

اصول بینایی کامپیوتر (۱- ۲۲۹۲۶) بهار ۱۴۰۱
تمرینات سری دوم
مهلت تحویل: ۱۷ اردیبهشت 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.

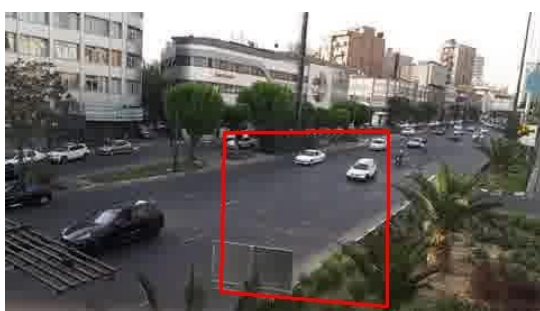
پانورامای ویدئو و پردازش ویدئو (۱۰۰ نمره)

در این تمرین به شما یک ویدئو در فایل video.mp4 از یکی از خیابان های شلوغ تهران داده شده است که در ادامه آن را ویدیوی خام می نامیم. هدف این است که ویدیویی به دست آوریم که پیش زمینه (foreground) ویدیوی خام، یعنی اجسامی که در این ویدئو حرکت می کنند (مثل ماشین ها، موتورها و انسان ها)، در آن مشخص شده باشد. در مسیر ساخت این ویدئو، یک تصویر پانوراما از کل محیطی که فیلم گرفته شده، تصویر پانوراما از پس زمینه (background)، و ویدیوی پس زمینه ی این ویدیوی خام نیز ساخته خواهد شد.

در این تمرین از ۹۰۰ فریم اول این ویدئو استفاده خواهیم کرد که متناظر با ۳۰ ثانیه اول آن می باشد. می توانید فریم های این ویدئو را با استفاده از ffmpeg یا هر ابزار دیگری استخراج کنید. ما از فریم شماره ۴۵۰ به عنوان فریم مرجع استفاده می کنیم. یعنی تمامی فریم های دیگر را به صفحه ی این فریم، با استفاده از یک نگاشت هوموگرافی، تصویر می کنیم. سعی شده در گرفتن ویدئو دوربین تنها چرخش داشته باشد و به همین دلیل فریم ها با نگاشت هوموگرافی به هم مربوط می شوند. عکس ها و ویدیوهایی که در این تمرین به شما برای نمونه و توضیحات بیشتر داده شده با رزولوشن کمتری از این ویدئو ایجاد شده اند. نتایج شما باید روی ویدیوی با رزولوشن کامل باشد.

بخش ۱ – ادغام (stitch) دو فریم اصلی (۵ نمره)

در این قسمت می خواهیم فریم ۲۷۰ را به صفحه ی فریم ۴۵۰ برده و آن ها را با هم ادغام کنیم. اگر این دو فریم را مشاهده کنید، متوجه می شوید که اشتراک زیادی با هم دارند. ابتدا با استفاده از SIFT یا هر روش دیگری، در این دو تصویر تعدادی نقطه ی مطلوب (interest point) استخراج کنید. سپس نقاط مطلوب متناظر با هم را در این دو عکس با استفاده از SIFT یا هر روش دیگری به دست آورید. در نهایت با استفاده از RANSAC یک نگاشت هوموگرافی H به دست آورید که با اعمال آن روی فریم ۲۷۰، به صفحه ی فریم ۴۵۰ برویم. برای محاسبه هوموگرافی با استفاده از RANSAC می توانید از کتابخانه های آماده استفاده کنید. درستی هوموگرافی به دست آمده را با کشیدن یک مستطیل در فریم ۴۵۰ و تصویر آن مربع تحت نگاشت H^{-1} در فریم ۲۷۰ آزمایش کنید. مثالی از این کار را در شکل زیر می توانید مشاهده کنید:



(ب) مستطیل در فریم ۲۷۰



(آ) مستطیل در فریم ۴۵۰

این تصاویر را با نام های res01-450-rect.jpg و res02-270-rect.jpg ذخیره کنید. برای ادغام این دو تصویر، فریم ۴۵۰ را روی وارپ (warp) شده فریم ۲۷۰ قرار دهید. فریم ۴۵۰ را در یک تصویر خالی به اندازه مناسب بزرگ که وارپ شده فریم ۲۷۰ هم در آن جای شود در جای مناسب قرار دهید. فریم ۲۷۰ را وارپ کنید و در جای مناسب در تصویر جدید زیر فریم ۴۵۰ قرار دهید، یعنی در قسمت های مشترک مقدار پیکسل ها از فریم ۴۵۰ باشند. تصویر به دست آمده یک پانوراما می باشد که از دو عکس تشکیل شده و فضای بزرگ تری نسبت به یک فریم از ویدئو را نشان می دهد. این تصویر بزرگ تر را با نام res03-270-450-panorama.jpg ذخیره کنید.

بخش ۲ - پانوراما با استفاده از ۵ فریم کلیدی (۱۵ نمره)

در این قسمت، با استفاده از فریم های ۹۰، ۲۷۰، ۴۵۰، ۶۳۰، و ۸۱۰ که آن ها را فریم های کلیدی می نامیم یک تصویر پانوراما درست کنید. هدف این است که هر فریم را به صفحه ی فریم ۴۵۰ تصویر کرده و آن ها را با هم ادغام کنیم. برای فریم های ۲۷۰ و ۶۳۰ می توانید از توضیحات بخش قبل استفاده کنید اما تصویر کردن فریم های ۹۰ و ۸۱۰ به طور مستقیم به فریم ۴۵۰ دشوار است چرا که این فریم ها با فریم ۴۵۰ اشتراک کمی دارند. به همین دلیل، شما باید نداشت هموگرافی مربوط به این فریم ها را در دو مرحله محاسبه کنید. یک هموگرافی از فریم ۹۰ به فریم ۲۷۰ و یکی از ۲۷۰ به ۴۵۰ (طبق توضیحات قسمت قبل) پیدا کنید و آن ها را در هم ضرب کنید. با این کار، یک نداشت هموگرافی از فریم ۹۰ به فریم ۴۵۰ خواهید داشت با این که این دو فریم اشتراک کمی نسبت به هم دارند. این کار را برای فریم ۸۱۰ نیز انجام دهید. در نهایت به این شکل برای هر یک از فریم های اصلی یک نداشت هموگرافی دارید که با اعمالش روی فریم مورد نظر، آن فریم به صفحه ی فریم ۴۵۰ خواهد رفت. این نداشت ها را روی فریم های مربوطه اعمال کرده و تصاویر حاصل را با هم ادغام کنید. فریم ها را طوری ادغام کنید که رنگ پانورامای حاصل یک دست باشد. میزان زیادی از نمره این بخش به ادغام کردن خوب این فریم ها بستگی دارد. نتیجه این بخش را با نام `res04-key-frames-panorama.jpg` ذخیره کنید.

بخش ۳ - تصویر ویدیوی خام به صفحه ی فریم مرجع (۱۰ نمره)

در این بخش، یک ویدیو می سازید که در آن برای هر فریم از ویدیوی خام یک نداشت هموگرافی به دست می آورید که آن فریم را به صفحه ی فریم مرجع (فریم ۴۵۰) می برد و آن را روی فریم مربوطه اعمال می کنید. کوچک ترین اندازه تصویری را به دست بیاورید که اگر با استفاده از همه ۹۰۰ فریم یک تصویر پانوراما بسازید که در آن همه فریم ها به صفحه فریم مرجع وارپ شده اند همه آنها در تصویر مورد نظر جای شوند. این تصویر را تصویر پانوراما می نامیم. برای هر فریم، یک تصویر خالی به اندازه تصویر پانوراما در نظر گرفته و وارپ شده آن فریم را در جای مناسب آن در تصویر خالی قرار دهید. برای این کار، نداشت هموگرافی بین آن و فریم مرجع را به دست بیاورید. برای فریم هایی که اشتراک کمی با فریم مرجع دارند باید از فریم های کلیدی کمک بگیرید. هموگرافی به دست آمده را روی آن فریم اعمال کنید و تصویر وارپ شده را در جای مناسب در تصویر پانورامای خالی قرار دهید. ویدیوی حاصل را با نام `res05-reference-plane.mp4` ذخیره کنید.

بخش ۴ - ساختن پانورامای پس زمینه (۱۵ نمره)

در این قسمت، هدف این است که اشیای متحرک را از ویدیو حذف کرده و یک پانوراما از پس زمینه ایجاد کنید. برای این کار، باید از پیکسل های تمام فریم های ویدیو استفاده کنید. در ویدیویی که در قسمت قبل تولید کردید، هر پیکسل در چندین فریم ظاهر می شود. شما باید با استفاده از رنگ هایی که این پیکسل در این فریم ها دارد، تخمین بزنید که رنگ پس زمینه ی این پیکسل چیست. ما از این مشاهده استفاده می کنیم که رنگ پس زمینه تقریباً مقداری ثابت است در حالی که رنگ پیش زمینه مقادیر متفاوتی دارد (چرا که پیش زمینه در حال حرکت و تغییر است و چیزهای مختلف رنگ یکسان ندارند). برای مثال، یک نقطه از خیابان که رنگ خاکستری دارد ممکن است در زمان های کوتاهی سفید، آبی، قرمز و یا هر رنگ دیگری بشود، ولی بیش از بقیه رنگ ها به رنگ خاکستری ظاهر میشود. برای هر پیکسل در ویدیوی بخش قبل، تمام رنگ هایی که این پیکسل در این ویدیو گرفته است را در نظر بگیرید و از بین آنها رنگ پس زمینه آن پیکسل را تخمین بزنید (طبیعتاً در زمان هایی که این پیکسل در ناحیه دید دوربین نبوده است نباید رنگ سیاه را برای این پیکسل به حساب بیاورید). روش خود را برای حدس زدن این رنگ توضیح دهید. این عملیات را برای تمام پیکسل ها انجام دهید. خروجی شما باید یک تصویر پانوراما باشد که تنها پیکسل های پس زمینه یا اشیای که حرکت نمی کنند در آن نمایش داده شده باشد. برای مثال، تصویر زیر بخشی از این تصویر پانورامای پس زمینه است. دقت کنید هرچه نتیجه شما در این قسمت بهتر باشد، باعث بهتر شدن نتیجه قسمت بعد شما خواهد شد. بنابراین خوب شدن نتیجه این بخش اهمیت بالایی دارد. تصویر پانورامای حاصل را با نام `res06-background-panorama.jpg` ذخیره کنید.



بخش ۵ – ساختن ویدیوی پس زمینه (۱۰ نمره)

در این قسمت، هدف حذف کردن پیش زمینه از ویدیو اصلی است، یعنی در ویدیو اصلی فقط پس زمینه دیده شود. برای این کار، پانورامای پس زمینه را که در بخش قبل تولید کردید، به مختصات فریم های دوربین تصویر کنید. در بخش قبل، برای هر فریم از ویدیو اصلی یک نگاشت هوموگرافی از آن فریم به فریم مرجع به دست آورده بودید. اگر معکوس این هوموگرافی را روی تصویر پانورامایی که در بخش قبل به دست آوردید اعمال کنید فریم بدون پیش زمینه به دست می آید. این کار را برای همه فریم ها انجام دهید. از فریم های حاصل یک ویدیو بسازید و با نام `res07-background-video.mp4` ذخیره کنید. ویدیوی حاصل مشابه ویدیوی خام (ویدیوی اولیه) می باشد با این تفاوت که تنها نقاط پس زمینه در آن هستند و شی متحرکی در آن وجود ندارد.

بخش ۶ – ساختن ویدیوی پیش زمینه (۱۵ نمره)

در این بخش، قسمت های پیش زمینه را مشخص می کنید. در ویدیوی پس زمینه، اشیاء متحرک حذف شده اند. در هر فریم اصلی، پیکسلی که رنگش به اندازه کافی با رنگ آن پیکسل در تصویر پس زمینه متفاوت است پیش زمینه حساب می شود. البته با این کار احتمالاً نویز هم خواهید داشت (تعدادی از پیکسل های پس زمینه به عنوان پیش زمینه تشخیص داده می شوند و برعکس). تلاش کنید با حذف این نویزها کیفیت تشخیص پیش زمینه را بهتر کنید. روش خود را به طور کامل توضیح دهید. پیکسل های پیش زمینه را محاسبه کنید و یک ویدیو همانند ویدیو اصلی تولید کنید که در آن پیکسل های پیش زمینه با رنگ قرمز قابل تشخیص باشند. می توانید مقدار کانال قرمز این پیکسل ها را ۱۰۰ واحد زیاد کنید. ویدیوی حاصل را با نام `res08-foreground-video.mp4` ذخیره کنید.

بخش ۷ – ساختن ویدیوی پهن تر (۱۰ نمره)

در بخش ۵ با استفاده از تصویر پانورامای پس زمینه و هوموگرافی هایی که برای هر فریم به دست آورده بودید یک ویدیوی پس زمینه تولید کردید. در این بخش هدف این است که یک ویدیوی پس زمینه پهن تر، مانند ویدیوی خروجی بخش ۵ ولی با زاویه دید افقی بیشتر، بسازید. ویدیوی خروجی شما باید حداقل ۵۰٪ پهن تر باشد (اندازه افقی آن بیشتر باشد). می توانید از همان ارتفاع ویدیوی اصلی استفاده کنید (ارتفاع را زیاد نکنید). روش خود را برای پهن تر کردن ویدیو به طور کامل توضیح دهید. ویدیوی خروجی این قسمت را با نام `res09-background-video-wider.mp4` ذخیره کنید.

بخش ۸ – حذف کردن لرزش دوربین (۲۰ نمره)

همانطور که با مشاهده ویدیو می توانید متوجه شوید، در هنگام ضبط ویدیو، مقداری لرزش در دوربین و در نتیجه در ویدیو وجود داشته است. در این قسمت هدف این است که ویدیویی به دست آوریم که در آن این لرزش حذف شده باشد. دقت کنید که در این مساله می توانید فرض کنید جابجایی نداریم و مرکز دوربین ثابت است و بین فریم های مختلف صرفاً دوران وجود دارد. مساله برای

حالتی که هم دوران و هم جابجایی داشته باشیم بسیار دشوارتر خواهد شد و همچنان یکی از موضوع های فعال برای تحقیقات میباشد. مجازید از هر روشی که به ذهنتان می رسد برای حل این سوال استفاده کنید ولی روش خود را باید کامل خودتان پیاده سازی کنید و استفاده از توابع آماده مخصوص این کار در این قسمت ممنوع می باشد. روش خود را کامل توضیح دهید. ویدیوی حاصل از حذف لرزش را با نام res10-video-shakeless.mp4 ذخیره کنید.