Project 2 Writeup

Lily and George

Contributions

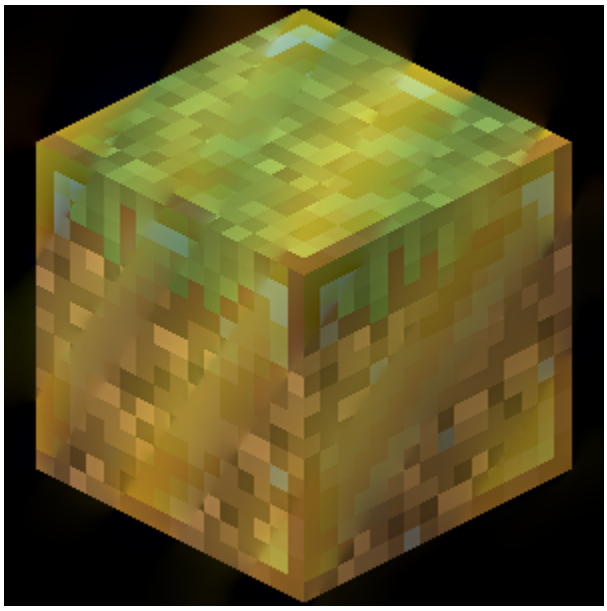
This project was a little wonky because both Lily and George had health issues midway through, but we still managed to contribute evenly. George wrote the pyramid blending, Lily wrote the reconstruction and combination of the pyramids. George wrote the hybrid images, and both of us worked on the "going" further portion of the project—we created a method to **automatically detect faces and align** the second image on top of the first, which is demonstrated in the presidential swap below.

Reference Attribution and Image Alignment

For this project, we obtained very similar Minecraft block images from a Minecraft wiki for the basic Laplacian blending. The photos of Joe and Obama used for the other tests were obtained from the official White House website (in the case of Barack from an archive of the site). For, demonstrative purposes, we obtained a picture of the sea from National Geographic and the sky from the Wikipedia image created by Mohammed Tawsif Salam. For alignment, none was necessary for the Minecraft blocks / demonstrative images because they are perfectly aligned already (used as an easy example). For the Presidents, we used an OpenCV face recognition model to translate the second image to align automatically with the first (see Python files starting with `face_detect...`. There was an attempt to scale the images properly as well, but time and the complexity of the task did not allow us to finish it.

Pyramid Blending Process

Include representative Laplacian pyramid blending outputs in your PDF.




To the left is a grass block blended with a gold block using a striped mask. To the right is Obiden, which is Obama's torso automatically blended with Joe Biden's slightly larger face using an OpenCV facial detection model to define a mask and the implementation of our Laplacian pyramid blending process.

Why does Laplacian pyramid blending blend low-frequency content over a larger distance than high-frequency content?

This example image demonstrates how Laplacian blends low-frequency content over a larger distance than high frequency content. Note how the color subtly and more smoothly changes from the different blues in the top right, while it fails to blend the high-frequency waves over

much distance at all in the bottom right. Laplacian pyramid blending is really good at breaking down the image into different bands of frequencies, and due to the algorithm's construction, weights the image representation more heavily on the lower frequencies, especially at the higher, smaller levels of the pyramid. So, when we stitch the images together, the alpha-blended images of low frequency content have more weight / are better reconstructed at the higher, smaller levels of the Laplacian images, while at the lower, larger levels, the high frequency content quickly changes over a short distance.

Hybrid Process

Include your hybrid image output in your PDF.

Pictured above is Joe Biden from a distance, and Barack Obama up close.

How did you arrive at good values for the constants σ_A, σ_B, k *for the hybrid image generation?*

Arriving at proper values for σ_A, σ_B, k was basic trial and error. First of all, we determine σ_B by multiplying σ_A and k, so we don't have to worry about that. We know that a larger σ_A will only retain the lowest frequency content in image A, and the smaller the scalar k, the more frequencies are removed from image B. We started with $k = 2$ as suggested in the project details, and decremented/incremented to find a suitable value of $\sigma_A = 2$.

If you display your hybrid image at full size on your computer screen, how close do you need to be in order to primarily see image B? How far away do you need to get before you only see features from image A?

You can be at most 4 feet away from our large screen to see image B. As you get further away, you begin to see image A. Yes, the distances are fairly consistent and the transition is seen across surveyed observers.

To what extent does each pyramid level resemble image A or image B, and why?

The pyramid starts resembling image B, then eventually transitions (as it gets smaller) to image A. This is because the Laplacian pyramid represents different frequency bands; at the lowest largest level, the highest frequency content is represented, and as the pyramid becomes higher and smaller, it represents more and more of the low frequency content. The Gaussian blur and downsampling operation acts as a low-pass filter, which incrementally removes high frequency content and saves low frequency content after each iteration, which explains why Barack becomes Joe.[1]

[1]Note: This was especially confusing to us because the pyramid building algorithm specifies that we subtract a low-sigma Gaussian blur from each level's downsampled image, which should imply that we subtract low-frequency content at each level. So how do we end up with the lowest frequency content at the highest level if at each iteration we subtract more low frequency content? Lily and I spent a lot of time discussing this question and reviewing our notes, without finding a well-formed answer. We'd appreciate your feedback on this one.