E27-Computer Vision

Weite Liu (Victor) and Fengjun Yang (Jack)

Work Credits

We collaborated extensively on all the tasks. Jack coded up the programs, while Victor debugged and prepared test cases. We tested our programs together. The first part of report on Laplacian pyramid blending is mainly written by Victor while the second part on hybrid image is mainly written by Jack.

## 1. Laplacian Pyramid Blending

### 1.1 Source Image Alignment and Mask Preparation

We used the face replacement of Bernie Sanders and Hillary Clinton as our test sample for Laplacian pyramid blending. The first step is to search for appropriate candidate photos online. Google's "search by image" function enables us to efficiently find the proper photos that have similar dimensions.


[Original photos of Bernie Sanders 544 px * 306 px, and Hillary Clinton 544 px * 302 px]

The next step is to crop the two images to be the same size and maintain the figures in each of the image to be relatively the same size and remain at relatively the same location with respect to the frame so that the mask could have a good match with the two faces. The final step is to align the two images so that after applying the mask, the faces of two figures would align well and give us a nice blending result. We tested the alignment manually and make adjustments by cropping and resizing the images using Photoshop.

[Processed Images of Bernie Sanders and Hillary Clinton, both 349 px * 302 px]

We chose to draw an ellipse as our alpha mask. In this case, we are replacing Clinton's face with Sanders's and an ellipse would cover the oval shape of face well enough. The parameters of mask were tested manually.

[Mask and Final Laplacian Pyramid Blending Result, 349 px * 302 px]

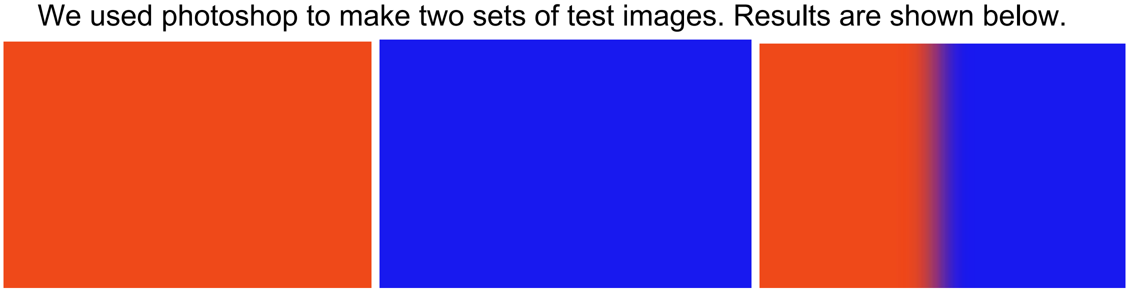
### 1.2 Depth of Laplacian Pyramid

The depth of our Laplacian pyramid is 8. The deeper the Laplacian pyramid is, the better the blending effect is. Namely, for example if the depth is only 1 or a small number, the effect is similar to just applying an alpha mask or a direct blending. We would clearly see a edge between Sanders's face and Clinton's face. And if the depth is 8 in this case, we won't see the sharpened seam of the ellipse and the facial color of Sanders, which is darker, would blend well with the facial color of Clinton, which is lighter. The final depth of 8 was determined by testing different parameters and we found for depth larger than 8, the difference of the final image is not significant.

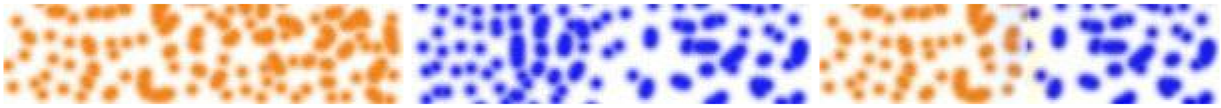
### 1.3 Low-frequency vs. High-frequency Blending

When doing Laplacian pyramid blending, we first build the Laplacian pyramid for each image. In this process, the image is subsampled and the high-frequency content will be left out and only low-frequency content will remain. Therefore, when doing blending for the higher levels and then reconstructing the image, the high-frequency content won't be recovered but rather the low-frequency content will be enlarged (like blending of facial colors of Sanders and Clinton in the example above). Therefore, low-frequency content will be blended over a larger distance than low-frequency content.

We used photoshop to make two sets of test images. Results are shown below.



[Blending of Two Low-frequency Images, a Wide Band in the Middle is Shown]



[Blending of Two High-frequency Images, a Clear Edge in the Middle is Shown]

## 2. Hybrid Image

### 2.1 Procedure

For hybrid image, we started by applying low-pass and high-pass filters on the images. Low-pass filter keeps low frequency content in the image and gets rid of high frequency contents, while high-pass filter does the opposite. Thus, the image processed with low-pass filter contains the general feature of an image but do not have much details, while the other image processed by the high-pass filter keeps the details but lost a certain proportion of general features.

### 2.2 Source Image

For this task, we used two grayscale images of Albert Einstein and Steve Jobs. We obtained the source images from the internet by intentional searching for images of our desired size (300px * 400px).

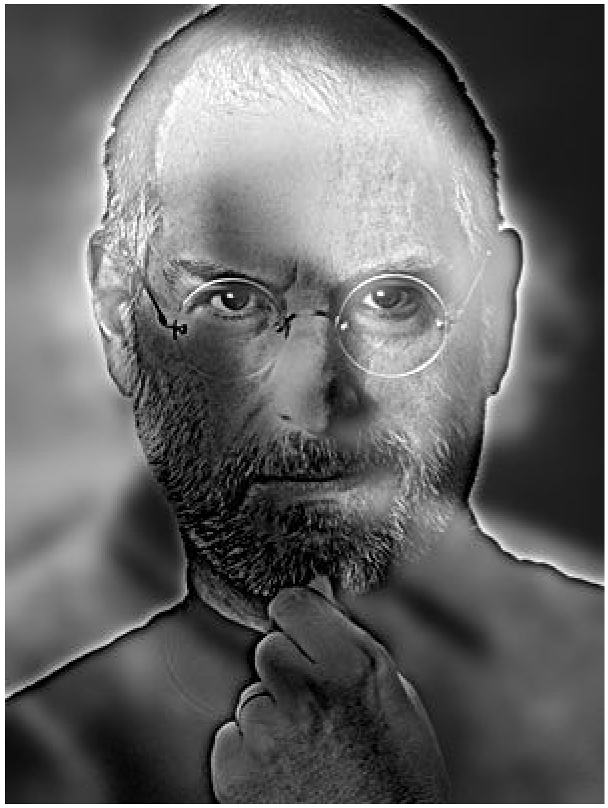
### 2.3 Tuning the Parameters

There are four parameters involved in the process of blending two images A and B. We refer to them as Ka, Kb, sigmaA, and sigmaB.

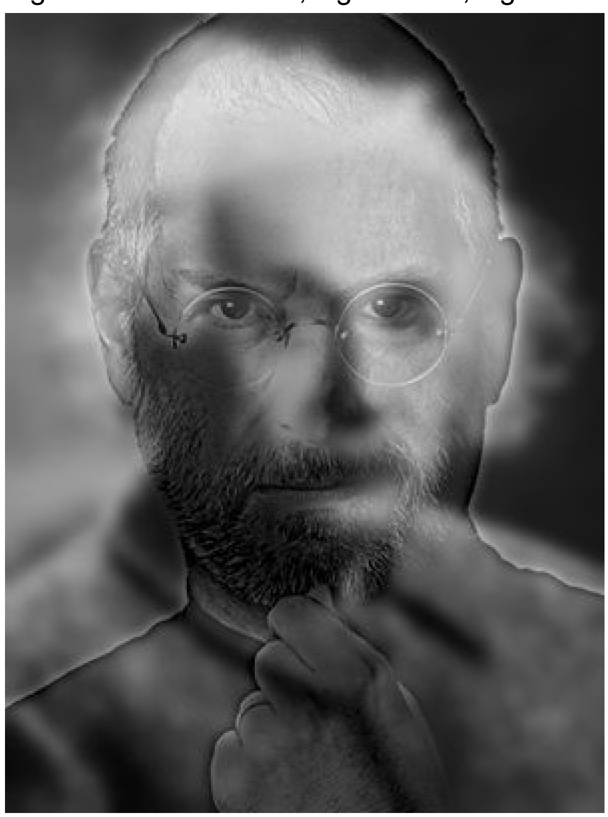
Ka and Kb are parameters that determine, when we are blending the two processed images together, how bright we want the two images to be respectively, i.e. how we want to weigh the two processed images. Using Ka > Kb means we want to stress the content in processed image A and vice versa.

SigmaA determines the amount of high-frequency contents, or details, we want to remove from image A, while sigmaB determines the amount of high-frequency contents we want to keep in image B. The greater sigmaA is, the more details we remove from image A, and the greater sigmaB is, the more details we preserve in image B.

We started with equal K values where Ka = Kb = 1, and a sigmaA = 5, slightly greater than our initial sigmaB = 3. The result is as follow is Image 2.3.1. In image 2.3 we can already find both a blurry outline of Einstein and a detailed face of Jobs. However, the content from the Jobs image in the foreground seems to be dominating and it is hard to see Einstein from this image. Thus we decided to tune down Kb, sigmaA, and sigmaB. After a few trial and errors, we arrived at final result, image 2.3.2.



[Image 2.3.1 Ka = Kb = 1, sigmaA = 5, sigmaB = 3]



[Image 2.3.1 Ka = 0.9, Kb = 0.4, sigmaA = 4, sigmaB = 3]

2.4 Analysis of Final Result

In general, the closer one is to the screen, the more details can one see, and the clearer image B would be. The farther one is to the screen, the less details can one see and the result will then seem more like image A.

The distance one need to be at to see image A and image B depends on the person's eyesight. When I have my glasses on and can see the details clearly, I can see the content of image B (Jobs) as far from the screen as 1 meter, while I take off my glasses and cannot see details even I am close to the screen, I can only see image A (Einstein) even I am just 30 centimeters away from the screen. Same results apply for my partner too.

The pyramid of the result of blended image also reflects the distribution of high-frequency and low-frequency contents in the result. At the top of the pyramid is the image with most high-frequency contents (details), and looks most like image B. As we go down the pyramid, we can find that the images preserves less and less high-frequency contents and look more and more like image A. The following example visualizes this fact.

[Image 2.4.1 The pyramid of depth = 4 for our hybrid result]