Hybrid Images

Angelic Phan

Abstract

This project explores the meaning of hybrid images, and the process it takes in rendering them. The idea for this project comes from the Oliva et al. research on hybrid images.[1] In their study, they examine how the overlay of two different images can produce a result in which the perception of the image is one object at a certain distance, and a different object at another distance. This illusion is created by combining a low-frequency image with a high-frequency image. Oliva et al. describes the process in which the hybrid image can be created in the frequency domain. However, this project is taking a different approach by constructing the image in the spatial domain. Either process will give the same results, but they employ slightly different mathematical formulas and methods.[2]

Keywords: hybrid image, low-frequency image, high-frequency image, spatial domain, frequency domain, Gaussian, Laplacian, convolution

Introduction

According to Oliva et al., a hybrid image is an image whose interpretation depends on the spatial distance in which you are observing the image.[3] From up close, you may see one object, but from far away, it may look like a completely different thing. This is the result of superimposing one low-frequency image and one high-frequency image together.[4] The low-frequency image is an image that has been filtered by a low-pass filter, such as a Gaussian filter (G_1).[5] This low-frequency image will thus be a blurred version of the original, making the details of the image less apparent. On the other hand, the high-frequency image will be one that enhances and keeps the more detailed features, such as the outline and edges of the subject in the image. Therefore, the high-frequency image is one that has been filtered by a high-pass filter such as a Laplacian filter ($1 – G_2$).[6] Since the high-frequency image contains the more detailed edges, this is the image that can be seen from up close. At the closer distance, the low-frequency image is blurry and harder to discern than the high-frequency image. At a further distance, the low-frequency image dominates over the high-frequency image, as the small detailed edges seem to fade away in the depths of the low-frequency image, and the blurry image begins to take shape. This is the illusion of the hybrid image.

---

[1] Oliva, Aude, Antonio Torralba & Phillippe G. Schyns. "Hybrid Images", SIGGRAPH, 2006, http://cvcl.mit.edu/publications/OlivaTorralb_Hybrid_Siggraph06.pdf.

[2] Another resource I have found that simplifies the explanation of the methodology and formulas used in the frequency domain is Mian, Ajmal, "CITS4402 Computer Vision: Week 4 Laboratory Exercises", The University of Western Australia, http://teaching.csse.uwa.edu.au/units/CITS4402/labs/week3/week3.html.

[3] Oliva et al., 527

[4] Oliva et al., 527

[5] Oliva et al., 527

[6] Oliva et al., 528

In this project, the Gaussian filter is created based on user input of the sigma value, σ. The sigma value represents the standard deviation of Gaussian, which determines how much to blur the input image by.[7] Thus, we will also discuss the method used for the Gaussian filter. Two sigma values are required. One for each input image, as each image is uniquely different. To find the best resulting hybrid image given two input images, different sigma values are tested using trial and error. Lastly, if we are to follow the method described in the Oliva et al. paper and construct the hybrid images in the frequency domain, we can multiply the input images by their respective filters and add the two resulting images together: $H = I_1 \bullet G_1 + I_2 \bullet (1 - G_2)$.[8] However, because we have chosen to produce our images in the spatial domain, we apply the filters to the input images by convolution: $H = I_1 \oplus G_1 + I_2 \oplus (1 - G_2)$. An important note, this project uses input images that will perfectly align with each other. Thus, this paper will not discuss the issue of aligning images together, and/or morphing images.[9]

Methodology

Gaussian Matrix

Given two input images, the goal of the project is to produce an acceptable hybrid image: $H = I_1 \oplus G_1 + I_2 \oplus (1 - G_2)$. Part of the process is to filter the input images by a low-pass filter. We use the Gaussian filtering for this step. Mentioned previously, this project creates its own Gaussian filter given a sigma value. Therefore, to obtain the same images as the ones produced in this project, the following steps are used for the creation of the Gaussian filtering matrix:

1. Given a sigma value σ, the filtering matrix will be of size: $2\sigma + 1$ x $2\sigma + 1$.[10]
2. For every element in the matrix, apply the following equation through convolution:[11]

$$G_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

   The x and y values are the indices for the center element in the filtering matrix. These values can be found by: round($(2\sigma + 1) / 2$), which takes half of the length of the width and height, and rounding to the median index.
3. Normalize the matrix by dividing each element by the sum of all the elements. As a result, the sum of all the values in the matrix will thus add up to 1.

[7] Liu, Feng. "Introduction to Visual Computing: Lecture 2", 2019, http://web.cecs.pdx.edu/~fliu/courses/cs410/notes/Lecture2.pdf, slide 41.
[8] Oliva et al., 528
[9] All of my example input images come from Unknown. "Project 1: Image Filtering and Hybrid Images", Georgia Tech, https://www.cc.gatech.edu/~hays/compvision/proj1/.
[10] Liu, slide 52. This project uses an adaptation of the general mask size formula. Liu originally uses: 2k + 1 x 2k + 1, where k ≈ 3σ. This adaptation allows for smaller filtering masks given non-fractional sigma values.
[11] Liu, slide 32.

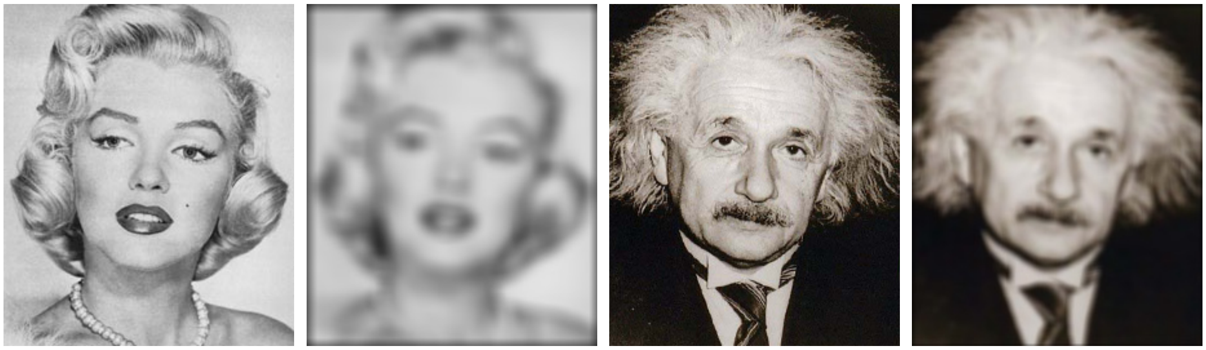

Figure 1: Example input images of Marilyn Monroe, Albert Einstein, and their respective low-frequency renderings. These results demonstrate step 1 of Image Rendering.

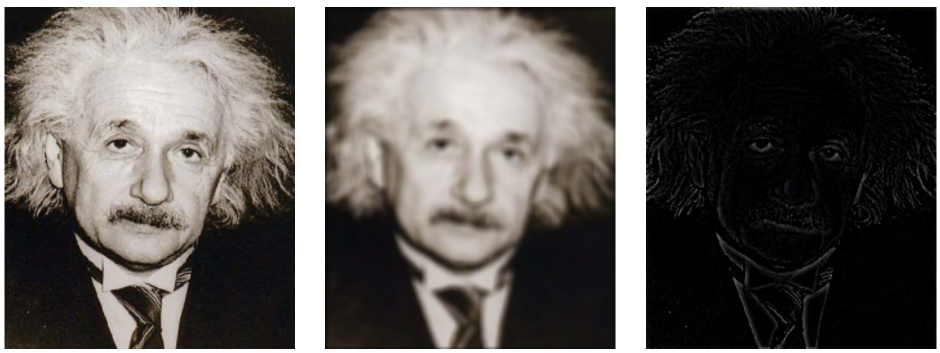

Figure 2: These images demonstrates step 2 of Image Rendering. To obtain the high-frequency image, we subtract its low-frequency image from the original image.

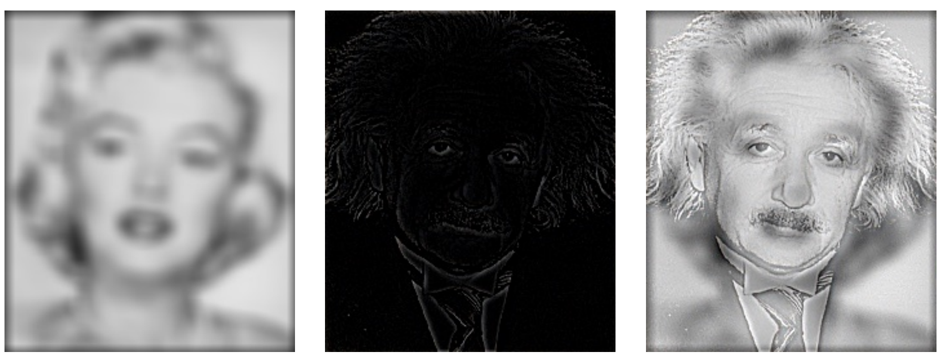

Figure 3: These images demonstrates step 3 of Image Rendering. To obtain the hybrid image, we add the high-frequency image to the low-frequency image.

Image Rendering

After generating the Gaussian filtering matrix, we can move on to working with our input images to apply the formula $H = I_1 \oplus G_1 + I_2 \oplus (1 - G_2)$:

1. Take one input image and apply the low-pass filter to it through convolution to obtain the low-frequency image: $LF_1 = I_1 \oplus G_1$. With image filtering, there is an edge case where the filtering matrix goes out of the bounds of the image matrix to calculate the values for the edge pixels. There are multiple ways in which we can treat the indices within the filtering matrix that goes

out of bounds. This project treats those indices as having RGB values of 0. Therefore, filtered images will have a noticeable boarder around the edges (Figure 1).

2. Take the second input image and apply the high-pass filter to it to obtain the high-frequency image: $HF_2 = I_2 \oplus (1 - G_2)$. Notice that we have not discussed the creation of the Laplacian filtering matrix. The same method employed for the creation of the Gaussian matrix can be used to find the Laplacian matrix, but with a different formula for the convolution step (Gaussian Matrix step 2). However, a simpler method is to subtract the low-frequency rendering from the original image to obtain its high-frequency rendering (Figure 2). This logic is valid, as the low-pass filter preserves the low-frequency components and removes the high-frequency components, while the high-pass filter does the opposite. By combining the two halves of a whole, we get the original image: $I = LF + HF$. Therefore, the formula for the high-frequency image can be represented as: $HF_2 = I_2 - (I_2 \oplus G_2)$.
3. Combine the low-frequency image with the high-frequency image to obtain the hybrid image: $H = LF_1 + HF_2$. This logic is like the one in step 2 where an image is a whole of two halves (a low-frequency image and a high-frequency image).

Continue to play around with the sigma values to obtain the most satisfactory result.

Results

The following are the results generated from the written program based off the methodology described above, and from sigma values found through trial and error. These are the images determined to provide the best results:

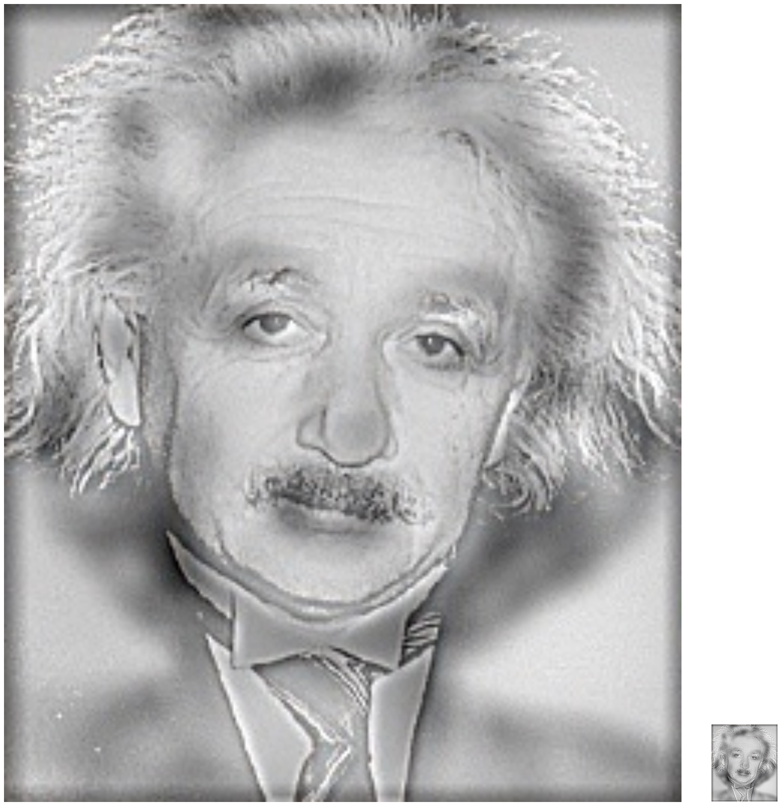

Figure 4: This is the resulting hybrid image given input images of Albert Einstein and Marilyn Monroe. From up close (or zoomed in), it is Einstein, but from far away (zoomed out), it is Monroe. This project found that a $\sigma = 3$ produced the best result for Albert Einstein, and $\sigma = 7$ produced the best result for Marilyn Monroe.

Figure 5: This is the resulting hybrid image given input images of a fish and a submarine. From up close, it is a fish, but from far away, it is a submarine. This project found that a $\sigma = 3$ produced the best result for the fish, and $\sigma = 5$ produced the best result for the submarine.

The hybrid image generated from input images of Albert Einstein and Marilyn Monroe turned out well (Figure 4). As did the hybrid image generated from input images of a fish and a submarine (Figure 5).

Figure 6: This is the resulting hybrid image given input images of a cat and a dog. The image on the left it is the hybrid image generated by Oliva et al.[12] The image in the middle is the result of the project. From up close, it is a cat, but from far away, it is a dog. This project found that a $\sigma = 7$ produced the best result for the cat, and $\sigma = 13$ produced the best result for the dog.

It is the hybrid image generated from the input images of a cat and dog that is slightly below expectations (Figure 6). The hybrid image turned out well, but in comparison to the image generated by Oliva et al., it is not quite a match. I suspect that a larger sigma value for the image of the dog will produce a more satisfying result. However, because convolution is a costly operation, I was unable to successfully run the program to completion with a sigma value of 15 for the image of the dog.

Conclusion

All the resulting hybrid images all turned out well. They were able to demonstrate that by combining a low-frequency image of one object and a high-frequency image of another, the resulting image will have two different perceptions based on the viewing distance. Furthermore, there is not much ambiguity in the resulting images. However, to improve or build upon the project, there are a few suggestions. First, in the creation of the Gaussian matrix, instead of using the method where we set the out-of-bounds RGB values to 0, we could mirror the pixels across the border edge, or try other methods to deal with the edge case. Second, it would be interesting to employ the method of hybridization in the frequency domain and test out its formulations. Third, we could introduce images in which the pair of input images do not perfectly align with each other; prompting us to implement a way in which we must analyze whether the input images need to be translated, flipped, rotated, resized, warped, or morphed to fit well with each other and produce a successful hybridization.

[12] Oliva, Aude, Antonio Torralba & Phillippe G. Schyns. "Hybrid Images gallery", Massachusetts Institute of Technology http://cvcl.mit.edu/hybrid_gallery/gallery.html.

References

Liu, Feng. "Introduction to Visual Computing: Lecture 2", 2019, http://web.cecs.pdx.edu/~fliu/courses/cs410/notes/Lecture2.pdf. Accessed February 7 2019.

Mian, Ajmal. "CITS4402 Computer Vision: Week 4 Laboratory Exercises", The University of Western Australia, http://teaching.csse.uwa.edu.au/units/CITS4402/labs/week3/week3.html. Accessed February 28 2019.

Oliva, Aude, Antonio Torralba & Phillippe G. Schyns. "Hybrid Images gallery", Massachusetts Institute of Technology http://cvcl.mit.edu/hybrid_gallery/gallery.html. Accessed February 7 2019.

Oliva, Aude, Antonio Torralba & Phillippe G. Schyns. "Hybrid Images", SIGGRAPH, 2006, http://cvcl.mit.edu/publications/OlivaTorralb_Hybrid_Siggraph06.pdf. Accessed February 7 2019.

Unknown. "Project 1: Image Filtering and Hybrid Images", Georgia Tech, https://www.cc.gatech.edu/~hays/compvision/proj1/. Accessed February 7 2019.