Project Report: Interactive Foreground/Background Segmentation Using Lazy Snapping

By: Saad Rasheed

1. Introduction

Image segmentation, the process of partitioning an image into distinct regions, is a fundamental task in computer vision. In this project, we tackle the problem of interactive foreground segmentation, where the goal is to accurately separate the foreground and background regions of an image. By leveraging the concept of Lazy Snapping [1], an interactive segmentation approach, we aim to develop a solution that utilizes partial human annotations to guide the segmentation process. Our objective is to implement a basic version of the Lazy Snapping algorithm and evaluate its performance in achieving precise figure-ground segmentation. The report provides an overview of the project, details the methodology employed, and presents the experimental results and analysis.

2. Methodology

The implemented methodology follows the Lazy Snapping algorithm for interactive foreground segmentation. The process can be summarized as follows:

1. K-means Clustering: A k-means clustering algorithm is implemented to group color pixels into clusters. The algorithm assigns each pixel to a cluster based on the Euclidean distance in the RGB color space. The desired number of clusters, k, is provided as input, and the algorithm iteratively computes cluster centroids to minimize the sum of squared distances.
2. Seed Pixel Extraction: Foreground and background seed pixels are extracted from auxiliary images. The seed pixels are marked with red brushstrokes for the foreground and blue brushstrokes for the background. The k-means clustering algorithm is utilized to obtain N clusters for each class. N, representing the number of clusters, can be experimentally set (e.g., N=64).
3. Computing Pixel Likelihood: The likelihood of a given pixel belonging to each cluster of a specific class (foreground or background) is calculated. This is achieved using an exponential function of the negative Euclidean distance between the pixel and the

cluster center in the RGB color space. The overall likelihood of a pixel belonging to a class is computed as a weighted sum of all cluster likelihoods.

4. Assigning Pixels to Classes: Based on the calculated pixel likelihoods, each pixel is assigned to either the foreground or background class. If the likelihood of a pixel belonging to the foreground class is greater than the likelihood of it belonging to the background class, the pixel is assigned to the foreground class, and vice versa.

The implemented methodology combines these steps to perform interactive foreground segmentation using the Lazy Snapping approach. The next section will present the results and analysis obtained through experimentation with various test images.

3. Results and Analysis

The implemented Lazy Snapping algorithm was evaluated on a set of test images: Mona Lisa, Lady, and Van Gogh. For each test case, different values of k were used in the k-means clustering step to explore the impact on the segmentation results.

3.1. Test Case: Mona Lisa

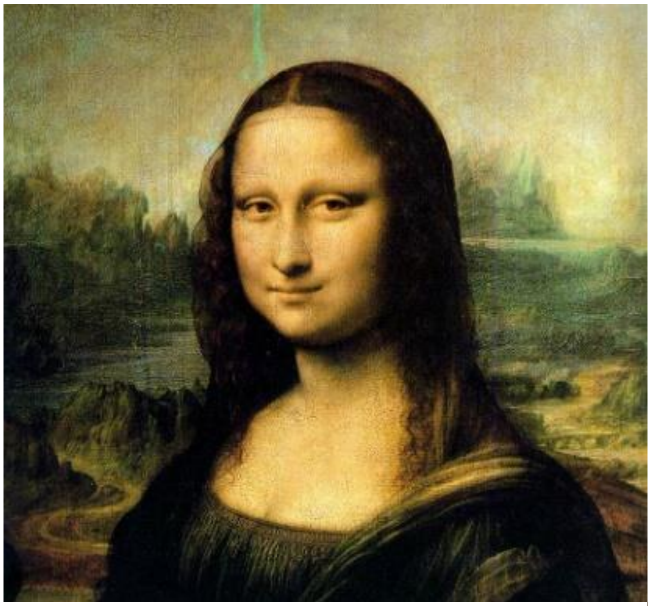
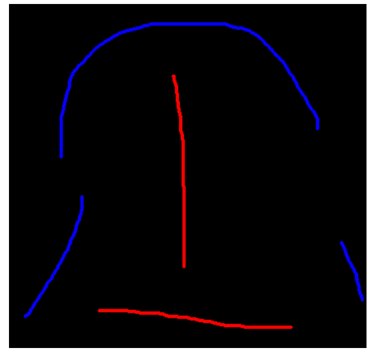


Figure 3.1(a): Mona Lisa

a. First Seed Segmentation

Figure 3.1(b): User Seed Segmentation for Mona Lisa



For the first test case of Mona Lisa, the segmentation was successful when k was set to 16 or 32. However, for higher values of k, the result was all background. The segmentation obtained with k = 16 or k = 32 identified both the face of Mona Lisa and the space above her head as background, likely due to similar color, illumination, and contrast levels.

Figure 3.1(c): k = 16

Figure 3.1(d): k = 32

b. Second Seed Segmentation

For the second test case of Mona Lisa, the algorithm could successfully segment Mona Lisa from the background for k values less than or equal to 64. Beyond that, the result was all background. The segmentation results were not satisfactory for k = 16, but improved for k = 32, k = 48, and k = 64, where the face and chest of Mona Lisa were identified as the foreground, while the rest of the area was classified as the background.

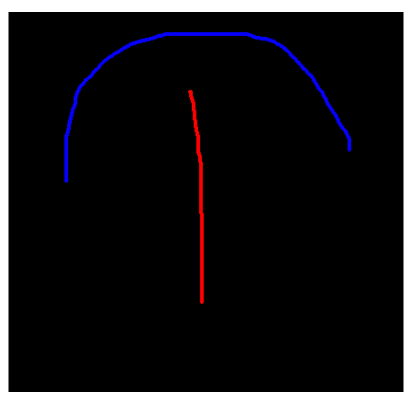


Figure 3.1(e): User Seed Segmentation for Mona Lisa (2)

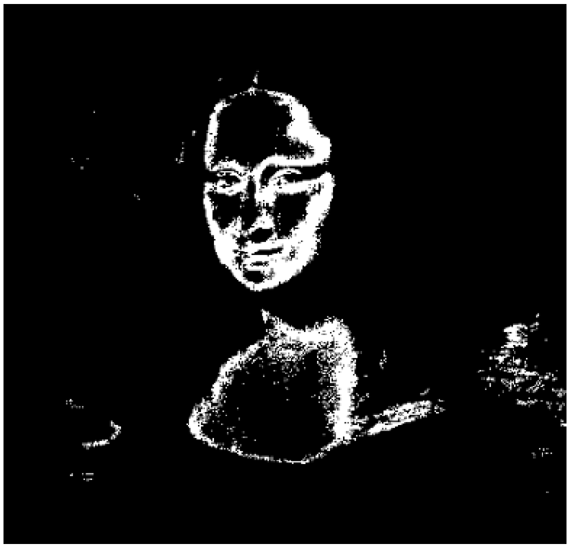


Figure 3.1(f): k = 32

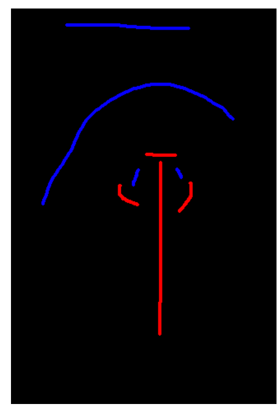
Figure 3.1(g): k = 48

Figure 3.1(h): k = 64

3.2. Test Case: Lady

Figure 3.2(a): Lady

Figure 3.2(b): User Seed Segmentation for Lady



a. First Seed Segmentation

For the first test case of Lady, the algorithm produced successful segmentation results only for k = 128, where the entire image was classified as the background. As the value of k decreased, the segmentation results improved.

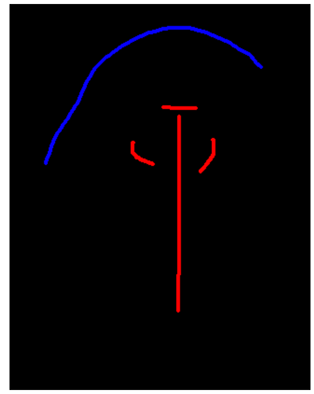
Figure 3.2(c): k = 64

Figure 3.2(d): k = 32

Figure 3.2(e): k = 16

b. Second Seed Segmentation

Figure 3.2(f): User Seed Segmentation for Lady (2)



Like the first test case, for the second test case of Lady, the algorithm generally produced satisfactory results. Only for k = 128 and k = 112, the entire image was classified as the background.

Figure 3.2(g): k = 64

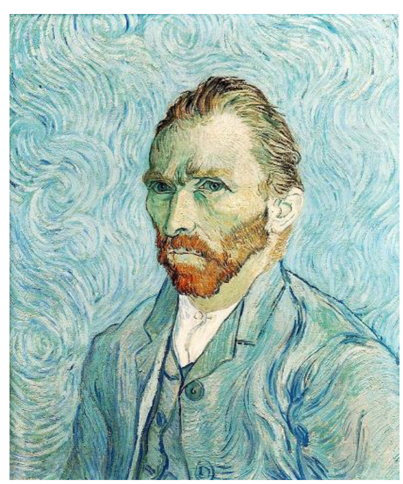
Figure 3.2(h): k = 32

Figure 3.2(i): k = 16

3.3. Test Case: Van Gogh

For the test case of Van Gogh, the segmentation results improved with decreasing values of k. After k = 64, all the resulting segmentations were classified as the background. However, for k values below 64, some parts of Van Gogh's face and his coat were correctly identified as the foreground, likely due to differences in color and illumination compared to the background.

Figure 3.3(a): Van Gogh



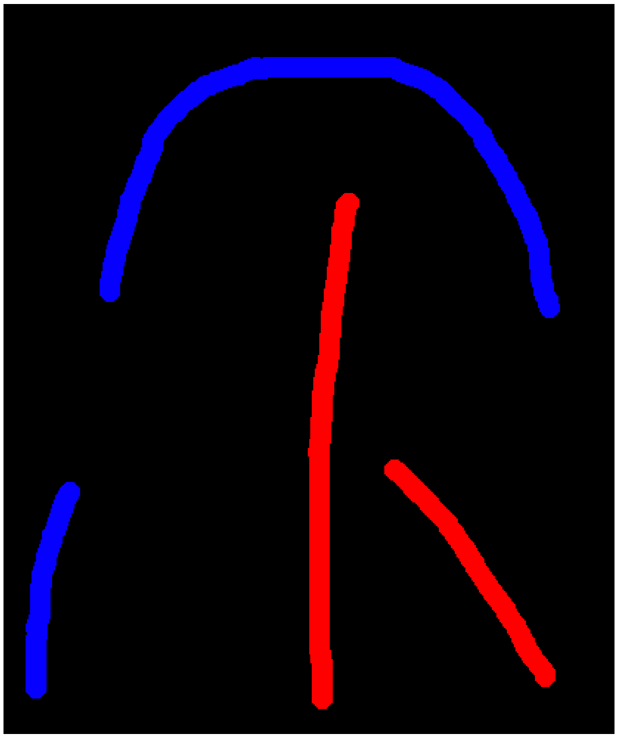


Figure 3.3(b): User Seed Segmentation

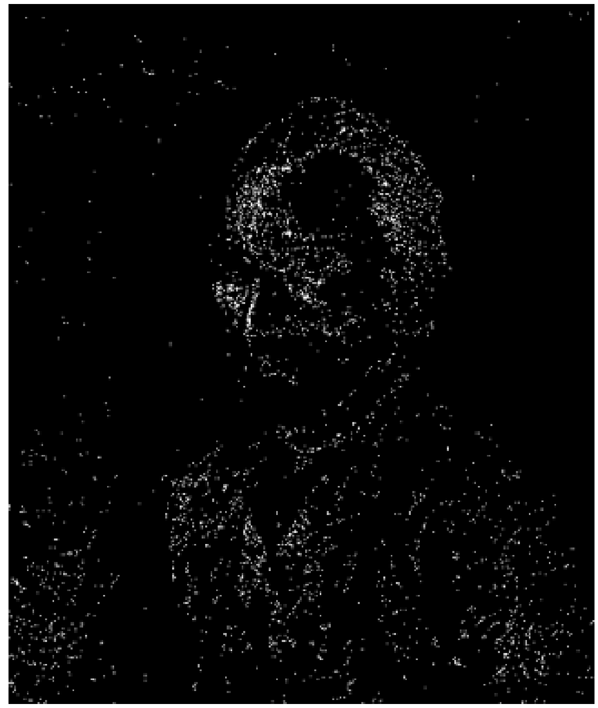


Figure 3.3(c): k = 64

Figure 3.3(d): k = 32

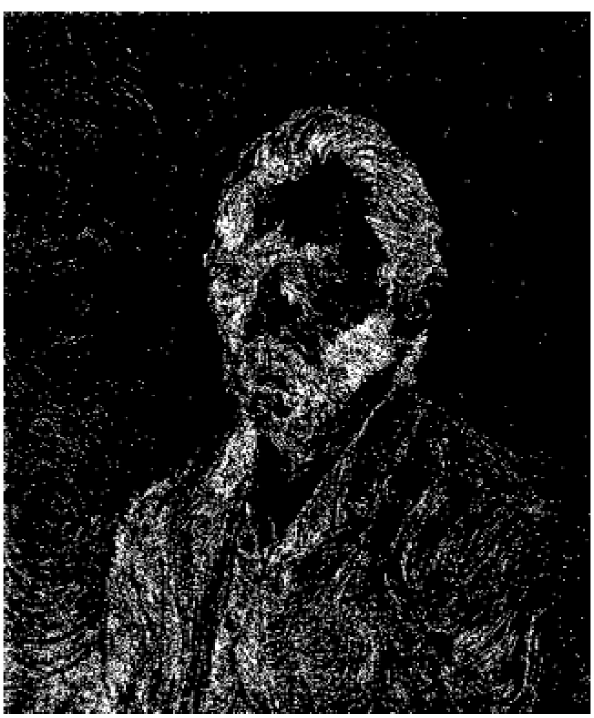
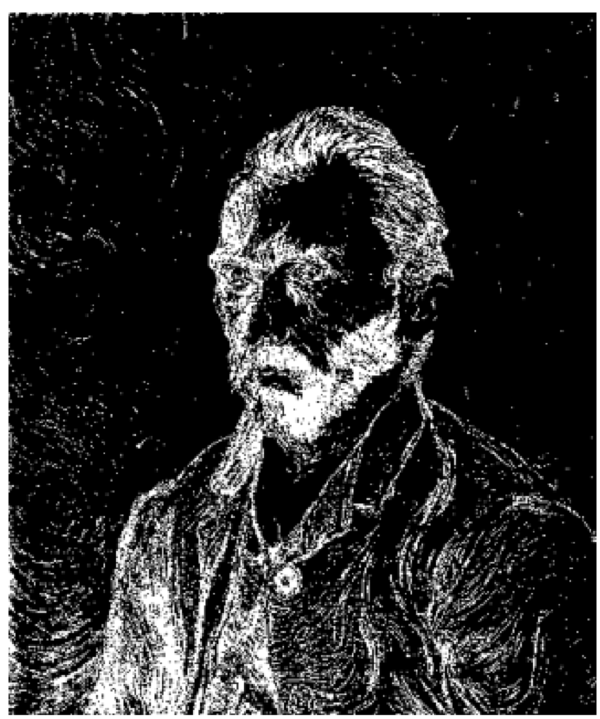


Figure 3.3(e): k = 16



4. Conclusion

In this project, we implemented a basic version of the Lazy Snapping algorithm for interactive foreground segmentation. By leveraging k-means clustering and pixel likelihood computation, we aimed to accurately separate foreground and background regions in test images. Through experimentation on test cases including Mona Lisa, Lady, and Van Gogh, we observed varying segmentation results for different values of k.

The results indicate that the performance of the algorithm is highly dependent on the choice of k and the characteristics of the input images. While satisfactory segmentations were achieved for certain test cases and k values, there were instances where the algorithm struggled to accurately distinguish between foreground and background regions, resulting in suboptimal segmentations.

To improve the performance of the algorithm, several avenues can be explored. First, refining the seed pixel extraction process by incorporating more robust techniques, such as region growing or user-guided selection, may enhance the accuracy of initial seed assignments. Additionally, exploring alternative distance metrics or incorporating feature extraction methods can improve the clustering process and the subsequent likelihood computation.

Furthermore, the use of advanced machine learning techniques, such as deep learning-based models, can potentially enhance the segmentation accuracy by capturing more intricate image features and relationships. These models can be trained on larger datasets with annotated foreground and background regions to learn more complex patterns and generalize better to diverse images.

In conclusion, while the implemented Lazy Snapping algorithm demonstrates the potential for interactive foreground segmentation, further research and development are necessary to enhance its performance and robustness. By incorporating advanced techniques and exploring the advancements in machine learning, more accurate and reliable segmentation results can be achieved, opening doors for broader applications in computer vision and image processing domains.

5. References

[1] Y. Li, J. Sun, C.-K. Tang, and H.-Y. Shum. Lazy snapping. In ACM SIGGRAPH 2004 Papers, pages 303–308, 2004.