

به نام خدا

دانشگاه صنعتی شریف

دانشکده مهندسی برق



دکتر عمادالدین فاطمی زاده - پردازش تصاویر دیجیتال

نیم سال دوم ۱۴۰۱ - ۱۴۰۰

تمرین عملی سری اول

لطفاً به نکات زیر توجه بفرمایید: (رعایت نکردن این موارد باعث کاهش نمره می شود.)

۱. نتایج و پاسخ های خود را در یک فایل با فرمت zip به نام HW2-Name-StudentNumber در سایت Quera قرار دهید. همچنین فایل پایتون یا متلب خود را به همان نام در قسمت مخصوص به خود آپلود کنید.
۲. کسب نمره کامل در هر سوال مستلزم تحویل کدها (۴۰ نمره) و توضیحات (۳۰ نمره) و نتایج (۳۰ نمره) می باشد.
۳. کدهای شما تماماً باید توسط خودتان نوشته شده باشند. هرگونه استفاده از کد دیگران، اعم از دوستان و اینترنت، به هر شکل ممکن، تقلب محسوب می شود و نمره تمام تمرینات جاری و تمام تمرینات قبلی صفر خواهد شد. با اجرای این کدها باید همان نتایجی که فرستاده اید قابل بازیابی باشند. برنامه شما باید به گونه ای باشد که بدون نیاز به هیچ تغییری قابل اجرا باشد، در غیر این صورت هیچ نمره ای تعلق نخواهد گرفت.
۴. برای تمام سؤالات، باید جزئیات روشی که استفاده کرده اید را توضیح دهید و نتایجی که گرفته اید را ارائه دهید. این توضیحات می تواند در یک فایل pdf و یا در یک فایل ipynb باشد. در توضیحات، باید اشاره کامل به کارهایی که انجام داده اید بنمایید به طوری که یک شخص آگاه از موارد درس بتواند به آسانی متوجه کاری که شما انجام داده اید شود.
۵. در طول ترم امکان ارسال با تاخیر پاسخ همه ی تمرین تا سقف شش روز و در مجموع بیست و یک روز وجود دارد. پس از گذشت این مدت، پاسخ های ارسال شده پذیرفته نخواهند بود. همچنین، به ازای هر روز تأخیر غیر مجاز بیست درصد از نمره تمرین به صورت ساعتی کسر خواهد شد.
۶. اگر از upyter notebook استفاده می کنید، میتوانید خروجی ها را پاک کنید تا حجم فایل تحویلی زیاد نشود.
۷. مهلت تحویل: ۱۹ فروردین ساعت ۲۳:۵۹
۸. نام طراح هر سوال در زیر آن نوشته شده است و شما میتوانید سؤالات خود را از طریق ایمیل یا تلگرام از طراح سوال بپرسید.

محمدامین علم الهدی: @Alam_Amin – Amin@ee.sharif.edu

سعید رضوی: @RazooIs – Saeedrazavi890@gmail.com

یاسمین مدقالچی: @Yasssssimed – Yasmed1379@yahoo.com

صدرا صبوری: @Sadrasabouri – Sadra@ee.sharif.edu

امیررضا حاتمی پور: @Arhp78 – Arhp78@gmail.com

سوال اول

طراح: سعید رضوی

در این سوال به دو روش sharpen کردن تصاویر در حوزه فرکانس می پردازیم. تصویر blur شده اولیه با نام q1.jpg را بخوانید.

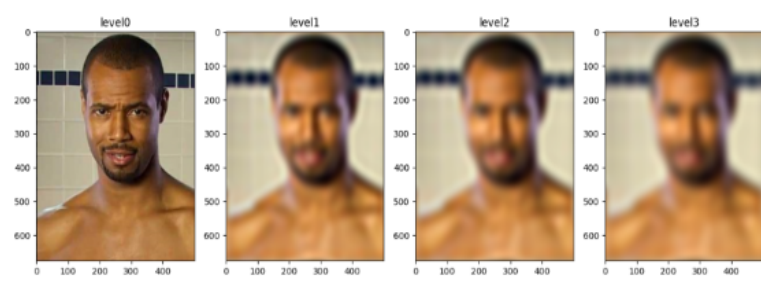
۱. روش اول: ابتدا نشان دهید که می توان تصویر را با استفاده از رابطه $F^{-1}(1 + kH_{HP})F$ شارپ تر کرد که F تبدیل فوریه تصویر مورد نظر (بر روی هر کانال دلخواه از تصویر)، H_{HP} یک فیلتر بالاگذر، و F^{-1} نماد تبدیل فوریه معکوس می باشد. (ترجیحا با نوشتن روابط ریاضی رابطه بالا را توجیه کنید، اگر با استفاده از روابط ریاضی موفق نبودید سعی کنید به صورت شهودی آن را توجیه کنید). پس از توجیه این رابطه به پیاده سازی آن می پردازیم. ابتدا تبدیل فوریه تصویر را به دست آورده و اندازه ی تبدیل فوریه را نمایش و با نام q1_res01.jpg ذخیره کنید. سپس فیلتر بالاگذر ساخته شده را با نام q1_res02.jpg ذخیره کنید. سپس خروجی $(1 + kH_{HP})F$ را با نام q1_res03.jpg ذخیره کنید و آن را با تبدیل فوریه اولیه اولیه عکس مقایسه کنید. در نهایت تصویر نهایی را با نام q1_res04.jpg ذخیره کنید و ذکر کنید به ازای چه مقدار k به نتیجه دلخواه خود رسیدید.

۲. روش دوم: ابتدا نشان دهید که می توان تصویر را با استفاده از رابطه $f + kF^{-1}\{4\pi^2(u^2 + v^2)F(u, v)\}$ که f عکس ورودی در حوزه ی زمان، $F(u, v)$ تبدیل فوریه تصویر مورد نظر (بر روی هر کانال دلخواه از تصویر)، (u, v) متغیرهای فوریه در حوزه فرکانس می باشد (هم ارز (x, y) در حوزه مکان) و F^{-1} نماد تبدیل فوریه معکوس می باشد، شارپ تر کرد (ترجیحا با نوشتن روابط ریاضی رابطه بالا را توجیه کنید، اگر با استفاده از روابط ریاضی موفق نبودید سعی کنید به صورت شهودی آن را توجیه کنید). حال به پیاده سازی رابطه بالا می پردازیم: اندازه $4\pi^2(u^2 + v^2)F(u, v)$ را نمایش دهید و با نام q1_res05.jpg ذخیره کنید. سپس $F^{-1}\{4\pi^2(u^2 + v^2)F(u, v)\}$ را که همان unsharp mask ما می باشد را با نام q1_res06.jpg ذخیره کنید. در نهایت تصویر نهایی را با نام q1_res07.jp ذخیره کنید و ذکر کنید به ازای چه مقدار k به نتیجه دلخواه خود رسیدید.

سوال دوم

طراح: سعید رضوی

در این سوال قصد داریم با استفاده از روش Laplacian stack (پشته لاپلاسین) به ترکیب دو عکس به صورتی بپردازیم که نتیجه حاصل بسیار طبیعی باشد. در ادامه این روش را توضیح می دهیم. برای به دست آوردن پشته ی لاپلاسین نیازمند به آشنایی با Gaussian stack (پشته ی گاوسی) هستیم. تعریف پشته ی گاوسی با n سطح (level) برابر است با مجموعه ی n تصویر blur شده از یک تصویر مشخص که در هر مرحله بالاتر، تصویر با شدت بیشتری blur شده است. (اگر بخواهیم تفسیری از پشته های گاوسی در حوزه زمان داشته باشیم، مثل این است که در هر مرحله تصویر را با یک فیلتر گاوسی با انحراف معیار بیشتری کانوالو کرده و در نتیجه هر مرحله که بالاتر میرویم، پشته گاوسی ما تصویری محوتر از تصویر اصلی است) نمونه ای از پشته های گاوسی را برای یک تصویر دلخواه در شکل زیر مشاهده می کنید:



دقت شود در این سوال blur کردن تصاویر و ساخت پشته‌های گاوسی باید در حوزه فرکانس انجام شود و مجاز به کانالو کردن تصاویر با فیلتر گاوسی در حوزه زمان نیستید. اما در ادامه می‌توانید پس از blur کردن در حوزه فرکانس آن را به حوزه مکان برگردانید و ادامه‌ی کار را در حوزه مکان انجام دهید. حال به تعریف Laplacian stack (پشته لاپلاسین) می‌پردازیم: لاپلاسین پشته در مرحله i ام از رابطه زیر به دست می‌آید:

$$l_{s_i} = G_{s_i} - G_{s_{i+1}}$$

که G_{s_i} پشته‌ی گاوسی تصویر در مرحله i ام می‌باشد. همانطور که می‌توانید حدس بزنید پشته لاپلاسین دارای جزئیات تصویر در هر مرحله است. نکته مهمی که باید به آن توجه داشت این است که پشته‌ی لاپلاسین در مرحله آخر همان پشته‌ی گاوسی در مرحله آخر است یعنی $l_{s_{last\ level}} = G_{s_{last\ level}}$ (توجه کنید دلیل برابری پشته‌ی لاپلاسین و پشته‌ی گاوسی در مرحله آخر چیست و چه کمکی به ترکیب تصاویر می‌کند؟). نکته مهم دیگر در ترکیب دو تصویر این است که برای این که مشخص کنیم چه بخشی از تصویر نهایی از تصویر اول و چه بخشی از تصویر نهایی از تصویر دوم تشکیل شده باید یک ماسک باینری (binary mask) مشخص کنیم (این ماسک باینری که یک ماسک سیاه و سفید است مشخص می‌کند که برای مثال ناحیه سفید رنگ تصویر باید از تصویر اول بیاید و ناحیه سیاه رنگ تصویر از تصویر دوم). در انتها پس از به دست آوردن پشته‌های لاپلاسین باید آن‌ها را در هر مرحله به صورت مناسب با یکدیگر جمع کرده. جمع پشته لاپلاسین دو عکس در هر مرحله از رابطه زیر بدست می‌آید:

$$l_{s_i\ blended\ image} = l_{s_i\ image1} G_{s_i\ mask} + l_{s_i\ image2} (1 - G_{s_i\ mask})$$

در رابطه بالا $G_{s_i\ mask}$ پشته‌ی گاوسی مرحله i ام ماسک می‌باشد. پس از تشکیل شدن پشته‌های لاپلاسین تصویر نهایی با جمع کردن تمام آنها با یکدیگر باید به جواب نهایی و تصویر ترکیب شده خود برسید. نمونه‌ای از ترکیب دو عکس با استفاده از این روش در ادامه آمده است. برای این سوال شما هم دو تصویر به دلخواه خودتان انتخاب کنید و با توضیحات بالا آن دو را به صورت دلخواه با هم ترکیب کنید.

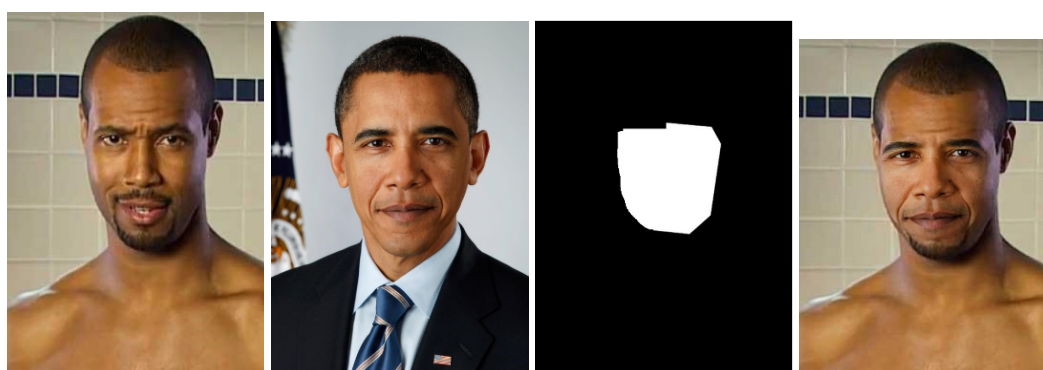


Image I

Image II

Mask

Blended Image

برای ساختن ماسک می‌توانید از هر نرم افزار دلخواهی مثل Paint، Photoshop استفاده کنید.

سوال سوم

طراح: یاسمین مدقالچی

در این سوال می‌خواهیم با عملکرد و نحوه‌ی پیاده‌سازی مدل‌های مختلف الگوریتم retinex آشنا شویم. هم چنین می‌خواهیم این پیاده‌سازی‌ها بر اساس مقاله Multi Scale Retinex که در پوشه‌ی درس قرار گرفته شده است انجام شود.

۱. عکس q3.jpeg را load کرده و نمایش دهید.

۲. ابتدا قسمت 2.1 مقاله (Single scale Retinex) را مطالعه کرده و به صورت مختصر در گزارش کار خود آن را توضیح دهید. الگوریتم SSR را به عنوان تابعی با همین نام پیاده کنید و عملکرد این تابع را بر روی عکس load شده بسنجید و خروجی تولید شده را در گزارش کار خود قرار دهید.

۳. قسمت 2.2 مقاله (Multi scale Retinex) را مطالعه کرده و تفاوت آن با قسمت قبل را توضیح دهید و هم چنین الگوریتم را در قالب تابعی به نام MSR پیاده کنید و عملکرد این تابع را بر روی عکس load شده بسنجید و خروجی آن را در گزارش کار خود بیاورید.

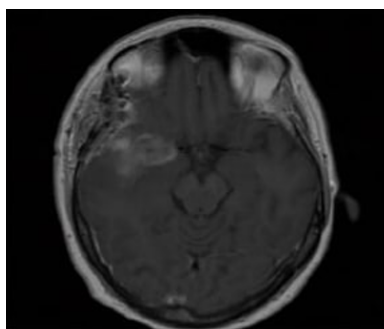
۴. الگوریتم MSRCR را در قسمت 2.3 مقاله مطالعه کرده و در مورد فرق‌های آن با SSR توضیح دهید و مانند قسمت‌های قبل آن را در قالب یک تابع پیاده کنید و نتایج را در گزارش کار خود بیاورید.

۵. الگوریتم MSRCP را نیز پیاده کنید و نتایج را در گزارش کار خود بیاورید.

سوال چهارم

طراح: امیررضا حاتمی‌پور

۱. هر کدام از روش‌های 1 HE، 2 AHE، 3 CLAHE را توضیح دهید. سپس فرض کنید تصویر ورودی به شکل زیر باشد.



حال مشخص کنید که خروجی‌های زیر هر کدام برای کدام یک از این سه روش است.

-
- 1 Histogram equalization
 2 Adaptive histogram equalization
 3 Contrast limited adaptive histogram equalization

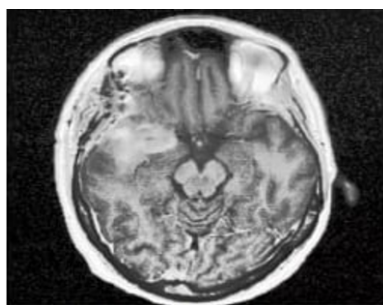


Image I

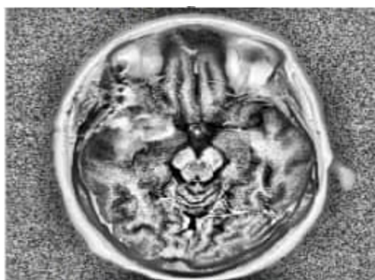


Image II

۲. برای تصویر q4.png روش های HE، AHE و CLAHE را پیاده کنید و خروجی ها را در فایل q4_1.png، q4_2.png و q4_3.png ذخیره کنید.

سوال پنجم طراح: صدرا صبوری

صفحه‌ی منطقه‌ای تصویری از مجموعه دواير هم‌مرکز است که یک در میان تغییر رنگ می‌دهند. در واقع در این تصویر فرکانس‌های مختلف در شعاع‌های متفاوت دیده می‌شوند. نمونه‌ای از این فایل به نام q5.png در اختیار شما قرار گرفته است. با توجه به سوالات زیر پاسخ دهید.

۱. به ترتیب فیلترهای سوبل و گوسی و لاپلاسین (با مقادیر دلخواه) اعمال کنید و نتایج را در q5_res01.png و q5_res02.png و q5_res03.png ذخیره کنید و تفاوت‌ها و شباهت‌های آن‌ها را بیان کنید.

۲. ابتدا نشان دهید که تبدیل فوریه یک فیلتر گوسی همچنان گوسی است. سپس تصویر اولیه را به حوزه فرکانس ببرید و تصویر مقارن شده آن را در مقیاس لگاریتمی در حوزه فرکانس نمایش دهید و با نام q5_res04.png ذخیره کنید. حال با ضرب کردن فیلتر گوسی (معادل فرکانسی فیلتر گوسی قسمت قبل) در تبدیل فوریه تصویر q5_res05.png را ذخیره کنید و سپس با ضرب متهم این فیلتر تصویر q5_res06.png را ذخیره کنید. هر دوی این تصاویر را از حوزه فرکانس برگردانید و نام تصاویر ایجاد شده را به ترتیب q5_res07.png و q5_res08.png بنامید. همین مراحل را برای یک فیلتر مربعی در حوزه فرکانس انجام دهید و تصاویر را به ترتیب q5_res09.png و q5_res10.png و q5_res11.png و q5_res12.png بنامیم. خروجی ها را با هم مقایسه کنید. علت به وجود آمدن مربع‌های کوچک در تصاویری که از فیلتر مربعی گذشتند چیست؟

۳. یک بار تصویر اولیه را با بزرگ‌نمایی ذخیره کنید (نام آن را q5_res13.png) و یکبار با کوچک‌نمایی (نام آن را q5_res14.png) قرار دهید و شباهت هر کدام از آن‌ها را با نتایج قسمت‌های قبل مقایسه کنید و دلیل این شباهت را بیان کنید.

سوال ششم طراح: محمدامین علم‌الهدی

در این سوال قصد داریم با استفاده از روش template matching، حرف B را از بین حروف دیگر الفبای انگلیسی تشخیص دهیم. برای ادامه‌ی تمرین، از دو تصویر زیر که به ترتیب در فایل های q7_1.png و q7_2.png آمده‌اند استفاده کنید.

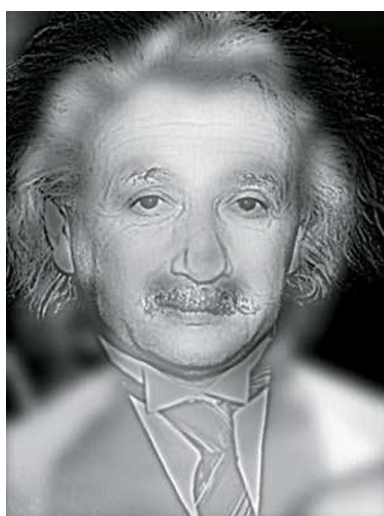


برای آشنایی بیشتر با انواع روش‌های مختلف temple matching می‌توانید از این لینک استفاده کنید. ۲ روش از روش‌های ذکر شده در لینک قبل را بر روی تصویر اعمال کنید و نتیجه‌ی آن به همراه توضیح روش انجام تمامی مراحل را در گزارش کار خود ذکر کنید. در صورتی که روش انتخابی شما به نتیجه‌ی مطلوب نرسید، سعی کنید دلیل آن را توضیح دهید.

سوال هفتم – امتیازی (۱۵ نمره از ۱۰۰ نمره)

طراح: محمدامین علم‌الهدی

در این سوال قصد داریم به کمک فیلترهای پایین‌گذر و بالا‌گذر، یک تصویر هیبریدی بسازیم. تصویر هیبریدی به تصویری گفته می‌شود که وقتی از نزدیک به آن نگاه می‌کنیم یک تصویر، و وقتی از دور به آن نگاه می‌کنیم تصویر دیگری می‌بینیم. به طور مثال، تصویر زیر یک تصویر هیبریدی است. هنگامی که از نزدیک به آن نگاه می‌کنید تصویر انیشتین را می‌بینید و وقتی از فاصله دورتر به آن نگاه کنید، تصویر دیگری را خواهید دید. در ادامه نمونه یکی از این تصاویر را می‌بینید. شما می‌توانید



به دلخواه دو تصویر انتخاب کنید. مطالعه این مقاله برای رسیدن به نتایج بهتر در این قسمت توصیه می‌شود.



برای به دست آوردن نتیجه بهتر، باید دو تصویر را هم اندازه کرده و با هم منطبق نمایید، یعنی قسمت‌های مشابه یا معادل را روی هم قرار دهید. برای مثال، اگر بخواهید دو تصویر از صورت دو شخص را باهم ادغام کنید، بهتر است در ابتدا دو تصویر را طوری تغییر دهید تا اجزای متناظر صورت دو شخص در یک مکان از تصویرشان قرارگیرد. می‌توانید این کار را با تطبیق چشم‌ها انجام دهید. چون شکل کلی صورت انسان‌ها یکسان می‌باشد، اگر چشم‌ها منطبق شده باشند می‌توان انتظار داشت که با تقریب نسبتاً خوبی بقیه قسمت‌ها نیز با هم منطبق شده‌اند. این کار را می‌توانید با انتخاب چند نقطه متناظر بین دو تصویر و استفاده از نگاشت‌های هندسی فضای دو بعدی (چرخش و انتقال) انجام دهید. اینکار را باید خودتان انجام دهید و نمی‌توانید از کدهای موجود در اینترنت استفاده نمایید. همچنین از انجام برخی پردازش‌ها بر روی تصاویر مثل نرمالایز کردن و تبدیل RGB به Grayscale غافل نشوید. توجه کنید که سائز دو تصویر نیز باید یکسان باشد، پس بعد از اینکه دو تصویر بر روی یکدیگر register شدند، سائز هر دو را یکی کنید.

هنگامی که از فاصله‌ی نزدیک به یک تصویر نگاه کنیم بیشتر جزئیات آن تصویر را می‌بینیم. علت این است که بیشتر جزئیات کوچک تصاویر در فرکانس‌های بالا ذخیره شده‌اند و جزئیات کلی تصویر در فرکانس‌های پایین ذخیره شده‌اند. بنابراین در تصویری که می‌خواهید از نزدیک دیده شود باید فرکانس‌های پایین را حذف کرده و فرکانس‌های بالا را نگه دارید (و بالعکس برای تصویری که می‌خواهید وقتی از دور به آن نگاه می‌کنید دیده شود). برای انجام این کار، ابتدا نیاز داریم که دو تصویر را به حوزه‌ی فرکانس ببریم. اینکار را می‌توانید توسط توابع آماده یا با استفاده از فرمول تبدیل فوریه 2 بعدی انجام دهید. مرحله‌ی بعد باید با استفاده از فیلترهای پایین‌گذر و بالا‌گذر، تصاویر را فیلتر کنیم. برای فیلتر، از فیلتر گوسی دو بعدی استفاده کنید (پارامترهای فیلتر مانند فرکانس cutoff و انحراف از معیار را خودتان انتخاب کنید و حتما در گزارش کار ذکر کنید). می‌توانید مقدار cutoff را برای هر دو فیلتر بالا‌گذر و پایین‌گذر یکسان در نظر بگیرید ولی برای کسب بهترین نتیجه، باید این دو مقدار را متفاوت در نظر بگیرید به طوری که cutoff برای فیلتر بالا‌گذر کوچک‌تر از مقدار برای فیلتر پایین‌گذر باشد. در اینصورت هر دو تصویر در دامنه فرکانس در یک نوار مشترک غیر صفر می‌شوند. برای ترکیب دو تصویر در این نوار، از میانگین‌گیری وزن‌دار استفاده نمایید (به محتوای فرکانسی هر تصویر یک ضریب اختصاص دهید و با هم در حوزه فرکانس جمعشان کنید). تصویر خروجی در مثال گفته شده به صورت زیر می‌باشد:



لطفا تمامی مراحل کار خود را به صورت کامل در گزارش کار توضیح دهید و سعی کنید عکس‌هایی که در گزارش کار قرار می‌دهید، caption مناسب داشته باشند. همچنین فایل تصاویر خواسته شده در ادامه را داخل فایل تحویلی خود با نام گذاری مناسب قرار دهید: ۱- عکس هر دو تصویر بعد از اعمال نگاشت‌ها و یکسان شدن سایز ۲- هر دو تصویر در حوزه فرکانس قبل از اعمال فیلتر ۳- تصویرهای فیلتر شده در حوزه فرکانس ۴- تصویر هیبرید نهایی