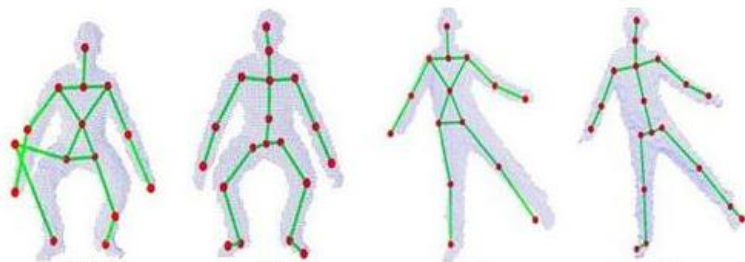


بسمه تعالی



دانشگاه صنعتی شریف
دانشکده مهندسی برق
گروه سیستم‌های دیجیتال



آزمایشگاه یادگیری و بینایی ماشین

دستور کار آزمایش هفتم: توصیفگرهای LBP، Gabor و HOG

زمان لازم برای انجام آزمایش: حداکثر دو جلسه

روش‌های LBP، Gabor و HOG

در این آزمایش با سه روش قدرتمند LBP، Gabor و HOG آشنا می‌شوید. هر یک از این روش‌ها در ادامه شرح داده شده و آزمایش‌هایی با آنها انجام می‌دهید.

اکنون، به روش استخراج ویژگی (LBP) Local Binary Pattern می‌پردازیم. این روش، یک روش بسیار ساده ولی در عین حال موثر برای استخراج ویژگی است. روش کار LBP به این صورت است که یک پنجره که رو تصویر می‌لغزد، اختیار می‌کند و آن پنجره را به سلول‌های مثلاً 8×8 پیکسل تقسیم می‌کند. سپس، برای هر پیکسل در یک سلول به صورت زیر عمل می‌کند. در اطراف هر پیکسل، یک پنجره 3×3 در نظر می‌گیرد. این روش برخلاف روش‌های قبلی که اکثراً با گرادیان سروکار داشتند، با شدت روشنایی پیکسل کار دارد. اگر شدت روشنایی پیکسل همسایه از پیکسل وسط بیشتر باشد، برای آن ۱ و در غیر اینصورت صفر قرار می‌دهیم. برای درک بهتر به شکل زیر توجه کنید.

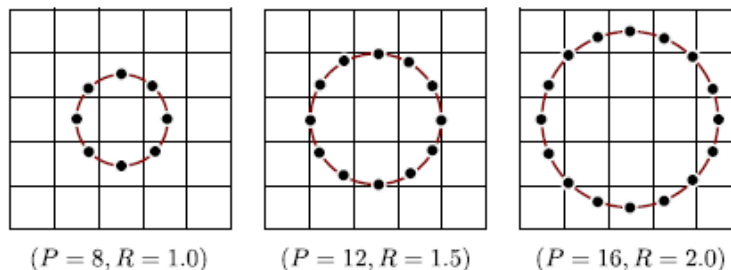
5	4	3
4	3	1
2	0	3

⁰ 1	¹ 1	² 1
⁷ 1		³ 0
⁶ 0	⁵ 0	⁴ 1

$$\text{LBP} = 10010111 = 151$$

این صفر و یک‌ها را کنار هم قرار می‌دهیم. حال، این رشته بیت باینری را آنقدر شیفت می‌دهیم که کوچکترین عدد حاصل شود. با این کار، نسبت به دوران مقاوم می‌شویم. مثلاً در شکل بالا، آن را به صورت 00101111 در می‌آوریم. پس برای هر پیکسل، یک عدد ۸ بیتی تشکیل می‌شود. هیستوگرام اعداد را روی هر سلول تشکیل می‌دهیم. بردارهای هیستوگرام تمام سلول‌ها را کنار هم قرار می‌دهیم تا بردار ویژگی پنجره یافت شود.

برخی از روش‌های LBP هم همسایه‌ها را به صورت‌های زیر در نظر می‌گیرند:



یک فیلتر LBP طراحی کرده و آن را بر تصویر داده شده اعمال کنید و بردار ویژگی آن را تشکیل



دهید.

حال، به روش گابور (Gabor) می‌پردازیم. این روش یک خانواده فیلتر است که ویژگی‌های مهم و تفکیک‌کننده‌ای از تصویر استخراج می‌کنند. فیلترهای گابور دوبعدی برای تصویر به کار می‌روند. رابطه‌ی این فیلترها به صورت زیر است:

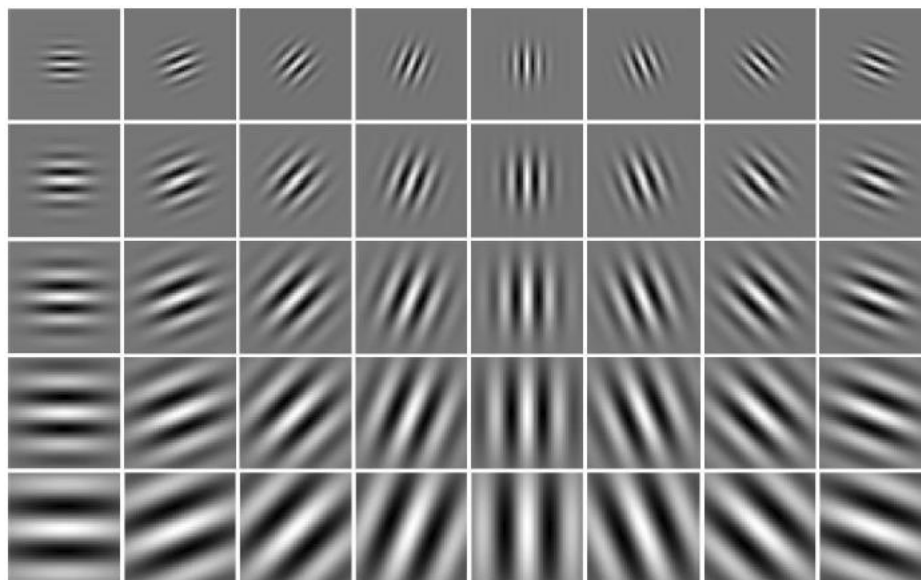
$$g(x, y; \lambda, \theta, \phi, \sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \sin\left(\frac{2\pi x'}{\lambda} + \phi\right)$$

$$x' = x\cos\theta + y\sin\theta$$

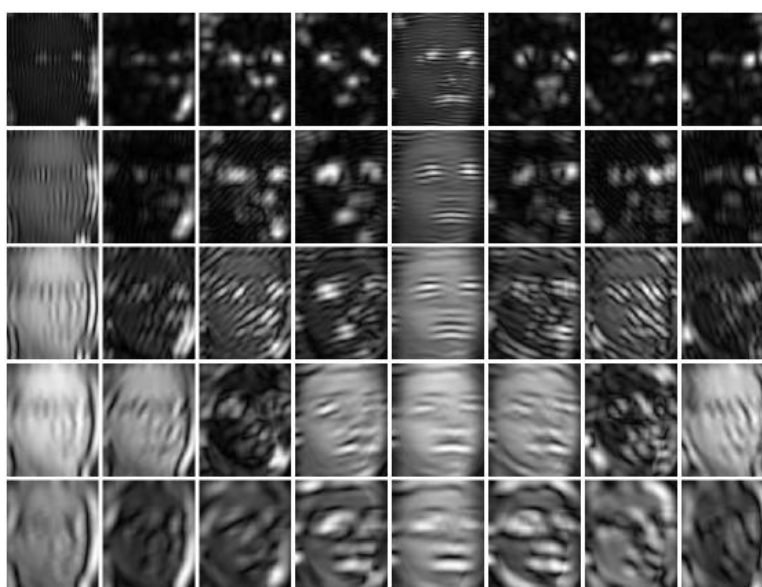
در رابطه بالا، σ انحراف معیار گوسی، θ زاویه فیلتر، λ طول موج سینوسی و ϕ هم اختلاف فاز فیلتر است. برای پیاده‌سازی، یک حلقه روی θ می‌اندازند و یک حلقه داخلی هم روی λ می‌اندازند. مقدار σ هم از روی λ بدست می‌آید.

اگر این فیلترها را رسم کنیم، یک بانک فیلتر به صورت زیر بدست می‌آید:

Gabor Filter Bank



این فیلترها بسیار مهم و تفکیک‌کننده هستند. برای درک اهمیت آنها، مثلاً به این موضوع توجه فرمایید که در روش شبکه عصبی کانولوشنال، معمولاً فیلترهای نهایی که پس از روزها آموزش بدست می‌آید، بسیار شبیه به فیلترهای گابور هستند. نمونه‌هایی از این خروجی این فیلترها را بر روی تصویر صورت مشاهده می‌کنید. این تصاویر همان ویژگی‌های استخراج شده هستند.





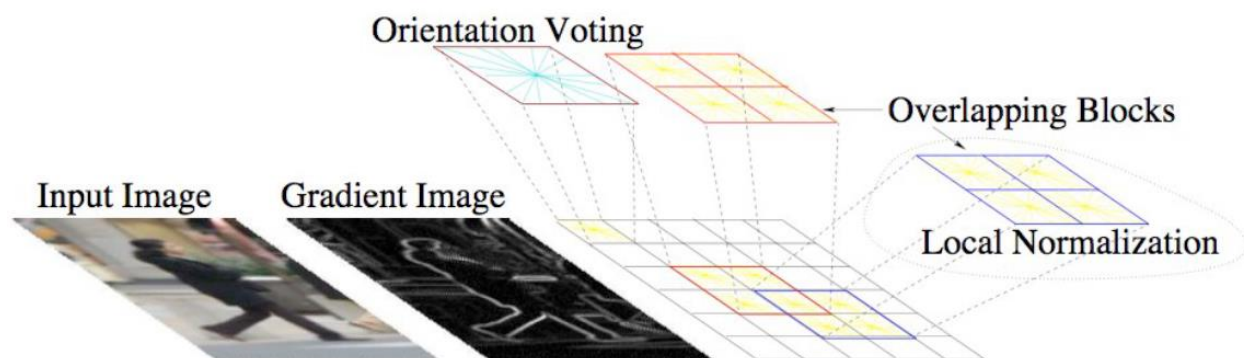
بر روی تصویر داده شده فیلترهای Gabor با زوایای ۴۵، ۹۰، ۱۳۵ و ۱۸۰ درجه اعمال کنید. خروجی‌ها را نمایش دهید. سپس، در زاویه ۴۵، مقدار پارامترهای تابع فیلتر گابور را تغییر دهید و نتایج را برای تغییرات پارامترها بررسی و تحلیل کنید.

در این آزمایش، با روش Histogram Of Gradient (HOG) آشنا می‌شوید. این روش مانند SIFT یک descriptor است و تلاش برای یافتن و بیان ویژگی‌های اشیاء در تصویر دارد. این روش البته بیشتر برای تشخیص عابر پیاده در تصویر استفاده شده است.

در این روش، ابتدا از تصویر ورودی، گرادیان گرفته می‌شود. سپس، یک پنجره با سایز مشخصی در نظر گرفته می‌شود که این پنجره روی تصویر می‌جهد. یک بلوک 16×16 پیکسل را که از چهار سلول کوچکتر 8×8 پیکسل تشکیل شده است، روی تصویر تحت این پنجره، طوری می‌لغزانیم که هر بلوک با بلوک مجاورش در دو سلول هم‌پوشانی داشته باشند. پس اگر پنجره مثلاً به 6×5 سلول تقسیم شده باشد، هر بلوک به صورت 2×2 سلول باشد و بلوکهای مجاور در دو سلول هم‌پوشانی داشته باشند، آن وقت 5×4 بلوک در این پنجره خواهیم داشت. سپس، برای زوایای گرادیان، یک هیستوگرام وزن‌دار روی هر سلول می‌زنیم و ۹ تا bin بین صفر تا ۱۸۰ درجه برای هیستوگرام در نظر می‌گیریم. سپس، برای هر بلوک، بردار ویژگی (descriptor) را به صورت زیر تشکیل می‌دهیم:

$$V = [\text{hist}(\text{cell}1), \dots, \text{hist}(\text{cell}4)]$$

که در آن، $\text{hist}(\text{cell}_i)$ یک بردار 1×9 است و لذا V یک بردار 1×36 است. سپس، این بردار V را نرمالیزه می‌کنیم یعنی بر اندازه آن تقسیم می‌کنیم. در شکل زیر، این مراحل را به صورت تصویری مشاهده می‌کنید.

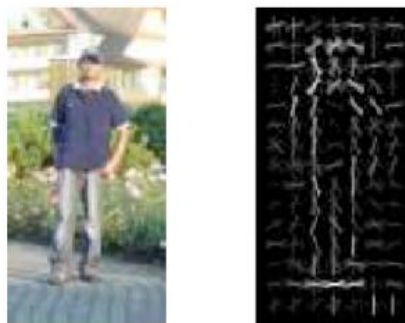


سپس، بردار ویژگی کلی به صورت زیر می‌شود (اگر N تا بلوک داشته باشیم):

$$V_{total} = [V_{block1}, \dots, V_{blockN}]$$

که V_{block_i} یک بردار 1×36 است و لذا V_{total} یک بردار $1 \times (36N)$ است. حال، این بردارهای ویژگی از تصاویر آموزش، توسط یک کلاس‌بند قوی مثل SVM آموزش داده می‌شود و از آن برای تشخیص انسان یا اشیاء آموزش داده شده در تصاویر تست استفاده می‌شود.

نمونه‌ای از این گرادیان‌های مربوط به روش HOG برای تشخیص انسان را در شکل زیر می‌بینید:



در تصویر داده شده با کمک روش HOG مکان فرد را در تصویر بیابید.



در این قسمت می‌خواهیم یک HOG مناسب برای هدف طبقه‌بندی احساس خوشحال و انزجار طراحی کنیم. ۳۵ تصویر از مجموعه disgust و ۶۵ تصویر از مجموعه happy را به صورت



تصادفی به عنوان مجموعه آموزش، و مابقی نمونه‌ها را به عنوان مجموعه تست جدا کنید. ماسک مناسب به تصاویر اعمال کنید و سپس توصیفگر HOG را برای هر تصویر بسازید. از آنجا که تنها می‌خواهیم طبقه‌بندی انجام دهیم و به دنبال مکان‌یابی نیستیم، کل تصویر را به عنوان یک پنجره در نظر بگیرید. انتخاب سائز هر سلول، سائز هر بلوک و میزان همپوشانی بلوک‌ها در اختیار خود شماست. بعد از تشکیل توصیفگرها، یک طبقه‌بند SVM مناسب (خطی یا غیر خطی بودن SVM نیز در اختیار شماست) با استفاده از مجموعه آموزش تعلیم دهید و نتیجه طبقه‌بندی روی مجموعه تست را به صورت ماتریس درهم ریختگی محاسبه کنید. (دقت نمایید که تنظیم پارامتر C و اگر از SVM غیر خطی استفاده می‌کنید انتخاب نوع کرنل را نیز باید بدون استفاده از نمونه‌های تست انجام دهید).

پیش گزارش

۱- توضیح دهید که در روش HOG نرمالیزه کردن بردار ویژگی چه فایده ای دارد.

۲- در روش LBP زمانی که میخواهند طول بردار ویژگی را کم کنند از روشی به نام Pattern Uniform کمک میگیرند. این روش را مطالعه کنید و توضیح دهید.