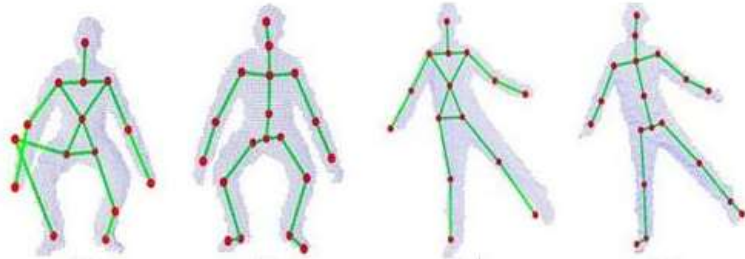


بسمه تعالی



دانشگاه صنعتی شریف
دانشکده مهندسی برق
گروه سیستم‌های دیجیتال



آزمایشگاه یادگیری و بینایی ماشین

دستور کار آزمایش ششم: استخراج نقاط کلیدی

زمان لازم برای انجام آزمایش: حداکثر یک جلسه

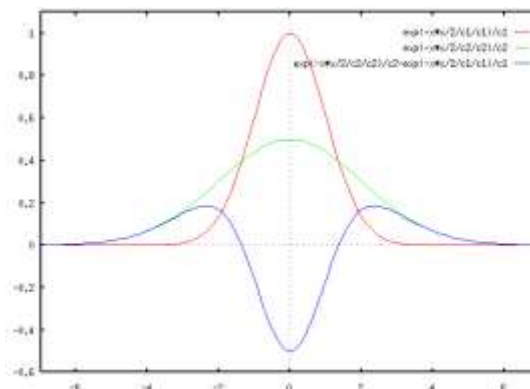
فیلتر LOG و نقاط SIFT

در این آزمایش، به دو روش پرکاربرد LOG و SIFT می‌پردازیم. این دو روش به تفصیل در ادامه تشریح و آزمایش می‌شوند.

ابتدا، به روش LOG می‌پردازیم. این روش یک فیلتر است که به عنوان blub detector (یابنده‌ی شکل لکه مانند در تصویر) عمل می‌کند. رابطه این فیلتر را در زیر مشاهده می‌کنید:

$$\nabla^2 g(x, y; \sigma) = \frac{-1}{2\pi\sigma^4} \left(2 - \frac{x^2 + y^2}{\sigma^2}\right) e^{-\frac{x^2 + y^2}{2\sigma^2}}$$

این فیلتر را می‌توان با تفاضل دو گوسی به صورت زیر تقریب زد. در شکل زیر، نمودارهای قرمز و سبز دو گوسی با انحراف معیارهای σ_1 و σ_2 هستند. حاصل تفاضل آنها نمودار آبی است که همان رفتار LOG را دارد. به این نمودار حاصل از تفاضل دو گوسی، DOG (Difference of Gaussian) گفته می‌شود.



اگر از رابطه توزیع دوبعدی گوسی $G(x, y; \sigma)$ نسبت به σ مشتق بگیریم، می‌بینیم که:

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

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 g$$

اگر مشتق را با تعریف مشتق جایگزین کنیم، در نهایت به تقریب زیر می‌رسیم:

$$\text{LOG}(\sigma) \approx \frac{1}{(K-1)\sigma^2} [G(x, y; K\sigma) - G(x, y; \sigma)]$$

(امتیازی): رابطه بالا را با توجه به توضیحات داده شده، بدست آورید.



توجه فرمایید که تقریب بالا به شرط $K < \sqrt{2}$ تقریب خوبی است. در کد زدن خود، از ترم $\frac{1}{(K-1)\sigma^2}$ صرف نظر کنید.

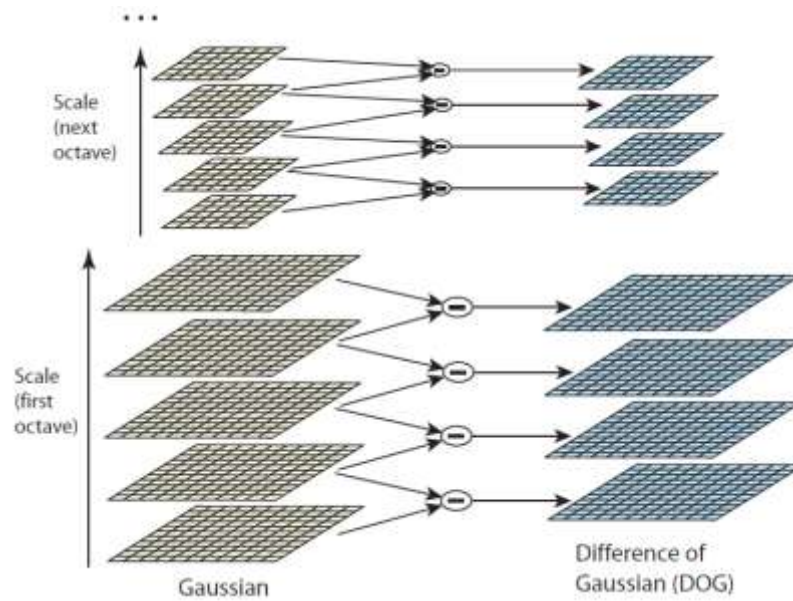
تصویر دلخواهی انتخاب و با کمک DOG، فیلتر LOG را روی آن پیاده کنید. تصویر لبه‌یابی شده را نمایش دهید و روی آن بحث نمایید. برای کد زدن آن، دو بار از دستور GaussianBlur با سایزهای Size(1,1) و Size(3,3) استفاده کنید. انحراف معیارهای آن را 0 قرار دهید تا خودشان محاسبه شوند. و نتیجه را از تفاضل خروجی این دو فیلتر بدست آورید.



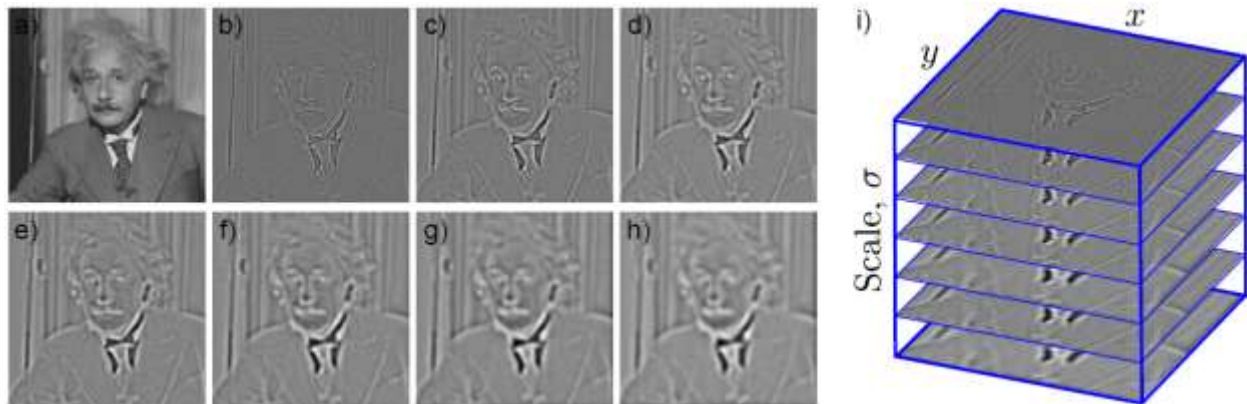
حال، به روش Scale Invariant Feature Transform (SIFT) می‌پردازیم. روش SIFT در سال ۲۰۰۴ توسط David Lowe ارائه شد که بسیار روش قوی‌ای است. ما در اینجا این روش را توضیح می‌دهیم. ولی برای مطالعه جزئی‌تر این روش به مقاله او با نام Distinctive Image Features from Scale-Invariant Keypoints مراجعه کنید.

روش SIFT یک روش برای یافتن ویژگی‌های اشیا در تصویر است که از این ویژگی‌ها می‌توان در یافتن و شناسایی همان اشیا در تصاویر دیگر هم بهره برد. جالبی این روش آن است که نسبت به دَوَران و scale مقاوم است.

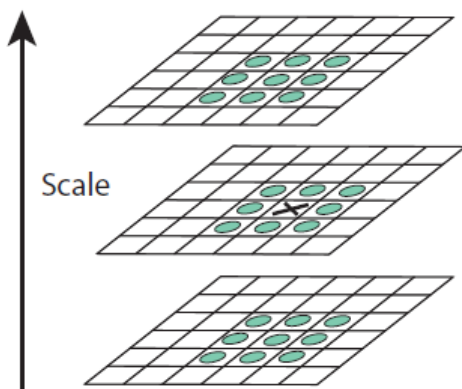
در این روش، در ابتدا، یک سری فیلتر گوسی با σ های گوناگون بر روی تصویر ورودی اعمال می‌شود. انحراف معیار این گوسی‌ها، K برابر K برابر زیاد می‌شود. هر چند تا گوسی را در کنار هم یک octave می‌نامیم. انحراف معیار گوسی‌های هر اکتاو، ۲ برابر انحراف معیار گوسی‌های اکتاو قبلی است و لذا، نیاز به محاسبه دوباره ندارد و کافی است نتایج اکتاو قبلی را فقط با نرخ 0.5 downsample کنیم. این فیلترهای گوسی به تصویر ورودی اعمال می‌شود و با دوبه‌دوی آنها، یک DOG ایجاد می‌شود. این توضیحات را در شکل زیر مشاهده می‌کنید.



نمونه‌ای از این فیلتر شده‌ی تصویر توسط فیلترهای DOG با انحراف معیارهای متفاوت را در زیر مشاهده می‌کنید.



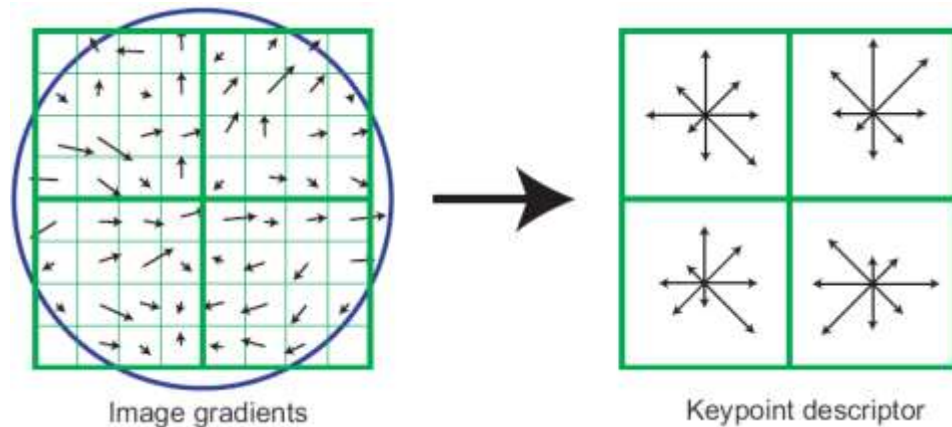
در مرحله بعدی، بین هر سه لایه از DOG های یافت شده، ماکزیمم می‌گیریم. به این صورت که، یک پنجره $3 \times 3 \times 3$ به صورت سه‌بعدی روی این سه لایه می‌رود و پیکسل وسط فقط در صورتی نامزد feature شدن می‌شود که از بقیه ۸ تایی اطرافش بزرگتر باشد.



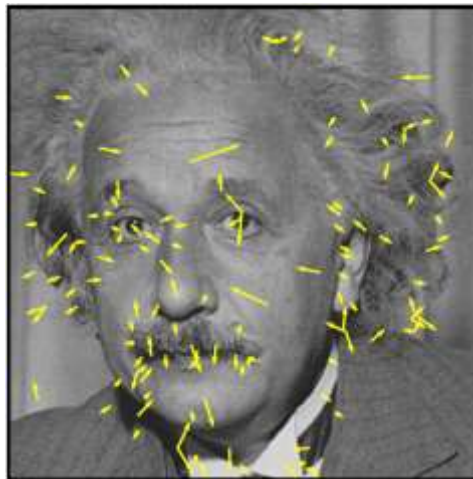
سپس، برای حذف ویژگی‌های پرت (outlier)، آن پیکسل‌هایی را که مقدار DOG آنها از یک حد آستانه (مثلاً 0.03) کمتر باشد، رد می‌کنیم. نمونه‌ای از حذف outlier را در شکل زیر مشاهده می‌کنید.



پس از آن که تمام نقاط ویژگی (پیکسل ویژگی) (keypoint) یافت شده است. حال، باید بردار ویژگی را تشکیل دهیم. برای این کار، از تصویر گرادیان ∇f می‌گیریم. یک پنجره 16×16 در اطراف هر پیکسل ویژگی در نظر می‌گیریم. سپس، با توجه به اندازه گرادیان و جهت گرادیان در این پنجره، یک هیستوگرام زاویه گرادیان تشکیل می‌دهیم. دقت کنید که هر پیکسل که می‌خواهد زاویه‌اش در این هیستوگرام وارد شود، با توجه به مقدار اندازه گرادیان آن، وزن می‌گیرد. کل 360 درجه را به 36 درجه‌ای تقسیم می‌کنیم. و سپس، حداکثر bin هیستوگرام را را به عنوان زاویه آن پیکسل ویژگی در نظر می‌گیریم.



نمونه‌ای از زوایای نقاط ویژگی را در زیر مشاهده می‌کنید.



حال، این جا دست نگه می‌داریم و آن زاویه را ذخیره می‌کنیم. سپس، همان بلوک 16×16 اطراف پیکسل ویژگی را به ۱۶ تا بلوک کوچکتر 4×4 تقسیم می‌کنیم. حال، به درون یکی از بلوک‌های کوچکتر (با سایز 4×4) می‌رویم. دوباره روی زاویه گرادیان در آن هیستوگرام (ولی این بار با ۸ تا bin) می‌زنیم. این کار را برای هر یک از این ۱۶ بلوک کوچکتر داخل بلوک بزرگتر، انجام می‌دهیم. هر یک از بلوک‌های کوچکتر به ما یک بردار هیستوگرام ۸ المانی داد. با کنار هم قرار دادن بردارهای این ۱۶ بلوک کوچک، یک بردار ویژگی با سایز $1 \times (16 \times 8)$ یعنی 128×1 بدست می‌آید.

حال، به زاویه برتر آن بلوک بزرگ برمی‌گردیم. تمام زاویه‌های (مولفه‌های) آن بردار ویژگی را منهای این زاویه می‌کنیم. همان‌طور که واضح است، این کار باعث می‌شود بردار ویژگی آن پیکسل ویژگی نسبت به زاویه کلی آن مقاوم باشد و لذا نسبت به دوران کلی مقاوم می‌شویم.

اکنون یک سری نقاط ویژگی با بردارهای با سایز 128×1 داریم که **keypoint descriptors** نام دارند. حال، برای تطبیق (یافتن) یک شیء خاص در دو تصویر متفاوت، در هر دو تصویر بردارهای ویژگی را کشف می‌کنیم. اکنون یک بردار ویژگی در تصویر دوم را در نظر می‌گیریم. فاصله اقلیدسی این بردار ویژگی را با تک‌تک بردارهای ویژگی در تصویر اول محاسبه کرده و کمترین فاصله اقلیدسی به ما ارتباط آن دو بردار ویژگی را نشان می‌دهد. نمونه‌ای از آن را در شکل زیر مشاهده می‌کنید. به این مرحله، **keypoint matching** گفته می‌شود.



همان‌طور که در شکل بالا دیدید، با این کار، شیء در تصویر دیگر، حتی اگر چرخیده باشد و یا اندازه‌اش عوض شده باشد، باز هم قابل پیدا شدن و شناسایی است.

ولی شاید به دلیل وجود نویز، این **matching** دچار خطا شود و یا چندین نقطه ویژگی بسیار نزدیک به بردار ویژگی آن یافت شود. لذا، اگر مثلاً f_1 و f_3 به ترتیب نقاط ویژگی شبیه نقطه ویژگی تصویر دوم باشد، و اگر نسبت $\frac{f_1}{f_3}$ بیشتر از 0.8 باشد (یعنی f_1 و f_3 خیلی به هم نزدیک باشند)، این نقطه ویژگی برای **matching** خوب نیست و دور ریخته می‌شود.

دو تصویر **img1.jpg** و **img2.jpg** را از درس افزار دانلود کنید. با استفاده از دستور مناسب، نقاط کلیدی آن‌ها را استخراج کنید و نمایش دهید. پارامترهای تابعی که برای استخراج نقاط کلیدی استفاده کرده‌اید را بررسی کنید و در مورد آن‌ها را توضیح دهید. این پارامترها را تغییر دهید و نتیجه را بررسی کنید.



با استفاده از تابع مناسب، برای هر کدام از نقاط کلیدی تصویر `img1.jpg`، نزدیک ترین نقطه کلیدی در `img2.jpg` را پیدا کنید. تناظرهای بدست آمده را نمایش دهید و تحلیل خود را بنویسید.



از هر کدام از دو تصویر `img1.jpg` و `img3.jpg`، ۱۰۰۰ نقطه کلیدی استخراج کنید. سپس ۲۰۰۰ نقطه بدست آمده را با استفاده از روش `k-means` به ۱۰ خوشه تقسیم کنید. حال هیستوگرام نقاط کلیدی هر خوشه را برای هر کدام از دو تصویر `img1.jpg` و `img3.jpg` رسم کنید و مشاهده خود را توضیح دهید.



فرض کنید مجموعه ای از تصاویر مشابه دو تصویر قسمت قبل، به عنوان مجموعه تصاویر آموزشی در اختیار دارید. با توجه به مشاهده ای که در قسمت قبل داشتید، روشی برای ساخت یک طبقه بند که این دو نوع تصویر را از هم تفکیک کند، بیان کنید. نحوه استفاده از این طبقه بند در فاز تست را نیز توضیح دهید. دقت کنید که مجاز به استفاده از تصاویر تست در هیچ قسمتی از فاز آموزش نیستیم.



پیش گزارش

۱ - روش SIFT چگونه خود را نسبت به دوران مقاوم میکند؟

۲ - روش خوشه بندی k-means را توضیح دهید. آیا خوشه ها به مقادیر اولیه مراکز خوشه ها وابسته هستند؟
چطور می توان از روش k-means برای خوشه بندی استفاده کرد به نحوی که خوشه ها به مقادیر اولیه مراکز خوشه ها وابسته نباشند؟