Assignment 2 - Pyramid Blending
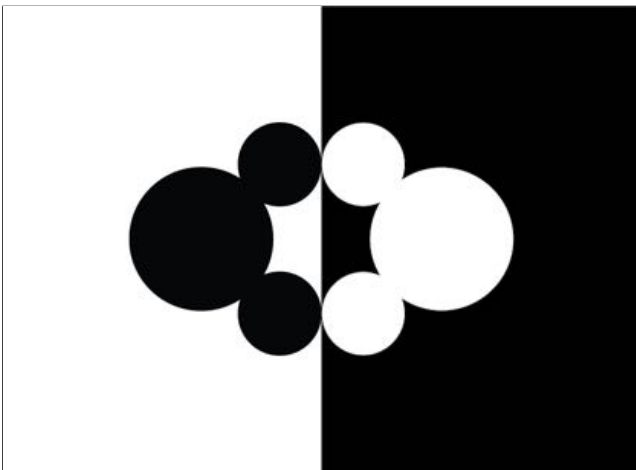
Andrew Parmar
CS6475 Fall 2022
aparmar32@gatech.edu

(a) Black Image

(a) White Image

(b) Completed Mask

(b) Final Blended Image

I. ORIGINAL BLACK, WHITE, AND MASK IMAGES

Image	Exposure	Aperture	ISO
Black	1/121	f/1.6	50
White	1/121	f/1.6	64

1) **Describe your original black and white images. What were your images of?**

In order to construct an original image with a controlled background. I chose to use colored art paper laid out in a grid on a table, and a wine bottle placed in the center of the grid. I used a different set of color paper to generate two unique images.

The image was taken in my living room with a large window directly positioned to the left. The color of the wine and the shape of the bottle created an interesting lighting effect on the paper as the light from the window passed through the bottle. To have the diffused lighting going in opposite directions for my two original images, I took one of the pictures by placing the wine bottle upside down and then rotating it 180 degrees before running it through the blending pipeline.

2) **How did you create your mask?**

I wanted a clean mask with multiple shapes. Adobe Illustrator was a good choice for this since it allowed for better vector shape manipulation than GIMP.

Importing my two images into illustrator and aligning them by making one image 50% translucent, I used trial and error to come to an alignment and crop that still gave me a 4:3 aspect ratio. The final image sizes were 834x610.

For my mask, I created a canvas of 834x610, as determined by the alignment and cropping process earlier. I split the canvas vertically in the center using two rectangles of equal size covering the left and right portions of the canvas. I colored the one to the right black and the other white.

For my mask, I wanted to use a well-known shape that would also provide some artistic effect. I chose the use the well-known Mickey Mouse head and ears logo. I started off creating my own version, but the proportions were not right. To mitigate this I searched and found the SVG Mickey Mouse head logo from Wikipedia Creative Commons. As this was already in SVG format I was able to directly download and import the vector shapes into my Illustrator canvas. After scaling, I rotated the logo by 90 degrees and also made a mirror copy. I vertically aligned these rotated logos to the center of the canvas and colored the logo on the white side black and the logo on the black side white. I wanted my final image

to have the outline of two Mickey Mouse logos on either side but of the opposite source image's color.
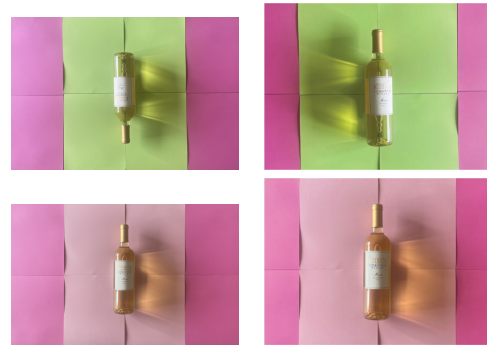

Fig. 3: Row 1: black image rotated + cropped + scaled; Row 2: white image cropped + scaled

II. DISCUSSION QUESTIONS

1) **How might you automate the blending point?**
Firstly this would be dependent on the desired outcome. If the blending process is intended for artistic/creative effect then it doesn't fully lend itself to automation.

However, one semi-automatic approach to generating a mask could be to use an edge detector to highlight all the edges in the source images. Once the edges are detected the user can use the edge image as a template to select the features that should be masked off.

2) **What is the significance of using a = 0.4 for the generating kernel?**
Using a = 0.4 ensures that the kernel generated has a functioning shape most like a Gaussian probability density distribution. If a = 0.5 were used instead the function shape would be triangular. A Gaussian shape is desired because it has the effect of low-pass filtering our target image. This in turn helps us smooth out our image for the creation of the next pyramid layer.

3) **Why does the output of expand_layer() have to be multiplied by 4?**
Experimenting with and without the multiplication by 4, I can see that without multiplying by 4 the Laplacian diff images are just dimmer versions of the original. The expanded image's matrix values are not large enough so the difference-image shows little variation from the original. The final resultant blended image has a hard boundary between the images at the mask line.

The reason for the lower values is that we first increase the size of our image by filling in blank rows and columns and then convolving over the increased size image. In the case of even rows and columns, only 9 of the 25 cells under the kernel contribute to the interpolation. In the case of odd cells, only 4 out of 25 cells under the kernel contribute. While the weights of values under a Gaussian normally add up to 1, in the case of this expanded image they only add up to 0.25. Multiplying by 4 then turns this weighting to 1 and thereby properly scales the image values up.

To determine this I manually added the actual values of the output of the $generatingKernel$ function over a grid of alternating rows and columns of zero and one.

4) **Do you think blending or using a cut (as discussed in Module 04-04) is a better approach for YOUR input images? Why?**

As I was going for artistic effect, blending worked out pretty well in my case. However, if my goal was to align the background images together perfectly then a cut approach would have yielded better results.

The main subject in my images, the wine bottle, was unintentionally misaligned during image capture. The end result after the blending process clearly shows the misalignment. A cut approach would have most likely produced a better boundary

III. Project Retrospective

If you were to do this project again, what approaches, concepts, or methods would you use to try and improve upon your final blended result? Discuss at least two.

As I was developing my original source images, I ran into a few issues with the background and placement of my main subject. Since I was using a table to create my multicolored grid, I had to place my object (wine bottle) sideways and then take the image from above. I didn't have a camera stand that could support this angle so I had to take the image free-handed. When switching out the grid pattern with different colors I had to again estimate the position of the subject, and take the image free-handed. In the end, this resulted in a noticeable misalignment of the wine bottle in my final results.

One potential solution is to I could use my monitor to display various backgrounds and have my subject in front of the monitor. With a tripod setup, I can then use the subject and my monitor as the backdrop to keep my subject stationary while trying our many different backgrounds. This is equivalent to using a green screen background. This would ensure that my subject is perfectly aligned between shots.

Second, I would choose a different subject and mask. My mask was selected primarily for artistic and lighting effects. I can see a better experiment being one where two separate subjects of similar size but different colors and materials are used for the blending process. Ideally, one of the textures would be smooth and the other rough, and likewise contrasting colors could also provide good visibility into the effects at the blending boundary. These adjustments would show the effects of the blending process much more effectively.

References

[1] P. J. Burt and E. H. Adelson, The Laplacian Pyramid as a Compact Image Code, 1983

[2] P. J. Burt and E. H. Adelson, A Multiresolution Spline With Application to Image Mosaics, 1983

[3] Ed Lessons, 4-1:3 , Accessed: Sept 11, 2022

[4] Ed Discussions, Assignment 2, https://edstem.org/us/courses/24462/discussion/1730514, Accessed Sept 11, 2022

[5] Philroc, Mickey Mouse Silhouette, https://commons.wikimedia.org/wiki/File:Mickey_Mouse_head_and_ears.svg, Accessed Sept 11, 2022