Using Image Pyramid for Image Blending

Elif Gürkan

Introduction

The technique of blending disparate images into a seamless composite is a foundational task that serves a multitude of applications, from artistic endeavors to practical implementations in software development and data visualization. The objective of this assignment was to explore and implement a multi-scale blending technique using image pyramids, specifically the Laplacian pyramid, to achieve a smooth and visually coherent transition between two distinct images.

The theoretical underpinning of this approach is grounded in the concept of spatial frequency representation of signals—in this case, the images. By decomposing images into different levels of detail, manipulate them at various scales, retaining control over how different features are merged. This process, akin to tuning the harmonics in an audio signal, allows for a refined synthesis of the final image that maintains the essence of both source images while creating a new, hybrid entity.

Our journey through this project involved the creation of Gaussian and Laplacian pyramids, which are instrumental in achieving a graduated blend. The Gaussian pyramid serves as a tool for downscaling the images, reducing their resolution while applying a Gaussian blur, which prepares the images for the subsequent step of creating the Laplacian pyramid. The Laplacian pyramid then captures the detail at each level of the Gaussian pyramid, allowing us to blend images by combining these details while controlling the spatial frequency components.

The culmination of the assignment was not only the creation of a visually appealing composite image but also the demonstration of the signal processing principles that underlie digital image manipulation. By sampling and visualizing the intensity values of the blended image, we can bridge the continuous world of visual perception with the discrete realm of digital processing, offering a tangible representation of how images are encoded and manipulated in the digital domain.

This report documents the methodology employed to execute the blending technique, the results obtained, and the insights gained from the exercise. It stands as a testament to the fascinating interplay between mathematical concepts and practical applications in the ever-evolving field of digital image processing.

Implementation Steps

Step 1: Kernel Generation

I developed a function to generate a Gaussian kernel, which is utilized for blurring during the downscaling and upscaling processes. This kernel helps in maintaining the continuity of the image signal and in reducing artifacts.

```
,def generatingKernel(a):

w_1d = np.array([[0.25 - a/2.0, 0.25, a, 0.25, 0.25 - a/2.0]])

return np.outer(w_1d, w_1d)
```

Step 2 : Image Reduction

I created a function to reduce the image size, which involves convolving the image with the Gaussian kernel and then subsampling it. This function was critical for building the Gaussian pyramid.

```
def reduce(image):

kernel = generatingKernel(0.4)

convolved = scipy.signal.convolve2d(image, kernel, mode='same')

reduced = convolved[::2, ::2]

return reduced
```

Step 3: Image Expansion

I created a function to formulate to expand the image size by upsampling and then convolving with the Gaussian kernel. This function was used to construct the Laplacian pyramid from the Gaussian pyramid in the next steps.

```
def expand(image):

upsampled = np.zeros((2 * len(image), 2 * len(image[0])))

upsampled[::2, ::2] = image

kernel = generatingKernel(0.4)

convolved = scipy.signal.convolve2d(upsampled, kernel, mode='same')

expanded = convolved * 4

return expanded
```

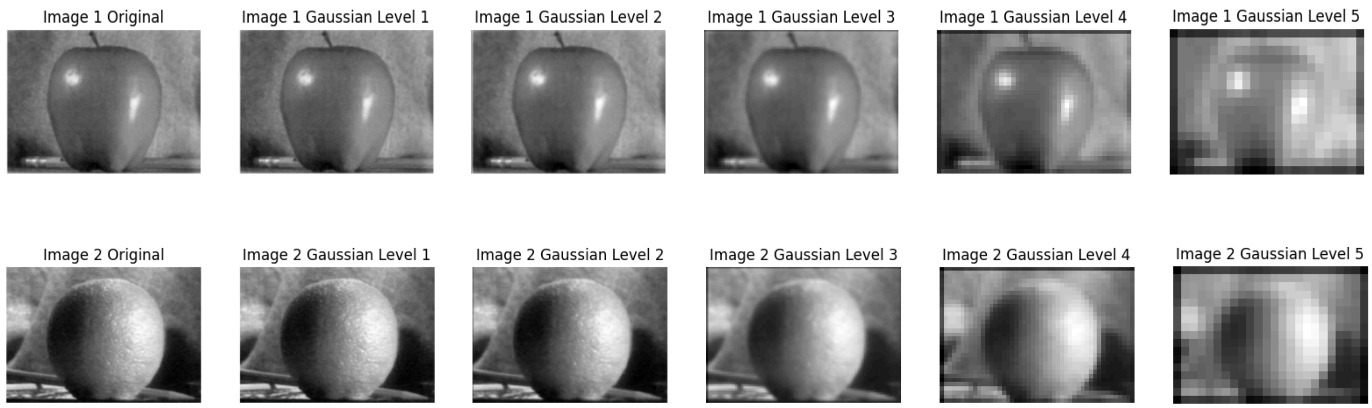
Step 4 : Gaussian Pyramid Construction

I constructed a Gaussian pyramid for each input image by iteratively applying the reduction function. This pyramid represents the image at multiple scales with increasing levels of Gaussian blur and reduced spatial resolution.

```
def gaussPyramid(image, levels):

output = [image]

for i in range(levels):

output.append(reduce(output[i]))

return output
```

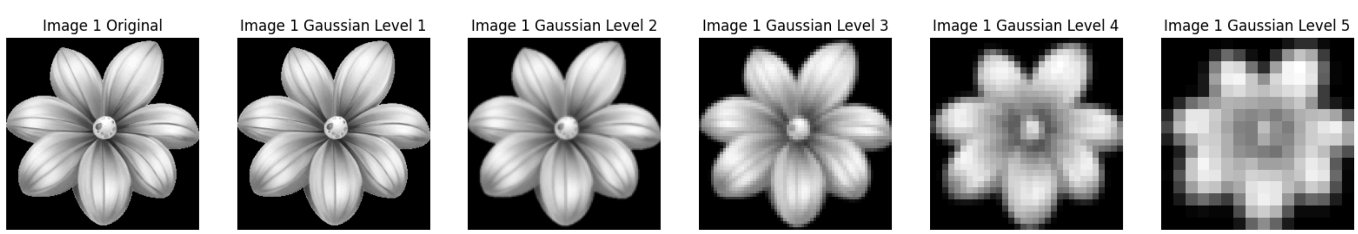
These are my some outputs with Gaussian Pyramid Construction:

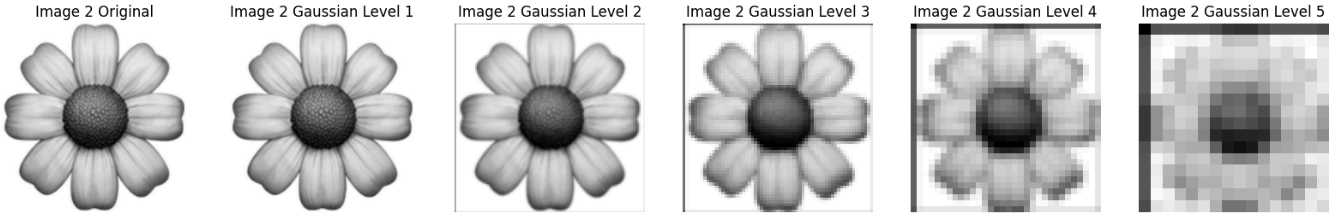
1. Apple and Orange



2. Cat and Dog

3. Flowers





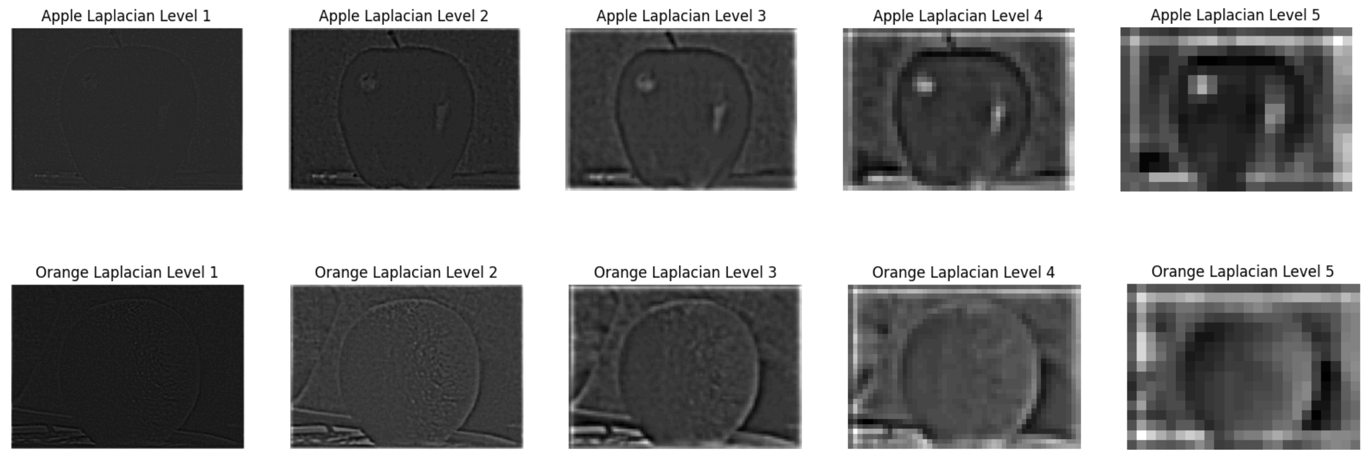
Step 5: Laplacian Pyramid Construction

I developed the Laplacian pyramid from the Gaussian pyramid. This step involves calculating the difference between each level of the Gaussian pyramid and its expanded upper level, effectively capturing the detail lost between scales.

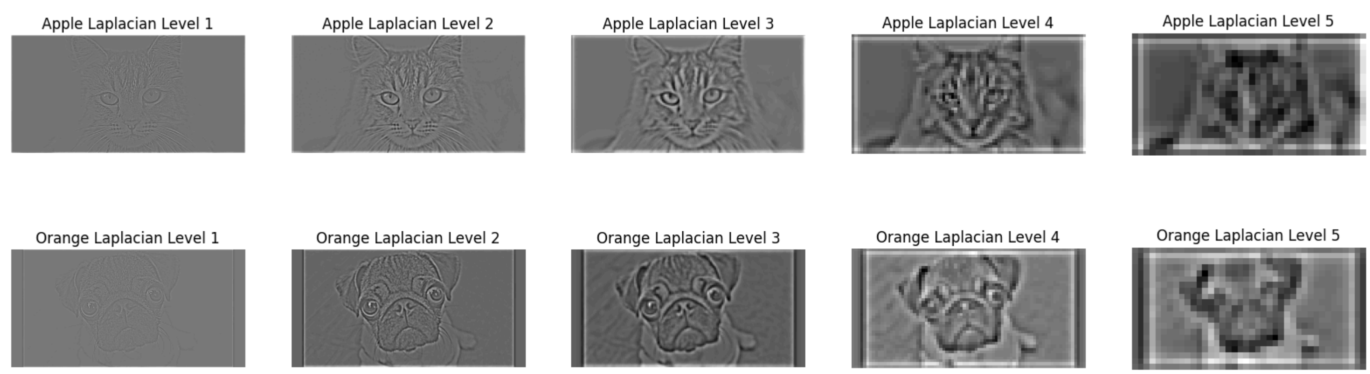
```
def laplPyramid(gaussPyr):

output = []

for i in range(len(gaussPyr)-1):

expanded = expand(gaussPyr[i+1])

if expanded.shape[0] != gaussPyr[i].shape[0]:

expanded = expanded[:gaussPyr[i].shape[0], :]

if expanded.shape[1] != gaussPyr[i].shape[1]:

expanded = expanded[:, :gaussPyr[i].shape[1]]

output.append(gaussPyr[i] - expanded)

output.append(gaussPyr[-1])

return output
```

These are my some outputs with Laplacian Pyramid Construction:

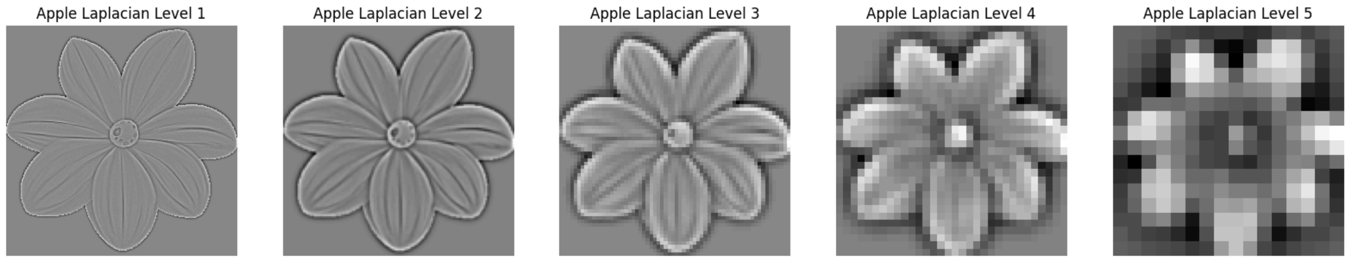
1. Apple and Orange

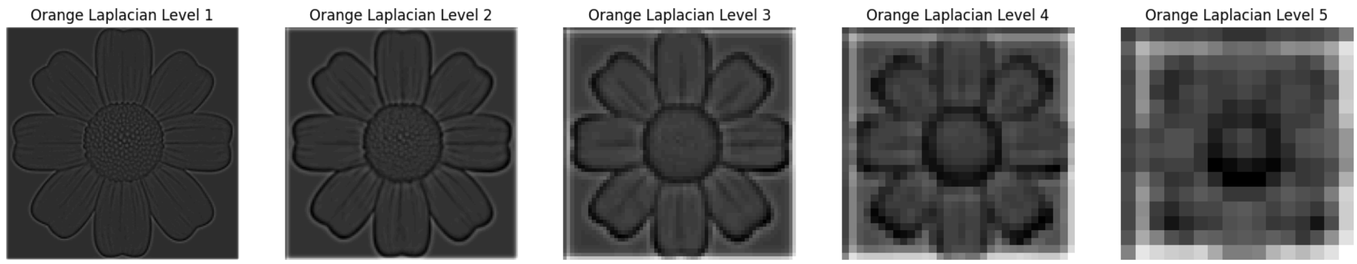


2. Cat and Dog

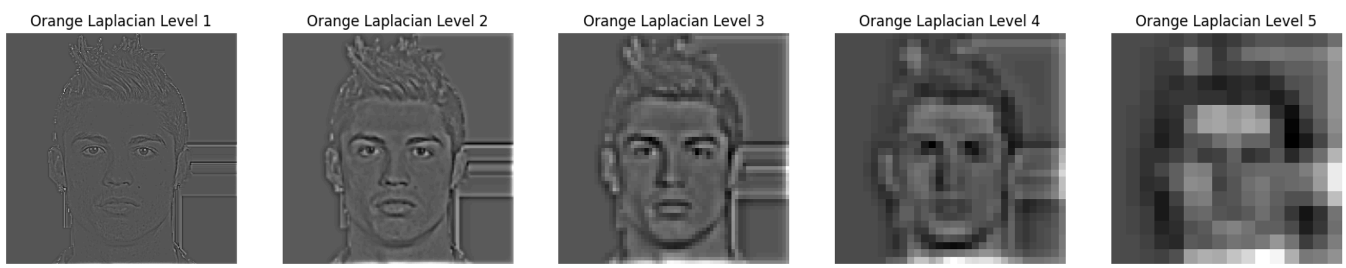


3. Flowers





4. Football Players



Step 7 : Blending Pyramids

I implement the blending of two Laplacian pyramids using the Gaussian pyramid of the mask. This step is where the actual blending of images occurs at multiple scales.The blending of pyramids enables us to merge images in a way that ensures a smooth transition between textures and details, avoiding harsh edges or abrupt changes. The result is a composite image that appears natural and cohesive, benefiting from the multi-scale representation of image details.A Laplacian pyramid is a data structure that decomposes an image into a series of levels that represent band-pass filtered versions of the image. Each level of the pyramid stores the difference between an image and a lower-resolution approximation of it. This difference, or the 'Laplacian', captures the image's details lost during the downscaling (Gaussian blur and subsampling). For blending purposes, I constructed two such pyramids: one for each image that I intend to blend together.The mask determines which parts of each image will be visible in the final blend. A Gaussian pyramid of the mask is created to have the same number of levels as the Laplacian pyramids of the images. At each level, this pyramid holds a blurred and subsampled version of the original mask, which matches the corresponding level of detail in the Laplacian pyramids.

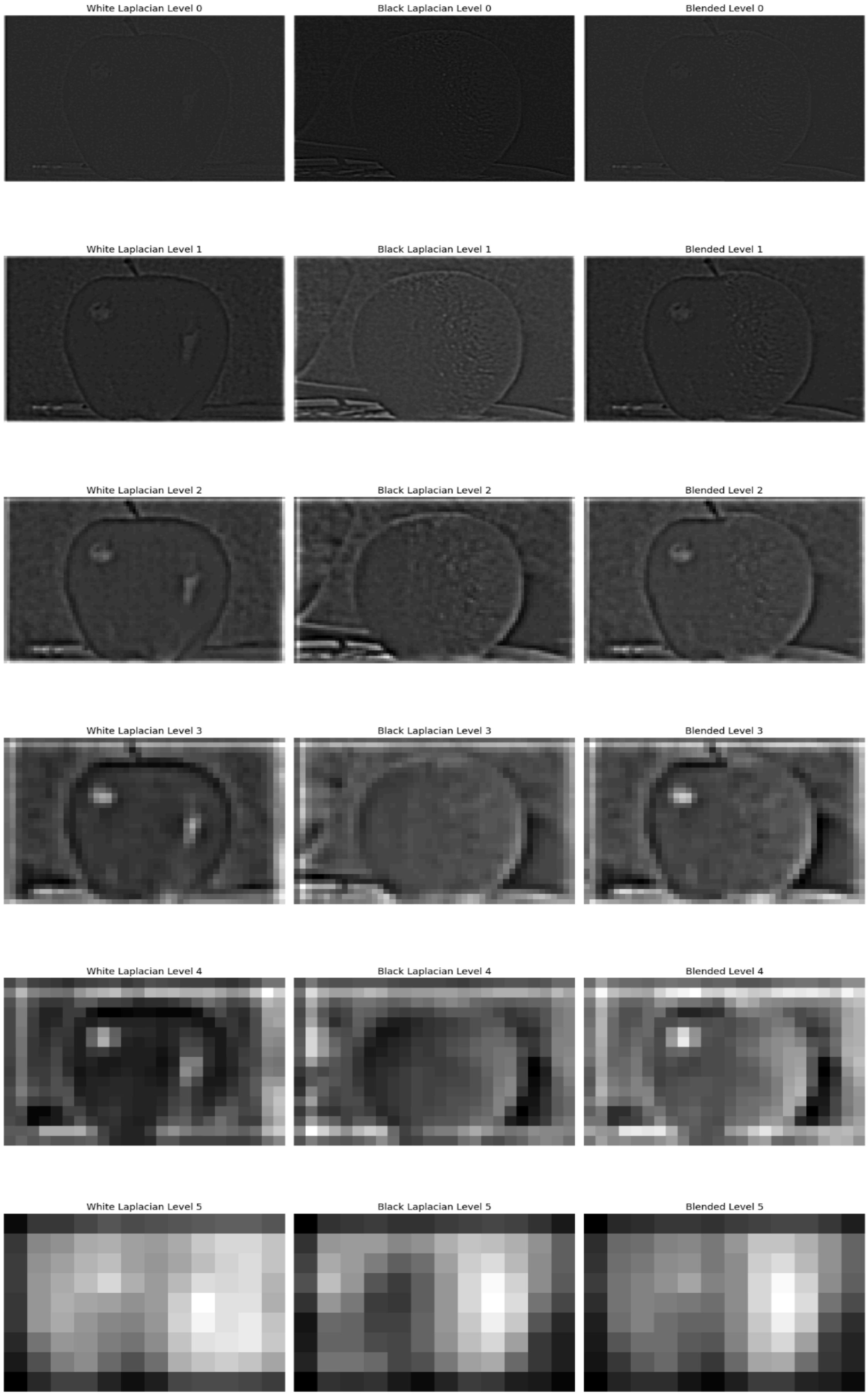
Blending Operation: The blending operation proceeds in a level-by-level manner:

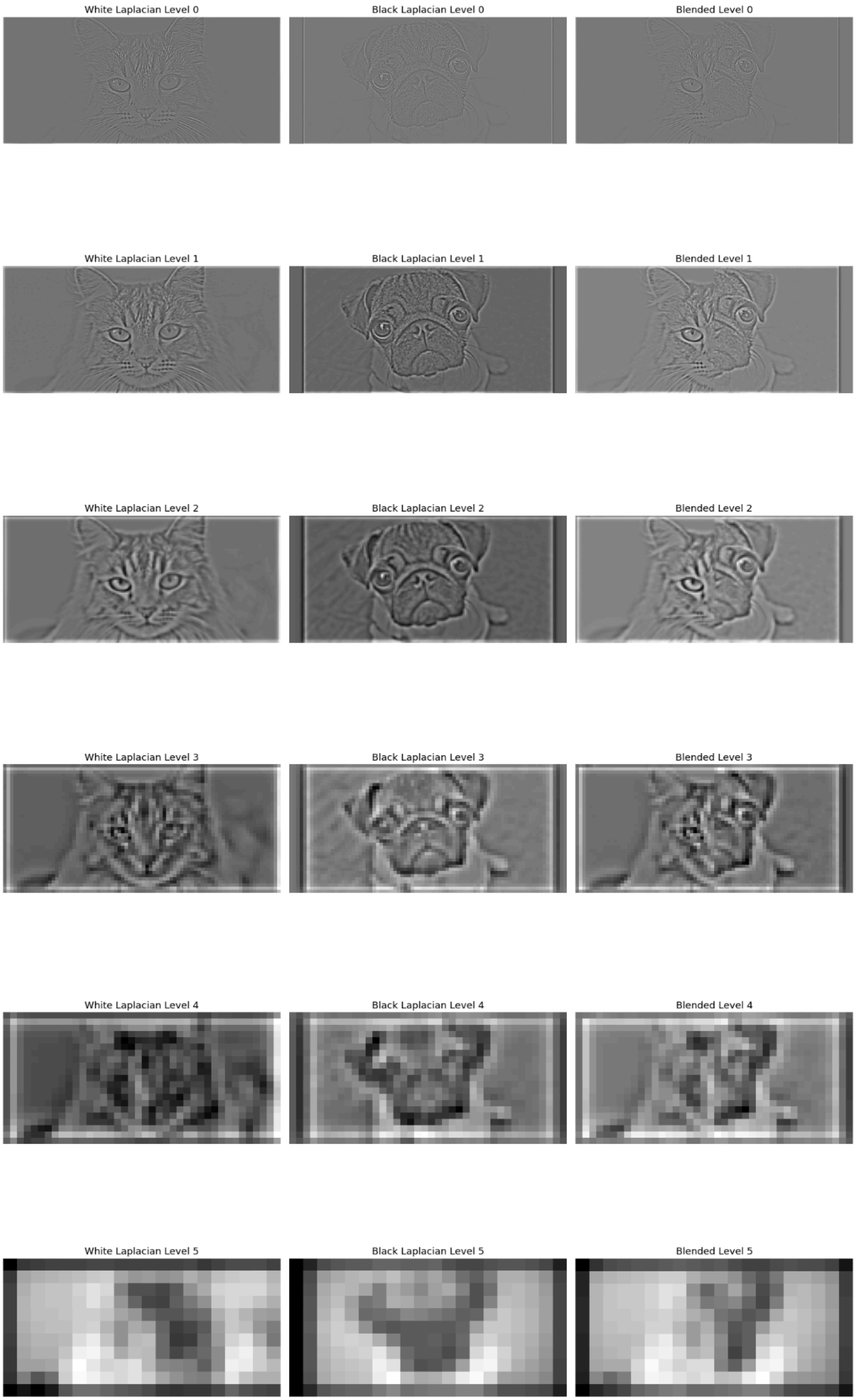
Level-wise Multiplication: At each level of the pyramids, I multiplied the Laplacian level of the first image by the Gaussian level of the mask. This operation selectively retains the details of the first image wherever the mask is white (or has high intensity)

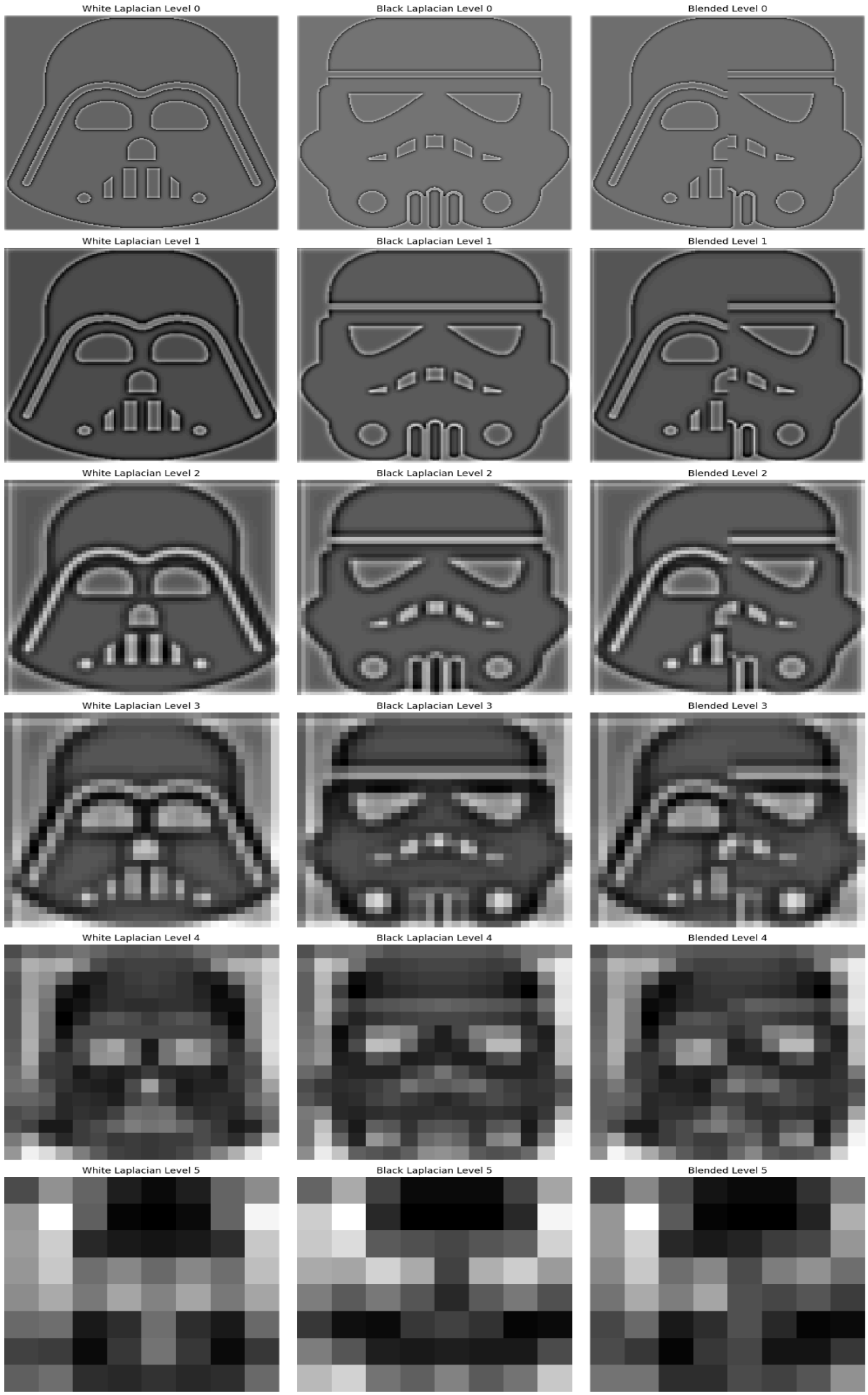
.Inverse Mask Multiplication: Simultaneously, I multiplied the Laplacian level of the second image by the inverse of the Gaussian level of the mask (i.e., 1 - mask). This retains the details of the second image where the mask is black (or has low intensity).

Combining Details: The two results from the above steps are added together to form a single blended level of the pyramid. This combined layer holds the blended details of both images, with the transition between them governed by the mask's gradient.

Creating the Blended Pyramid: The blended levels from all the layers are collected to form a complete blended pyramid. Each level of this pyramid now contains the seamlessly blended details corresponding to the spatial frequency band it represents.


White Laplacian Level 0
Black Laplacian Level 0
Blended Level 0
White Laplacian Level 1
Black Laplacian Level 1
Blended Level 1
White Laplacian Level 2
Black Laplacian Level 2
Blended Level 2
White Laplacian Level 3
Black Laplacian Level 3
Blended Level 3
White Laplacian Level 4
Black Laplacian Level 4
Blended Level 4
White Laplacian Level 5
Black Laplacian Level 5
Blended Level 5


White Laplacian Level 0
Black Laplacian Level 0
Blended Level 0
White Laplacian Level 1
Black Laplacian Level 1
Blended Level 1
White Laplacian Level 2
Black Laplacian Level 2
Blended Level 2
White Laplacian Level 3
Black Laplacian Level 3
Blended Level 3
White Laplacian Level 4
Black Laplacian Level 4
Blended Level 4
White Laplacian Level 5
Black Laplacian Level 5
Blended Level 5


White Laplacian Level 0
Black Laplacian Level 0
Blended Level 0
White Laplacian Level 1
Black Laplacian Level 1
Blended Level 1
White Laplacian Level 2
Black Laplacian Level 2
Blended Level 2
White Laplacian Level 3
Black Laplacian Level 3
Blended Level 3
White Laplacian Level 4
Black Laplacian Level 4
Blended Level 4
White Laplacian Level 5
Black Laplacian Level 5
Blended Level 5

Step 8 : Pyramid Collapse

The process of pyramid collapse is the final and crucial step in multi-scale image blending using pyramids. It involves reconstructing a single, high-resolution image from the blended pyramid, which contains the composite details of the images across multiple scales. The collapse of the blended pyramid is a reverse operation of the pyramid construction process, synthesizing a full-scale image from the series of successively smaller, band-pass filtered images.

To reconstruct the final composite image from the blended pyramid: I started from the top: the process began with the smallest, highest-level image in the blended Laplacian pyramid. This image contains the finest details from the blending operation, representing the highest spatial frequencies.

Then I continued with upsampling and expanded this image to a higher resolution, which is the next lower level in the pyramid. This step involves upsampling the image and then applying a Gaussian filter to interpolate the missing pixel values. The result is an image that matches the size of the next level in the pyramid but still contains only high-frequency details.

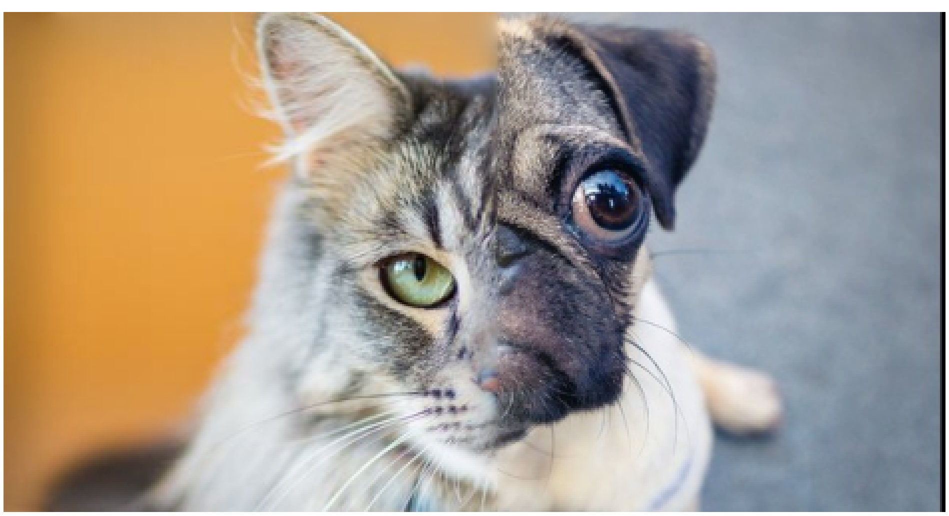
Then the pyramid adds the upsampled image to the next level down. This level holds the difference between two levels of the Gaussian pyramid and thus represents a specific band of spatial frequencies. By adding the upsampled image to this level, you reintroduce the higher frequency details that were missing.The iterative process repeats the process of upsampling and adding for each successive level of the pyramid. With each iteration, the image gains more details across the spectrum of spatial frequencies. Continuing this process until I reached the base level of the pyramid. The base level is the lowest frequency representation and provides the overall structure to which the finer details are added

During the process, it’s essential to ensure that the size of the upsampled image matches the next level in the pyramid before adding them together. If there is a size mismatch due to even-to-odd dimension changes during downsampling, you may need to crop or pad the upsampled image to match the size precisely.

My final outputs:

1. Apple and Orange Blending:

2. Cat and Dog Blending



3. Football Players Blending

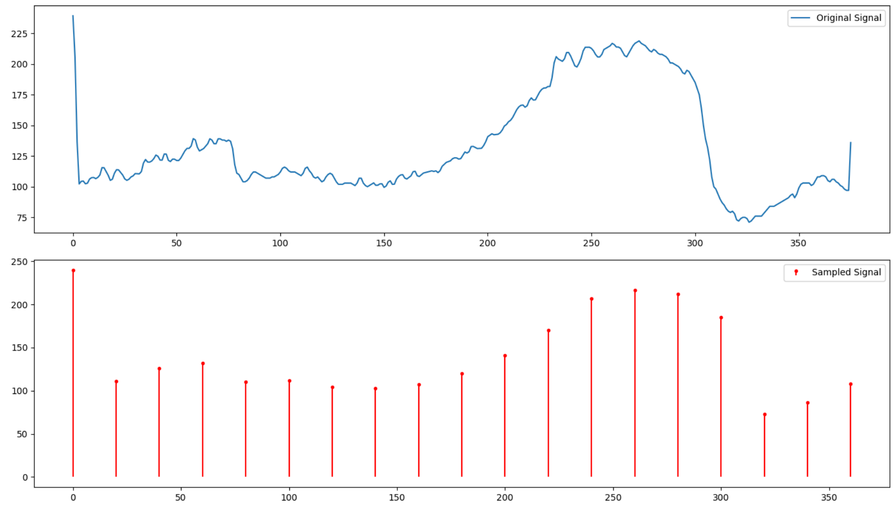
4. Butterfly Blending

5. Flower Blending

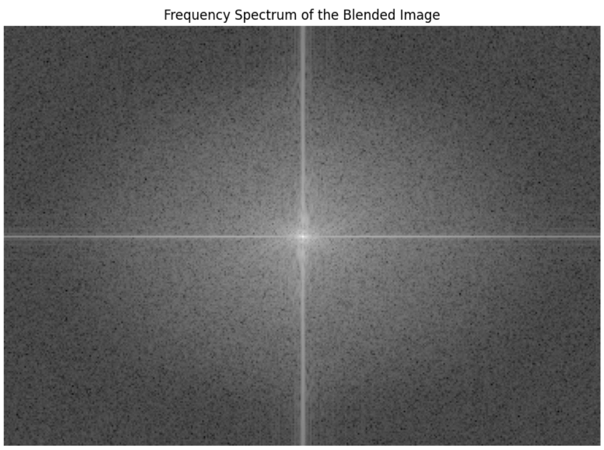
6. Character Blending

Step 9: Analysis

1.Apple and Orange Blending

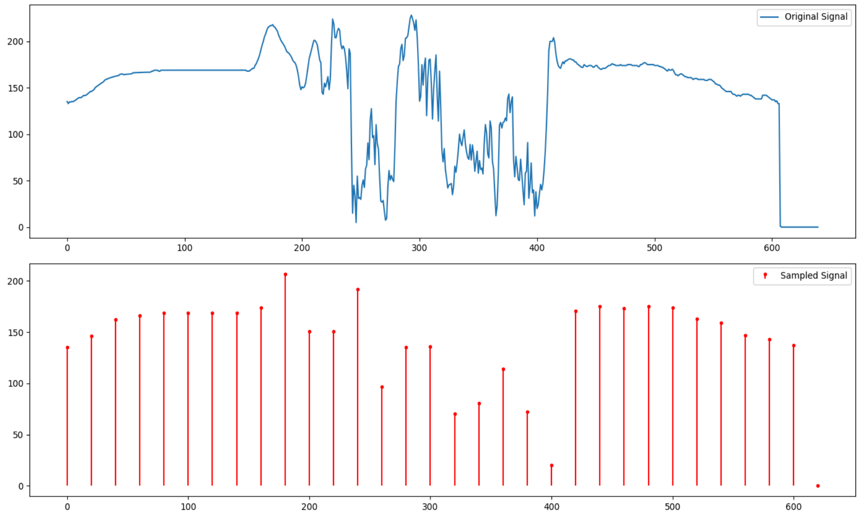


The blending operation has been successful in creating a smooth transition between the apple and orange images, as evidenced by the smooth gradients in the original signal plot.The sampled signal demonstrates the effectiveness of the sampling strategy and underscores the importance of choosing an appropriate sampling rate to capture the essential details of the image.The smooth transitions in the curve indicate the areas where the blending between the apple and orange images is gradual.Peaks and troughs correspond to areas of high contrast within the image, such as edges or texture details.

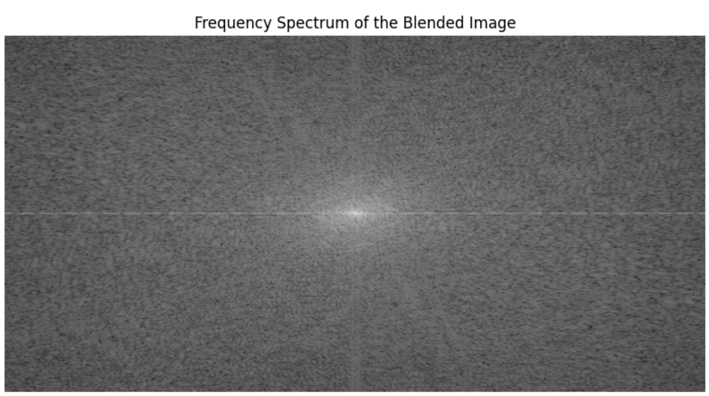


The predominance of low-frequency components in the blended image's frequency spectrum confirms that the blend between the two images is smooth, without harsh transitions or stark contrasts.The frequency spectrum also reinforces the need for careful blending to avoid creating high-frequency artifacts that could detract from the visual coherence of the final image.Low-frequency components are associated with the smooth, unchanging areas of the image, suggesting that the blending between the apple and orange images created regions with gradual tonal variations rather than abrupt changes.The relative darkness indicates that these high-frequency components are less dominant than the low-frequency ones, which is expected in a well-blended image where sharp transitions are minimized.

2. Cat and Dog Blending

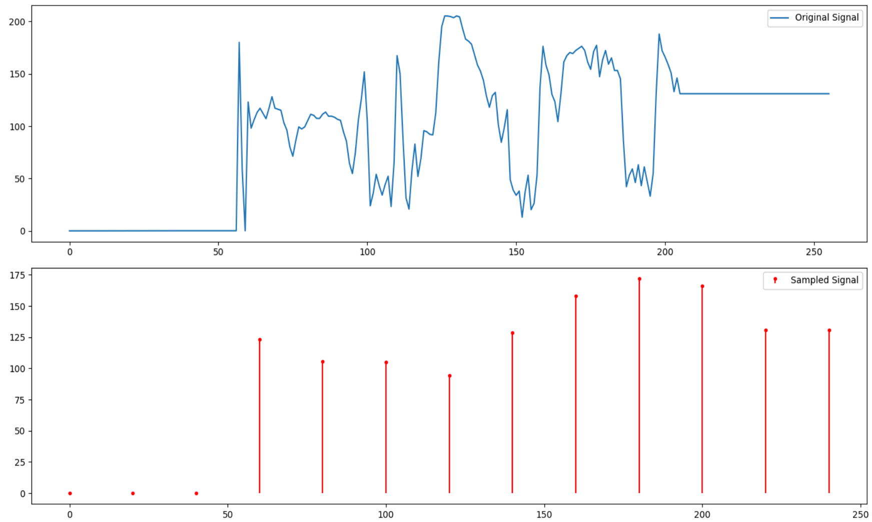


The original signal plot shows a more complex intensity profile than before, indicating that the horizontal line from the blended image crosses through regions with varying detail and contrast.The variability in the signal suggests that the line may cross different textures or edges within the blended image, leading to the peaks and valleys seen in the plot.The abrupt drop towards the end might indicate an edge or boundary within the image where the intensity changes drastically.The sampling captures the general trend of the signal but may miss smaller details between the samples, which is evident in areas where the original signal has sharp peaks or rapid changes.Some high-frequency details are not captured due to the lower sampling rate (seen as the distance between the red stems), which could lead to loss of detail or aliasing in a reconstructed signal.The complexity of the original signal indicates that the chosen line of the image contains a mixture of smooth gradients and detailed textures.The sampled signal plot demonstrates the importance of choosing a proper sampling rate to capture the critical features of the signal without losing significant information.



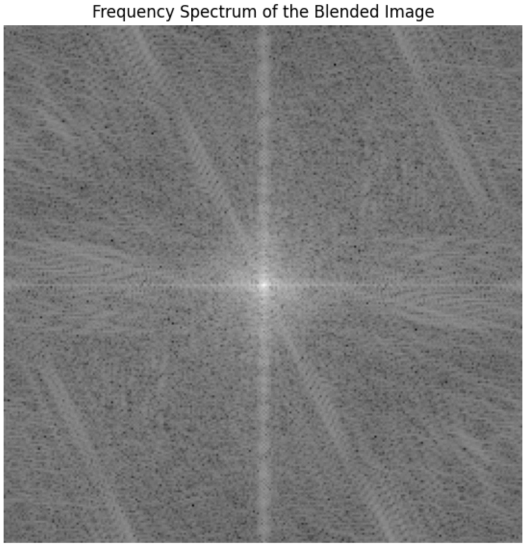
The frequency spectrum appears to have a bright central region and less contrast in the outer regions compared to the previous spectrum.The concentrated brightness along the horizontal and vertical axes in the center suggests the presence of strong horizontal and vertical features in the blended image.The less contrasty, more homogenized appearance of the spectrum implies that the high-frequency details are more evenly spread out or that the image has less high-frequency content overall. This could indicate a more thorough blending or a smoother image with fewer sharp edges and fine textures.

3. Football Players Blending



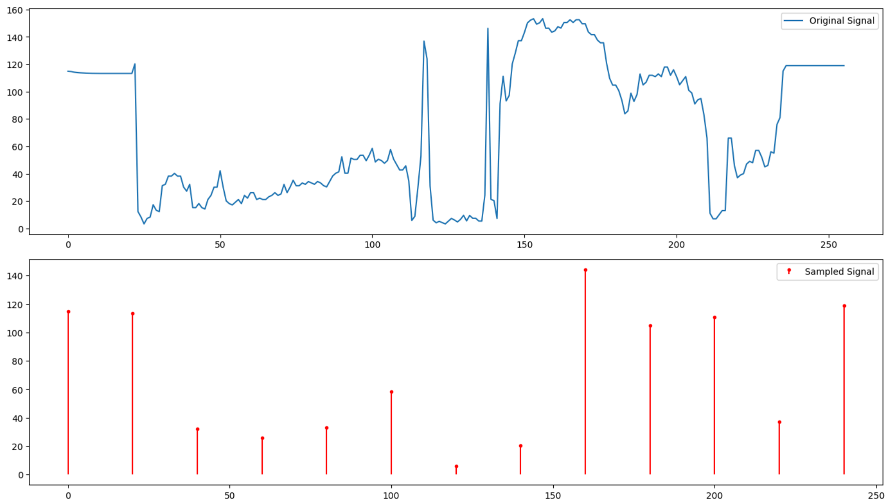
The plot shows a signal with a dynamic range of intensity values and several notable peaks and troughs.The sharp spike at the beginning could indicate a boundary or a high-contrast edge within the image.The fluctuations in the middle of the plot, where there is a high frequency of peaks, suggest regions with significant detail or texture, possibly where the two images converge in the blending process.The sampled signal indicates selected intensity values at regular intervals.The choice of sampling rate (the distance between red stems) is critical; here, it appears to be quite sparse. This results in missing some of the finer details present in the original signal, evident by the absence of red stems where there are peaks in the original signal.This undersampling could result in a loss of detail or aliasing, where the sampled signal does not accurately represent the original signal's details.The signal sampling indicates that a denser sampling rate might be necessary to capture the full detail of the blended image's signal, particularly in areas with high-frequency changes.The original signal's complexity reinforces the idea that the blending process involves regions with varying levels of detail, affecting the texture and overall appearance of the final image.

Frequency Spectrum of the Blended Image



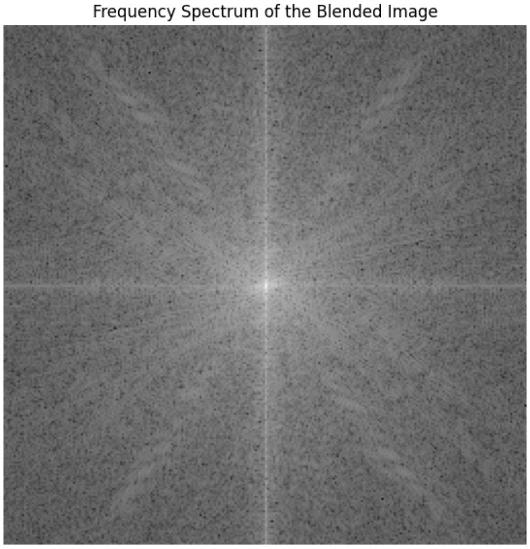
The frequency spectrum shows a central bright region, typical of low-frequency components, and numerous radial lines extending outward, indicative of edges or directional textures in the image.The radial lines suggest that there are features of various orientations in the image, which could correspond to textures or edges that are not aligned solely horizontally or vertically.The overall texture pattern in the spectrum suggests a complex image with varied spatial frequencies due to the blending of two different images with their own unique textures and edges.The spectrum's complexity reflects the composite nature of the blended image, which combines different spatial frequencies from the source images.

4. Butterfly Blending



The original signal graph shows intensity values with significant variation, including sharp spikes which are indicative of high contrast edges or points of rapid intensity change in the image.These sharp features in the signal are typical of distinct boundaries within the image, such as the edges of objects or transitions between different textures.The sampled signal graph indicates that the chosen sampling points capture the overall trend but miss the finer details present in the original signal. This is especially apparent in the regions between the sampling points where there are sharp changes in the original signal.The sampling rate appears to be low relative to the frequency of changes in the original signal, leading to a loss of information. This undersampling could result in aliasing where high-frequency components are misrepresented as lower frequencies.The sharp features in the original signal reflect the presence of distinct image features within the horizontal cross-section of the blended image.The sampled signal indicates that a higher sampling rate might be necessary to capture all relevant details accurately, especially considering the presence of sharp transitions in the original signal.

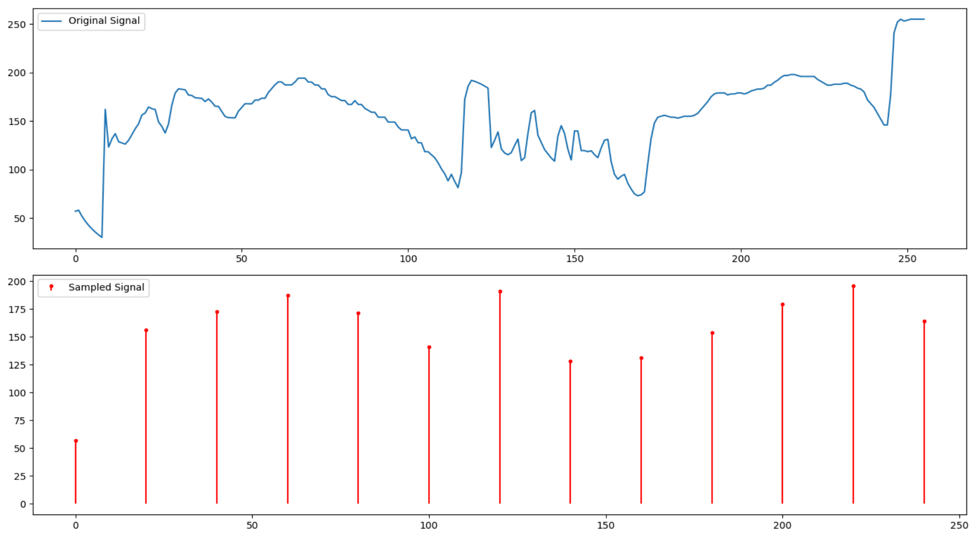
Frequency Spectrum of the Blended Image



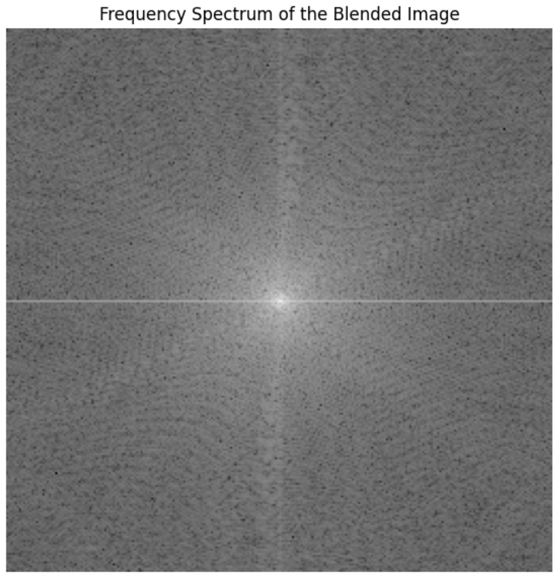
The frequency spectrum graph presents a bright center and faint radial lines extending from the center to the edges, suggesting the presence of a wide range of spatial frequencies with a dominance of low-frequency components.The symmetrical spread of the spectrum indicates that the image contains structures and textures in various orientations. The radial lines suggest the presence of periodic or repetitive structures within the blended image.The bright center tapering off into fainter radial lines suggests that the low-frequency components (broad structures and smooth transitions) are more prevalent than high-frequency components (fine details and sharp edges), but there is a balance between

.The dominance of low-frequency content and the presence of radial lines in the frequency spectrum indicates that the blending has been done in a way that preserves both broad features and finer textures from the original images.The balance of frequencies suggests that the blending is neither too smooth, which would obscure details, nor too rough, which would make the blend appear unnatural.

5. Flower Blending

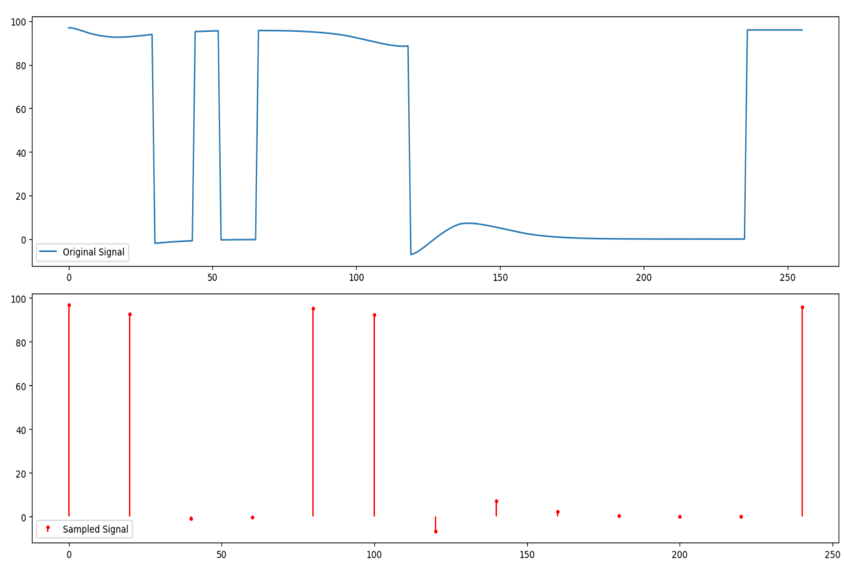


The original signal plot shows a range of intensity values, with several significant peaks that likely correspond to the bright features or edges of the flowers against a contrasting background.The gradual slopes between the peaks suggest areas of the image where the color or intensity changes more smoothly, possibly representing the petals or other less textured parts of the flowers.The sharp peaks followed by rapid declines indicate high-contrast areas within the image, which are typically the edges of the flowers or regions where there is a sudden change in color or brightness.The sampled signal captures the general trend of the original signal but with fewer data points. This reduction in data points means that some details of the original signal are not represented in the sampled signal.The peaks in the sampled signal correspond to the locations of the original signal that have been sampled. However, the finer details between these samples, such as smaller peaks and troughs, are not captured due to the sparsity of the sampling rate.This illustrates the concept of aliasing, where the true detail of the signal may be misrepresented due to insufficient sampling frequency.The original signal's variability and the sampled signal's sparsity underscore the importance of choosing a sampling rate that captures the important details in the image without losing significant information due to undersampling.



The frequency spectrum of the blended image showcases a bright central core, which is indicative of a strong low-frequency component. This suggests that there are large areas of uniform color or gradual changes in intensity in the blended image, which could be due to the smooth blending of the flower petals and background.The radial streaks emanating from the center to the edges of the spectrum are characteristic of images with edges at various orientations, aligning with the presence of petals and other structures within the flower images.The relatively homogenous distribution of the higher frequency signals (graininess of the image) indicates that the details of the flowers are preserved across the image. However, there are no strong directional frequencies, which suggests that the image does not have dominant directional textures.The frequency spectrum demonstrates that the blended image has both strong low-frequency components, representing the broad structures and smooth areas, and a wide distribution of high-frequency components, representing the detailed textures within the flower image.The results suggest a well-blended image that balances the preservation of details with smooth transitions, indicative of a successful blending technique that merges the source images into a cohesive composite.

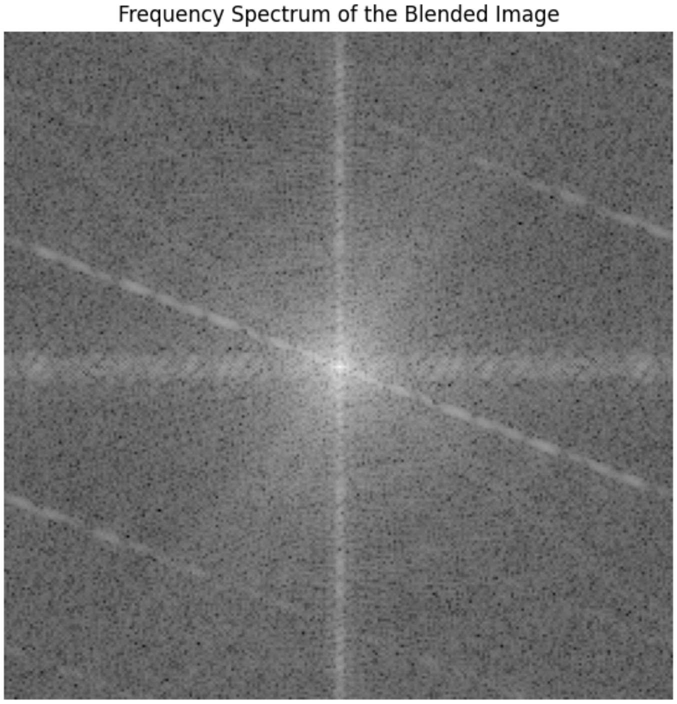
5. Character Blending



The original signal shows a staircase-like pattern, which suggests the presence of uniform intensity plateaus, possibly representing the flat areas of the characters or the background.The sharp vertical rises and drops indicate edges or transitions between different characters or between the characters and the background.The sampled signal captures the general levels of the original signal but misses the detailed transitions between the steps.

This is indicative of aliasing, where the nuances of the signal are not captured due to a sampling rate that is too low to accurately represent the sharp transitions.The tall stems correspond to the high points of the signal, which likely represent the edges of the characters, while the details within the plateaus are not represented.The original signal's stepped pattern suggests that the characters or symbols in the blended image have distinct, uniform intensity areas, which could indicate a successful blend where the individual characters maintain their distinct appearance.The underrepresentation of details in the sampled signal implies that a higher sampling rate would be necessary to capture the full detail present in the blended image. This highlights the importance of choosing an appropriate sampling rate to accurately represent the content of an image, especially one with sharp edges and flat regions like typographic characters or symbols.

Frequency Spectrum of the Blended Image



The frequency spectrum graph shows a bright central region, indicating the predominance of low-frequency components. These components are typical of smooth, uniform areas in the image, suggesting that the characters have been blended in such a way that they do not contain sharp or high-contrast transitions.The clear horizontal and vertical lines extending from the center of the spectrum suggest that there are strong, uniform edges in the image, which could be due to the edges of the characters against a contrasting background.The lack of distinct radial lines or directional textures implies that the characters are likely to be uniformly textured or colored, without complex patterns that would otherwise show up in the frequency spectrum.The frequency spectrum indicates that the character blending has resulted in a predominantly low-frequency image with distinct edges, as evidenced by the bright center and clear lines in the frequency spectrum.

References:

http://amroamroamro.github.io/mexopencv/opencv/pyramids_blending.html

https://www.youtube.com/watch?v=vJyttOxXyiM

https://towardsdatascience.com/blend-images-and-create-watermark-with-opencv-d24381b81bd0

https://docs.opencv.org/4.x/dc/dff/tutorial_py_pyramids.html

Pyramid Course Slides of BBM 413