

Amirkabir University of Technology
(Tehran Polytechnic)



Department of
Computer Engineering

درس: بینایی ماشین

تمرین چهارم

نجمه محمدباقری

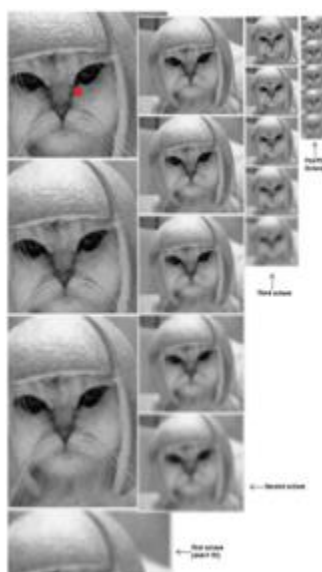
۹۹۱۳۱۰۰۹

سوال اول:

محاسبه‌ی ویژگی سیفت در ۶ گام انجام می‌شود.

گام اول: ساخت فضای مقیاس.

در این گام از تصویر اولیه تعدادی تصویر در مقیاس‌های مختلف ساخته می‌شود. برای اینکار از فیلترهای گوسی با انحراف معیارهای مختلف استفاده می‌شود. در ابتدا تصویر اولیه به درجات مختلف تار می‌شود. سپس اندازه تصویر نصف و تارسازی تکرار می‌شود. بدفعات لازم اندازه تصویر نصف و تارسازی تکرار می‌شود. به تصاویر هم اندازه که میزان تارشدگی متفاوتی دارند یک اکتاو گفته می‌شود. نمونه‌ای از این گام در تصویر ۱ مشاهده می‌شود.



شکل ۱ ساخت فضای مقیاس

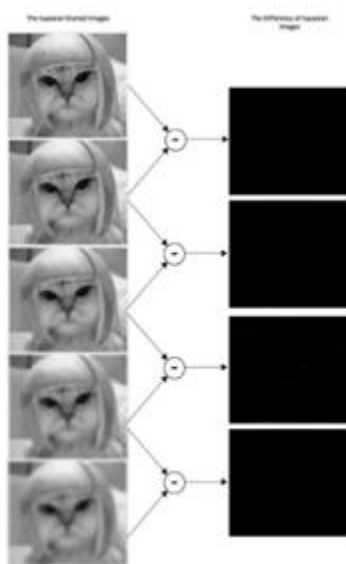
اگر مقیاس تاری در یک تصویر σ باشد در تصویر بعدی $k\sigma$ است. در جدول زیر برای ۴ اکتاو و ۵ تصویر مقیاس‌ها نشان داده شده است. در صورتی که اولین مقیاس ۰.۷ باشد.

octave	scale →				
	0.707107	1.000000	1.414214	2.000000	2.828427
1	1.414214	2.000000	2.828427	4.000000	5.656854
2	2.828427	4.000000	5.656854	8.000000	11.313708
3	5.656854	8.000000	11.313708	16.000000	22.627417

شکل ۲ مقیاس‌های تاری

گام دوم: تقریب عملگر LOG با تفاضل نتایج فیلترهای گوسی.

به دلیل اینکه محاسبه‌ی دقیق LOG پرهزینه است از تقریب آن استفاده می‌شود. برای تقریب زدن از تفاضل نتایج فیلترهای گاوسی استفاده می‌شود. در شکل ۳ یک نمونه از خروجی این عمل مشاهده می‌شود.

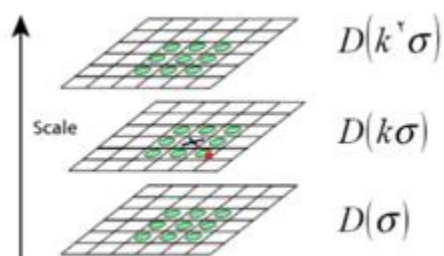


شکل ۳ تقریب LOG

تصاویر تفاضل از مقیاس هم مستقل است.

گام سوم: پیدا کردن مکان حداقل‌ها و حداکثرها در تصاویر تفاضل .

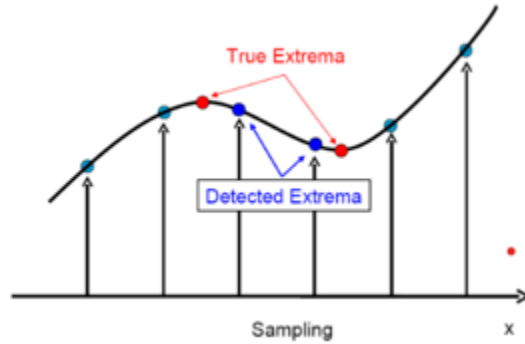
برای هر نقطه یک همسایگی $3 \times 3 \times 3$ در نظر می‌گیریم و ۲۶ مقایسه انجام می‌دهیم. نقطه‌ی مدنظر بعنوان نقطه‌ی شاخص انتخاب می‌شود اگر از همه‌ی ۲۶ همسایه‌ی خود بزرگتر یا کوچکتر باشد.



شکل ۴ نمایی از همسایگی موردنظر برای یافتن نقاط حداقل و حداکثر

لازم به ذکر است که در بالاترین و پایین‌ترین تصویر هر اکتاو این عمل انجام نمی‌شود. چون همسایگی لازم را نمیتوانیم تشکیل دهیم.

بعد از این لازم است زیرپیکسل‌های نقاط حداکثر و حداقل پیدا شوند. بدلیل اینکه ممکن است نقاط حداکثر و حداقل در جایی بین پیکسل‌ها باشند، لازم است نزدیک‌ترین پیکسل به آن نقاط را تعیین کنیم. روش کار؟ استفاده از بسط تیلور تا درجه‌ی دوم آن و حل آن در تساوی با صفر.



شکل ۵ تعیین نقاط زیرپیکسلی

گام چهارم: حذف لبه‌ها و پاسخ‌های کنتراست پایین از بین نقاط شاخص

نقاطی که مقدارشان از یک حد آستانه‌ای کمتر است را حذف می‌کنیم. برای حذف لبه‌ها میتوان مانند گوشه‌یاب هریس از دترمینان و تریس ماتریس هیسن استفاده کرد. به صورتی که اگر رابطه‌ی زیر برقرار بود، نقطه حذف شود.

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r}$$

گام پنجم: تخصیص یک وضعیت به هر نقطه شاخص

اندازه و زاویه‌ی گرادیان در هر نقطه شاخص با روابط زیر محاسبه می‌شود. (میدانیم که نقطه‌ی شاخص مربوط به کدام مقیاس است).

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \tan^{-1} \left(\frac{(L(x, y+1) - L(x, y-1))}{(L(x+1, y) - L(x-1, y))} \right)$$

سپس باید وضعیت را در اطراف نقطه‌ی شاخص مشخص کنیم. سه گام ایده برای انجام این کار وجود دارد:

۱. تجميع اطلاعات اندازه‌ها و جهت گرادیان‌ها در محدوده‌ی اطراف یک نقطه‌ی شاخص.
۲. تعیین مهمترین جهت یا جهت‌های گریان.

۳. تخصیص جهت(های) مهم به نقطه‌ی شاخص.

اندازه‌ی محدوده‌ی انتخابی به مقیاس تصویر نقطه‌ی شاخص وابسته است. هرچه مقیاس بزرگتر باشد محدوده نیز باید بزرگتر باشد. پنجره‌ای با اندازه‌ی ۱.۵ برابر مقیاس تاری پیشنهاد شده است.

پس از تخصیص جهت محاسبات بعدی انجام می‌شود تا تغییر ناپذیری نسبت به چرخش ایجاد شود.

برای هر محدوده اطراف نقطه‌ی شاخص یک هیستوگرام ۳۶ بازه‌ای از جهت‌ها تشکیل می‌دهیم. رای هر نقطه با اندازه‌ی گرادیان آن ضرب در تابع گوسی (با مرکز نقطه شاخص و انحراف معیاری ۱.۵ برابر مقیاس نقطه شاخص) مشخص می‌شود.

بلندترین قله‌ی هیستوگرام بیانگر جهت غالب است. در صورت لازم می‌توان بیش از یک جهت و نقطه‌ی کلیدی نگه داشت.

گام ششم: ساخت توصیفگر

برای هر نقطه شاخص بدست آمده، یک ناحیه 16×16 پیکسلی از تصویر هموارشده با مقیاس نقطه را در نظر می‌گیریم. مقدار گرادیان این نقاط را با دامنه یک فیلتر گوسی واقع در نقطه شاخص با انحراف معیاری برابر نصف بعد پنجره ضرب می‌کنیم تا اثر گرادیانهای دورتر کمتر شود. برای تغییر ناپذیری در برابر چرخش زاویه بردارهای گرادیان نسبت به وضعیت نقطه شاخص سنجیده میشوند (زاویه نقطه شاخص از همه ی زاویه ها کم میشود). در هر همسایگی 4×4 ناحیه، یک هیستوگرام ۸ بازه‌ای از زاویه های گرادیان ها میسازیم. رای هر پیکسل برابر اندازه گرادیان وزن داده شده است. نتیجه این محاسبات ۱۶ هیستوگرام ۸ بازه‌ای است که یک بردار ۱۲۸ عنصری میدهد.

سوال دوم:

برای این کار ابتدا یک تصویر از تویی که می‌خواهیم آن را دنبال کنیم گرفته و ویژگی‌های سیفت آنرا استخراج می‌کنیم. سپس در هر فریم بررسی میکنیم که آیا تطبیقی برای هر نقطه‌ی کلیدی در تصویر اولیه با این فریم وجود دارد یا خیر.

اگر تطبیقی وجود داشت این تطبیق را بررسی می‌کنیم که میزان شباهتش با نقطه‌ی کلیدی چقدر است، اگر از حد آستانه‌ای بیشتر بود (برای فاصله، کمتر) این نقطه از فریم به نقطه‌ی متناظرش در تصویر اولیه تطبیق می‌خورد. پس از آنکه نقاطی که از فریم با تصویر اولیه تطبیق خوردند، یکی از نقاط که بیشترین شباهت را دارد در

هر فریم در نظر گرفته و مکان پیکسلی آن را به مکان پیکسلی فریم قبل وصل می‌کنیم. برای رسم یک بردار می‌توان از تابع آماده در کتابخانه‌ی cv2 استفاده کرد.

سوال سوم و چهارم:

خروجی‌ها در فایل ذخیره شده است.

تصویر اولیه مورد استفاده :



روش سیفت: نسبت به چرخش و حرکات سریع نیز خوب عمل می‌کرد.

روش فریک: با چرخش جسم، نمیتوانست نقاط کلیدی را تشخیص دهد.

بطور کلی عملکرد روش سیفت بهتر بود اما سرعت فریک بیشتر بود.