

اصول بینایی کامپیوتر (۱- ۲۲۹۲۶) بهار ۱۴۰۱
تمرینات سری اول
مهلت تحویل: ۲۷ فروردین ۱۴۰۱ (شنبه) ساعت ۱۲ شب

لطفاً به نکات زیر توجه بفرمایید.

۱. نتایج و پاسخ های خود را در یک فایل zip رار دهید.

۲. کسب نمره در یک سؤال که نیاز به برنامه نویسی دارد مستلزم تحویل سه مورد نتایج، کدها، و گزارش می باشد. چنانچه حتی یکی از این سه مورد تحویل داده نشود، نمره آن سؤال صفر می شود.

۳. نمره کل شما در یک سؤال به کیفیت نتیجه به دست آمده و کامل بودن گزارش بستگی دارد.

۴. نتایج خود را حتماً باید ذخیره کرده و بفرستید، حتی با وجود اینکه برنامه ای که آن نتایج را تولید می کند را می فرستید.

۵. فایل هایی که در صورت سوال ها بوده اند مانند فایل توضیحات سؤالات و تصاویری که در صورت سؤال ها بوده اند را به همراه پاسخ های خود بفرستید. اضافه کردن آنها باعث افزایش حجم فایل پاسخ های شما می شود.

۶. حتماً کدهای استفاده شده که منجر به نتایج فرستاده شده است را باید بفرستید. با اجرای این کدها باید همان نتایجی که فرستاده اید قابل بازیابی باشند. برنامه شما باید بدون نیاز به تغییری قابل اجرا باشد. در صورت اجرا نشدن برنامه شما به هر دلیلی و یا به دست نیامدن نتیجه ای که فرستاده اید به هر دلیلی نمره آن سؤال صفر خواهد بود. در صورت استفاده از فایل های متعدد، تمام آن ها را به همراه پاسخ های خود بفرستید تا برنامه شما قابل اجرا باشد. در چنین مواردی می توانید فایل ها را با نام های دلخواه خود ذخیره نمایید ولی فایل اصلی باید با نام اشاره شده در صورت سؤال ذخیره شود و طوری باشد که با اجرای آن تمام قسمت های برنامه مورد نظر اجرا شود. کدهای شما تماماً باید توسط خودتان نوشته شده باشند. هرگونه استفاده از کد دیگران، اعم از دوستان و اینترنت، به هر شکل ممکن، اعم از کپی کردن یا همکاری کردن، تقلب محسوب می شود.

۷. برای تمام سؤالات، باید جزئیات روشی که استفاده کرده اید را در گزارش توضیح دهید. گزارش باید فقط در قالب فایل pdf باشد. گزارش به معنی کامنت نوشتن در کدها نیست. از کپی و پیست کدهای خود در گزارش و کامنت گذاشتن روی آنها خودداری کنید. گزارش شما نباید حاوی کدها باشد. باید روش استفاده شده را در گزارش طوری توضیح دهید که هر شخص آشنا با مطالب درس بتواند کاری که شما انجام داده اید را پیاده سازی نماید.

۸. در صورتیکه در انجام دادن تمرینات خود از هم فکری دیگران استفاده نموده اید باید نام آنها را در ابتدای گزارش خود ذکر نمایید.

۱. Harris Corner Detection and Matching (۲۵ نمره)

تصاویر im01.jpg و im02.jpg را به ترتیب تصاویر اول و دوم می نامیم. با استفاده از روش هریس (Harris) تعدادی نقطه مطلوب (interest point) در هریک از تصاویر پیدا کنید. در اطراف هر نقطه یک همسایگی در نظر بگیرید و ناحیه داخل آن را توسط یک بردار ویژگی (features vector) توصیف کنید. بردار ویژگی نقاط بین دو تصویر را با هم مقایسه کرده و نقاط متناظر (corresponding points) را پیدا کنید.

برای هر یک از تصاویر مراحل می شود را انجام دهید.

- مشتق تصویر در هر پیکسل را در راستای عمودی (I_x) و راستای افقی (I_y) به دست آورید. در تصاویر رنگی، برای هر پیکسل، در هر سه کانال رنگی مشتق در راستای عمودی را به دست بیاورید. سپس از بین این سه مقدار، مقداری که قدر مطلق آن بزرگترین است را به عنوان مشتق در راستای عمودی در آن پیکسل در نظر بگیرید. برای مشتق در راستای افقی هم به صورت مشابه عمل کنید. می توانید از توابع و کتابخانه های آماده استفاده نمایید.
- مقادیر I_x^2 ، I_y^2 و $I_x I_y$ را برای هر پیکسل محاسبه کنید.
- بزرگی گرادیان تصویر را در هر پیکسل به دست آورده و نمایش دهید. بزرگی گرادیان را می توانید از رابطه $\sqrt{I_x^2 + I_y^2}$ به دست آورید. دقت نمایید که برای نمایش بزرگی گرادیان محدوده مقادیر گرادیان را باید در بازه ای قرار دهید که قابل نمایش باشند.
- مقادیر گرادیان تصویر اول را با نام res01_grad.jpg و تصویر دوم را با نام res02_grad.jpg ذخیره نمایید.
- یک فیلتر گوس در نظر گرفته و روی ماتریس های I_x^2 ، I_y^2 و $I_x I_y$ اعمال نمایید. نتایج را با نام های S_x^2 ، S_y^2 و S_{xy} نمایش دهید. برای این کار می توانید از توابع و کتابخانه های آماده استفاده نمایید. مقدار مناسب **انحراف معیار** و **اندازه فیلتر** استفاده شده را به صورت تجربی به دست آورده و در گزارش خود ذکر نمایید. بدین ترتیب، برای هر پیکسل مقدار درایه های ماتریس تنسور ساختار (structure tensor) که میانگین وزن دار داخل یک همسایگی در اطراف آن پیکسل می باشند را به دست می آورید.
- دترمینان (det) و اثر (trace) تنسور ساختار را برای هر پیکسل به دست آورید. دترمینان $S_x^2 - S_{xy}^2$ و اثر $S_x^2 + S_y^2$ است.
- مقدار تابع هریس $(R = \det - k (\text{trace})^2)$ را برای هر پیکسل به دست آورید. ماتریس حاصل را نمایش دهید. دقت نمایید که محدوده مقادیر را باید در بازه ای قرار دهید که قابل نمایش باشد. مقادیر مربوط به تصویر اول را با نام res03_score.jpg و تصویر دوم را با نام res04_score.jpg ذخیره نمایید.
- در قسمت قبل، مقدار k را به صورت تجربی انتخاب نموده و در گزارش خود ذکر نمایید. برای انتخاب مقدار مناسب، می توانید تابع $z = x y - k (x + y)^2$ را برای مقادیر مختلف k بررسی نمایید.
- یک مقدار آستانه (threshold) در نظر گرفته و مقادیر بیشتر از آن را نگه داشته و مقادیر دیگر را حذف نمایید. حاصل را نمایش دهید. مقادیر مربوط به تصویر اول را با نام res05_thresh.jpg و تصویر دوم را با نام res06_thresh.jpg ذخیره نمایید. مقدار آستانه را در گزارش خود ذکر نمایید.
- از روش non-maximum suppression که پیکسل با بیشترین مقدار در بین پیکسل های یک مؤلفه را به دست می آورد استفاده نموده و برای تعدادی نقطه مجاور هم که بیشتر از مقدار آستانه بوده اند فقط یک نقطه انتخاب کرده و بقیه را حذف نمایید. نقاط باقیمانده را روی تصاویر اصلی نشان داده و ذخیره نمایید. نتیجه مربوط به تصویر اول را با نام res07_harris.jpg و تصویر دوم را با نام res08_harris.jpg ذخیره نمایید.
- برای هریک از نقاط به دست آمده، یک همسایگی به اندازه $n \times n$ در نظر بگیرید و پیکسل های داخل این همسایگی را در یک بردار n^2 بعدی قرار دهید. مقدار مطلوب n را به صورت تجربی به دست آورده و در گزارش خود ذکر نمایید. بدین صورت، برای هر یک از نقاط مطلوب یک بردار ویژگی به دست می آید.

حال که در هریک از تصاویر تعدادی نقطه مطلوب به دست آمده است، نقاط متناظر در دو تصویر را پیدا کنید. برای این کار، مراحل که در ادامه توضیح داده می شود را انجام دهید.

- بردارهای ویژگی تمام نقاط دو تصویر را با هم مقایسه نمایید. برای مقایسه بردارهای ویژگی می توانید از فاصله L_1 یا L_2 یا کسینوسی استفاده کنید. فاصله را انتخاب کرده و در گزارش خود ذکر کنید.
- برای هر نقطه از تصویر اول (q)، دو نقطه از تصویر دوم که کمترین فاصله را با آن دارند در نظر بگیرید. نزدیک ترین نقطه را p_1 و دومین نزدیک ترین را p_2 بنامید. فاصله آن ها تا q را d_1 و d_2 بنامید.
- اگر نسبت d_2 به d_1 کمتر از یک مقدار آستانه شد، q متناظر p_1 می شود، در غیر این صورت، q هیچ نقطه متناظری ندارد. مقدار آستانه را انتخاب نموده و در گزارش خود ذکر نمایید.
- همین کار را برای نقاط از تصویر دوم به اول انجام دهید.
- چنانچه یک جفت نقطه هم در تناظر از تصویر اول به دوم و هم در تناظر از تصویر دوم به اول متناظر به دست آمده بود، آن را به عنوان تناظر خوب نگهداشته و بقیه را حذف نمایید.
- چنانچه یک نقطه از یک تصویر از بیش از یک نقطه از تصویر دیگر متناظر شده بود، تمام آن ها را حذف نمایید.
- نقاط مطلوب باقیمانده را در تصاویر خودشان نمایش دهید. نتیجه مربوط به تصویر اول را با نام res09_corres.jpg و تصویر دوم را با نام res10_corres.jpg ذخیره نمایید.
- دو تصویر را کنار هم قرار داده و تعدادی از نقاط متناظر به دست آمده را روی آن ها نمایش داده و بین دو نقطه متناظر در دو تصویر یک خط بکشید. تعداد نقاط را به اندازه ای انتخاب نمایید که نتیجه حاصل قابل ملاحظه باشد. نتیجه را با نام res11.jpg ذخیره نمایید.

روش خود را به طور کامل در فایل pdf گزارش توضیح دهید. فایل کد اصلی خود را با نام q1.py ذخیره نمایید. دقت بفرمایید که تعداد ۱۱ تصویر که در بالا ذکر شده اند را ذخیره نمایید. همچنین تعداد ۷ پارامتر که شامل انحراف معیار، اندازه فیلتر، n ، k ، نوع تابع فاصله، و دو مقدار آستانه می باشند باید در گزارش شما ذکر شده باشند.

۲. Perspective (۳۰ نمره)

قصد داریم آرم دانشگاه صنعتی شریف را در وسط زمین فوتبال با چمن های کوتاه تر بسازیم. اگر طرح آرم را به شکل اصلی روی زمین بسازیم، هنگام فیلم برداری تلویزیونی توسط دوربین هایی که از کنار زمین فیلم برداری می کنند این طرح به دلیل پرسپکتیو به شکل واقعی دیده نمی شود. هدف ساختن آرم تغییر یافته در وسط زمین به گونه ای می باشد که در فیلم برداری توسط دوربین مرکزی زمین، آرم به شکل اصلی خود دیده شود. منظور از دوربین مرکزی زمین دوربینی است که در وسط راستای طولی زمین و در کنار زمین قرار دارد و در جهت های مختلف می چرخد.

قصد داریم تصویری از آرم تغییر یافته به دست آورده و به پیمانکار بدهیم تا طبق آن تصویر مناسب روی زمین ساخته شود. برای این کار، فرض می کنیم تصویر آرم دانشگاه که در فایل logo.png می باشد توسط دوربین فیلم برداری دیده شده است. این تصویر را طوری تغییر می دهیم که طرحی که روی زمین باید ساخته شود به دست آید. این طرح مشابه تصویری می باشد که توسط دوربینی که به صورت عمود به پایین در وسط زمین قرار بگیرد گرفته می شود. این دوربین فرضی را هم ارتفاع مکان دوربین تلویزیون در نظر می گیریم، فاصله کانونی آن را مساوی فاصله کانونی دوربین قرار می دهیم، و تصویر آن را به اندازه کافی بزرگ در نظر می گیریم تا تمام آرم در آن جای شده و نقطه وسط زمین در وسط تصویر قرار گیرد. دوربین تلویزیونی در فاصله ۴۰ متری از نقطه وسط زمین و در امتداد خط عرضی وسط زمین و در ارتفاع ۲۵ متری قرار دارد. فاصله کانونی دوربین ۵۰۰ پیکسل می باشد.

تصویر مطلوب را به دست آورده و با نام res12.jpg ذخیره نمایید. روش خود را به طور کامل در فایل pdf گزارش توضیح دهید. فایل کد اصلی خود را با نام q2.py ذخیره نمایید.

تصاویر im03.jpg و im04.jpg را به ترتیب تصاویر اول و دوم می نامیم. ساختمان تصویر اول قسمتی از تصویر دوم است. تصویر اول تقریباً از رو به رو گرفته شده است ولی تصویر دوم زاویه پرسپکتیو زیادی دارد. هدف در این مسأله ایجاد یک تصویر جدید از تصویر دوم می باشد به طوری که تصویر حاصل از زاویه دید تصویر اول شود.

مراحل زیر را انجام دهید:

- تعدادی نقطه مطلوب در دو تصویر به دست آورید. برای این کار از SIFT استفاده کنید. می توانید از توابع و کتابخانه های موجود استفاده کنید.
- نقاط به دست آمده را با رنگ سبز روی هر یک از تصاویر نمایش داده و ذخیره نمایید. هر دو تصویر را در یک فریم کنار هم نمایش دهید: تصویر اول سمت چپ، تصویر دوم سمت راست. نقاط به دست آمده را روی تصاویر با رنگ سبز نمایش دهید. نتیجه را با نام res13_corners.jpg ذخیره نمایید.
- نقاط متناظر بین دو تصویر را با SIFT به دست آورده و در دو تصویر با رنگ آبی نشان دهید. همانند قسمت قبل، دو تصویر را در یک فریم نمایش داده، تمام نقاط مطلوب را با رنگ سبز نمایش دهید، و نقاطی که متناظر شده اند را با رنگ آبی نمایش دهید. نقاط آبی را به اندازه کافی بزرگ و قابل دیدن نمایش دهید. نتیجه را با نام res14_correspondences.jpg ذخیره نمایید.
- تمام نقاط متناظر را به هم وصل نمایید. در تصویر به دست آمده در مرحله قبل، نقاطی که با هم متناظر شده اند را با یک خط آبی به هم وصل کنید. نتیجه را با نام res15_matches.jpg ذخیره نمایید. با کمی دقت به احتمال زیاد قادر خواهید بود تعدادی نقطه که به اشتباه به هم متناظر شده اند را پیدا کنید.
- در نتیجه حاصل مرحله قبل، به احتمال زیاد خطوط یکدیگر را پوشانده اند و ابتدا و انتهای خطوط به آسانی قابل تشخیص نیست. در یک تصویر مشابه، تعداد ۲۰ نقطه متناظر اتفاقی (random) از بین تمام نقاط انتخاب نموده و تنها آن ها را نمایش داده و به هم وصل نمایید. نتیجه را با نام res16.jpg ذخیره نمایید.
- با استفاده از تمام نقاط متناظر و با استفاده از RANSAC ماتریس هوموگرافی از تصویر دوم به اول را محاسبه کنید. از توابع و کتابخانه های موجود استفاده نمایید. در توضیحات خود ذکر نمایید که تعداد تکرار در RANSAC چند بار بوده است. ماتریس هوموگرافی به دست آمده را در گزارش خود بنویسید.
- نقاط متناظری که به عنوان نقاط inlier در نهایت به دست آمده اند را روی تصویر با رنگ قرمز نمایش دهید. تمام نقاط متناظر را با آبی نمایش داده و با خط آبی به هم وصل کنید و آنهایی که inlier هستند را با قرمز نشان دهید. نتیجه را با نام res17.jpg ذخیره نمایید.
- بررسی نمایید که آیا تمام نقاط inlier تناظر های درستی هستند؟ و آیا تمام آنها روی دیوار ساختمان قرار دارند؟ چنانچه مثالی از تناظر نادرست پیدا کردید در یک تصویر با نام res18_mismatch.jpg ذخیره نمایید.
- معکوس ماتریس هوموگرافی به دست آمده را روی چهار گوشه تصویر اول اعمال کنید و متناظر آنها را روی تصویر دوم به دست آورید. چهارضلعی که این چهار نقطه تشکیل می دهند را روی تصویر دوم بکشید. اگر هوموگرافی درست به دست آمده باشد، این چهارضلعی باید ناحیه دیده شده در تصویر اول را مشخص کنید. در یک فریم همانند بالا در سمت چپ تصویر اول را نمایش دهید و در سمت راست تصویر دوم که این چهارضلعی در آن رسم شده است را نمایش دهید. نتیجه را با نام res19.jpg ذخیره نمایید.
- ماتریس هوموگرافی را روی تصویر دوم اعمال نمایید و تصویر حاصل را با نام res20.jpg ذخیره نمایید. دقت کنید که ماتریس هوموگرافی را روی تمام تصویر دوم اعمال کنید.
- در یک فریم، در سمت چپ تصویر اول را نمایش دهید، در سمت راست تصویر دوم وارپ شده (تصویری که هوموگرافی روی آن اعمال شده است) به همراه چهارضلعی گفته شده در دو مرحله قبل را نمایش دهید. نتیجه را با نام res21.jpg ذخیره نمایید.

روش خود را به طور کامل در فایل pdf گزارش توضیح دهید. فایل کد اصلی خود را با نام q3.py ذخیره نمایید.

۴. محاسبه هوموگرافی و پیاده سازی RANSAC (۲۵ نمره)

سؤال قبل را یک بار دیگر انجام دهید، با این تفاوت که این بار برای محاسبه هوموگرافی و استفاده از RANSAC مجاز به استفاده از کتابخانه و کدهای موجود نیستید و خود شما باید آنها را پیاده سازی نمایید. برای محاسبه هوموگرافی از روش DLT که در کلاس تدریس شد استفاده کنید. تصاویر خواسته شده در سؤال قبل را این بار به همان ترتیب با نام های res22.jpg تا res30.jpg ذخیره نمایید. مقدار آستانه و تعداد تکرار در RANSAC را در گزارش خود ذکر کنید. روش خود را به طور کامل در فایل pdf گزارش توضیح دهید. فایل کد اصلی خود را با نام q4.py ذخیره نمایید.