

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِيْمِ



دانشگاه آزاد اسلامی واحد شبستر

پروژه برای دریافت درجه کارشناسی «B.Sc.»

رشته مهندسی کامپیوتر، گرایش نرم افزار

عنوان

بررسی و پیاده سازی عملکردهای کاربردی شبکه های عصبی در پردازش تصویر

استاد راهنما

جناب آقای مهندس جلالی

نگارش

سعید سلیمانی وايقان

بهار ۱۳۹۱

تقدیم به:

ای پدر از تو هر چه می گوییم باز هم کم می آورم

خورشیدی شدی و از روشنایی ات جان گرفتم و در نامیدی ها نازم را

کشیدی و لبریزم کردی از شوق

اکنون حاصل دستان خسته ات رمز موفقیتم شد

به خودم تبریک می گوییم که تو را دارم و دنیا با همه بزرگیش مثل تو را

ندارد.....

و تو ای مادر، ای شوق زیبایی نفس کشیدن

ای روح مهربان هستی ام

تو رنگ شادی هایم شدی و لحظه ها را با تمام وجود از من دور کردی و

عمری خستگی ها را به جان خربدی تا اکنون توانستی طعم خوش

پیروزی را به من بچشانی

چیدم گلی ز باغ ادب تا بروز عید

در بارگاه میر ادب پرور آورم

حیف است با خسان گل دانش کنی نثار

من گل نثار مردم دانشور آورم

شهریار

سپاس گزاری

نمی توانم معنایی بالاتر از تقدیر و تشکر بر زبانم جاری سازم و سپاس خود را در وصف استادان خویش آشکار نمایم، که هر چه گوییم و سرایم، کم گفته ام.

فهرست مطالب

عنوان	صفحه
چکیده	۱
فصل اول	
۱-۱ مقدمه ای بر شبکه های عصبی	۲
۱-۱-۱ شبکه های عصبی	۲
۱-۱-۲ ساختار شبکه های عصبی	۴
۱-۱-۳ کاربردهای شبکه های عصبی مصنوعی ANN	۶
۱-۱-۴ قوانین برانگیختگی	۸
۱-۱-۵ شبکه عصبی مصنوعی ساده	۱۰
۱-۱-۶ شبکه های عصبی بر مبنای روش آموزش	۱۱
۱-۱-۷ مزایای شبکه های عصبی	۱۳
۱-۱-۸ کاربرد شبکه های عصبی	۱۴
۱-۱-۹ معایب ANNها	۱۵
۱-۲ نورون مصنوعی	۱۵
۱-۳ ترون پیچیده تر	۱۷
۱-۴ نورون ساده	۱۸
۱-۵ پرسپترون	۱۸
۱-۶ پرسپترون های ساده	۱۹

۱۹.....	۱-۵-۲ قدرت پرسپترون
۲۰.....	۱-۵-۳ دنباله های پرسپترون
۲۱.....	۱-۵-۴ شبکه های پرسپترون چند لایه

فصل دوم

۲۴.....	۲-۱ مقدمه ای بر پردازش تصویر
۲۴.....	۲-۱-۱ آشنایی با مفاهیم اولیه در پردازش تصویر
۲۶.....	۲-۱-۲ پردازش تصاویر
۲۶.....	۲-۱-۳ کاربردهای پردازش تصویر
۲۷.....	۲-۱-۴ مقدمه پردازش تصویر
۲۹.....	۲-۱-۵ تصویر دیجیتالی
۳۰.....	۲-۱-۶ مفاهیم اولیه در پردازش تصویر
۳۰.....	۲-۱-۶-۱ مفهوم پیکسل در پردازش تصویر
۳۱.....	۲-۱-۶-۲ مفهوم عمق بیتی
۳۱.....	۲-۱-۶-۳ مفهوم بعد یک تصویر
۳۱.....	۲-۱-۶-۴ چگونگی تشکیل رنگ در چشم انسان
۳۳.....	۲-۲ پردازش تصاویر رنگی
۳۳.....	۲-۳-۲ انواع مدل رنگ
۳۴.....	۱-۳-۲ مدل رنگ RGB
۳۶.....	۲-۳-۲ مدل رنگ CMY
۳۶.....	۳-۳-۲ مدل رنگ YIQ

۳۷.....	HIS ۲-۳-۴ مدل رنگ
۳۸.....	۴-۲ انواع پردازش تصویر
۳۹.....	۴-۴-۲ پردازش قاب
۳۹.....	RayTracing ۲-۵
۴۱.....	۲-۲ عارزش پیکسل
۴۲.....	۲-۷ اتصالات پیکسل ها در تصاویر
۴۳.....	۲-۸ درون یابی
۴۴.....	۲-۸-۱ Nearest Neighbor Interpolation
۴۴.....	۲-۸-۲ Bilinear Interpolation
۴۴.....	۲-۸-۳ Bicubic Interpolation
۴۵.....	۲-۸-۴ Fractal Interpolation
۴۵.....	۲-۹-۱ انواع تصویر
۴۵.....	۲-۹-۱ تصاویر دودوئی (Binary)
۴۶.....	۲-۹-۲ تصاویر grayscale
۴۷.....	۲-۹-۳ تصاویر RGB
۴۷.....	۲-۹-۴ خاکستری (Grayscale) کردن تصویر
۴۹.....	۲-۱۰ تفریق دو تصویر
۵۱.....	۲-۱۱ جمع دو تصویر
۵۲.....	۲-۱۲ مکمل کردن تصویر
۵۴.....	۲-۱۳ میانگین گیری از تصویر

۵۵	۱۴-۲ هیستوگرام تصویر
۵۷	۱۵-۲ تعدل هیستوگرام
۵۸	۱۶-۲ فشرده سازی تصاویر
۵۹	۱۶-۲ روشن JPEG
۶۰	۱۶-۲ روشن MPEG
۶۰	۱۶-۲ روشن MP3
۶۰	۱۶-۲ روشن MPEG2
۶۰	۱۶-۲ روشن MPEG4
۶۱	۱۷-۲ تصاویر آنالوگ
۶۲	۱۸-۲ تصاویر رقومی (دیجیتالی)
۶۳	۱۹-۲ مقادیر پیکسلها
۶۳	۲۰-۲ دقت تصویر
۶۴	۲۱-۲ فیلتر کردن تصویر
۶۷	۲۱-۲ فیلترهای پایین گذر میانگین
۶۸	۲۱-۲ فیلتر میانه (Median)
۶۹	۲۱-۲ فیلترهای تیز کننده حوزه مکانی
۷۰	۲۱-۲-۳ فیلتر تیز کننده
۷۱	۲۱-۲-۴ فیلتر تیز کننده لابلس ۲
۷۲	۲۲-۲ فراتفکیک پذیری
۷۴	۲۲-۲-۱ پیکربندی تکنیک فراتفکیک پذیری

۷۴	۲۳-۲ شکل شناسی
۷۵	۲۴-۲ عملگرهای مجموعه ای پایه
۷۷	۲۵-۲ گسترش باینری مورفولوژیکی
۷۸	۲۶-۲ سایش باینری مورفولوژیکی
۸۰	۲۷-۲ بستن و بازکردن باینری مورفولوژیکی
۸۲	۲۸-۲ استخراج اسکلت بندی اشیاء
۸۴	۲۹-۲ الگوریتم اسکلت بندی
۸۶	۳۰-۲ نویز در تصویر
۸۷	۳۰-۱ کاهش نویز با Neat Image

فصل سوم

۹۰	۱-۳ مقدمه
۹۲	۲-۳ تشخیص الگو و تشخیص چهره
۹۴	۱-۲-۳ الگوها و کلاس های الگوها
۹۵	۲-۲-۳ مسائل اساسی در طراحی سیستم تشخیص الگو
۹۷	۳-۲-۳ یادگیری و تمرین دادن
۹۸	۴-۲-۳ کلیات یک سیستم تشخیص الگو
۹۹	۳-۳ تشخیص چهره
۹۹	۱-۳-۳ کلیات یک سیستم تشخیص چهره عام
۱۰۳	۴-۳ روش های استخراج ویژگی در تشخیص چهره
۱۰۳	۱-۴-۳ تشخیص چهره با استفاده از چهره - ویژه ها

۱۰۶	۲-۴-۳ کلیات سیستم تشخیص چهره
۱۰۹	۳-۴-۳ محاسبه ی چهره - ویژه ها
۱۱۰	۳-۵ تشخیص چهره با استفاده از شبکه های عصبی
۱۱۱	۳-۵-۱ مدل نرون
۱۱۲	۳-۵-۲ توابع انتقال
۱۱۴	۳-۵-۳ نرونی با ورودی برداری
۱۱۵	۴-۵-۳ معماری های شبکه
۱۱۵	۳-۵-۴ یک لایه از نرون ها
۱۱۷	۳-۵-۵ چندین لایه از نرون ها
۱۱۸	۳-۵-۶ یادگیری شبکه های عصبی
۱۲۱	۳-۵-۷ دسته بندی با استفاده از شبکه های عصبی
۱۲۳	مراجع
۱۲۶	پیوست

فهرست جداول

عنوان	صفحه
جدول ۱-۲ صحت قبل از اعمال قوانین برانگیختگی	۹
جدول ۱-۳ صحت بعد از اعمال قوانین برانگیختگی	۹
جدول ۲-۱: نمونه برخی از ارزش پیکسل	۴۱

فهرست اشکال

عنوان	صفحه
شکل ۱-۱ مشخصات مسائل در خور شبکه های عصبی مصنوعی (ANN) Artificial Neural Network	۷
شکل ۲-۱: نحوه تشکیل ورودی های نرون	۱۰
شکل ۳-۱ نحوه عملکرد پسپترون	۱۹
شکل ۴-۱ شبکه پرسپترون با یک لایه پنهان	۲۲
شکل ۱-۲ منحنی حساسیت چشم برای مشاهده کننده استاندارد در مقابل دریافت نور رنگی	۳۲
شکل ۲-۲ طیف رنگ ها نسبت به طول موج	۳۲
شکل ۳-۲ مکعب واحد در در مقادیر G و R و B	۳۵
شکل ۴-۲ پردازش هندسی بر روی تصویر	۳۸
شکل ۵-۲ مخلوط کردن دو یا چند تصویر با یکدیگر	۳۹
شکل ۶-۲: اتصالات تصاویر دو بعدی	۴۲
شکل ۷-۲: اتصالات تصاویر سه بعدی	۴۳
شکل ۸-۲: نحوه نمایش تصاویر دودویی	۴۶
شکل ۹-۲: نحوه نمایش تصاویر grayscale	۴۶
شکل ۱۰-۲: نحوه نمایش تصاویر RGB	۴۷
شکل ۱۱-۲: نمایش تصویر RGB به همراه تصویر Grayscale	۴۹
شکل ۱۲-۲ تفریق دو تصویر	۵۰

..... ۵۱	شکل ۱۳-۲ جمع دو تصویر
..... ۵۳ شکل ۱۴-۲ الف: مکمل تصویر
..... ۵۳ شکل ۱۴-۲ ب: تصویر ماموگرام
..... ۵۴ شکل ۱۵-۲ میانگین گیری از تصویر
..... ۵۶ شکل ۱۶-۲: نرمال سازی هیستوگرام
..... ۵۷ شکل ۱۷-۲ تعدل سازی هیستوگرام بر روی تصویر
..... ۶۲ شکل ۱۸-۲ تصویر آنالوگ (عکس هوایی که نیاز به اصلاح و پردازش ندارد)
..... ۶۲ شکل ۱۹-۲ تصاویر رقومی، پیکسلی
..... ۶۴ شکل ۲۰-۲ دقت تصویر ۳ بیتی و دقت تصویر ۸ بیتی
..... ۶۵ شکل ۲۱-۲ تصویرداری نویز فلفل - نمکی
..... ۶۶ شکل ۲۲-۲ فیلتر تیز کننده بر روی تصویر
..... ۶۷ شکل ۲۳-۲ فیلتر میانگین 3×3
..... ۶۹ شکل ۲۴-۲ فیلتر میانه بر روی تصویر
..... ۷۰ شکل ۲۵-۲ اعمال فیلتر پایین گذر و بالا گذر بر روی تصویر
..... ۷۱ شکل ۲۶-۲ فیلتر تیز کننده با یک ماسک 3×3
..... ۷۱ شکل ۲۷-۲ فیلتر لاپلاسین ترکیبی
..... ۷۲ شکل ۲۸-۲ فیلتر تیزکننده لاپلاس ۲ بر روی تصویر
..... ۷۵ شکل ۲۹-۲ پیاده سازی عملگر اجتماع برای دو تصویر باینری با or کردن پیکسل های متناظر
..... ۷۵ دو تصویر

شکل ۳۰-۲ پیاده سازی عملگر اشتراک برای دو تصویر با and کردن پیکسل های متناظر دو

..... تصویر ۷۶

..... شکل ۳۱-۲ تفاضل به شکل اشتراک تصویر دوم و مکمل تصویر اول ۷۶

..... شکل ۳۲-۲ پیاده سازی عملگر مکمل ۷۶

..... شکل ۳۳-۲ تصویر اصلی قبل از گسترش باینری مورفولوژیکی ۷۷

..... شکل ۳۴-۲ حاصل گسترش تصویر با گسترش باینری مورفولوژیکی ۷۸

..... شکل ۳۵-۲ تصویر اصلی قبل از سایش باینری مورفولوژیکی ۷۹

..... شکل ۳۶-۲ حاصل گسترش تصویر با عنصر ساختمانی 3×3 ۷۹

..... شکل ۳۷-۲ تصویر اصلی قبل از اعمال موفولوژیکی ۸۰

..... شکل ۳۸-۲ باز کردن مورفولوژیکی در تصویر ۸۱

..... شکل ۳۹-۲ بستن مورفولوژیکی در تصویر ۸۱

..... شکل ۴۰-۲ اسکلت بنده بر روی تصویر ۸۲

..... شکل ۴۱-۲ روش Hilditch با همسایه های هشتگانه ۸۳

..... شکل ۴۲-۲ اسکلت استخراج شده برای تعدادی تصویر با استفاده از الگوریتم Hilditch ۸۵

..... شکل ۴۳-۲ کاهش نویز با Neat Image ۸۷

..... شکل ۳-۱ دو کلاس مجزا. بردار الگو از دو کمیت weight و height تشکیل شده است ۹۶

..... شکل ۲-۳ : دیاگرام بلوکی یک سیستم تشخیص الگوی تطبیقی ۹۸

..... شکل ۳-۳ : یک سیستم تشخیص چهره‌ی عام ۱۰۰

..... شکل ۴-۳ (a) مجموعه‌ی تمرین تصاویر چهره به صورت نمونه (b) تصویر چهره‌ی میانگین

..... مجموعه‌ی تمرین ۱۰۴

..... ۱۰۵	بدست آمده است
..... ۱۰۹	شکل ۶-۳ دیاگرام بلوکی سیستم تشخیص چهره‌ی ارائه شده
..... ۱۱۱	شکل ۷-۳ یک نرون با ورودی اسکالر و بدون بایاس
..... ۱۱۲	شکل ۸-۳ یک نرون با ورودی اسکالر و بایاس
..... ۱۱۲	شکل ۹-۳ تابع انتقال hard-limit
..... ۱۱۳	شکل ۱۰-۳ تابع انتقال خطی
..... ۱۱۳	شکل ۱۱-۳ تابع انتقال log-sigmoid
..... ۱۱۴	شکل ۱۲-۳ یک نرون با بردار ورودی R عنصری
..... ۱۱۵	شکل ۱۳-۳ یک نرون با بردار ورودی R عنصری (نشانه‌گذاری سطح بالاتر)
..... ۱۱۶	شکل ۱۴-۳ یک شبکه‌ی یک لایه‌ای با R عنصر ورودی و S نرون
..... ۱۱۷	شکل ۱۵-۳ یک شبکه‌ی یک لایه‌ای با R عنصر ورودی و S نرون (با نشانه‌گذاری سطح بالاتر)
..... ۱۱۸	شکل ۱۶-۳ شبکه‌ی سه لایه‌ای
..... ۱۲۰ مرحله نشان می دهند
..... ۱۲۲	شکل ۱۷-۳ مثالی از فرآیند انتشار عقب. خط چین‌ها جهت مشتق یک تابع خط را در هر

چکیده:

موضوعی که در این پایان نامه مورد بررسی قرار گرفته شبکه های عصبی است، که شامل موضوعاتی از قبیل شبکه عصبی چیست، چرا از شبکه های عصبی استفاده کنیم، شامل چه مواردی است، از چه بخش هایی تشکیل شده، چه کاربردی دارد و اینکه چگونه کار می کند، می باشد. از آنجا که شبکه های عصبی نسل جدیدی از شبکه می باشد لذا دارای دنیایی از اطلاعات جدید در زمینه شبکه و علم آن است. از جمله نتایج ارزشمند این مقاله می توان به آشنایی با شبکه عصبی و طریقه کار آن، که پایه و مقدمه ای برای آشنایی هر چه بیشتر با این علم جدید می باشد.

امروزه شبکه های عصبی در کاربردهای مختلفی نظیر مسائل تشخیص الگو، مسائل دسته بندی و کنترل به کار می روند. در کنترل یا مدل سازی سیستم هایی که ساختار داخلی ناشناخته یا بسیار پیچیده ای دارند شبکه عصبی خود تابع کنترل را یاد خواهد گرفت. در حال حاضر تعداد بسیار زیادی از انواع مختلف شبکه های عصبی مصنوعی وجود دارند که به صورت خلاصه عبارتند از: شبکه های پرسپترون چند لایه، کوهونن^۱، هاپفیلد^۲... که این شبکه ها نیز خود با روش های مختلفی آموزش می بینند. از جمله کاربردهای این بحث می توان از شناسایی الگوها، پردازش تصویر و رویت، هوش مصنوعی، کنترل رباتها و موارد بسیار دیگر نام برد.

ما در اینجا پس از مقدمه به مسائل در خور شبکه های عصبی مصنوعی و نیز کاربردهای آن خواهیم پرداخت، در ادامه پرسپترون^۳ ها را که یکی از مهمترین الگوریتم های شبکه های عصبی مصنوعی می باشد معرفی می کنیم.

واژه های کلیدی: شبکه های عصبی، نرون ها، پرسپترون، پردازش چهره و تصویر

¹ kohonen

² Hopfield

³ Perceptron

فصل اول

۱- مقدمه‌ای بر شبکه‌های عصبی

۱-۱ شبکه های عصبی

شبکه های عصبی مصنوعی^۱ الگویی برای پردازش اطلاعات می باشند که با تقلید از شبکه های عصبی بیولوژیکی مثل مغز انسان ساخته شده اند. عنصر کلیدی این الگو ساختار جدید سیستم پردازش اطلاعات آن می باشد و از تعداد زیادی عناصر (نرون) با ارتباطات قوی داخلی که هماهنگ با هم برای حل مسائل مخصوص کار می کنند تشکیل شده اند. شبکه های عصبی مصنوعی با پردازش روی داده های تجربی، دانش یا قانون نهفته در ورای داده ها را به ساختار شبکه منتقل می کند که به این عمل یادگیری می گویند. اصولاً^۲ توانایی یادگیری مهمترین ویژگی یک سیستم هوشمند است. سیستمی که بتواند یاد بگیرد منعطف تر است و ساده تر برنامه ریزی میشود، بنابراین بهتر میتواند در مورد مسایل و معادلات جدید پاسخگو باشد.

شبکه های عصبی شبیه به مغز انسان اطلاعاتی را پردازش می کنند. شبکه از تعداد زیادی سلول های عصبی تشکیل شده با پردازشی بسیار بزرگ و بههم پیوسته که در حل موازی مسائل ویژه مشغول به کارند.

یادگیری شبکه های عصبی از طریق مثالهاست. آنها برای انجام یک کار خاص برنامه ریزی نشده اند. مثالها باید با دقت بسیار بالایی انتخاب شوند والا زمان مفید هدر خواهد رفت و یا حتی ممکن است شبکه به طور ناقص دایر شود و در اینجا راهی برای فهمیدن اینکه سیستم معیوب است یا خیر وجود ندارد مگر اینکه خطای رخ دهد.

¹Artificial Neural Network

²Neuron

شبکه های عصبی مصنوعی یک ترکیبی از مجموعه نرون هاست و البته نرونهای مصنوعی ای که بسیار شبیه به نرونهای زیستی کار می کنند. و بدین گونه است که ورودی های زیادی با وزنهای مختلف می گیرد و یک خروجی که به ورودی وابسته است تولید می کند. نرونهای زیستی می توانند در حال برانگیزش باشند یا نباشند(وقتی یک نرون برانگیخته می شود ضربه علائم خروجی آن مقداری کمتر از ۱۰۰ هرتز است).

شبکه های عصبی استفاده وسیعی در شناسایی الگوها دارند زیرا از خودشان قابلیت آن را دارند که بطور عمومی به ورودی های غیرمنتظره نیز پاسخ دهند. در طول ساخت نرونها می آموزند که چگونه الگوهای ویژه گوناگون را تشخیص دهند. اگر الگویی پذیرفته شود در حالی که در طول اجرا ورودی با خروجی مرتبط نباشد، نرون از مجموعه ای از الگوهایی که سابقاً آموخته خروجی را که شبیه به الگو می باشد و کمترین تفاوت را با ورودی دارد انتخاب می کند. این روال عموماً فرآخوانی می شود[1].

مثال:

وقتی که ورودی نرون ۱۱۱۱ باشد چهار ورودی بر حسب برانگیزش مرتب شده اند و وقتی ورودی های ۰۰۰۰ را داریم نرون برای برانگیزش مرتب نیست. قاعده عمومی این است که نرونها مایلند برانگیخته شوند وقتی که ورودی ها ۰۱۱۱، ۰۱۱۰، ۰۱۱۱، ۰۱۰۱ یا ۱۱۱۱ باشند و در صورتی که ورودی آنها ۰۰۰۱، ۰۰۱۰، ۰۰۱۱ یا ۱۰۰۰ باشند مایل به برانگیخته شدن نیستند.

شناسایی الگوهای پیچیده سطح بالا می تواند به وسیله شبکه ای از نرونها انجام شود و بدین ترتیب نام آن را شبکه های عصبی مصنوعی گذاشتند. اکنون شبکه های عصبی کاربردهای زیادی دارند(در منطق و کلام و شناسایی عکسها) البته شناسایی الگوهای تواند به طور موفقیت آمیز بر روی کامپیوترهای عمومی انجام شود. این شبکه های عمومی که برای شناسایی الگوها استفاده میشوند-Feed-

نامیده می شدند زیرا آنها یک بازخورد^۱ داشته و به طور ساده ورودی ها را با خروجی ها می آمیختند. اما اینشناسایی الگوها به تدریج کاملتر شده به طوری که بر روی کامپیوترهای عمومی با سیستم خاص خودشان به سختی انجام می شده پس برای شناسایی الگوها شبکه های Feed-Forward کافی نبودند.

در شبکه های عصبی خروجی هر نرون به ورودی نرونها مجاورش متصل شده است. شبکه های عصبی نمی توانند معجزه کنند اما اگر به درستی استفاده شوند نتایج شگفت انگیزی خواهند داشت.

۱-۲-۱ ساختار شبکه های عصبی:

یک شبکه عصبی شامل اجزاء سازنده لایه ها و وزنها می باشد. رفتار شبکه نیز وابسته به ارتباط بین اعضاء است. در حالت کلی در شبکه های عصبی سه نوع لایه نرونی وجود دارد:

- لایه ورودی: دریافت اطلاعات خامی که به شبکه تغذیه شده است.
- لایه های پنهان: عملکرد لایه های پنهان به وسیله ورودیها و وزن ارتباط بین آنها و لایه های پنهان تعیین می شود. وزنها بین واحدهای ورودی و پنهان تعیین می کند چه وقت یک واحد پنهان باید فعال شود
- لایه خروجی: عملکرد واحد خروجی بسته به فعالیت واحد پنهان و وزن ارتباط بین واحد پنهان و خروجی میباشد.

^۱Feed-Back

مرسومترین نوع شبکه عصبی MLP¹ یا درک چند لایه ای می باشد. این نوع از شبکه عصبی بعنوان شبکه ناظر² نیز شناخته می شود. به این علت که باید ابتدا هدف و خروجی سیستم مشخص شود. هدف از این نوع شبکه ایجاد یک مدلی است که با گرفتن ورودی، خروجی را بطور صحیح نمایش دهد با استفاده از اطلاعات پیش فرض. در نتیجه این مدل مناسبی است مخصوصا هنگامی که نتوان به طور عملی خروجی از سیستم گرفت می توان با دادن ورودی و یک سری اطلاعات پیش فرض بطور تئوری خروجی را مشخص کرد.[2]

در سازماندهی تک لایه که در آن تمام واحدها به یک لایه اتصال دارند بیشترین مورداستفاده را دارد و پتانسیل محاسباتی بیشتری نسبت به سازماندهیهای چند لایه دارد. در شبکه های چند لایه واحدها به وسیله لایه ها شماره گذاری می شوند (به جای دنبال کردن شماره گذاری سراسری).

هر دو لایه از یک شبکه به وسیله وزنها و در واقع اتصالات با هم ارتباط می یابند. در شبکه های عصبی چند نوع اتصال و یا پیوند وزنی وجود دارد:

● پیشرو: بیشترین پیوندها از این نوع است که در آن سیگنالها تنها در یک جهت حرکت می

کنند. از ورودی به خروجی هیچ بازخورده (حلقه) وجود ندارد. خروجی هر لایه بر همان

لایه تاثیری ندارد.

● پسخور: داده ها از گره های لایه بالا به گره های لایه پایین بازخورانده می شوند.

● جانبی: خروجی گره های هر لایه به عنوان ورودی گره های همان لایه استفاده می شوند.

¹multilayer perceptron
²supervised network

۱-۳-۱ کاربردهای شبکه های عصبی مصنوعی ANN

می توان موارد زیر را از کاربردهای شبکه های عصبی مصنوعی ذکر کرد:

- پردازش تصویر و دید^۱ پردازش علائم^۲: شامل رخت شناسی و تجزیه و تحلیل علائم مربوط به

زمین لرزه ها و ...

- شناسایی الگوهای^۳: شامل شناسایی چهره، اثر انگشت، تشخیص نوع صدا و نوع صحبت کردن،

دستخط و ...

- پزشکی^۴: شامل تجزیه و تحلیل و تشخیص علائم دستگاه ضربان نگار قلب الکتروکاردیوگرافیک)،

تشخیص امراض گوناگون و ...

- سیستم های نظامی^۵: شامل ردیابی مین های زیردریایی، دسته بندی صدای هنجر و مخل

در رادارها.

- سیستم های تجاری^۶: شامل تجزیه و تحلیل انبار مغازه ها، ارزیابی واقعی املاک و ...

- برنامه ریزی، کنترل و جستجو^۷): شامل اجرای موازی Planning, control, and search (

مسائل و کنترل رباتها.

هوش مصنوعی^۸: شامل برخی سیستم های طبی و اجرای سیستم های خبره.

- سیستم های قدرت^۹: شامل برآورد وضعیت سیستم، ردیابی سریع و دسته بندی ردیابی، ردیابی

خطا و ترمیم آن، پیش بینی و برآورد تخمین امنیت.

¹Image processing and computer vision

²Signal processing

³Pattern recognition

⁴Medicine

⁵Military systems

⁶Financial systems

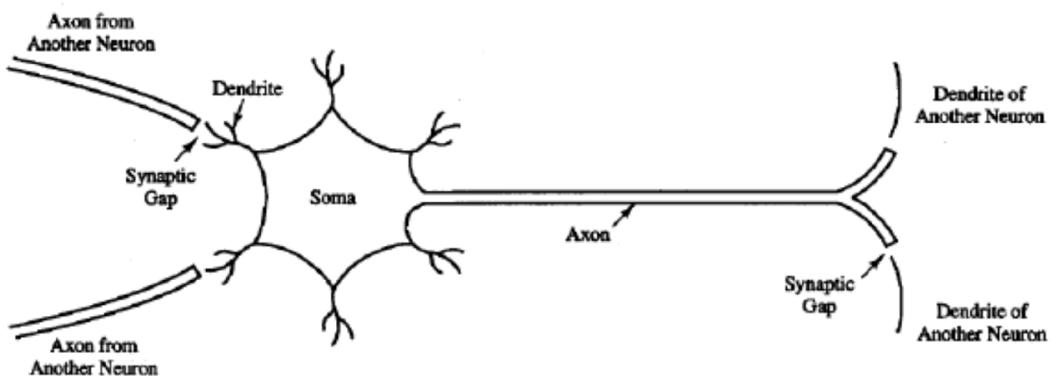
⁷Artificial intelligence

⁸Power systems

انسانها از زمانهای بسیار دور سعی بر آن داشتند که بیوفیزیولوژی مغز را دریابند چون همواره مسئله هوشمندی انسان و قابلیت یادگیری، تعمیم، خلاقیت، انعطاف پذیری و پردازش موازی در مغز برای بشر جالب بوده و بکارگیری این قابلیتها در ماشینها بسیار مطلوب می‌نمود. روش‌های الگوریتمیک برای پیاده‌سازی این خصایص در ماشینها مناسب نمی‌باشد در نتیجه می‌باشد روش‌ها مبتنی بر همان مدل‌های بیولوژیکی باشد. ANN درست مثل انسانها با استفاده از مثال‌ها آموزش می‌بیند؛ همانطور که یک بچه با دیدن انواع مختلف از یک حیوان قادر به تشخیص آن می‌باشد.

به هر حال یک شبکه عصبی مصنوعی ANN ایده‌ای است برای پردازش اطلاعات که از سیستم عصبی زیستی الهام گرفته شده و مانند مغز به پردازش اطلاعات می‌پردازد. عنصر کلیدی این ایده، ساختار جدید سیستم پردازش اطلاعات است. این سیستم از شمار زیادی عناصر پردازشی فوق العاده بهم پیوسته تشکیل شده که برای حل یک مسئله با هم هماهنگ عمل می‌کند.^[3]

به هر یک از این عناصر پردازشی فوق العاده یک نرون (neurons) گفته می‌شود.



شکل ۱-۱ مشخصات مسائل در خور شبکه‌های عصبی مصنوعی (ANN) Artificial Neural Network

تقلید از ساختارهای محاسباتی سیستم زیستی ممکن است ایده اصلی نمونه‌های محاسباتی برای ساخت کلاسهایی از مسائل باشد. از جمله این مسائل می‌توان از مسائل مشکل NP که شامل مسائل

طبقه بندی شده، زمانبندی شده و جستجو نام برد، کلاس مسائل شناسایی الگوها، افراد و موضوعات

مشخص را در دیدار و تماس با آنها می‌شناسد و کلاس مربوط به داده‌های ناقص، اشتباه، متناقض، فازی و

احتمالی این مسائل توسط همه یا برخی از موارد زیر توصیف می‌شوند:

یک فضای مسئله با بعد بزرگ، پیچیده، ناشناخته با اثرات متقابل پیچیده ریاضی وار بین متغیرها و

یک فضای راه حل که ممکن است خالی باشد(شامل یک راه حل یکتا یا بیشتر، شامل تعدادی از راه

حلهای مفید)به نظر می‌رسد ANN ها راه حل هایی برای مسائلی که با ورودی های حسی بیشتر

درگیرنده ارائه می‌دهد(صحبت کردن، دیدن، شناسایی دستخط و..)..

۱-۵ قوانین برانگیختگی:

قوانین آتش یا برانگیختگی یک مفهوم مهم برای شبکه‌های عصبی و ایجاد انعطاف پذیری در آنها

می‌باشد. قانون به تمام الگوهای ورودی مربوطی شود نه تنها الگوهای آموزش داده شده است. با استفاده

از تکنیک فاصله همینگ به یک قانون ساده برانگیختگی اشاره می‌کنیم. به ترتیب زیر:

مجموعه‌ای از الگوهای آموزشی آماده می‌شود. آنهایی که باعث آتش می‌شوند ۱ و آنهایی که مانع

شوند، را دریافت می‌کنند. الگوهایی که در مجموعه جمع آوری شده نمی‌باشند باعث برانگیختگی می

شوند اگر در مقایسه با الگوهای قبلی ۰ و ۱ عناصر مشترک بیشتری داشته باشند. اگر گره‌ای به وجود

بیاید الگو در وضعیتی تعریف نشده باقی می‌ماند. به عنوان مثال یک نرون با ۳ ورودی را در نظر بگیرید

که در حالتی که ورودیهای X1,X2,X3 در حالت ۱۰۱ یا ۱۱۱ هستند خروجی ۱ و در حالتی که ۰۰۰ یا

۰۰۱ باشند ۰ را نمایش می‌دهد. قبل از اعمال قوانین برانگیختگی جدول صحت به صورت زیر است.

X1:	0	0	0	0	1	1	1	1
X2:	0	0	1	1	0	0	1	1
X3:	0	1	0	1	0	1	0	1
OUT:	0	0	0/1	0/1	0/1	1	0/1	1

جدول ۱-۲ صحت قبل از اعمال قوانین برانگیختگی

به عنوان مثال در حالتی که قانون به کار گرفته شود ورودی ۰۱۰ را اعمال می کنیم که با ۰۰۰ در ۱ عنصر با ۰۰۱ در ۲ عنصر با ۱۰۱ در ۳ عنصر و با ۱۱۱ در ۲ اختلاف دارد در نتیجه نزدیکترین الگو ۰۰۰ می باشد که متعلق به خروجی ۰ است. در حالیکه ۱۱۱ دارای اختلاف فاصله مساوی از دو الگوی از قبل شناخته شده می باشد که هریک دارای خروجی متفاوت است در نتیجه خروجی به صورت تعريف نشده (۱/۰) باقی می ماند. با اعمال قانون به ستونها جدول صحت زیر به دست می آید.

X1:	0	0	0	0	1	1	1	1
X2:	0	0	1	1	0	0	1	1
X3:	0	1	0	1	0	1	0	1
OUT:	0	0	0	0/1	0/1	1	1	1

جدول ۱-۳ صحت بعد از اعمال قوانین برانگیختگی

تفاوت بین دو جدول را عمومی سازی(عمومیت بخشی یا تعمیم دهی) گویند. در نتیجه قانون

برانگیختگی به نرون مفهوم شبیه سازی را رسانده و آنرا قادر به پاسخ دهی محسوس به الگوهایی که در

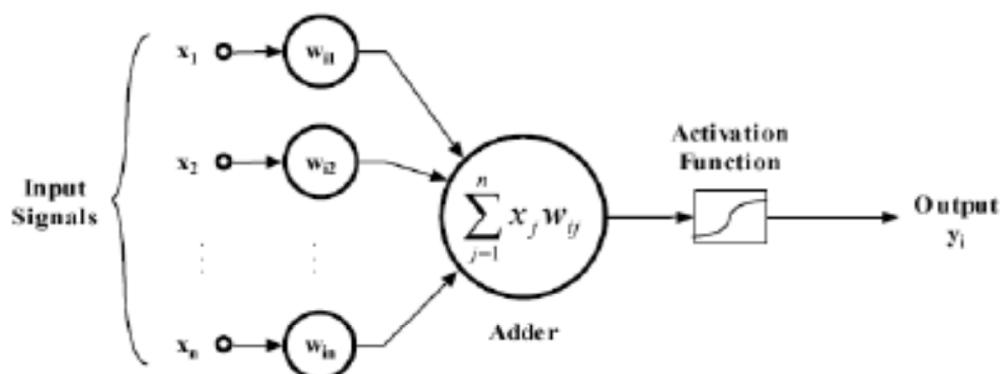
طول آموزش با آنها برخورد نداشته می کند.[4]

۱-۱- شبکه عصبی مصنوعی ساده:

فرانک روزن بلات، با اتصال این نرون‌ها به طریقی ساده پرسپترون را ایجاد و ابداع کرد، و برای

نخستین بار این مدل را در کامپیوترهای دیجیتال شبیه‌سازی و آن‌ها را به طور رسمی تحلیل نمود.

- پرسپترون برداری از ورودیها را دریافت می کند.
- ترکیب خطی این ورودی‌ها را محاسبه می کند.
- اگر حاصل از یک مقدار آستانه بیشتر بود آتش می کند. (خروجی پرسپترون برابر ۱ می شود.)



شکل ۱-۲: نحوه تشکیل ورودی‌های نرون

سیگنال های ورودی X_1 تا X_n معادل سیگنال های عصبی ورودی و وزن های W_1 تا W_n معادل مقادیر اتصالات سیناپسی ورودی های نرون می باشند که جمعاً ورودی های نرون را تشکیل داده است.

[5]

خروجی پرسپترون وسط رابطه زیر مشخص می شود:

$$y_i = ActivationFunction\left(\sum_{j=1}^n x_i w_{ij}\right)$$

$$\begin{cases} 1 \text{ if } w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n > 0 \\ -1 \text{ otherwise} \end{cases}$$

۷-۱-۱ شبکه های عصبی بر مبنای روش آموزش

شبکه های عصبی بر مبنای روش آموزش به چهار دسته تقسیم می شوند:

۱-۱-۱-۱ وزن ثابت: آموزشی در کار نیست و مقادیر وزن ها به هنگام نمی شود. و از کاربردهای آن می

توان نام برد:

• بهینه سازی اطلاعات (کاهش حجم، تفکیک پذیری و فشرده سازی)

• حافظه های تناظری

۲-۱-۱-۱ آموزش بدون سرپرست: وزن ها فقط بر اساس ورودی ها اصلاح می شوند و خروجی

مطلوب وجود ندارد تا با مقایسه خروجی شبکه با آن و تعیین مقدار خطأ وزن ها اصلاح شود. وزن ها فقط

بر اساس اطلاعات الگوهای ورودی به هنگام می‌شوند. هدف استخراج مشخصه‌های الگوهای ورودی بر اساس راهبرد خوشه یابی و یا دسته‌بندی و تشخیص شباهت‌ها (تشکیل گروه‌هایی با الگوی مشابه) می‌باشد، بدون اینکه خروجی یا کلاس‌های متناظر با الگوهای ورودی از قبل مشخص باشد. این یادگیری معمولاً بر پایه شیوه برترین هم خوانی انجام می‌گیرد. شبکه بدون سرپرست وزن‌های خود را بر پایه خروجی حاصل شده از ورودی تغییر می‌دهد تا در برخورد بعدی پاسخ مناسبی را برای این ورودی داشته باشد. در نتیجه شبکه یاد می‌گیرد چگونه به ورودی پاسخ بدهد. اصولاً هدف این است که با تکنیک نورون غالب نورونی که بیشترین تحریک آغازین را دارد برگزیده شود. بنابر این در شبکه‌های بدون سرپرست یافتن نورون غالب یکی از مهمترین کارها است.

۱-۱-۷-۳-آموزش با سرپرست: به ازای هر دسته از الگوهای ورودی خروجی‌های متناظر نیز به شبکه نشان داده می‌شود و تغییر وزن‌ها تا موقعی صورت می‌گیرد که اختلاف خروجی شبکه به ازای الگوهای آموزشی از خروجی‌های مطلوب در حد خطای قابل قبولی باشد. در این روش‌ها یا از خروجی‌ها به وزن‌ها ارتباط وجود دارد یا خلا به صورت پس انتشار از لایه خروجی به ورودی توزیع شده‌است و وزن‌ها اصلاح می‌شوند. هدف طرح شبکه‌ای است که ابتدا با استفاده از داده‌های آموزشی موجود، آموزش ببیند و سپس با ارائه بردار ورودی به شبکه که ممکن است شبکه آن را قبلاً فراگرفته یا نگرفته باشد کلاس آن را تشخیص دهد. چنین شبکه‌ای به طور گسترده برای کارهای تشخیص الگو به کار گرفته می‌شود.

۱-۱-۷-۴-آموزش تقویتی: کیفیت عملکرد سامانه به صورت گام به گام نسبت به زمان بهبود می‌یابد. الگوهای آموزشی وجود ندارد اما با استفاده از سیگنالی به نام نقاد بیانی از خوب و یا بد بودن رفتار سامانه بدست می‌آید (حالته بین یادگیری با سرپرست و بدون سرپرست).[6]

دلایل استفاده از شبکه های عصبی

شبکه‌های عصبی با توانایی قابل توجه خود در استنتاج نتایج از داده‌های پیچیده می‌توانند در استخراج الگوها و شناسایی گرایش‌های مختلفی که برای انسان‌ها و کامپیوتر شناسایی آنها بسیار دشوار استفاده شوند.

۱-۱-امزایی شبکه‌های عصبی

۱-۸-۱ یادگیری تطبیقی: توانایی یادگیری اینکه چگونه وظایف خود را بر اساس اطلاعات داده شده به آن و یا تجارت اولیه انجام دهد در واقع اصلاح شبکه را گویند.

۲-۸-۱-۱ خود سازماندهی: یک شبکه عصبی مصنوعی به صورت خودکار سازماندهی و ارائه داده‌هایی که در طول آموزش دریافت کرده را انجام دهد. نورون‌ها با قاعدهٔ یادگیری سازگار شده و پاسخ به ورودی تغییر می‌یابد.

۳-۸-۱-۱ عملگرهای بی‌درنگ: محاسبات در شبکه عصبی مصنوعی می‌تواند به صورت موازی و به وسیله ساخت افزارهای مخصوصی که طراحی و ساخت آن برای دریافت نتایج بهینه قابلیت‌های شبکه عصبی مصنوعی است انجام شود.

۱-۱-۴-۴ تحمل خطاباً: با ایجاد خرابی در شبکه مقداری از کارایی کاهش می‌یابد ولی برخی امکانات آن با وجود مشکلات بزرگ همچنان حفظ می‌شود.

۱-۱-۸-۵ دسته بندی شبکه‌های عصبی قادر به دسته بندی ورودی‌ها برای دریافت خروجی مناسب می‌باشند.

۱-۱-۶-۶ تعمیم دهی: این خاصیت شبکه را قادر می‌سازد تا تنها با برخورد با تعداد محدودی نمونه،

یک قانون کلی از آن را به دست آورده، نتایج این آموخته‌ها را به موارد مشاهده از قبل نیز تعمیم دهد.

توانایی که در صورت نبود آن سامانه باید بی نهایت واقعیت‌ها و روابط را به خاطر بسپارد.

۱-۱-۷-۱ پایداری - انعطاف پذیری: یک شبکه عصبی هم به حد کافی پایدار است تا اطلاعات

فراگرفته خود را حفظ کند و هم قابلیت انعطاف و تطبیق را دارد و بدون از دست دادن اطلاعات قبلی

می‌تواند موارد جدید را بپذیرد [7].

۱-۱-۹-۱ کاربرد شبکه‌های عصبی:

شبکه‌های عصبی مصنوعی دارای دامنه کاربرد وسیعی می‌باشند از جمله سامانه‌های آنالیز ریسک،

کترل هوایپما بدون خلبان، آنالیز کیفیت جوشکاری، آنالیز کامپیوتر، آزمایش اتاق اورژانس،

اکتشاف روغن و گاز، سامانه‌های تشخیص ترمز کامیون، تخمین ریسک وام، شناسایی طیفی، تشخیص

دارو، فرآیندهای کترول صنعتی، مدیریت خطاب، تشخیص صدا، تشخیص هپاتیت، بازیابی اطلاعات راه دور،

شناسایی مین‌های زیردریایی، تشخیص اشیاء سه بعدی و دست نوشته‌ها و چهره و... در کل می‌توان

کاربردهای شبکه‌های عصبی را به صورت زیر دسته بندی کرد: تناظر (شبکه الگوهای مغشوش و به هم

ریخته را بازشناسی می‌کند)، خوش‌یابی، دسته بندی، شناسایی، بازسازی الگو، تعمیم دهی (به دست

آوردن یک پاسخ صحیح برای محرك ورودی که قبلاً به شبکه آموزش داده نشده)، بهینه سازی. امروزه

شبکه‌های عصبی در کاربردهای مختلفی نظیر مسائل تشخیص الگو که خود شامل مسائلی مانند تشخیص

خط، شناسایی گفتار، پردازش تصویر و مسائلی از این دست می‌شود و نیز مسائل دسته بندی مانند دسته

بندي متون یا تصاویر، به کار می‌روند. در کترول یا مدل سازی سامانه‌هایی که ساختار داخلی ناشناخته یا

بسیار پیچیده‌ای دارند نیز به صورت روز افزون از شبکه‌های عصبی مصنوعی استفاده می‌شود. به عنوان مثال می‌توان در کنترل ورودی یک موتور از یک شبکه عصبی استفاده نمود که در این صورت شبکه عصبی خود تابع کنترل را یاد خواهد گرفت[8].

۱-۱-۱ معايب ANNها

دستورات مشخصی برای طراحی شبکه جهت یک کاربرد اختاری وجود ندارد.

- دقیق نتایج بستگی زیادی به اندازه مجموعه آموزش دارد.
- آموزش شبکه ممکن است مشکل یا حتی غیرممکن باشد.
- پیش‌بینی عملکرد آینده شبکه (عمومیت یافتن) آن به سادگی امکان‌پذیر نیست.[9]

۲-۱ نورون مصنوعی:

یک نورون مصنوعی سامانه‌ای است با تعداد زیادی ورودی و تنها یک خروجی. نورون دارای دو حالت می‌باشد، حالت آموزش و حالت عملکرد. در حالت آموزش نورون یاد می‌گیرد که در مقابل الگوهای ورودی خاص برآنگیخته شود و یا در اصطلاح آتش کند. در حالت عملکرد وقتی یک الگوی ورودی شناسایی شده وارد شود، خروجی متناظر با آن ارائه می‌شود. اگر ورودی جزء ورودی‌های از پیش شناسایی شده نباشد، قوانین آتش برای برآنگیختگی یا عدم آن تصمیم گیری می‌کند.

هر نورون طبیعی از سه قسمت اصلی تشکیل شده است:

▪ بدن سلول^۱

▪ دندrite^۲

¹some

²dendrite

■ اکسون¹

دندریت ها به عنوان مناطق دریافت سیگنال های الکتریکی، شبکه های تشکیل یافته از فیبرهای سلولی هستند که دارای سطح نامنظم و شاخه های انشعابی بی شمار می باشند. دندریت ها سیگنال های الکتریکی را به هسته سلول منتقل می کنند. بدنه سلول، انرژی لازم را برای فعالیت نرون فراهم کرده و بر روی سیگنال های دریافتی عمل می کنند، که با یک عمل ساده جمع و مقایسه با یک سطح آستانه مدل می گردد.

اکسون برخلاف دندریت ها از سطحی هموارتر و تعداد شاخه های کمتری برخوردار می باشد. اکسون طول بیشتری دارد و سیگنال های الکتروشیمیایی دریافتی از هسته سلول را به نرون های دیگر منتقل می کند. محل تلاقي یک اکسون از یک سلول به دندریت های سلول های دیگر را سیناپس می گویند. توسط سیناپس ها ارتباطات ما بين نرون ها برقرار می شود. به فضای مابین اکسون و دندریت ها فضای سیناپسی گویند. در حقیقت دندریت ها به عنوان ورودی نرون و اکسون به عنوان خروجی و فضای سیناپسی محل اتصال اين دو می باشد. زمانی که سیگنال عصبی از اکسون به نرون ها و دیگر عناصر بدن مثل ماهیچه ها می رسد باعث تحریک آنها می شود. نرون ها از هر یک از اتصالات ورودی خود یک ولتاژ کم دریافت می کند(توضیح سیگنال های عصبی ورودی) و آن ها را با هم جمع می زند اگر این حاصل جمع به یک مقدار آستانه رسید نرون آتش می کند و روی اکسن خود یک ولتاژ خروجی ارسال می کند که این ولتاژ به دندریت هایی که به این اکسون متصل اند رسیده و باعث یکسری فعل و انفعالات شیمیایی در اتصالات سیناپسی می شود و می تواند باعث آتش کردن نرون های دیگر شود. تمامی فعالیت های مغزی انسان توسط همین آتش کردن ها انجام می شود. حافظه کوتاه مدت انسان جرقه های لحظه

¹axon

ای الکتریکی می باشد و حافظه بلند مدت به صورت تغییرات الکتروشیمیایی در اتصالات سیناپسی ذخیره می شود که عموما منجر به تغییر یون ها می شود.

گمان میرود که مغز انسان از تعداد 10^{11} نرون تشکیل شده باشد که هر نرون با تقریبا 10^4 نرون دیگر در ارتباط است.

سرعت سوئیچنگ نرونها در حدود 10^{-3} ثانیه است و آدمی قادر است در $1 \cdot 10^4$ ثانیه تصویر یک انسان را باز شناسائی نماید. اینقدر فوق العاده باید از پردازش موازی توزیع شده در تعدادی زیادی از نرونها حاصل شده باشد [10].

۱-۳ نرون پیچیده تر:

ساختار نرون قبلی کاری جز آنچه کامپیوترهای سنتی انجام می دهند نمی کرد. مدل مک کلوک و پیتر (MCP) نرونی با سطح بالاتر است که اختلاف آن با مدل قبلی در ورودیهای وزن دار است. تاثیری که هر ورودی در تصمیم گیری دارد وابسته به وزن ورودی مخصوص به آن می باشد. وزن یک ورودی عددی است که وقتی در ورودی ضرب شود یک ورودی با وزن را تشکیل دهد. این ورودیهای با وزن با هم جمع می شوند. اگر حاصل جمع از ارزش آستانه ای تجاوز کند نرون آتش می کند یا در اصطلاح برانگیخته می شود.

در اصطلاح ریاضی، نرون برانگیخته می شود اگر و تنها اگر:

$$X_1W_1 + X_2W_2 + \dots > T$$

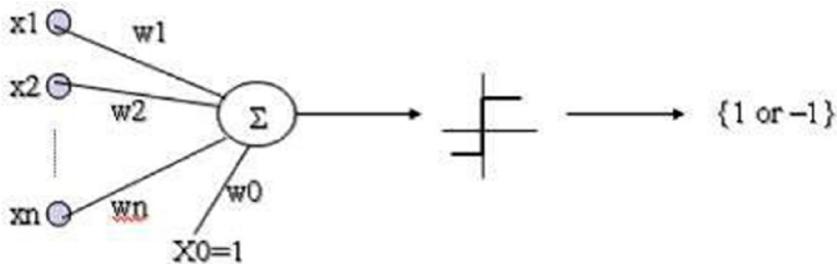
این روش باعث می شود نرون قدرتمند و انعطاف پذیر باشد. نرون MCP توانایی سازگاری با وضعیت خاص به وسیله‌ی تغییر وزنها و یا حد آستانه را دارد. الگوریتمهای متنوعی موجود می باشند که باعث تطبیق نرون می شوند، بیشترین استفاده از قوانین Delta و Back Error Propagation است. اولی در شبکه‌های پیشرو و دیگری در شبکه‌های پسخور استفاده می شود.[11]

۱-۴- نرون ساده:

یک نرون مصنوعی سیستمی است با تعداد زیادی ورودی و تنها یک خروجی. نرون دارای دو مد می باشد، مد آموزش و مد عملکرد. در مد آموزش نرون یاد می گیرد که در مقابل الگوهای ورودی خاص برانگیخته شود و یا در اصطلاح آتش کند. در مد عملکرد وقتی یک الگوی ورودی شناسایی شده وارد شود خروجی متناظر با آن ارائه می شود. اگر ورودی جزء ورودیهای از پیش شناسایی شده نباشد قوانین آتش برای برانگیختگی یا عدم آن تصمیم گیری می کند.[12]

۱-۵- پرسپترون

به طور خلاصه می توان گفت که یک پرسپترون یک برداری از ورودی ها با ورودی های حقیقی را گرفته و یک ترکیب خطی از این ورودی ها را محاسبه می کند. اگر حاصل از یک مقدار آستانه بیشتر بود خروجی پرسپترون برابر ۱ و در غیر این صورت معادل ۱ - خواهد بود.[13]



شکل ۱-۳-۱ نحوه عملکرد پسپترون

۱-۵-۱ پرسپترون های ساده

یک خانواده ساده از شبکه های عصبی مدل پرسپترونی باشد. در یک دسته بندی تک خروجی، تعداد n ورودی و یک خروجی دارد. با هر ورودی یک ضریب وزنی W_i و با هر خروجی یک مقدار آستانه q مرتبط است.

پرسپترون های کوچکتر باشد خروجی می کنند:

ورودی های پرسپترون یک بردار ورودی از n مقدار حقیقی است.

پرسپترون مجموع وزنها را محاسبه می کند $a = \sum_{i=1}^n W_i X_i$. این مقدار با مقدار آستانه q مقایسه می شود، اگر این مقدار از مقدار آستانه کوچکتر باشد خروجی ۰ است و در غیر این صورت ۱ است [14].

۱-۵-۲ قدرت پرسپترون

به وسیله تنظیم اعداد ورودی، وزن آنها و مقدار آستانه می توان یک پرسپترون برای انجام نسبتاً خوب محاسبات گوناگون طراحی کرد. برای مثال تابع منطقی بولین مانند AND، OR و NOT را می توان به وسیله پرسپترون طراحی کرد و هر مدار منطقی دیگر را به وسیله گیتهای AND و NOT یا AND و OR وسیله های پرسپترون نظری طراحی کرد.

طراحی کرد. دسته های زیادی از پرسپترونها ممکن است خروجی های دسته های دیگر را به عنوان ورودی خود درخواست کنند.

به عنوان مثالی از پرسپترونها می توان یک تشخیص دهنده قالب متن را نام برد. حرف A در آرایه ای $q = e^{-25}$ به رمز درمی‌آید. این متن (حرف) به وسیله یک پرسپترون با ۲۵ ورودی تشخیص داده می شود که در آن وزنها مقادیری برابر با مقادیر عددی داخل آرایه را می گیرند و مقدار آستانه برابر است با $e^{-1} < q$: که در آن e^{-1} خروجی پرسپترون است. اگر و فقط اگر ورودیان از ۱ و -۱- های بیاورد که عیناً در آرایه آمده است [15].

۱-۵-۳- دنباله های پرسپترون

یکی از خصوصیات جالپرسپترون هایین است که آنها می توانند به وسیله مثالهای مثبت و منفی (صحیح و اشتباه) برای انجام توابع دسته‌بندی شده مخصوص بارها مرتب شوند.

حال به یک مثال ساده از پرسپترون دو ورودی X_1 و X_2 ، که تشخیص می دهد که کدامیک از دو کلاس، عناصر متعلق به خودش را دارد. ما فرض می کیم که این پرسپترون دو طرح از کارکترهای چاپ شده از یک متن را بررسی کند، خروجی ۱ است اگر و فقط اگر کارکتر رقم ۸ باشد. فرض کنیم که X_1 بیانگر تعداد حفره های کارکتر است و X_2 درجه راستی سمت چپ کارکتر را نشان می دهد.

اگر پرسپترون در اول کار با وزنهایی برابر ۰ و مقدار آستانه را برابر ۱۰ مقداردهی شود یک رده بنده از همه مثالهای منفی انجام داده می شود. با قرار دادن رده بنده های نادرست از ۸ مقادیر ورودی از مثال ۸ با بعضی فاکتورها مثل d جمع می شوند و تولیدات جدید با وزنهای متناظر ایجاد می شوند.

فرض کنیم $a = d$ پس وزن ورودی ها از ۰ به ۱ و ۲ رشد پیدا می کند. حال در اینجا $a = b$ دست می

آید که هنوز از مقدار آستانه ۱۰ کوچکتر است. مثال هنوز به رده بندی صحیحی نرسیده است و این قدم

دنباله باید تکرار شود. بعد از دو قدم وزنها برابر ۲ و ۴ میشوند که مقدار $a = 10$ را نتیجه میدهد که برابر

مقدار آستانه است و مثال مثبت از ۸ به طور صحیح دسته بندی شده است. از آنجا که ضرایب وزنی تغییر

کرده بودند لازم است که در همه مثالها رده بندی ها بازنگشان^۱ شوند. این را می توان به سادگی دید که مثال

B رده بندی نادرستی است زیرا با وزنها ۲ و ۴ داریم $a = 24$ ولی این حرف مورد نظر ما نیست، چون

این مرحله را پیش رفته ایم لازم است که d.1 از W1 و d.2 از W2 کم شود تا رده بندی نادرستی از B

ثابت شود. به هر حال یک رده بندی از ۸ را دوباره بیرون میدهد.

بعدها موقع بروز خطای وزن ها را برای درست کردن خطاهای رده بندی اصلاح می شوند. اگر مثالها

دارای خاصیت صحیحی باشند وزنها در مجموعه ای از مقادیری که به درستی روی هر ورودی کار می

کنند قرار می گیرند[16].

۴-۵-۱ شبکه های پرسپترون چند لایه:

شبکه هایی که از یک نرون ساخته شده اند دارای محدودیت هایی می باشند. این شبکه ها توانایی

پیاده سازی توابع غیر خطی را ندارند. به عنوان مثال توسط این شبکه ها نمی توان تابع XOR را پیاده

سازی نمود. برای حل این مشکل شبکه های دیگری پیشنهاد شد که ما به یکی از آنان که بیشترین استفاده

را داردمی پردازیم. مدل عمومی شبکه های پرسپترون، شبکه جلوروند با روال تعلیم انتشار به عقب است.

شبکه های جلوروند، شبکه هایی هستند که ورودی های لایه اول نرون های آن به لایه های بعدی متصل

¹Reset

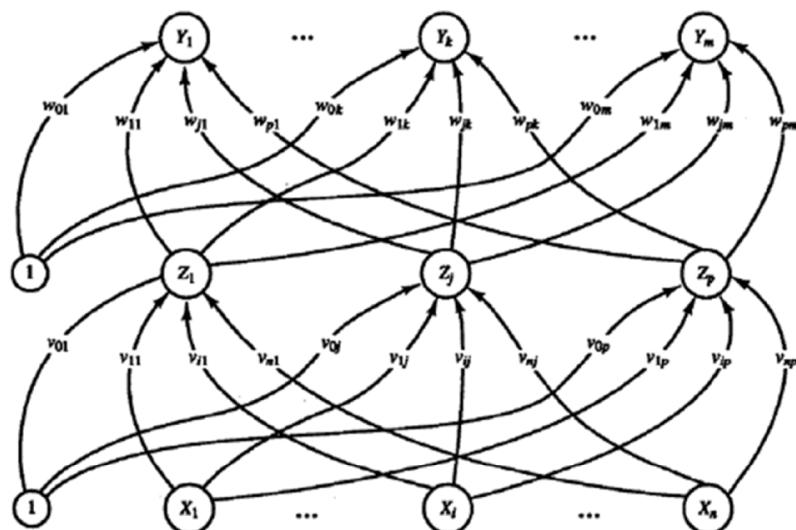
بوده و در هر سطح این مسئله صادق بوده تا به لایه خروجی برسد. روال انتشار به عقب بدین معنی است

که پس از مشخص شدن خروجی شبکه، ابتدا وزن های لایه آخر تصحیح شده و بعد به ترتیب لایه های

قبلی تصحیح می شوند.

شبکه های پرسپترون از یک لایه ورودی، تعدادی لایه پنهان و یک لایه خروجی تشکیل شده است.

در شکل زیر یک شبکه پرسپترون با یک لایه پنهان نشان داده شده است.



شکل ۱-۴ شبکه پرسپترون با یک لایه پنهان

در این شکل لایه X لایه ورودی، لایه Z لایه پنهان و لایه Y لایه خروجی می باشد. در این شبکه ها

شرایط زیر وجود دارد:

نرون های هر لایه تنها به نرون های لایه بعدی متصل می باشد.

هر نرون به تمامی نرون های لایه بعد متصل است.

نرون های لایه ورودی عملی را انجام نمی دهند و اوزان ها ثابت و برابر ۱ می باشد. این نرون ها فاقد تابع فشرده سازی می باشند.

در شبکه های پرسپترون چند لایه، تعداد لایه های پنهان می تواند هر تعداد باشد. البته در بیشتر کاربردها یک لایه پنهان کفايت می کند. در بعضی مواقع نیز دو لایه پنهان یادگیری شبکه را ساده تر می کند. [17]

برای مدلسازی شبکه های عصبی مصنوعی ابزارهای زیادی وجود دارد از این جمله:

- SNN(Stuttgart Neural Network Simulator)
- Matlab Standard NN Toolbox
- Netlab Toolbox For Matlab

فصل دوم

۱-۱ مقدمه ای بر پردازش تصویر

دیجیتال دانش جدیدی است که سابقه آن به پس از اختراع رایانه های دیجیتال 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 رد پای پردازش تصویر را در بسیاری از علوم و صنایع مشاهده نمود. علاقه به روش های پردازش تصویر از دو محدوده کاربردی اصلی نشات می گیرد که آن محدوده ها عبارتند از بهبود اطلاعات تصویری به منظور تعبیر انسانی و پردازش داده های صحنه برای ادراک ماشینی مستقل است.

۱-۱-۱ آشنایی با مفاهیم اولیه در پردازش تصویر

پردازش تصویر

به مجموعه عملیات و پردازش هایی که در راستای آنالیز تصویر در زمینه های مختلف انجام می شود، علم پردازش تصویر گفته میشود. پردازش تصویر می تواند کیفیت تصویر را کاهش یا افزایش دهد.

۲-۱ پردازش تصاویر^۱:

امروزه بیشتر به موضوع پردازش تصویر دیجیتال گفته می شود که شاخه ای از دانش رایانه است که با پردازش سیگنال دیجیتال که نماینده تصاویر برداشته شده با دوربین دیجیتال یا پویش شده توسط پویشگر هستند سر و کار دارد.

^۱Image processing

پردازش تصاویر دارای دو شاخه عمدۀ بهبود تصاویر و بینایی ماشین است. بهبود تصاویر دربرگیرنده روش‌هایی چون استفاده از فیلتر محوکننده و افزایش تضاد برای بهتر کردن کیفیت دیداری تصاویر و اطمینان از نمایش درست آنها در محیط مقصد(مانند چاپگر یا نمایشگر رایانه) است، در حالی که بینایی ماشین به روش‌هایی می‌پردازد که به کمک آنها می‌توان معنی و محتوای تصاویر را درک کرد تا از آنها در کارهایی چون رباتیک و محور تصاویر استفاده شود.

در معنای خاص آن پردازش تصویر عبارتست از هر نوع پردازش سیگنال که ورودی یک تصویر است مثل عکس یا صحنه‌ای از یک فیلم. خروجی پرداششگر تصویر میتواند یک تصویر یا یک مجموعه از نشانه‌ای ویژه یا متغیرهای مربوط به تصویر باشد. اغلب تکنیک‌های پردازش تصویر شامل برخورد با تصویر به عنوان یک سیگنال دو بعدی و بکاربستن تکنیک‌های استاندارد پردازش سیگنال روی آنها می‌شود. پردازش تصویر اغلب به پردازش دیجیتالی تصویر اشاره می‌کند ولی پردازش نوری و آنالوگ تصویر هم وجود دارد. علم پردازش تصویر در چند دهه اخیر از هر دو جنبه نظری و عملی پیشرفت‌های چشمگیری داشته است. سرعت این پیشرفت به اندازه‌ای بوده است که هم اکنون، به راحتی می‌توان رد پای پردازش تصویر 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 نمی‌باشند. اگر چه ذکر تمام جزئیات کاربردهای پردازش تصویر در یک مقاله امکان پذیر نمی‌باشد ولی سعی شده است که به طور کلی اکثر زمینه‌های کاربرد آن بیان شود. در این مقاله چهارده زمینه‌ی مختلف کاربرد پردازش تصویر بیان شده است که عبارتند از: صنعت، پزشکی، علوم نظامی و امنیتی، زمین‌شناسی، فضانوردی و نجوم، شهرسازی، هنر و سینما، فناوری‌های علمی، سیاست و روانشناسی، کشاورزی، هواشناسی، باستان‌شناسی، اقتصاد و تبلیغات[18].

۳-۱-۲ کاربردهای پردازش تصویر:

- کاربردهای صنعتی مانند کنترل کیفیت بسته بندی دارو در یک شرکت داروسازی.
- کاربردهای امنیتی مانند تشخیص حرکت، تشخیص اثر انگشت، تشخیص چهره، تشخیص دست خط یا امضا.
- کاربردهای پزشکی مانند تلقای ویژگی های تصاویر اشعه X، تولید تصاویر MRI از مغز و یا تصاویر CTScan مربوط به زمین.
- کاربردهای نظامی مانند تشخیص و هدف یابی خودکار اهداف متحرک یا ثابت توسط موشک های هوایی به زمین.

کاربردهای پردازش تصویر در حمل و نقل

تشخیص شماره پلاک خودرو و نرم افزار شمارش خودروهای عبوری از عرض خیابانی شک یکی از مؤثر ترین مولفه ها در مدیریت و برنامه ریزی دسترسی به آمار دقیق می باشد. در صورت وجود آمار دقیق و سریع می توان از روش های کنترل بهینه استفاده کرد و بهره وری را افزایش داد. به عنوان مثال اگر آمار دقیقی از میزان مصرف یک محصول غذایی وجود داشته باشد با برنامه ریزی مناسب می توان زمینه تولید و عرضه اصولی آن را فراهم کرد. لذا احتمال نابسامانی در بازار و متضرر شدن کشاورز و مصرف کننده کاهش می یابد. چنان که بیان شد مهمترین فاکتور در برنامه ریزی دسترسی به آمار مناسب است اما تهیه آمار فرایند پیچیده و وقت گیر است و معمولا هزینه زیادی را در بر دارد. به عنوان مثال به دلایلی از جمله کنترل ترافیک یا کنترل میزان روشنایی خیابان باید خودروهای عبوری از خیابان شمارش شوند. این کار اگر به صورت دستی یا انسانی انجام شود، هزینه زیادی نیاز دارد، امکان سهل انگاری انسانی نیز وجود

دارد پس استفاده از یک دستگاه مناسب که توانایی شمارش خودروهای عبوری را داشته باشد تنها گزینه ممکن است. با توجه به نیاز فوق نرم افزاری تهیه شده است که با استفاده از تصاویر گرفته شده از عرض خیابان خودروهای عبوری را تشخیص می‌دهد و تعداد آنها را شمارش می‌کند. این نرم افزار امکان استفاده در روز یا شب را دارا می‌باشد.^[19]

۴-۱-۲ مقدمه پردازش تصویر:

چشم به عنوان یکی از حیاتی ترین حسگرهای انسان نقش بسزایی در زندگی ما دارد. امروزه با پیشرفت چشمگیری که در ساخت پردازنده‌ها صورت گرفته است، این امکان نیز برای ما فراهم شده تا در ساخت روبات‌ها و سیستم‌های کنترلی از دوربین به عنوان یک چشم مصنوعی استفاده کنیم.

پنج کاربرد عمده پردازش تصویر را می‌توان رباتیک، سیستم‌های دفاعی، مهندسی پزشکی، کنترل صنعتی و گرافیک کامپیوترا عنوان کرد. در سیستم‌های رباتیک معمولاً از پردازش تصویر برای هدایت ربات و تشخیص اشیا استفاده می‌شود. در سیستم‌های دفاعی برای یافتن هدف و یا رهگیری یک هدف متحرک پردازش تصویر یکی از قابل اعتمادترین روش‌های موجود می‌باشد. در مورد کاربردهای گرافیکی نیز یکی از معروف‌ترین نرم افزارهای مبتنی بر پردازش تصویر فتوشاپ می‌باشد که همگی با کاربردهای این نرم افزار آشنا هستیم. تشخیص نوع بیماری نیز رایج ترین کاربرد پردازش تصاویر پزشکی است. در نهایت، امروزه خطوط تولید صنعتی برای کنترل کیفیت محصولات تولید شده و همچنین کنترل حرکات خط تولید از سیستم‌های مبتنی بر پردازش تصویر بهره جسته‌اند.

یک تصویر از لحظه ورود به سیستم پردازش تصویرتا تولید تصویر خروجی، به ترتیب مراحل زیر را طی می‌کند:

♦ در یافت تصویر ورودی:

در این مرحله تصویر از ورودی خوانده شده و وارد سیستم می‌گردد. تصویر ورودی می‌تواند بر روی ابزار ذخیره سازی بوده و یا از یک دوربین گرفته شود. به عنوان مثال در سیستم تشخیص اثر انگشت تصویر ورودی از طریق اسکنر اثر انگشت وارد سیستم می‌شود.

♦ پیش پردازش تصویر:

اهداف کلی این مرحله را می‌توان ارتقاء تصویر و حذف مولفه‌های غیر ضروری از تصویر دانست. به عنوان مثال خشکی پوست، جراحت و یا عدم تمیزی پوست مواردی هستند که قبل از پردازش تصویر انگشت باید با استفاده از تکنیک‌های پردازش تصویر به رفع آنها پرداخت.

♦ پردازش تصویر:

هدف کلی در این مرحله شناسایی ویژگی‌هایی از تصویر است که بتوان از آنها برای کاربر مورد نظر خود استفاده کرد. شناسایی نقاط ویژه و انحناها از جمله ویژگی‌هایی هستند که در سیستم تشخیص اثر انگشت می‌توانند استخراج گردد.

♦ آنالیز تصویر:

در این مرحله با استفاده از ویژگی‌های استخراج شده به آنالیز تصویر می‌پردازیم. به عنوان مثال پس از شناسایی نقاط ویژه و انحناها در سیستم تشخیص اثر انگشت با آنالیز کردن تصویر سعی می‌کنیم شخص متناظر با اثر انگشت را پیدا کنیم. آنالیز تصویر معمولاً از تکنیک‌های هوش مصنوعی همانند شبکه های عصبی، درخت‌های تصمیمی، کلاس‌بندی و... استفاده می‌کند.^[20]

۲-۱-۵ تصویر دیجیتالی

تصویر دیجیتالی

یک تصویر را می‌توان توسطتابع دوبعدی $f(x,y)$ نشان داد که در آن X و Y رامختصات مکانی و مقدار f در هر نقطه را شدت روشنایی تصویر در آن نقطه می‌نامند. اصطلاح سطح خاکستری نیز به شدت روشنایی تصاویر مونوکروم اطلاق می‌شود. تصاویر رنگی نیز از تعدادی تصویر دوبعدی تشکیل می‌شود. زمانی که مقادیر X و Y و مقدار $f(x,y)$ با مقادیر گسسته و محدود بیان شوند، تصویر را یک تصویر دیجیتالی می‌نامند. دیجیتال کردن مقادیر X و Y را Sampling و دیجیتال کردن مقدار $f(x,y)$ را quantization گویند. برای نمایش یک تصویر $N \times M$ از یک آرایه دو بعدی (ماتریس) که سطر و N ستون دارد استفاده می‌کنیم.

مقدار هر عنصر از آرایه نشان دهنده شدت روشنایی تصویر در آن نقطه است. هر عنصر آرایه یک مقدار ۸ بیتی است که می‌تواند مقداری بین ۰ و ۲۵۵ داشته باشد. مقدار صفر نشان دهنده رنگ تیره (سیاه) و مقدار ۲۵۵ نشان دهنده رنگ روشن (سفید) است.

به عنوان مثال تصویری که سایز آن 288×265 است از یک ماتریس که دارای ۲۸۸ سطر و ۲۶۵ ستون است برای نمایش تصویر استفاده می‌کند. هر پیکسل از این تصویر نیز مقداری بین ۰ و ۲۵۵ دارد. نقاط روشن مقادیری نزدیک به ۲۵۵ و نقاط تیره مقادیر نزدیک به ۰ دارد. همه توابع پردازش تصویر از این مقادیر استفاده کرده و اعمال لازم را بر روی تصویر انجام می‌دهند.[21]

۶-۱-۲ مفاهیم اولیه در پردازش تصویر:

۶-۱-۱ مفهوم پیکسل در پردازش تصویر:

پیکسل^۱ که بر گرفته شده از عناصر تصویر است نقطه های سیار ریز مربع شکلی هستند که از تجمع آنها، تصویر روی صفحه نمایش یا روی کاغذ توسط چاپگر شکل می گیرد. همان طور که بیت، کوچک ترین واحد اطلاعات قابل پردازش توسط کامپیوتر است، پیکسل نیز کوچک ترین عنصر سخت افزار و نرم افزار نمایشی یا چاپی است که برای شکل گرفتن تصاویر مورد استفاده قرار می گیرد. اگر برای هر پیکسل دو رنگ (معمولًا سیاه و سفید) در نظر گرفته شود، توسط یک بیت قابل کد دهی است و در صورتیکه بیش از دو بیت برای ارائه یک پیکسل استفاده شود، محدوده رنگ ها یا سایه های خاکستری وسیع تری، قابل ارائه خواهد شد.^[22]

۶-۱-۲ مفهوم عمق بیتی:

مقدار هر پیکسل در یک یا چند بیت اطلاعات، ذخیره می شود. برای تصاویر ساده تک رنگ، یک بیت برای نشان دادن هر پیکسل کافیست اما در تصاویر رنگی و سایه های خاکستری، هر پیکسل نیاز به بیش از یک بیت اطلاعات دارد.

هر چه از بیتهاي بيشتری برای نشان دادن یک پیکسل استفاده شود رنگ ها و سایه های خاکستری

¹ Pixel

² Picture Elements

بیشترین می توان نشان داد. غلظت نقاط یا همان وضوح تصویر^۱ را تعیین می کند. این ویژگی با واحد نقطه در اینچ (dpi) و یا با تعداد ردیف ها و ستون ها سنجیده می شود (مانند 480×640).

۳-۶-۱-۲ مفهوم بعد یک تصویر:

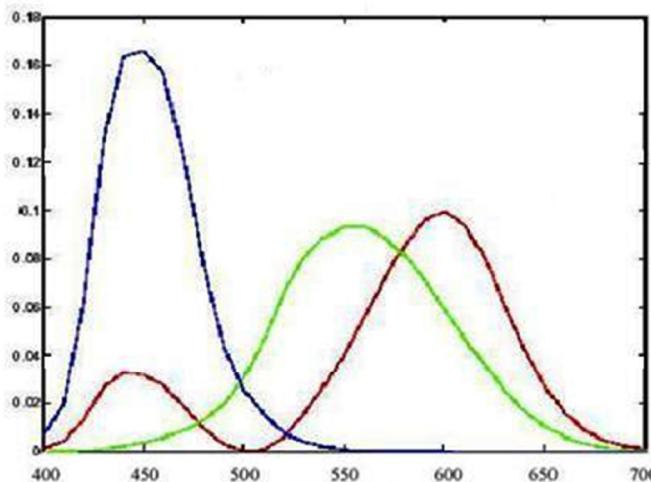
تصویر های مبتنی بر نقشه بیتی، همواره به صورت شبکه های مرتع شکلبرگ می باشد. این شبکه ها مانند صفحه شطرنج یا موزاییک های کف آشپزخانه می باشد. این شبکه های مرتع شکل بزرگ از مرتع های کوچکتری تشکیل شده اند. یکی از مشخصه هایی که همواره می توان در مورد شبکه ها بیان نمود، این است که دارای ابعاد می باشند. صفحه شطرنج همواره 8×8 است، اما شبکه پیکسل های تشکیل دهنده صفحه نمایش کامپیوتر مثلا 800×600 است.

ابعاد شبکه در واقع معادل تعداد مرتع هایی است که طول و عرض تصویر را تشکیل داده اند و ربطی به اندازه واقعی تصویر ندارند.[23]

۴-۶-۱-۲ چگونگی تشکیل رنگ در چشم انسان:

منحنی حساسیت چشم برای مشاهده کننده استاندارد در مقابل دریافت نور رنگی به صورت شکل زیرمی باشد:

¹ Resolution

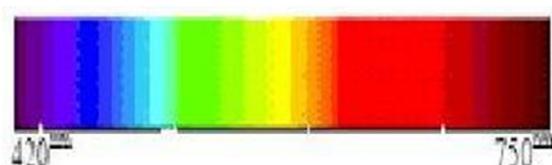


شکل ۱-۲ منحنی حساسیت چشم برای مشاهده کننده استاندارد در مقابل دریافت نور رنگی

در واقع جهت مشاهده و درک هر رنگ، سه انرژی جداگانه مربوط به سه رنگ اصلی دریافت شده و از ترکیب آنها با یکدیگر رنگ تصویر ایجاد میشود. هر سیستم گیرنده، منحنی حساسیت رنگی مختص به خود را دارد لذا ترکیب سه انرژی رنگ قرمز، $G(x,y)$ انرژی رنگ سبز و $B(x,y)$ انرژی رنگ آبی باعث ایجاد رنگ می شود.

بدین ترتیب جهت مشخص نمودن هر تصویر رنگی باید از سه ماتریس جهت مقادیر قرمز، سبز و آبی برای هر پیکسل استفاده کنیم. رنگ هر پیکسل توسط ترکیب سه رنگ اصلی در سه ماتریس داده شده به دست می آید.[24]

طیف رنگ ها را نسبت به طولموج می توان در شکل زیر مشاهده کرد:



شکل ۲-۲ طیف رنگ ها نسبت به طول موج

۲-۲ پردازش تصاویر رنگی:

استفاده از رنگ در پردازش تصاویر دوانگیزه اصلی دارد:

دلیل اول تحلیل خودکار تصویررنگی، توصیف گر توانایی است که در اغلب موارد شناسایی و استخراج شئ از صحنه را ساده تر میکند.

دلیل دومدر مواردی که تحلیل تصویر به وسیله انسان انجام می شود چشم انسان قادر است هزاران سایه و شدت رنگ را در مقایسه با حدود ۲۴ سایه خاکستری تشخیص دهد.

پردازش تصویر رنگی در حوزه اصلی به دو دسته تقسیم می شود:

پردازش تمام رنگی و شبه رنگی که در گروه اول تصاویر موردنظر معمولاً با یک Sensor تمام رنگی نظیر دوربین تلویزیون رنگی با پیمایش گر رنگی^۱ برداشته می شود در گروه دوم به هر شدت تک رنگ خاص یا محدوده ای از شدت ها یک سایه رنگی متنسب می شود.

۳-۲ انواع مدل رنگ:

هدف از انتخاب مدل رنگ تسهیل مشخص سازی رنگ ها در یک استاندارد است که معمولاً روش مورد قبولی می باشد در اصل مدل رنگ تعیین یک سامانه مختصات سه بعدیو زیر فضای درون آن سامانه است که در آن سامانه هر رنگ تنها با یک پیکسل بیان میشود.

بیشتر مدل های رنگی که اکنون استفاده می شوند به سمت سخت افزار (مانند مانیتور و چاپگرهای رنگی) یا کاربردهایی گرایش دارند که هدف آنها کار با رنگ استمانند تولید گرافیک های رنگی برای

^۱color scanner

انیمیشن.

عمومی ترین مدل های ساخت افزار گرا عبارت اند از:

مدل RGB (قرمز، سبز، آبی) برای نمایشگرهای رنگی و یک گروه وسیع از دوربین‌های رنگی.

مدل CMY (آبی فیروزه‌ای، بنفش، زرد) برای چاپگرهای رنگی.

مدل YIQ که یک استاندارد پخش تلویزیون رنگی است. در این مدل Y متناظر با لومیناتوس است، I و Q دو مؤلفه رنگ هستند که به ترتیب هم فاز^۱ و متعامد^۲ خوانده می‌شود. مدل HIS (اصل رنگ، اشاع، مقدار) از مدل‌های هم‌ستند که به طور مکرر برای کار با تصویر رنگی استفاده می‌شوند.

مدل‌های رنگی YIQ و RGB و HIS متداول ترین مدل‌هایی هستند که برای پردازش تصویر استفاده می‌شوند.

۱-۳-۲ مدل رنگ: RGB

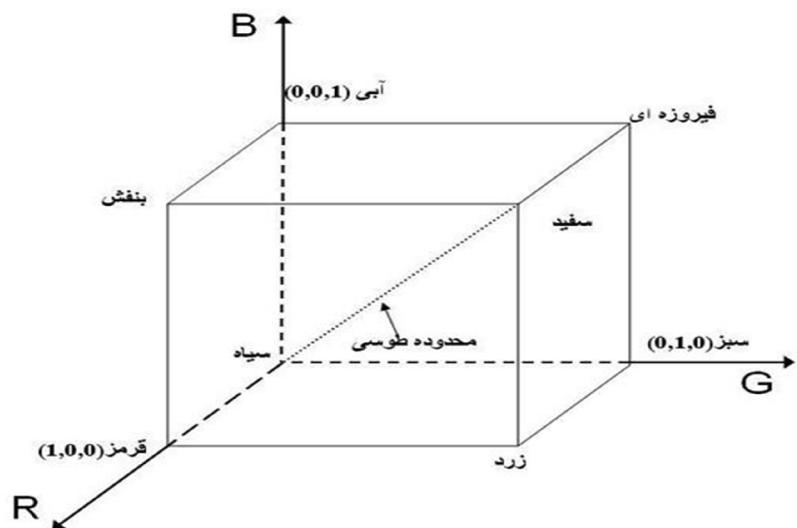
این مدل براساس سامانه مختصات‌کارتزین است. زیر فضای رنگی مورد علاقه، مکعب تصویر زیر می‌باشد، که در آن مقادیر RGB در سه گوشه، آبی فیروزه‌ای، بنفش، و زرد در سه گوشه دیگر، سیاه در مبدأ و سفیددر دورترین گوشه از مبدأ قرار دارد.

در این مدل محدوده خاکستری از سیاه تا سفیددر طول خط واصل این دو نقطه قرار دارد و سایر رنگ‌ها نیز نقاطی در روی یا درون مکعبه‌ستند، که با بردارهایی که از مبدأ می‌گذرد تعریف می‌شوند.

¹Inphase

²Quadrature

برای تسهیل مدل، فرض براین است که تمام مقادیر رنگ تراز شده اند طوری که مکعب تصویر زیر مکعب واحد باشد، یعنی فرضی شود که تمام مقادیر G و B و R در محدوده $[0, 1]$ قرار دارند.



شکل ۳-۲ مکعب واحد در در مقادیر G و B و R

هر تصویر در مدل رنگ RGB سه صفحه مستقل دارد که هر صفحه برای رنگ اولیه می باشد. وقتی این سه صفحه به نمایشگر RGB داده شوند، روی صفحه ها ترکیب می شوند تا یک تصویر رنگی از مدل RGB جهت پردازش تصویر معنا می یابد. همچنین اغلب دوربین های رنگی که برای تصویر برداری دیجیتالی به کار می روند از این قالب استفاده می کنند که این موضوع خود به تنها یی مدل RGB را مدل مهمی در مبحث پردازش تصویر می سازد. یکی از بهترین مثال های کاربرد مدل RGB پردازش داده های چند فریمی (frame) هوایی یا ما هواره ای است.

۲-۳-۲ مدل رنگ CMY :

فیروزه ای، بنفش و زرد رنگ های ثانویه نوری یا رنگ های اولیه مادیهستند. برای مثال وقتی بر سطح پوشیده از ماده رنگی فیروزه ای، نور سفید تابیده میشود هیچ نور قرمزی از آن منعکس نمی شود یعنی فیروزه ای نور قرمز را از نور سفید تفریق می کند بیشتر وسایلی که مواد رنگی را بر روی کاغذ می نشانند، نظیر چاپگرها و کپی برداری های رنگی به داده های CMY نیاز دارند یا این که در داخل خود، داده های CMY را به RGB تبدیل می کنند که با رابطه زیر این عمل امکان پذیر می شود فرض بر ایناست که تمام مقادیر رنگی در محدوده [۰,۱] تراز شده اند.

$$\begin{bmatrix} G \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

۳-۳-۲ مدل رنگ YIQ :

این مدلدر پخش عمومی تلویزیون رنگی تجاری استفاده می شود. در واقع YIQ تغییر تصویر RGB بهمنظور افزایش بازده انتقال و حفظ سازگاری با استانداردهای تلویزیون تک رنگ می باشد. مولفه Y تمام اطلاعات ویدیوئی مورد نیاز تلویزیون تک رنگ می باشد تبدیل RGB به صورت زیر

تعریف می شود:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

جهت به دست آوردن مقادیر RGB از مقادیر YIQ به عرضبند (یا در حالت دیجیتالی) بیشتری برای نمایش Y و عرض باند (یا بیت ها) کمتری برای نمایش I و Q نیاز دارد.

مزیت اصلی مدل YIQ در پردازش تصویر ناشی از خاصیتی است که اطلاعات لومیناتوس (Y) و اطلاعات رنگی I و Q ناهمبسته^۱ هستند.

۴-۳-۲ مدل رنگ HIS:

اصل رنگ است، که رنگ خالص را توصیف می کند (زرد، نارنجی یا قرمز خالص) در حالی که اشباع میزانی که یک رنگخالص با نور سفید ترقیق شده است، را می دهد. سودمندی مدل رنگ HIS مدیون دو عامل اصلی است:

الف) مولفه شدت I از اطلاعات رنگ تصویر مجزا است.

ب) مولفه های اصل رنگ و اشباع، رابطه نزدیکی با روش دریافت رنگ توسط انسان دارند.

این ویژگی ها، مدل HIS را به ابزار ایده آلی برای تولید الگوریتم های پردازش تصویر که مبنی بر بعضی خواص احساس رنگ سامانه بینایی انسان هستند، تبدیل می کند.^[25]

¹Decoupled

۴-۲ انواع پردازش تصویر:

برای پردازش تصویر روش های مختلفی وجود دارد که ۴ روش از آنها را مورد بررسی قرار می

دهیم.

- پردازش نقطه ای^۱

پردازش نقطه ای (پیکسل به پیکسل) ارزش واقعی پیکسل ها را تغییر می دهد.

- پردازش سطحی^۲

پردازش سطحی ارزش پایه و واقعی پیکسل و پیکسل های کنار آن پیکسل را تغییر می دهد.

- پردازش هندسی^۳

پردازش هندسی مکان ها یا ترتیب پیکسل ها را تغییر می دهد که در تصاویر زیر مشاهده می شود

تصویر پردازش شده به حالت کروی در آمده است.



شکل ۴-۲ پردازش هندسی بر روی تصویر

¹Point Process

²Area Process

³Geometric Process

۴-۴-۲ پردازش قاب^۱

اینروش پردازش تصویر با مخلوط کردن^۲ دو یا چند تصویر با یکدیگر تصویر جدیدی به وجود می آورد که در تصاویر زیر مشاهده می شود.[26]



شکل ۵-۲ مخلوط کردن دو یا چند تصویر با یکدیگر

RayTracing ۵-۲

در واقع برگرفته از ایده یونانیان باستان در کشیدن نقاشی های^۳ بعدی می باشد و یکی از مهمترین application های گرافیک کامپیوترا می باشد. در واقع یونانیان باستان از صفحه ای مشبک که در جلوی تصویر مورد نظر قرار می دادند برای کشیدن تصاویر استفاده می کردند. به این صورت که در هر قسمت از صفحه مشبک تصویری را که از همان بخش دیده می شد ترسیم می کردند و در نهایت و پس از برداشتن مشبک تصویر مورد نظر بر روی کاغذ

¹Frame Process
²combine

ترسیم گشته بود.

• نحوه عملکرد

در واقع با استفاده از همین اصل در RayTracing ما با فرض اینکه نور از چشم ما به اجسام برخورد می کند، در اصل حرکت Ray یا همان پرتو نوری را دنبال می کنیم و محل های برخورد آن را با اشیا صفحه نمایش یافته و سپس با الگوریتم بازگشتنی مسیر مورد نظر را به منظور به دست آوردن رنگ مورد نظر برای شی موجود که با Ray برخورد داشته است، برمیگردیم.

حال می توان عدد بدست آمده از این مرحله را نرمال سازی نمود و سپس آن را در صفحه مانیتور به عنوان رنگ مورد نظر به نمایش درآورد.

در ادامه باید خاطر نشان کرد که این کار برای تک تک پیکسلهای تصویر مورد نظر باید صورت پذیرد و لذا عملیات بسیار پرهزینه ای چه از لحاظ زمانی و چه از لحاظ سخت افزار مورد نظر می باشد.

• کاربردها

عملیات RayTracing امروزه طیف وسیعی از application های گرافیک کامپیوتری را به خود اختصاص داده است که چند مورد از انواع آنها را در اینجا ذکر می کنیم:

۱-استفاده از موتورهای پیشرفته Raytracing در نرم افزارهایی که اطلاعات تصویر را گرفته و تصویر را به کاربر ارائه می دهند از قبیل Max.۳

۲-استفاده از آنها در رender کردن^۱ Frame by Frame انیمیشن ها و بازی های گرافیکی و حتی فیلم های سینمایی که در آنها از جلوههای تصویری استفاده گردیده است.

¹Render

۳- استفاده از آنها در پزشکی برای ترسیم تصاویر شبیه سازی شده مربوط به بافت های مختلف بدن و DNA.

۴- استفاده از آنها در ترسیم اشکال شبیه سازی شده از فضا و سایر کاربردهای متعدد دیگر [27].

۶- ارزش پیکسل:

ارزش پیکسل مقدار درخشنندگی هر پیکسل را نشان می دهد.

به عنوان مثال برای تصاویر دودوئی Pixel Value (Binary image) یک بیت بوده که نشان دهنده پیش نمایا پس نما (۰، ۱) است. برای یک عکس grayscale (خاکستری) Pixel Value یک عدد تک رقمی است که نشان دهنده درخشنندگی یا رنگ هر پیکسل است. عمومی ترین Pixel Value برای تصاویر، تصاویر با ایتی است که به ۸ بیت می رسد که اعداد از ۰ تا ۲۵۵ تغییر می کنند. برای سیاه و ۰ برای سفید و رنگ های دیگر راما بین این اعداد می توان به وجود آورد.

جدول زیر برخی از Pixel Value ها را نشان می دهد:

جدول ۲-۱: نمونه برخی از ارزش پیکسل

Bits of Resolution	Number of Gray Levels	Black Value	White Value
1	2	0	1
4	16	0	15
6	64	0	63
8	256	0	255
12	4096	0	4095
16	65536	0	65535

۷-۲ اتصالات پیکسل ها در تصاویر:

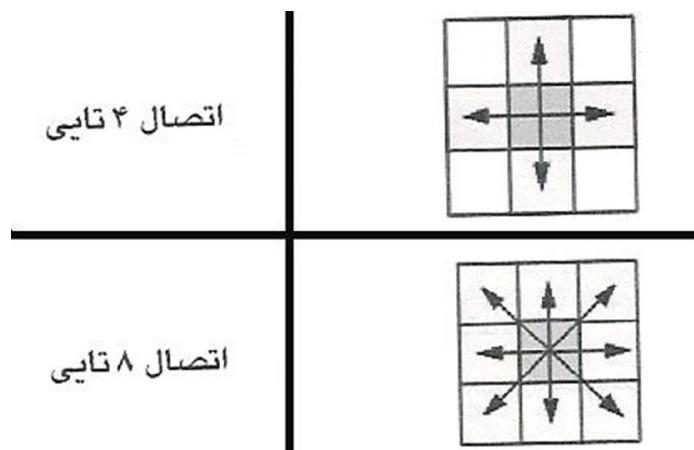
اتصال بین پیکسل ها مفهوم مهمی است که در تعیین مرز اشیا باجزای نواحی تصاویر استفاده می شود. جهت تعیین اتصال بین دو پیکسل باید نوع همسایگی پیکسل ها مشخص گردد. به عنوان مثال، آیا همسایه ها چهار گانه هستند و آیا سطوح خاکستری آنها معیار شباهت مشخص شده ای را برآورده می کنند؟

برای مثال در تصاویر دودوئی با مقادیر صفر و یک ممکن است دو پیکسل، همسایه چهار گانه باشند اما متصل گفته نمی شوند مگر اینکه دارای مقادیر یکسانی باشند.

اتصالات پیکسل ها، تعیین کننده اتصال هر پیکسل با پیکسل مجاور و همچنین نوع اتصالات هر پیکسل می باشد. به تنظیمات گروهی پیکسل ها و اتصالات آنها در تصاویر دودوئی، Object یا [28]. Connected component گفته می شود.

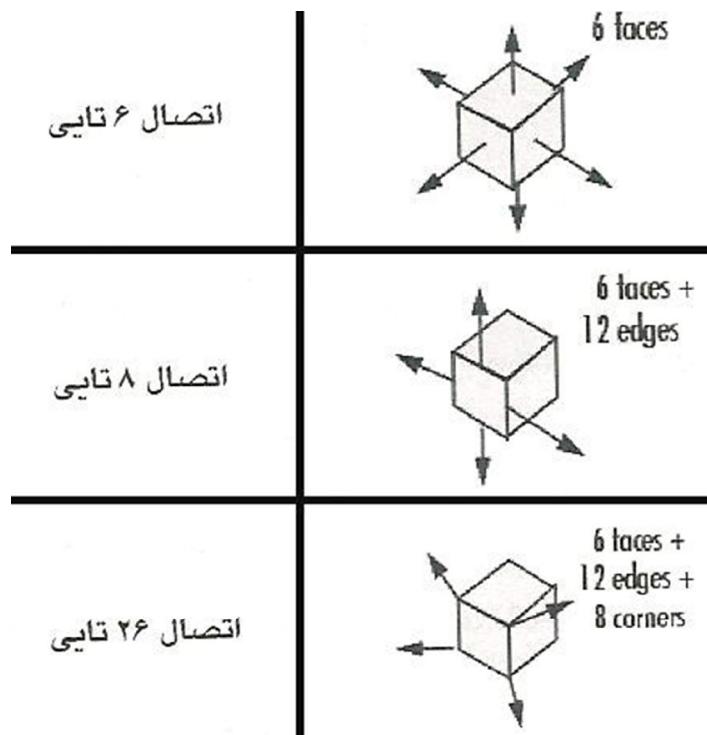
تعیین اتصالات پیکسل های تصویر:

لیست زیر شامل اتصالات تصاویر دو بعدی است:



شکل ۲-۶: اتصالات تصاویر دو بعدی

لیست زیر شامل اتصالات تصاویر سه بعدی است:



شکل ۷-۲: اتصالات تصاویر سه بعدی

۲-۸ درون یابی^۱

درون یابی که گاهی resampling نیز نامیده می شود یک روش گرافیکی برای افزایش یا کاهش تعداد پیکسل های یک تصویر دیجیتالی است. دوربین های دیجیتالی از این روش برای زوم دیجیتال و یا افزایش مصنوعی ابعاد تصویر، نسبت به تصویر اصلی ثبت شده توسط حسگر خود استفاده می کنند. کیفیت تصویر نهایی حاصل از درون یابی به پیچیدگی الگوریتم ساخت تصویر بستگی دارد. حالت های متفاوت درون یابی عبارتند از:

¹Interpolation

Nearest Neighbor Interpolation ۱-۸-۲

این روش ساده‌ترین عملیات درون‌یابی محسوب می‌شود که اساساً پیکسل‌های بزرگ‌تری را ایجاد می‌کند و رنگ هر پیکسل در تصویر جدید، با رنگ نزدیک‌ترین پیکسل در تصویر اصلی مطابقت دارد. به عنوان مثال اگر شما یک تصویر را ۲۰۰ درصد بزرگ کنید، یک پیکسل تصویر به یک فضای چهار پیکسلی (دو پیکسل در دو پیکسل) با رنگ پیکسل اولیه تبدیل می‌شود. بسیاری از دوربین‌های دیجیتالی برای زوم‌کردن روی یک سوزه از این روش استفاده می‌کنند. زیرا این روش هیچ تأثیری بر اطلاعات رنگی تصویر نمی‌گذارد. اصولاً بهتر است از این روش برای افزایش ابعاد تصویر استفاده نشود. چرا که موجب ایجاد حالت شطرنجی در تصویر می‌گردد.

Bilinear Interpolation ۲-۸-۲

این روش ارزش یک پیکسل رنگی را براساس چهار پیکسل در جهت‌های عمودی و افقی پیکسل در تصویر اصلی معین می‌کند. تصویر جدید دارای خاصیت Anti-aliasing است و تقریباً هیچ اثری از پیکسل‌های شطرنجی در آن دیده نمی‌شود.

Bicubic Interpolation ۳-۸-۲

این روش دارای پیچیده‌ترین الگوریتم درون‌یابی است و تصویری که با این روش به دست می‌آید، لبه‌های بسیار نرم‌تری پیدا می‌کند. در این حالت پیکسل جدید براساس ارزش تخمینی ۱۶ پیکسل (چهار پیکسل در چهار پیکسل) برآورده شود. شایان ذکر است که اغلب دوربین‌ها، چاپگرهای و نرم‌افزارهای گرافیکی برای تغییر تعداد پیکسل‌های یک تصویر از این روش استفاده می‌کنند.

* نرم افزار Photoshop CS دو نوع Smoother و Sharper این درون یابی را نیز ارائه می دهد .

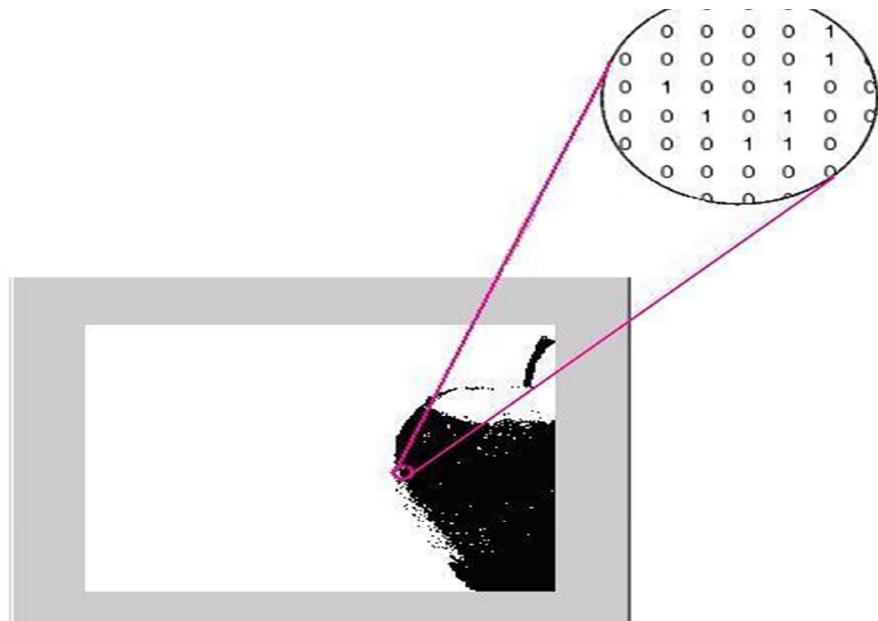
Fractal Interpolation ۴-۸-۲

این روش برای بزرگ کردن تصویر در ابعاد خیلی بزرگ (به عنوان مثال چاپ روی بدن) یک اتوبوس) روش بسیار مناسبی به شمار می رود. زیرا با دقت خوبی می تواند شکل اجزای تصویر را حفظ کند. لبه های بسیار تمیزتر و صاف تری را نیز بر جا می گذارد. در این حالت مناطق محیطی که در لبه های تصویر در روش های قبل وجود داشتند، به ندرت دیده می شوند. البته روش های دیگری هم وجود دارند که جز در برنامه های گرافیکی پیشرفته، به ندرت از آنها استفاده می شود [29] .

۹-۲ انواع تصویر:

۱-۹-۲ تصاویر دودوئی (Binary)

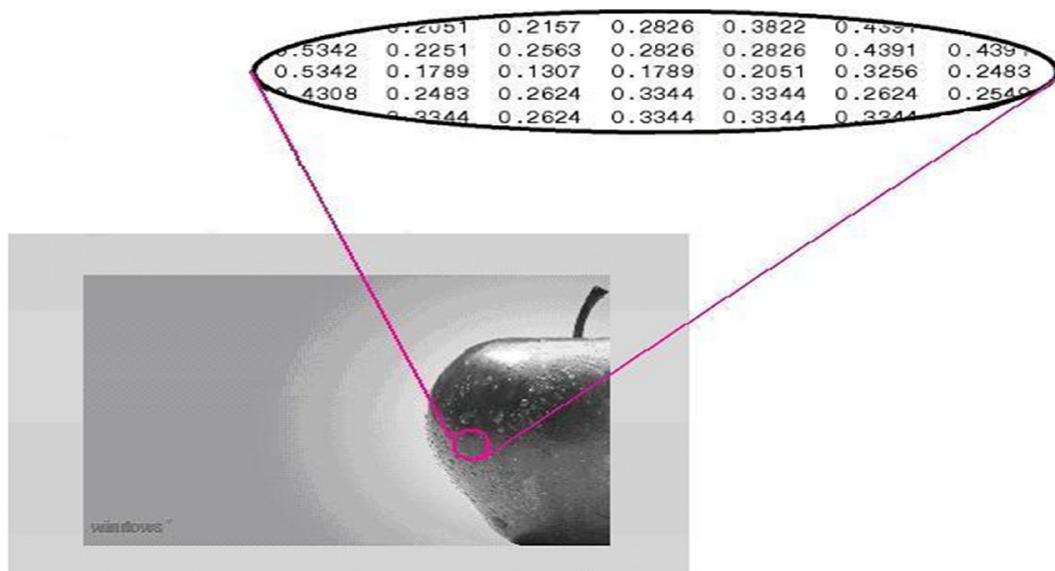
این دسته از تصاویر تنها دارای دو ارزش مجزا می باشند، که شاملیک آرایه منطقی به صورت ۰ با عنوان Off Pixel و آرایه منطقی دیگری به صورت ۱ با عنوان On Pixel می باشد، که نحوه نمایش این دسته از تصاویر به صورت زیر میباشد:



شکل ۲-۸: نحوه نمایش تصاویر دودویی

۲-۹-۲ تصاویر grayscale

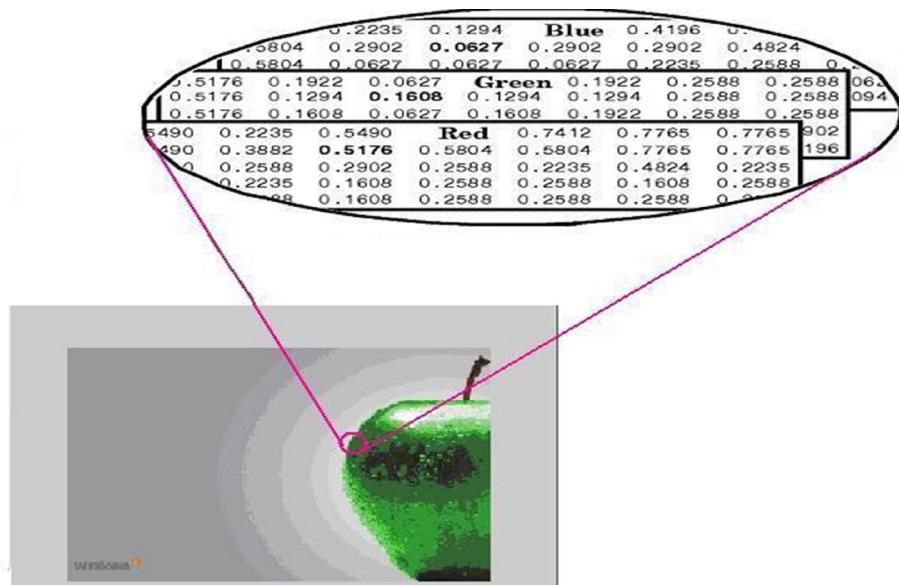
فرمت این دسته از تصاویر به صورت خاکستری است.



شکل ۲-۹: نحوه نمایش تصاویر grayscale

۹-۳ تصاویر RGB :

فرمت ایندسته از تصاویر به صورت رنگی است.



شکل ۱۰-۲: نحوه نمایش تصاویر RGB

۴-۹-۲ خاکستری (Grayscale) کردن تصویر:

در رایج ترین مدل رنگ گرافیک کامپیوتری، رنگ ها از ترکیب سه رنگ قرمز، سبز و آبی به وجود می آیند که در مجموع ۱۶۵۸۱۳۷۵ رنگ متفاوت توسط این سه مولفه می توان تولید کرد. این مدل رنگ در گرافیک کامپیوتری با نام RGB شناخته می شود. در کنار مدل رنگ RGB مدل های دیگری همچون Grayscale و HSV، HSI، CMYK نیز وجود دارد که هریک از آن ها به روش متفاوتی به نمایش رنگ ها می پردازند. در این بین مدل رنگ خاکستری از اهمیت ویژه ای برای ما برخوردار است. چرا که در بیشتر کاربردها نیازی به یک تصویر رنگی نمی باشد و داشتن تنها یک تصویر خاکستری کافی خواهد بود. در میان عموم، تصویر خاکستری با نام تصویر سیاه و سفید شناخته می شود (البته استفاده از نام سیاه

سفید به جای خاکستری اشتباه است و در اینجا فقط برای روشن تر شدن مطلب از این نام استفاده کرده ایم) یک تصویر RGB متشكل از سه ماتریس است که هریک از آنها مقادیر Red (قرمز)، Green (سبز) و Blue (آبی) تصویر رنگی را نگه می دارند. نمایش تصویر بر روی صفحه نمایش نیز از ترکیب مقادیر درآیه های متناظر در سه ماتریس انجام می پذیرد. همانطور که در ابتدا یادآور شدیم، در بیشتر کاربردها نیازی به تصویر رنگی نداریم و استفاده از یک تصویر خاکستری کافی خواهد بود. یک پیکسل زمانی مقدار خاکستری خواهد داشت که مولفه های R و G و B آن مقادیر یکسانی داشته باشند. با توجه به این تعریف، در مواردی که تصویر ورودی یک تصویر RGB است، برای خاکستری کردن آن از فرمول زیر می توانیم استفاده می کنیم:

$$S_G(x, y) = S_B(x, y) = [R(x, y) + G(x, y) + B(x, y)] / 3 = S_R(x, y)$$

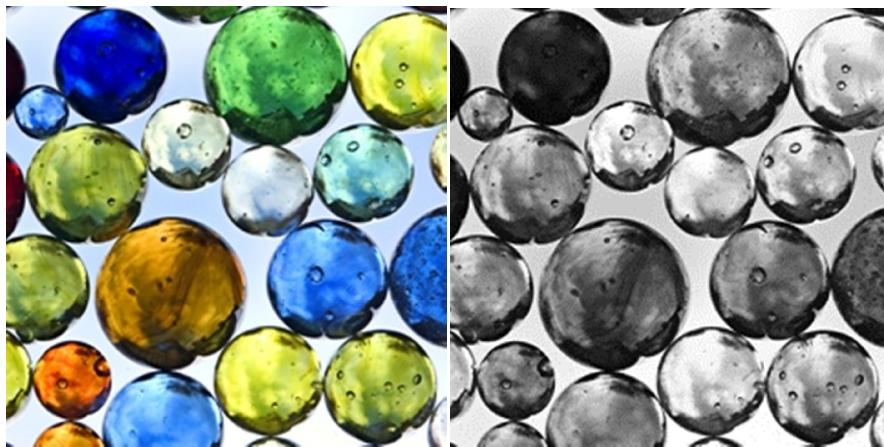
در این فرمول S_X مولفه های تصویر خروجی و R، G و B هر یک به ترتیب ماتریس مولفه های سبز قرمز و آبی تصویر ورودی می باشند. شبیه کد زیر نحوه خاکستری کردن یک تصویر RGB با سایز $N * M$ را نشانمی دهد:

```

Procedure Grayscale( output , input As Bitmap )
Begin
For I = 1 to M Do
For J = 1 To N Do
Temp = ( input.R( I,J ) + input.G( I,J ) + input.B( I,J ) ) / 3
Output.R ( I,J ) = Temp
Output.G ( I,J ) = Temp
Output.B ( I,J ) = Temp
End For
End For
End

```

شکل زیر یک تصویر RGB را به همراه تصویر خاکستری آن نشان می دهد:



شکل ۱۱-۲: نمایش تصویر RGB به همراه تصویر Grayscale

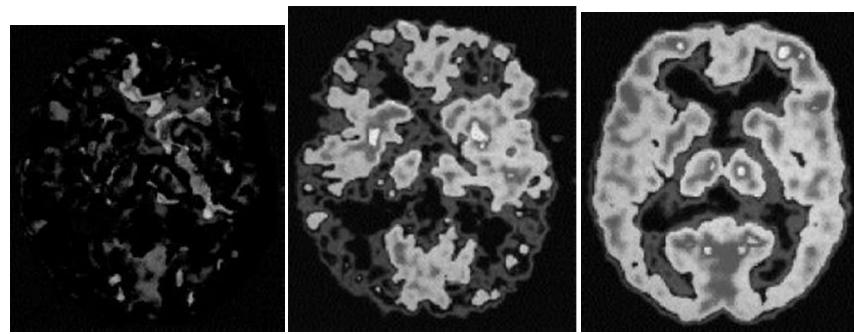
از آنجا که مقادیر ماتریس های R , G , B پس از خاکستری کردن تصویر باهم برابر هستند، بنابراین می توان تنها از یک ماتریس برای نشان دادن مقادیر استفاده کرد. در الگوریتم های بعدی که برای پردازش تصویر ارائه خواهند شد از تصاویر خاکستری استفاده خواهیم کرد. بنابراین عملیات نیز بر روی تنها یک ماتریس انجام خواهد پذیرفت.[30]

۱۰-۲ تفریق دو تصویر:

تفریق دو تصویر هم اندازه بدین مفهوم است که شدت روشنایی پیکسل های متناظر دو تصویر را از هم کم کنیم. فرض کنید می خواهیم تغییرات مغز افرادی را بررسی کنیم که دچار بیماری آלצהیمر هستند. برای این منظور می توانیم تصویری از یک مغز سالم را در مراحل مختلف با تصویر مغز فردی که

دچار بیماری آلزایمر است، مقایسه کنیم. با اعمال عملگر فوق بر روی دو تصویر مذکور نقاطی از مغز که در آن نقاط مغز دچار تغییر شده است مشخص می شوند.

تصویر زیر عکس اسکن PET مغز نرمال و بیمار را به همراه حاصل تفریق آن ها نشان می دهد:



تصویر مغز فرد سالم

تصویر مغز فرد دچار بیماری آلزایمر

حاصل تفریق دو تصویر

شکل ۱۲-۲ تفریق دو تصویر

شبه کد زیر نحوه پیاده سازی عملگر تفریق را نشان می دهد:

```
Procedure Subtract (image1, image2: Byte[][])
Begin
result: Byte[][]
sub: integer
For i = 1 To Height Do
For j = 1 To Width Do
sub = image1[i, j] - image2[i, j]
result[i, j] = sub < 0 ? 0: sub
End For
End For
Return result
End
```

هنگام تفريقي مقادير پيكسل ها، مقادير منفي را به مقدار صفر تبديل می کنيم. همچنين می توانيم هنگام تفريقي از قدر مطلق تفريقي نيز استفاده کنيم به عنوان يك مثال کاربردي از تفريقي دو تصوير می توان به شناسايي حرکت در سيسitem های دوربین مدار بسته اشاره کرد. زمانی می گویيم حرکت رخ داده است که در بين دو فریم متوالی گرفته شده از دوربین اختلاف وجود داشته باشد. و همانطور که در ابتدا يادآور شدیم، برای محاسبه اختلاف بين دو تصوير از عملگر تفريقي استفاده می کنيم. بنابراین با تفريقي فریم فعلی و فریم قبلی گرفته شده از دوربین می توانيم اختلاف موجود در دو تصوير را پیدا کنيم. از ديگر کاربردهای تفريقي دو تصوير می توان به حذف پشت زمینه ثابت از تصوير اشاره کرد.[31]

۱۱-۲ جمع دو تصوير:

جمع دو تصوير بدین مفهوم است که در دو تصوير، شدت روشناني پيكسل های متناظر دو تصوير را از با هم جمع کنيم. تصاویر زیر نتيجه اعمال عملگر جمع را نشان می دهند:



تصویر اول

تصویر دوم

حاصل جمع دو تصویر

شكل ۱۳-۲ جمع دو تصوير

شبیه کد زیر نحوه پیاده سازی عملگر جمع را نشان می دهد:

```
Procedure Add (image1, image2: Byte[][])
Begin
result: Byte[][]
sub: integer
For i = 1 To Height Do
For j = 1 To Width Do
sub = image1[i, j] + image2[i, j]
result[i, j] = sub > 255 ? 255: sub
End For
End For
Return result
End
```

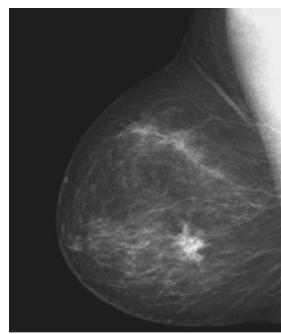
یکی از رایج ترین کاربردهای جمع دو تصویر افزودن پشت زمینه به تصویر است. برای این منظور نیاز به دو تصویر پشت زمینه و تصویر شی داریم که تصویر شی باید در مقابل پرده ای با رنگ ثابت گرفته شود. از جمع کردن دو تصویر مذکور تصویری با پشت زمینه دلخواه به دست می آید.[32]

۱۲-۲ مکمل کردن تصویر:

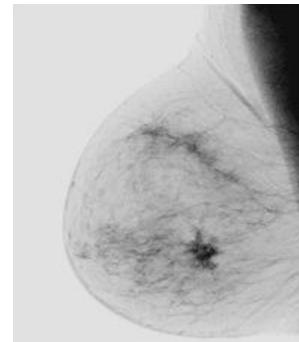
در یک تصویر سیاه و سفید که فقط دو سطح شدت روشنایی ۰ و ۱ دارد، منظور از مکمل کردن تصویر، ۱ کردن پیکسلهای با مقدار ۰ و ۰ کردن پیکسل های با مقدار ۱ است. در تصویری که از L سطح روشنایی ([L .. ۱]) برای نمایش شدت روشنایی پیکسل ها استفاده می کند، مکمل یک پیکسل از فرمول زیر به دست می آید:

$$\text{Pixel}[i, j] = L - 1 - \text{Pixel}[i, j]$$

به شکل های زیر توجه کنید:



شکل ۱۴-۲ ب: تصویر ماموگرام



شکل ۱۴-۲ الف: مکمل تصویر

شکل سمت چپ یک تصویر ماموگرام است و شکل سمت راست نیز مکمل آن را نشان می دهد.

همانطور که در این دو شکل به وضوح دیده می شود، تشخیص اجزای تصویر مکمل شده آسانتر از تصویر اصلی است. در چنین مواردی که اجزای اصلی تصویر در یک تصویر با محدوده سیاه بزرگ قرار گرفته باشد، از مکمل تصویر استفاده می کنیم.

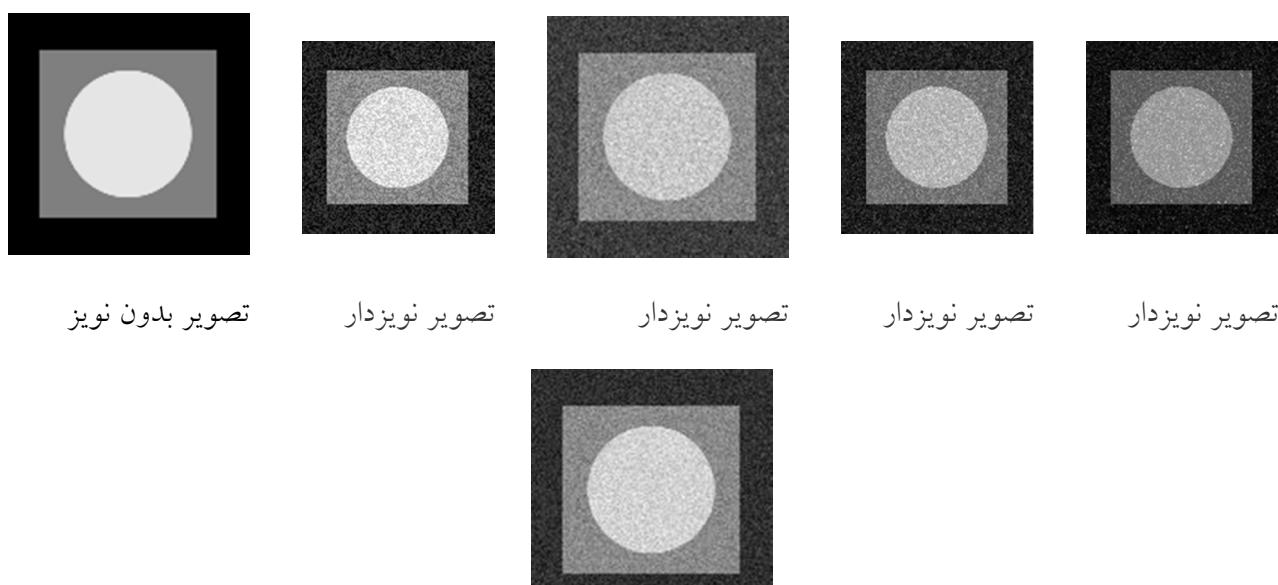
شبه کد زیر نحوه مکمل کردن تصویر را نشان می دهد:

```
Procedure Complement ( input: Image )
Begin
For 1 To Height of image Do
For 1 To Width of image Do
Inout.pixels[ i , j ] = 255 - Inout.pixels[ i , j ]
End For
End For
End
```

فرض کنیم می خواهیم یک روبات تعقیب خط با استفاده از دوربین ایجاد کنیم. همانطور که می دانید یکی از قوانین ربات تعقیب خط مکمل شدن رنگ خط و پش زمینه در مقاطعی از مسیر است. برای مرتفع کردن این شرط مسابقه می توان از عملگر مکمل کردن تصویر استفاده کرد.[33]

۱۳-۲ میانگین گیری از تصویر:

فرض کنید چند تصویر یکسان داریم که بر روی هر کدام از آنها نویزهای مختلفی وجود دارد و می خواهیم کیفیت این تصاویر را ارتقا دهیم. در چنین مواردی می توان از میانگن گیری از همه تصاویر استفاده کنیم. بدین صورت که مقادیر پیکسل های متناظر در همه تصاویر را باهم جمع کرده و سپس به تعداد کل تصاویر تقسیم کنیم. بدیهی است که هر قدر تعداد تصاویر برای میانگین گیری بیشتر باشد، تصویر حاصل از میانگین گیری آنها نیز بیشتر به واقعیت نزدیک خواهد بود. به عنوان مثال مجموعه تصاویر زیر، تصویر بدون نویز، تصاویر نویز دار و حاصل میانگین تیری از تصاویر نویزرا نشان می دهد:



شکل ۱۵-۲ میانگین گیری از تصویر

شبه کد زیر نحوه پیاده سازی عملگر میانگین را نشان می دهد:

```
Procedure Average( bmp1 , bmp2: Image)
Begin
Result: Image
For 1 To Height Do
For 1 To Width Do
Result.Pixels[ i , j ] = ( bmp1.Pixels[ i , j ] + bmp2.Pixels[ i , j ] ) /2
End For
End For
End
```

هنگام تفرق مقادیر پیکس ها، مقادیر منفی را به مقدار صفر تبدیل می کنیم. همچنین می توانیم

هنگام تفرق از قدرمطلق تفرق نیز استفاده کنیم. به عنوان یک مثال کاربردی از تفرق دو تصویر می توان

به شناسایی حرکت در سیستم های دوربین مدار بسته اشاره کرد. زمانی می گوییم حرکت رخ داده است

که در بین دو فریم متوالی گرفته شده از دوربین اختلاف وجود داشته باشد. و همانطور که در ابتدا یادآور

شده، برای محاسبه اختلاف بین دو تصویر از عملگر تفرق استفاده می کنیم. بنابراین با تفرق فریم فعلی

و فریم قبلی گرفته شده از دوربین می توانیم اختلاف موجود در دو تصویر را پیدا کنیم. از دیگر

کاربردهای تفرق دو تصویر می توان به حذف پشت زمینه ثابت از تصویر اشاره کرد.[34]

۱۴-۲ هیستوگرام تصویر:

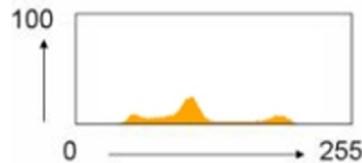
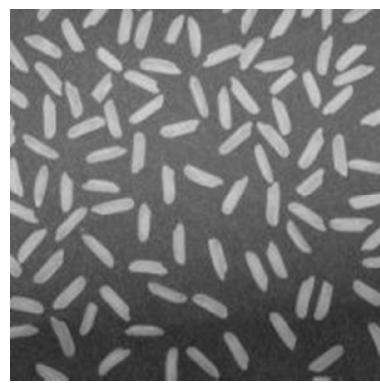
هیستوگرام تصویر نموداری است که توسط آن تعداد پیکسل های هر سطح روشنایی در تصویر

ورودی مشخصی شود. فرض کنید تصویر ورودی یک تصویر خاکستری با ۲۵۶ سطح روشنایی باشد،

بنابراین هریک از پیکسل های تصویر مقداری در بازه [۰...۲۵۵] می توانند داشته باشند.

برای به دست آوردن هیستوگرام تصویر، کافی است با پیمایش کل پیکسل های تصویر، تعداد پیکسل های هر سطح روشناهی را محاسبه می کنیم. هیستوگرام نرمال نیز از تقسیم کردن مقادیر هیستوگرام به تعداد کل پیکسل های تصویر به دست می آید. نرمال سازی هیستوگرام موجب می شود که مقادیر هیستوگرام در بازه $[0, 1]$ قرار گیرند.

شكل زیر تصویری را به همراه هیستوگرام نرمال آن نشان می دهد.



شکل ۱۶-۲: نرمال سازی هیستوگرام

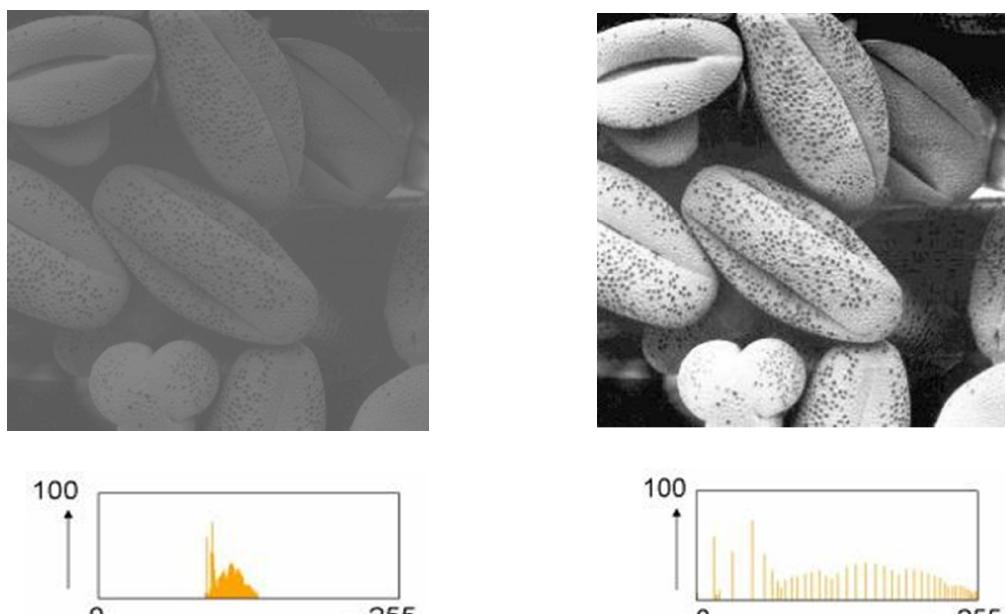
یکی از کاربردهای هیستوگرام در فوکوس خودکار دوربین های دیجیتالی است. بدین صورت که لنز دوربین از ابتدا تا انتهای حرکت کرده و در هر گام از حرکت خود تصویری از صحنه می گیرد. سپس کنترast تصویر گرفته شده را با استفاده از هیستوگرام آن محاسبه می کند. پس از آن که لنز به انتهای حرکتی خود رسید، محلی از حرکت لنز که در آن تصویر دارای بیشترین کنترast خود بوده است، به عنوان محل لنز تعیین می گردد.

روش مذکور یکی از ساده ترین روش های فوکوس خودکار دوربین می باشد و همانطور که می توان حدس زد این الگوریتم در صحنه هایی که رنگ تیره و روشن باهم وجود داشته باشد، دچار اشکالاتی خواهد بود و باید تغییراتی در آن اعمال کرد (مفهوم کتراست در ادامه بیان شده است).[35]

۱۵-۲ تعدیل هیستوگرام:

یکی دیگر از کاربردهای هیستوگرام در افزایش کتراست تصاویر با کتراست پایین است. زمانی که می گوییم کتراست تصویری کم است این بدان معنا است که اختلاف بین کمترین و بیشتری شدت روشنایی تصویر کم است. تعدیل سازی هیستوگرام موجب می شود که کتراست تصویر ورودی تا حد ممکن افزایش یابد.

به عنوان مثال شکل زیر تصویری را قبل و بعد از تعدیل سازی هیستوگرام نشان می دهد:



تصویر ورودی و هیستوگرام آن

تصویر خروجی پس از تعدیل هیستوگرام

شکل ۲ ۱۷-۲ تعدیل سازی هیستوگرام بر روی تصویر

الگوریتم زیر روش تعدیل سازی هیستوگرام را نشان می دهد:

1 - هیستوگرام تصویر را محاسبه می کنیم. فرض کنید مقادیر هیستوگرام در آرایه hist قرار گیرد.

2 - با استفاده از فرمول زیر فراوانی هیستوگرام را محاسبه می کنیم:

$$\text{histCum}[i] = \text{histCum}[i-1] + \text{hist}[i]$$

3 - از فرمول زیر استفاده کرده و هیستوگرام تعدیل شده را محاسبه می کنیم:

$$\text{eqHist}[i] = \text{Truncate}((L * \text{histCum}[i]) - N) / N$$

4 - در مرحله نهایی مقادیر جدید پیکسل ها را به صورت زیر مقدار دهی می کنیم:

$$\text{Result}[i, j] = \text{eqHist}[\text{input}[i, j]]$$

که در این فرمول L تعداد سطوح خاکستری و N تعداد کل پیکسل ها، Result تصویر خروجی و

input تصویر ورودی را نشان می دهد.[36]

۱۶-۲ فشردهسازی تصاویر:

برای ذخیرهسازی تصاویر باید حجم اطلاعات را تا جایی که ممکن است کاهش داد و اساس تمام

روش های فشردهسازی کنار گذاردن بخش هایی از اطلاعات و داده ها است.

ضریب یا نسبت فشردهسازی است که میزان و درصد کنار گذاشتن اطلاعات را مشخص می کند. این

روش ذخیرهسازی و انتقال اطلاعات را آسان تر می کند و پهنای باند و فرکانس مورد نیاز کاهش می یابد.

امروزه روش هایی متعدد و پیشرفته برای فشردهسازی وجود دارد. فشردهسازی تصویر از این اصل مهم

تبیعت می کند که چشم انسان حد فاصل دو عنصر تصویری نزدیک به هم را یکسان دیده و تمایز آنها را

نمی‌تواند تشخیص دهد. همچنین اثر نور و تصویر برای مدت زمان معینی در چشم باقی مانده و از بین نمی‌رود که این ویژگی در ساخت تصاویر متحرک مورد توجه بوده است.

^۱JPEG روش^۲

از این روش در فشرده‌سازی عکس و تصاویر گرافیکی ساکن استفاده می‌شود. JPEG اولین و ساده‌ترین روش در فشرده‌سازی تصویر است به همین دلیل در ابتدا سعی شد برای فشرده‌سازی تصاویر متحرک مورد استفاده قرار گیرد. برای این منظور تصاویر به صورت فریم به فریم مانند عکس فشرده می‌شدند و با ابداع روش MOTION JPEG برای ارتباط دادن این عکس‌ها به هم تلاش شد که با مشکلاتی همراه بود.

^۱ MPEG روش^۲

این روش در ابتدای سال ۹۰ ابداع شد و در آن اطلاعات تصویر با سرعت حدود ۱/۵ مگابایت بر ثانیه انتقال پیدا می‌کرد که در تهیه تصاویر ویدئویی استفاده می‌شد. با این روش امکان ذخیره حدود ۶۵۰ مگابایت اطلاعات معادل حدود ۷۰ دقیقه تصویر متحرک در یک دیسک به وجود آمد. در MPEG بیت‌های اطلاعات به صورت سریال ارسال می‌شوند و به همراه آنها بیت‌های کنترل و هماهنگ‌کننده نیز ارسال می‌شوند که موقعیت و نحوه قرارگیری بیت‌های اطلاعاتی را برای انتقال و ثبت اطلاعات صدا و تصویر تعیین می‌کنند.

^۱JOINT PHOTOGRAPHIC EXPERT GROUP

^۲MOVING PICTURE EXOERT GROUP

MP3 روش ۳-۱۶

MP3 نیز روشی برای فشرده سازی اطلاعات صوتی به ویژه موسیقی است که از طریق آن حجم

زیادی از اطلاعات صوتی در فضای نسبتاً کوچکی ذخیره میشود.

MPEG2 روش ۴-۱۶

در روش MPEG2 از ضریب فشرده سازی بالاتری استفاده میشود و امکان دسترسی به اطلاعات ۳

تا ۱۵ مگابایت بر ثانیه است از این روش در دیوی دی های امروزی استفاده می شود در اینجا نیز هر فریم تصویری شامل چندین سطر از اطلاعات دیجیتالی است.

MPEG4 روش ۵-۱۶

از این روش برای تجهیزاتی که با انتقال سریع یا کند اطلاعات سرو کار دارند استفاده میشود. این روش توانایی جبران خطا و ارائه تصویر با کیفیت بالا را دارد. مسئله خطا و جبران آن در مورد تلفن های همراه و کامپیوتر های خانگی و لپ تاپ ها و شبکه ها از اهمیت زیادی برخوردار است. در شبکه های کامپیوتری باید تصویر برای کاربرانی که از مودم های سریع یا کند استفاده می کنند به خوبی نمایش داده شود، در چنین حالتی روش MPEG4 مناسب است. از این روش در دوربین های تلویزیونی نیز استفاده میشود. ایده اصلی این روش تقسیم یک فریم ویدئویی به یک یا چند موضوع است که مطابق قاعده خاصی کنار هم قرار میگیرند مانند درختی که از روی برگ های آن بتوان به شاخه تنہ یا ریشه آن دست

یافت. هر برگ میتواند شامل یک موضوع صوتی یا تصویری باشد. هر کدام از این اجزا به صورت مجزا و جداگانه قابل کپی و یا انتقال هستند. این تکنیک را با آموزش زبان می‌توان مقایسه کرد.

همان‌طوری که در آموزش زبان کلمات به صورت مجزا و جداگانه قرار داده می‌شوند و ما با مرتب کردن آن جملات خاصی می‌سازیم و می‌توانیم در چند جمله، کلمات مشترک را فقط یک‌بار بنویسیم و هنگام مرتب کردن آن‌ها به کلمات مشترک رجوع کنیم، در اینجا هم هر یک از این اجزا یک موضوع خاص را مشخص می‌کند و ما می‌توانیم اجزا مشترک را فقط یک‌بار به کار ببریم و هنگام ساختن موضوع به آنها رجوع کنیم. هر یک از موضوعات هم می‌توانند با موضوعات دیگر ترکیب و مجموعه جدیدی را بوجود آورند. این مسئله باعث انعطاف‌پذیری و کاربرد فراوان روش **MPEG4** می‌شود. برای مثال به صحنه بازی تنیس توجه کنید. در یک بازی تنیس میتوان صحنه را به دو موضوع بازیکن و زمین بازی تقسیم کرد زمین بازی همواره ثابت است بنا بر این بعنوان یک موضوع ثابت همواره تکرار می‌شود ولی بازیکن همواره در حال حرکت است و چندین موضوع مختلف خواهد بود. این مسئله سبب کاهش پهنه‌ای باند اشغالی توسط تصاویر دیجیتالی می‌شود. توجه داشته باشید که علاوه بر سیگنال‌های مربوط به این موضوعات سیگنال‌های هماهنگ کننده‌ای هم وجود دارند که نحوه ترکیب و قرارگیری صحیح موضوعات را مشخص می‌کند[37].

۱۷-۲ تصاویر آنالوگ:

تصاویری مانند عکس‌های هوایی که توسط سیستم‌های عکس برداری (دوربین) به دست می‌آیند. از آنجایی که در این عکس‌ها از فیلم عکاسی استفاده شده است، پس هیچ پردازشی نیاز ندارد.



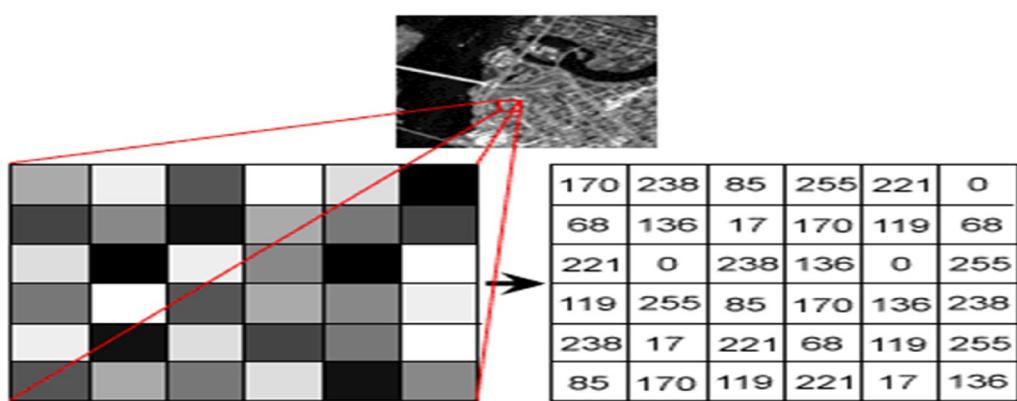
شکل ۱۸-۲ تصویر آنالوگ (عکس هوایی که نیاز به اصلاح و پردازش ندارد).

۱۸-۲ تصاویر رقومی (دیجیتالی):

تصاویر سنجش شده که از تعداد زیادی مربعات کوچک(پیکسل) تشکیل شده اند. هر پیکسل

دارای یک شماره رقمنی (Digital Number) میباشد که بیانگر میزان روشنایی آن پیکسل است.

به این نوع تصاویر، تصاویر رستری هم میگویند. تصاویر رستری دارای سطر و ستون میاشند [38].



شکل ۱۹-۲ تصاویر رقومی، پیکسلی

(تصویر بالا(رقمنی).پایین وسمت چپ(پیکسلها).سمت راست و پایین(شماره‌های هر پیکسل (DN))

۱۹-۲ مقادیر پیکسلها:

مقدار انرژی مغناطیسی که یک تصویر رقومی به هنگام تصویر برداری کسب میکند، رقم های دوتایی^۱ یا بیتها^۲ را تشکیل میدهند که از قوه صفر تا ۲ ارزش گذاری شده است.

هر بیت، توان یک به قوه ۲ (2^1 بیت=۲) میباشد. حداکثر تعداد روشنایی بستگی به تعداد بیت ها دارد.

بنابراین ۸ بیت یعنی ۲۵۶ شماره رقومی که دامنه ای از ۰ تا ۲۵۵ دارد. به همین دلیل است که وقتی شما تصویر رستری از سنجنده خاصی مانند TM را وارد نرم افزاری میکنید تغییرات میزان روشنایی را بین ۰ تا ۲۵۵ نشان میدهد.

۲۰-۲ دقت تصویر:

دقت تصویر بستگی به عدد پیکسل ها دارد. با یک تصویر ۲ بیتی، حداکثر دامنه روشنایی ۲۲ یعنی ۴ میباشد که دامنه آن از ۰ تا ۳ تغییر میکند. در این حالت تصویر دقت (تفکیک پذیری لازم) را ندارد. تصویر ۸ بیتی حداکثر دامنه ۲۵۶ دارد و تغییرات آن بین ۰ تا ۲۵۵ است. که دقت بالاتری دارد.

¹ Digit binary

² Bits

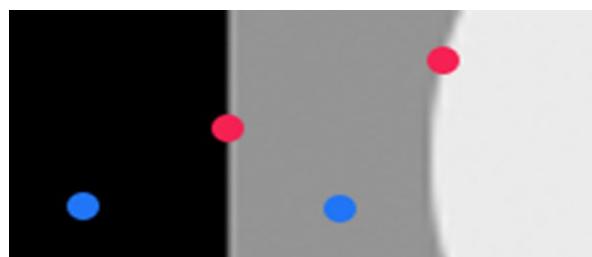


شکل ۲۰-۲ دقت تصویر ۳ بیتی دقت تصویر ۸ بیتی

۲۱-۲ فیلتر کردن تصویر:

در کل، پیکسل های تصویر را به دو دسته می توان تقسیم کرد: پیکسل های تیز و پیکسل های آرام

به عنوان مثال شکل زیر را در نظر بگیرید:



پیکسل های زیر ناحیه های قرمز در این تصویر را در دسته پیکسل های تیز قرار می دهیم. پیکسل هایی از تصویر را که اختلاف شدت روشنایی آنها (مقدار پیکسل) با پیکسل های همسایه خود زیاد باشند، پیکسل های تیز می نامیم. پیکسل های تیز معمولاً یا نشان دهنده لبه های اشیای موجود در تصویر هستند یا نشان دهنده نویزهای یک تصویر هستند و در حالت کلی می توان گفت که جزئیات تصویر با استفاده از پیکسل های تیز استخراج می شوند.

به عنوان مثال، تصاویری که هنگام شب توسط موبایل یا یک وب کم گرفته می‌شوند، دارای نویز معروف فلفل-نمکی هستند. تصویر زیر گویای این مطلب است:



شکل ۲۱-۲ تصویردارای نویز فلفل-نمکی

در مقابل پیکسل های تیز، پیکسل های آرام تصویر قرار دارند (ناحیه های آبی رنگ در مثال فوق). پیکسل هایی از تصویر که اختلاف شدت روشنایی آنها با پیکسل های همچوار خود کم است، پیکسل های آرام تصویر می‌نامیم. حال با توجه به این مطالب به شرح فیلتر کردن تصویر می‌پردازیم. در حوزه مکانی فیلترها به دو دسته تقسیم می‌شوند:

۱- فیلترهای آرام کننده

۲- فیلترهای تیزکننده

فیلترهای آرام کننده با اعمال تغییر در شدت روشنایی پیکسل های تیز تصویر موجب تولید تصاویر آرام تری می‌شود. همچنین این فیلترها تاثیر چندانی بر روی پیکسلهای آرام تصویر نمی‌گذارند. آرام شدن تصویر بدین معنی است که بخشی از جزئیات تصویر از بین می‌رود. به عنوان مثال اعمال یک فیلتر آرام کننده بر روی یک تصویر نویزدار موجب کاهش شدت نویزها خواهد شد. یک مثال کاربردی ساده از فیلتر

های آرام کننده رامی توان حذف چین و چروک ها از تصاویر صورت بیان کرد. ساده ترین فیلترهای آرام کننده تصویر فیلترهای میانگین هستند. یک فیلتر میانگین m^*n ، ماسکی است که مقدار هریک از عناصر آن برابر با عدد ۱ (m.n) / است. به عبارت دیگر مقدار هر پیکسل در تصویر خروجی برابراست با میانگین مقدار شدت روشنایی پیکسل جاری و پیکسل های همسایه پیکسل جاری.

برخلاف فیلترهای آرام کننده، فیلترهای تیز کننده با اعمال تغییر بر روی پیکسل های آرام تصویر موجب استخراج جزئیات بیشتری از تصویر می شوند. در برخی موارد می بینیم که تصویر داری حالتی بلور مانند است (جزئیات نمایان نیستند). در چنین مواردی برای حذف خاصیت بلوری، از فیلترهای تیز کننده تصویر استفاده می کنیم:



تصویر اصلی



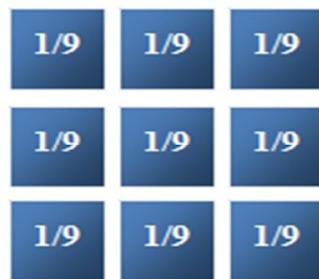
تصویر تیز شده

شکل ۲۲-۲ فیلتر تیز کننده بر روی تصویر

در شکل فوق برای افزایش جزئیات تصویر و به دست آوردن تصویر واضح تر، فیلتر تیزکننده ای را بر روی تصویر سمت چپ اعمال کرده و تصویر سمت راست را تولید کرده ایم.

۲۱-۱ فیلترهای پایین گذر میانگین:

ساده‌ترین نوع فیلترهای پایین گذر فیلترهای میانگین می‌باشند. یک فیلتر میانگین $m * n$ شامل ضرایبی مثبت می‌باشند که همه این ضرایب بر عکس مجموع کل ضرایب فیلتر ضرب می‌شوند به عبارت دیگر همه عناصر ماسک مقداری برابر با $mn/1$ خواهند داشت. به عنوان مثال یک فیلتر میانگین $3 * 3$ به صورت زیر مشخص می‌شود:



شکل ۲۳-۲ فیلتر میانگین $3 * 3$

تصویر آرام شده مذکور با استفاده از همین ماسک صورت گرفته است. برای بلور کردن بیشتر تصویر از ماسک‌های بزرگتری با سایز $n * m$ باید استفاده کرد که مقادیر همه عناصر ماسک برابر با مقدار $mn/1$ خواهد بود. بدیهی است هرچه اندازه ماسک بزرگتر باشد، زمان بیشتری برای پردازش تصویر نیاز خواهد بود. به نظر شما از نظر محاسباتی ۲ بار کانولوت کردن ماسک $3 * 3$ زمان بیشتری را مطلبید یا اعمال کردن ۱ بار ماسک $5 * 5$ ؟ در زیر فیلتر آرام کننده دیگری که به فیلتر پایین گذر گاوسین^۱ مشهور است، نمایش داده شده است:

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

¹ Gaussian

شبه کد زیر نحوه تولید فیلتر آرام کننده گاوسین را نشان می دهد:

```
Function Gaussian( maskWidth, maskHeight )  
Begin  
filter = double[maskHeight, maskWidth]  
For I = 1 to maskWidth Do  
For J = 1 To maskHeight Do  
filter[J,I] = g(I,J)  
End For  
End For  
Return filter  
End
```

همانندیگر فیلترهای حوزه مکانی، فیلتر گاوسین نیز با استفاده از عملگر کانولوشن بر روی تصویر اعمال می شود.

۲-۲۱-۲ فیلتر میانه (Median)

فیلتر پایین گذر دیگری که نسبت به فیلترهای دیگر به زمان پردازش بیشتری نیاز دارد، فیلتر میانه است. فیلتر پایین گذر میانه از یک همسایگی $n * m$ استفاده می کند و روش کار آن نیز به این صورت است که کل همسایگی ها را به صورت صعودی مرتب کرده و عنصر وسط اعداد مرتب شده را انتخاب و جایگزین پیکسل مرکزی می کند. لازم به ذکر است که فیلتر پایین گذر میانه برای حذف نویز فلفل نمکی^۱ می تواند مورد استفاده قرار گیرد. شکل زیر تصویر نویزداری (نویز فلفل نمکی) را نشان می دهد که با استفاده از فیلتر میانه ارتقا یافته است.

¹Salt&Pepper



شکل ۲-۲۴ فیلتر میانه بر روی تصویر

۲-۲۱-۳ فیلترهای تیز کننده حوزه مکانی:

برای انجام تبدیلات در هر دو حوزه مکانی از فیلترها استفاده می‌کنیم. فیلترها در به دو دسته تقسیم می‌شوند:

- فیلترهای پایین گذر^۱: با عبور از پیکسل های با فرکانس کم، بر روی پیکسل های با فرکانس بالا

تغییرات ایجاد می‌کنند.

- فیلترهای بالاگذر^۲: با عبور از پیکسل های با فرکانس بالا، بر روی پیکسل های با فرکانس پایین

تغییرات ایجاد می‌کنند.

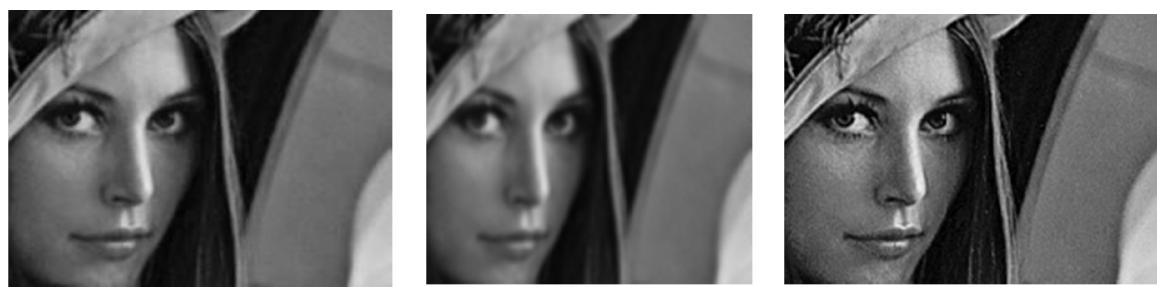
پیکسل فرکانس پایین پیکسلی است که اختلاف شدت روشنایی آن پیکسل با پیکسل های همسایه اش کم باشد (نقاط آرام تصویر). در مقابل پیکسل فرکانس بالا پیکسلی است که اختلاف شدت روشنایی آن پیکسل با پیکسل های همسایه اش زیاد باشد (لبه ها و نویزها).

¹Low Pass Filter

²High Pass Filter

با توجه به مطالب فوق می توان نتیجه گرفت که نتیجه اعمال فیلترهای پایین گذر، تصویری آرام خواهد بود (لبه ها و نویزها تا حدودی آرام می شوند) و با اعمال فیلتر بالاگذر نیز تصویری با جزئیات بیشتر به دست می آید.

شكل زیر تصویری را پس از اعمال فیلتر پایین گذر و بالا گذر نشان می دهد:

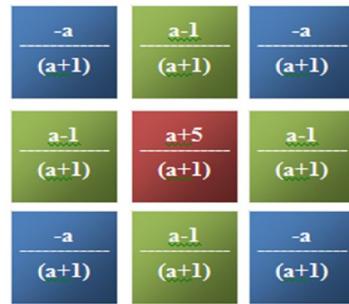


شکل ۲۵-۲ اعمال فیلتر پایین گذر و بالا گذر بر روی تصویر

همانطور که هنگام توضیح عملگر کانولوشن بررسی کردیم، فیلترهای ماسکی را با کانولوت کردن ماسک بر روی تصویر می توان اعمال کرد. در ادامه به بررسی برخی از فیلترهای تیزکننده رایج می پردازیم.

۳-۳-۲۱-۲ فیلتر تیزکننده:

همانطور که بررسی کردیم، فیلترهای تیز کننده برای نمایان کردن بیشتر جزئیات تصویر به کار می روند. ساده ترین فیلتر تیزکننده یک ماسک 3×3 است که مقادیر آن به صورت زیر تعیین می گردند:



شکل ۲-۲۶ فیلتر تیز کننده با یک ماسک 3×3

فیلترهای تیز کننده لاپلاسین دو نوع فیلتر تیزکننده دیگر نیز وجود دارد که از آن ها می توان برای تیز کردن تصاویر استفاده کرد. این فیلتر ها که آن ها را فیلترهای لاپلاسین ترکیبی می نامیم به صورت زیر هستند:



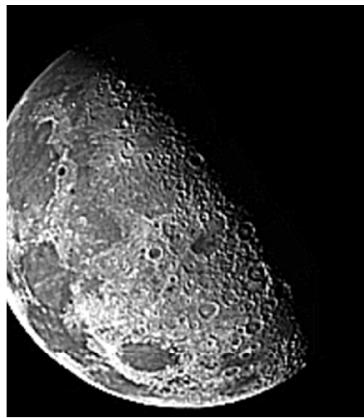
شکل ۲-۲۷ فیلتر لاپلاسین ترکیبی

۴-۳-۲۱-۲ فیلتر تیزکننده لاپلاس

فیلترهای تیزکننده لاپلاس نسبت به روش اول جزئیات بیشتری از تصویر را مشخص می کنند. به عنوان مثال تصاویر زیر نتیجه اعمال فیلترهای تیزکننده مذکور را نشان می دهند:



تصویر اصلی



لابلس ترکیبی ۲



لابلس ترکیبی ۱



فیلتر تیزکننده اول

شکل ۲۸-۲ فیلتر تیزکننده لابلس ۲ بر روی تصویر

۲۲-۲ فراتفکیک پذیری^۱

یکی از تکنیک‌های مطرح شده در بعد نرم‌افزاری، جهت افزایش کیفیت تصویر چه از لحاظ تعداد پیکسل‌ها و چه از لحاظ کاهش مقدار نویز، تکنیک فراتفکیک پذیری (سوپر رزولوشن) می‌باشد. این تکنیک از لحاظ نامگذاری بدلیل آنکه قادر خواهیم بود از محدوده توانایی سیستم تصویر برداری فراتر رویم، فرا

^۱Super resolution

تفکیک پذیری نامیده می‌شود؛ در این تکنیک تلفیق چندین تصویر با رزولوشن پائین تر باعث تولید تصویر نهایی با رزولوشن بالاتر می‌گردد. نکته کلیدی در این تکنیک، استفاده از در هم آمیختگی می‌باشد زیرا که هر تصویر برداشت شده از صحنه مورد نظر تنها بخشی از اطلاعات فرانس بالای صحنه را دریافت نموده است و این اطلاعات فرانس بالا در تمام تصاویر پخش شده است، بنابراین می‌توان از این اطلاعات توزیع شده استفاده نمود و تصویری با رزولوشن و کیفیت بالاتر ایجاد نمود. مطالعات انجام شده نشان می‌دهد که پدیده در هم آمیختگی به دلیل محدود بودن تعداد پیکسل های حسگر های دوربین های دیجیتالی می‌باشد.

سوال بنیادی این می‌باشد که چه عاملی فراتفکیک پذیری را ممکن می‌سازد. این پرسش را توسط مثالی که در ادامه بدان خواهیم پرداخت، توضیح خواهیم داد؛ چنانچه حسگر دوربینی با ابعاد 4×4 از صحنه خاصی تصویر برداری نماید، با افزایش تعداد سلول‌های حسگر دوربین به تعداد 16×16 ، تصویر برداشت شده دارای رزولوشن بیشتری خواهد بود. حال اگر توسط همان حسگر چهار تصویر از یک صحنه یکسان برداشت نمائیم که اختلاف آنها در حد مقداری صحیح از واحد پیکسل باشد، فراتفکیک پذیری ممکن نخواهد بود، ولی چنانچه چهار تصویر دریافتی اختلافی در حد کسری از واحد پیکسل داشته باشند، فراتفکیک پذیری ممکن می‌گردد؛ زیرا که اختلاف چهار تصویر فوق در حد کسری از واحد پیکسل، اطلاعات اضافه‌ای را از صحنه برداشت شده ایجاب می‌کند که پتانسیل افزایش رزولوشن را تقویت می‌نماید [39].

۲۲-۱-پیکربندی تکنیک فراتفکیک پذیری

اکثر روش‌های فراتفکیک پذیری را می‌توان به دو بخش تقسیم نمود: بخش ثبت تصویر^۱ و بخش بازسازی تصویر^۲. دقت بسیار بالایی در بخش ثبت تصویر لازم است (در حد کسری از واحد پیکسل) تا بتوان در بخش بازسازی، تصویری با رزولوشن بالا را بطور صحیح ایجاد نمود. اگر پارامترهای ثبت تصویر بطور غلط تخمین زده شده باشند، معمولاً بهتر است که یکی از تصاویر را توسط روش‌های درونیابی به اندازه مطلوب تغییر دهیم، تا اینکه اطلاعات چندین تصویر را بطور غلط، با یکدیگر تلفیق نماییم. پس از آنکه تصاویر ثبت شدند، جهت بدست آوردن تصویری با رزولوشن بالا از نمونه‌های نمونه برداری شده بصورت غیریکنواخت، یک روش بازسازی تاثیر ناپذیر از نویز (Robust)، لازم می‌باشد. بخش اصلی تر، بخش ثبت تصویر می‌باشد که در تکنیک فراتفکیک پذیری از اهمیت بالایی برخوردار می‌باشد، زیرا که در این بخش تفاوت اصلی میان درونیابی و تکنیک فراتفکیک پذیری بطور آشکار مشخص می‌گردد[40].

۲۳-۲ شکل شناسی:

یکی از مهمترین مراحل در پیش پردازش تصویر، پردازش شکل شناسی می‌باشد. در این قسمت تنها به بررسی شکل شناسی برای تصاویر دودویی خواهیم پرداخت. منظور از تصاویر دودویی، تصاویر با دو سطح روشنایی ۰ یا ۱ می‌باشد که در آن منظور از ۰ رنگ سیاه و منظور از ۱ رنگ سفید می‌باشد.

شکل شناسی عموماً از عملگرهای مجموعه‌ای استفاده می‌کند و از آن بیشتر برای استخراج نقاط کلیدی تصویر، حذف نقاط غیر مفید تصویر و موارد مشابه دیگر استفاده می‌کنیم. در این قسمت ابتدا

¹RegistrationImage

²Image Reconstruction

عملگرهای مجموعه ای پایه را برای پردازش شکل شناسی بررسی می کنیم. سپس به بررسی مهمترین

عملگرهای مجموعه ای خواهیم پرداخت که ویژه پردازش شکل شناسی می باشند.[40]

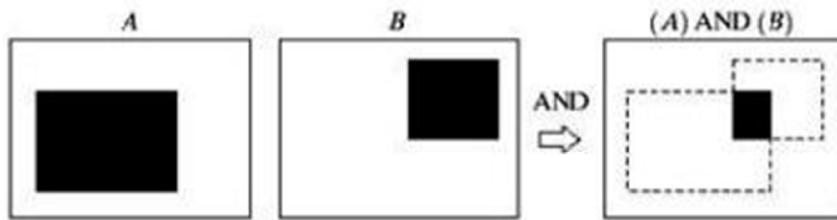
۲۴-۲ عملگرهای مجموعه ای پایه:

این عملگر ها شامل عملگر اجتماع، اشتراک، تفاضل دو مجموعه می باشد. اگر هر تصویر باینری را یک مجموعه در نظر بگیریم، اجتماع دو تصویر باینری هم اندازه، همانطور که در شکل زیر مشاهده می کنید تصویری خواهد بود که در این تصویر هر پیکسلی که در تصویر اول یا تصویر دوم مقدار ۱ داشته باشد، مقدار ۱ خواهد داشت. برای پیاده سازی عملگر اجتماع برای دو تصویر باینری، می توانیم پیکسل های متناظر را در دو تصویر باهم OR بیتی کنیم.



شکل ۲۹-۲ پیاده سازی عملگر اجتماع برای دو تصویر باینری با OR کردن پیکسل های متناظر دو تصویر

اشتراک دو تصویر باینری هم اندازه، نیز تصویری خواهد بود که در این تصویر هر پیکسلی که در تصویر اول و دوم مقدار ۱ داشته باشد، مقدار ۱ خواهد داشت. برای اشتراک دو تصویر باینری هم اندازه، نیز تصویری خواهد بود که در این تصویر هر پیکسلی که در تصویر اول و دوم مقدار ۱ داشته باشد، مقدار ۱ خواهد داشت. برای پیاده سازی عملگر اشتراک برای دو تصویر باینری، می توانیم پیکسل های متناظر را در دو تصویر با هم AND بیتی کنیم.



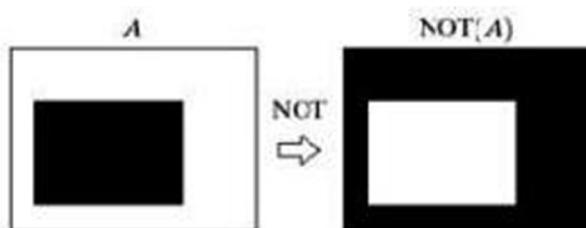
شکل ۳۰-۲ پیاده سازی عملگر اشتراک برای دو تصویر با and کردن پیکسل های متناظر دو تصویر

تفاضل دو تصویر باینری هم اندازه، تصویری خواهد بود که در این تصویر پیکسل هایی از تصویر اول با مقدار ۱ که در تصویر دوم مقدارشان ۱ نباشد، مقدار ۱ خواهند داشت. تفاضل را به شکل اشتراک تصویر دوم و مکمل تصویر اول نیز می توان تعریف کرد.



شکل ۳۱-۲ تفاضل به شکل اشتراک تصویر دوم و مکمل تصویر اول

همچنین عملگر تک عملوندی مکمل نیز عملگری است که پس از اعمال آن بر روی تصویر باینری، در تصویر حاصل شده، مقادیر ۱ به ۰ و مقادیر ۰ به ۱ تغییر می یابند. برای پیاده سازی عملگر مکمل می توانیم هر پیکسل از تصویر بانری را NOT بیتی کنیم.



شکل ۳۲-۲ پیاده سازی عملگر مکمل

۲۵-۲ گسترش باینری مورفولوژیکی:

همانطور که از نام عملگر پیداست، این عملگر باعث گسترش نقاط ۱ در تصویر می‌شود. در عملگر گسترش نیز از یک ماسک (نفاب، پنجره) همانند آنچه که در "تعریف عمل فیلتر کردن" است – استفاده می‌کنیم. در اینجا به جای ماسک به آن عنصر ساختمانی می‌گوییم که مقادیر عنصر ساختمانی ۱ یا صفر می‌تواند باشد.

گسترش تصویر A با عنصر ساختمانی B به صورت زیر تعریف می‌شود:

$$A \oplus B = \{w | reflection(B) \cap A \neq NULL, w \in A\}$$



شکل ۲۳-۲ تصویر اصلی قبل از گسترش باینری مورفولوژیکی

که در اینجا *reflection* عنصر ساختمانی B را حول مرکز خود قرینه می‌کند. به عبارت دیگر گسترش A با عنصر ساختمانی B بدین معنی است که اگر عنصر ساختمانی B را بر روی پیکسل‌های A حرکت دهیم، و در هربار حرکت اشتراک عنصر ساختمانی با محدوده زیر عنصر ساختمانی در تصویر A

نهی نباشد، مقدار پیکسل مرکزی که عنصر ساختمانی بر روی آن قرار گرفته است، برابر ۱ خواهد شد.

شکل زیر خروجی تصویری را پس از گسترش تصویر با عنصر ساختمانی 3×3 تمام ۱ نشان می دهد.[43]



شکل ۳۴-۲ حاصل گسترش تصویر با گسترش باینری مورفولوژیکی

۲۶-۲ سایش باینری مورفولوژیکی:

همانطور که از نام عملگر پیداست، این عملگر باعث سایش نقاط ۱ در تصویر می شود. همانند

عملگر گسترش، در عملگر سایش نیز از یک عنصر ساختمانی استفاده می کنیم که مقادیر عنصر ساختمانی ۱

یا صفر می تواند باشد. به ازای هر پیکسل، مرکز عنصر ساختمانی را روی پیکسل قرار داده و عملگر سایش

را با توجه به مقادیر عنصر ساختمانی در مورد آن پیکسل اعمال می کنیم. سایش تصویر A با عنصر

ساختمانی B به صورت زیر تعریف می شود:

$$A \ominus B = \{w | (B) \subseteq A, w \in A\}$$



شکل ۲-۳۵ تصویر اصلی قبل از سایش باینری مورفولوژیکی

به عبارت دیگر گسترش A با عنصر ساختمانی B بدین معنی است که اگر عنصر ساختمانی B را بر روی پیکسل های A حرکت دهیم، و در هربار حرکت همه نقاطی که در زیر مقادیر ۱ از عنصر ساختمانی قرار گرفته اند نیز، مقدار یک داشته باشند، مقدار پیکسل حاصل حاصل نیز ۱ خواهد بود. شکل زیر خروجی تصویر را پس از سایش تصویر با عنصر ساختمانی 3×3 تمام ۱ نشان می دهد. [42]



شکل ۲-۳۶ حاصل گسترش تصویر با عنصر ساختمانی 3×3

۲۷-۲ بستن و بازکردن باینری مورفولوژیکی:

اعمال عملگر بستن بر روی تصویر باینری موجب می شود تا حفره های کوچک تصویر پرشوند.

به عنوان مثال در تصاویر باینری که انگشت گرفته می شود. به دلیل خشکی پوست دست ممکن است ترک های در تصویر دیده شوند. برای پر کردن این ترک ها می توان از عملگر بستن مورفولوژیکی استفاده کرد.

بستن تصویر از ترکیب عملگر سایش و گسترش و به صورت زیر به دست می آید:

$$A \cdot B = (A \oplus B) \ominus B$$

اعمال عملگر باز کردن بر روی تصویر باینری موجب می شود تا اتصالات باریک تصویر حذف شده و تصویری آرام تر به دست آید. باز کردن تصویر از ترکیب عملگر سایش و گسترش و به صورت زیر به دست می آید:

$$A \circ B = (A \ominus B) \oplus B$$

مجموعه تصاویر زیر نشان می دهد که چطور می توان با ترکیب عملگرهای باز کردن و بستن نویزها و اشکالات یک تصویر را بر طرف کرد. همچنین یکی از پرکاربرترین زمینه هایی که از عملگر های مورفولوژیکی استفاده می شود، در تشخیص کاراکترها می باشد.



شکل ۳۷-۲ تصویر اصلی قبل از اعمال موفولوژیکی



شکل ۲-۳۸ باز کردن مورفولوژیکی در تصویر



شکل ۲-۳۹ بستن مورفولوژیکی در تصویر

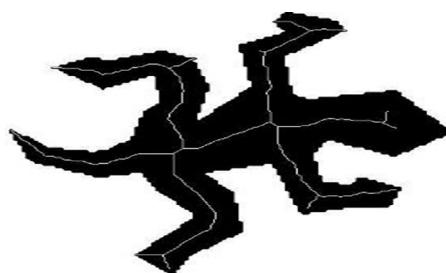
در برخی موارد به علت پایین بودن کیفیت تصویر اسکن شده و یا پایین بودن کیفیت کاغذ مورد نظر در برخی از کاراکتر ها با شکستگی در بدن حرف یا کلمه مواجه هستیم. در چنین مواردی می توان ارتقای تصویر از عملگر های مورفولوژیکی استفاده کرد. توجه داشته باشید که شکل و اندازه عنصر ساختمانی نقش حیاتی در اعمال این عملگر ها بازی می کند. به عنوان مثال تصاویر بالا با اعمال یک عنصر ساختمانی ۳ ۳ تمام ۱ به وجود امده اند. بدیهی است که تغییر اندازه و شکل عنصر ساختمانی نتایج دیگری می تواند تولید کند. [44]

۲۸-۲ استخراج اسکلت بندی اشیاء:

استخراج اسکلت بندی یک الگو به معنی باریک کردن الگو به نحوی است که شکل کلی الگو از بین نمود. از این رو اسکلت یک الگو همانند خود الگو باید قابل تشخیص بوده و توسط آن بتوان شکل کلی الگو را حدس زد. بنابراین اسکلت به دست آمده برای یک الگو باید دارای خصوصیات زیر باشد:

- به اندازه کافی باریک باشد

- متصل باشد: یعنی فرآیند اسکلت بندی نباید باعث تکه شدن اسکلت الگو شود زمانی که این دو شرط برقرار شد، الگوریتم اسکلت بندی متوقف می شود. شکل زیر تصویری را به همراه اسکلت آن نشان می دهد (اسکلت تصویر همان خطوط به هم پیوسته سفید رنگ داخل تصویر است).



شکل ۲۰-۲ اسکلت بندی بر روی تصویر

در حالت کلی الگوریتم های استخراج اسکلت الگو بر پایه دو استراتژی مختلف هستند.

در استراتژی اول از همه پیکسل های تصویر و در استراتژی دوم از پیکسل های کانتور تصویر برای این منظور استفاده می شود.

روشی که ما در اینجا بررسی خواهیم کرد روش مبتنی بر پیکسل های تصویر بوده و به نام Hilditch معروف است که بر روی تصاویر دو سطحی (باینری) اعمال می کنیم.^[45]

تعاریف اولیه:

همسایه های هشتگانه پیکسل P1 را در نظر بگیرید:

P ₉	P ₂	P ₃
P ₈	P ₁	P ₄
P ₇	P ₆	P ₅

شکل ۴۱-۲ روش Hilditch با همسایه های هشتگانه

هدف این است که در مورد حذف پیکسل P1 یا ماندن آن به عنوان یک پیکسل بر روی اسکلت تصمیم بگیریم. برای این منظور همسایه های هشتگانه آن را در جهت عقربه های ساعت مرتب کرده و دو

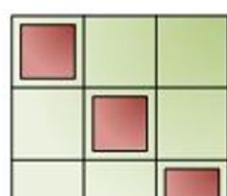
تابع زیر را برای P1 تعریف می کنیم:

• تابع B: تعداد همسایه های غیر صفر P1 را برمی گرداند.

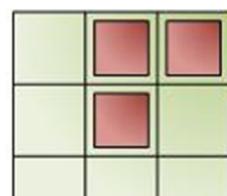
• تابع A: تعداد الگوهای ۱،۰ در لیست P2,P3,P4,P5,P6,P7,P8,P9,P2 را برمی گرداند. الگوی

۱ بدين معنی است که با در نظر گرفتن دو همسایه پشت سرهم مثل P6 و P7، مقدار P6 صفر بوده و

مقدار ۱ داشته باشد. به عنوان مثال:



$$B(P_1) = 2, A(P_1) = 2$$



$$B(P_1) = 2, A(P_1) = 1$$

۲۹-۲ الگوریتم اسکلت بندی:

الگوریتم Hilditch از یک پنجره 3×3 استفاده کرده و تازمانی تغییری در تصویر به وجود آید، به طور مداوم همه پیکسل های تصویر را بررسی می کند و به ازای هر پیکسل در صورتی که چهار شرط زیر برقرار شود، پیکسل جاری از تصویر حذف می گردد:

$$\begin{array}{ll} 2 \leq B(p_1) \leq 6 & \bullet \\ A(p_1) = 1 & \bullet \\ A(p_2) <> 1 \text{ یا } p_2 \cdot p_4 \cdot p_8 = 0 & \bullet \\ A(p_4) <> 1 \text{ یا } p_2 \cdot p_4 \cdot p_6 = 0 & \bullet \end{array}$$

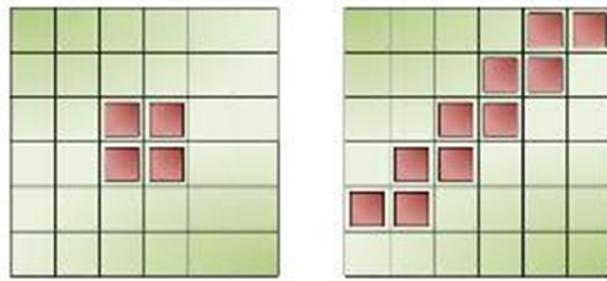
شبه کد زیر نحوه اعمال این الگوریتم را نشان می دهد:

۱-تا زمانی که تغییر دیگری حاصل نشود:

a. کل تصویر پیمایش شده و پیکسل هایی که هر چهار شرط فوق برای آنها صادق باشد کاندیدای حذف شدنی شوند.

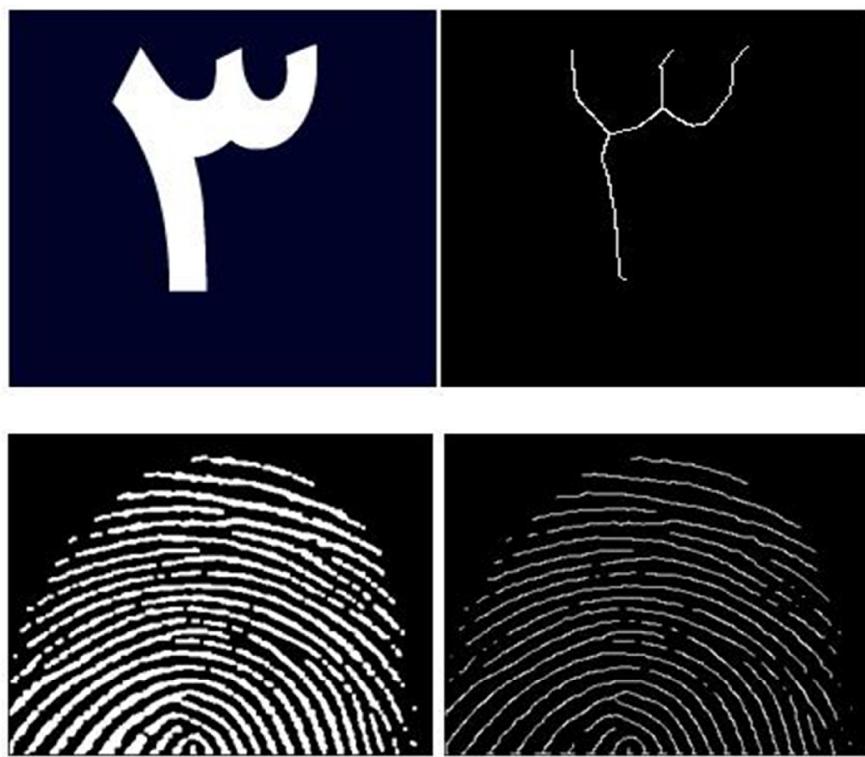
b. پیکسل های کاندیدای حذف شدن، از تصویر حذف می شوند.

توجه داشته باشید که الگوریتم Hilditch کاملی نمی باشد چرا که بر روی برخی از انواع الگوها به درستی عمل نمی کند. به عنوان مثال الگوریتم Hilditch باعث ساییده شدن کامل الگوهای زیر می شود:



در زیر اسکلت استخراج شده برای تعدادی تصویر با استفاده از الگوریتم Hilditch نشان داده شده

است:



شکل ۴۲-۲ اسکلت استخراج شده برای تعدادی تصویر با استفاده از الگوریتم Hilditch

۳۰-۲ نویز در تصویر:

نویزها باعث می شوند که سطح جزئیات ثبت شده در عکس دیجیتالی یا فیلمی کاهش یابد، بنابراین کاهش دادن این نویزها سبب می شود که کیفیت عکس نهایی یا خروجی چاپی افزایشیابد. مشکلی که وجود دارد اینست که تکنیک های رایج برای کاهش یا حذف نویز تصاویر به همان اندازه منجر به سافت شدن تصویر می شود. سافت شدن تصویر در عکس هایی که از موضوعاتی مانند آسمان یا سطح آب گرفته شده اند مشکل چندانی ایجاد نمی کند اما مثلا در تصویر یک منظره از درختان حتی تلاش های محتاطانه برای کاهش نویز موجب صدمه دیدن جزئیات شاخ و برگ های درختان موجود در عکس می شود.

نویزها در عکاسی دیجیتال واسکن انواعی دارند که دو مورد آنها از بقیه شایعتر هستند **Random** و **Fixed Pattern Noise**.

نویزهای اتفاقی غالبا در ایزوهای بالا دیده می شوند و همانگونه که از اسم شان مشخص است، محل ثابتی ندارند. یعنی اگر دو یا چند عکس با شرایط یکسان و با ایزو بالا تهیه شود، محل نویزها در عکس ها مختلف متفاوت است. اما نویزهای ثابت غالبا در نوردهی های طولانی و موارد دیگری که با گرم شدن سنسور همراه است، دیده می شوند.

حذف نویزهای ثابت راحت تر است چون محل آنها ثابت است و اگر یک فریم سیاه(با بستن دربوش لنز) بگیریم و تفاصل آن را با عکس اصلی تهیه کنیم، نویزها حذف خواهند شد. کاری که برخی دوربین ها به طور خودکار نیز انجام می دهند. اما چالش اصلی در حذف نویزهای اتفاقی است؛ چه در دوربین، اسکنر و چه در نرم افزار. غالباً این روش ها منجر به کاهش جزئیات عکس و یا ایجاد آثار جانبی،

بخصوص در لبه‌های اجسامی گرددند. در واقع کاربر باید بین این مشکلات و نویز، تعادلی برقرار نماید و تصویرنهايی، حاصل اين تصميم کاربر خواهد بود.

۱-۳۰-۲ کاهش نویز با Neat Image

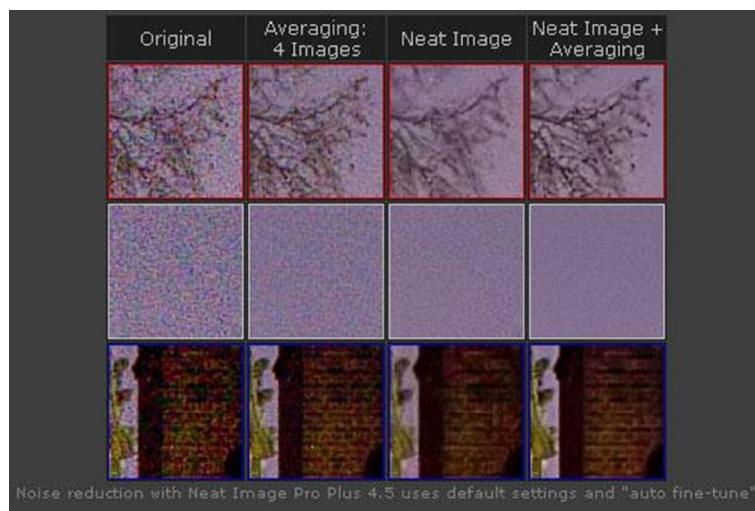
برنامه های کاهش نویز مانند Neat Image بهترین بازار های آماده موجود برای کاهش نویز هستند لذا در مقایسه بعدی از آن به عنوان معیار استفاده می کنیم:



شکل ۴۳-۲ کاهش نویز با Neat Image

همانطور که دیده می شود Neat Image بهترین عملکردها در کاهش نویز در بخش آسمان صاف داشته است اما در دو بخش دیگر استفاده از آن باعث از بین رفتن جزئیات ریز در شاخه های درخت و خطوط بند کشی آجرهای دیوار شده است. شارپ کردن عکس می تواند باقیمانده جزئیات عکس را ارتقا دهد و در بهبود شفافیت عکس ونمایش کلی آن عملکرد تاثیرگذاری داشته باشد اما نمی تواند جزئیات از دست رفته را باز گردد.

نیز یک تکنیک ابتدایی است که در اکثر نسخه های فتوشاپ موجود است. طرز کار Median Filter این فیلتر به این صورت است که حجم هر پیکسل را با محاسبه میانگین حجم پیکسل های همسایه آن محاسبه و اعمال می کند. این فیلتر برای رفع نویزهای خیلی ریز مناسب است اما موجب ایجاد نوسانات بیشتری در عکس می شود و جزئیات در سطح پیکسل ها را نابود می کند. بهر حال استفاده از Neat Image بهترین راه حل برای عکس هایی است که نمی توان روی آنها روش میانگین گیری را اجرا کرد. اما برای کسب بهترین نتیجه می توان دو تکنیک را به این صورت استفاده کرد: استفاده از تکنیک میانگین گیری برای افزایش SNR عکس تا آنجا که ممکن است و سپس استفاده از Neat Image برای کاهش نویز های باقیمانده:



دیده می شود که چگونه استفاده از تکنیک میانگین گیری همراه با Neat Image باعث شده است که علاوه بر ارتقای نمایش جزئیات در قسمت دیوار آجری، شفافیت آن آسیب ندیده و حفظ شود و نویزها نیز کمتر به چشم بیایند.

معایب تکنیک میانگین گیری هم افزایش حجم فایل عکس و طولانی تر شدن زمان عکسبرداری است (زیرا باید چندین عکس یکسان از یک سوژه گرفته شود). در ضمن این تکنیک برای عکس هایی که دارای

هستند کاربردیندارد، اگر دقت کنید می بینید که پیکسل براق fixed pattern noise یا banding noise

سفیدی که در سمت چپ پایین عکس وجود دارد اثر میانگین گیری از بین نرفته است.

برای تهیه عکسهای مورد نیاز برای میانگین گیری، دوربین چه در حین زمان نوردهی برای گرفتن عکس و چه در زمان بین گرفتن یک عکستا عکس بعدی، نباید کوچکترین حرکتی داشته باشد و عکس های تهیه شده از سوژه باید دقیقاً یکسان و منطبق بر هم باشند.^[46]

فصل سوم

۱-۳ مقدمه

نیاز به سیستم های اطلاعاتی بهبود یافته بیشتر از قبل مورد توجه قرار گرفته است چرا که اطلاعات، عنصری اساسی در تصمیم گیری است و جهان در حال افزایش دادن مقدار اطلاعات در فرم های مختلف با درجه هایی از پیچیدگی است. یکی از مسائل اصلی در طراحی سیستم های اطلاعاتی مدرن، تشخیص الگو به طور اتوماتیک است. تشخیص به عنوان یک صفت اصلی انسان بودن است. یک الگو، توصیفی از یک شیء است. یک انسان دارای یک سیستم اطلاعاتی سطح بالاست که یک دلیل آن داشتن قابلیت تشخیص الگوی پیشرفته است. بر طبق طبیعت الگوهای مورد تشخیص، عملیات تشخیص در دو گونه‌ی اصلی تقسیم می‌شوند:

- تشخیص آیتم های واقعی: این ممکن است به عنوان تشخیص سنسوری معرفی شود که تشخیص الگوهای سمعی و بصری را دربر می‌گیرد. این فرآیند تشخیص، شناسایی و دسته بندی الگوهای فضایی و الگوهای زمانی را در بر می‌گیرد. مثال هایی از الگوهای فضایی کارکترها، اثر انگشت ها، اشیاء فیزیکی و تصاویر هستند. الگوهای زمانی شامل فرم های موجی گفتار، سری های زمانی و ... هستند.
- تشخیص آیتم های انتزاعی: از طرف دیگر، یک بحث قدیمی یا یک راه حل برای مسئله، قابل تشخیص است. این فرآیند شامل تشخیص آیتم های انتزاعی است و می‌توان آن را تشخیص انتزاعی خواند.

در واقع تشخیص انسان، یک پرسش تخمین زدن احتمال نسبی ای است که یک داده‌ی ورودی به یکی از مجموعه‌های جمعیت‌های آماری شناخته شده، که وابسته به تجربیات گذشته است، می‌تواند مرتبط باشد. بنابراین مسئله‌ی تشخیص الگو ممکن است به عنوان یک مسئله‌ی تشخیص اختلاف داده‌ی ورودی با جمعیت‌ها به وسیله‌ی جستجوی ویژگی‌ها یا صفات ثابت در میان اعضای یک جمعیت مورد توجه قرار گیرد.

چهره نقش اساسی را در شناسایی افراد و نمایش احساسات آنها در سطح جامعه دارد. توانایی انسان در تشخیص چهره‌ها قابل توجه است ما می‌توانیم هزاران چهره‌ی یاد داده شده در طول عمرمان را تشخیص دهیم و در یک نگاه چهره‌های آشنا را حتی پس از سالها جدایی شناسایی کنیم. این مهارت در مقابل تغییرات در شرایط دیداری مانند حالت چهره، سن و همچنین تغییراتی در عینک، ریش یا سبک مدل موها ایستادگی می‌کند. تشخیص چهره یک موضوع مهم در کاربردهایی همچون سیستم‌های امنیتی، کترل کارت اعتباری و شناسایی مجرمان شده است. برای مثال، قابلیت مدل کردن یک چهره‌ی خاص و تمیز دادن آن از یک تعداد فراوان از مدل‌های چهره‌ی ذخیره شده، شناسایی مجرمان را به صورت گسترده‌ای بهبود خواهد بخشید. اگرچه درست است که انسان‌ها در تشخیص چهره توانا هستند اما نحوه‌ی کدینگ و دی کدینگ چهره‌ها در مغز انسان کاملاً آشکار نیست. تشخیص چهره‌ی انسان برای بیش از بیست سال مورد مطالعه قرار گرفته است. توسعه‌ی یک مدل محاسباتی برای تشخیص چهره کاملاً دشوار است و دلیل آن پیچیدگی چهره‌ها و ساختار چند بعدی بینایی است. بنابراین تشخیص چهره یک فعالیت سطح بالا در بینایی کامپیوتر است و می‌تواند بسیاری از تکنیک‌های بینایی اولیه را در بر گیرد.

۲-۳ تشخیص الگو و تشخیص چهره

چهره نقش اساسی را در شناسایی افراد و نمایش احساسات آنها در سطح جامعه دارد. توانایی انسان در تشخیص چهره ها قابل توجه است ما می توانیم هزاران چهره‌ی یاد داده شده در طول عمرمان را تشخیص دهیم و در یک نگاه چهره‌های آشنا را حتی پس از سالها جدایی شناسایی کنیم. این مهارت در مقابل تغییرات در شرایط دیداری مانند حالت چهره، سن و همچنین تغییراتی در عینک، ریش یا سبک مدل موها ایستادگی می کند.

تشخیص چهره یک موضوع مهم در کاربردهایی همچون سیستم های امنیتی، کترل کارت اعتباری و شناسایی مجرمان شده است. برای مثال، قابلیت مدل کردن یک چهره‌ی خاص و تمیز دادن آن از یک تعداد فراوان از مدل‌های چهره‌ی ذخیره شده، شناسایی مجرمان را به صورت گسترده‌ای بهبود خواهد بخشید.

اگرچه درست است که انسان‌ها در تشخیص چهره توانا هستند اما نحوه‌ی کدینگ و دی کدینگ چهره‌ها در مغز انسان کاملاً آشکار نیست. تشخیص چهره‌ی انسان برای بیش از سال مورد مطالعه قرار گرفته است. توسعه‌ی یک مدل محاسباتی برای تشخیص چهره کاملاً دشوار است و دلیل آن پیچیدگی چهره‌ها و ساختار چند بعدی بینایی است. بنابراین تشخیص چهره یک فعالیت سطح بالا در بینایی کامپیوتر است و می‌تواند بسیاری از تکنیک‌های بینایی اولیه را در بر گیرد.

مرحله‌ی اول تشخیص چهره‌ی انسان، استخراج ویژگی‌های آشکار از تصاویر چهره‌هاست. در اینجا یک سوال بوجود می‌آید که تا چه اندازه ویژگی‌های چهره قابلیت اندازه‌گیری شدن را دارند.

بررسی های محققین در چندین سال گذشته بر آن اشاره دارد که ویژگیهای خاصی از چهره برای شناسایی چهره ها توسط انسان تشخیص داده می شود.

این بخش بر روی پایه های اصلی تشخیص الگو و چهره تمرکز دارد.

نگاه کلی

نیاز به سیستم های اطلاعاتی بهبود یافته بیشتر از قبل مورد توجه قرار گرفته است چرا که اطلاعات عصری اساسی در تصمیم سازی است و جهان در حال افزایش دادن مقدار اطلاعات در فرم های مختلف با درجه هایی از پیچیدگی است . یکی از مسائل اصلی در طراحی سیستم های اطلاعاتی مدرن، تشخیص الگو به طور اتوماتیک است.

تشخیص به عنوان یک صفت اصلی انسان بودن است . یک الگو، توصیفی از یک شیء است . یک انسان دارای یک سیستم اطلاعاتی سطح بالاست که یک دلیل آن داشتن قابلیت تشخیص الگوی پیشرفتی است . بر طبق طبیعت الگوهای مورد تشخیص، عملیات تشخیص در دو گونه‌ی اصلی تقسیم می شوند:

- تشخیص آیتم های واقعی :

این ممکن است به عنوان تشخیص سنسوری معرفی شود که تشخیص الگوهای سمعی و بصری را دربر می گیرد . این فرآیند تشخیص، شناسایی و دسته بندی الگوهای فضایی و الگوهای زمانی را در بر می گیرد . مثال هایی از الگوهای فضایی کارکترها، اثر انگشت ها، اشیاء فیزیکی و تصاویر هستند . الگوهای زمانی شامل فرم های موجی گفتار، سری های زمانی و ... هستند.

- تشخیص آیتم های انتزاعی :

از طرف دیگر، یک بحث قدیمی یا یک راه حل برای مسئله، قابل تشخیص است . این فرآیند شامل تشخیص آیتم های انتزاعی است و می توان آن را تشخیص انتزاعی خواند.

در واقع تشخیص انسان، یک پرسش تخمین زدن احتمال نسبی ای است که یک داده ی ورودی به یکی از مجموعه های جمعیت های آماری شناخته شده، که وابسته به تجربیات گذشته است، می تواند مرتبط باشد . بنابراین مسئله ی تشخیص الگو ممکن است به عنوان یک مسئله ی تشخیص اختلاف داده ی ورودی با جمعیت ها به وسیله ی جستجوی ویژگی ها یا صفات ثابت در میان اعضای یک جمعیت مورد توجه قرار گیرد [47] .

۱-۲-۳ الگوها و کلاس های الگوها

تشخیص الگو می تواند به عنوان دسته بندی داده ها ی ورودی در کلاس های شناخته شده به وسیله ی استخراج ویژگیهای مهم یا صفات داده تعریف شود.

یک کلاس الگو، یک دسته ی متمایز شده به وسیله ی برخی صفات و ویژگی های مشترک است . ویژگی های یک کلاس الگو، صفات نوعی هستند که بین همه ی الگوهای متعلق به آن کلاس مشترک هستند . این ویژگی ها اغلب به عنوان ویژگی های اینتراست شناخته می شوند . ویژگی هایی که تفاوت های بین کلاس های الگو را بیان می کنند اغلب به عنوان ویژگی های اینترست شناخته می شوند.

یک الگو، توصیفی از یکی از اعضای دسته است که ارائه دهنده ی کلاس الگو می باشد . برای راحتی، الگوها معمولاً به وسیله ی یک بردار نمایش داده می شوند.

مانند:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix}$$

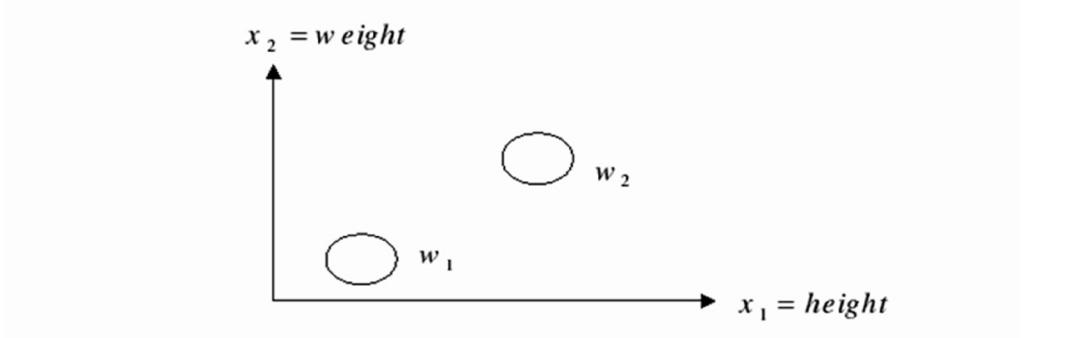
که در آن عنصر x_i یک ویژگی آن الگو را ارائه می کند. اغلب مفید است که یک بردار الگو به عنوان یک نقطه در فضای اقلیدسی n بعدی دیده شود.

۲-۲-۳ مسائل اساسی در طراحی سیستم تشخیص الگو

به طور کلی طراحی یک سیستم تشخیص الگو چندین مسئله ای اصلی را در بر می گیرد:

• اول از همه، ما بایستی در مورد نمایش داده های ورودی تصمیم بگیریم.

چرا که بایستی اثیائی که قرار است مورد تشخیص قرار گیرند، اندازه گیری شوند. این یک مسئله sensing است. هر کمیت اندازه گیری شده یک ویژگی الگو یا شیء را توصیف می کند و مجموعه ای این ویژگی ها در یک بردار قرار می گیرند. این بردار الگو که داده ای ورودی را توصیف می کند، بایستی قالب بندی شود. بردارهای الگو همه یا بخشی از اطلاعات اندازه گیری شده موجود را در مورد الگوهای در بر می گیرد. مجموعه ای الگوهای متعلق به کلاسی یکسان به گروهی از نقاط در یک فضای اندازه گیری شده نگاشت می شوند. یک مثال ساده از این مورد در شکل ۱ برای دو کلاس W_1 و W_2 نشان داده شده است.



شکل ۳-۱ دو کلاس مجزا بردار الگو از دو کمیت $height$ و $weight$ تشکیل شده است.

- دومین مسئله در تشخیص الگو، استخراج ویژگیها یا صفات خاصی از داده‌ی ورودی دریافته شده و کاهش ابعاد بردارهای الگوست. این مورد اغلب به عنوان مسئله‌ی پیش‌پردازش و استخراج ویژگی معرفی می‌شود. عناصر ویژگیهای (ایتراست) برای همه‌ی کلاس‌های الگو مشترک هستند می‌توانند حذف شوند.

اگر یک مجموعه‌ی کامل از ویژگیهای تشخیص برای هر کلاس از داده‌های اندازه‌گیری شده تعیین شود. تشخیص و دسته‌بندی الگوهای دشواری کمتری را در برخواهد داشت. تشخیص اتوماتیک ممکن است به یک فرآیند تطبیق ساده یا یک جدول جستجو کاهش یابد. به هر حال در بسیاری از مسائل تشخیص الگو، در عمل، تعیین یک مجموعه کامل از ویژگیهای تشخیص اگر غیرممکن نباشد دشوار است.

- مسئله‌ی سوم در طراحی سیستم تشخیص الگو تعیین رویه‌های تصمیم‌گیری است که در فرآیند شناسایی و دسته‌بندی مورد نیاز واقع می‌شود. پس از آنکه داده‌های مشاهده شده از الگوهای جمع آوری شد و در فرم نقاط الگو یا بردارهای اندازه‌گیری در فضای الگو بیان شد، ما ماشینی را می‌خواهیم تا تصمیم بگیرد که این داده به کدام کلاس الگو تعلق دارد. اجازه دهید سیستم قادر به شناسایی M کلاس

مختلف الگو باشد . در این صورت فضای الگو می تواند شامل M ناحیه باشد که هر کدام نقاط الگوی

یک کلاس را در بر می گیرد . مسئله‌ی تشخیص حالا می تواند به عنوان تولید کننده‌ی محدوده‌های

تصمیم که M کلاس را روی بردارهای اندازه گیری از همدیگر جدا می کند در نظر گرفته شود.

به طور کلی، این محدوده‌های تصمیم به وسیله‌ی توابع تصمیم تعیین می شوند.

۳-۲-۳ یادگیری و تمرین دادن

توابع تصمیم به روش‌های متنوعی قابل تولید هستند . زمانی که دانش قبلی در مورد الگوهایی که

بایستی تشخیص داده شوند، موجود باشد، تابع تصمیم براساس این اطلاعات ممکن است با دقت تعیین

شود . زمانی که تنها دانشی کیفی در مورد الگوها موجود باشد، حدس‌هایی مستدل از فرم‌های تابع تصمیم

می توان داشت . در این مورد محدوده‌های تصمیم ممکن است از پاسخ صحیح دور شود . وضعیت کلی

تر آنست که دانش قبلی کمی در مورد الگوهای مورد تشخیص موجود باشد . در این شرایط ماشین‌های

تشخیص الگو با استفاده از یک رویه‌ی یادگیری یا تمرین دادن طراحی بهتری خواهند داشت . به صورت

ابتداً، توابع تصمیم وقت فرض می شوند و از طریق دنباله‌ای از مراحل تمرینی تکراری، این توابع

تصمیم به سمت فرم‌های بهینه و راضی کننده پیش می روند.

این مهم است به ذهن بسپاریم که تمرین و یادگیری فقط در طول فاز طراحی سیستم تشخیص الگو

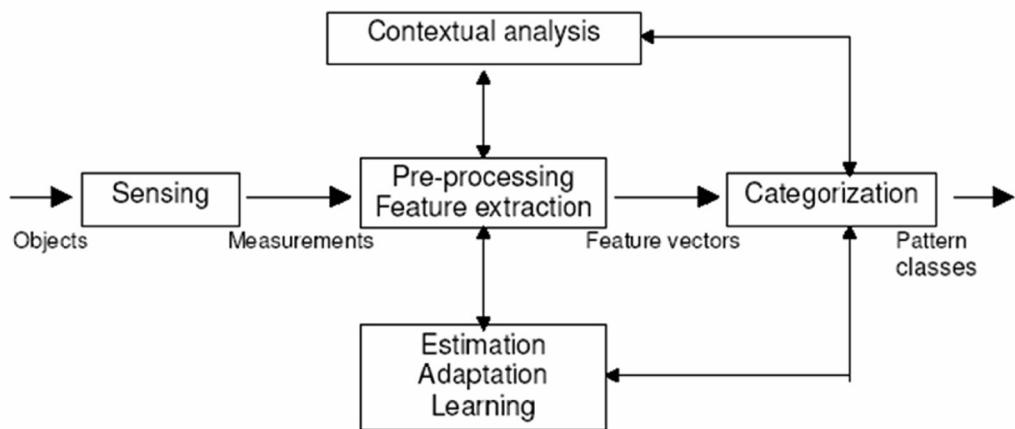
انجام می شوند. هنگامی که نتایج قابل قبول با مجموعه‌ی الگوهای تمرینی به دست آمد، سیستم برای

وظیفه‌ی اجرائی واقعی خود بر روی نمونه‌های محیطی به کار گرفته می شود . کیفیت کارآئی تشخیص به

طور گستردۀ ای به وسیله‌ی تشابه الگوهای تمرینی و داده‌های واقعی که سیستم در طول عملیات مواجه خواهد شد، تعیین می‌شود.

۴-۲-۳ کلیات یک سیستم تشخیص الگو

در شکل ۲-۳، دیاگرام بلوکی یک سیستم تشخیص الگوی تطبیقی نشان داده می‌شود.



شکل ۲-۳ : دیاگرام بلوکی یک سیستم تشخیص الگوی تطبیقی

تشخیص صحیح به میزان اطلاعات موجود در اندازه‌گیری‌ها و نحوه‌ی استفاده از این اطلاعات وابسته خواهد بود . در برخی کاربردها، اطلاعات زمینه برای بدست آوردن تشخیص دقیق الزامی است . برای نمونه، در تشخیص کارکترهای دست نویس خمیده و دسته بندي اثر انگشت‌ها، اطلاعات زمینه با اهمیت هستند . زمانیکه ما قصد طراحی یک سیستم تشخیص الگو داریم که در مقابل distortion مقاوم باشد همچنین در مقابل large pattern deviations پذیر باشد و نیز قابلیت خود تنظیمی را داشته باشد، ما با مسئله‌ی تطبیق دادن مواجه شده‌ایم [48].

۳-۳ تشخیص چهره

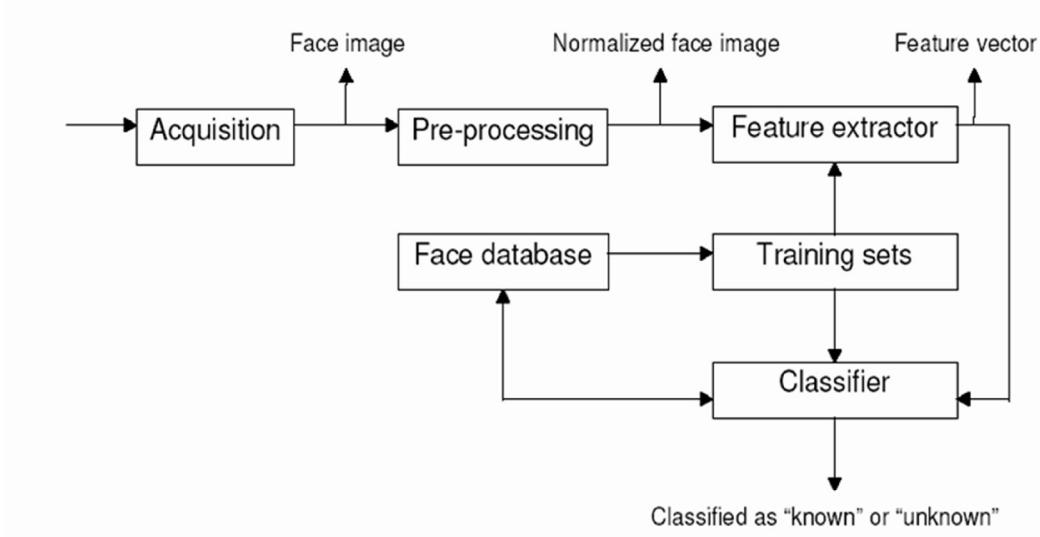
تشخیص چهره یک عمل تشخیص الگو است که به طور خاص بر روی چهره ها انجام می شود.

این عمل عبارتست از دسته بندی یک چهره به عنوان «شناخته شده» و «ناشناس» پس از مقایسه با چهره های ذخیره شده توسط افراد شناخته شده.

مدل های محاسباتی تشخیص چهره بایستی پاسخگوی چندین مسئله ی دشوار باشند. این سختی از آنجا ناشی می شود که چهره ها بایستی به گونه ای ارائه شوند که اطلاعات موجود در چهره را برای تشخیص یک چهره ی خاص از دیگر چهره ها به بهترین نحو بکار ببرند. در این مورد، چهره ها باعث یک مسئله ی دشوار می شوند چرا که همه ی چهره ها از آنجا که مجموعه ویژگی های یکسانی مانند چشمان، بینی و دهان دارند، مشابه با یکدیگر هستند.

۱-۳-۳ کلیات یک سیستم تشخیص چهره عام

در شکل ۳-۳، کلیات یک سیستم تشخیص چهره ی عام نشان داده شده است. این کلیات، ویژگیهایی از سیستم تشخیص الگوی عام که قبل از بحث شد، در بر دارد.



شکل ۳-۳: یک سیستم تشخیص چهره‌ی عام

شش بلوک عملیاتی اصلی وجود دارند که مسئولیت‌های آنها در زیر آمده است:

۱-۳-۱ مژول دریافت : این نقطه‌ی ورود به فرایند تشخیص چهره است . این مژولی است که تصویر چهره‌ی مورد نظر به سیستم ارائه می شود . به عبارت دیگر، در این مژول، از کاربر درخواست می شود تا یک تصویر چهره به سیستم تشخیص چهره ارائه کند . یک مژول دریافت می تواند تصویر چهره را از چندین محیط متفاوت تحویل بگیرد : یک تصویر چهره ممکن است یک فایل تصویر بر روی یک دیسک مغناطیسی باشد، ممکن است به وسیله‌ی یک frame grabber گرفته شده باشد یا به کمک یک اسکنر از روی کاغذی اسکن شده باشد.

۲-۳-۱ مژول پیش - پردازش : در این مژول، بوسیله تکنیک‌های بینایی ، تصاویر چهره نرم‌السازی می شوند و اگر نیاز بود، برای افزایش کارائی تشخیص سیستم بهبود می یابند . برخی با تمام مراحل پیش - پردازش زیر ممکن است در یک سیستم تشخیص چهره پیاده سازی شوند :

۱-۳-۲-۱ نرمال سازی اندازه‌ی تصویر: این کار معمولاً انجام می‌شود تا اندازه‌ی تصویر گرفته

شده به یک اندازه تصویر پیش فرض مانند 128×128 تغییر کند، این اندازه‌ی تصویر همان اندازه‌ی تصویری است که سیستم تشخیص چهره با آن کار می‌کند.

۲-۳-۱-۲ یکسان سازی هیستوگرام : این معمولاً بر روی تصاویر خیلی تاریک یا خیلی روشن برای

بهبود کیفیت تصویر و بهبود کارآئی تشخیص چهره انجام می‌شود که گستره‌ی contrast تصویر را اصلاح می‌کند که به عنوان نتیجه، برخی ویژگیهای مهم چهره آشکارتر می‌شوند.

۳-۱-۳-۲ فیلترینگ میانه : برای تصاویر دارای نویز، به خصوص تصاویری که از طریق دوربین

عکاسی و یا frame grabber گرفته شده است، فیلترینگ میانه می‌تواند تصویر را بدون از دست رفتن اطلاعات تمیز کند.

۴-۱-۳-۲ فیلترینگ بالا گذر: استخراج کننده‌های ویژگی ای که مبتنی بر روی کلیات چهره هستند،

ممکن است از تصویری که لبه یابی روی آن صورت گرفته است نتیجه‌ی بهتری بدeneد. فیلترینگ بالا گذر بر روی جزئیاتی مانند لبه‌ها تاکید دارند که در نتیجه کارآئی تشخیص لبه را افزایش می‌دهد.

۵-۱-۳-۲ حذف کردن پس زمینه : به منظور دست یابی به خود اطلاعات چهره، پس زمینه‌ی چهره

قابل حذف است. این موضوع برای سیستم‌های تشخیص چهره‌ای مهم‌تر است که اطلاعات موجود در کل تصویر، استفاده می‌شود. همچنین توجه داشته باشیم که در حذف پس زمینه، مازول پیش-پردازش بایستی قادر به تعیین محدوده‌ی چهره باشد.

۶-۱-۳-۲ نرمال سازی چرخشی و انتقالی : در برخی موارد، ممکن است بر روی تصویر چهره‌ای

کار شود که سر شخص، چرخش یا انتقال یافته باشد. سر، نقش کلیدی در تعیین ویژگیهای چهره بازی

می کند . به خصوص برای سیستم های تشخیص چهره ای که مبتنی بر چهره های تمام رخ هستند، ممکن است مطلوب باشد که ماژول پیش - پردازش وجود چرخش یا انتقال را مشخص کند و اگر ممکن بود، به نرمال سازی انتقال ها و چرخش ها در مکان سر بپردازد.

۷-۲-۱-۳-۳ نرمال سازی نور پردازی : تصاویر چهره‌ی گرفته شده در شرایط نوری متفاوت می‌توانند کارآئی تشخیص چهره را پایین بیاورند به خصوص برای سیستم های تشخیص چهره‌ی مبتنی بر آنالیز جزء اصلی که اطلاعات کل تصویر برای تشخیص مورد استفاده قرار می‌گیرند.

۳-۱-۳-۳ ماژول استخراج ویژگی : پس از انجام پیش - پردازش (اگر لازم بود)، تصویر چهره‌ی نرمال شده به منظور پیدا کردن ویژگیهای کلیدی ای که قرار است برای دسته بندی از آنها استفاده شود، به ماژول استخراج ویژگی فرستاده می‌شود . به عبارت دیگر، این ماژول مسئول ایجاد یک بردar ویژگی است که به حوبی تصویر چهره را ارائه کند.

۴-۱-۳-۳ ماژول دسته بندی : در این ماژول، با کمک یک دسته بند الگو، ویژگیهای استخراج شده از تصویر چهره با آنها مقایسه کرده و این دسته بندی را ارائه می‌کند. پس از انجام این مقایسه، تصویر چهره به عنوان «شناخته شده» و «ناشناخته شده» دسته بندی می‌شود .

۳-۱-۵-۳ مجموعه‌ی تمرین : مجموعه‌های تمرینی در طول «فاز یادگیری» فرآیند تشخیص چهره استفاده می‌شوند . ماژول های استخراج ویژگی و دسته بندی به منظور بدست آوردن کارآئی بهینه‌ی تشخیص به وسیله‌ی استفاده از مجموعه‌های تمرینی، پارامترهایشان را تنظیم می‌کنند.

۳-۱-۶ پایگاه داده‌ی چهره : پس از آنکه یک چهره به عنوان « ناشناخته » دسته بندی شد.

تصاویر چهره می‌توانند با بردارهای ویژگی شان برای مقایسه‌های بعدی به پایگاه داده اضافه شوند.

ماژول دسته بندی به طور مستقیم از پایگاه داده‌ی چهره استفاده می‌کند.

ماژول‌های استخراج ویژگی و دسته بندی، دو ماژولی هستند که وجه تمایز بین بیشتر سیستم‌های

تشخیص چهره‌ی متفاوت را باعث می‌شوند. تقریباً می‌توان گفت که تفاوت کارائی سیستم‌های

تشخیص چهره، در امر تشخیص چهره‌ها از کارائی این دو ماژول آن‌ها متاثر است. در بخش‌های بعدی

به چهار روش از روش‌های متفاوتی که سیستم‌های تشخیص چهره در این دو ماژول به کار می‌برند، می-

پردازیم [49].

۴-۳ روش‌های استخراج ویژگی در تشخیص چهره

۱-۴-۳ تشخیص چهره با استفاده از چهره - ویژه‌ها

به زبان تئوری اطلاعات، در تشخیص چهره اطلاعات داخل یک تصویر چهره، با حداکثر کارآئی

استخراج و کدگذاری می‌شوند و سپس با یک پایگاه داده از مدل‌هایی که به صورت مشابه کدگذاری شده

اند مقایسه می‌شوند. یک رویکرد ساده برای استخراج اطلاعات موجود در یک تصویر چهره، به دست

آوردن اختلاف در مجموعه‌ی تصاویر چهره‌های سپس می‌توان با استفاده از این اطلاعات به کدگذاری

و مقایسه‌ی تصاویر چهره‌ی افراد پرداخت.

به زبان ریاضی، اجزای اصلی توزیع چهره‌ها، یا بردار ویژه‌های ماتریس کواریانس مجموعه‌ی

تصاویر چهره، به گونه‌ای رفتار می‌کنند که یک تصویر به عنوان یک نقطه (یا بردار) در یک فضای با

تعداد ابعاد بسیار بالا در نظر گرفته می‌شود.

بردار ویژه ها می توانند به عنوان مجموعه ای از ویژگیها فرض شوند که باهم اختلاف بین تصاویر چهره ها را مشخص می کنند . در اینجا چون بردار ویژه ها در مورد تصاویر چهره هستند، چهره ویژه نامیده می شوند.

نمونه ای از تصاویر چهره در شکل (a) ۴-۳ و چهره ویژه های مربوط به آن ها در شکل (b) نشان داده شده است.



شکل ۴-۳ (a) مجموعه ای تمرین تصاویر چهره به صورت نمونه (b) تصویر چهره ای میانگین مجموعه ای تمرین



شکل ۵-۳ هفت چهره ویژه‌ی با بالاترین مقادیر ویژه که از مجموعه‌ی تمرین شکل ۴-۳ بدست آمده است.

هر چهره‌ای می‌تواند دقیقاً به وسیله‌ی ترکیب خطی چهره - ویژه‌ها ارائه شود. همچنین هر چهره می‌تواند با استفاده از «بهترین» چهره - ویژه‌ها، آنها یکه بیشترین مقادیر ویژه را دارند، تقریب زده شود. M چهره-ویژه‌ی بهترین، یک زیر فضای M بعدی را به وجود می‌آورد که ما فضای «چهره»‌ی همه‌ی تصاویر ممکن می‌نامیم.

Sirovich و Kirby تکنیکی را برای ارائه‌ی تصاویر چهره‌ها به طور کارا با استفاده از آنالیز جزء اصلی توسعه دادند آن‌ها بهترین مختصات سیستم را برای فشرده سازی تصویر محاسبه کردند که هر محور مختصات واقعاً یک تصویر است که آن را «تصویر - ویژه» نامیدند. آن‌ها بیان کردند، حداقل در تئوری، هر مجموعه‌ای از تصاویر به صورت تقریبی می‌تواند به وسیله‌ی ذخیره کردن یک مجموعه‌ی کوچک از وزن‌ها برای هر چهره و مجموعه‌ای کوچک از تصاویر استاندارد (تصویر - ویژه‌ها) بازسازی شود. وزن‌ها‌ی توصیف کننده‌ی هر چهره به وسیله‌ی تصویر کردن هر تصویر چهره بر روی تصویر - ویژه‌ها به دست می‌آیند.

در اینجا ما روشی که توسط M.Turk و A.Pentland برای توسعه‌ی یک سیستم تشخیص چهره‌ی مبتنی بر رویکرد چهره - ویژه پیشنهاد شده را دنبال خواهیم کرد . آن‌ها بیان کردنده که تعداد زیادی از تصاویر چهره به وسیله‌ی مجموع وزن داده شده‌ی مجموعه‌ی کوچکی از ویژگیهای خاص یا چهره - ویژه‌ها بازسازی می‌شوند . شاید یک روش کارا برای یادگیری و تشخیص چهره‌ها، به دست آوردن ویژگیهای خاص به وسیله‌ی تجربه و تشخیص چهره‌ها به وسیله‌ی مقایسه‌ی وزن‌های ویژگی های مورد نیاز برای بازسازی تقریبی آنها با وزنهای مرتبط با چهره‌های شناخته شده باشد . بنابراین، هر چهره به وسیله‌ی مجموعه‌ی کوچکی از ویژگی‌ها یا وزن‌های تصویر - چهره‌ی مورد نیاز برای توصیف و بازسازی آن‌ها مشخص می‌شود .

۲-۴-۳ کلیات سیستم تشخیص چهره

سیستم تشخیص چهره‌ی ارائه شده در طول فرآیند تشخیص چهره از سه فاز اصلی عبور می‌کند . سه واحد عملیاتی اصلی در این فازها قرار گرفته‌اند و آنها در شکل ۶-۳ ترسیم شده‌اند . ویژگی‌های این فازها به همراه سه واحد عملیاتی در زیر ارائه شده‌اند:

۱-۲-۴-۳ فاز شکل دهی پایگاه داده‌ی چهره : در این فاز، دریافت و پیش - پردازش تصاویر صورتی که قصد دارند به پایگاه داده‌ی چهره اضافه شوند، انجام می‌پذیرد . تصاویر در یک پایگاه داده‌ی چهره در سیستم ذخیره می‌شوند که این پایگاه داده‌ی چهره را کتابخانه‌ی چهره می‌نامیم چون در آن زمان خصوصیات پایگاه داده‌ی رابطه‌ای را ندارد . کتابخانه‌ی چهره در ابتدا خالی است . به منظور شروع فرآیندهای تشخیص چهره، در ابتدا کتابخانه‌ی خالی چهره با چهره‌ها پر می‌شود . فرض کنیم سیستم

تشخیص چهره بر روی تصاویر 128×128 کار می کند . به منظور انجام عملیات تبدیل اندازه و بهبود تصاویر، یک ماژول «پیش - پردازش» در سیستم وجود دارد . این ماژول به صورت اتوماتیک هر تصویر چهره را به تصویر 128×128 (در صورت لزوم) تبدیل می کند . دیگر عملیات پیش - پردازش که قبل ذکر شد نیز در این ماژول قرار دارد . در کتابخانه ی چهره برای هر چهره نیاز به دو عنصر داریم :

یک عنصر همان تصویر چهره است (که البته می توان فشرده سازی نیز روی آن انجام داد هر چند سرعت کم تر می شود .) و دیگری بردار اوزان مربوط به هر تصویر چهره است . بردار اوزان عناصر کتابخانه ی چهره خالی هستند تا زمانی که یک عملیات تمرین انتخاب شود و چهره - ویژه ها تشکیل شوند .

۲-۴-۳ فاز تمرین : پس از اضافه کردن تصاویر چهره ها به کتابخانه ی چهره، سیستم برای انجام تمرین و تشکیل چهره - ویژه ها آماده است . تصاویر چهره ای که می خواهیم در تمرین حضور داشته باشند از کتابخانه ی چهره انتخاب می شوند . از آنجائی که عناصر کتابخانه ی چهره نرمال شده هستند، هیچگونه عمل پیش - پردازشی در این مرحله صورت نمی پذیرد . پس از انتخاب مجموعه ای تمرین، چهره - ویژه ها تشکیل و برای استفاده های بعد، ذخیره می شوند . چهره - ویژه ها از روی مجموعه ای تمرین محاسبه می شوند و فقط M تصویر مربوط به بیشترین مقادیر ویژه نگهداری می شوند . این M تصویر چهره، «فضای چهره » ی M بعدی را نشان می دهدن . همان طور که چهره های جدید تجربه می شوند، چهره ویژه ها بروز رسانی و یا دوباره محاسبه می شوند . اوزان مربوطه در فضای وزن M بعدی برای هر عضو کتابخانه محاسبه می شود (به وسیله ای تصویر کردن تصویر چهره ای آن بر روی «فضای چهره » ای که به وسیله ای چهره - ویژه ها بوجود آمده است) . اکنون بردار وزن مربوط به هر عضو کتابخانه که در ابتدا خالی بود، به روز رسانی می شود . سیستم حالا برای فرآیند تشخیص آمده است .

۳-۲-۴-۳ فاز تشخیص و یادگیری : پس از انتخاب یک مجموعه‌ی تمرین و به وجود آوردن

بردارهای وزن عناصر تصویر چهره، اکنون سیستم آماده‌ی فرآیند تشخیص چهره است . کاربر فرآیند

تشخیص را با انتخاب یک تصویر آغاز می‌کند . بر اساس در خواست کاربر و اندازه‌ی تصویر دریافتی،

مراحل پیش - پردازش برای نرمال سازی این تصویر با مشخصات کتابخانه‌ی چهره صورت می‌پذیرد .

هنگامی که تصویر نرمال می‌شود، بردار وزن آن به کمک چهره - ویژه‌ها که قبلاً در فاز تمرین ساخته و

ذخیره شده بود، ساخته می‌شوند . پس از به دست آوردن بردار وزن، با بردار وزن هر عضو کتابخانه‌ی

چهره مقایسه می‌شود . همچنین کاربر یک آستانه شباهت را معین می‌کند که بر اساس آن سیستم

تصاویری که کمتر از این آستانه با تصویر چهره ورودی شباهت داشته باشند به عنوان تطابق ارائه نمی‌کند .

اگر حداقل یک تصویر چهره وجود دارد که در تشابه با تصویر ورودی مقدار آستانه را بر آورده کند،

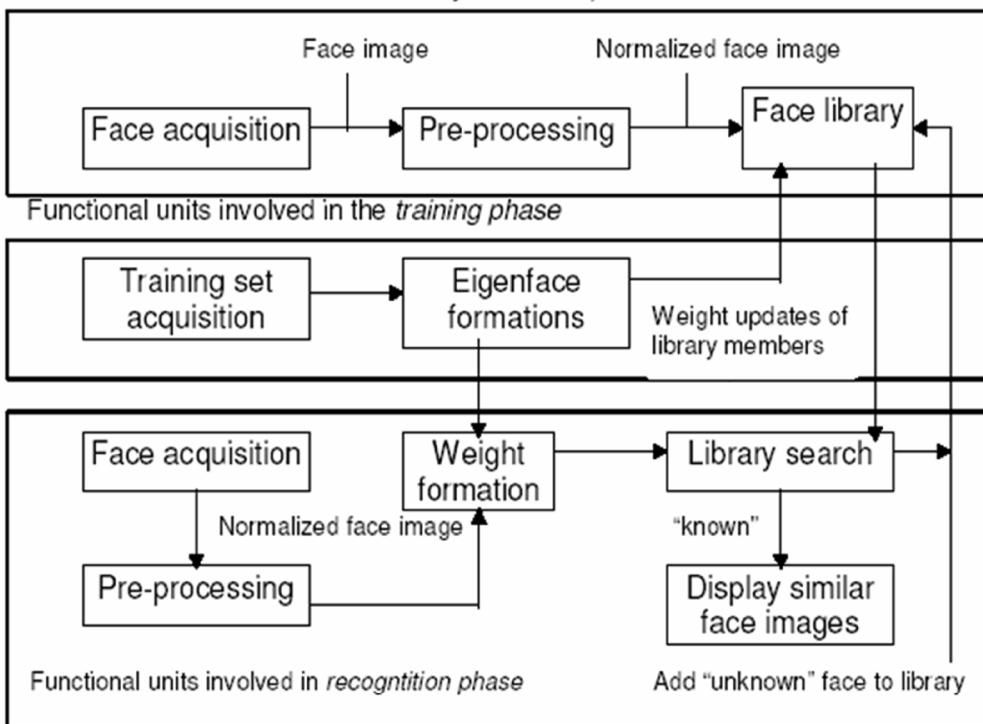
تصویر چهره به عنوان «شناخته شده» معرفی می‌شود . اگر هیچ تصویری مقدار آستانه را برآورده نکند،

تصویر به عنوان «شناخته نشده» معرفی می‌شود . پس از آنکه تصویر به عنوان ناشناخته معرفی شد

تصویر چهره‌ی جدید به همراه بردار وزن آن می‌تواند به کتابخانه‌ی چهره اضافه شود تا بعداً مورد استفاده

قرار گیرد (یادگیری برای تشخیص) .

Functional units involved in the *library formation phase*



شکل ۶-۳ دیاگرام بلوکی سیستم تشخیص چهره‌ی ارائه شده

۳-۴-۳ محاسبه‌ی چهره - ویژه‌ها

- \wedge intensity $I(x, y)$ یک آرایه‌ی دو بعدی $N \times N$ از مقادیر \wedge می‌باشد. همچنین یک تصویر ممکن است به عنوان یک بردار N به توان ۲ بعدی در نظر گرفته شود، بنابراین یک تصویر با اندازه‌ی 256×256 ، برداری 65536 بعدی یا به طور معادل یک نقطه در فضای 65536 بعدی می‌باشد.

تصاویر چهره‌ها، به خاطر شباهت در پیکربندی کلی، به طور تصادفی در این فضای بزرگ توزیع نمی‌شوند و بنابراین به وسیله‌ی یک زیر فضای به نسبت کم بعد قابل توصیف هستند. ایده‌ی اصلی آنالیز

جزء اصلی (یا بسط Karhunen – Loeve) پیدا کردن بردارهایی است که بهترین مقدار را برای توزیع تصاویر چهره در فضای تصویر کامل دارند.

این بردارها، زیر فضایی از تصاویر چهره را تعریف می کنند که ما فضای «چهره» می نامیم . طول هر بردار N به توان ۲ است، که یک تصویر $N \times N$ را توصیف می کند، که یک ترکیب خطی از تصاویر چهره ای اصلی است . چون این بردارها، بردارهای ویژه ای ماتریس کواریانس مربوط به تصاویر چهره ای اصلی هستند و چون آنها در ظاهر شبیه چهره هستند، ما آنها را چهره - ویژه می نامیم [50] .

۳-۵ تشخیص چهره با استفاده از شبکه های عصبی

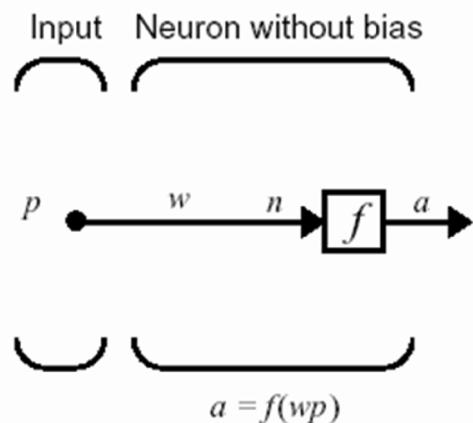
جستجو برای هوش مصنوعی (ماشین) یا سیستم هایی با رفتار انسانی چالشی فرا روی محققان شاخه های مختلف بوده است. بحث هایی بر روی تعریف یک سیستم هوشمند وجود دارد اما تنها سیستمی که به عنوان هوشمند مورد توجه قرار می گیرد، مغز انسان است و این جاست که یک ایده ای مهندسی معکوس مطرح می شود . چرا بایستی چیز هایی را که هم اکنون وجود دارند، اختراع کنیم . در میانه های قرن بیستم، ایده ای برای شبیه سازی مغز انسان از طریق طراحی پایین به بالای سیستم با شروع از نرون های به هم متصل و رسیدن به شبکه هایی که اعمال اصلی سیستم عصبی انسان را شبیه سازی می کنند، مطرح شد.

با گذر زمان، شبکه های عصبی خودشان را به عنوان مدل محاسباتی مورد استفاده مطرح کردند . شبکه های عصبی معمولا در انجام وظایف کامپیوتر های عام (عملیات با اعداد، اعمال منطقی و ...) نادقيق

هستند، اما در زمینه هایی که کامپیوتر ها در آنجا به مشکل بر می خورند، کارآیی بیشتری دارند مانند زمینه هایی که الگوریتم مشخصی وجود ندارد مانند پیش بینی، تشخیص چهره و....

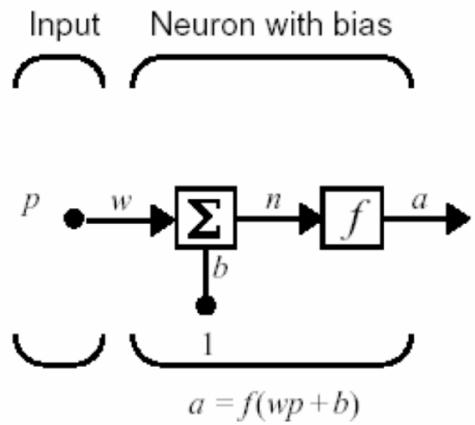
۱-۵-۳ مدل نرون

یک نرون با ورودی اسکالر و بدون بایاس در شکل ۶-۳ نشان داده شده است.



شکل ۷-۳ یک نرون با ورودی اسکالر و بدون بایاس

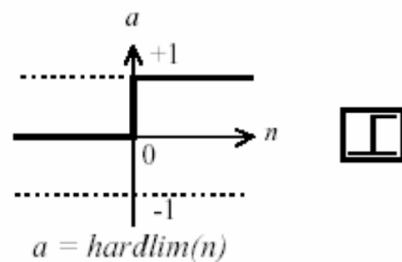
ورودی اسکالر p از طریق ارتباطی که مقدار آن را در وزن اسکالر w ضرب می کند به صورت حاصلضرب wp که اسکالر است، منتقل می شود. در این مرحله، ورودی وزن داده شده wp به عنوان آرگومان به تابع f منتقل می شود و این تابع خروجی a را تولید می کند. نرون شکل ۲ - ۵ یک مقدار b دارد. این مقدار به حاصلضرب wp اضافه شده است و تابع f را به مقدار b به سمت چپ محور X ها شیفت می دهد.



شکل ۳-۸ یک نرون با ورودی اسکالر و بایاس

۲-۵-۳ توابع انتقال

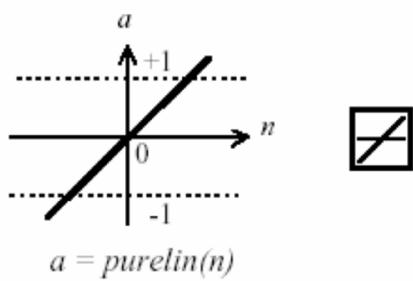
سه تابع انتقال که معمولاً مورد استفاده قرار می‌گیرد در شکل ۳-۸، ۹-۳ و ۱۰-۳ نشان داده شده است.



شکل ۳-۹ تابع انتقال hard-limit

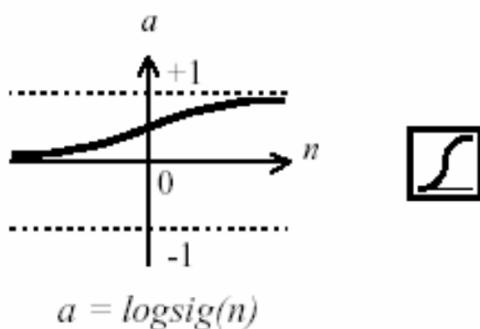
تابع انتقال hard-limit خروجی نرون را به مقادیر صفر و یک محدود می کند . اگر آرگومان ورودی تابع کمتر صفر بود، خروجی تابع مقدار صفر می باشد و آرگومان ورودی بزرگتر و یا مساوی صفر باشد، خروجی تابع یک می باشد . نرون هایی با این تابع انتقال در دسته بندی مورد استفاده قرار می گیرند.

تابع انتقال خطی در شکل ۱۰-۳ نشان داده شده است . نرون هایی با این تابع انتقال در تقریب زننده های خطی استفاده می شوند.



شکل ۱۰-۳ تابع انتقال خطی

تابع سیگموئید در شکل ۱۱-۳ نشان داده شده است. مقدار خروجی این تابع در بازه 0 بین مقادیر صفر و یک محصور است . این تابع انتقال معمولا در شبکه های انتشار عقب مورد استفاده قرار می گیرد.



شکل ۱۱-۳ تابع انتقال log-sigmoid

۳-۵ نرونی با ورودی برداری

یک نرون با بردار ورودی R عنصری در شکل ۱۲-۳ نشان داده شده است. بردار ورودی p

$$P_1, p_2, \dots, p_R$$

در بردار وزن W

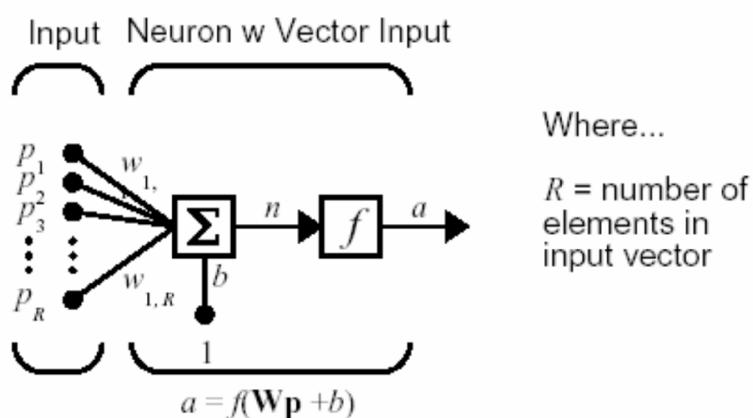
$$W_{1,1}, W_{1,2}, \dots, W_{1,R}$$

به روش ضرب اسکالر، ضرب می شود.

$$n = w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b$$

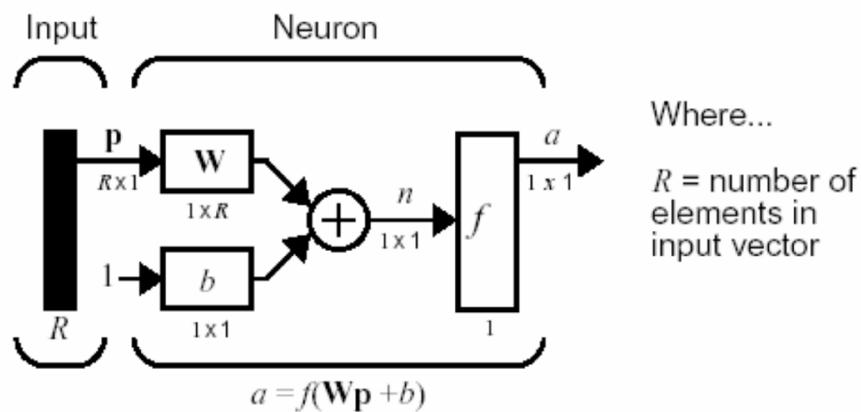
نتیجه‌ی این عملیات ضرب که یک مقدار عددی است ابتدا با مقدار بایاس جمع می شود. سپس

این حاصل جمع به عنوان آرگومان ورودی به تابع f انتقال می یابد [51].



شکل ۱۲-۳ یک نرون با بردار ورودی R عنصری

در شکل ۱۳-۳، نرون شکل ۱۲-۳ را با نشانه های سطح بالاتر می بینید.



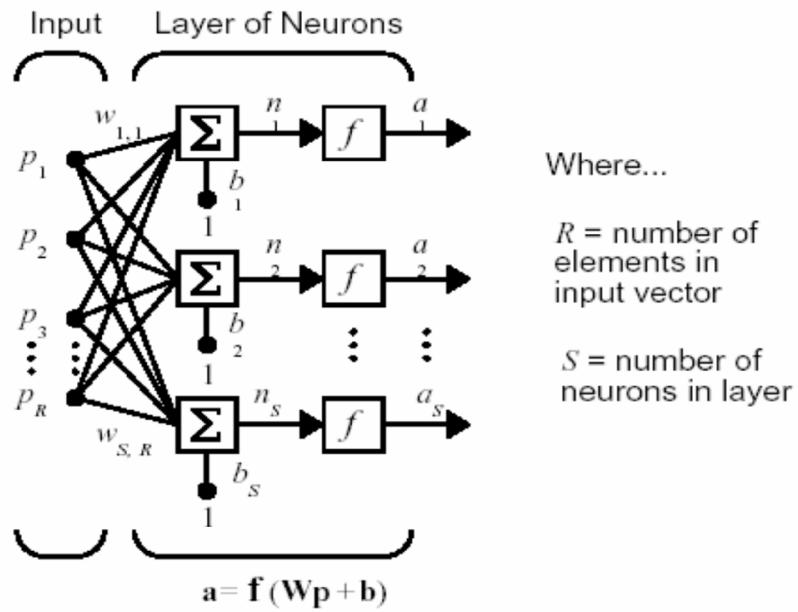
شکل ۱۳-۳ یک نرون با بردار ورودی R عنصری (نشانه گذاری سطح بالاتر)

۴-۵-۳ معماری های شبکه

دو یا چند نرون می توانند در یک لایه ترکیب شوند، یک شبکه می تواند از یک یا چندین لایه داشته باشد . ابتدا یک لایه از نرون ها را بررسی می کنیم.

۵-۵-۳ یک لایه از نرون ها

در شکل ۱۴-۳، یک شبکه ی یک لایه ای با R عنصر ورودی و S نرون را مشاهده می کنید.



شکل ۱۴-۳ یک شبکه‌ی یک لایه‌ای با R عنصر ورودی و S نرون

در این شبکه، هر عنصر بردار ورودی p به هر ورودی نرون از طریق ماتیس وزن W مرتبط می‌شود. نرون i ام جمع کننده‌ای دارد که ورودی‌های وزن داده شده را با مقدار بایاس برای تولید مقدار خروجی (i) جمع می‌کند. n های جمع آمده از جمع کننده‌ها به عنوان ورودی به توابع انتقال فرستاده می‌شوند و در نهایت مقادیر i به عنوان خروجی لایه محسوب می‌شوند و بردار a را می‌سازند. الزامی وجود ندارد که تعداد ورودی‌ها به یک لایه با تعداد نرون‌های آن لایه برابر باشد. وزن‌های لایه در قالب ماتریس W قابل بیان است.

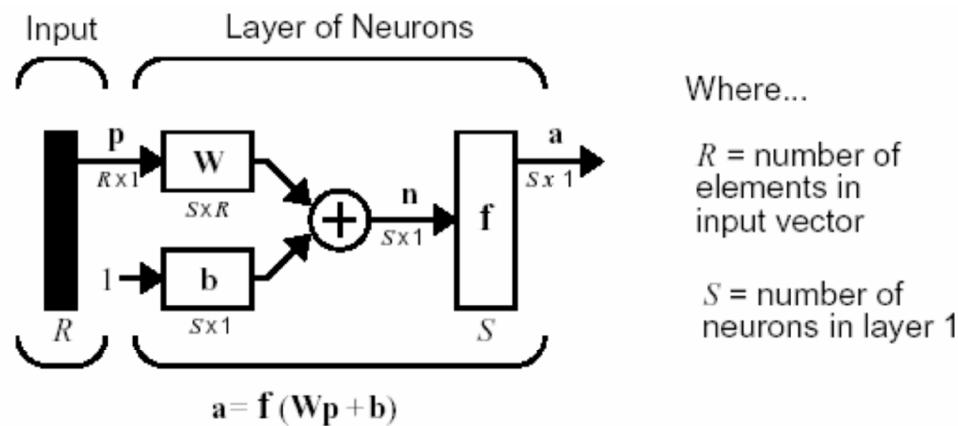
$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,R} \\ w_{2,1} & w_{2,2} & \dots & w_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S,1} & w_{S,2} & \dots & w_{S,R} \end{bmatrix}$$

Where...

R = number of elements in input vector

S = number of neurons in layer

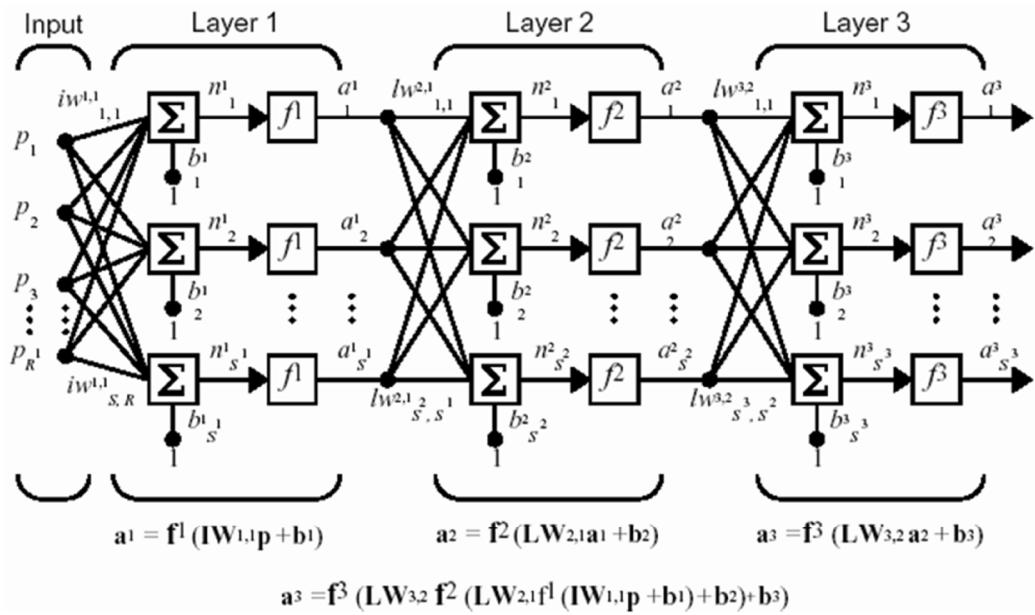
توجه کنید که اندیس سطر عناصر ماتریس W به نرون مقصد وزن اشاره دارد و اندیس ستون عناصر به منبعی که ورودی وزن است، اشاره دارد . بنابراین اندیس های $1,2$ w بر آن اشاره دارد که شدت سیگنال از عنصر ورودی دوم به اولین نرون، $1,2$ w است . شبکه‌ی تک لایه‌ی مورد بحث با نشانه گذاری کوتاه‌تری در شکل ۱۵-۳ نشان داده شده است.



شکل ۱۵-۳ یک شبکه‌ی یک لایه‌ای با R عنصر ورودی و S نرون (با نشانه گذاری سطح بالاتر)

۱۵-۴-۳ چندین لایه از نرون‌ها

یک شبکه می‌تواند چندین لایه داشته باشد . هر لایه یک ماتریس وزن W ، یک بردار بایاس b و یک بردار خروجی a دارد . برای تشخیص بین ماتریس‌های وزن، بردار‌های خروجی و...، برای هر یک از لایه‌ها در شکل‌ها، شماره‌ی لایه‌ی مربوطه را به متغیر‌های مورد نظر اضافه می‌کنیم . می‌توان استفاده از این نشانه‌ی لایه را در شبکه‌ی سه لایه‌ای که در شکل ۱۶-۳ نشان داده شده است، ملاحظه نمود[52] .



شکل ۱۶-۳ شبکه‌ی سه لایه ای

۷-۵-۳ یادگیری شبکه‌های عصبی

شبکه‌های عصبی قادر به یادگیری از نمونه‌های ارائه شده است. در شبکه‌های عصبی، فرآیند یادگیری، تنظیم وزن‌های ارتباطات بین نرون‌هاست. از نقطه نظری دیگر، فرآیند یادگیری ممکن است به عنوان نوعی جستجو برای پیدا کردن مینیمم کلی تابع خطای شبکه تفسیر شود.

دو رویکرد برای این مسئله وجود دارد. اولی یادگیری ناظارت یافته (یادگیری با آموزگار است.). در مرحله‌ی یادگیری، برای هر نمونه‌ی یادگیری هر دوی ورودی و خروجی مورد نظر آن ورودی را داریم. بنابراین، تعدادی جفت‌های تمرینی داریم (بردارهای ورودی و خروجی ممکن است به عنوان «سوال» و «پاسخ صحیح» دیده شوند.). این جفت‌ها به طور تکراری از شبکه گذرانده می‌شوند و

در هر تکرار الگوریتم های یادگیری به کار برده می شوند و وزن ها تنظیم می شوند تا شبکه بتواند به ورودی ها، پاسخ صحیح دهد (مثال نوعی این کلاس یادگیری، الگوریتم «انتشار به عقب خطأ» است.).

دومین رویکرد یادگیری بدون ناظارت است . این فرم یادگیری تنها بردار ورودی را نیاز دارد و معمولاً نوعی از دسته بنده را بر طبق فاصله‌ی (مثل فاصله‌ی همینگ) یک بردار ورودی با بردارهای ورودی دیگر انجام می دهد . به هر حال نتایج این نوع یادگیری بایستی آنالیز شود چرا که دسته بنده ممکن است نادرست باشد.

اکنون برای نمونه، تعریفی رسمی از الگوریتم انتشار به عقب (برای یک شبکه‌ی سه لایه) ارائه می شود.

۱. بردار ورودی به شبکه ارائه می شود.
۲. انجام می شود با این هدف که هر نرون، خروجی خود را با دنبال کردن

$$O_i = \frac{1}{1 + e^{(\sum_{i=1}^n x_i w_i + b)}}$$

بر روی نرون های لایه‌ی قبلی محاسبه کند.

۳. خطأ در لایه‌ی خروجی برای هر نرون با استفاده از خروجی مطلوب (y_i) محاسبه می شود :

$$err^o_j = o_j (1 - o_j) (y_i - o_j)$$

۴. خطأ به لایه‌ی میانی انتشار عقب می شود

$$err^h_i = h_i (1 - h_i) \sum_{j=1} \text{err}^o_j w_{ij}^o$$

۵. با داشتن مقادیر محاسبه شده‌ی err_i^h و err_j^o ، وزن‌ها از لایه‌ی میانی به لایه‌ی خروجی و از وزودی به لایه‌ی میانی به ترتیب طبق فرمول‌های

$$W_{ij}^o(t+1) = W_{ij}^o(t) + \alpha err_j^o h_i$$

و

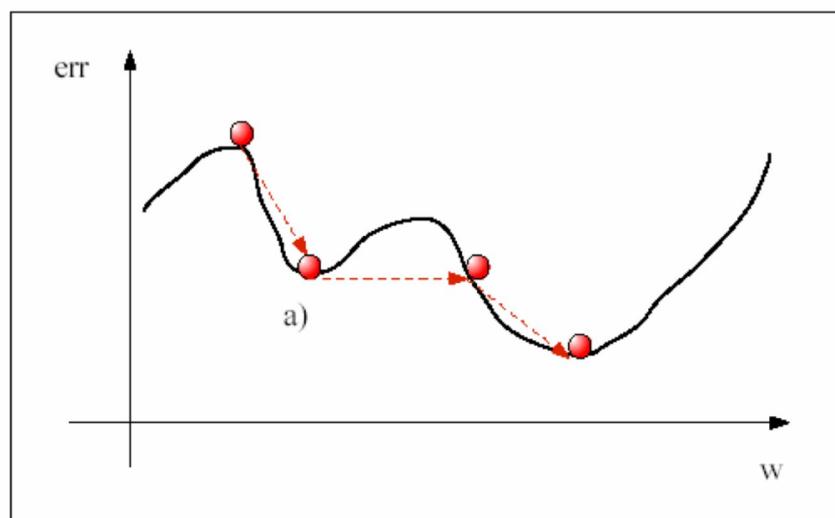
$$W_{ij}^h(t+1) = W_{ij}^h(t) + \alpha err_j^h x_i$$

می‌توانند تنظیم شوند که ضریب یادگیری و نرون i در لایه‌ی ورودی است.

۶. تمامی مراحل قبل تا زمانی که خطای کلی شبکه

$$E = \frac{1}{2} \sum_{i=1} (y_i - o_i)$$

بر روی جفت‌های تمرین زیر سطح معینی نباشد، ادامه می‌یابد.



شکل ۱۷-۳ مثالی از فرآیند انتشار عقب. خط چین‌ها جهت مشتق یکتابع خطأ را در هر مرحله نشان می‌دهند.

مسئله ای که در فرآیند یادگیری شبکه ممکن است رخ دهد قابلیت تعمیم شبکه است. انتشار به عقب مجموعه های تمرینی را به شبکه یاد می دهد اما در مواردی مجموعه‌ی تمرین، تمام محدوده‌ی مسئله را ارائه نمی دهد. در این وضعیت ما بخشی از مثال‌های مسئله را برای یادگیری به شبکه ارائه می کنیم و توقع داریم تا شبکه این دانش را تعمیم دهد. این دلیلی است که که مدل شبکه‌های عصبی موفق شمرده می شوند (به دلیل توانایی در تعمیم). به عبارت دیگر، در زمان ساخت شبکه‌ای برای تشخیص نوری اعداد دست نویس، نیازی به فراهم کردن همه‌ی شکل‌های ممکن اعداد نیست و ممکن است مجموعه‌ی تمرینی متشكل از تعدادی از نمونه‌های متدالول تر مهیا کنیم و فرض می شود که شبکه یک « دید کلی» از هر عدد را یاد می گیرد به نسبت آن که هر نمونه را در حافظه‌ی خود نگه دارد. بنابراین ممکن است شبکه قادر به ارائه‌ی پاسخی صحیح به الگوهایی باشد که تا کنون مشاهده نکرده است [53].

۸-۵-۳ دسته بندی با استفاده از شبکه‌های عصبی

یکی از کاربرد‌های شبکه‌های عصبی دسته بندی است. همانطورکه در فصل‌های قبل نیز گفته شده، در فرآیند دسته بندی، کلاسی که یک الگوی ورودی عضو آن است مشخص می گردد. در اینجا به نحوه‌ی سامان‌دهی شبکه‌ی عصبی و جفت‌های تمرین برای دسته بندی الگو‌ها می پردازیم.

اگر جفت‌های تمرین دسته بندی را داشته باشیم، این بدان معناست که به عنوان ورودی، الگو را داریم و به عنوان خروجی، کلاسی که عضو آن است. ورودی را می توانیم به صورت یک بردار به شبکه ارائه کرد که در این صورت تعداد ورودی‌های شبکه برابر طول بردار خواهد بود. نکته‌ای که می ماند آنست که خروجی‌ها را چگونه طراحی کنیم تا دسته بندی را برآورده سازد.

برای رسیدن به این هدف، فرمتی برای بردار خروجی معین می کنیم . این فرمت می تواند اینگونه باشد که به ازای هر کلاس، یک عنصر در بردار خروجی داریم . اگر در پاسخ به یک ورودی، مقدار یکی از این عناصر، یک شود و عناصر دیگر مقدار صفر داشته باشند، الگوی ارائه شده مربوط به کلاسی است که مقدار عنصر آن کلاس در بردار خروجی یک شده است. در شکل ۱۸-۳ این کد گذاری را برای تصاویر چهره‌ی ورودی به شبکه ملاحظه می کنید.



class1 : 1000 class2 : 0100 class3 : 0010 class4 : 0001

شکل ۱۸-۳ کد گذاری خروجی برای کلاس‌های چهره

در این مورد، چهار کلاس برای شناسایی چهار فرد داریم که هر کلاس به یک فرد منتسب است . شاید سر راست ترین گزینه برای تولید بردار از تصویر چهره‌ی دو بعدی به جهت ارائه به شبکه، پیمایش سطر به سطر یا ستون به ستون تصویر باشد و از کنار هم قرار دادن سطرها یا ستونها، یک بردار به دست آید . در تشخیص چهره، این بردار ورودی که به شبکه عصبی ارائه می شود، همان بردار ویژگی است . ویژگی‌هایی که در مراحل قبل فرآیند تشخیص چهره از تصویر استخراج شده‌اند و برای دسته بندی به شبکه‌ی عصبی ارائه می شوند. شبکه‌ی عصبی همانطور که قبلاً بحث شد، فرآیند یادگیری را پشت سر می گذراند تا برای دسته بندی چهره‌ها آماده شود[54] .

مراجع

- [1] Simon Hayek Neural Networks: A Comprehensive Foundation
- [2] H. Beigy and M. R. Meybodi, "Optimization of topology of neural networks Proc. of 4th Annual Int. Computer Society of Iran Computer Conf. CICC-98, Tehran, Iran, PP. 417-428 (In Persian), 1999.
- [3] M. R. Meybodi and H. Beigy, "Neural Network engineering desired size for three layer feedforward neural network." Technical Reports, Computer Eng. Dept. Amirkabir University of Technology, Tehran, Iran (In Persian), 1999.
- [4] Gabor Szeto, Color-Space Converter: RGB to YCrCb, August 27, 2007 [5] Scott Tan Yeh Ping (yptan@stanford.edu),(etc), FACE DETECTION THROUGH TEMPLATE MATCHING AND COLOR SEGMENTATION, 2003
- [6] Cheryl Lam,(etc), Face Detection, 2003
- [8] www.forum.pca.ir
- [7] www.mathnetworks.com
- [9] www.modir21.com/wp.../11/NeuralNetwork
- [10] **artificial.ir/intelligence/thread10295-2.html**
- [11] www.magzsaz.com/magz
- [12] www.izhikevich.org/publications/spikes.htm
- [13] www.civilica.com/Paper-WFP01-WFP01_046.html
- [14] itproject.blogfa.com/post-25.aspx
- [15] www.civilica.com/Paper-ICEE16-ICEE16_519.html
- [16] www.prozhe.com/tag/ -**perceptron**
- [47] Angela Chau,(etc), The Detection of Faces in Color Images, Dept. of Electrical
- [17] www.noormags.com/view/fa/articlepage/609279
- Rafael C. Gonzalez Digital Image Processing,[18] Richard E. Woods
- [19] Wenmiao Lu and Shaohua Sun, Face Detection in Color Images, Department of Electrical Engineering Stanford University, 2003

- [20] Turk, M., and Pentland, A., "Eigenfaces for recognition", Journal of Cognitive Neuroscience, Vol. 3, pp. 71-86, (1991).
- [21] www.electronix.ir/forum/index.php?topic=54.0
- [22] Guodong Guo, Stan Z. Li, and Kapluk Chan , “Face Recognition by Support Vector Machines ”
- [23] ALAN L. STEWART, The face-detection effect: Configuration enhances detection Stevens Institute of Technology, Hoboken, New Jersey,2003
- [24] Peter Brende,(etc), Face Detection, 2003
- [25] Diedrick Marius,(etc), Face Detection Using Color Thresholding, and Eigenimage Template Matching Digital Image Processing Project, 2003
- [26] Michael Bax,(etc), FACE DETECTION AND GENDER RECOGNITION, , 2003
- [27] www.mojzazan.com/forum/showthread.php?tid
- [28] [http:// bias.csr.unibo.it/](http://bias.csr.unibo.it/) FVC2000 and FVC2002/ databases.asp
- [29] [www.cambridgeincolour.com/tutorials/**image-interpolation**.htm](http://www.cambridgeincolour.com/tutorials/image-interpolation.htm)
- [30] [www.georeference.org/doc/**image_types**.htm](http://www.georeference.org/doc/image_types.htm)
- [31] icsystem.ir/Dars-pdf/pardazeshTasvir
- [32] www.prozhe.com/tag add constant to image
- [33] www.metan.ir/14362.html
- [34] ai.amini.eu/tutorial/dip/dipAverage.aspx
- [35] [ai.amini.eu/tutorial/dip/dip**Histogram**.aspx](http://ai.amini.eu/tutorial/dip/dipHistogram.aspx)
- [36] www.eca.ir/forum2/index.php?topic=14907.0
- [37] www.prozhe.com/tag JPEG Compressor
- [38] www.electronix.ir/forum/index.php?topic=54.0
- [39] [ai.amini.eu/tutorial/dip/dip**Filtering**.aspx](http://ai.amini.eu/tutorial/dip/dipFiltering.aspx)
- [40] www.civilica.com/Papers-PHOTONICS03-0-100
- [41] [ai.amini.eu/tutorial/dip/dipCrisp**SetImage**.aspx](http://ai.amini.eu/tutorial/dip/dipCrispSetImage.aspx)
- [42] ai.amini.eu/tutorial/dip/dipMorphErode.aspx
- [43] ai.amini.eu/tutorial/dip/dipMorphDilate.aspx

- [44] icsystem.ir/Dars-pdf/pardazeshTasvir
- [45] www.mojzasan.com/forum/showthread.php
- [46] www.iran-forum.ir
- [47] Webb, R. Andrew, "Statistical Pattern Recognition," John Wiley & Sons , 2nd Edition, July 2002.
- [48] A Survey Ming-Hsuan Yang, Member, IEEE,(etc), Detecting Faces in Images IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 24, NO. 1, JANUARY 2002
- [49] Siddharth Joshi,(etc),Digital Image Processing FACE DETECTION, 2003
- [50] Robust Face Detection , Centre for Multimedia Signal Processing Department of Electronic and Information Engineering DigitalImageProcessing-Gonzalez, Hong KongThe Hong Kong Polytechnic University,2008
- [51] [37] Matsugu, M., Mori, K., Ishii, M., & Mitarai, Y. (2002). Convolutional spiking neural network model for robust face detection. Proceedings of the Ninth International Conference on Neural Information Processing, Singapore, 660–664.
- [52] Liang Lu-Hong, Aihai-Zhou, et al, "A Survey of Human Face Detection", Chinese Journal of Computer,2002,25(5), pp. 449-458.
- [53] Fukushima, K. (1980). Neocognitron: a self-organizing neural networks for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics, 36, 193–202.
- [54] Henry A. Rowley,(etc), Neural Network-Based Face Detection, Copyright 1998 IEEE
- [55] papers2.persiangig.com/Papers/CMP/CMP2-1.pdf
- [56] www.ismvip.ir/MVIP%202003/papers/MV15
- [57] papers2.persiangig.com/Papers/CMP/CMP2-1
- [58] www.magiran.com/view.asp?Type
- [59] <http://www.codeproject.com/Articles/39060/Face-Detection-in-C>

پیوست

پ-۱ آشکار سازی چهره انسان در تصاویر

آشکار سازی چهره انسان در تصاویر، کاربردهای وسیعی در نظارت ویدئویی، تعامل انسان و کامپیوتر و ردیابی انسان دارد. بعضی از کارهای انجام شده در این مورد به آشکار سازی چهره در تصاویر سطح خاکستری محدود شده است، که در آن آشکارسازی چهره در تصاویر از رویه رو با تاکید بر ویژگی های صورت نظیر چشم ها و گوشه های لب انجام گرفته است. کارهای دیگری نظیر شناسایی چهره یک فرد در تصویر محدود شده است، در حالیکه اشکار سازی چهره در تصاویری که دارای چندین چهره انسان هستند، مورد نیاز است. بر عکس در اینجا هدف قابلیت اشکار سازی چهره در محدوده وسیعی از جهات مورد دید یا از دیدگاه های مختلف با تاکید بر ویژگی های عمومی چهره است. ویژگی های استفاده شده برای اشکار سازی چهره، رنگ پوست صورت انسان است. اولین قدم برای رسیدن به این منظور اشکار سازی ناحیه پوست در تصویر است. یکی از روش هایی که در اشکارسازی رنگ پوست مورد استفاده قرار گرفته است مبتنی بر ارائه مدلی برای رنگ پوست است. با توجه به اینکه رنگ پوست در فضای رنگی ycber تشکیل یک کلاستر یا خوشه را می دهد، به همین جهت یک رویکرد مستقیم به نظر میرسد. [55]

پ-۲ مشخصات رنگ پوست انسان

در بسیاری از کاربردهای نظیر ذخیره سازی و انتقال ویدیویی معمولاً رنگ ها به مولفه های روشنایی و رنگی تفکیک می شوند. یکی از دلایل این تقسیم بندی، حساسیت کمتر چشم انسان به تغییرات مولفه های رنگی است. تجربیات روانشناسی نشان می دهد که ادراک رنگ ها سه مشخصه است: شدت نور، اشباع و رنگ. در فضای رنگی ycber شدت نور متناظر با مقدار روشنایی y است، در حالی که اشباع و رنگ در مولفه های رنگی cb , cr نگهداری می شوند. رنگ پوست انسان تشکیل طبقه خاصی از رنگ ها را می

دهد که از اکثر رنگ های اشیای طبیعی دیگر تفکیک پذیر است. اگر چه رنگ های پوست از شخصی به شخصی و از نژادی به نژادی دیگر فرق می کند، لیکن دارای توزیع خیلی کوچک در صفحه رنگی است. این بدان معناست که رنگ های پوست نسبتاً در اشباع و رنگ ثابت اند. رنگ هایی به عنوان رنگ پوست طبقه بندی می شوند که اشباع و رنگ آن ها در محدوده خاصی قرار داشته باشند. با حذف مولفه روشنایی از رنگ ها، اختلاف بین رنگ های پوست افراد و نژاد های مختلف و تاثیر شرایط نورانی کاهش داده می شود. در اینجا از مشخصات رنگ پوست انسان با به کارگیری مولفه های **cbcr** استفاده شده و با استفاده از توزیع رنگ پوست در فضای رنگی **cbcr** پروسه دسته بندی بین رنگ پوست و غیر پوست (چهره و غیر چهره) انجام گرفته است. [56]

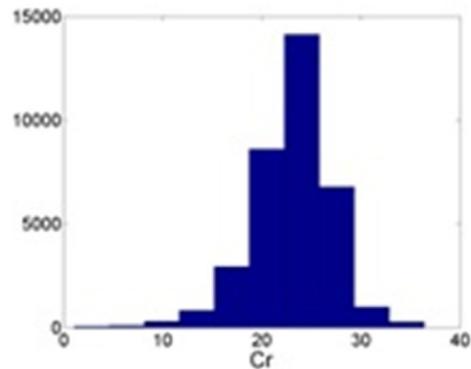
پ-۳ تولید رنگ پوست در فضای رنگی **cbcr**

خروجی اکثر دوربین های رنگی، تصاویر رنگی در فضای رنگ **RGB** را در اختیار قرار می دهند.

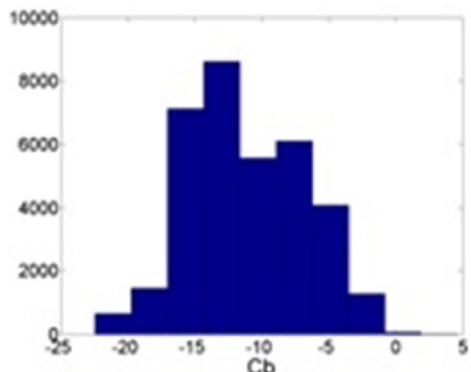
برای تبدیل فضای رنگی **RGB** به فضای رنگی **ycbcr** می توان از رابطه زیر استفاده نمود:

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ Cb &= -0.169R - 0.332G + 0.500B \\ Cr &= 0.500R - 0.419G - 0.081B \end{aligned}$$

و هیستوگرام توزیع آنها در شکل زیر نشان داده شده است:



شکل پ-۱ هیستوگرام توزیع cr



شکل پ-۲ هیستوگرام توزیع cb

ما می توانیم با انتخاب یک معیار فاصله با آستانه یک واحد در فضای رنگی cbcr، نمونه های غیر پوسی که دارای فاصله کمتر از یک واحد در فضای رنگی cbcr از نمونه های پوسی بودند را حذف کنیم. بدین ترتیب بین دو کلاس پوسی و غیر پوسی حداقل یک واحد فاصله در فضای رنگی انتخاب شده

وجود دارد و در نتیجه شبکه عصبی امکان تفکیک دو کلاس را خواهد داشت.[57]

برای آشکار سازی رنگ از کد زیر استفاده می کنیم:

```
public int[] FaceDetection(Form1 Main, PictureBox PicFirst,  
PictureBox Picresult, TextBox Twskin, PictureBox face, TextBox Tbskin)  
{  
    Testdetection = 0;  
    Graphics hh = PicFirst.CreateGraphics();  
    Bitmap PreResult = new Bitmap(PicFirst.Image);  
    Bitmap Result = new Bitmap(PicFirst.Width, Picresult.Height);  
    Bitmap HisResult = new Bitmap(PicFirst.Width, Picresult.Height);  
    Bitmap Template50 = new Bitmap("sample50.bmp");  
    Bitmap Template25 = new Bitmap("sample25.bmp");  
    Bitmap facee = new Bitmap(face.Width, face.Height);  
    for (int i = 0; i < PicFirst.Width; i++)  
    {  
        for (int j = 0; j < PicFirst.Height; j++)  
        {  
            //Getting red layer  
            R = PreResult.GetPixel(i, j).R;  
            //Getting green layer  
            G = PreResult.GetPixel(i, j).G;  
            //Getting blue layer  
            B = PreResult.GetPixel(i, j).B;  
            //Formula for convert RGB 2 YCBCR  
            //Y = 0.299 * R + 0.587 * G + 0.114 * B;  
            //Cb = -0.169*R - 0.332*G + 0.500*B;  
            Cr = 0.500 * R - 0.419 * G - 0.081 * B;  
            //Check for "-" values(Because we can't use "-" values for RGB color)  
            if (Cr < 0)  
                //Convert to Positive values  
                Cr = 0;  
            HisResult.SetPixel(i, j, Color.FromArgb((int)Cr, (int)Cr, (int)Cr));  
        }  
    }  
}
```

پ-۴ آشکار سازی چهره

خروجی شبکه عصبی مقداری در محدوده ۰-۱ به هر نقطه تصویر نسبت می دهد، بدین ترتیب هر نقطه تصویر دارای مقداری است که نشان می دهد این نقطه به پوست شباهت دارد یا خیر. با در نظر گرفتن آستانه ای برابر با صفر می توان تصویر ورودی را به یک تصویر باینری تبدیل نمود که در ان نقاط پوستی و غیر پوستی مشخص می شوند.^[58]



این کد برای آشکار سازی تصویر مورد استفاده قرار می گیرد:

```
if (Cr > 10)
    //Convert to Positive values
    Cr = 255;
    //if (Cr > 20)//
    //Convert to Positive values
    //Cr = Cr + 128;
    //Set the pixel by cr value
    Result.SetPixel(i, j, Color.FromArgb((int)Cr, (int)Cr, (int)Cr));
}
```

این کد برای تشخیص لبه تصویر مورد استفاده قرار می گیرد:

```

//Create filter for my image
FiltersSequence processingFilter = new FiltersSequence();
//Detection of the edges
processingFilter.Add(new Edges());
WPresentdetection=0;
BPresentdetection1= 0;
//Apply filter for my pic
Picresult.Image =processingFilter.Apply( Result);
//Detect face with big template

```



پ-۵ استخراج لبه الگوی صورت

با استفاده از استخراج لبه می توان ویژگی های مهم یک تصویر را استخراج نمودوسپس ان را به

قالب یک چهره در اورد.[59]

```

for (int d = 0; d <150-50; d+=20)
{
Testdetection = 0;
WPresentdetection = 0;
BPresentdetection1 = 0;
for (int c = 0; c <150-50; c+=20)
{
Testdetection = 0;
WPresentdetection = 0;
BPresentdetection1 = 0;
for (int w = 0; w < Template50.Width; w++)

```

```

{
for (int h = 0; h < Template50.Height; h++)
{
if (Template50.GetPixel(w, h).B == 255)
{
if (Result.GetPixel(d + w, c + h).B == Template50.GetPixel(w, h).B)
{
WPresentdetection++;
}
//if (Template50.GetPixel(w, h).B ==255)
//{
//if (Result.GetPixel(d + w, c + h).B == Template50.GetPixel(w, h).B)
//{
//BPresentdetection1++;
//}
//}
}
}
TWskin.Text = WPresentdetection.ToString();
TBskin.Text = BPresentdetection1.ToString();
if (WPresentdetection >377)//&&BPresentdetection1>1579)
{
hh.DrawRectangle(new Pen(Color.Red), d, c,60, 60);
for (int i = 0; i <60; i++)
{
for (int j = 0; j<60; j++)
{
r = PreResult.GetPixel(i + d, j + c).R;
facee.SetPixel(i, j, Color.FromArgb((int)r, (int)r, (int)r));
face.Image= facee;
}
}
TWskin.Text = WPresentdetection.ToString();
TBskin.Text = BPresentdetection1.ToString();
WPresentdetection = 0;
BPresentdetection1 = 0;
Testdetection = 1;
break;
}
// Detect face with small template
//else
//{
//Testdetection = 0;

```

```

//WPresentdetection = 0;
//for (int w = 0; w < Template25.Width; w++)
//{
//{
//for (int h = 0; h < Template25.Height; h++)
//{
//if (Result.GetPixel(d + w, c +h).B == 255)
//{
//if (Result.GetPixel(d + w, c + h).B == Template25.GetPixel(w, h).B)
//{
//WPresentdetection++;
//}
//}
//}
//}
//}
//if (WPresentdetection > 50)
//{
//hh.DrawRectangle(new Pen(Color.Red), d, c, 30, 30);
//for (int i = 0; i < 30; i++)
//{
//for (int j = 0; j < 30; j++)
//{
//r = PreResult.GetPixel(i + d, j + c).R;
//facee.SetPixel(i, j, Color.FromArgb((int)r, (int)r, (int)r));
//face.Image = facee;
//}
//}
//}
//TWskin.Text = WPresentdetection.ToString();
//WPresentdetection = 0;
//Testdetection = 2;
//break;
//}
//}
}
}

if (Testdetection == 1 || Testdetection == 2)
{
WPresentdetection = 0;
BPresentdetection1=0;
Testdetection = 0;
break;
}
}

```

پ-۶ واژه نامه

فارسی

انگلیسی

شبکه های عصبی مصنوعی

Artificial Neural Network

سلول های عصبی

Neuron

بازخورد

Feed-Back

شبکه ناظر

supervised network

پردازش تصویر و دید

Image processing and computer vision

پردازش علائم

Signal processing

شناسایی الگوهای

Pattern recognition

پزشکی

Medicine

سیستم های نظامی

Military systems

سیستم های تجاری

Financial systems

هوش مصنوعی

Artificial intelligence

سیستم های قدرت

Power systems

بدنی سلول

Some

دندانه

Dendrite

اکسون

Axon

بازنشان

Reset

پردازش تصاویر	Image processing
پیکسل	Pixel
عناصر تصویر	Picture Elements
وضوح تصویر	Resolution
پیمایش گر رنگی	color scanner
هم فاز	Inphase
متعامد	Quadrature
ناهمبسته	Decoupled
پردازش نقطه ای	Point Process
پردازش سطحی	Area Process
پردازش هندسی	Geometric Process
پردازش قاب	Frame Process
مخلوط کردن	Combine
رندر کردن	Render
درونویابی	Interpolation
رقم های دوتایی	Digit binary
بیتها	Bits
فلفل نمکی	Pepper&Salt
فیلترهای پایین گذر	Low Pass Filter

فیلترهای بالاگذر

High Pass Filter

فراتفکیک پذیری

Super resolution

ثبت تصویر

RegistrationImage

بازسازی تصویر

Image Reconstruction

انحراف، اعوجاج

Distortion

مغایرت

Contrast

باینری

Binary

Abstract

This thesis examines the context in which neural networks. What to include topics such as neural networks, Why neural networks, what is included, what is the use And how does it work, is of the neural network is a new generation We have a world of information and knowledge in the field of networking. Among the results of this paper can be valuable in understanding neural networks and how they work,

an introduction to basic and more familiar with this new knowledge. Today neural networks in different applications such as pattern recognition problems, are used to control and classification problems. Or control systems in modeling the internal structures are unknown or very complex neural network will learn to control their function. Currently, there are many different types of artificial neural networks, which are summarized as: multilayer perceptron networks, Kvhvnn, Hapfyld ... The network also face their own ways. Including a discussion of the applications of pattern recognition, image processing and view, artificial intelligence, control, robotics and many others named.

So here we are issues worthy of introduction to artificial neural networks and its applications, we, in the Adamhprsptrvn algorithms, artificial neural networks is one of the most important illustrated.

Keywords: networked neurons, neurons, Perceptron, image processing and image



ISLAMIC AZAD UNIVERSITY

Shabestar Branch

**Thesis for B.Sc. degree
In Software Engineering**

Subject:

**Implementation of neural network applications
functions in image processing**

Thesis Advisor:

Amirhosein Jalali (Eng)

By:

Saeid Soleimani vaygan

Summer 2012