Homework 2 Questions

Instructions

- One exercise graded for completion only.
- Five graded questions.
- Write code where appropriate.
- Feel free to include images or equations.
- Please make this document anonymous.
- This assignment is **fixed length**, and the pages have been assigned for you in Gradescope. As a result, **please do NOT add any new pages**. We will provide ample room for you to answer the questions. If you *really* wish for more space, please add a page *at the end of the document*.
- **We do NOT expect you to fill up each page with your answer.** Some answers will only be a few sentences long, and that is okay.

Exercise

E1: Let's look again at the webcam Fourier decomposition demo which Srinath showed in class. Let's run it in our CSCI 1430 virtual environment *preferably on a computer with a webcam*, from within the 'questions' directory.

```
$ python liveFFT2.py
```

This file contains five parts for you to explore and see the amplitude image, the phase image, and the effect of the reconstructed image.

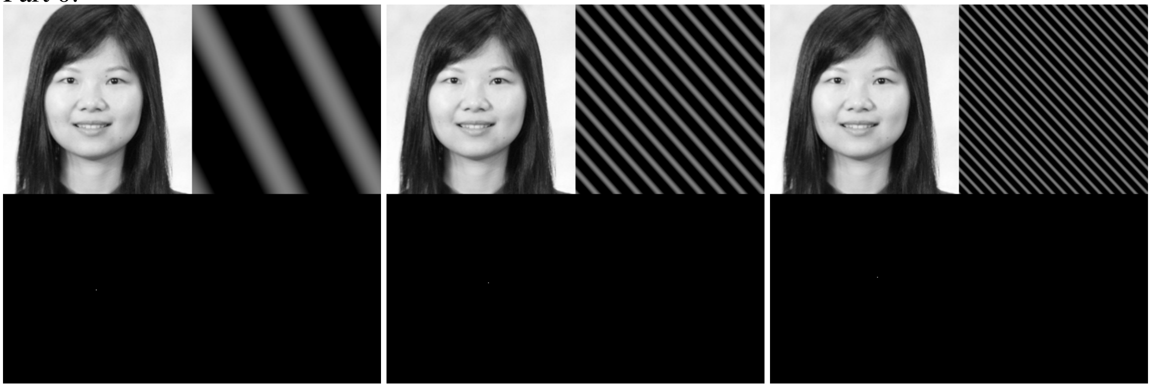
- Part 0: Scanning the basis and observing the output image.
- Part 1: Reconstructions from different numbers of basis frequencies.
- Part 2: Replacing amplitude and phase with that from a different image.
- Part 3: Replacing amplitude and phase with that from a noise image.
- Part 4: Manipulating the amplitude and phase images.

Uncomment the different parts and explore the camera feed decomposition! Please include the results of your experimentation, e.g., two-to-three screenshots of what you discover. We'll be grading for completion, not correctness. *Note:* For anonymous grading, try not to put yourself in the camera frame. Show your favourite vector calculus book, wear a mask, use your cat, etc. Extra credit for creative effort.

Exercise Results

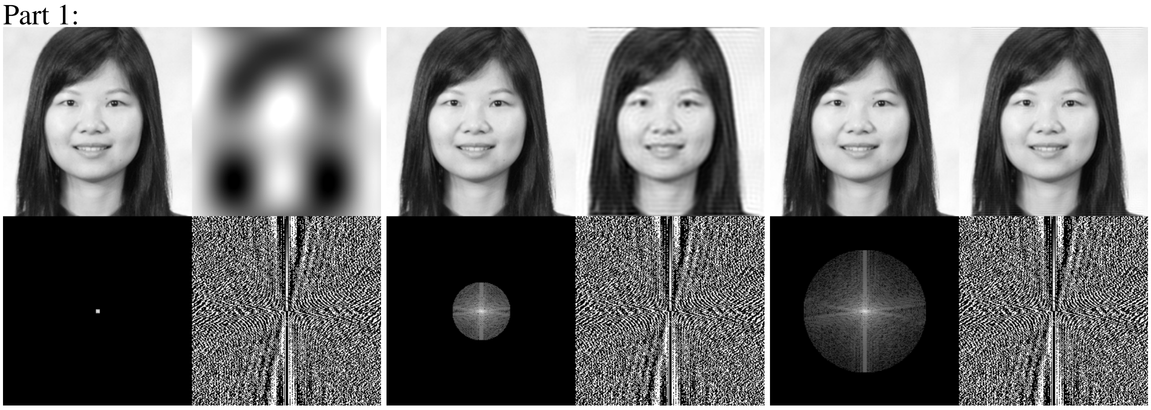
Please include images of your results from the exercise here, e.g., two-to-three screenshots of your findings.

Part 0:



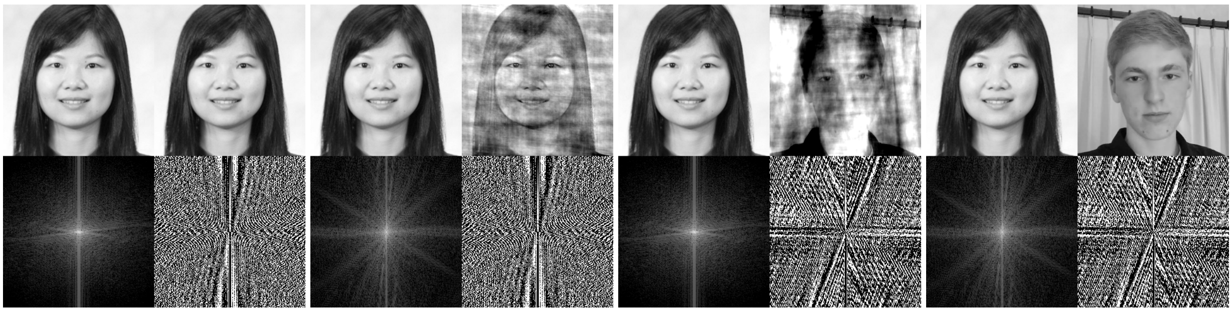
As the amplitude shift through the values, the white dot in the amplitude image moves along a circle and the circle becomes larger and larger. which means the amplitude increases. The reconstructed image has denser black and white lines as the amplitude increases. The lines orient with the white dot in the amplitude image like a circle.

Part 1:



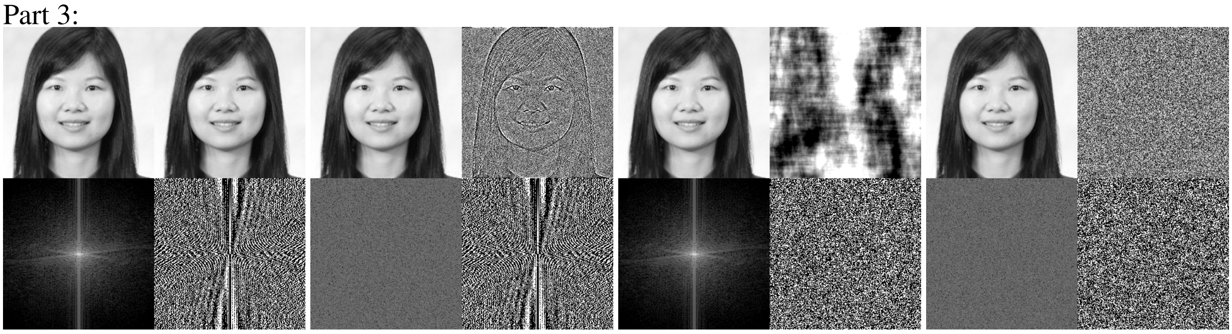
Here we have a circular mask over the amplitude image to modify amplitudes of certain frequencies to zero and leave the others nonzero. In this case, we can see that as number of nonzero amplitude frequencies we left in the amplitude image increases, the image becomes clearer and vice versa. Because we need various frequencies to construct a clear image, as the number of frequencies in the basis decrease, the image become more blur.

Part 2:



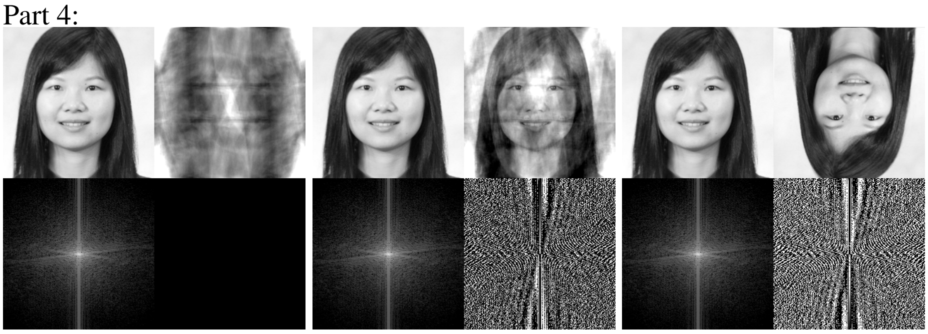
The first image is the original. The second one has amplitude from Jack image. We can see that it blurs the image but still keep most features from Yuanning's image. The third one uses Jack's phase and Yuanning's amplitude, which causes the reconstructed image to have Jack's face shown. The last one uses both Jack's amplitude and phase, which reconstructed the image totally as the Jack's original image.

Part 3:

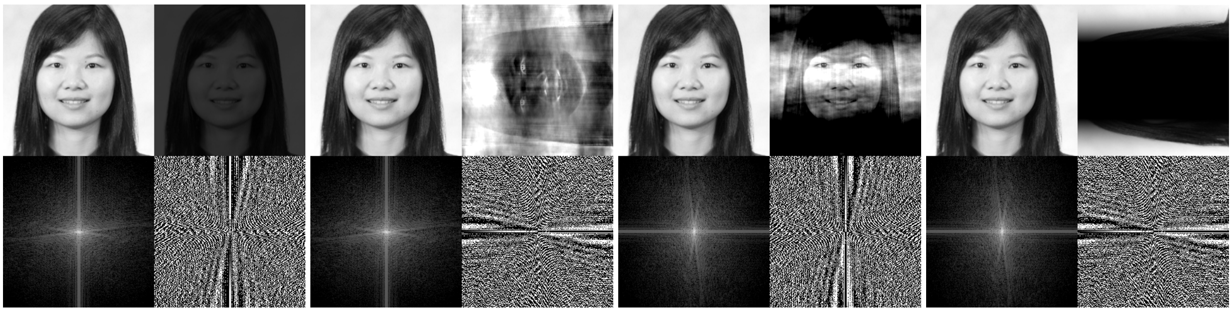


The first image is the original image. The second image is reconstructed with noise amplitude. We can see that the reconstructed image has only edges from Yuanning's face. The third image is noise phase. We can see that the noise is in the reconstructed image with Yuanning's image amplitude and we can easily recognize some hair and face areas from the reconstructed image. The last one use amplitude and phase from noise, which reconstructed an image of noise.

Part 4:



The first image is the zero phase, and we can see the reconstructed image have no outlines of Yuanning's face. The second one is the half face and we can see a rough outline of Yuaning's phase. The third image has full phase but with negative absolut values, so we can see an inverted Yuanning's face.



The first image is the rotation of phase's values only, so the brightness is changed from bright to dark and backwards in the reconstructed image. The second image is the rotation of direction of the phase, so the face's direction is rotated. The third image is a rotation of amplitude, and we can see blackness is in bottom and top while whiteness stays in the middle. The last image is a rotation of both phase and amplitude, And we can see that it is an overlapping effect of image 2 and image 3.

Extra space

Questions

Q1: Imagine we wished to find points in one image which matched to the same world point in another image—so-called feature point correspondence matching. We are tasked with designing an image feature point algorithm which could match world points in the following three pairs of images.

RISHLibrary: One Two — *Chase:* One Two — *LaddObservatory:* One Two

Please use the included python script `plot_corners.py` to find corners using Harris corner detection for each image above.

For each pair, discuss the differences in the returned corners (if any), and what aspects of the images may have caused these differences. Then discuss what real world challenges exist in matching features between images.

A1 (a): Answer for RISHLibrary pair

RISHLibrary-1 find corers on the shoes and RISHLibrary-2 find corners on the leg of chair. Since the two images contain different content from the RISHLibrary, they have different points of corners.

A1 (b): Answer for Chase pair

Chase-1 contain points of corners founded on the sign of the statue while Chase-2 contain no points of corners founded by the algorithm. The difference between the results comes from the clarity of the images. Chase-1 has a higher resolution than the Chase-2 and there are lots of overlapping effects in the Chase-2 image which may also cause difficulties in recognizing corners.

A1 (c): Answer for LaddObservatory pair

Corners for LaddObservatory-1 are mainly in the right part of the roof and bottom part of the roof (border between roof and body). Corners for LaddObservatory-2 contain not only those in the LaddObservatory-1, but also many points in the crowd of people. Also, 2 extra corners in the roof top are found in LaddObservatory-2. Since the content of the 2 images are different, LaddObservatory-2 is taken from a further point than LaddObservatory-1, it contains more pixels from crowd of people. However, both find similar corners in the roof because both have roof in their scene. But because of distance, Ladd-2 contains more lower frequencyies than Ladd-1; hence they have different points of corners.

Q2: The Harris Corner Detector is commonly used in computer vision algorithms to find interest points from which to extract stable features for image matching.

How do the eigenvalues of the ‘M’ second moment matrix vary with local image brightness, and how might we interpret the eigenvalues geometrically?

A2: Your answer here.

If the local image brightness increases in all directions, that means both eigenvalues are large. If the local image brightness is a flat region, which does not change very much, that means both eigenvalues are small. If the local image is an edge, that means one eigenvalue is much larger than the other.
We can view the local image as an ellipse; the small eigenvalue indicates the direction of the fastest change in brightness, while the big eigenvalue indicates the direction of the slowest change of the brightness.

Q3: Goal: Understand who can collect and control surveillance data and reflect on the nature of the relationship between the subjects of that data and other stakeholders.

Please read this 2008 Brown Daily Herald article describing the gradual increase in surveillance cameras on Brown's campus. Then, read this 2020 Brown Daily Herald article, which mentions that there are now 800 surveillance cameras on campus, about as many cameras as full-time faculty. Note, that since 2015 Axis Communication, Brown's primary supplier of surveillance cameras has advertised its cameras' abilities to perform onboard facial recognition.

(a) Were you aware of Brown's high number of surveillance cameras before reading these articles? Do you have any concerns? Provide an argument both for and against this surveillance system. (4–5 sentences)

A3 (a): Your answer here.

I am not aware of Brown's high number of surveillance cameras before reading these articles. I have concerns about the camera's break of people's privacy. In terms of the goodness of the camera, it is a good way to digitally record and monitor the activities on campus, which protect students safety and public/private assets on campus. It also performs a good role in predict alarms and provide providence for events with affordable costs and waste no human resources. In terms of the badness of the surveillance camera, it might intrude people's privacy because the policies and locations of these cameras are kept as secrets.

Drones are often used by individuals, companies, and governments for photography, with some organizations demonstrating delivery applications, too. In the U.S., the federal government has started using drones for domestic surveillance. From the Electronic Frontier Foundation:

"Drones already in use by law enforcement can carry various types of equipment including live-feed video cameras, infrared cameras, heat sensors, and radar. Some military versions can stay in the air for hours or days at a time, and their high-tech cameras can scan entire cities, or alternatively, zoom in and read a milk carton from 60,000 feet. They can also carry WiFi crackers and fake cell phone towers that can determine your location or intercept your texts and phone calls. Drone manufacturers even admit they are made to carry "less lethal" weapons such as tasers or rubber bullets."

Drone surveillance may improve security, but have also been used to perpetrate human rights violations.

(b) Imagine in 2 years' time, in response to the growing backlash against the proliferation of surveillance cameras, Brown's Department of Public Safety (DPS) is considering purchasing a small drone fleet and is phasing out its surveillance cameras, so instead of 800 surveillance cameras there are only a handful of drones that will patrol campus. Who should be involved in making this decision and what factors should they consider? (3–4 sentences)

A3 (b): Your answer here.

I think tech leaders and scientists should be involved to consider technical accessibility and vulnerable cases of usages of surveillance drones. Economists should also participate into consider the affordability and sustainability of the change. Administrators need to consider the laws and policies about drones. Also, students and common visitors need to participate to consider their convenience and privacy issues.

Drones do not require cameras if they are remotely controlled within line of sight. However, drones use increasingly-sophisticated computer vision to aid control. In Homework 2, we learn about feature detection, description,and matching algorithms to provide correspondence between images. When applied to video, this allows drones to estimate their motion relative to the world. Please watch this video for a demonstration—at one minute in, we see the real-time feature matching in the bottom left.

As such, cameras and image correspondence are foundational technologies for reliable drone localization and navigation. Yet, adding cameras to drones also makes them potentially capable as surveillance platforms.

(c) With respect to cameras, find and cite a source for a solution we might consider to reconcile a desire for useful drones with a desire for privacy (3–4 sentences)
These might be technological, sociological, legal, etc.—wide scope.

A3 (c): Your answer here.

This article talks about using laws and some techniques to protect privacy from drone surveillance. It discussed about the restrictions of retaining images of identifiable people. Also, it suggests that drones can be placed only if it is proved to only track information on a specified crime issue, otherwise a warrant is needed.

Q4: Given a interest point location, the SIFT algorithm converts a 16×16 patch around the interest point into a 128×1 feature descriptor of the gradient magnitudes and orientations therein. Write pseudocode *with matrix/array indices* for these steps.

Notes: Do this for just one interest point at one scale; ignore the overall interest point orientation; ignore the Gaussian weighting; ignore all normalization post-processing; ignore image boundaries; ignore sub-pixel interpolation and just pick an arbitrary center within the 16×16 for your feature descriptor. Please just explain in pseudocode how to go from the 16×16 patch to the 128×1 vector. You are free to simplify the gradient computation.

A4: Your answer here.

```
# You can assume access to the image, x and y gradients, and their
                                    magnitudes/orientations.
image = imread('rara.jpg')
grad_x = filter(image, 'sobelX')
grad_y = filter(image, 'sobelY')
grad_mag = sqrt( grad_x .^2 + grad_y.^2 )
grad_ori = atan2( grad_y, grad_x )

# Takes in a interest point x,y location and returns a feature
                                    descriptor
def SIFTdescriptor(x, y)
    descriptor = zeros(128,1)
    patch = image[x-8:x+8, y-8:y+8] # get a 16 * 16 patch
    cells = divide(patch) # get 16 4*4 cells
    num_bins = 8
    for i in range(len(cells)):
        grad_ori, grad_mag = getMagAndori(cell)
        histogram = zeros(1, num_bins)
        for n in range(len(grad_ori)):
            index = floor(grad_ori[n] / ori_step) + 1
            histogram[index] = histogram[index] + grad_mag[n]
        descriptor[i*8:i*8+7] = histogram
    return descriptor
```

Q5: Distance metrics for feature matching.

(a) Explain the differences between the geometric interpretations of the Euclidean distance and the cosine similarity metrics. What does this tell us about when each should be used?

A5 (a): Your answer here.

Euclidean distance describes the distance between 2 points or 2 vectors. Cosine similarity metrics describes the angle between 2 vectors. The euclidean distance is influenced heavily by the vector magnitude while cosine similarity keeps same. In this sense, we should use cosine similarity metrics if we care more about orientation instead of magnitude. Also, we should use euclidean distance if we care more about magnitude other than orientation.

(b) Given a distance metric, what is a good method for feature descriptor matching and why?

A5 (b): Your answer here.

A good method for feature descriptor matching should focus on orientation, such as cosine similarity. Since images vary in size, brightness and lots of other factors, the magnitude can vary in a large range. In this sense, two similar features can be very far away from each other in magnitude such as euclidean distance. Therefore, a good method should care more about qualities such as angles, orientations that are not influenced by the size of the image.

Feedback? (Optional)

Please help us make the course better. If you have any feedback for this assignment, we'd love to hear it!