

تمرین شماره ۰۵ بینایی ماشین

فاطمه توکلی - ۴۰۰۱۳۱۰۱۶

سؤال ۱:

مراحل ساخت توصیفگر سیفت به ترتیب و صورت زیر انجام میشود:

• ساخت فضای مقیاس

چند مقیاسی در طبیعت هم رایج است، برای مثال برای مشاهده یک درخت به برگها و جزئیات آن کاری نداریم. از همین جهت برای حذف جزئیاتی که به آنها احتیاج نداریم تصاویر با ابعاد مختلف و فیلتر گوسی مختلف را ایجاد میکنیم. فیلتر گوسی به صورت ضرب در تصویر اولیه حاصل میشود.

برای ابعادهای متفاوت اصولاً از ۴ اکتاو استفاده میشود. یعنی ابتدا خود تصویر بعد از آن تصویر با ابعاد نصف، یک چهارم و یک هشتم تولید میشود.

در هر کدام از این اکتاو ها یا ابعادهای متفاوت عموماً ۵ مقیاس تار با فیلتر گوسی اعمال میشود.

این مقیاس ها به نحوی است که اگر k را به عنوان ضریبی در نظر بگیریم و به طور مثال مقدار $\sqrt{2}$ را برای آن بگذاریم سیگمای فیلتر گوسی σ ، $\sqrt{2}\sigma$ ، 2σ ، $2\sqrt{2}\sigma$ ، 4σ در نظر گرفته میشود و این فیلتر بر روی تصویر اعمال میشود. حال که ۲۰ تصویر تولید شده است به مرحله بعدی می رویم.

• تقریب LOG

اپراتور LOG در تعیین نقاط شاخص عملکرد خوبی را از خود نشان داده است از همین جهت استفاده از این اپراتور برای پیدا کردن نقاط کلیدی مفید است اما از طرفی عملیاتی کردن آن به دلیل اینکه روی بخش بخش تصویر اعمال میشود و تعداد زیادی فیلتر در فیلتر بانک است و مقادیر باید محاسبه شود زمان بر هستند و نسبت به مقیاس نیز همچنان حساس است. از همین رو به دنبال روشی که از نظر محاسباتی کمتر زمان بر است میگردد. برای اینکه نشان دهد که تفاضل بین تصویرهای گوسی ایجاد شده در واقع تخمینی از LOG است از معادله حرارت استفاده میکند. معادله حرارت به صورت زیر است که در آن دما همان واریانس در نظر گرفته میشود.

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma}$$

که با ضرب مخرج در سمت چپ به صورت زیر میشود:

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k - 1)\sigma^2 \nabla^2 G.$$

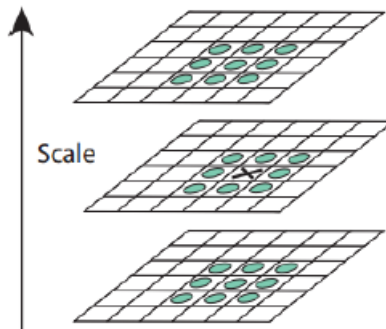
در اینجا مختصات نقاط با x, y به نمایش در آمده و k نیز ضریب برای تولید مقیاس های تار مختلف است G نیز تصاویر گوسی هستند.

در اینجا عامل حساسیت به مقیاس که سیگما است با ضرب از بین رفته است و چون به دنبال حداکثرها و حداقلها هستیم بنابر این ضریب k در این معادله تاثیر گذار نیست. پس DOG را که تخمین به صورت تفاضلی از LOG است را محاسبه میکنیم و تصاویر اصلی را به شکل زیر میسازیم:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma). \end{aligned}$$

• یافتن نقاط حداکثر و حداقل

در اینجا که تصویر D تولید شده است به دنبال پیدا کردن نقاط حداکثر و حداقل برای پیدا کردن نقاط کلیدی میگردیم. برای اینکار یک همسایگی ۳ در ۳ در ۳ حول هر نقطه در نظر میگیریم به صورت زیر:



بعد از در نظر گرفتن این همسایگی نقطه مرکزی را با تمام همسایه‌ها بررسی میکنیم، اگر این نقطه از تمام همسایه‌های خود کمتر یا مقدار بیشتری داشت آن را به عنوان نقطه کلیدی در نظر میگیریم. باید توجه داشت که این همسایگی برای مقیاس‌های اول و آخر اتفاق نمیافتد چون مقیاس پایین یا بالایی ندارند تا همسایگی در نظر گرفته شود.

حال اگر این نقاط حداقل و حداکثری بین پیکسل‌ها باشد در آن هنگام باید از بسط دو بعدی تیلور استفاده شود تا تغییرات تخمین زده شو و نقاط حداقل و حداکثری تشخیص داده شود و بعد از آن نزدیکترین پیکسل را به عنوان نقاط کلیدی برگرداند. (در پیدا کردن نقاط حداقلی و حداکثری با بسط تیلور نقاط زیر پیکسلی یا بین پیکسلی پیدا میشود.) برای بسط تیلور از فرمول زیر استفاده میشود:

$$D(\mathbf{x}) = D + \frac{\partial D}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x}$$

این تخمین با دو تصویر تفاضلی قابل محاسبه است بنابراین دلیل ۵ تصویر با مقیاس متفاوت که تصویر تفاضلی ابتدایی و انتهای قابل استفاده نیستند دلیل مناسب بودن تعداد ۵ برای مقیاس است. بعد از پیدا کردن این نقاط به مرحله بعد میرویم.

• حذف نقاط کنتراست پایین و لبه‌ها

در این مرحله به حذف نقاط کنتراست پایین و لبه‌ها میپردازیم در این حذف کردن می‌خواهیم نقاطی که بیشتر شاخص هستند باقی بمانند. در نقاطی که کنتراست پایین دارند تفاوت آنها با اطرافشان کم است و برای نقاطی که در لبه هستند با اینکه نقاط کلیدی هستند اما آنها برای باقیماندن نقاط گوشه که اهمیت بیشتر دارند حذف میشوند. در گام پیدا کردن مقدار تفاوت نسبت به اطراف که از حدی کمتر است ممکن است نیاز به محاسبه نقاط زیر پیکسلی داشته باشیم تا این تفاوت را به درستی دریابیم. برای حذف نقاط بر روی لبه نقاطی که در یک امتداد پاسخ قوی دارند و در امتداد دیگر ندارند حذف میشوند. برای اینکار میتوان از گوشه یاب هریس نیز استفاده کرد و از دترمینان و تریس ماتریس هسین به این اطلاعات پی برد. برای اینکار به صورت زیر عمل میشود که اگر ماتریس هسین از مشتقات تصاویر به صورت زیر باشد:

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

و تریس و دترمینان این ماتریس به صورت زیر:

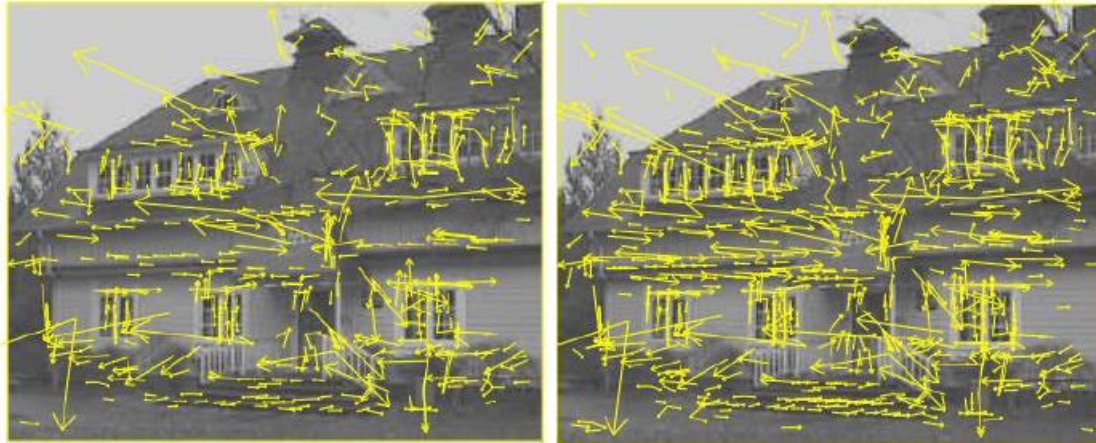
$$\text{Tr}(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta,$$

$$\text{Det}(\mathbf{H}) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta.$$

با محاسبه مقدار زیر به گوشه یا لبه بودن نقاط و در نهایت نگه داشتن آن پی میبریم ($\alpha = r\beta$):

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

برای ۲ مقدار ۱۰ را در نظر گرفته است و اگر نامساوی بالا برقرار باشد یعنی لبه داشته و آن را حذف میکنیم. برای نقاط با کنتراست پایین هم $D(x)$ را بدست آورده و با یک مقدار آستانه بررسی میکنیم. با انجام دو مرحله زیر تصویر راست به تصویر چپ در شکل زیر تبدیل میشود:

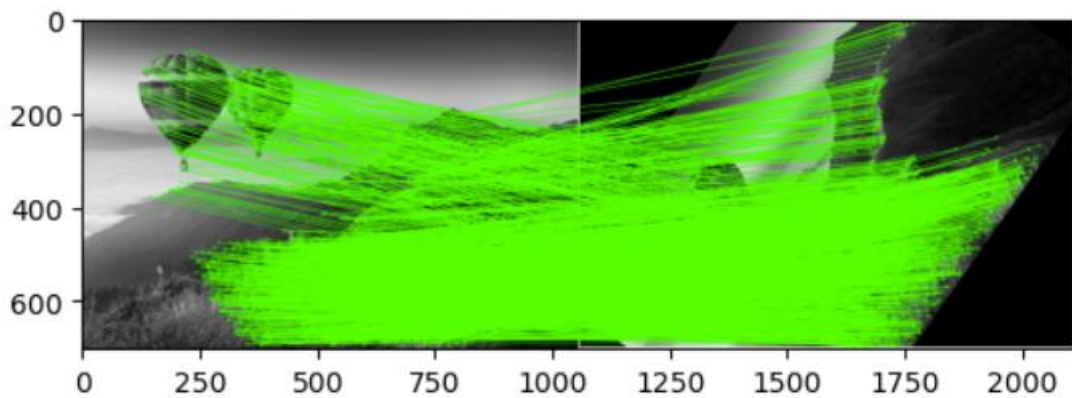
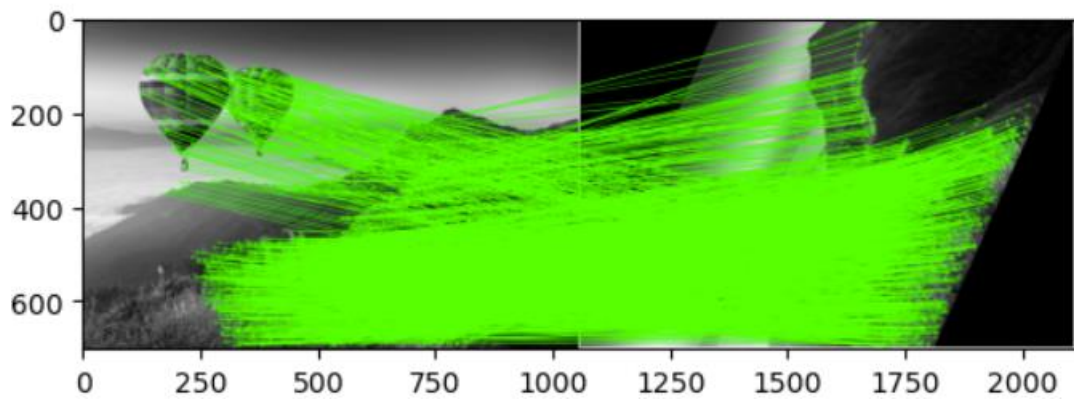


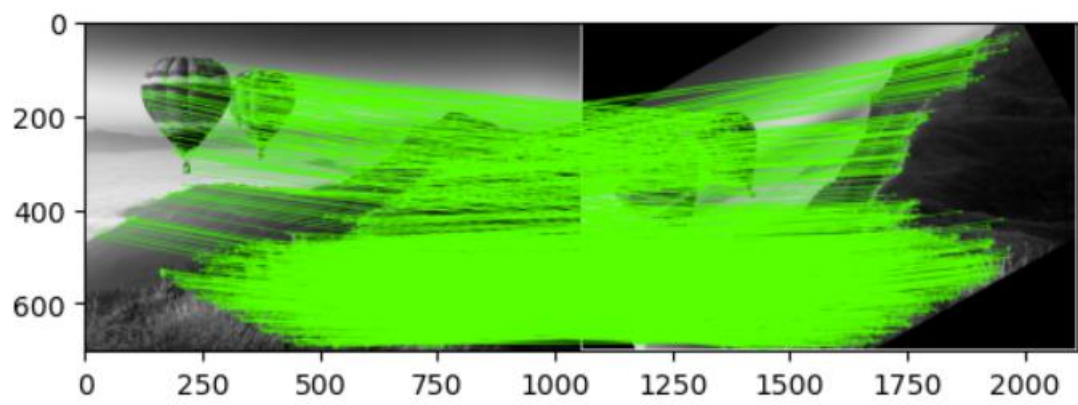
• تخصیص یک وضعیت و جهت به هر نقطه

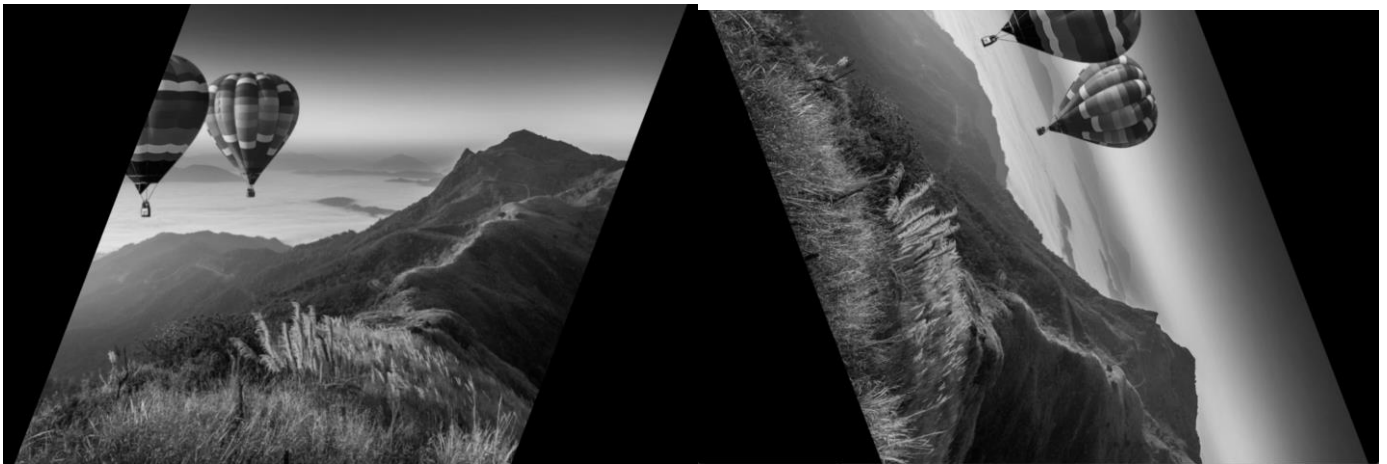
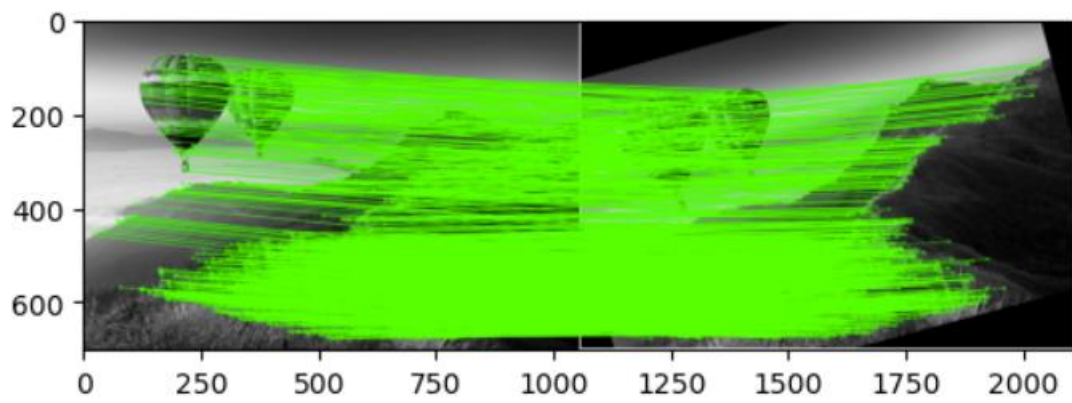
تصویر با مقیاس اولیه را برمیداریم و بعد از آن قدر مطلق و زاویه گرادیان آن نقطه را محاسبه میکنیم. برای اینکه اطلاعاتی که برای نقطه شاخص نشان میدهیم شامل اطلاعات اطراف آن نیز باشد برای همین در محدوده اطراف این نقطه به اندازه ۱.۵ برابر مقیاس به دنبال مهمترین جهتها و اندازهها جستجو میکنیم و این مقدار را برای نقطه شاخص در نظر میگیریم. برای تعیین جهت و اندازه مهم، برای این محدوده یک هستوگرام ۱۰ بازهای برای زاویه در نظر میگیریم و برای هر نقطه را برابر با اندازه گرادیان در جهتی که دارد قرار میدهیم. بعد از ایجاد هستوگرام یک منحنی به هستوگرام با استفاده از نقطه ماکزیمم و دو نقطه اطراف آن برازش میکنیم. از این منحنی مکان دقیق ماکزیمم را دریافته و آن را ثبت میکنیم. در پیدا کردن ماکزیممها و قلهها اگر قلهای ۸۰ درصد بلندترین قله بود با همان مختصات، اما مقادیر گرادیان و جهت متفاوت ثبت میشود.

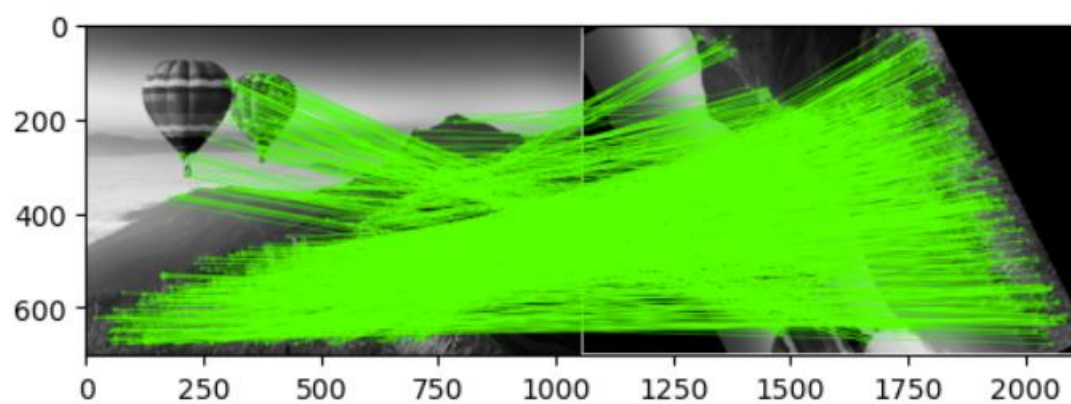
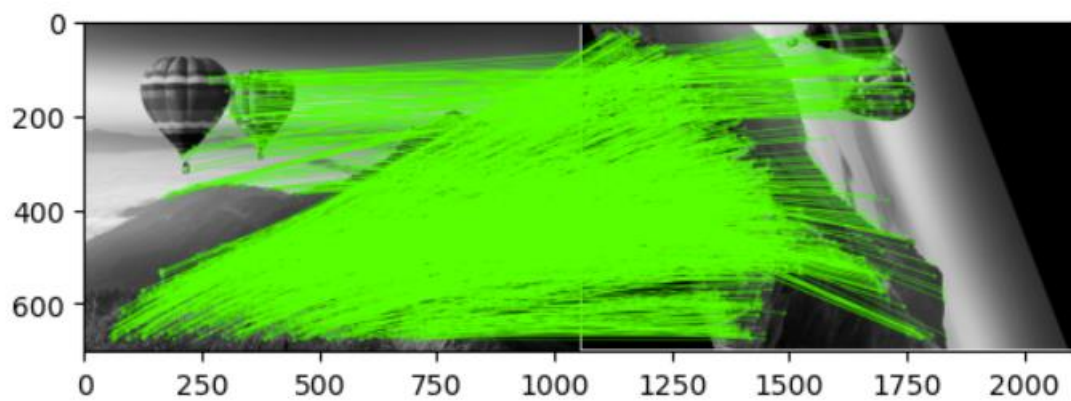
• ساخت توصیفگر

در این مرحله برای ساخت توصیفگر ابتدا برای هر نقطه کلیدی یک همسایگی ۱۶ در ۱۶ در نظر میگیریم و در گام بعد مقدار گرادیان این نقاط اطراف را با یک فیلتر گوسی به مقیاس نصف مقیاس اصلی در نظر میگیریم تا اثر گرادیان در فاصله دورتر کمتر شود. بعد از آن نیز برای اینکه زاویه با چرخش حفظ شود از این مقادیر مقدار داخل نقطه شاخص کم میشود. بعد از آن یک همسایگی ۴ در ۴ بر روی ناحیه تشکیل میدهیم و برای هر خانه ایجاد شده یک هستوگرام ۸ بازهای برای زاویه با اندازه رای گرادیان برای هر پیکسل ایجاد میکنیم. در آخر ۱۶ هستوگرام با ۸ بازه یک عدد ۱۲۸ بعدی را از مقدار کد کرده این ناحیه برای نقطه شاخص را به عنوان توصیف این نقطه به ما میدهد. و در نهایت توصیف نهایی ایجاد میشود. البته برای اینکه نسبت به روشنایی یکنواخت حساس نباشد ۱۲۸ را به طول واحد نرمالیزه میکنیم و برای روشناییهای غیر یکنواخت مقادیر بیشتر از ۰.۲ را به ۰.۲ کاهش داده و بعد آن را نرمالیزه میکنیم.









سرعت sift از freak بهتر است ، اما به طور کلی همانطور که در این تمرین نیز مشاهده شد دقت و عملکرد تشخیص ویژگی و بازگردانی تصاویر در هر دو روش خوب بوده است.
همچنین هردو به یک نسبت به چرخش مقاوم بودند.