

اصول بینایی کامپیوتر (۱- ۲۲۹۲۶) بهار ۱۴۰۱
تمرینات سری سوم
مهلت تحویل: ۷ خرداد ۱۴۰۱ (شنبه) ساعت ۱۲ شب

لطفاً به نکات زیر توجه بفرمایید.

۱. نتایج و پاسخ های خود را در یک فایل zip ارائه 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 بازبینی 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.

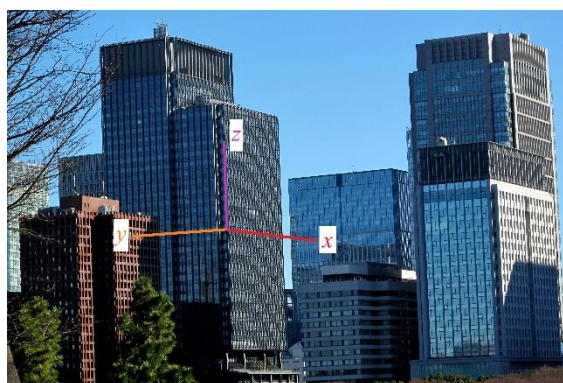
۸. در صورتیکه در انجام دادن تمرینات خود از هم فکری دیگران استفاده نموده اید باید نام آنها را در ابتدای گزارش خود ذکر نمایید.

۱. نقاط و خطوط محو شدن (Vanishing Points and Lines)

در این تمرین، در تصویر vns.jpg، نقاط محو شدن و خط محو شدن افق را به دست می آورید، با استفاده از آنها فاصله کانونی و نقطه اساسی (principal point) دوربین را به دست می آورید، زاویه صفحه تصویر و یا به طور متناظر سنسور دوربین با زمین و خط افق را به دست می آورید، و دوربین را طوری می چرخانید که صفحه تصویر عمود به زمین شده و ضلع افقی آن موازی خط افق شود.

بخش ۱ – به دست آوردن نقاط محو شدن و خط محو شدن افق (۱۰ نمره)

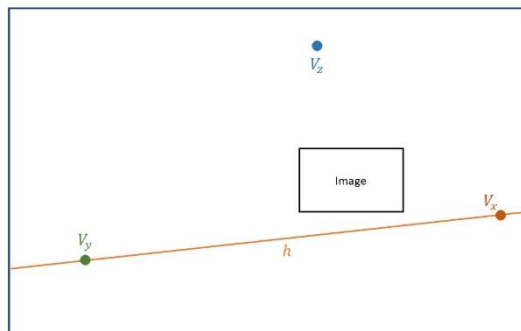
نقاط محو شدن متناظر سه راستای عمود بر هم در صحنه ای که از آن تصویر برداری شده است را به دست آورید. یک نقطه فرضی در فضای سه بعدی به عنوان مرکز مختصات در نظر بگیرید. نیازی به تعیین دقیق و مشخص کردن این نقطه نیست. همانند تصویر زیر، محور در امتداد افقی به سمت راست را محور x ، در امتداد افقی به سمت چپ را محور y ، و در امتداد عمودی به بالا را محور z در نظر بگیرید. نقاط محو شدن آنها را به ترتیب V_x ، V_y ، و V_z بنامید.



برای به دست آوردن هریک از نقاط محو شدن، باید تصویر تعدادی خط که موازی آن راستا در صحنه هستند را در تصویر به دست آورده و محل تلاقی این نقاط را بیابید. این خطوط را می توانید یا دستی مشخص کنید و یا با استفاده از کتابخانه های موجود به دست بیاورید. مختصات سه نقطه به دست آمده را در گزارش خود بنویسید. روش پیدا کردن محل تلاقی یک مجموعه خط در صفحه را در گزارش خود به طور کامل توضیح دهید.

با استفاده از دو نقطه محو شدن V_x و V_y خط افق که در واقع خط محو شدن زمین است را به دست بیاورید. خط افق را h نامیده و آن را با حاصلضرب خارجی مختصات همگن V_x و V_y به دست آورید ($h = V_x \times V_y$). این خط را به صورت $ax + by + c = 0$ که در آن $a^2 + b^2 = 1$ است در گزارش خود بنویسید.

خط افق را روی تصویر کشیده و نتیجه را با نام res01.jpg ذخیره نمایید. در یک تصویر سفید، سه نقطه محو شدن، خط افق، و محدوده تصویر را کشیده و با نام res02.jpg ذخیره نمایید، همانند تصویر زیر. روش خود را به طور کامل در فایل pdf گزارش توضیح دهید. فایل کد اصلی خود را با نام q1p1.py ذخیره نمایید.



بخش ۲ – به دست آوردن فاصله کانونی و نقطه اساسی و زاویه دوربین (۱۰ نمره)

با استفاده از سه نقطه محو شدن که در بخش قبل به دست آوردید، فاصله کانونی و مختصات نقطه اساسی را به دست آورده و در گزارش خود ذکر کنید. نحوه به دست آوردن این دو مقدار را در گزارش خود به طور کامل توضیح دهید. نقطه اساسی را روی تصویر مشخص نموده و مقدار فاصله کانونی را به صورت عنوان (title) بالای تصویر نوشته و حاصل را با نام res03.jpg ذخیره نمایید.

با داشتن مقدار فاصله کانونی و مختصات نقطه اساسی و فرض کردن مربعی بودن پیکسل های دوربین و صفر بودن مقدار اریبی (S)، می توانید ماتریس کالیبراسیون دوربین (K) را به دست آورید. حال با باز آفکَنش (back-projection) نقاط محو شدن می توانید راستای محورهای دستگاه مختصات در صحنه را به دست آورید. به دست آورید که دوربین حول محور Z خود چند درجه باید بچرخد تا تراز شود: محور x آن (راستای افقی یا ضلع افقی تصویر) موازی زمین و یا متناظراً موازی خط افق شود. همچنین، به دست آورید که، پس از تراز کردن دوربین، دوربین چند درجه باید حول محور x بچرخد تا صفحه تصویر یا متناظراً سنسور دوربین عمود بر زمین شود. این مقادیر را در گزارش خود ذکر کرده و نحوه به دست آوردن آنها را به طور کامل توضیح دهید.

روش خود را به طور کامل در فایل pdf گزارش توضیح دهید. فایل کد اصلی خود را با نام q1p2.py ذخیره نمایید.

بخش ۳ – صاف کردن دوربین (۵ نمره)

فرض کنید دوربین را بدون جه به جا کردن حول محور Z و سپس محور x آن به مقدار مناسب می چرخانیم تا محور x آن موازی زمین و سنسور عمود بر زمین شود. از آنجاییکه تفاوت بین این حالت با تصویر اصلی تنها در دوران است، می توان بین تصویر گرفته شده در این حالت و تصویر اصلی یک هوموگرافی به دست آورد. این هوموگرافی را محاسبه کرده و روی تصویر اصلی اعمال نمایید. تصویر حاصل را با نام res04.jpg ذخیره نمایید. هوموگرافی به دست آمده را در گزارش خود بنویسید. روش خود را در توضیحات به طور کامل توضیح دهید.

روش خود را به طور کامل در فایل pdf گزارش توضیح دهید. فایل کد اصلی خود را با نام q1p3.py ذخیره نمایید.

۲ – هندسه اپیپولار (Epipolar Geometry) (۱۵ نمره)

در این تمرین، بین تصاویر 01.jpg و 02.jpg ماتریس فاندامنتال به دست می آورید و ویژگی های هندسه اپیپولار بین آنها را بررسی می نمایید. تعدادی نقطه متناظر بین این دو تصویر با استفاده از روش دلخواه خود به دست بیاورید. ماتریس فاندامنتال (F) را با استفاده از RANSAC و با استفاده از کتابخانه های موجود محاسبه نمایید. ماتریس فاندامنتال به دست آمده را در گزارش خود بنویسید. دو تصویر را در کنار هم در یک قاب (فریم) قرار داده و روی آنها نقاط مطلوب (interest) که پرت نیستند (inlier) را با رنگ سبز و نقاط پرت (outlier) را با رنگ قرمز نمایش دهید. نتیجه را با نام res05.jpg ذخیره نمایید.

نقاط اپیپول در دو تصویر را به دست آورید. نقطه اپیپول در تصویر اول (e) را با در نظر گرفتن رابطه $Fe = 0$ به دست آورید. از آنجاییکه ماتریس F^t ماتریس فاندمنتال از تصویر دوم به اول است، بنابراین واضح است که رابطه $F^t e' = 0$ نیز برقرار است که در آن e' اپیپول تصویر دوم است. از این رابطه می توانید اپیپول تصویر دوم را به دست آورید. مختصات نقاط اپیپول به دست آمده را در گزارش خود بنویسید. اپیپول تصویر اول را روی آن نشان داده و نتیجه را با نام res06.jpg ذخیره نمایید. اپیپول تصویر دوم را روی آن نشان داده و نتیجه را با نام res07.jpg ذخیره نمایید.

برای ۱۰ جفت نقطه متناظر، خطوط اپیپولار آنها را به دست آورید. نحوه به دست آوردن خطوط اپیپولار را در گزارش خود توضیح دهید. دو تصویر را کنار هم در یک فریم نمایش داده و ۱۰ نقطه متناظر و خطوط اپیپولار آنها را روی تصاویر مربوطه رسم کنیم. نتیجه را با نام res08.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.

۳- سه بعدی سازی (3D Reconstruction) (۱۰ نمره)

یک شیء و یا ساختمان انتخاب نموده و حداقل ۳۰ تصویر از آن بیندازید. با استفاده از نرم افزارهای موجود، بازسازی سه بعدی این تصاویر که شامل به دست آوردن ابر نقاط (point cloud) و بافت روی مدل سه بعدی است را انجام دهید. می توانید از نرم افزارهایی مانند VisualSFM، OpenSFM، OpenMVG و Meshroom استفاده نمایید. تصاویری که گرفته اید و نتیجه خود را ذخیره نمایید.

۴- تشخیص صحنه با استفاده از سبد کلمات (Scene Recognition Using Bag of Words) (۵۰ نمره)

در این تمرین، مساله تشخیص صحنه با استفاده از سبد کلمات را انجام خواهید داد. داده مورد نظر در یک فایل زیپ با نام data.zip در اختیار شما قرار داده شده است. شما می توانید این داده را از این [آدرس](#) دانلود نمایید. این داده شامل ۱۵ دسته (class) مختلف است که به دو قسمت آموزش (train) و تست (test) تقسیم شده است. در داده آموزش تعداد متفاوتی از هر دسته قرار داد که در مجموع ۲،۹۸۵ تصویر را شامل می شود. در داده تست از هر دسته ۱۰۰ تصویر وجود دارد.

بخش ۱- نمایش ساده و نزدیک ترین همسایه (۱۰ نمره)

در این قسمت نمایش بسیار ساده ای از تصاویر در نظر گرفته می شود به این صورت که تمام سطرها و یا تمام ستون های هر تصویر به دنبال هم قرار داده شده و به صورت یک بردار در نظر گرفته می شوند. از آنجاییکه تصاویر اندازه های مختلفی دارند و همچنین برداری کردن آنها منجر به بردارهای بسیار طولانی می شود، بهتر است که تمام تصاویر را به یک اندازه مشخص کوچکی تبدیل کنید و سپس برداری کنید. اندازه 16×16 برای این تمرین توصیه می شود. اندازه های مختلفی را در نظر بگیرید و این بخش را تا انتها انجام داده و دقت را روی داده تست به دست آورده و مطلوب ترین اندازه از نظر اینکه هم دقت به اندازه کافی بالا باشد و هم میزان محاسبات به اندازه کافی کم باشد را انتخاب کنید. از این پس این اندازه را به عنوان اندازه مطلوب در نظر می گیریم. این اندازه را در گزارش خود ذکر کنید.

اندازه تمام تصاویر آموزش و همچنین تست را به اندازه مطلوب تبدیل کنید و آنها را به صورت برداری در نظر بگیرید. حال، با استفاده از روش نزدیک ترین همسایه (nearest neighbor) دسته تصاویر تست را به دست بیاورید. برای هر تصویر تست که به صورت بردار در نظر گرفته می شود، نزدیک ترین بردار از بین داده های آموزش را پیدا کنید. این کار را می توانید با فاصله L_1 یا L_2 انجام دهید. فاصله مناسب را می توانید با در نظر گرفتن میزان دقت آنها انتخاب کنید. فاصله ای که در نظر می گیرید را در گزارش خود ذکر کنید. به این صورت، برای هریک از داده های تست یک دسته اختصاص داده می شود. این دسته های به دست آمده را با دسته های درست داده های تست مقایسه کرده و دقت را به دست بیاورید. تعداد کل تصاویر تستی که دسته آنها درست تشخیص داده شده است تقسیم بر تعداد کل تصاویر تست (۱۵۰۰) برابر دقت می باشد. این مقدار را در ۱۰۰ ضرب کنید تا دقت به صورت درصد به دست آید.

به جای استفاده از روش نزدیک ترین همسایه می توانید از روش k نزدیک ترین همسایه (kNN) استفاده نمایید. مقادیر مختلف k را بررسی نموده و بهترین مقدار را در گزارش خود ذکر نمایید.

بهترین دقت به دست آمده با این روش را در گزارش خود ذکر کنید. اندازه تصاویر، نوع محاسبه فاصله، و مقدار k 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 q4p1.py ذخیره نمایید.

بخش ۲ - سبد کلمات و نزدیک ترین همسایه (۲۰ نمره)

ابتدا لغتنامه را باید به دست آورید. از هر کدام از تصاویر آموزش تعدادی بردار ویژگی با روش دلخواه خود به دست آورده و همه آنها را با هم در نظر بگیرید. با استفاده از روش k -means این بردارها را که نقاطی در فضای n بعدی هستند به k خوشه تقسیم کرده و مرکز این خوشه ها را به عنوان لغات بصری (visual words) لغتنامه در نظر بگیرید. می توانید تعداد خوشه ها را ۵۰ یا ۱۰۰ در نظر بگیرید. می توانید تعداد مختلفی برای لغات بصری در نظر بگیرید و در نهایت آن تعداد که بهترین دقت را حاصل می کند انتخاب کنید. به این صورت لغتنامه که حاوی k لغت بصری است تشکیل می شود.

حال، برای هریک از تصاویر آموزش و تست یک بردار هیستوگرام به دست بیاورید به این صورت که بردار ویژگی های هر تصویر را به دست بیاورید و برای هر بردار ویژگی نزدیک ترین کلمه بصری را پیدا کنید و تعداد آن را در هیستوگرام یک عدد افزایش دهید. به این صورت، یک هیستوگرام از لغات بصری برای هر تصویر به دست می آید. با استفاده از kNN دسته هر تصویر تست را همانند بخش قبل به دست آورده و دقت را محاسبه نمایید. بهترین مقدار k در kNN را به دست آورده و در گزارش خود ذکر کنید.

روش خود را به طور کامل در فایل pdf گزارش توضیح دهید. فایل کد اصلی خود را با نام q4p2.py ذخیره نمایید.

بخش ۳ - سبد کلمات و SVM (۲۰ نمره)

بخش قبل را تکرار کنید با این تفاوت که این بار به جای kNN از SVM استفاده کنید. دقت به دست آمده را گزارش کنید. ماتریس کانفیوژن را برای نتایج این مرحله تشکیل داده و در تصویر با نام res09.jpg ذخیره نمایید. روش خود را به طور کامل در فایل pdf گزارش توضیح دهید. فایل کد اصلی خود را با نام q4p3.py ذخیره نمایید.