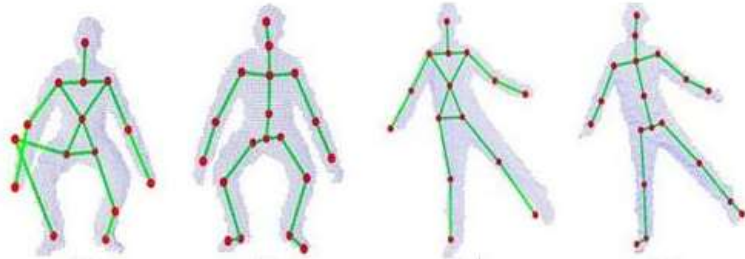


بسمه تعالی



دانشگاه صنعتی شریف
دانشکده مهندسی برق
گروه سیستم‌های دیجیتال



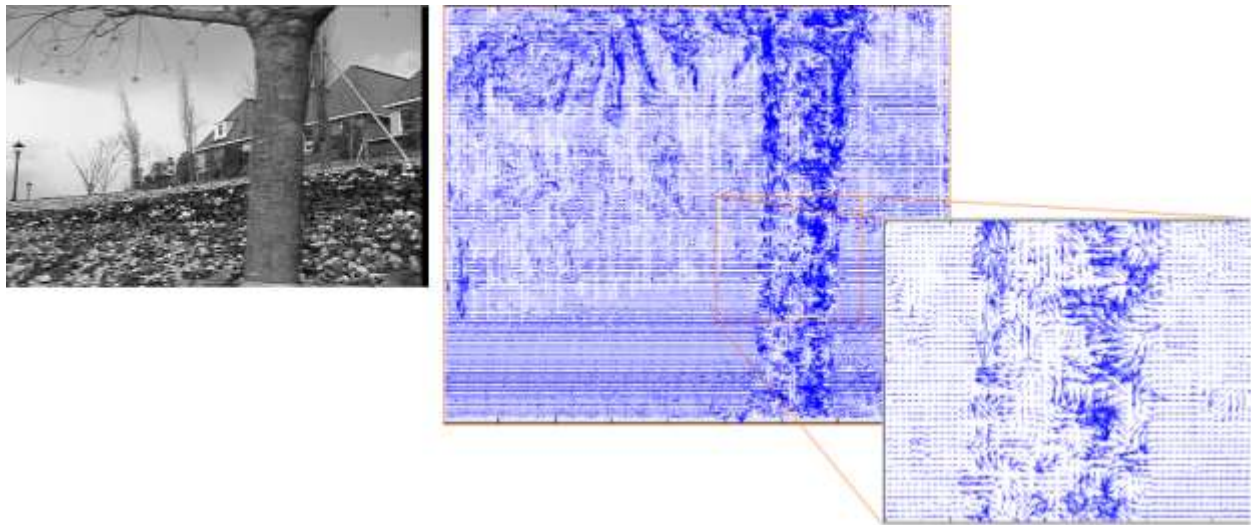
آزمایشگاه یادگیری و بینایی ماشین

دستور کار آزمایش هشتم: شار نوری

زمان لازم برای انجام آزمایش: حداکثر دو جلسه

شار نوری برای حذف پس‌زمینه

در این آزمایش به شار نوری (Optical flow) می‌پردازیم. شار نوری به معنای سرعت مکانی هر پیکسل در دو فریم متوالی از یک فیلم است. به عبارت ساده‌تر می‌خواهیم ببینیم که هر پیکسل، چه قدر از یک فریم به فریم دیگر جابه‌جا شده است و لذا با این کار، کل جابه‌جایی‌های انجام شده بین دو فریم را نمایش دهیم. یافتن شار نوری روش‌های گوناگونی دارد از جمله روش Horn-Schunck و Lucas-Kanade که ما در اینجا به توضیح روش Lucas-Kanade بسنده می‌کنیم. در زیر نمونه‌ای از شار نوری (حرکت دوربین در هنگام فیلم برداری) را مشاهده می‌کنید. همان‌طور که دیده می‌شود، جابه‌جایی را با بردار نشان می‌دهیم و بردارهای با طول صفر (نقطه) به معنی عدم جابه‌جایی است.



در روش‌های یافتن شار نوری، اولین فرضی که انجام می‌شود، این است که فرض می‌شود روشنایی هر پیکسل بین دو فریم متوالی ثابت است و تغییر نمی‌کند. این فرض در صورتی که بین دو فریم متوالی، مثلاً روشنایی روز عوض نشود یا سایه‌ای روی آن پیکسل نیفتد، فرض معقولی است.

$$f(x, y, t) = f(x + dx, y + dy, t + dt)$$

در عبارت بالا، x, y, t و f به ترتیب نمادهای مختصات اول پیکسل، مختصات دوم پیکسل، زمان فریم و روشنایی پیکسل است.

همچنین، فرض می‌کنیم که دو فریم خیلی به هم نزدیک هستند و جابه‌جایی خیلی کم است. لذا می‌توان با تقریب خوبی، $f(x + dx, y + dy, t + dt)$ را تا جمله اول بسط تیلور تقریب زد.

$$f(x + dx, y + dy, t + dt) = f(x, y, t) + \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \frac{\partial f}{\partial t} \Delta t$$

حال اگر طرفین را بر Δt تقسیم کنیم، داریم:

$$\frac{f(x + dx, y + dy, t + dt) - f(x, y, t)}{\Delta t} = f_x u + f_y v + f_t$$

که در آن، $f_x = \frac{\partial f}{\partial x}$ و $f_t = \frac{\partial f}{\partial t}$ و $u = \frac{\partial x}{\partial t}$ و $v = \frac{\partial y}{\partial t}$ است. توجه فرمایید این پنج عبارت، به ترتیب مشتق روشنایی پیکسل نسبت به مکان x ، مشتق روشنایی پیکسل نسبت به مکان y ، مشتق مکان x پیکسل نسبت به زمان، مشتق مکان y پیکسل نسبت به زمان و مشتق روشنایی پیکسل نسبت به زمان است. عبارات f_x و f_y همان مولفه‌های گرادیان تصویر در فریم t هستند که قبلاً با گرادیان تصویر آشنا شدید.

در رابطه بالا، طبق فرض اول دال بر ثابت ماندن روشنایی پیکسل در دو فریم متوالی، به صورت زیر ساده می‌شود:

$$f_x u + f_y v = -f_t$$

در روش Lucas-Kanade، یک معیار پیوستگی مطرح می‌شود که فرض می‌کند که رابطه بالا برای هر پنجره 9×9 از تصویر برقرار است و این عبارت را برای کل پنجره در نظر می‌گیرد. لذا برای هر پنجره مذکور از پیکسل‌ها، سعی در یافتن u و v ای دارد که عبارت زیر را کمینه کند:

$$\min \sum_{i=-1}^1 \sum_{j=-1}^1 \left(f_x(i, j)u + f_y(i, j)v + f_t(i, j) \right)^2$$

دقت شود که اگرچه این عبارت برای کل پنجره اعمال می‌شود، ولی نتیجه u و v ای که از این کمینه کردن بدست می‌آید، فقط برای پیکسل وسط این پنجره محسوب می‌شود. اگر از عبارت بالا، نسبت به u و v مشتق بگیریم و مساوی صفر قرار دهیم (مساله بهینه‌سازی)، خواهیم یافت:

$$\begin{bmatrix} u \\ v \end{bmatrix} = (A^T A)^{-1} A^T b$$

که در آن، داریم:

$$A = \begin{bmatrix} f_{x1} & f_{y1} \\ \vdots & \vdots \\ f_{x9} & f_{y9} \end{bmatrix}$$

$$b = \begin{bmatrix} f_{t1} \\ \vdots \\ f_{t9} \end{bmatrix}$$

(امتیازی): عبارت بالا را اثبات نمایید.



برای درک بهتر و ساده‌تر رابطه بالا، معیار پیوستگی را به صورت زیر می‌نویسیم:

$$\begin{cases} f_{x1}u + f_{y1}v + f_{t1} = 0 \\ \vdots \\ f_{x9}u + f_{y9}v + f_{t9} = 0 \end{cases}$$

$$\rightarrow \begin{bmatrix} f_{x1} & f_{y1} \\ \vdots & \vdots \\ f_{x9} & f_{y9} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_{t1} \\ \vdots \\ f_{t9} \end{bmatrix} \rightarrow A \begin{bmatrix} u \\ v \end{bmatrix} = b \rightarrow A^T A \begin{bmatrix} u \\ v \end{bmatrix} = A^T b \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = (A^T A)^{-1} A^T b$$

به روش حل دستگاه معادلات که در بالا ذکر شد، روش **least square** می‌گویند. اگر دقت کنید، متوجه شباهت این رابطه با رابطه رگرسیون (آزمایش اول) می‌شوید. حال، که این u و v یافت شد، در واقع شار نوری و جابه‌جایی پیکسل‌ها یافت شده است. اگر این u و v را به صورت بردارهای متناسب با اندازه و جهت u و v رسم کنیم، شار نوری را رسم کرده‌ایم.

با استفاده از وبکم لپ‌تاپ خود به صورت بلادرنگ **optical flow** را با دو روش **dense** و **sparse**



پیاده‌سازی کنید. در حالت **optical flow**، **dense**، تمام پیکسل‌های صفحه محاسبه می‌

کنید. **angle** و **magnitude** ای که برای نقاط مختلف بدست آمده را روی یک روش نمایش رنگ مانند **HSV** انتقال دهید و نمایش دهید. در حالت **sparse** شما اول چند نقطه را که مستعد **tracking** هستند (نقاط گوشه نقاط مناسبی می‌باشد) را پیدا می‌کنید و این نقاط را به تابع مرتبط می‌دهید و اگر نقاط بدست آمده شرایط **flow** برایش برقرار بود یک خط بین دو نقطه در مختصات تصویر در زمان متوالی رسم می‌کنید.



در این قسمت می خواهیم ویدیوهای خوشحال و متعجب را از طریق توصیفگرهایی که با استفاده از شارنوری بدست می آوریم، طبقه بندی کنیم. ابتدا تصاویر هر ویدیو را به طریق مناسبی ماسک کنید و ابعاد آن ها را به نصف ابعاد اولیه تغییر دهید.

برای ساخت توصیفگر هر ویدیو، ابتدا شارنوری را برای تمام فریم های آن ویدیو بدست آورید. سپس هیستوگرام زوایای شارنوری هر فریم را محاسبه کنید. تعداد انبارک های هیستوگرام ها را ۸ تا در نظر بگیرید.

برای هر ویدیو، هیستوگرام های شارنوری فریم های مختلف را یک بار از طریق میانگین گیری و بار دیگر از طریق محاسبه بیشینه خلاصه (pool) کنید. (در نتیجه دو هیستوگرام برای هر ویدیو بدست می آید). سپس توصیفگر نهایی را با کنار هم قرار دادن این دو هیستوگرام بسازید. در انتها، برای انجام آزمایش، ۷۰ درصد ویدیوهای موجود در پوشه های **happy videos** و **surprise videos** را به عنوان مجموعه آموزش و مابقی را به عنوان تست جدا کنید. با استفاده از یک طبقه بند مناسب، آموزش و تست را انجام دهید و نرخ طبقه بندی تست را بدست آورید.

پیش گزارش

۱ - در مورد شار نوری با کمک هرم (pyramid) تحقیق کنید و دلیل استفاده از هرم را بیان نمایید. نیازی نیست روابط ریاضی آن را بیان نمایید. به صورت کیفی (ولی نه ناقص) آن را توضیح داده و دلیل آن را شرح دهید.