‌ای و پیروی از روالی شبیه به روال اثبات عبارات ریاضی

1) Modus Ponens

2) Modus Tollens

3) Projection

4) Universal specialization

5) Derived

6) Theorems

7) Proof

انجام می شود. روش دیگر مبتنی بر تطبیق اجزای عباراتی می باشد که در شکل گزاره ای هستند. هر دو روش در مثال زیر استفاده می شوند.

مثال: فرض کنید که واقعیت های زیر را می دانیم: (۱) سطل زباله پشت میز تحریر است، و (۲) صندلی کنار میز تحریر است. فرض می کنیم که در این مورد قوانین فیزیکی زیر برقرارند: (۳) اگر x پشت y باشد، آنگاه x دیدنی نیست. واقعیات (۱) و (۲) خاص مسأله هستند، اما فرض (۳) از نوع دانشی است که مستقل از مسأله است. یعنی، تا زمانی که شرایط معینی برآورده شوند، نظیر این که x کوچکتر از y باشد، x نسبت به محل ناظر کاملاً پشت لا باشد، ولا شیخی کدر باشد، برقرار است. ما می خواهیم تنها با استفاده از دو واقعیت و یک فرض نتیجه بگیریم (ایبات کنیم) که سطل زباله دیدنی نیست.

آن دو واقعیت عبارتند از:

BEHIND(wastebasket, desk)

و

NEXT-TO(chair, desk).

براساس صورت مسأله، می دانیم که این دو واقعیت با پیوند دهنده منطقی " \wedge " به هم پیوند می خورند:

BEHIND(wastebasket, desk) \wedge NEXT-TO(chair, desk).

قانون فیزیکی در شکل گزاره ای آن عبارتست از:

$$(\forall x, y)[\text{BEHIND}(x, y) \Rightarrow \text{INVISIBLE}(x)]$$

که به آسانی می توان با استفاده از این واقعیت (از بخش قبلی) که $B \Rightarrow A \Leftrightarrow \sim A \vee B$ است، آن را به شکل غیر گزاره ای تبدیل کرد. به عبارت دیگر، عبارت قبل در شکل غیر گزاره ای آن عبارتست از

$$(\forall x, y)[\sim \text{BEHIND}(x, y) \vee \text{INVISIBLE}(x)].$$

اکنون تمام آنچه را که درباره 'این مسأله' می دانیم، به شکل ترکیب عطفی بیان می کنیم:

۷۶۷ تعبیر ۵.۹

$$(الف) (\forall x, y)[BEHIND(wastebasket, desk) \wedge NEXT-TO(chair, desk) \wedge [\sim BEHIND(x, y) \vee INVISIBLE(x)]].$$

با قرار دادن wastebasket به جای x و desk به جای y داریم:

$$(ب) BEHIND(wastebasket, desk) \wedge NEXT-TO(chair, desk) \wedge [\sim BEHIND(wastebasket, desk) \vee INVISIBLE(wastebasket)].$$

با استفاده از قاعده 'پروجکشن عبارت زیر را تیجه می گیریم^۱:

$$(ب) BEHIND(wastebasket, desk) \wedge [\sim BEHIND(wastebasket, desk) \vee INVISIBLE(wastebasket)]$$

با استفاده از یکی از قوانین توزیعی معرفی شده در بخش قبل، داریم که $(A \wedge \sim A \vee B) = (A \wedge B)$ است. بنابراین عبارت ساده شده زیر بدست می آید:

$$(ت) BEHIND(wastebasket, desk) \wedge INVISIBLE(wastebasket).$$

با کاربرد مجدد قاعده 'پروجکشن' داریم که

$$(ت) INVISIBLE(wastebasket).$$

بنابراین ثابت کرده ایم که عبارت اولیه داده شده در (الف) کاملاً معادل عبارت (ت) است. به عبارت دیگر، از اطلاعات داده شده تیجه گرفته ایم که سطل زیاله دیدنی نیست.

اکنون می خواهیم تیجه فوق را با استفاده از نمایش گزاره ای اثبات کنیم. در این روش باید ثابت کنیم که نقیض عبارت گزاره ای حکم، با واقعیت هاناساز گار است، بنابراین ثابت می شود که عبارت گزاره ای مورد نظر درست یا معتبر است. این روند به ظاهر معکوس روشنی است که معمولاً اثبات کننده های قضایای مکانیکی به کار می برند.

براساس تعریف هایی که قبل از آن شد، می توانیم دانش خود درباره این مسأله را به شکل گزاره ای زیر بیان کنیم:

- (الف) $\Rightarrow BEHIND(wastebasket, desk),$
- (ب) $\Rightarrow NEXT-TO(chair, desk),$
- (ب) $(\forall x, y)[BEHIND(x, y) \Rightarrow INVISIBLE(x)]$

(ت) INVISIBLE(wastebasket) ⇒ .

و

به خاطر آورید که می خواهیم نادرستی نقیض گزاره' INVISIBLE(wastebasket) را اثبات کنیم، که براساس تعریف قبل، به صورت \Rightarrow INVISIBLE(wastebasket) نمایش داده می شود.

پس از این که عناصر مسأله به شکل گزاره ای بیان شدند، هدف این است که طرف های چپ و راست عبارات شرطی مختلف را با هم تطبیق کنیم تا این که به یک گزاره' تهی که تناقض است، برسیم. تطبیق با جای گذاری متغیرها و به منظور یکسان سازی اتم ها انجام می شود. پس از تطبیق، گزاره حاصل که جداکننده^۱ خوانده می شود، از طرفهای چپ و راست تطبیق نشده تشکیل می شود. اگر wastebasket را جایگزین X و desk را جایگزین لا کنیم، طرف چپ (پ) با طرف راست (الف) تطبیق می کند. جداکننده عبارتست از

(ث) INVISIBLE(wastebasket)

در هر حال، تفاوت (ت) و (ث) جمله' تهی است، زیرا طرف چپ (ت) و طرف راست (ث) یکسان هستند. این نتیجه یک تناقض است، که نشان می دهد نقیض \Rightarrow INVISIBLE(wastebasket) نادرست است، بنابراین درستی قضیه' INVISIBLE(wastebasket) را ثابت می کند. □

یک نتیجه گیری مهم در حساب گزاره ای این است که قضایای درست در زمان محدودی قابل اثبات هستند. روایی الگوریتمی (اما ناکارآمد) برای پیدا کردن چنین اثبات هایی در ابتدا به وسیله' Herbrand[1930] پیشنهاد شد. روشهای بسیار کارآمدتر برای اثبات قضیه در سه دهه بعد بواسیله Robinson[1965] کشف شد. وی یک قاعده استنتاج که تفکیک^۲ خوانده می شود، را پیشنهاد کرد. Robinson نشان داد که قاعده تفکیک کامل بودن^۳ را حفظ می کند، با این مفهوم که تمام قضایای صحیح اثبات پذیر هستند؛ و صحیح بودن^۴ را حفظ می کند، با این مفهوم که هیچ قضیه' ناصحیحی اثبات پذیر نیست. همان طور که مثال قبل نشان می دهد، قاعده' تفکیک ساده است. پس از این که عناصر مسأله به شکل گزاره ای بیان شدند، این روش مقدمه ها و نتیجه های جملات استنتاجی مختلف که می توانند با هم تطبیق شوند، را جستجو می کند. تطبیق با جای گذاری متغیرها و با هدف یکسان سازی اتم ها انجام

می شود. همان طور که قبل اذکر شد، پس از تطبیق، جمله حاصل که جداگانه خوانده می شود، مشکل از طرف های چپ و راست تطبیق نشده است. حال اثبات قضیه شامل جداسازی^۱ عبارات با هدف تولید عبارت تهی که یک تناقض است، می باشد. در عمل، قسمت چشمگیری از متون چاپ شده در مورد اثبات قضیه، به ایجاد الگوریتم هایی برای کشف کارآمد جداگانه ها در یک پایگاه دانش معین می پردازند.

مثال: فرض کنید که اطلاعات زیر قسمتی از پایگاه دانش یک سامانه تعبیر تصاویر هوایی باشد:

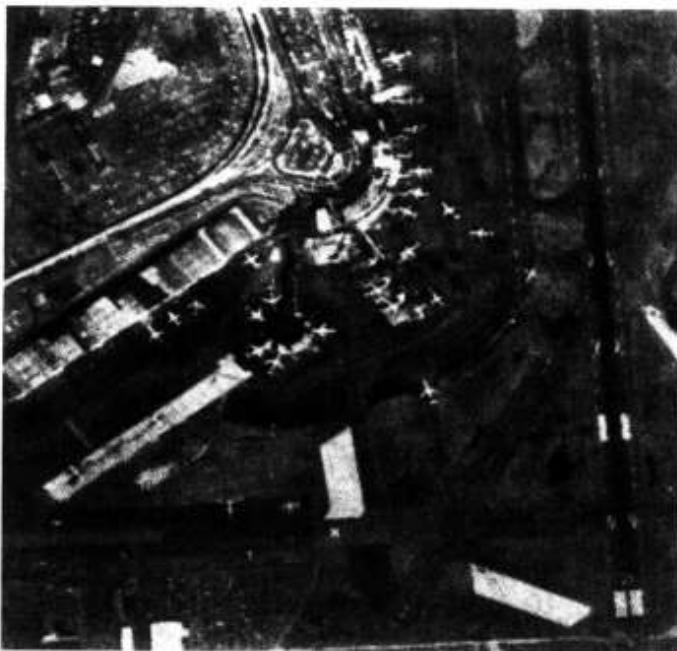
۱. تمام تصاویر فرودگاه های تجاری دارای (یک یا چند) باند پرواز هستند.
۲. تمام تصاویر فرودگاه های تجاری دارای چند هوایپما هستند.
۳. تمام تصاویر فرودگاه های تجاری دارای (یک یا چند) ساختمان هستند.
۴. در هر فرودگاه تجاری، دست کم یک ساختمان پایانه وجود دارد.
۵. ساختمانی که محصور بین هوایپماها باشد و هوایپماها رو به آن باشند، ساختمان پایانه است.

این اطلاعات را به "مدلی" از فرودگاه تجاری به شکل شرطی زیر فشرده می کنیم:

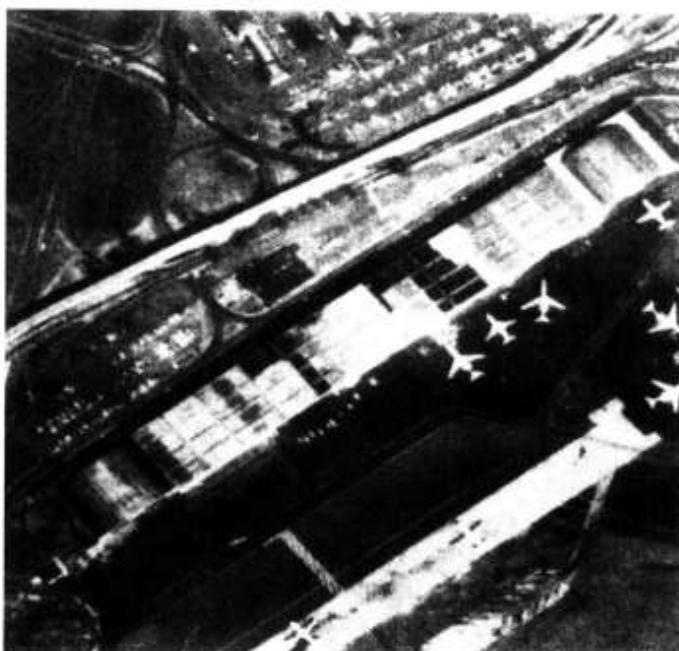
$$(\forall x)[\text{CONTAINS}(x, \text{runway}(s)) \wedge \text{CONTAINS}(x, \text{airplanes}) \wedge \\ \text{CONTAINS}(x, \text{building}(s)) \wedge \text{POINT-TO}(\text{airplanes}, \\ \text{building}(s))] \Rightarrow \text{COMAIRPORT}(x).$$

اطلاعات داده شده در جمله^(۴) مستقیماً در مدل استفاده نمی شوند. در عوض، مفهوم آن به طور ضمنی در دو شرطی که الزام می کنند (یک یا چند) ساختمان وجود داشته باشد و هوایپماها نیز رو به ساختمان (ها) باشند، وجود دارد. شرط اخیر ساختمانی را به عنوان پایانه تعیین می کند.

فرض کنید تصویر هوایی شکل ۹. ۳۵(الف) به ما داده شده است (نمای نزدیک آن در شکل ۹. ۳۵(ب) منظر بهتری از جزئیات تصویر به ما می دهد) و فرض کنید «موتور تشخیصی» داریم که قادر به تشخیص اشیای گوناگون در تصاویر هوایی است. به دو طریق می توانیم شکل ۹. ۳۵(الف) را تعبیر کنیم: (۱) می توانیم پرسیم، "این تصویر متعلق به چیست؟" یا (۲) می توانیم پرسیم، "آیا تصویری از فرودگاه تجاری است؟". به طور کلی باروش های فعلی نمی توانیم به سوال اول پاسخ گوییم. در حالت کلی پاسخ به سوال دوم نیز بسیار مشکل است، اما اگر حوزه بحث را محدود کنیم، پاسخ دادن به آن



(الف)



(ب)

شکل ۳۵.۹ (الف) تصویر هوایی فرودگاه ملی واشنگتن؛ (ب) جزئیات.

(McKeown, Harvey & McDermott[1985] از)

ساده تر می شود. بویژه، روش مدل گرا^۱ که در آن اشیاء مورد علاقه به صورتی که در بالا نشان داده شد، مدل شده اند، مزیت مهمی دارد که می توان با استفاده از آن عمل موتور تشخیص راهداشت کرد. در اینجا موتور تشخیص باید بتواند سه گروه اشیا را تشخیص دهد: باندهای پرواز، هواپیماها، و ساختمان ها. اگر ارتفاع عکسبرداری معلوم باشد، که معمولاً چنین است، باز هم جستجو برای این سه نوع اشیا ساده تر می شود، زیرا از مقیاس های متناسب با ابعاد اشیاء می توان برای هدایت فرآیند تشخیص استفاده کرد. خروجی تشخیص دهنده ای که روی شکل ۹ .(الف) کار کرده و از مدل فوق برای فرودگاه تجاری استفاده می کند، اطلاعات رابه شکل زیر ارائه می دهد: CONTAINS (image, runway)، CONTAINS (image, buildings)، و CONTAINS (image, airplanes). پردازش بیشتر اشیای تشخیص داده شده، درستی یا نادرستی عبارت (airplanes, buildings) POINT (airplanes, buildings) را تعیین می کند. اگر نادرست باشد، روال مبتنی بر این مدل ساده متوقف می شود. در غیر این صورت کار را ادامه می دهد تا با بررسی درستی عبارت COMAIRPORT (image) تعیین کند که شکل ۹ .(الف) تصویر یک فرودگاه تجاری است یا خیر. این کار را با استفاده از اثبات قضیه بوسیله 'تفکیک'^۲، و با اطلاعات زیر که از تصویر استخراج شده اند، شروع می کنیم:

- i. $\Rightarrow \text{CONTAINS}(\text{image}, \text{runway})$,
- ii. $\Rightarrow \text{CONTAINS}(\text{image}, \text{airplanes})$,
- iii. $\Rightarrow \text{CONTAINS}(\text{image}, \text{buildings})$,
- iv. $\Rightarrow \text{POINT-TO}(\text{airplanes}, \text{buildings})$,

و در نهایت نقیض عبارتی را که می خواهیم اثبات کنیم، به دست می آوریم:

v. COMAIRPORT(image) \Rightarrow .

با ذکر این که اگر image را جایگزین x کنیم، یکی از عبارات سمت چپ مدل با طرف راست (i) تطبیق می کند، کار را شروع می کنیم. آنگاه جداگانه عبارتست از:

$$[\text{CONTAINS}(\text{image}, \text{airplanes}) \wedge \text{CONTAINS}(\text{image}, \text{buildings}) \wedge \text{POINT-TO}(\text{airplanes}, \text{buildings})] \Rightarrow \text{COMAIRPORT}(\text{image}).$$

1) A model- driven approach

به طور مشابه، یکی از عبارات سمت چپ جداکننده با طرف چپ (ii) تطبیق می کند. آنگاه جداکننده جدید عبارتست از:

$$\begin{aligned} & [\text{CONTAINS}(\text{image}, \text{buildings}) \wedge \text{POINT-TO}(\text{airplanes}, \text{buildings})] \\ & \Rightarrow \text{COMAIRPORT}(\text{image}). \end{aligned}$$

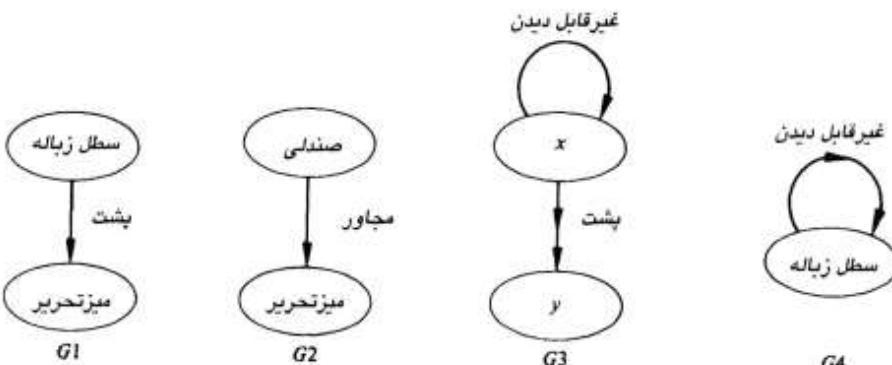
سپس کاربرد (iii) و (iv) جداکننده زیر را می دهد:

$$\Rightarrow \text{COMAIRPORT}(\text{image}).$$

در نهایت، تفکیک این نتیجه و (v) گزاره 'تهی را می دهد، بنابراین تناقضی را تولید می کند. این تناقض درستی عبارت COMAIRPORT(image) را اثبات کرده و نشان می دهد این تصویر در واقع تصویری از یک فرودگاه تجاری است (یعنی، با مدل ما از فرودگاه تجاری مطابقت می کند).

۴.۵.۹ شبکه های معنایی

شبکه های معنایی گراف های بر چسب دار جهت داری هستند که در آنها معمولاً گره ها نمایانگر اشیا یا متغیرها و کمانها نمایانگر ارتباط بین گره ها هستند. شبکه هایی از این نوع چند مزیت دارند، از جمله این که روش کارآمد شهودی و تصویری برای نمایش دانش هستند. به علاوه، چون نمایش پایه ای یک گراف است، از روش های تطبیق و بر چسب زنی گراف می توان برای کار با عناصر یک حوزه 'مسأله' استفاده کرد. شکل ۹.۳۶، نمایش شبکه 'معنایی' واقعیت ها در مسأله سطل زباله را نشان می دهد.



شکل ۹.۳۶ بخش های نمودار معنایی برای مسأله سطل زباله

شبکه های معنایی جایگزین مناسبی برای اجرای روش های تعییر می باشند. گرچه راه های متعددی برای استفاده از شبکه ها به این منظور وجود دارند، ولی در این بخش روی شبکه هایی که برای نمایش و پردازش دانش بیان شده به شکل عبارات گزاره ای به کار می روند، تمرکز می کنیم. این روش مفاهیم پایه ای استفاده از شبکه های معنایی را روشن می سازد و به حفظ پیوستگی بحث کمک می کند.

ساخت شبکه معنایی از عبارات گزاره ای ساده است. روش این است که هر عبارت گزاره ای را به صورت قسمتی از شبکه' معنایی نمایش دهیم (یعنی، هر عبارت گزاره ای را به صورت زیر گراف جداگانه ای نمایش دهیم). گره های شبکه دارای متغیرها و ثابت ها (عمولاً آشیاء) هستند، و کمان های برچسب دار جهت دار ارتباطات دودویی بین گره هارا نشان می دهند. هر کمان اگر نمایانگر رابطه ای در قسمت نتیجه (راست) عبارت گزاره ای باشد، به صورت پیکان تک نوک کشیده می شود، و اگر نمایانگر رابطه ای در قسمت شرط (چپ) عبارت گزاره ای باشد، به صورت پیکان دو نوک کشیده می شود. شکل ۹ . ۳۶، قسمت های گرافی متناظر با عبارات (الف)- (ت) مثال سطل زباله را نشان می دهد. ما رابطه' یکانی، نظری (x)، INVISIBLE(x)، را با کلک نوشتاری (x) به رابطه' دوتایی تبدیل می کنیم (به منظور یکنواختی در نمایش های گرافی).

استنتاج کردن با شبکه های معنایی به شکل ۹ . ۳۶ ساده و از نظر مفهومی خوشایند است. با تطبیق متغیرها و ثابت ها، گراف ها در هم ادغام می شوند و پیکان های دو نوک، پیکان های تک نوک را همراه با گره های انشعاب شان اگر معادل باشند، لغو^۱ می کنند. همانند بخش قبلی، هدف این است که با به کار گیری یک عبارت گزاره ای نقیض شده^۲ در شروع فرآیند استنتاج، به یک بخش خالی برسیم. شکل ۹ . ۳۷. نتیجه ادغام بخش های گرافی G1 و G3 (که با G13 نشان داده می شود) از شکل ۹ . ۳۶ را نشان می دهد. ادغام G13 و G4 یک بخش خالی را نتیجه می دهد، که نشان می دهد (مانند گذشته ثابت می کند) سطل زباله قابل دیدن نیست.

مثال: اجازه دهید به مسأله فرودگاه باز گردیم. شکل ۹ . ۳۸ (الف) مدل ما از فرودگاه تجاری را به صورت شبکه' معنایی نمایش می دهد. دو گره' متفاوت برای "هوایپماها" نشان داده می شوند. به طور معادل درست خواهد بود که هر گره' هوایپما را با کمانی که از x به آن می رود و برچسب CONTAINS (حاوی) دارد و کمان دیگری با برچسب POINT-TO (رو به) که از گره' هوایپما به گره' ساختمان ها

1) Cancel

1) Negated



شکل ۳۷.۹ اجرای استنتاج با ادغام بخش های گراف

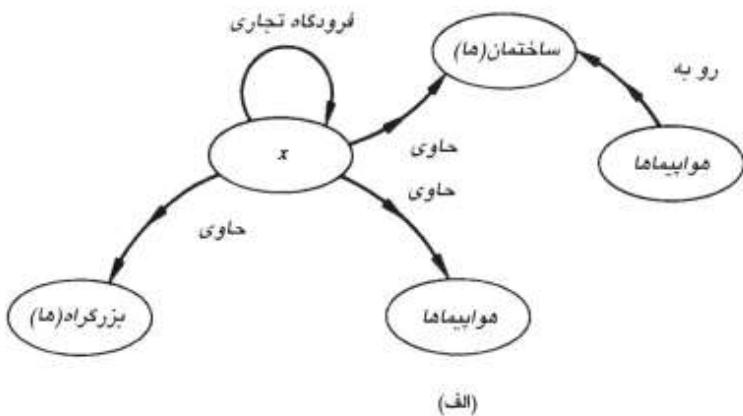
می زود، نشان دهیم. در هر حال، برای برقراری مدل باید هر دو شرط برآورده شوند، بنابراین نشان دادن دو گره با برچسب "هوایماها" ثبت وقایع را ساده می کند.

شکل ۹ . (۳۸) (ب) واقعیات استخراج شده از تصویر شکل ۹ . (الف) رانشان می دهد. به
پیکان های تک نوک که نتیجه نمایش های گزاره ای (۱) - (۷) هستند، توجه کنید. شکل ۹ . (۳۸) (ت)
نقیض واقعیتی را که می خواهیم ثابت کنیم، نشان می دهد [نمایش گزاره ای (۷)]. شکل ۹ . (۳۸) (ت)
مدل را پس از این که داده های تصویری در آن جایگزین شده و ادغام های مناسب انجام شدند، نمایش
می دهد. در نهایت، شکل ۹ . (۳۸) (ث) نشان می دهد که ادغام بخش های گراف در شکل های ۹ . (۳۸) (پ)
و (ت) یک بخش خالی را نتیجه داده و بنابراین ثابت می کند که تصویر مورد بررسی در واقع تصویر یک
فرودگاه تجاری است.

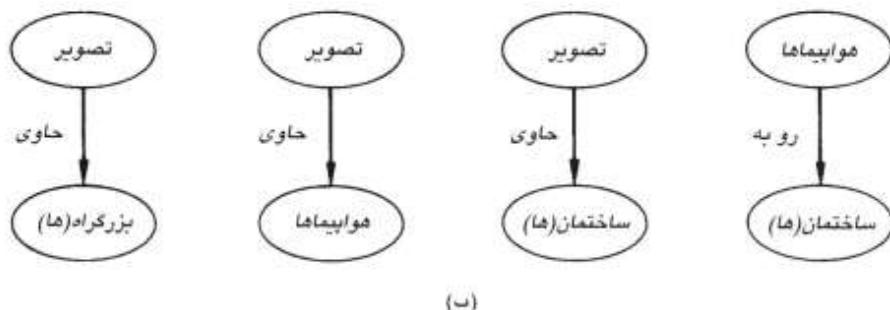
۵.۵.۹ سامانه‌های فرآوری (خبره)

سامانه های فرآوری راهی جایگزین برای انواع استنتاج هایی که در بخش های ۹ . ۵ . ۳ و ۴ . ۵ . ۹ تشریح شدند، پیشنهاد می کنند. سامانه های فرآوری نیز همانند حساب گزاره ای و شبکه های معنایی به تطبیق نیاز دارند تا بدین وسیله شناسایی کنند که کدام استنتاج را انجام دهنند. با این حال کنش های این سامانه پس از پیدا کردن یک تطبیق بسیار کلی تر هستند. در واقع، کنش هایی با پیچیدگی دلخواه، ممکن

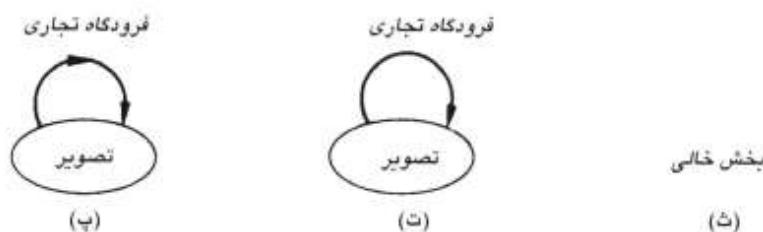
این انعطاف‌پذیری پاسخ، علاقه زیادی به توسعه سامانه فرآوری ایجاد کرده است. به طور خاص سامانه‌های خوب در کاربردهای گوناگون، از بردازش تصویر و بازرسی صنعتی، گرفته تا تشخوص‌های



(الف)



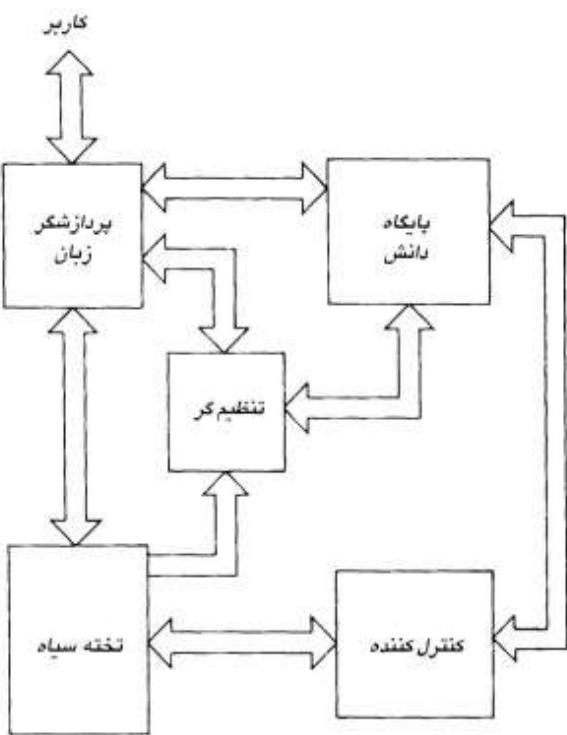
(ب)



شکل ۳۸.۹ (الف) مدل شبکه' معنایی فروگاه تجاری؛ (ب) داده های استخراج شده از تصویر هوایی شکل ۳۵ (الف)؛ (پ) تدقیق واقعیتی که باید اثبات شود؛ (ت) نتیجه جایگذاری داده ها در مدل و ادغام؛ (ث) نتیجه' ادغام (پ) و (ت).

پژوهشکی و کنترل فرآیند، به طور موقعی استفاده شده اند. سامانه های خبره همان سامانه های انسانی - ماشینی با مهارت خاص در حل مسأله هستند. بنابراین نام سامانه خبره¹ نشان دهنده این واقعیت است که این سامانه ها معمولاً مبتنی بر دانش افراد خبره در یک زمینه' معین هستند.

1) expert system



شکل ۹.۳۹. اجزای پایه‌ای سامانه خبره

شکل ۹.۳۹ اجزای اساسی یک سامانه خبره را نشان می‌دهد. پردازشگر زبان^۱ به عنوان واسطه ارتباط بین کاربر و سامانه می‌باشد. کاربر از طریق یک زبان مسأله گرا^۲، که معمولاً در یک قالب محدود شبه انگلیسی است، با سامانه خبره ارتباط برقرار می‌کند. برای تکمیل ارتباط زبانی ممکن است یک واسطه گرافیکی نیز استفاده شود. پردازشگر زبان، ورودی‌های فراهم شده بوسیله 'کاربر را تعبیر می‌کند و اطلاعات تولید شده بوسیله 'سامانه را قالب‌بندی می‌کند. تنظیم گر^۳ کنش‌های سامانه را برای کاربر شرح می‌دهد. مثلاً به سوالاتی درباره 'این که چرا یک نتیجه به دست آمد یا چرا نتیجه' دیگر رد شد، پاسخ می‌دهد. همچنین تنظیم گر نقش بسیار مهمی در طراحی اولیه و اشکال زدایی سامانه خبره بر عهده دارد.

'تخنه سیاه'^۴ در حال کار است که برای ثبت مواردی نظری داده‌های فعلی درباره 'یک مسأله

1) *Language processor*2) *Problem-oriented*3) *Justifier*4) *Blackboard*

خاص، فرض های در حال کار، و تصمیم های میانی استفاده می شود. اصطلاح های حافظه' موقت^۱ و حافظه' کوتاه مدت^۲ نیز غالباً برای اشاره به تخته سیاه به کار می روند. همان طور که در بخش ۹.۵.۲.۵.۹ تعریف شد، پایگاه دانش^۳ (زیر مجموعه ای از پایگاه دانش شکل ۹.۱) حاوی دانش های رویه ای^۴، دیداری و جهانی است. درجه' همپوشانی دانش مورد استفاده در سامانه خبره با دانش مورد استفاده در سایر عناصر سامانه پردازش تصویر، به کاربرد بستگی دارد. همچین پایگاه دانش، حاوی قواعد مورد نیاز سامانه در حل مسأله می باشد. شیوه کنترل، حاوی دانش کلی حل مسأله و راهبردهای حل تناقض^۵ است. دانش کلی حل مسأله، که گاهی موتور استنتاج^۶ خوانده می شود، شالوده' چگونگی تطبیق قواعد دانش، الگوریتم ها، واقعیت های تاکنون شناخته شده درباره' یک کار را به منظور رسیدن (با استنتاج) یک جواب را تشکیل می دهد. این بدان معناست که شیوه کنترل، در انتخاب و کاربرد ابزارهای حل یک مسأله به عنوان راهبر^۷ عمل می کند. وقتی دو یا چند قاعدة همزمان شروع (چکانده) می شوند، روش های رفع تناقض در شیوه کنترل مورد نیاز هستند. روش حل تناقض به کاربرد بستگی دارد، اما بعضی روش های کلی عبارتند از: مرتب سازی قاعدة^۸، که در آن قاعدة' دارای بالاترین حق تقدم برای آتش^۹ (که بعداً تعریف می شود) انتخاب می شود، و مرتب سازی زمانی^{۱۰}، که در آن آخرین قاعدة' استفاده شده مجاز است که آتش شود. قواعد در سامانه خبره معمولاً به شکل

اگر (شرط^{۱۱}) آنگاه (کنش ها^{۱۲})

هستند که شرایط و کنش ها به صورت عبارات عطفی بیان می شوند. به عبارت دیگر، قاعدة فوق بدین صورت خوانده می شود: اگر شرط های او^{۱۳} و ... و m برقرار باشند، آنگاه کنش های او^{۱۴} و ... و k اجرا می شوند. ^{۱۵} مثلاً، یک قاعدة' نمونه برای ادغام دو ناحیه' تصویری را در نظر بگیرید:

R(107) اگر ابعاد ناحیه ها کوچک باشند، و نواحی مجاور هم باشند، و اختلاف بین

واریانس های سطح خاکستری شان کوچک باشد،

آنگاه دو ناحیه را ادغام کنید.

^۱ کلمه "و" در این متن عملگر AND منطقی است، که در بخش های ۹.۵.۵.۴ و ۹.۳.۵ با "آنگاه" نشان داده می شود. در بیان قواعد سامانه خبره، کاربرد "و" متدائل است، زیرا به زیان متن تزدیکتر است.

1) Temporary memory

2) Short-term memory

3) Knowledge base

4) Procedural

5) Conflict-resolution

6) Inference engine

7) Strategist

8) Rule ordering

9) Firing

10) Recency ordering

11) Condition

11) Action

مفاهیم "کوچک" و "مجاور" باید از قبل تعریف و در جایی که قواعد در هنگام نیاز بتوانند به آنها دست یابند، ذخیره شوند. این نوع اطلاعات در پایگاه دانش ذخیره می شود. فرض می شود که نواحی، شناسایی شده و توصیف گرهایشان نیز محاسبه شده اند. یعنی؛ فرض می کنیم که بخش بندی (فصل ۷) و توصیف (فصل ۸) اجرا شده اند. در نهایت، این قاعده سه شرط و تنها یک کنش دارد. قواعدی با چند شرط که به یک کنش متبوع می شوند، در سامانه های پردازش تصویر متداول هستند.

وقتی که تمام شرایط یک قاعده با هم برآورده شوند، گفته می شود که آن قاعده چکانده شده^۱ است. وقتی کنش ها اجرا شوند، گفته می شود که قاعده آتش^۲ شده است. این واقعیت که یک قاعده چکانده شده است، بدین معنی نیست که آن قاعده به طور خودکار آتش شده است. با مجموعه ویژه ای از شرایط می توان بیش از یک قاعده را چکاند، که در این حالت باید میستم وارد مرحله حل تناقض شود تا تعیین کند کدام قاعده برای آتش شدن تقدیم دارد. فرض بر این است که در یک لحظه تنها یک قاعده امکان آتش دارد.

شباهت بین قاعده های سامانه خبره به شکلی که اخیراً توصیف شد و عباراتی به شکل گزاره ای که قبل از تشریح شد، باید آشکار باشد. در واقع، تعیین و خواندن قواعد خبره معمولاً بسیار آسان تر است. مثلاً قانون فیزیکی در مساله سطل زباله که اخیراً تشریح شد، را می توان در شکل قاعده ای به صورت:

اگر شئی \neq پشت شئی لا باشد،

آنگاه شئی \neq قابل دیدن نیست.

بیان کرد. وقتی واقعیت های معینی داده شوند، سامانه خبره تلاش می کند که آن واقعیت ها را با شرایط پایگاه قاعده تطبیق کند. اگر تمام شرایط یک قاعده تطبیق کنند، آن قاعده چکانده می شود. اگر هیچ تناقضی رخ ندهد، کنش مربوط به آن قاعده فعال شده و بنابراین آن قاعده را آتش می کند. همانند قاعده فوق، ممکن است که خود کنش چیزی بیش از یک عبارت خبری نباشد. با توجه به واقعیت "سطل زباله پشت میز تحریر است"، سامانه خبره سطل زباله را با \neq و میز تحریر را با \neq تطبیق داده و قاعده را می چکاند. در این مورد، کنش، عبارت خبری "سطل زباله غیر قابل دیدن است" خواهد بود. درجه پذیرش عبارات در قالب های شبه انگلیسی (یا هر زیان طبیعی دیگر) به کار آمدی سامانه خبره بستگی

دارد. چون سعی براین است که سامانه های خبره "سهل الاستفاده" باشند، تعامل با آنها معمولاً بسیار ساده‌تر از سامانه های مبتنی بر منطق گزاره‌ای یا شبکه‌های معنایی است.

مثال: اجازه دهید که مدل سامانه خبره مسأله تعییر فرودگاه را که قبلاً تشریح شد، در نظر بگیریم. این مدل در بیان به شکل قاعده‌ای، بدین صورت است:

اگر \times تصویر باشد و

\times حاوی باند(های) پرواز باشد و

\times حاوی هوایپماها باشد و

\times حاوی ساختمان(ها) باشد و

(بعضی) هوایپماها را و به ساختمان(ها) باشند،

آنگاه \times تصویر یک فرودگاه تجاری است.

در عمل، روال استفاده از این سامانه ویژه برای تعییر تصویر هوایی با چنین پرسشی شروع می‌شود که: آیا تصویر، مربوط به یک فرودگاه است؟ در این مورد، خود پرسش، اولین واقعیت را که " \times تصویر باشد" شروع می‌شوند، تطبیق کند. در این مثال ساده، تنها نیاز است که تعیین کنیم آیا سه شرط دیگر برآورده می‌شوند یا خیر. کنترل کننده سعی می‌کند که شرایط آن دسته از قواعدی را که با "اگر \times تصویر باشد" شروع می‌شوند، تطبیق کند. را که برای جستجوی نمونه‌های مواردی در تصویر که برای برآورده شدن شرایط قاعده نیاز هستند، مدیریت خواهد کرد. برای نمونه، کنترل کننده، روند تشخیص برای پیدا کردن باندهای پرواز در تصویر آغاز می‌کند. این روند به نوبه خود برای استخراج و مشخص کردن نواحی منتخبی که احتمالاً باند پرواز هستند، به الگوریتم‌های بخش بندی و توصیف نیاز دارد. پس از پیدا کردن باندهای پرواز، سامانه، جستجو برای پیدا کردن ساختمان‌های اساسی شروع می‌کند، و بدین ترتیب الى آخر. اگر تمام شرایط برآورده شوند، قاعده(ها) چکانده می‌شوند. اگر تنها یک قاعده چکانده شود، هیچ حل تناقضی نیاز نخواهد بود و سامانه تا مرحله آتش کردن قاعده و در این مورد، تولید نتیجه "تصویر از یک فرودگاه تجاری است."، به کار آدame می‌دهد.



مثال قبلی مراحل اساسی را که یک سامانه خبره نوعی در تعیین برآورده شدن یکی از قواعد هاییش طی می‌کند، نشان می‌دهد. در سامانه های پیچیده^۱ مبتنی بر قاعده که صدها، یا حتی هزارها، قاعده دارند، مشکلات اصلی عبارتند از: جستجوی قاعده هایی که می خواهیم تطبیق کنند، ردگیری این که در هر

لحظه کدام قاعده ها در هر حال پردازش هستند، مدیریت منابع موجود برای پردازش تصویر، و اجرای حل تناقض در هنگام نیاز. در بخش مرجع که در ادامه می آید، مثالهایی از سامانه های خبره را ذکر می کنیم که در انواع گوناگونی از کاربردهای پردازش تصویر محدود، اما عملی، موفقیت آمیز بوده اند.

۶.۹ نکات پایانی

گرچه مطالب این فصل مقدماتی هستند، ولی عناوین مطرح شده در آن برای فهم حالت فعلی تشخیص و تعبیر تصویر اساسی هستند. در هر حال، همان طور که احتمالاً متوجه شده اید، ماتنها قسمتی از مطالب را گفته ایم. سالیان درازی است که مباحث تشخیص و تعبیر تصویر در مباحث گوناگون در حال بررسی و توسعه هستند. در نتیجه، هزارها مقاله و صدھا کتاب حاوی روش های فراوان نظری و تجربی وجود دارند که برای اجرای وظایف گوناگون تحلیل خودکار تصویر پیشنهاد شده اند.

با وجود این میزان فعالیت زیاد، هنوز هم این مبحث مسائل حل نشده ای دارد. بویژه، راه حل مسائل در تحلیل تصویر خاص مسأله هستند، بنابراین قابلیت توسعه و پیشرفت با استفاده از نتایج سایر کارها، محدود می شود. پیش بینی می شود در آینده نیز طراحی سامانه های تحلیل تصویر نیازمند ترکیبی از هنزو علم باشد. از آن نظر، مطالب مطرح شده در این کتاب را می توان به عنوان پایه ای قوی برای حل مسائل پردازش تصویر در طیف وسیعی از زمینه ها در نظر گرفت.

مراجع

- متون عمومی برای شرح مقدماتی و مطالب بخش ۹.۱ عبارتند از: Ballard & Brown[1982]
- مطالب Gonzalez[1991]، Abidi, Eason & Chang[1989]، Fu, Gonzalez & Lee[1987]
- بخش ۹.۲ تا انتهای بخش ۹.۳.۲ بر مبنای نظریه 'تشخیص الگو' هستند. مراجع این بخش ها Duda& Hart[1973]، Gonzalez & Thomason[1978]، Tou & Gonzalez[1974] و Fu[1982] هستند.
- مراجع پایه ای در مورد شبکه های عصبی (بخش ۹.۳.۳)، کتابهای McClelland & Rumelhart[1986]، Harston, Maren& Pap[1990]، Pao[1989]، McClelland & Ramelhart[1986]
- Skapura & Freeman[1991] و Khanna[1990] می باشند. شماره 'ویژه ای از مجله Computer' با عنوان "شبکه های عصبی مصنوعی" حاوی چند مقاله' مروری^۱ جالب است.

مسائل ۷۸۱

مثالی که درباره تشخیص شکلها ای اعوجاج باقته است، برگرفته از Gupta و همکاران [1990] می باشد. مطلب بخش ۹ . ۳ . ۳ . مقدماتی است. درواقع، مدل شبکه عصبی مورد استفاده در متن کتاب یکی از مدل های متعددی است که در طول سال ها پیشنهاد شده اند. کتاب هایی که قبل اذکر شدند، همراه با مقاله ای از Lippman [1987]، مرور خوبی بر مدل های گوناگونی که اکنون استفاده می شوند، داشته اند. در هر حال مدل مطرح شده در این فصل نماینده ای از سایر مدل ها بوده و به طور گسترده ای در پردازش تصویر استفاده می شود.

مطلوب بخش ۹ . ۱ . از [1980] Guzman & Bribiesca و مطلب بخش ۹ . ۴ . از Yang & Sze [1981] است. مراجع بخش ۹ . ۳ . ۴ . Thomason & Gonzalez [1987] و Fu [1982] رواج آموزش خودکاره محدود از روی رشته های نمونه، از [1972] Feldman & Biermann می باشد. برای مطالعات اضافی درباره مطلب بخش ۹ . ۱ . ۵ . Chang [1989] را ببینید. دسته بندی های گوناگون دانش که در بخش ۹ . ۲ . تشریح شدند، از Adimari و همکارانش [1988] می باشد. مراجع بخش های ۹ . ۳ . ۵ . ۶ . ۹ . ۴ . ۵ . عبارتند از: Nilsson [1971, 1980]، Brown & Ballar [1982]، Chang [1989]. برای مثالی جامع از شبکه معنایی، Niemann و همکارانش [1990] را ببینید. مراجع کلی برای بخش ۹ . ۵ . ۵ . ۵ .، Nilsoon [1980]، Hayes-Roth و همکارانش [1983] هستند. همانطور که در بخش ۹ . ۵ . بیان شد، سامانه های خبره "سبتاً پیچیده" ای برای تعبیر تصویر ساخته شده اند. در مقاله های Wixon, Wilson & McKeown [1985]، Mc Dermott, Harvey & McKeown [1989]، Talashi & Nagao [1980]، و Huertas و همکارانش [1990] نوع کاربرد، کتاب [1980]، و مقاله ای از Lewis & Goodson [1990] در رابطه با ارتباط سایر مثالهای نوعی از کاربرد سامانه های خبره برای تعبیر تصاویر هوایی پیچیده ارائه شده است. در ارتباط با این نوع کاربرد، کتاب [1980]، و مقاله ای از Lee & Yuan [1987] در رابطه با استخراج تحلیل تصاویر تراشه های LSI (تجمع در مقیاس بزرگ)، Brzakovic و همکارانش [1991] در رابطه با آشکارسازی لبه.

مسائل

۹ . ۱ . الف) توابع تصمیم دسته بندی کننده حداقل فاصله برای الگوهای شکل ۹ . ۲ را محاسبه کنید.

پردازش تصویر رقemi ۷۸۲

ب) سطوح تصمیم ایجاد شده با توابع تصمیم قسمت (الف) را رسم کنید.

۴.۹ نشان دهید که معادلات $(9.4-3)$ و $(9.5-3)$ نیز همان کار را انجام می دهند.

۳.۹ نشان دهید که سطح حاصل از معادله $(9.6-3)$ ، عمود منصف خطی است که نقاط m و m' را به هم متصل می کند.

۴.۹ نشان دهید که چگونه با استفاده از M بانک مقاومتی (M ، تعداد دسته ها است)، و یک اتصال جمع کننده در هر بانک (برای جمع کردن جریان ها)، و یک انتخابگر بیشینه که می تواند بیشینه M ورودی، که ورودیها جریان هستند، را انتخاب کند، می توان دسته بندی کننده حداقل فاصله تشریح شده در رابطه با شکل ۹.۷ را پیاده سازی کرد.

۵.۹ نشان دهید که ضریب همبستگی معادله $(9.8-3)$ مقادیری در بازه $[1, 1]$ دارد.
(راهنمایی: t_1 را به شکل برداری بیان کنید.)

۶.۹ در یک آزمایش، تصاویر دودویی حباب هایی که شکل بیضی مانند دارند، تولید می شود. حبابها در سه اندازه هستند، که مقدار متوسط محورهای اصلی بیضی ها $(1.3, 0.7)$ ، $(1.0, 0.5)$ و $(0.75, 0.25)$ هستند. طول این محورها $\pm 10\%$ حول مقدار متوسطشان تغییر می کند. (تصویری نوعی که حاوی این حبابها است، در صفحه بعد دیده می شود). سامانه تحلیل تصویری طرح کنید که بتواند بیضی های ناقص یا روی هم افتاده را حذف کند و سپس بیضی های منفرد باقی مانده را در یکی از سه دسته معین شده قرار دهد. جواب خود را به شکل نمودار بلوکی نشان دهید، و جزئیات خاص راجع به عمل هر بلوک را بیان کنید.



مسئله را با استفاده از دسته بندی کننده حداقل فاصله حل کنید، و به روشنی بیان کنید که

چگونه نمونه های آموزشی را بدست می آورید و چگونه از این نمونه ها برای آموزش دسته بندی کننده استفاده می کنید.

۷.۹

فرض کنید که دسته های الگوی زیر تابع چگالی احتمال گوسی دارند:

$$\omega_1: \{(4, 4)^T, (6, 4)^T, (6, 6)^T, (4, 6)^T\} \quad \omega_2: \{(0, 0)^T, (2, 0)^T, (2, 2)^T, (0, 2)^T\}$$

(الف) فرض کنید که $P(\omega_1) = P(\omega_2) = 1/2$ و معادله مرز تصمیم بین این دو دسته را بدست آورید.

(ب) مرز را رسم کنید.

۸.۹

مسئله ۹ . ۷ را تکرار کنید، اما از دسته های الگوی زیر استفاده کنید:

$$\omega_1: \{(-2, 0)^T, (0, -2)^T, (2, 0)^T, (0, 2)^T\} \quad \omega_2: \{(-1, 0)^T, (0, -1)^T, (1, 0)^T, (0, 1)^T\}$$

توجه کنید که این دسته ها جداپذیر خطی نیستند.

۹.۹

مسئله ۹ . ۶ را تکرار کنید، اما از دسته بندی کننده بیز استفاده کنید (چگالی ها را گوسی فرض کنید). به روشنی بیان کنید که چگونه نمونه های آموزشی را به دست می آورید و چگونه از این نمونه ها برای آموزش دسته بندی کننده استفاده می کنید.

۱۰.۹

تابع تصمیم بیز $p(x/\omega_i)P(\omega_i) = p(x/\omega_i)$ ، $d_i(x) = j$ ، با استفاده ازتابع $T(x) = 1 - 0.5d_i(x)$ بدست آمدند. ثابت کنید که این تابع تصمیم، احتمال خطوارا کمینه می کنند. (راهنمایی: احتمال خطای e $p(e) = 1 - p(c)$ است، که $p(c)$ احتمال صحیح بودن است.). برای بردار الگوی x که به دسته ω_i تعلق دارد، $p(c/x) = p(\omega_i/x)$ را بیابید و نشان دهید که وقتی $p(x/\omega_i)P(\omega_i)$ بیشینه باشد، $p(c/x) = p(e)$ کمینه است.

(الف) الگوریتم پرسپترون را به دسته های الگوی زیر اعمال کنید:

$$\omega_1: \{(0, 0, 0)^T, (1, 0, 0)^T, (1, 0, 1)^T, (1, 1, 0)^T\}$$

$$\omega_2: \{(0, 0, 1)^T, (0, 1, 1)^T, (0, 1, 0)^T, (1, 1, 1)^T\}$$

$$\text{فرض کنید } c = 1 \text{ و } w(1) = (-1, -2, -2, 0)^T$$

(ب) سطح تصمیم بدست آمده در (الف) را رسم کنید. دسته های الگو و طرف مثبت سطح تصمیم را نشان دهید.

۱۲.۹

الگوریتم پرسپترون داده شده در معادلات $(34-3) - (36-3)$ را می توان با

ضرب الگوهای دسته' w در ۱- به شکل مختصرتری بیان کرد، که در این مورد مراحل تصحیح الگوریتم بدین صورت است که اگر $0 > w^T(k)y(k)$ ، آنگاه $w(k+1) = w(k) + c$ و در غیر این صورت، $w(k+1) = w(k) + cy(k)$. این بیان یکی از چند بیان الگوریتم پرسپترون است و می توان آن را با شروع از معادله' کلی نزول گرادیان، یعنی

$$w(k+1) = w(k) - c \left[\frac{\partial J(w, y)}{\partial w} \right]_{w=w(k)}$$

به دست آورد که در معادله' فوق $0 > c(w, y)$ تابع معیار است، و مشتق جزئی نیز در $w = w(k)$ ارزیابی می شود. نشان دهد که رابطه' الگوریتم پرسپترون را می توان از روای $\text{argmin}_{w \in \mathbb{R}^n} J(w, y) = w^T y - \frac{1}{2} \|w\|^2$ بدست آورد، که $J(w, y) = \frac{1}{2} \|w - y\|^2$ بقدر مطلق آرگومان است. (نتجه: مشتق جزئی $y^T w$ نسبت به w برابر y است.)

ثابت کنید که اگر مجموعه های الگوی آموزشی، جداپذیر خطی باشند، الگوریتم آموزش پرسپترون داده شده با معادلات $(9.3-3) - (9.3-4)$ در تعداد مراحل محدودی همگرا می شود. [راهنمایی: الگوهای دسته' w را در ۱- ضرب کنید و آستانه' نامنفی T را در نظر بگیرید، تا این که الگوریتم آموزش پرسپترون (با $c=1$) بدین صورت بیان شود: اگر $T > w^T(k)y(k)$ ، آنگاه $w(k+1) = w(k) + y(k)$ و در غیر این صورت $\|a\|^2\|b\|^2 \geq (a^T b)^2$. ممکن است نامعادله' کوشی-شوارتز $w(k+1) = w(k) + y(k)$ مورد نیاز تان باشد.]

۱۴.۹ ساختار وزن های یک شبکه' عصبی که قادر به انجام دقیقاً همان کار دسته بندی کننده' حداقل فاصله' برای دو دسته الگو در فضای \mathbb{R}^n بعدی باشد، را تعیین کنید.

۱۵.۹ ساختار وزن های یک شبکه عصبی که قادر به انجام دقیقاً همان کار دسته بندی کننده' بیز برای دو دسته الگو در فضای \mathbb{R}^n بعدی باشد، را تعیین کنید. دسته های توزیع گوسی با میانگین های متفاوت اما ماتریس های کوواریانس یکسان دارند.

۱۶.۹ (الف) در چه شرایطی شبکه های عصبی مسائل ۹.۱۴ و ۹.۱۵ یکسان هستند؟
 (ب) اگر شبکه های عصبی پیشخور چند لایه' تشریح شده در بخش ۹.۳ با قاعده' دلتای

تعیین یافته و با تعداد نمونه های کافی آموزش داده شوند، آیا این شبکه' عصبی خاص را به ما می دهند؟

۱۷.۹ دو دسته الگو در فضای دو بعدی طوری توزیع شده اند که الگوهای دسته ۱ به طور تصادفی در محیط دایره ای به شعاع r_1 قرار گرفته اند. به طور مشابه الگوهای دسته ۲ به طور تصادفی در محیط دایره ای به شعاع r_2 قرار گرفته اند، که $r_2 = 2r_1$. ساختار شبکه' عصبی با حداقل تعداد لایه ها و گره ها که برای دسته بندی صحیح الگوهای این دو دسته مورد نیاز است، را تعیین کنید.

۱۸.۹ مسئله ۶ را تکرار کرده اما از شبکه' عصبی استفاده کنید. به روشنی بیان کنید که چگونه نمونه های آموزشی را به دست می آورید و چگونه از این نمونه ها برای آموزش دسته بندی کننده استفاده می کنید. ساده ترین شبکه' عصبی را، که به نظر شما، می تواند این مسئله را حل کند، انتخاب کنید.

۱۹.۹ نشان دهید که عبارت $O_i - h'_i(I_i) = O_i(I_i)$ که در معادله' (۹.۳-۷۱) داده شده و در آن $h'_i(I_i) = \partial h_i(I_i) / \partial I_i$ ، از معادله' (۹.۳-۵۰) با $= 1$ نتیجه می شود.

۲۰.۹ نشان دهید که معیار فاصله' $D(A, B)$ در معادله' (۹.۴-۱)، خواص داده شده در معادله' (۹.۴-۲) را برآورده می کند.

۲۱.۹ نشان دهید که $M = \max(|A|, |B|)$ در معادله' (۹.۴-۳) صفر است اگر و تنها اگر A و B رشته های یکسانی باشند.

۲۲.۹ (الف) خودکاره' محدودی را تعیین کنید که قادر به تشخیص رشته های الگو به شکل a^{ab} باشد.
ب) برای جواب به قسمت (الف)، گرامر منظم متناظر را به دست آورید. (مسئله را با بازرسی^۱ حل نکنید).

۲۳.۹ یک گرامر درختی گسترشی برای تولید تصاویری که در هر دو جهت مکانی ۱' ها و ۰' های متناوب (الگوی شترنجی) دارند، ارائه دهید. فرض کنید که عنصر سمت چپ بالا ۱ بوده و تمام تصاویر در سمت چپ پایین با عنصر ۱ پایان می یابند.

۲۴.۹ روال یادگیری داده شده با معادلات (۹.۴-۹) - (۹.۴-۱۴) را برای یادگیری خودکاره'

محدودی که قادر به تشخیص رشته هایی به شکل $a^{n+1}ab$ ، با $n > 0$ ، است، به کار برید. با مجموعه 'نمونه' $\{aba, abba, abbba\}$ شروع کنید. اگر این مجموعه برای کشف تکرار منظم نماد a/b سیله' الگوریتم کافی نیست، رشته های نمونه' بیشتری را اضافه کنید تا این که بتوانند این کار را انجام دهد.

نیاز داشت که خودکاره' درختی داده شده در رابطه با شکل ۹.۳۰، درخت داده شده در شکل ۹.۳۱ (ب) را می پذیرد.

درستی تمام برابری های فهرست شده در زیربخش "چند تساوی مهم" از بخش ۹.۵.۴ را ثابت کنید.

دانش را که برای دسته بندی حباب های تصاویر مسئله ۹.۶ به چهار گروه پایه ای: (۱) حباب های بیضی کامل؛ (۲) حباب های بیضی ناکامل؛ (۳) حباب های ترکیبی کامل؛ و (۴) حباب های ترکیبی ناکامل، مورد نیاز است، با جملات شبه انگلیسی معین کنید. فرض کنید تصویرها بدرستی به تصاویر دودویی که در آنها هر پیکسل یا شیئی و یا زمینه است، بخش بندی شده اند. در (۲) و (۴)، "ناکامل" به معنی حباب هایی است که قسمتی از آنها خارج از ناحیه' تصویر است.

پاسخ خود به مسئله ۹.۲۷ را به عنوان مبنای برای یک سامانه مبتنی بر دانش که منطق گزاره ها را به کار می برد و می تواند تمام حباب ها را به چهار دسته' مذکور دسته بندی کند، استفاده کنید. تمام توابع پردازش تصویر مورد نیاز در سامانه منطقی مربوط را به روشنی معین کنید.

مسئله ۹.۲۸ را تکرار کنید، اما شبکه' معنایی را به کار برید. تمام توابع پردازش تصویر موردنیاز سامانه را دقیقاً معین کنید.

براساس پاسخستان به مسئله ۹.۲۷، قواعد سامانه خبره ای که قادر به دسته بندی تمام حباب ها به چهار دسته' مذکور است، را ارائه کنید. تمام توابع پردازش تصویر موردنیاز سامانه خبره تان در انجام این کار را دقیقاً معین کنید.

کارخانه معینی پرچم های کوچک آمریکا را برای بازی های ورزشی، به طور انتبه تولید می کند. گروه تضمین کیفیت مشاهده کردنده که در خلال دوره های با تولید حداکثر، بعضی

ماشین های چاپ بین یک تا سه ستاره و یک یا دو نوار کامل را (به طور تصادفی) چاپ نمی کنند. با صرف نظر از این خطاهای پرچم ها کامل هستند. گرچه پرچم هایی که خطای دارند، درصد کوچکی از مجموع تولید هستند، اما مدیر کارخانه تصمیم می گیرد که مسأله را حل کند. پس از بررسی زیاد، او به این نتیجه می رسد که خودکارسازی تصویری، انتصادی ترین راه برای حل مسأله است. مشخصات پایه ای به صورت زیر هستند: پرچم ها تقریباً به ابعاد 3×5 اینچ هستند، و به طور طولی با سرعت تقریباً 20 اینچ در ثانیه (جدا از هم و با ± 15 درجه نوسان در جهت) در خط تولید حرکت می کنند، که فاصله بین هر دو پرچم تقریباً دو اینچ است. (در تمام موارد، عبارت «تقریباً» به معنای $\pm 5\%$ است). مدیر کارخانه طراحی یک سامانه پردازش تصویر برای هر خط تولید را به شما و آگذار می کند. به شما گفته می شود که هزینه و سادگی، عوامل مهم در میزان تحقق پذیری راه حل شما هستند. سامانه کاملی براساس مدل شکل ۹.۱ طراحی کنید و راه حل خود (شامل فرض ها) را به صورت گزارش کتبی مختصر (اما واضح) به مدیر کارخانه ارائه کنید.

ضمیمه الف

تولید تصاویر نیم تن^۱

همان طور که در بخش ۱.۴ شرح داده شد، روش مطلوب برای نمایش تصویر رقمی استفاده از نمایشگری است که در آن شدت هر نقطه باشد پیکسل متناظر در تصویر ورودی متناسب باشد. وقتی، تنها وسیله خروجی در دسترس، وسایل دو سطحی نظیر نمایشگر دودویی، چاپگر خطی یا چاپگر لیزری [تکرنگ]^۲ باشد، روش های چاپ خاکستری که در اینجا شرح داده شدند، می توانند به عنوان جایگزین مورد استفاده قرار گیرند. وقتی این روش ها در مورد تصاویر رمز شده 'ضمیمه' ب بکار روند، راه مفیدی برای آزمایش روش های مشروح در متن کتاب و سپس نمایش نتایج روی چاپگرهای متداول موجود به شمار می روند.

روشی که نیم تن کردن^۳ خوانده می شود، روش متداولی است که برای چاپ عکس های خاکستری در روزنامه ها، مجلات و کتابها مورد استفاده قرار می گیرد. نظریه حاکم بر عمل نیم تن کردن بسیار ساده است. واحد تفکیک کوچکی (مثلاً یک مربع 0.02×0.02 اینچ) را که دارای جزئیات ریز سیاه روی زمینه سفید است، در نظر بگیرید. وقتی چشم از فاصله 'دید عادی' (مثلاً ۱۰ اینچ) به آن نگاه کند، از محتویات ناحیه انتگرال می گیرد طوری که ناظر تنها متوسط شدت ناحیه را احساس می کند. این خاصیت انتگرال گیری چشم، اساس نیم تن کردن است. برای تحقیق این روش باید در سطح واحد تفکیک، دایره ای جوهری چاپ کنیم که ابعادش باشد تصویر در آن واحد نسبت عکس داشته باشد. شکل الف. ۱ بزرگ شده یک الگوی نیم تن نوعی را نمایش می دهد. توجه کنید که در نواحی روشن تصویر، نقطه ها بسیار

¹⁾ Halfton¹⁾ Halftoning



شکل الف. تصویر نیم تن بزرگ شده

کوچکتر هستند و در نواحی تاریکتر به تناسب بزرگتر می شوند. نیم تن های روزنامه ای، تفکیکی در حدود ۸۰ نقطه در اینچ (DPI) دارند، در حالی که نیم تن های کتاب و مجله حدود دو برابر آن تفکیک را دارند.

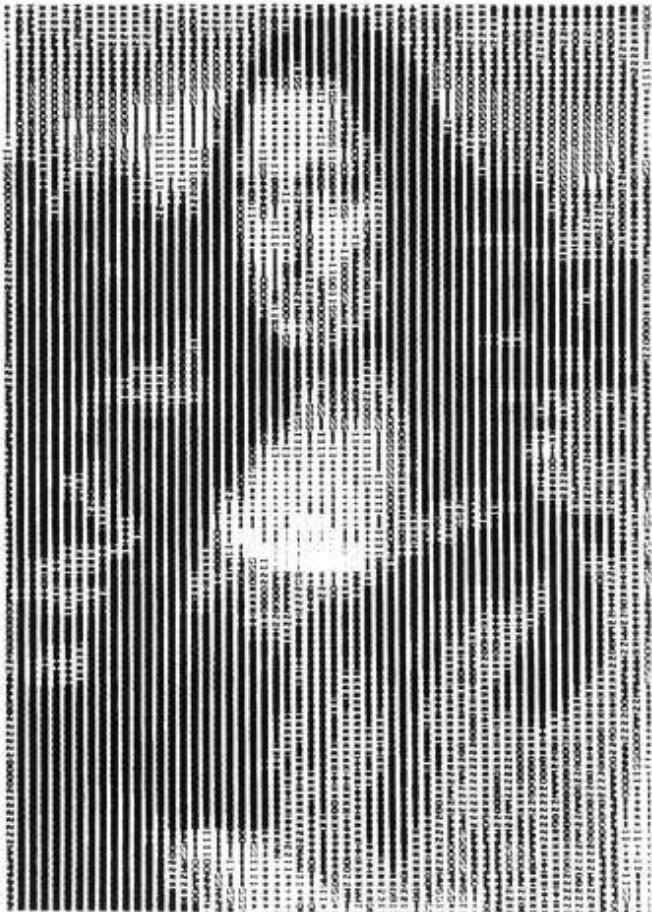
الگوی نیم تن را می توان به راه های مختلفی روی یک وسیله دو سطحی تقریب زد. یکی از ساده ترین روش ها این است که از چاپ روی هم نویسه ها استفاده کنیم. شکل الف. ۲ مجموعه ای از نویسه ها برای

```
MMMMMMMHIIIIIXHXOZWMNOS=I*++=:-.-  
WWWWWW# # *++---- = - -  
####00+-  
000  
+
```

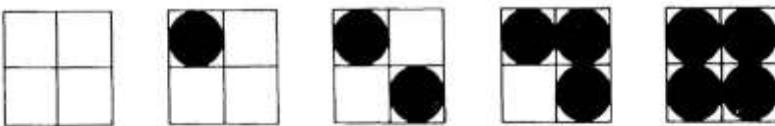
Gray levels

شکل الف. ۲ نویسه های روی هم چاپ شده که برای تولید ۳۲ سطح خاکستری متفاوت مورد استفاده قرار گرفتند. نویسه سی و دوم نویسه سفید (خالی) است.

رسیدن به این نتیجه را نشان می دهد. در این مورد خاص، ترکیبات مختلف پنج نویسه برای تقریب ۳۲ سطح خاکستری استفاده می شوند. واضح است که علاوه بر نویسه های نشان داده شده در شکل الف. ۲، از سایر ترکیبات نویسه های نیز می توان استفاده کرد، البته با این شرط که نویسه های مورد استفاده، انتقال نسبتاً آرامی از میاه به سفید تولید نمایند. مزیت اصلی چاپ روی هم نویسه ها فشردنگی آن است، بدین معنا که تنها از یک مکان نویسه برای تولید تعداد نسبتاً زیادی از سطوح خاکستری استفاده می شود. ایراد اصلی آن، کندی آن است. معمولاً وقتی وسیله 'خروجی، چاپگر خطی باشد، روش چاپ روی هم ترجیح داده می شود. شکل الف. ۳ یک تصویر خاکستری را که بوسیله 'چاپگر خطی با استفاده از روش چاپ روی هم ایجاد شده است، نشان می دهد.



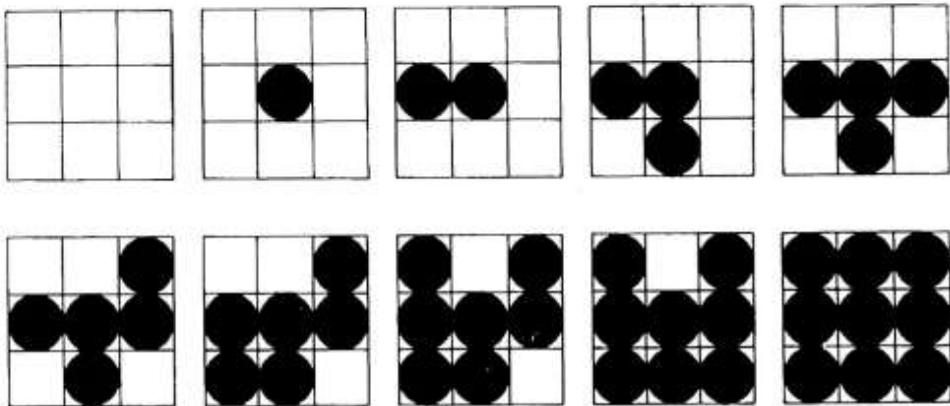
شکل الف. ۳۰ تصویری که با چاپ روی هم بدست آمد.



شکل الف. ۴. ناحیه تفکیک 2×2 و الگوهای نقطه ای مورد استفاده برای تولید پنج سطح خاکستری متفاوت

روش دیگر برای نیم تن کردن توسط یک وسیله دو سطحی، تشکیل الگوهای مختلف نقطه های سیاه درون یک ناحیه' مربعی سفید است، تابدین وسیله نظریه حاکم بر روش نیم تن برای چاپ رسانه ها که قبل از شرح داده شد، را تقریب بزنیم. همان طور که در شکل الف. ۴ دیده می شود، می توان از ناحیه' 2×2 پیکسلی یک وسیله دو سطحی برای تولید پنج سطح خاکستری استفاده کرد. مزیت این روش نسبت به چاپ روی هم، سرعت بیشتر آن است. ایراد اصلی آن نیز این است که تفکیک تصویر در راستای هر محور نصف می شود، زیرا در وسیله' چاپ، برای تولید سطح خاکستری هر پیکسل ورودی، یک ناحیه مربعی چهار پیکسلی مورد نیاز است.

از یک ناحیه 3×3 می توان ۱۰ سطح خاکستری را با استفاده از، مثلاً، الگوهای نقطه ای شکل الف. ۵ تولید کرد. به روشنی جریمه پرداخته شده برای افزایش تعداد سطوح خاکستری، کاهش دو- سوم تفکیک تصویر در طول هر محور نمایش است. بنابراین هر گروه 11×11 از پیکسل های دو سطحی می تواند $1 + 2^2$ سطح خاکستری تولید کند. طراحی الگوهای نقطه ای برای استفاده در این روش شامل دو ملاحظه مهم



شکل الف. ۵. ناحیه تفکیک 3×3 و الگوهای نقطه ای آن که برای تولید ۱۰ سطح خاکستری متفاوت مورد استفاده قرار گرفتند.

است: اولاً نقطه های الگو باید طوری چیده شوند که در تصویر چاپ شده از الگوهای مشخص و آشنا احتراز شود. مثلاً اگر سه نقطه متناظر با سطح خاکستری ۳ در شکل الف. ۵ به صورت یک خط افقی در مرکز ناحیه 3×3 چیده شوند، یک ناحیه بزرگ با سطح خاکستری ثابت ۳ به صورت دو خط سفید افقی، یک خط سیاه بعد آن، دو خط سفید افقی دیگر بعد آن و به همین ترتیب الى آخر دیده می شود. الگوهای نشان داده شده در شکل الف. ۵ چنین اثراتی را کاهش می دهدن. دوم این که اگر پیکسلی در الگو برای سطح خاکستری آ سیاه باشد، برای تمام سطوح \angle زیز سیاه خواهد بود. این قید، منحنی سازی اشتباه^۱ در تصویر خروجی را کاهش می دهد (بخش ۲ . ۳ را بینید).

شکل الف. ۶ (ب) تصویری را که با الگوهای نقطه ای شکل الف. ۵ چاپ شده است، را نشان می دهد. تصویر اولیه نشان داده شده در شکل الف. ۶ (الف) با استفاده از این الگوهای نقطه ای، رمز گردید و آنگاه با چاپگر لیزری معمولی چاپ شد. تصویر اولیه، تفکیک ۵۱۲ \times ۵۱۲ پیکسلی با 256 سطح خاکستری داشت. پس از کاربردن ناحیه 3×3 خاکستری که برای ایجاد ۱۰ سطح خاکستری مورد نیاز بودند، تفکیک تصویر رمز شده به 170×170 کاهش یافت. گرچه در شکل الف. ۶ (ب) پدیده منحنی سازی اشتباه به میزانی دیده می شود، این تقریب برای کارهای پردازش تصویر در نبود وسیله خروجی خاکستری با کیفیت بالا قابل قبول است.



(الف)



(ب)

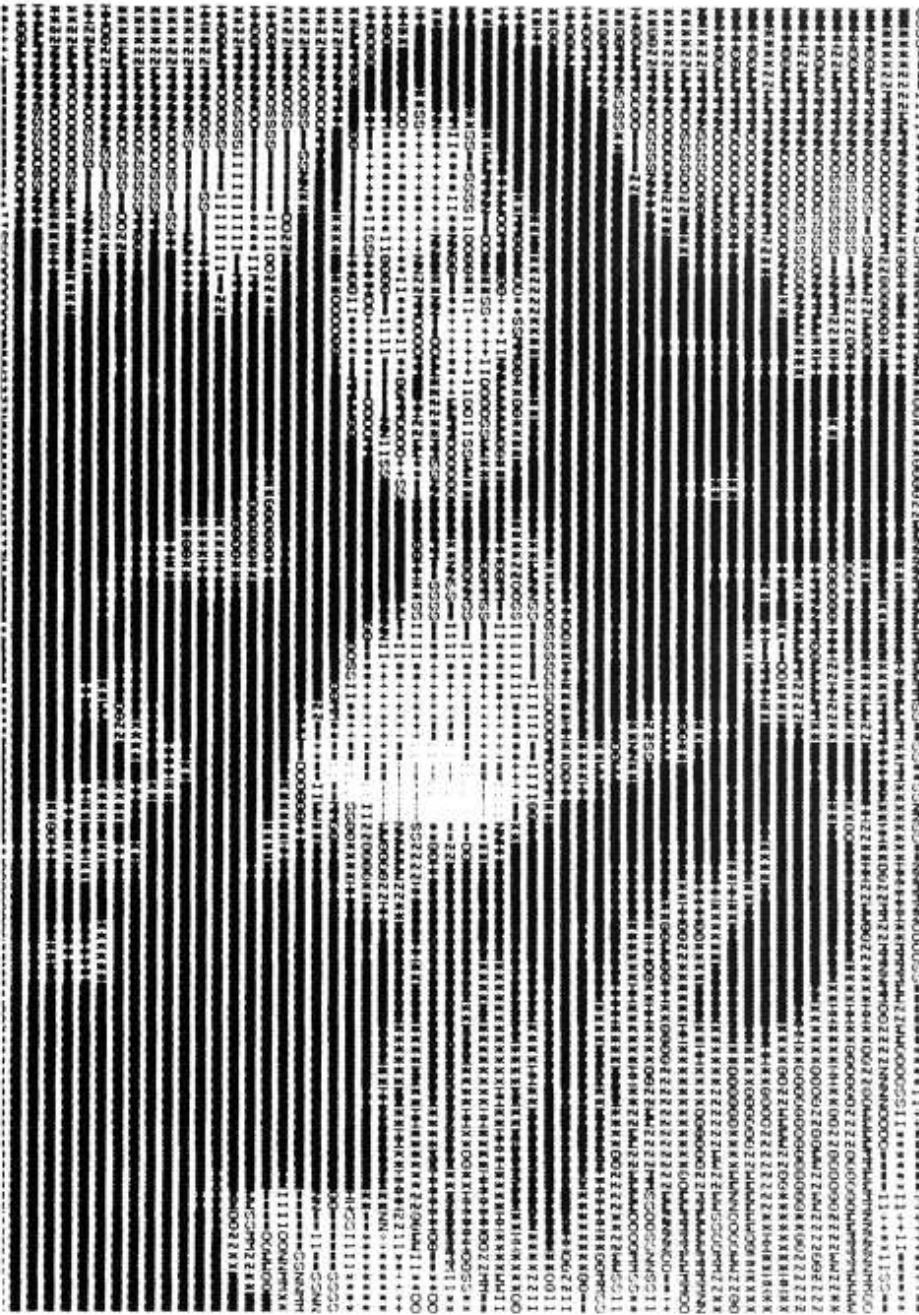
شکل الف. ۶ (الف) یک تصویر رقمنی؛ (ب) نتیجه به دست آمده با الگوهای نیم تن شکل الف. ۵

1) False contouring

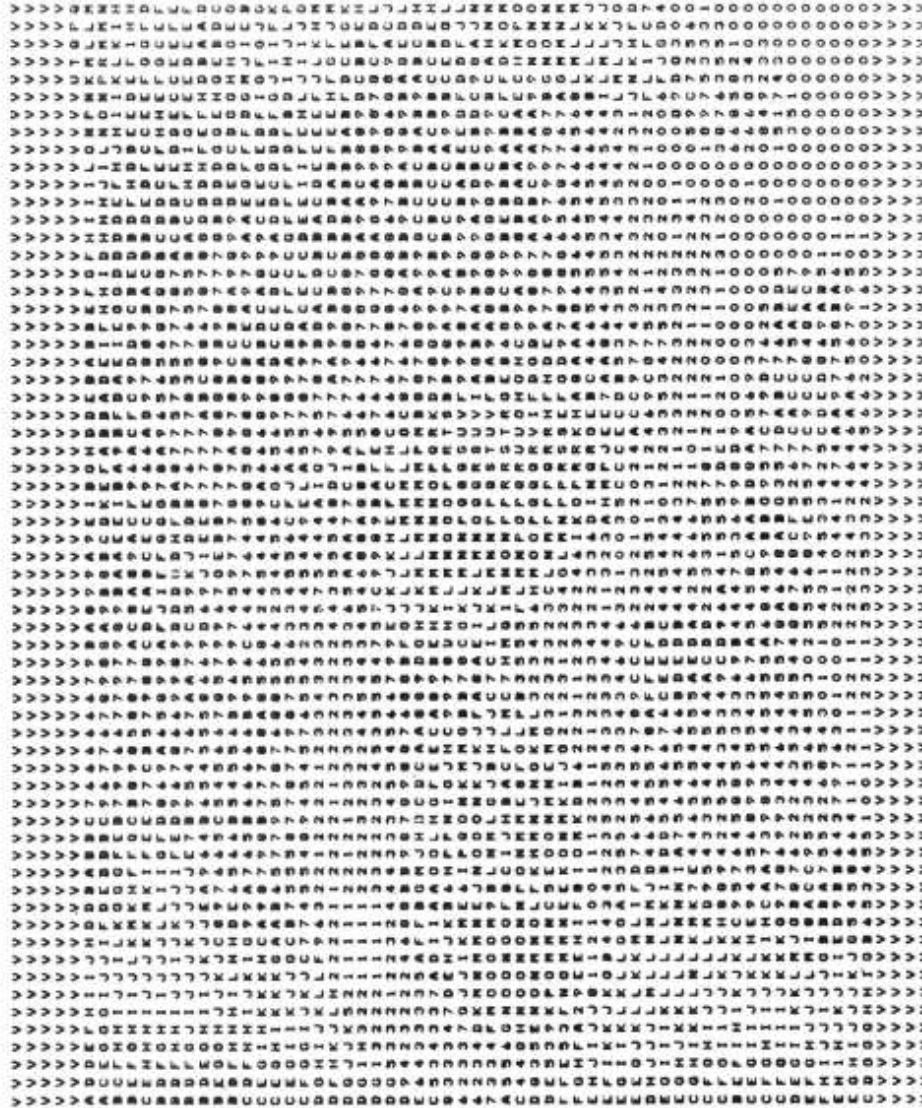
ضمیمه ب

تصاویر رمز شده

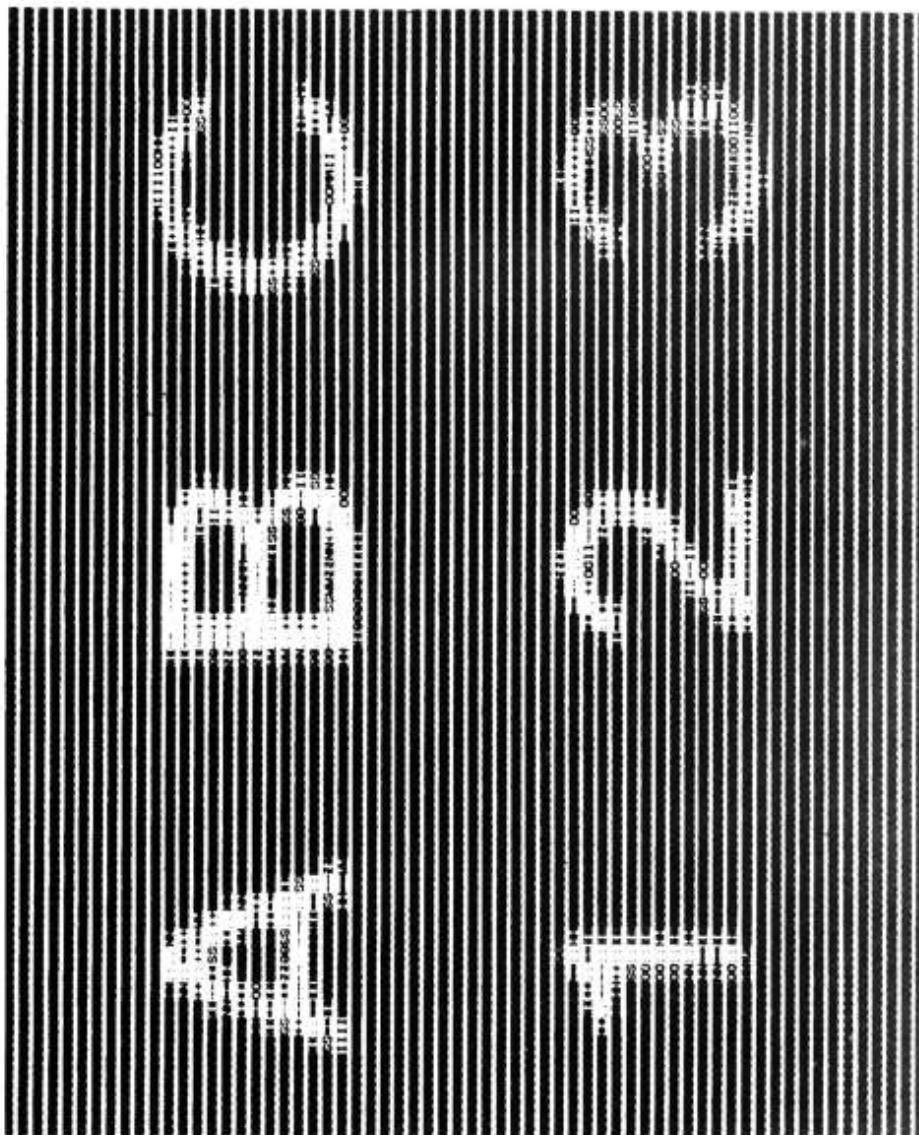
تصاویر ۳۲ سطحی 64×64 پیکسلی را می توان به عنوان داده های آزمون برای بسیاری از مفاهیم پردازش تصویر مطرح شده در متن به کار برد. همراه هر تصویر آرایه 'رمز شده ای نشان داده می شود که حاوی نویسه های الفبا- عددی هر پیکسل تصویر است. محدوده 'این نویسه ها ۰-۹ و A-V است که متناظر با ۳۲ سطح خاکستری می باشد. اولین قدم پس از ورود تصویر رمز شده به رایانه این است که نویسه های الفبا- عددی را به سطوح عددی در محدوده '۰ تا ۳۱ تبدیل کنیم. از آرایه 'عددی حاصل یا می توان در شکل اولیه اش استفاده کرد، یا این که، به عنوان مثال، با افزودن نویز به هر پیکسل، آن را تخریب کرد. این انعطاف پذیری امکان می دهد که داده های ورودی گوناگونی تولید کنیم که می توان از آنها برای نمایش اثرات الگوریتم های پردازش تصویر استفاده کرد. تابع قبل و بعد از پردازش را می توان روی یک وسیله 'دو سطحی استاندارد، نظیر چاپگر خطی ، با استفاده از برنامه ای که مبتنی بر مفاهیم معرفی شده در ضمیمه 'الف است، نمایش داد.

شکل ب. *Mona Lisa*

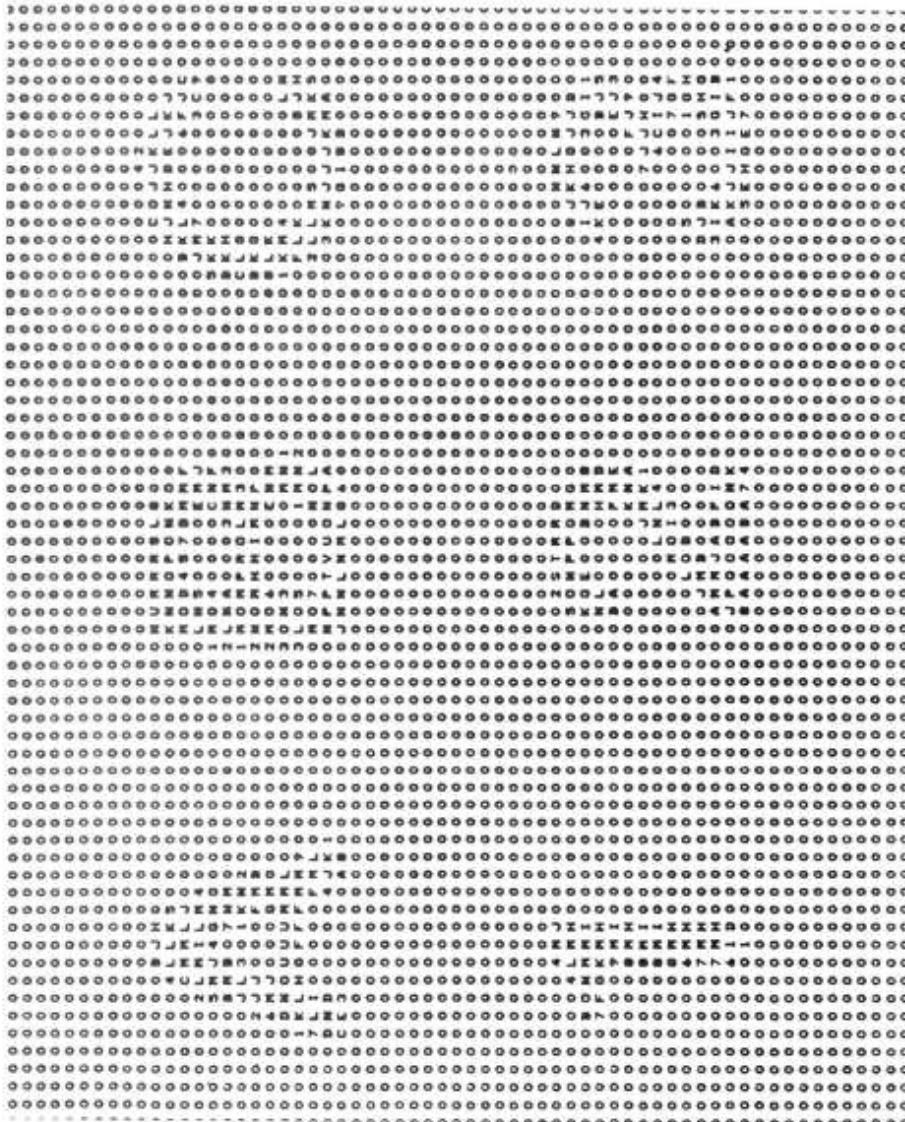
تصاویر رمز شده ۷۹۷



شکل ب. ۲. تصویر رمز شده 'Mona Lisa'



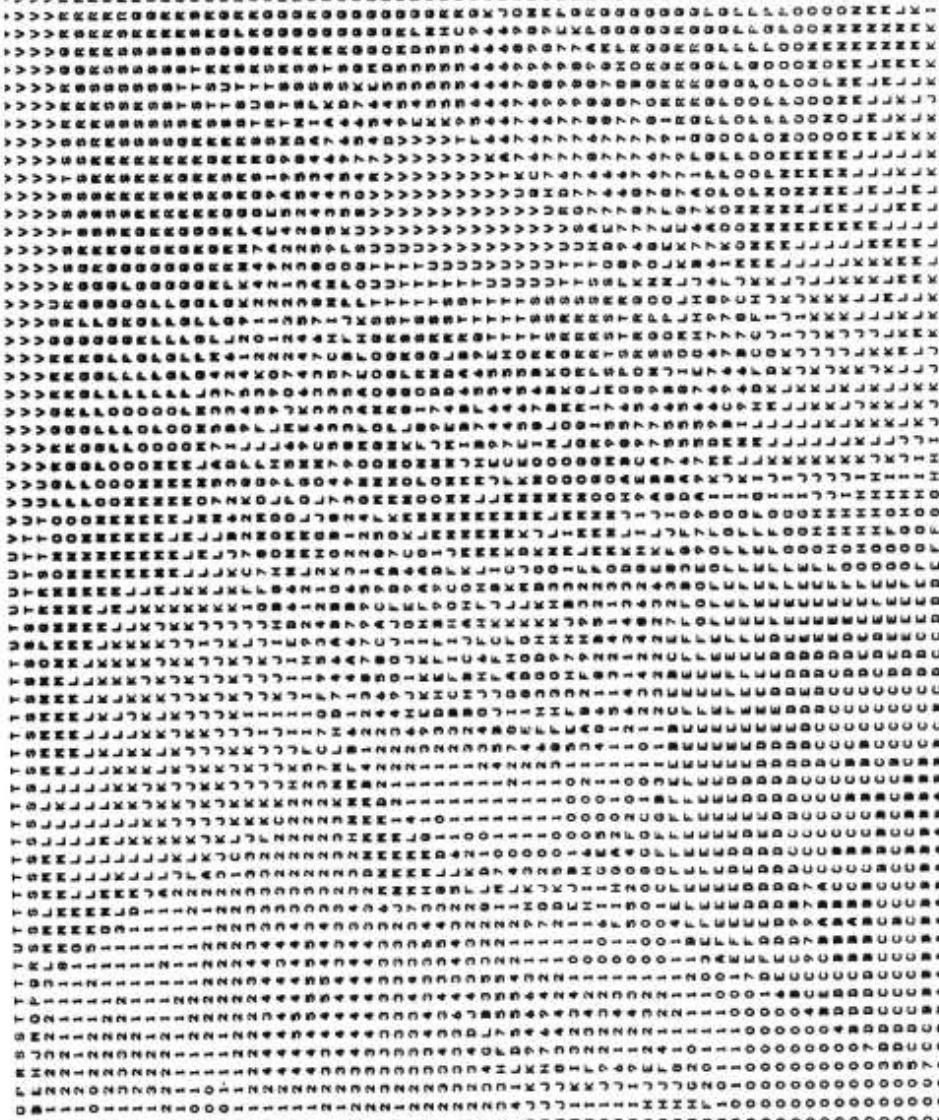
شکل ب ۳۰ نویسه ها

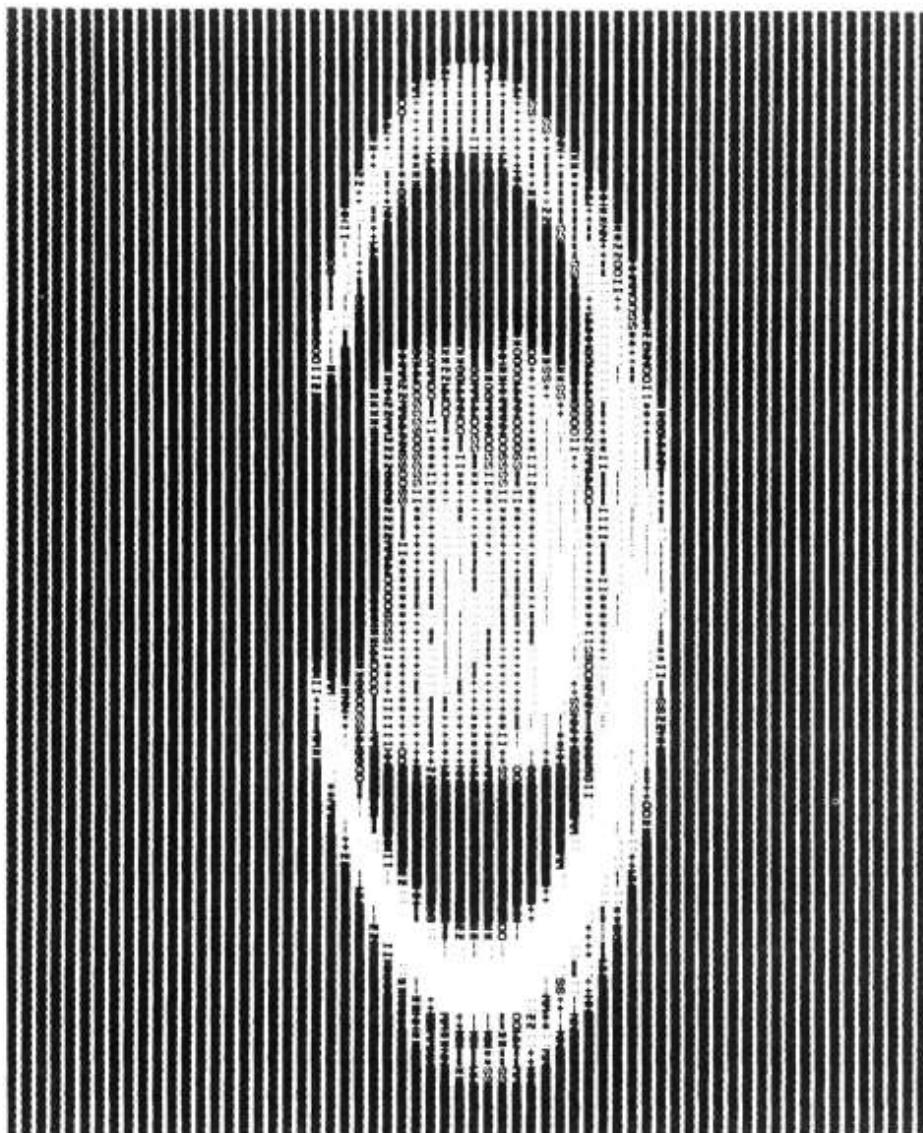


شکل ب. ۴ نویسه های رمز شده



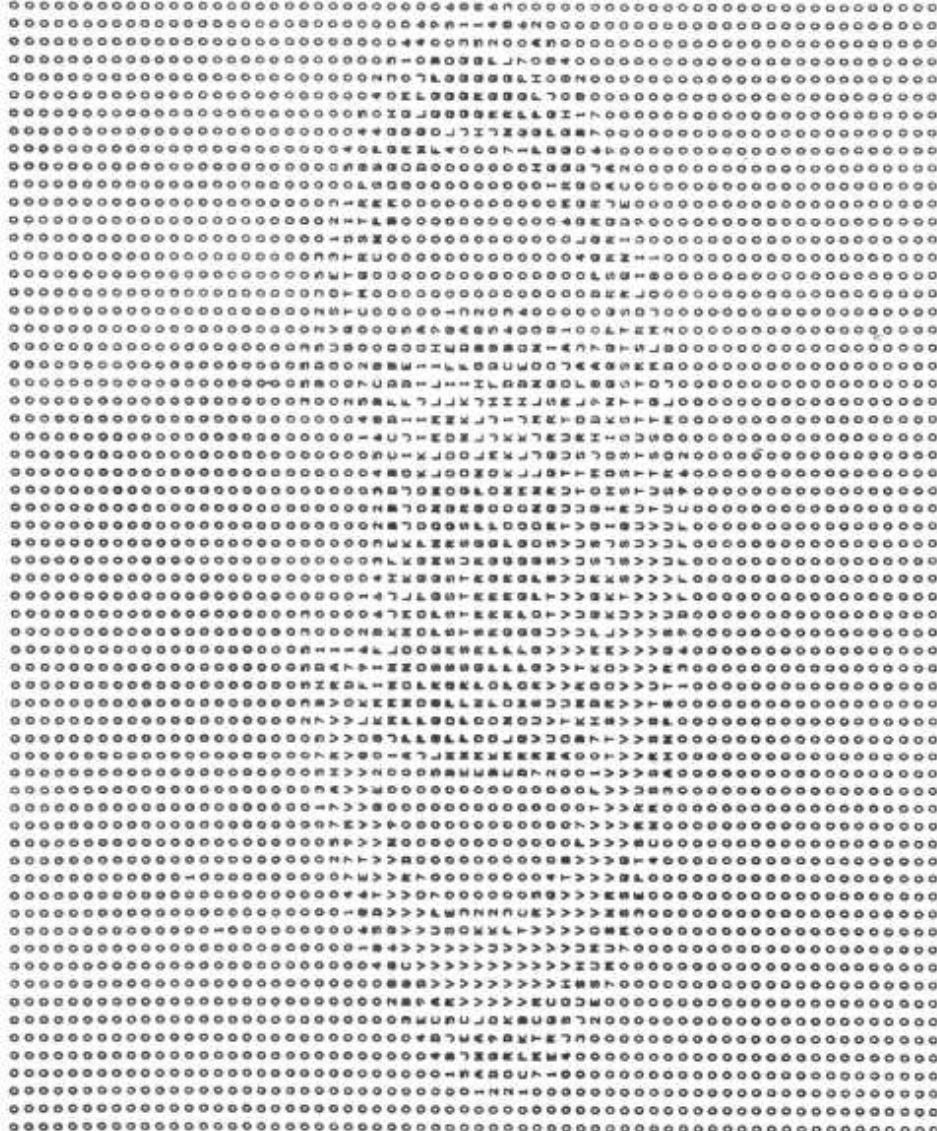
شکل ب . ۵ عکس *Lincoln*

شکل ب . عکس رمز شده *Lincoln*

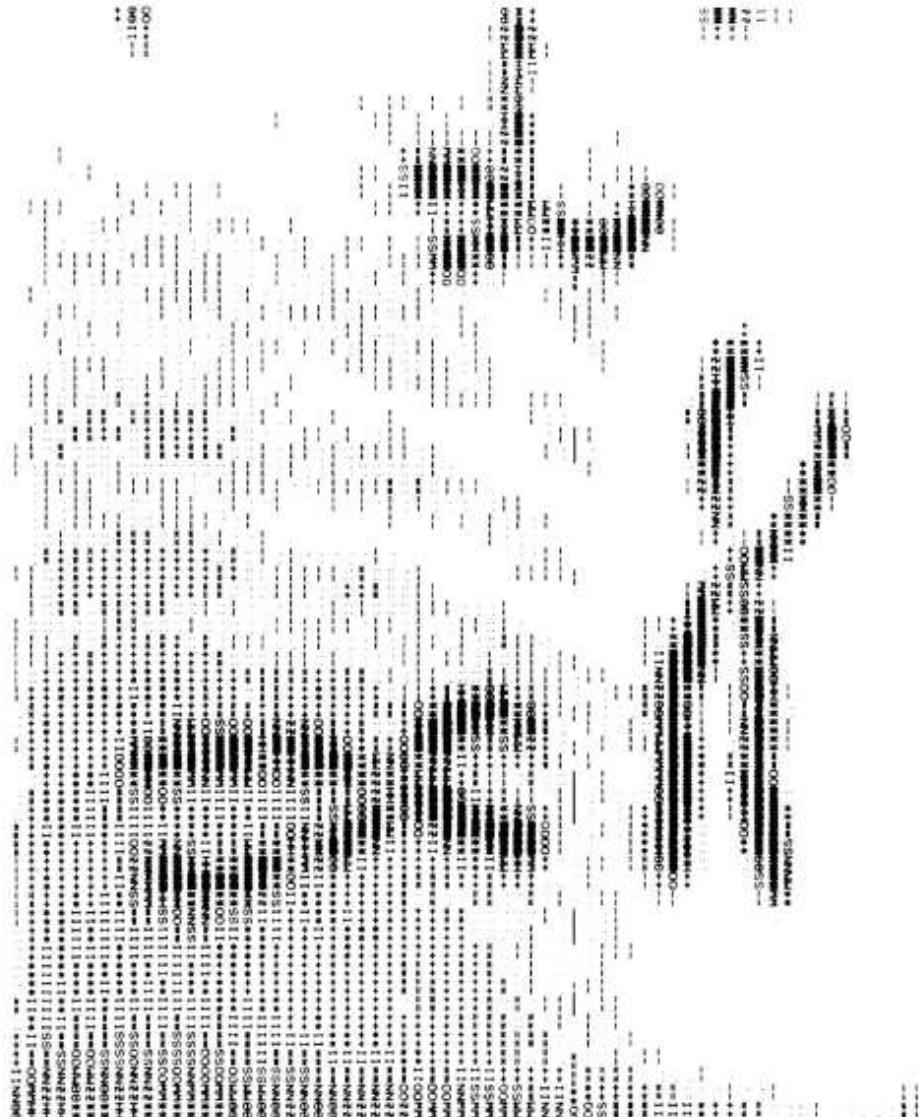


شکل ب . ۷ حل

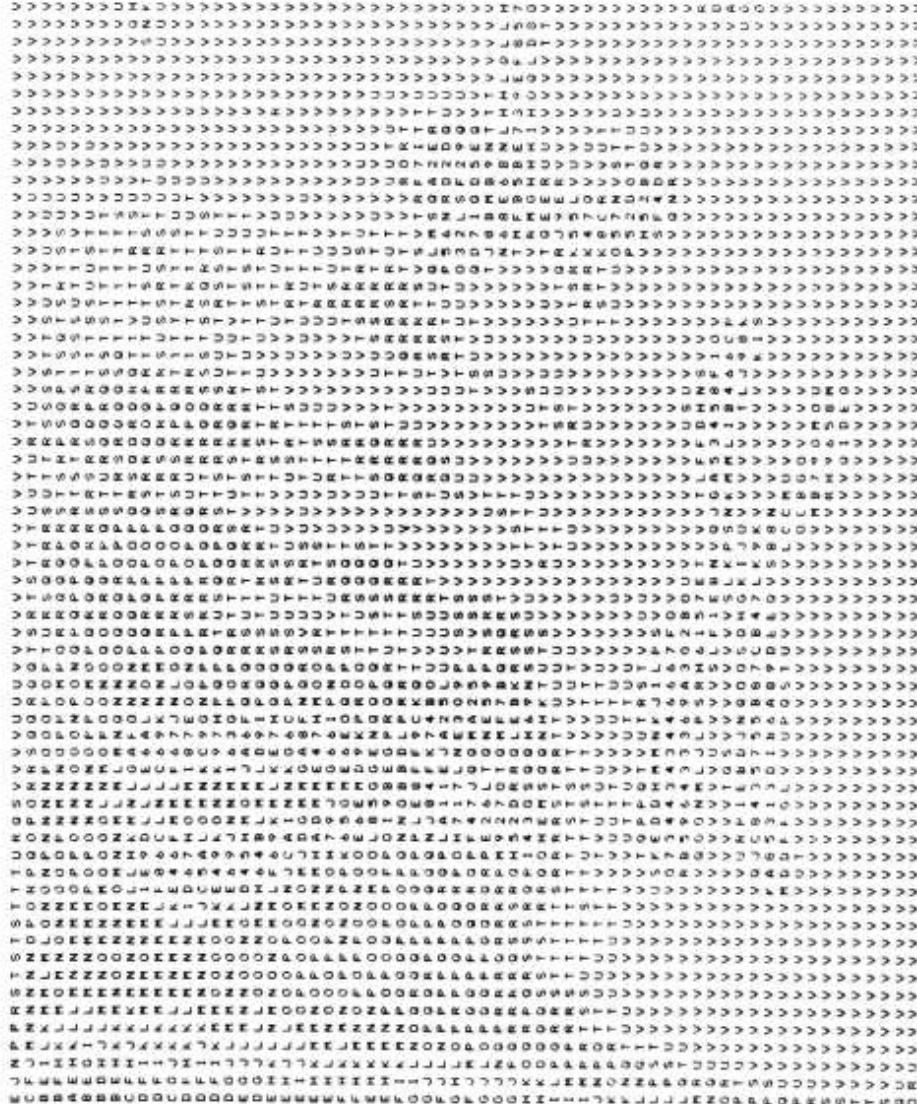
A + ۳ تصویر رمز شده



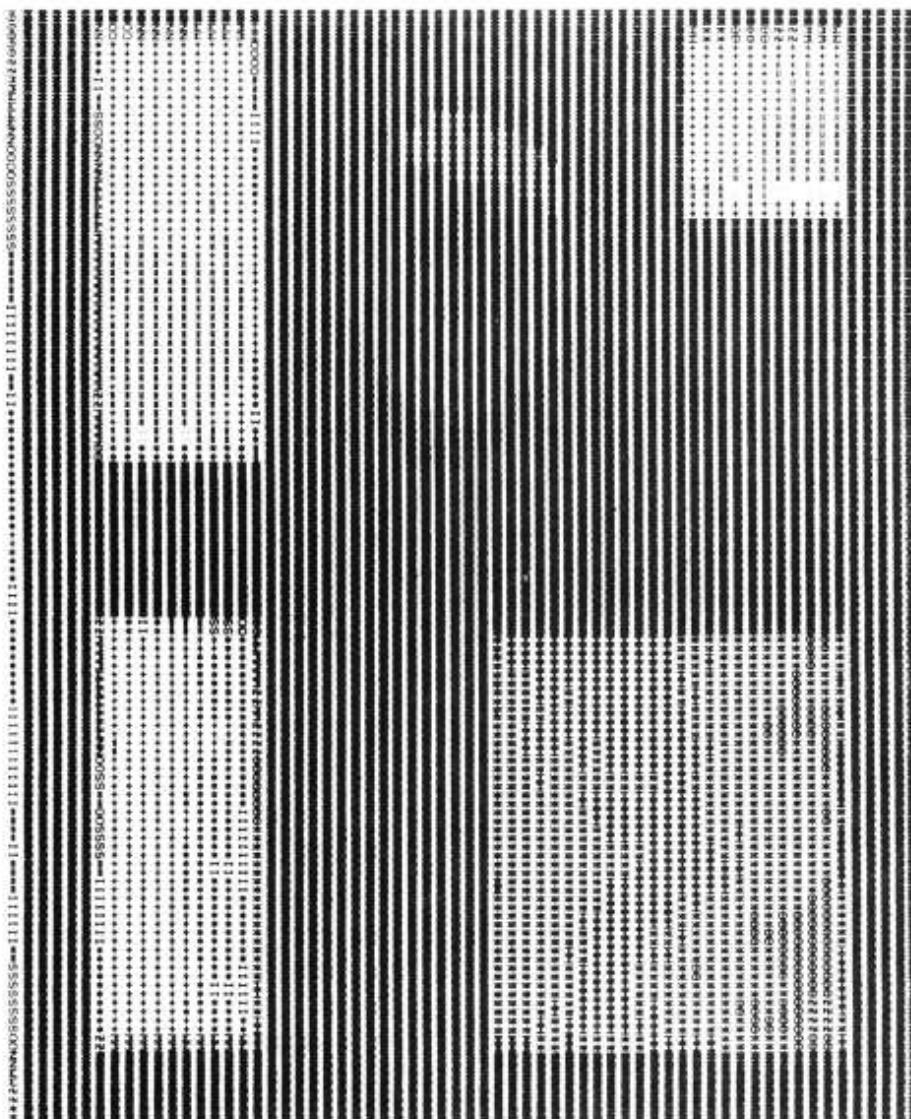
شکل ب A + ۳ تصویر رمز شده زحل



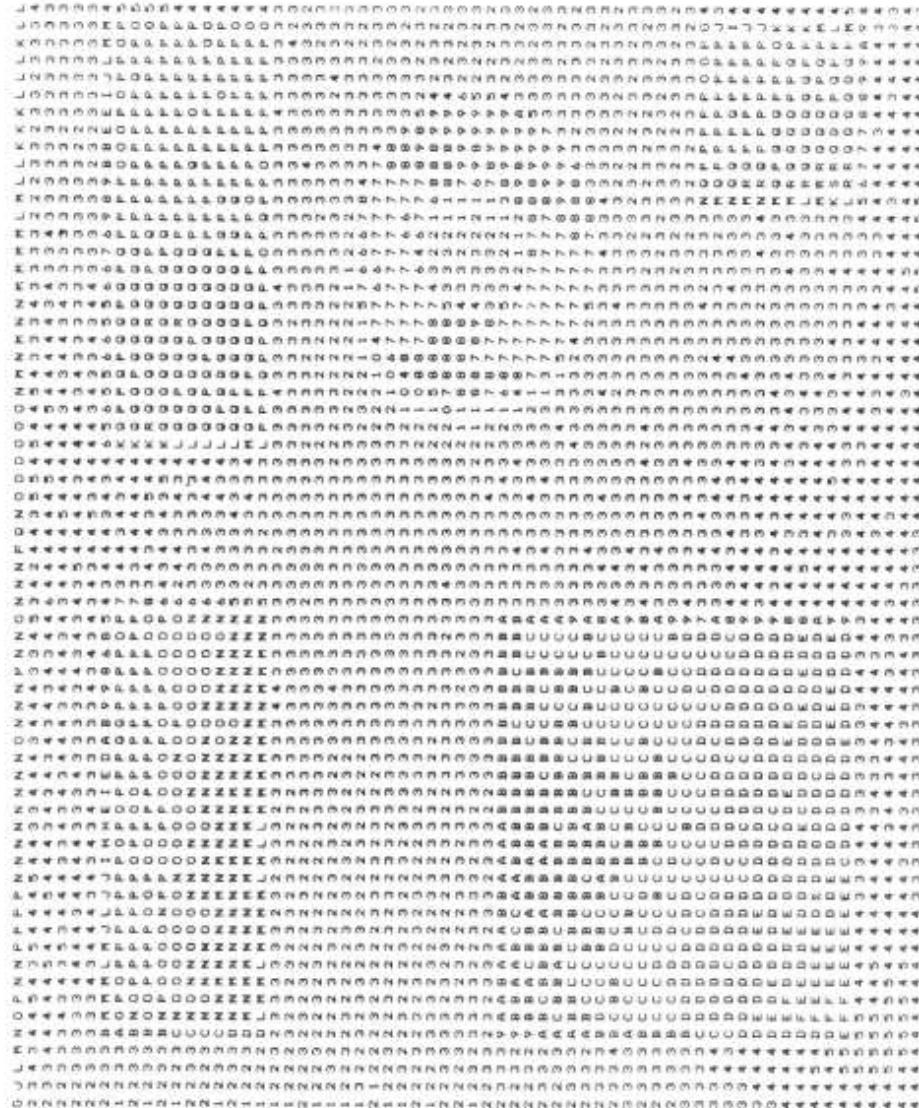
شکل ب . ۹ کروموزوم ها



شکل ب ۱۰۰ تصویر رمز شده کروموزوم ها



شکل ب . ۱۱ . اشکال هندسی



شکل ب ۱۲۰ تصویر رمز شده اشکال هندسی

کتاب‌شناسی

- Abidi, M.A., Eason, R.O., and Gonzalez, R.C. [1991]. "Autonomous Robotics Inspection and Manipulation Using Multisensor Feedback." *IEEE Computer*, vol. 24, no. 4, pp. 17–31.
- Abramson, N. [1963]. *Information Theory and Coding*, McGraw-Hill, New York.
- Adimari, M., Masciangelo, S., Borghesi, L., and Vernazza, G. [1988]. "A Knowledge-Based Approach to Industrial Scene Analysis: Shadows and Reflexes Detection." In *Image Analysis and Processing II*, Cantoni, V. et al., eds., Plenum Press, New York, pp. 101–110.
- Adiv, G. [1985]. "Determining Three-Dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-7, no. 4, pp. 384–401.
- Aggarwal, J.K., and Badler, N.I., eds. [1980]. "Motion and Time-Varying Imagery." *IEEE Trans. Pattern Anal. Mach. Intell.*, Special Issue, vol. PAMI-2, no. 6, pp. 493–588.
- Ahmed, N., Natarajan, T., and Rao, K.R. [1974]. "Discrete Cosine Transforms." *IEEE Trans. Comp.*, vol. C-23, pp. 90–93.
- Ahmed, N., and Rao, K.R. [1975]. *Orthogonal Transforms for Digital Signal Processing*, Springer-Verlag, New York.
- Ambler, A.P., et al. [1975]. "A Versatile System for Computer Controlled Assembly." *Artificial Intell.*, vol. 6, no. 2, pp. 129–156.
- Anderson, G.L., and Netravali, A.N. [1976]. "Image Restoration Based on a Subjective Criterion." *IEEE Trans. Syst. Man. Cyb.*, vol. SMC-6, no. 12, pp. 845–853.
- Andrews, H.C. [1970]. *Computer Techniques in Image Processing*, Academic Press, New York.
- Andrews, H.C. [1974]. "Digital Image Restoration: A Survey." *Computer J.*, vol. 7, no. 5, pp. 36–45.
- Andrews, H.C., and Hunt, B.R. [1977]. *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, N.J.
- Andrews, H.C., Tescher, A.G., and Kruger, R.P. [1972]. "Image Processing by Digital Computer." *IEEE Spectrum*, vol. 9, no. 7, pp. 20–32.
- Ang, P.H., Ruetz, P.A., and Auld, D. [1991]. "Video Compression Makes Big Gains." *IEEE Spectrum*, vol. 28, no. 10, pp. 16–19.
- Anuta, P.F. [1969]. "Digital Registration of Multispectral Video Imagery." *Soc. Photo-Optical Instrum. Engs.*, vol. 7, pp. 168–175.
- Bajcsy, R. [1973]. "Computer Description of Textured Surfaces." *Proc. 1973 Int. Conf. Artificial Intell.*, Stanford, Calif., pp. 572–579.
- Bajcsy, R., and Lieberman, L. [1976]. "Texture Gradient as a Depth Cue." *Comput. Graph. Image Proc.*, vol. 5, no. 1, pp. 52–67.
- Ballard, D.H. [1981]. "Generalizing the Hough Transform to Detect Arbitrary Shapes." *Pattern Recognition*, vol. 13, no. 2, pp. 111–122.
- Ballard, D.H., and Brown, C.M. [1982]. *Computer Vision*, Prentice-Hall, Englewood Cliffs, N.J.

- Bates, R.H.T. and McDonnell, M.J. [1986]. *Image Restoration and Reconstruction*, Oxford University Press, New York.
- Baumert, L.D., Golomb, S.W., and Hall, M., Jr. [1962]. "Discovery of a Hadamard Matrix of Order 92." *Bull. Am. Math. Soc.*, vol. 68, pp. 237-238.
- Baylon, D.M., and Lim, J.S. [1990]. "Transform/Subband Analysis and Synthesis of Signals." *Tech. Report*, MIT Research Laboratory of Electronics, Cambridge, Mass.
- Bell, E.T. [1965]. *Men of Mathematics*, Simon and Schuster, New York.
- Bellman, R. [1970]. *Introduction to Matrix Analysis*, 2nd ed., McGraw-Hill, New York.
- Bengtsson, A., and Eklundh, J.O. [1991]. "Shape Representation by Multiscale Contour Approximation." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 1, pp. 85-93.
- Berger, T. [1971]. *Rate Distortion Theory*, Prentice-Hall, Englewood Cliffs, N.J.
- Bernstein, R. [1976]. "Digital Image Processing of Earth Observation Sensor Data." *IBM J. Res. Dev.*, vol. 20, no. 1, pp. 40-56.
- Biberman, L.M. [1973]. "Image Quality." In *Perception of Displayed Information*, Biberman, L.M. ed., Plenum Press, New York.
- Biermann, A.W., and Feldman, J.A. [1972]. "On the Synthesis of Finite-State Machines from Samples of Their Behavior." *IEEE Trans. Comput.*, vol. C-21, no. 6, pp. 592-597.
- Billingsley, F.C., Goetz, A.F.H., and Lindsley, J.N. [1970]. "Color Differentiation by Computer Image Processing." *Photo. Sci. Eng.*, vol. 14, no. 1, pp. 28-35.
- Bisignani, W.T., Richards, G.P., and Whelan, J.W. [1966]. "The Improved Grey Scale and Coarse-Fine PCM Systems: Two New Digital TV Bandwidth Reduction Techniques." *Proc. IEEE*, vol. 54, no. 3, pp. 376-390.
- Blackman, E.S. [1968]. "Effects of Noise on the Determination of Photographic System Modulation Transfer Function." *Photogr. Sci. Eng.*, vol. 12, pp. 244-250.
- Blackman, R.B., and Tukey, J.W. [1958]. *The Measurement of Power Spectra*, Dover Publications, New York.
- Blahut, R.E. [1987]. *Principles and Practice of Information Theory*, Addison-Wesley, Reading, Mass.
- Blum, H. [1967]. "A Transformation for Extracting New Descriptors of Shape." in *Models for the Perception of Speech and Visual Form*, Wathen-Dunn, W., ed., MIT Press, Cambridge, Mass.
- Blume, H., and Fand, A. [1989]. "Reversible and Irreversible Image Data Compression Using the S-Transform and Lempel-Ziv Coding." *Proc. SPIE Medical Imaging III: Image Capture and Display*, vol. 1091, pp. 2-18.
- Bouman, C., and Liu, B. [1991]. "Multiple Resolution Segmentation of Textured Images." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 2, pp. 99-113.
- Brayer, J.M., Swain, P.H., and Fu, K.S. [1977]. "Modeling of Earth Resources Satellite Data." *Syntactic Pattern Recognition Applications*, (K.S. Fu, ed.), Springer-Verlag, New York.
- Bribiesca, E. [1981]. "Arithmetic Operations Among Shapes Using Shape Numbers." *Pattern Recog.*, vol. 13, no. 2, pp. 123-138.
- Bribiesca, E., and Guzman, A. [1980]. "How to Describe Pure Form and How to Measure Differences in Shape Using Shape Numbers." *Pattern Recog.*, vol. 12, no. 2, pp. 101-112.
- Brice, C.R., and Fennema, C.L. [1970]. "Scene Analysis Using Regions." *Artificial Intelligence*, vol. 1, pp. 205-226.
- Brigham, E.O. [1974]. *The Fast Fourier Transform*, Prentice-Hall, Englewood Cliffs, N.J.

۴۱۱ کتاب شناسی

- Brill, E.L. [1968]. "Character Recognition Via Fourier Descriptors." WESCON, Paper 25/3, Los Angeles, Calif.
- Brown, J.L., Jr. [1960]. "Mean-Square Truncation Error in Series Expansions of Random Functions." *J. SIAM*, vol. 8, pp. 18–32.
- Brummer, M.E. [1991]. "Hough Transform Detection of the Longitudinal Fissure in Tomographic Head Images." *IEEE Trans. Biomed. Images*, vol. 10, no. 1, pp. 74–83.
- Brzakovic, D., Patton, R., and Wang, R. [1991]. "Rule-based Multi-template Edge Detection." *Comput. Vision, Graphics, Image Proc: Graphical Models and Image Proc.*, vol. 53, no. 3, pp. 258–268.
- Budak, A. [1974]. *Passive and Active Network Analysis and Synthesis*, Houghton Mifflin, Boston.
- Campbell, J.D. [1969]. "Edge Structure and the Representation of Pictures." Ph.D. dissertation, Dept. of Elec. Eng., University of Missouri, Columbia.
- Candy, J.C., Franke, M.A., Haskell, B.G., and Mounts, F.W. [1971]. "Transmitting Television as Clusters of Frame-to-Frame Differences." *Bell Sys. Tech. J.*, vol. 50, pp. 1889–1919.
- Cannon, T.M. [1974]. "Digital Image Deblurring by Non-Linear Homomorphic Filtering." Ph.D. Thesis, University of Utah.
- Carlson, A.B. [1968]. *Communication Systems*, McGraw-Hill, New York.
- Castleman, K.R. [1979]. *Digital Image Processing*, Prentice-Hall, Englewood Cliffs, N.J.
- Chang, S.K. [1989]. *Principles of Pictorial Information Systems Design*, Prentice-Hall, Englewood Cliffs, N.J.
- Chaudhuri, B.B. [1983]. "A Note on Fast Algorithms for Spatial Domain Techniques in Image Processing." *IEEE Trans. Syst. Man Cyb.*, vol. SMC-13, no. 6, pp. 1166–1169.
- Chen, D., and Wang, L. [1991]. "Texture Features Based on Texture Spectrum." *Pattern Recog.*, vol. 24, no. 5, pp. 391–399.
- Chen, P.H., and Wintz, P.A. [1976]. "Data Compression for Satellite Images." TR-EE-76-9, School of Electrical Engineering, Purdue University, West Lafayette, Ind.
- Chow, C.K. and Kaneko, T. [1972]. "Automatic Boundary Detection of the Left Ventricle from Cineangiograms." *Comp. and Biomed. Res.*, vol. 5, pp. 388–410.
- Clark, J.J. [1989]. "Authenticating Edges Produced by Zero-Crossing Algorithms." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 8, pp. 830–831.
- Clark, R.J. [1985]. *Transform Coding of Images*, Academic Press, New York.
- Cochran, W.T., Cooley, J.W., et al. [1967]. "What Is the Fast Fourier Transform?" *IEEE Trans. Audio and Electroacoustics*, vol. AU-15, no. 2, pp. 45–55.
- Computer [1974]. Special issue on digital image processing, vol. 7, no. 5.
- Computer [1988]. Special issue on artificial neural systems, vol. 21, no. 3.
- Cooley, J.W., Lewis, P.A.W., and Welch, P.D. [1967a]. "Historical Notes on the Fast Fourier Transform." *IEEE Trans. Audio and Electroacoustics*, vol. AU-15, no. 2, pp. 76–79.
- Cooley, J.W., Lewis, P.A.W., and Welch, P.D. [1967b]. "Application of the Fast Fourier Transform to Computation of Fourier Integrals." *IEEE Trans. Audio and Electroacoustics*, vol. AU-15, no. 2, pp. 79–84.
- Cooley, J.W., Lewis, P.A.W., and Welch, P.D. [1969]. "The Fast Fourier Transform and its Applications." *IEEE Trans. Educ.*, vol. E-12, no. 1, pp. 27–34.
- Cooley, J.W., and Tukey, J.W. [1965]. "An Algorithm for the Machine Calculation of Complex Fourier Series." *Math. of Comput.*, vol. 19, pp. 297–301.
- Cornsweet, T.N. [1970]. *Visual Perception*, Academic Press, New York.

- Cowart, A.E., Snyder, W.E., and Ruedger, W.H. [1983]. "The Detection of Unresolved Targets Using the Hough Transform." *Comput. Vision Graph Image Proc.*, vol. 21, pp. 222-238.
- Crimmins, T.R., and Brown, W.R. [1985]. "Image Algebra and Automatic Shape Recognition." *IEEE Trans. Aerospace and Electron. Syst.*, vol. AES-21, no. 1, pp. 60-69.
- Cross, G.R., and Jain, A.K. [1983]. "Markov Random Field Texture Models." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-5, no. 1, pp. 25-39.
- Cumani, A., Guiducci, A., and Grattoni, P. [1991]. "Image Description of Dynamic Scenes." *Pattern Recog.*, vol. 24, no. 7, pp. 661-674.
- Cutrona, L.J., and Hall, W.D. [1968]. "Some Considerations in Post-Facto Blur Removal." In *Evaluation of Motion-Degraded Images*, NASA Publ. SP-193, pp. 139-148.
- Cutrona, L.J., Leith, E.N., and Palermo, C.J. [1960]. "Optical Data Processing and Filtering Systems." *IRE Trans. Info. Theory*, vol. IT-6, no. 3, pp. 386-400.
- Danielson, G.C., and Lanczos, C. [1942]. "Some Improvements in Practical Fourier Analysis and Their Application to X-Ray Scattering from Liquids." *J. Franklin Institute*, vol. 233, pp. 365-380, 435-452.
- Davenport, W.B., and Root, W.L. [1958]. *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York.
- Davis, L.S. [1975]. "A Survey of Edge Detection Techniques." *Comput. Graphics Image Proc.*, vol. 4, pp. 248-270.
- Davis, L.S. [1982]. "Hierarchical Generalized Hough Transforms and Line-Segment Based Generalized Hough Transforms." *Pattern Recog.*, vol. 15, no. 4, pp. 277-285.
- Davis, P.J. [1979]. *Circulant Matrices*, John Wiley & Sons, New York.
- Davission, L.D. [1972]. "Rate-Distortion Theory and Application." *Proc. IEEE*, vol. 60, pp. 800-808.
- Delp, E.J., and Mitchell, O.R. [1979]. "Image Truncation using Block Truncation Coding." *IEEE Trans. Comm.*, vol. COM-27, pp. 1335-1342.
- Deutsch, R. [1965]. *Estimation Theory*, Prentice-Hall, Englewood Cliffs, N.J.
- Dijkstra, E. [1959]. "Note on Two Problems in Connection with Graphs." *Numerische Mathematik*, vol. 1, pp. 269-271.
- Dougherty, E.R. [1992]. *An Introduction to Morphological Image Processing*, SPIE Press, Bellingham, Wash.
- Doyle, W. [1962]. "Operations Useful for Similarity-Invariant Pattern Recognition." *J. ACM*, vol. 9, pp. 259-267.
- Duan, J.R., and Wintz, P.A. [1974]. "Information Preserving Coding for Multispectral Scanner Data." TR-EE-74-15, School of Electrical Engineering, Purdue University, West Lafayette, Ind.
- Duda, R.O., and Hart, P.E. [1972]. "Use of the Hough Transformation to Detect Lines and Curves in Pictures." *Comm. ACM*, vol. 15, no. 1, pp. 11-15.
- Duda, R.O., and Hart, P.E. [1973]. *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York.
- Dudani, S.A., and Luk, A. [1977]. "Locating Straight-Edge Segments on Outdoor Scenes." *Proc. Conf. Pattern Recog. Image Proc.*, vol. 2, pp. 367-380.
- Dyer, C.R. [1983]. "Gauge Inspection Using Hough Transforms." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, no. 6, pp. 621-623.

- Dyer, C.R., and Rosenfeld, A. [1979]. "Thinning Algorithms for Grayscale Pictures." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-1, no. 1, pp. 88-89.
- Elias, P. [1952]. "Fourier Treatment of Optical Processes." *J. Opt. Soc. Am.*, vol. 42, no. 2, pp. 127-134.
- Elliott, D.F., and Rao, K.R. [1983]. *Fast Transforms: Algorithms and Applications*, Academic Press, New York.
- Elsgolc, L.E. [1962]. *Calculus of Variations*, Addison-Wesley, Reading, Mass.
- Equitz, W.H. [1989]. "A New Vector Quantization Clustering Algorithm." *IEEE Trans. Acous. Speech Signal Processing*, vol. ASSP-37, no. 10, pp. 1568-1575.
- Essman, J., and Wintz, P.A. [1973]. "The Effects of Channel Errors in DPCM Systems and Comparison with PCM Systems." *IEEE Trans. on Comm.*, vol. 21, no. 8, pp. 867-877.
- Evans, R.M. [1959]. *An Introduction to Color*, John Wiley & Sons, New York.
- Falconer, D.G. [1970]. "Image Enhancement and Film Grain Noise." *Opt. Acta*, vol. 17, pp. 693-705.
- Falconer, D.G. [1977]. "Target Tracking with the Hough Transform." *Proc. Asilomar Conf. Circ. Syst. Comput.*, vol. 11, pp. 249-252.
- Fine, N.J. [1949]. "On the Walsh Functions." *Trans. Am. Math. Soc.*, vol. 65, pp. 373-414.
- Fine, N.J. [1950]. "The Generalized Walsh Functions." *Trans. Am. Math. Soc.*, vol. 69, pp. 66-77.
- Fischler, M.A. [1980]. "Fast Algorithms for Two Maximal Distance Problems with Applications to Image Analysis." *Pattern Recog.*, vol. 12, pp. 35-40.
- Foley, J.D., and Van Dam, A. [1982]. *Fundamentals of Interactive Computer Graphics*, Addison-Wesley, Reading, Mass.
- Fox, E.A. [1991]. "Advances in Interactive Digital Multimedia Systems." *Computer*, vol. 24, no. 10, pp. 9-21.
- Fram, J.R., and Deutsch, E.S. [1975]. "On the Quantitative Evaluation of Edge Detection Schemes and Their Comparison with Human Performance." *IEEE Trans. Computers*, vol. C-24, no. 6, pp. 616-628.
- Freeman, H. [1961]. "On the Encoding of Arbitrary Geometric Configurations." *IEEE Trans. Elec. Computers*, vol. EC-10, pp. 260-268.
- Freeman, H. [1974]. "Computer Processing of Line Drawings." *Comput. Surveys*, vol. 6, pp. 57-97.
- Freeman, H., and Shapira, R. [1975]. "Determining the Minimum-Area Encasing Rectangle for an Arbitrary Closed Curve." *Comm. ACM*, vol. 18, no. 7, pp. 409-413.
- Freeman, J.A., and Skapura, D.M. [1991]. *Neural Networks: Algorithms, Applications, and Programming Techniques*, Addison-Wesley, Reading, Mass.
- Frei, W., and Chen, C.C. [1977]. "Fast Boundary Detection: A Generalization and a New Algorithm." *IEEE Trans. Computers*, vol. C-26, no. 10, pp. 988-998.
- Frendendall, G.L., and Behrend, W.L. [1960]. "Picture Quality—Procedures for Evaluating Subjective Effects of Interference." *Proc. IRE*, vol. 48, pp. 1030-1034.
- Frieden, B.R. [1972]. "Restoring with Maximum Likelihood and Maximum Entropy." *J. Opt. Soc. Am.*, vol. 62, pp. 511-518.
- Frieden, B.R. [1974]. "Image Restoration by Discrete Deconvolution of Minimal Length." *J. Opt. Soc. Am.*, vol. 64, pp. 682-686.

- Fu, K.S. [1974]. *Syntactic Methods in Pattern Recognition*, Academic Press, New York.
- Fu, K.S. [1982]. *Syntactic Pattern Recognition and Applications*, Prentice-Hall, Englewood Cliffs, N.J.
- Fu, K.S., and Bhargava, B.K. [1973]. "Tree Systems for Syntactic Pattern Recognition," *IEEE Trans. Comput.*, vol. C-22, no. 12, pp. 1087-1099.
- Fu, K.S., Gonzalez, R.C., and Lee, C.S.G. [1987]. *Robotics: Control, Sensing, Vision, and Intelligence*, McGraw-Hill, New York.
- Fu, K.S., and Mui, J.K. [1981]. "A Survey of Image Segmentation." *Pattern Recog.*, vol. 13, no. 1, pp. 3-16.
- Fu, K.S., and Rosenfeld, A. [1976]. "Pattern Recognition and Image Processing." *IEEE Trans. Computers*, vol. C-25, no. 12, pp. 1336-1346.
- Gattis, J., and Wintz, P.A. [1971]. "Automated Techniques for Data Analysis and Transmission." TR-EE-71-37, School of Electrical Engineering, Purdue University, West Lafayette, Ind.
- Gaven, J.V., Jr., Tavitian, J., and Harabedian, A. [1970]. "The Informative Value of Sampled Images as a Function of the Number of Gray Levels Used in Encoding the Images." *Phot. Sci. Eng.*, vol. 14, no. 1, pp. 16-20.
- Gentleman, W.M. [1968]. "Matrix Multiplication and Fast Fourier Transformations." *Bell System Tech. J.*, vol. 47, pp. 1099-1103.
- Gentleman, W.M., and Sande, G. [1966]. "Fast Fourier Transform for Fun and Profit." *Fall Joint Computer Conf.*, vol. 29, pp. 563-578, Spartan, Washington, D.C.
- Gharavi, H., and Tabatabai, A. [1988]. "Sub-band Coding of Monochrome and Color Images." *IEEE Trans. Circuits Sys.*, vol. 35, no. 2, pp. 207-214.
- Giardina, C.R., and Dougherty, E.R. [1988]. *Morphological Methods in Image and Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J.
- Gish, H., and Pierce, J.N. [1968]. "Asymptotically Efficient Quantizer." *IEEE Trans. Info. Theory*, vol. IT-14, pp. 676-683.
- Gooley, M.J.E. [1969]. "Hexagonal Parallel Pattern Transformations." *IEEE Trans. Comput.*, vol. C-18, pp. 733-740.
- Goldmark, P.C., and Hollywood, J.M. [1951]. "A New Technique for Improving the Sharpness of Television Pictures," *Proc. IRE*, vol. 39, pp. 1314-1322.
- Golomb, S.W., and Baumert, L.D. [1963]. "The Search for Hadamard Matrices." *Am. Math. Monthly*, vol. 70, pp. 27-31.
- Gonzalez, R.C. [1972]. "Syntactic Pattern Recognition—Introduction and Survey." *Proc. Natl. Elec. Conf.*, vol. 27, pp. 27-31.
- Gonzalez, R.C. [1985]. "Computer Vision." *Yearbook of Science and Technology*, McGraw-Hill, New York, pp. 128-132.
- Gonzalez, R.C. [1985]. "Industrial Computer Vision." In *Advances in Information Systems Science*, Tou, J.T., ed., Plenum, New York, pp. 345-385.
- Gonzalez, R.C. [1986]. "Image Enhancement and Restoration." In *Handbook of Pattern Recognition and Image Processing*, Young, T.Y. and Fu, K.S., eds., Academic Press, New York, pp. 191-213.
- Gonzalez, R.C., Barrero, A., and Thomason, M.G. [1978]. "A Measure of Scene Content." *Proc. Pattern Recog. Image Proc. Conf.*, vol. 1, pp. 385-389.
- Gonzalez, R.C., Edwards, J.J., and Thomason, M.G. [1976]. "An Algorithm for the Inference of Tree Grammars." *Int. J. Comput. Info. Sci.*, vol. 5, no. 2, pp. 145-163.
- Gonzalez, R.C., and Fitts, B.A. [1975]. "Gray-Level Transformations for Interactive Image Enhancement." *Proc. Second Conf. Remotely Manned Syst.*, pp. 17-19.

- Gonzalez, R.C., and Fittes, B.A. [1977]. "Gray-Level Transformations for Interactive Image Enhancement." *Mechanism and Machine Theory*, vol. 12, pp. 111-122.
- Gonzalez, R.C., and Safabakhsh, R. [1982]. "Computer Vision Techniques for Industrial Applications." *Computer*, vol. 15, no. 12, pp. 17-32.
- Gonzalez, R.C., and Thomason, M.G. [1978]. *Syntactic Pattern Recognition: An Introduction*. Addison-Wesley, Reading, Mass.
- Gonzalez, R.C., Woods, R.E., and Swain, W.T. [1986]. "Digital Image Processing: An Introduction." *Digital Design*, vol. 16, no. 4, pp. 15-20.
- Good, I.J. [1958]. "The Interaction Algorithm and Practical Fourier Analysis." *J. R. Stat. Soc. (Lond.)*, vol. B20, pp. 361-367; *Addendum*, vol. 22, 1960, pp. 372-375.
- Goodman, J.W. [1968]. *Introduction to Fourier Optics*, McGraw-Hill, New York.
- Goodson, K.J., and Lewis, P.H. [1990]. "A Knowledge-Based Line Recognition System." *Pattern Recog. Letters*, vol. 11, no. 4, pp. 295-304.
- Graham, C.H., ed. [1965]. *Vision and Visual Perception*, John Wiley & Sons, New York.
- Graham, D.N. [1967]. "Image Transmission by Two-Dimensional Contour Coding." *Proc. IEEE*, vol. 55, pp. 336-346.
- Graham, R.E. [1958]. "Predictive Quantizing of Television Signals." *IRE Wescon Conv. Rec.*, vol. 2, pt. 2, pp. 147-157.
- Graham, R.L., and Yao, F.F. [1983]. "Finding the Convex Hull of a Simple Polygon." *J. Algorithms*, vol. 4, pp. 324-331.
- Gray, R.M. [1984]. "Vector Quantization." *IEEE Trans. Acous. Speech Signal Processing*, vol. ASSP-1, no. 2, pp. 4-29.
- Green, W.B. [1983]. *Digital Image Processing—A Systems Approach*, Van Nostrand Reinhold, New York.
- Gries, D. [1971]. *Compiler Construction for Digital Computers*, John Wiley & Sons, New York.
- Grosky, W.I., and Jain, R. [1983]. "Optimal Quadtrees for Image Segments." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, no. 1, pp. 77-83.
- Guillemin, E.A. [1949]. *The Mathematics of Circuit Analysis*, John Wiley & Sons, New York.
- Gupta, L., Mohammad, R.S., and Tammana, R. [1990]. "A Neural Network Approach to Robust Shape Classification." *Pattern Recog.*, vol. 23, no. 6, pp. 563-568.
- Gupta, L., and Srinath, M.D. [1988]. "Invariant Planar Shape Recognition Using Dynamic Alignment." *Pattern Recog.*, vol. 21, pp. 235-239.
- Habibi, A. [1971]. "Comparison of Nth Order DPCM Encoder with Linear Transformations and Block Quantization Techniques." *IEEE Trans. Comm. Tech.*, vol. COM-19, no. 6, pp. 948-956.
- Habibi, A. [1972]. "Two-Dimensional Bayesian Estimate of Images." *Proc. IEEE*, vol. 60, pp. 878-883.
- Habibi, A. [1974]. "Hybrid Coding of Pictorial Data." *IEEE Trans. Comm.*, vol. COM-22, no. 5, pp. 614-624.
- Habibi, A., and Wintz, P.A. [1971]. "Image Coding by Linear Transformations and Block Quantization." *IEEE Trans. Comm. Tech.*, vol. COM-19, pp. 50-62.
- Hadamard, J. [1893]. "Resolution d'une Question Relative aux Determinants." *Bull. Sci. Math.*, Ser. 2, vol. 17, Part 1, pp. 240-246.
- Haddon, J.F., and Boyce, J.F. [1990]. "Image Segmentation by Unifying Region and Boundary Information." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 10, pp. 929-948.

- Hall, E.L. [1972]. "Automated Computer Diagnosis Applied to Lung Cancer." *Proc. 1972 Int. Conf. on Cybernetics Soc.*, New Orleans.
- Hall, E.L. [1974]. "Almost Uniform Distributions for Computer Image Enhancement." *IEEE Trans. Computers*, vol. C-23, no. 2, pp. 207-208.
- Hall, E.L. [1979]. *Computer Image Processing and Recognition*, Academic Press, New York.
- Hall, E.L. et al. [1971]. "A Survey of Preprocessing and Feature Extraction Techniques for Radiographic Images." *IEEE Trans. Comput.*, vol. C-20, no. 9, pp. 1032-1044.
- Hall, E.L., and Frei, W. [1976]. "Invariant Features for Quantitative Scene Analysis." Final Report, Contract F 08606-72-C-0008, Image Processing Institute, University of Southern California, Los Angeles.
- Hamming, R.W. [1950]. "Error Detecting and Error Correcting Codes." *Bell Sys. Tech. J.*, vol. 29, pp. 147-160.
- Hammond, J.L., and Johnson, R.S. [1962]. "Orthogonal Square-Wave Functions." *J. Franklin Inst.*, vol. 273, pp. 211-225.
- Haralick, Haralick, R.M. [1979]. "Statistical and Structural Approaches to Texture." *Proc. 4th Int. Joint Conf. Pattern Recog.*, pp. 45-60.
- Haralick, R.M., and Lee, J.S.J. [1990]. "Context Dependent Edge Detection and Evaluation." *Pattern Recog.*, vol. 23, no. 1-2, pp. 1-20.
- Haralick, R.M., and Shapiro, L.G. [1985]. "Survey: Image Segmentation." *Comput. Vision, Graphics, Image Proc.*, vol. 29, pp. 100-132.
- Haralick, R.M., Sternberg, S.R., and Zhuang, X. [1987]. "Image Analysis Using Mathematical Morphology." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no. 4, pp. 532-550.
- Haralick, R.M., Shanmugan, R., and Dinstein, I. [1973]. "Textural Features for Image Classification." *IEEE Trans Syst. Man Cyb.*, vol. SMC-3, no. 6, pp. 610-621.
- Harmuth, H.F. [1968]. "A Generalized Concept of Frequency and Some Applications." *IEEE Trans. Info. Theory*, vol. IT-14, no. 3, pp. 375-382.
- Harmuth, H.F. [1970]. *Transmission of Information by Orthogonal Signals*, Springer-Verlag, New York.
- Harris, J.L. [1964]. "Resolving Power and Decision Theory." *J. Opt. Soc. Am.*, vol. 54, pp. 606-611.
- Harris, J.L. [1966]. "Image Evaluation and Restoration." *J. Opt. Soc. Am.*, vol. 56, pp. 569-574.
- Harris, J.L. [1968]. "Potential and Limitations of Techniques for Processing Linear Motion-Degraded Images." In *Eval. Motion Degraded Images*, NASA Publ. SP-193, pp. 131-138.
- Hart, P.E., Nilsson, N.J., and Raphael, B. [1968]. "A Formal Basis for the Heuristic Determination of Minimum-Cost Paths." *IEEE Trans. Syst. Man Cyb.*, vol. SMC-4, pp. 100-107.
- Hayes-Roth, F., Waterman, D.A., and Lenat, D.B. [1983]. *Building Expert Systems*, Addison-Wesley, Reading, Mass.
- Healy, D.J., and Mitchell, O.R. [1981]. "Digital Video Bandwidth Compression Using Block Truncation Coding." *IEEE Trans. Comm.*, vol. COM-29, no. 12, pp. 1809-1817.
- Hebb, D.O. [1949]. *The Organization of Behavior: A Neuropsychological Theory*, John Wiley & Sons, New York.
- Hecht, E., and Zajac, A. [1975]. *Optics*, Addison-Wesley, Reading, Mass.
- Helstrom, C.W. [1967]. "Image Restoration by the Method of Least Squares." *J. Opt. Soc. Am.*, vol. 57, no. 3, pp. 297-303.

۸۱۷ کتاب شناسی

- Henderson, K.W. [1964]. "Some Notes on the Walsh Functions." *IEEE Trans. Electronic Computers*, vol. EC-13, no. 1, pp. 50–52.
- Herbrand, J. [1930]. "Recherches sur la Théorie de la Démonstration." *Travaux de la Société des Sciences et des Lettres de Varsovie, Classe III, Sciences Mathématiques et Physiques*, no. 33.
- Horn, B.K.P. [1986]. *Robot Vision*, McGraw-Hill, New York.
- Horner, J.L. [1969]. "Optical Spatial Filtering with the Least-Mean-Square Error Filter." *J. Opt. Soc. Am.*, vol. 59, pp. 553–558.
- Horowitz, M. [1957]. "Efficient Use of a Picture Correlator." *J. Opt. Soc. Am.*, vol. 47, p. 327.
- Horowitz, S.L., and Pavlidis, T. [1974]. "Picture Segmentation by a Directed Split-and-Merge Procedure." *Proc. 2nd Int. Joint Conf. Pattern Recog.*, pp. 424–433.
- Hotelling, H. [1933]. "Analysis of a Complex of Statistical Variables into Principal Components." *J. Educ. Psychol.*, vol. 24, pp. 417–441, 498–520.
- Hough, P.V.C. [1962]. "Methods and Means for Recognizing Complex Patterns." U.S. Patent 3,069,654.
- Hsu, C.C., and Huang, J.S. [1990]. "Partitioned Hough Transform for Ellipsoid Detection." *Pattern Recog.*, vol. 23, no. 3–4, pp. 275–282.
- Hu, M.K. [1962]. "Visual Pattern Recognition by Moment Invariants." *IRE Trans. Info. Theory*, vol. IT-8, pp. 179–187.
- Huang, T.S. [1965]. "PCM Picture Transmission." *IEEE Spectrum*, vol. 2, no. 12, pp. 57–63.
- Huang, T.S. [1966]. "Digital Picture Coding." *Proc. Natl. Electron. Conf.*, pp. 793–797.
- Huang, T.S. [1968]. "Digital Computer Analysis of Linear Shift-Variant Systems." in *Evaluation of Motion-Degraded Images*, NASA Publ. SP-193, pp. 83–87.
- Huang, T.S. [1968]. "Run-length Coding and its Extensions." *EG&G Tech. Report*, No. B-3742. (Also in *Picture Bandwidth Compression*, Huang, T.S., and Tretiak, O.J., eds., [1972], Gordon and Breech, New York).
- Huang, T.S., ed. [1975]. *Picture Processing and Digital Filtering*, Springer, New York.
- Huang, T.S. [1981]. *Image Sequence Analysis*, Springer-Verlag, New York.
- Huang, T.S., and Hussian, A.B.S. [1972]. "Facsimile Coding by Skipping White." *IEEE Trans. Comm.*, vol. COM-23, no. 12, pp. 1452–1466.
- Huang, Y., and Schultheiss, P.M. [1963]. "Block Quantization of Correlated Gaussian Random Variables." *IEEE Trans. Commun. Syst.*, vol. CS-11, pp. 289–296.
- Huang, T.S., and Tretiak, O.J. (eds.). [1972]. *Picture Bandwidth Compression*, Gordon and Breech, New York.
- Huang, T.S., Yang, G.T., and Tang, G.Y. [1979]. "A Fast Two-Dimensional Median Filtering Algorithm." *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. ASSP-27, pp. 13–18.
- Huertas, A., Cole, W., and Nevatia, R. [1990]. "Detecting Runways in Complex Airport Scenes." *Comput. Vision, Graphics, Image Proc.*, vol. 51, no. 2, pp. 107–145.
- Huffman, D.A. [1952]. "A Method for the Construction of Minimum Redundancy Codes." *Proc. IRE*, vol. 40, no. 10, pp. 1098–1101.
- Hummel, R.A. [1974]. "Histogram Modification Techniques." Technical Report TR-329, F-44620-72C-0062, Computer Science Center, University of Maryland, College Park, Md.
- Hunt, B.R. [1971]. "A Matrix Theory Proof of the Discrete Convolution Theorem." *IEEE Trans. Audio and Electroacoust.*, vol. AU-19, no. 4, pp. 285–288.

- Hunt, B.R. [1973]. "The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer." *IEEE Trans. Comput.*, vol. C-22, no. 9, pp. 805–812.
- Hunter, R., and Robinson, A.H. [1980]. "International Digital Facsimile Coding Standards." *Proc. IEEE*, vol. 68, no. 7, pp. 854–867.
- IEEE Trans. Circuits and Syst.* [1975]. Special issue on digital filtering and image processing, vol. CAS-2, pp. 161–304.
- IEEE Trans. Computers* [1972]. Special issue on two-dimensional signal processing, vol. C-21, no. 7.
- IEEE Trans. Comm.* [1981]. Special issue on picture communication systems, vol. COM-29, no. 12.
- IES Lighting Handbook* [1972]. Illuminating Engineering Society Press, New York.
- Jain, A.K. [1975]. "A Fast Karhunen–Loeve Transform for a Class of Random Processes." *IEEE Trans. Commun.*, vol. COM-24, pp. 1023–1029.
- Jain, A.K. [1979]. "A Sinusoidal Family of Unitary Transforms." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-1, no. 4, pp. 356–365.
- Jain, A.K. [1981]. "Image Data Compression: A Review." *Proc. IEEE*, vol. 69, pp. 349–389.
- Jain, A.K. [1989]. *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, N.J.
- Jain, A.K., and Angel, E. [1974]. "Image Restoration, Modeling, and Reduction of Dimensionality." *IEEE Trans. Computers*, vol. C-23, pp. 470–476.
- Jain, J.R., and Jain, A.K. [1981]. "Displacement Measurement and Its Application in Interframe Image Coding." *IEEE Trans. Comm.*, vol. COM-29, pp. 1799–1808.
- Jain, R. [1981]. "Dynamic Scene Analysis Using Pixel-Based Processes." *Computer*, vol. 14, no. 8, pp. 12–18.
- Jain, R. [1983]. "Segmentation of Frame Sequences Obtained by a Moving Observer." Report GMR-4247, General Motors Research Laboratories, Warren, Mich.
- Jang, B.K., and Chin, R.T. [1990]. "Analysis of Thinning Algorithms Using Mathematical Morphology." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 6, pp. 541–551.
- Jayant, N.S., ed. [1976]. *Waveform Quantization and Coding*, IEEE Press, New York.
- Kahaner, D.K. [1970]. "Matrix Description of the Fast Fourier Transform." *IEEE Trans. Audio Electroacoustics*, vol. AU-18, no. 4, pp. 442–450.
- Kak, A.C., and Slaney, M. [1988]. *Principles of Computerized Tomographic Imaging*, IEEE Press, New York.
- Karhunen, K. [1947]. "Über Lineare Methoden in der Wahrscheinlichkeitsrechnung." *Ann. Acad. Sci. Fennicae*, Ser. A137. (Translated by I. Selin in "On Linear Methods in Probability Theory," T-131, 1960, The RAND Corp., Santa Monica, Calif.)
- Ketcham, D.J. [1976]. "Real-Time Image Enhancement Techniques." *Proc. Soc. Photo-Optical Instrum. Eng.*, vol. 74, pp. 120–125.
- Khanna, T. [1990]. *Foundations of Neural Networks*, Addison-Wesley, Reading, Mass.
- Kimme, C., Ballard, D.H., and Sklansky, J. [1975]. "Finding Circles by an Array of Accumulators." *Comm. ACM*, vol. 18, no. 2, pp. 120–122.
- Kirsch, R. [1971]. "Computer Determination of the Constituent Structure of Biological Images." *Comput. Biomed. Res.*, vol. 4, pp. 315–328.
- Kiver, M.S. [1965]. *Color Television Fundamentals*, McGraw-Hill, New York.
- Klinger, A. [1972]. "Patterns and Search Statistics." In *Optimizing Methods in Statistics*, Rustagi, J.S., ed., Academic Press, New York, pp. 303–339.

۸۱۹ کتاب شناسی

- Klinger, A. [1976]. "Experiments in Picture Representation Using Regular Decomposition." *Comput. Graphics Image Proc.*, vol. 5, pp. 68–105.
- Knowlton, K. [1980]. "Progressive Transmission of Gray-Scale and Binary Pictures by Simple, Efficient, and Lossless Encoding Schemes." *Proc. IEEE*, vol. 68, no. 7, pp. 885–896.
- Koch, C., Luo, J., and Mead, C. [1988]. "Computing Motion Using Analog and Resistive Networks." *Computer*, vol. 21, no. 3, pp. 52–63.
- Kodak Plates and Films for Scientific Photography* [1973]. Publication no. P-315, Eastman Kodak Co., Rochester, N.Y.
- Kohler, R.J., and Howell, H.K. [1963]. "Photographic Image Enhancement by Superposition of Multiple Images." *Photogr. Sci. Eng.*, vol. 7, no. 4, pp. 241–245.
- Koschman, A. [1954]. "On the Filtering of Nonstationary Time Series." *Proc. 1954 Natl. Electron. Conf.*, p. 126.
- Kovasznay, L.S.G., and Joseph, H.M. [1953]. "Processing of Two-Dimensional Patterns by Scanning Techniques." *Science*, vol. 118, pp. 475–477.
- Kovasznay, L.S.G., and Joseph, H.M. [1955]. "Image Processing." *Proc. IRE*, vol. 43, pp. 560–570.
- Kramer, H.P., and Mathews, M.V. [1956]. "A Linear Coding for Transmitting a Set of Correlated Signals." *IRE Trans. Info. Theory*, vol. IT-2, pp. 41–46.
- Kushnir, M., Abe, K., and Matsumoto, K. [1985]. "Recognition of Handprinted Hebrew Characters Using Features Selected in the Hough Transform Space." *Pattern Recog.*, vol. 18, no. 2, pp. 103–114.
- Langdon, G.C., and Rissanen, J.J. [1981]. "Compression of Black–White Images with Arithmetic Coding." *IEEE Trans. Comm.*, vol. COM-29, no. 6, pp. 858–867.
- Langford, M.J. [1984]. *The Darkroom Handbook*, Random House, New York.
- Lantuéjoul, C. [1980]. "Skeletonization in Quantitative Metallography." In *Issues of Digital Image Processing*, Haralick, R.M., and Simon, J.C. (eds.), Sijthoff and Noordhoff, Groningen, The Netherlands.
- Lawley, D.N., and Maxwell, A.E. [1963]. *Factor Analysis as a Statistical Method*, Butterworth, London.
- Ledley, R.S. [1964]. "High-Speed Automatic Analysis of Biomedical Pictures." *Science*, vol. 146, no. 3461, pp. 216–223.
- Ledley, R.S., et al. [1965]. "FIDAC: Film Input to Digital Automatic Computer and Associated Syntax-Directed Pattern Recognition Programming System." In *Optical and Electro-Optical Information Processing Systems*, Tippet, J., Beckowitz, D., Clapp, L., Koester, C., and Vanderburgh, A., Jr., eds., MIT Press, Cambridge, Mass., Chap. 33.
- Lee, C.C. [1983]. "Elimination of Redundant Operations for a Fast Sobel Operator." *IEEE Trans. Syst. Man Cybern.*, vol. SMC-13, no. 3, pp. 242–245.
- Lee, S.U., Chung, S.Y., and Park, R.H. [1990]. "A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation." *Comput. Vision, Graphics, Image Proc.*, vol. 52, no. 2, pp. 171–190.
- Legault, R.R. [1973]. "The Aliasing Problems in Two-Dimensional Sampled Imagery." In *Perception of Displayed Information*, Biberman, L.M., ed., Plenum Press, New York.
- Lema, M.D., and Mitchell, O.R. [1984]. "Absolute Moment Block Truncation Coding and Its Application to Color Images." *IEEE Trans. Comm.*, vol. COM-32, no. 10, pp. 1148–1157.
- Levine, M.D. [1985]. *Vision in Man and Machine*, McGraw-Hill, New York.

- Limb, J.O., and Rubinstein, C.B. [1978]. "On the Design of Quantizers for DPCM Coders: A Functional Relationship Between Visibility, Probability, and Masking." *IEEE Trans. Comm.*, vol. COM-26, pp. 573-578.
- Linde, Y., Buzo, A., and Gray, R.M. [1980]. "An Algorithm for Vector Quantizer Design." *IEEE Trans. Comm.*, vol. COM-28, no. 1, pp. 84-95.
- Lipkin, B.S., and Rosenfeld, A., eds. [1970]. *Picture Processing and Psychopictorics*. Academic Press, New York.
- Lippmann, R.P. [1987]. "An Introduction to Computing with Neural Nets." *IEEE ASSP Magazine*, vol. 4, pp. 4-22.
- Loève, M. [1948]. "Fonctions Aléatoires de Second Ordre." in P. Lévy, *Processus Stochastiques et Mouvement Brownien*, Hermann, Paris.
- Lohman, A.W., and Paris, D.P. [1965]. "Space-Variant Image Formation." *J. Opt. Soc. Am.*, vol. 55, pp. 1007-1013.
- Lu, H.E., and Wang, P.S.P. [1986]. "A Comment on 'A Fast Parallel Algorithm for Thinning Digital Patterns.'" *Comm. ACM*, vol. 29, no. 3, pp. 239-242.
- Lu, S.Y., and Fu, K.S. [1978]. "A Syntactic Approach to Texture Analysis." *Comput. Graph. Image Proc.*, vol. 7, no. 3, pp. 303-330.
- MacAdam, D.P. [1970]. "Digital Image Restoration by Constrained Deconvolution." *J. Opt. Soc. Am.*, vol. 60, pp. 1617-1627.
- Maragos, P. [1987]. "Tutorial on Advances in Morphological Image Processing and Analysis." *Optical Engineering*, vol. 26, no. 7, pp. 623-632.
- Maren, A.J., Harston, C.T., and Pap, R.M. [1990]. *Handbook of Neural Computing Applications*. Academic Press, New York.
- Mark, D.M., and Abel, D.J. [1985]. "Linear Quadtrees from Vector Representations of Polygons." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, no. 3, pp. 344-349.
- Marr, D. [1982]. *Vision*, Freeman, San Francisco.
- Marr, D., and Hildreth, E. [1980]. "Theory of Edge Detection." *Proc. R. Soc. Lond.*, vol. B207, pp. 187-217.
- Martelli, A. [1972]. "Edge Detection Using Heuristic Search Methods." *Comput. Graphics Image Proc.*, vol. 1, pp. 169-182.
- Martelli, A. [1976]. "An Application of Heuristic Search Methods to Edge and Contour Detection." *Comm. ACM*, vol. 19, no. 2, pp. 73-83.
- Max, J. [1960]. "Quantizing for Minimum Distortion." *IRE Trans. Info. Theory*, vol. IT-6, pp. 7-12.
- McClelland, J.L., and Rumelhart, D.E., eds. [1986]. *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, vol. 2: *Psychological and Biological Models*, the MIT Press, Cambridge, Mass.
- McCulloch, W.S., and Pitts, W.H. [1943]. "A Logical Calculus of the Ideas Imminent in Nervous Activity." *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115-133.
- McFarlane, M.D. [1972]. "Digital Pictures Fifty Years Ago." *Proc. IEEE*, vol. 60, no. 7, pp. 768-770.
- McGlamery, B.L. [1967]. "Restoration of Turbulence-Degraded Images." *J. Opt. Soc. Am.*, vol. 57, no. 3, pp. 293-297.
- McKeown, D.M., Harvey, W.A., and McDermott, J. [1985]. "Rule-Based Interpretation of Aerial Imagery." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, no. 5, pp. 570-585.
- McKeown, D.M., Wilson, A.H., and Wixson, L.E. [1989]. "Automatic Knowledge Ac-

- quisition for Aerial Image Interpretation." *Comput. Vision, Graphics, Image Proc.*, vol. 46, no. 1, pp. 37-81.
- Mees, C.E.K., and James, T.H. [1966]. *The Theory of the Photographic Process*, Macmillan, New York.
- Merlin, P.M., and Farber, D.J. [1975]. "A Parallel Mechanism for Detecting Curves in Pictures." *IEEE Trans. Comput.*, vol. C-24, no. 1, pp. 96-98.
- Meyer, E.R., and Gonzalez, R.C. [1983]. "Spatial Techniques for Digital Image Enhancement and Restoration." *Proc. First South Afr. Symp. Digital Image Proc.*, Univ. of Natal, Durban, South Africa, pp. 137-182.
- Meyer, F., and Beucher, S. [1990]. "Morphological Segmentation." *J. Visual Comm. and Image Representation*, vol. 1, no. 1, pp. 21-46.
- Meyer, H., Rosdolsky, H.G., and Huang, T.S. [1973]. "Optimum Run Length Codes." *IEEE Trans. Comm.*, vol. COM-22, no. 6, pp. 826-835.
- Minsky, M., and Papert, S. [1969]. *Perceptrons: An Introduction to Computational Geometry*, the MIT Press, Cambridge, Mass.
- Mokhtarian, F., and Mackworth, A. [1986]. "A Scale-Based Description and Recognition of Planar Curves and Two-Dimensional Shapes." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, no. 1, pp. 34-43.
- Moon, P. [1961]. *The Scientific Basis of Illuminating Engineering*, Dover, New York.
- Mueller, P.F., and Reynolds, G.O. [1967]. "Image Restoration by Removal of Random Media Degradations." *J. Opt. Soc. Am.*, vol. 57, pp. 1338-1344.
- Muerle, J.L., and Allen, D.C. [1968]. "Experimental Evaluation of Techniques for Automatic Segmentation of Objects in a Complex Scene." In *Pictorial Pattern Recognition*, (G.C. Cheng et al., eds.), Thompson, Washington, D.C.
- Nagao, M., and Matsuyama, T. [1980]. *A Structural Analysis of Complex Aerial Photographs*, Plenum Press, New York.
- Nahim, P.J. [1974]. "The Theory of Measurement of a Silhouette Description for Image Processing and Recognition." *Pattern Recog.*, vol. 6, no. 2, pp. 85-95.
- Narasimhan, R., and Fornango, J.P. [1963]. "Some Further Experiments in the Parallel Processing of Pictures." *IEEE Trans. Elec. Computers*, vol. EC-12, pp. 748-750.
- Narendra, P.M., and Fitch, R.C. [1981]. "Real-Time Adaptive Contrast Enhancement." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, no. 6, pp. 655-661.
- Nelson, C.N. [1971]. "Prediction of Densities in Fine Detail in Photographic Images." *Photogr. Sci. Eng.*, vol. 15, pp. 82-97.
- Netravali, A.N. [1977]. "On Quantizers for DPCM Coding of Picture Signals." *IEEE Trans. Info. Theory*, vol. IT-23, no. 3, pp. 360-370.
- Netravali, A.N., and Limb, J.O. [1980]. "Picture Coding: A Review." *Proc. IEEE*, vol. 68, no. 3, pp. 366-406.
- Niemann, H., Sagerer, G.F., Schröder, S., and Kummert, F. [1990]. "ERNEST: A Semantic Network for Pattern Understanding." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 9, pp. 883-905.
- Nilsson, N.J. [1965]. *Learning Machines: Foundations of Trainable Pattern-Classifying Systems*, McGraw-Hill, New York.
- Nilsson, N.J. [1971]. *Problem Solving Methods in Artificial Intelligence*, McGraw-Hill, New York.
- Nilsson, N.J. [1980]. *Principles of Artificial Intelligence*, Tioga, Palo Alto, Calif.
- Noble, B. [1969]. *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, N.J.

- O'Gorman, F., and Clowes, M.B. [1976]. "Finding Picture Edges Through Collinearity of Feature Points." *IEEE Trans. Comput.*, vol. C-25, no. 4, pp. 449-454.
- O'Handley, D.A., and Green, W.B. [1972]. "Recent Developments in Digital Image Processing at the Image Processing Laboratory of the Jet Propulsion Laboratory." *Proc. IEEE*, vol. 60, no. 7, pp. 821-828.
- Ohlander, R.B. [1975]. "Analysis of Natural Scenes." Ph.D. dissertation, Dept. of Computer Science, Carnegie-Mellon Univ., Pittsburgh.
- O'Neill, E.L. [1956]. "Spatial Filtering in Optics." *IRE Trans. Info. Theory*, vol. IT-2, no. 2, pp. 56-65.
- O'Neil, J.B. [1971]. "Entropy Coding in Speech and Television Differential PCM Systems." *IEEE Trans. Info. Theory*, vol. IT-17, pp. 758-761.
- Oppenheim, A.V., and Schafer, R.W. [1975]. *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J.
- Oppenheim, A.V., Schafer, R.W., and Stockham, T.G., Jr. [1968]. "Nonlinear Filtering of Multiplied and Convolved Signals." *Proc. IEEE*, vol. 56, no. 8, pp. 1264-1291.
- Paez, M.D., and Glisson, T.H. [1972]. "Minimum Mean-Square-Error Quantization in Speech PCM and DPCM Systems." *IEEE Trans. Comm.*, vol. COM-20, pp. 225-230.
- Panter, P.F., and Dite, W. [1951]. "Quantization Distortion in Pulse Code Modulation with Nonuniform Spacing of Levels." *Proc. IRE*, vol. 39, pp. 44-48.
- Pao, Y.H. [1989]. *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley, Reading, Mass.
- Papoulis, A. [1962]. *The Fourier Integral and Its Applications*, McGraw-Hill, New York.
- Papoulis, A. [1965]. *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York.
- Papoulis, A. [1968]. *Systems and Transforms with Applications in Optics*, McGraw-Hill, New York.
- Park, R.H., and Choi, W.Y. [1990]. "Comments on 'A Three-Module Strategy for Edge Detection.'" *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 2, pp. 23-24.
- Parker, J.R. [1991]. "Gray Level Thresholding in Badly Illuminated Images." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 8, pp. 813-819.
- Pattern Recognition* [1970]. Special issue on pattern recognition in photogrammetry, vol. 2, no. 4.
- Pavlidis, T. [1972]. "Segmentation of Pictures and Maps Through Functional Approximation." *Comp. Graph. Image Proc.*, vol. 1, pp. 360-372.
- Pavlidis, T. [1977]. *Structural Pattern Recognition*, Springer-Verlag, New York.
- Pavlidis, T., and Liow, Y.T. [1990]. "Integrating Region Growing and Edge Detection." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 3, pp. 225-233.
- Pearson, D.E. [1975]. *Transmission and Display of Pictorial Information*, John Wiley & Sons (Halsted Press), New York.
- Pennebaker, W.B., Mitchell, J.L., Langdon, G.G., Jr., and Arps, R.B. [1988]. "An Overview of the Basic Principles of the Q-coder Adaptive Binary Arithmetic Coder." *IBM J. Res. Dev.*, vol. 32, no. 6, pp. 717-726.
- Perez, A., and Gonzalez, R.C. [1987]. "An Iterative Thresholding Algorithm for Image Segmentation." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no. 6, pp. 742-751.

- Perona, P., and Malik, J. [1990]. "Scale-Space and Edge Detection Using Anisotropic Diffusion." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 7, pp. 629–639.
- Perrin, F.H. [1960]. "Methods of Appraising Photographic Systems." *J. SMPTE*, vol. 49, pp. 151–156 and 239–249.
- Persoon, E., and Fu, K.S. [1977]. "Shape Discrimination Using Fourier Descriptors." *IEEE Trans. Systems Man Cyb.*, vol. SMC-7, no. 2, pp. 170–179.
- Petrou, M., and Kittler, J. [1991]. "Optimal Edge Detector for Ramp Edges." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 5, pp. 483–491.
- Phillips, D.L. [1962]. "A Technique for the Numerical Solution of Certain Integral Equations of the First Kind." *J. Assoc. Comp. Mach.*, vol. 9, pp. 84–97.
- Piech, M.A. [1990]. "Decomposing the Laplacian." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 8, pp. 830–831.
- Pitas, I., and Vanetsanopoulos, A.N. [1990]. *Nonlinear Digital Filters: Principles and Applications*, Kluger, Boston.
- Pokorny, C.K., and Gerald, C.F. [1989]. *Computer Graphics: The Principles Behind the Art and Science*, Franklin, Beedle & Associates, Irvine, Calif.
- Pratt, W.K. [1971]. "Spatial Transform Coding of Color Images." *IEEE Trans. Comm. Tech.*, vol. COM-19, no. 6, pp. 980–991.
- Pratt, W.K. [1974]. "Correlation Techniques of Image Registration." *IEEE Trans. Aerospace and Elec. Syst.*, vol. AES-10, no. 3, pp. 353–358.
- Pratt, W.K. [1978]. *Digital Image Processing*, John Wiley & Sons, New York.
- Pratt, W.K. [1991]. *Digital Image Processing*, 2nd ed., John Wiley & Sons, New York.
- Preparata, F.P., and Shamos, M.I. [1985]. *Computational Geometry: An Introduction*, Springer-Verlag, New York.
- Preston, K. [1983]. "Cellular Logic Computers for Pattern Recognition." *Computer*, vol. 16, no. 1, pp. 36–47.
- Prewitt, J.M.S. [1970]. "Object Enhancement and Extraction." in *Picture Processing and Psychopictorics*, Lipkin, B.S. and Rosenfeld, A., eds., Academic Press, New York.
- Price, K.E. [1976]. "Change Detection and Analysis in Multispectral Images." Dept. of Computer Science, Carnegie-Mellon Univ., Pittsburgh.
- Pritchard, D.H. [1977]. "U.S. Color Television Fundamentals—A Review." *IEEE Trans. Consumer Electronics*, vol. CE-23, no. 4, pp. 467–478.
- Proc. IEEE* [1967]. Special issue on redundancy reduction, vol. 55, no. 3.
- Proc. IEEE* [1972]. Special issue on digital picture processing, vol. 60, no. 7.
- Proc. IEEE* [1980]. Special issue on the encoding of graphics, vol. 68, no. 7.
- Proc. IEEE* [1985]. Special issue on visual communication systems, vol. 73, no. 2.
- Proctor, C.W., and Wintz, P.A. [1971]. "Picture Bandwidth Reduction for Noisy Channels." TR-EE 71-30, School of Electrical Engineering, Purdue University, West Lafayette, Ind.
- Rabbani, M., and Jones, P.W. [1991]. *Digital Image Compression Techniques*, SPIE Press, Bellingham, Wash.
- Rajala, S.A., Riddle, A.N., and Snyder, W.E. [1983]. "Application of the One-Dimensional Fourier Transform for Tracking Moving Objects in Noisy Environments." *Comput. Vis. Graph. Image Proc.*, vol. 21, pp. 280–293.
- Ramer, U. [1975]. "Extraction of Line Structures from Photographs of Curved Objects." *Comput. Graphics Image Proc.*, vol. 4, pp. 81–103.

- Ready, P.J., and Wintz, P.A. [1973]. "Information Extraction, SNR Improvement, and Data Compression in Multispectral Imagery." *IEEE Trans. Comm.*, vol. COM-21, no. 10, pp. 1123–1131.
- Rino, C.L. [1969]. "Bandlimited Image Restoration by Linear Mean-Square Estimation." *J. Opt. Soc. Am.*, vol. 59, pp. 547–553.
- Riseman, E.A., and Arbib, M.A. [1977]. "Computational Techniques in Visual Systems. Part II: Segmenting Static Scenes." IEEE Computer Society Repository, R77–87.
- Robbins, G.M., and Huang, T.S. [1972]. "Inverse Filtering for Linear Shift-Variant Imaging Systems." *Proc. IEEE*, vol. 60, pp. 862–872.
- Roberts, L.G. [1965]. "Machine Perception of Three-Dimensional Solids." In *Optical and Electro-Optical Information Processing*, Tippett, J.T., ed., MIT Press, Cambridge, Mass.
- Robinson, G.S. [1976]. "Detection and Coding of Edges Using Directional Masks." University of Southern California, Image Processing Institute, Report no. 660.
- Robinson, J.A. [1965]. "A Machine-Oriented Logic Based on the Resolution Principle." *J. ACM*, vol. 12, no. 1, pp. 23–41.
- Roese, J.A., Pratt, W.K., and Robinson, G.S. [1977]. "Interframe Cosine Transform Image Coding." *IEEE Trans. Comm.*, vol. COM-25, pp. 1329–1339.
- Rosenblatt, F. [1959]. "Two Theorems of Statistical Separability in the Perceptron." In *Mechanisation of Thought Processes: Proc. of Symposium No. 10*, held at the National Physical Laboratory, November 1958, H.M. Stationery Office, London, vol. 1, pp. 421–456.
- Rosenblatt, F. [1962]. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*, Spartan, Washington, D.C.
- Rosenfeld, A. [1969]. *Picture Processing by Computer*, Academic Press, New York.
- Rosenfeld, A. [1972]. "Picture Processing." *Comput. Graph. Image Proc.*, vol. 1, pp. 394–416.
- Rosenfeld, A. [1973]. "Progress in Picture Processing: 1969–71," *Comput. Surv.*, vol. 5, pp. 81–108.
- Rosenfeld, A. [1974]. "Picture Processing: 1973," *Comput. Graph. Image Proc.*, vol. 3, pp. 178–194.
- Rosenfeld, A. et al. [1965]. "Automatic Cloud Interpretation," *Photogrammetr. Eng.*, vol. 31, pp. 991–1002.
- Rosenfeld, A., and Kak, A.C. [1982]. *Digital Picture Processing*, 2nd ed., Academic Press, New York.
- Roth, W. [1968]. "Full Color and Three-Dimensional Effects in Radiographic Displays." *Investigative Radiol.*, vol. 3, pp. 56–60.
- Rudnick, P. [1966]. "Note on the Calculation of Fourier Series." *Math. Comput.*, vol. 20, pp. 429–430.
- Rumelhart, D.E., Hinton, G.E., and Williams, R.J. [1986]. "Learning Internal Representations by Error Propagation." In *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, vol. 1: *Foundations*, Rumelhart, D.E., et al. eds., MIT Press, Cambridge, Mass., pp. 318–362.
- Rumelhart, D.E., and McClelland, J.L., eds. [1986]. *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, vol. 1: *Foundations*, MIT Press, Cambridge, Mass.
- Runge, C. [1903]. *Zeit. für Math. and Physik*, vol. 48, p. 433.

- Runge, C. [1905]. *Zeit. für Math. and Physik*, vol. 53, p. 117.
- Runge, C., and König, H. [1924]. "Die Grundlehren der Mathematischen Wissenschaften." *Vorlesungen über Numerisches Rechnen*, vol. 11, Julius Springer, Berlin.
- Rushforth, C.K., and Harris, R.W. [1968]. "Restoration, Resolution, and Noise." *J. Opt. Soc. Am.*, vol. 58, pp. 539-545.
- Sahoo, P.K., Soltani, S., Wong, A.K.C., and Chan, Y.C. [1988]. "A Survey of Thresholding Techniques." *Comput. Vision, Graphics, Image Proc.*, vol. 4, pp. 233-260.
- Saito, N., and Cunningham, M.A. [1990]. "Generalized E-Filter and its Application to Edge Detection." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 8, pp. 814-817.
- Sakrison, D.J., and Algazi, V.R. [1971]. "Comparison of Line-by-Line and Two-Dimensional Encoding of Random Images." *IEEE Trans. Info. Theory*, vol. IT-17, no. 4, pp. 386-398.
- Salari, E., and Siy, P. [1984]. "The Ridge-Seeking Method for Obtaining the Skeleton of Digital Images." *IEEE Trans. Syst. Man Cyb.*, vol. SMC-14, no. 3, pp. 524-528.
- Sawchuk, A.A. [1972]. "Space-Variant Image Motion Degradation and Restoration." *Proc. IEEE*, vol. 60, pp. 854-861.
- Schalkoff, R.J. [1989]. *Digital Image Processing and Computer Vision*, John Wiley & Sons, New York.
- Schonfeld, D., and Goutsias, J. [1991]. "Optimal Morphological Pattern Restoration from Noisy Binary Images." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 1, pp. 14-29.
- Schowengerdt, R.A. [1983]. *Techniques for Image Processing and Classification in Remote Sensing*, Academic Press, New York.
- Schreiber, W.F. [1956]. "The Measurement of Third Order Probability Distributions of Television Signals." *IRE Trans. Info. Theory*, vol. IT-2, pp. 94-105.
- Schreiber, W.F. [1967]. "Picture Coding." *Proc. IEEE*, (Special issue on Redundancy Reduction), vol. 55, pp. 320-330.
- Schreiber, W.F., and Knapp, C.F. [1958]. "TV Bandwidth Reduction by Digital Coding." *Proc. IRE National Convention*, pt. 4, pp. 88-99.
- Schutten, R.W., and Vermeij, G.F. [1980]. "The Approximation of Image Blur Restoration Filters by Finite Impulse Responses." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 2, pp. 176-180.
- Schwartz, J.W., and Barker, R.C. [1966]. "Bit-Plane Encoding: A Technique for Source Encoding." *IEEE Trans. Aerosp. Elec. Systems*, vol. AES-2, no. 4, pp. 385-392.
- Schwarz, R.E., and Friedland, B. [1965]. *Linear Systems*, McGraw-Hill, New York.
- Scoville, F.W. [1965]. "The Subjective Effect of Brightness and Spatial Quantization." *Q. Rep.*, no. 78, MIT Research Laboratory of Electronics, Cambridge, Mass.
- Seidman, J. [1972]. "Some Practical Applications of Digital Filtering in Image Processing." *Proc. Conf. Comput. Image Proc. Recog.*, University of Missouri, Columbia, vol. 2, pp. 9-1-9-1-16.
- Selin, I. [1965]. *Detection Theory*, Princeton University Press, Princeton, N.J.
- Serra, J. [1982]. *Image Analysis and Mathematical Morphology*, Academic Press, New York.
- Serra, J., ed. [1988]. *Image Analysis and Mathematical Morphology*, vol. 2, Academic Press, New York.
- Sezan, M.I., Rabbani, M., and Jones, P.W. [1989]. "Progressive Transmission of Images Using a Prediction/Residual Encoding Approach." *Opt. Eng.*, vol. 28, no. 5, pp. 556-564.

- Shack, R.V. [1964]. "The Influence of Image Motion and Shutter Operation on the Photographic Transfer Function." *Appl. Opt.*, vol. 3, pp. 1171-1181.
- Shamos, M.I. [1978]. "Computational Geometry." Ph.D. Thesis, Yale University, New Haven, Conn.
- Shanks, J.L. [1969]. "Computation of the Fast Walsh-Fourier Transform." *IEEE Trans. Comput.*, vol. C-18, no. 5, pp. 457-459.
- Shannon, C.E. [1948]. "A Mathematical Theory of Communication." *The Bell Sys. Tech. J.*, vol. XXVII, no. 3, pp. 379-423.
- Shariat, H., and Price, K.E. [1990]. "Motion Estimation with More Than Two Frames." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 5, pp. 417-434.
- Shaw, A.C. [1970]. "Parsing of Graph-Representable Pictures." *J. ACM*, vol. 17, no. 3, pp. 453-481.
- Sheppard, J.J., Jr. [1968]. *Human Color Perception*, Elsevier, New York.
- Sheppard, J.J., Jr., Stratton, R.H., and Gazley, C., Jr. [1969]. "Pseudocolor as a Means of Image Enhancement." *Am. J. Optom. Arch. Am. Acad. Optom.*, vol. 46, pp. 735-754.
- Shih, F.Y.C., and Mitchell, O.R. [1989]. "Threshold Decomposition of Gray-Scale Morphology into Binary Morphology." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 1, pp. 31-42.
- Shore, J.E. [1973]. "On the Application of Haar Functions." *IEEE Trans. Comm.*, vol. COM-21, pp. 209-216.
- Simon, J.C. [1986]. *Patterns and Operators: The Foundations of Data Representations*, McGraw-Hill, New York.
- Sklansky, J., Chazin, R.L., and Hansen, B.J. [1972]. "Minimum-Perimeter Polygons of Digitized Silhouettes." *IEEE Trans. Comput.*, vol. C-21, no. 3, pp. 260-268.
- Slepian, D. [1967a]. "Linear Least-Squares Filtering of Distorted Images." *J. Opt. Soc. Am.*, vol. 57, pp. 918-922.
- Slepian, D. [1967b]. "Restoration of Photographs Blurred by Image Motion." *BSTJ*, vol. 46, pp. 2353-2362.
- Slepian, D., and Pollak, H.O. [1961]. "Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty-I." *Bell Sys. Tech. J.*, vol. 40, pp. 43-64.
- Smith, A.R. [1978]. "Color Gamut Transform Pairs." *Proc. SIGGRAPH '78*, published as *Computer Graphics*, vol. 12, no. 3, pp. 12-19.
- Smith, S.L. [1963]. "Color Coding and Visual Separability in Information Displays." *J. Appl. Psychol.*, vol. 47, pp. 358-364.
- Snider, H.L. [1973]. "Image Quality and Observer Performance." In *Perception of Displayed Information*, Biberman, L.M., ed., Plenum Press, New York.
- Som, S.C. [1971]. "Analysis of the Effect of Linear Smear." *J. Opt. Soc. Am.*, vol. 61, pp. 859-864.
- Sondhi, M.M. [1972]. "Image Restoration: The Removal of Spatially Invariant Degradations." *Proc. IEEE*, vol. 60, no. 7, pp. 842-853.
- Stark, H., ed. [1987]. *Image Recovery: Theory and Application*, Academic Press, New York.
- Stevens, S.S. [1951]. *Handbook of Experimental Psychology*, John Wiley & Sons, New York.
- Stockham, T.G., Jr. [1972]. "Image Processing in the Context of a Visual Model." *Proc. IEEE*, vol. 60, no. 7, pp. 828-842.
- Storer, J.A., and Reif, J.H., eds. [1991]. *Proceedings of DDC '91*, IEEE Computer Society Press, Los Alamitos, Calif.

۸۴۷ کتاب شناسی

- Stumpff, K. [1939]. *Tafeln und Aufgaben zur Harmonischen Analyse und Periodogrammrechnung*, Julius Springer, Berlin.
- Sze, T.W., and Yang, Y.H. [1981]. "A Simple Contour Matching Algorithm." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, no. 6, pp. 676-678.
- Tanimoto, S.L. [1979]. "Image Transmission with Gross Information First." *Comput. Graphics Image Proc.*, vol. 9, pp. 72-76.
- Tasto, M., and Wintz, P.A. [1971]. "Image Coding by Adaptive Block Quantization." *IEEE Trans. Comm. Tech.*, vol. COM-19, pp. 957-972.
- Tasto, M., and Wintz, P.A. [1972]. "A Bound on the Rate-Distortion Function and Application to Images." *IEEE Trans. Info. Theory*, vol. IT-18, pp. 150-159.
- Teh, C.H., and Chin, R.T. [1989]. "On the Detection of Dominant Points on Digital Curves." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 8, pp. 859-872.
- Thomas, J.B. [1969]. *Statistical Communication Theory*, John Wiley & Sons, New York.
- Thomas, L.H. [1963]. "Using a Computer to Solve Problems in Physics." *Application of Digital Computers*, Ginn, Boston.
- Thomason, M.G., and Gonzalez, R.C. [1975]. "Syntactic Recognition of Imperfectly Specified Patterns." *IEEE Trans. Comput.*, vol. C-24, no. 1, pp. 93-96.
- Thompson, W.B. (ed.) [1989]. "Special Issue on Visual Motion." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 5, pp. 449-541.
- Thompson, W.B., and Barnard, S.T. [1981]. "Lower-Level Estimation and Interpretation of Visual Motion." *Computer*, vol. 14, no. 8, pp. 20-28.
- Titchmarsh, E.C. [1948]. *Introduction to the Theory of Fourier Integrals*, Oxford University Press, New York.
- Tomita, F., Shirai, Y., and Tsuji, S. [1982]. "Description of Texture by a Structural Analysis." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, no. 2, pp. 183-191.
- Toriwaki, J.I., Kato, N., and Fukumura, T. [1979]. "Parallel Local Operations for a New Distance Transformation of a Line Pattern and Their Applications." *IEEE Trans. System, Man, Cyb.*, vol. SMC-9, no. 10, pp. 628-643.
- Tou, J.T., and Gonzalez, R.C. [1974]. *Pattern Recognition Principles*, Addison-Wesley, Reading, Mass.
- Toussaint, G.T. [1982]. "Computational Geometric Problems in Pattern Recognition." In *Pattern Recognition Theory and Applications*, Kittler, J., Fu, K.S., and Pau, L.F., eds., Reidel, New York, pp. 73-91.
- Trivedi, M.M., Chen, C., and Cress, D.H. [1990]. "Object Detection by Step-Wise Analysis of Spectral, Spatial, and Topographic Features." *Comput. Vision, Graphics, Image Proc.*, vol. 51, no. 3, pp. 235-255.
- Twomey, S. [1963]. "On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature." *J. Assoc. Comput. Mach.*, vol. 10, pp. 97-101.
- Udpikar, V.R., and Raina, J.P. [1987]. "BTC Image Coding Using Vector Quantization." *IEEE Trans. Comm.*, vol. COM-35, no. 3, pp. 352-356.
- Umeyama, S. [1988]. "An Eigendecomposition Approach to Weighted Graph Matching Problems." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, no. 5, pp. 695-703.
- VanderBrug, G.J., and Rosenfeld, A. [1977]. "Two-Stage Template Matchings." *IEEE Trans. Comput.*, vol. C-26, no. 4, pp. 384-394.
- Van Valkenburg, M.E. [1955]. *Network Analysis*, Prentice-Hall, Englewood Cliffs, N.J.
- Vuylsteke, P., and Kittler, J. [1990]. "Edge-Labeling Using Dictionary-Based Relaxation." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 2, pp. 165-181.

- Walsh, J.L. [1923]. "A Closed Set of Normal Orthogonal Functions." *Am. J. Math.*, vol. 45, no. 1, pp. 5–24.
- Walsh, J.W.T. [1958]. *Photometry*, Dover, New York.
- Warshall, S. [1962]. "A Theorem on Boolean Matrices." *J. ACM*, vol. 9, no. 1, pp. 11–12.
- Webb, J.A., and Aggarwal, J.K. [1981]. "Visually Interpreting the Motion of Objects in Space." *Computer*, vol. 14, no. 8, pp. 40–49.
- Wechsler [1980]. "Texture Analysis—A Survey." *Signal Process*, vol. 2, pp. 271–280.
- Wechsler, W., and Sklansky, J. [1977]. "Automatic Detection of Ribs in Chest Radiographs." *Pattern Recogn.*, vol. 9, no. 1, pp. 21–28.
- Weinberg, L. [1962]. *Network Analysis and Synthesis*, McGraw-Hill, New York.
- Welch, T.A. [1984]. "A Technique for High-Performance Data Compression." *IEEE Computer*, vol. 17, no. 6, pp. 8–19.
- Weszka, J.S. [1978]. "A Survey of Threshold Selection Techniques." *Comput. Graphics Image Proc.*, vol. 7, pp. 259–265.
- Whelchel, J.E., Jr., and Guinn, D.F. [1968]. "The Fast Fourier-Hadamard Transform and its Use in Signal Representation and Classification." *Eascon 1968 Convention Record*, pp. 561–573.
- White, J.M., and Rohrer, G.D. [1983]. "Image Thresholding for Optical Character Recognition and Other Applications Requiring Character Image Extraction." *IBM J. Res. Devel.*, vol. 27, no. 4, pp. 400–411.
- Widrow, B. [1962]. "Generalization and Information Storage in Networks of 'Adaline' Neurons." In *Self-Organizing Systems 1962*, Yovitz, M.C. et al. (eds.), Spartan, Washington, D.C., pp. 435–461.
- Widrow, B., and Hoff, M.E. [1960]. "Adaptive Switching Circuits." *1960 IRE WESCON Convention Record*, Part 4, pp. 96–104.
- Widrow, B., and Stearns, S.D. [1985]. *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J.
- Wilkins, L.C., and Wintz, P.A. [1970]. "Studies on Data Compression, Part I: Picture Coding by Contours; Part II: Error Analysis of Run-Length Codes." TR-EE 70-17, School of Electrical Engineering, Purdue University, West Lafayette, Ind.
- Williamson, J. [1944]. "Hadamard's Determinant Theorem and the Sum of Four Squares." *Duke Math. J.*, vol. 11, pp. 65–81.
- Wintz, P.A. [1972]. "Transform Picture Coding." *Proc. IEEE*, vol. 60, no. 7, pp. 809–820.
- Witten, I.H., Neal, R.M., and Cleary, J.G. [1987]. "Arithmetic Coding for Data Compression." *Comm. ACM*, vol. 30, no. 6, pp. 520–540.
- Wolfe, G.J., and Mannos, J.L. [1979]. "Fast Median Filter Implementation." *Proc. Soc. Photo-Optical Inst. Eng.*, vol. 207, pp. 154–160.
- Wong, R.Y., and Hall, E.L. [1978]. "Scene Matching with Invariant Moments." *Comput. Graph. Image Proc.*, vol. 8, pp. 16–24.
- Wood, R.C. [1969]. "On Optimum Quantization." *IEEE Trans. Info. Theory*, vol. IT-15, pp. 248–252.
- Woods, J.W., and O'Neil, S.D. [1986]. "Subband Coding of Images." *IEEE Trans. Acous. Speech Signal Proc.*, vol. ASSP-35, no. 5, pp. 1278–1288.
- Woods, R.E., and Gonzalez, R.C. [1981]. "Real-Time Digital Image Enhancement." *Proc. IEEE*, vol. 69, no. 5, pp. 643–654.
- Yachida, M. [1983]. "Determining Velocity Maps by Spatio-Temporal Neighborhoods from Image Sequences." *Comput. Vis. Graph. Image Proc.*, vol. 21, no. 2, pp. 262–279.

۱۲۹ کتاب شناسی

- Yamazaki, Y., Wakahara, Y., and Teramura, H. [1976]. "Digital Facsimile Equipment 'Quick-FAX' Using a New Redundancy Reduction Method." *NTC '76*, pp. 6.2-1-6.2-5.
- Yates, F. [1937]. "The Design and Analysis of Latin Squares." Commonwealth Bureau of Soils, Farnham Royal, Burks, England.
- Yu, S.S., and Tsai, W.H. [1990]. "A New Thinning Algorithm for Gray-Scale Images." *Pattern Recog.*, vol. 23, no. 10, pp. 1067-1076.
- Yuan, M., and Li, J. [1987]. "A Production System for LSI Chip Anatomizing." *Pattern Recog. Letters*, vol. 5, no. 3, pp. 227-232.
- Zahn, C.T., and Roskies, R.Z. [1972]. "Fourier Descriptors for Plane Closed Curves." *IEEE Trans. Comput.*, vol. C-21, no. 3, pp. 269-281.
- Zhang, T.Y., and Suen, C.Y. [1984]. "A Fast Parallel Algorithm for Thinning Digital Patterns." *Comm. ACM*, vol. 27, no. 3, pp. 236-239.
- Ziv, J., and Lempel, A. [1977]. "A Universal Algorithm for Sequential Data Compression." *IEEE Trans. Info. Theory*, vol. IT-23, no. 3, pp. 337-343.
- Ziv, J., and Lempel, A. [1978]. "Compression of Individual Sequences Via Variable-Rate Coding." *IEEE Trans. Info. Theory*, vol. IT-24, no. 5, pp. 530-536.
- Zucker, S.W. [1976]. "Region Growing: Childhood and Adolescence." *Comput. Graphics Image Proc.*, vol. 5, pp. 382-399.



FERDOWSI UNIVERSITY OF MASHHAD

Publication No. 369

DIGITAL IMAGE PROCESSING

Rafael C. Gonzalez

University of Tennessee
Perceptics Corporation

Richard E. Woods

Perceptics Corporation

Translator

M. Khademi

D. Jaefari

FERDOWSI UNIVERSITY PRESS



Rafael C. Gonzalez , Richard E. Woods

Digital Image Processing

Translated by:

M. Khademi (Ph.D) , D. Jafari

کتاب پردازش تصویر رقص مبتنی بر دو ویرایش بسیار موفق قبلی و نیز بیست سال تجربه آموزشی و صنعتی مولفان در زمینه پردازش تصویر است. این کتاب به عنوان مقدمه ای درباره مقاهم و روشهای پردازش تصویر و نیز پیمانی برای مطالعه بیشتر در این مبحث متتنوع و در حال توسعه سریع، بسیار مفید است.

در کتاب عنوانی گوناگونی از ارتقا و بازیابی تصویر گرفته تا رمزگذاری، بخش پندی، توصیف و تغییر تصویر پوشش داده می شوند. این مباحث بوسیله تعداد زیادی تصاویر پردازش شده با رایانه پشتیبانی می شوند. یک کتاب راهنمای نیز برای مدرسان موجود است که حاوی حل تمام سوالات، توضیحات بیشتر در مورد مباحث و تعدادی تعریف پروژه میباشد.

عنوان جدید

- یک فصل جدید در مورد تشخیص و تغییر تصویر
- بخشهای جدید با موضوعات شکل شناسی تصویر، شبکه های عصبی و تحلیل تصویر مبتنی بر دانش
- مباحث جامع در مورد مدل های رنگی برای پردازش تصویر
- نکاهی جدید به فشرده سازی داده های تصویری ، که کارهای جدید در این مبحث مهم تغییر فشرده سازی تصویر استاد (برای دورنگار) و استانداردهای فشرده سازی داده را پوشش میدهد
- مرجع شناسی گسترده تر و مسائل بیشتر در انتهای هر فصل
- تجدید نظر در مورد چند مبحث معمولی و مثال های جدید با استفاده از تصاویر با تکنیک بالاتر