Mean Shift Algorithm for Image Segmentation

Computer Vision

Abhimanyu Anand, i6332095

15 May 2023

1 Introduction

Image segmentation is a fundamental task in computer vision that involves dividing an image into distinct regions or segments based on their visual properties. The Mean Shift algorithm is a popular approach for image segmentation that offers effective clustering and grouping of pixels. It operates by iteratively shifting data points towards the centroid of any cluster of points.

The Mean Shift algorithm has gained significant attention in computer vision [1] due to its ability to handle complex image structures and adapt to varying data distributions. It does not require prior knowledge of the number of segments and can effectively capture non-linearly separable data. The algorithm's convergence property ensures locally optimal segments.

In this report, we provide an analysis of the Mean Shift algorithm for image segmentation. We discuss its theoretical foundations, including density estimation techniques and the iterative shifting process. We also explore optimization techniques to enhance the algorithm's efficiency, such as basin of attraction and distance-based association. Additionally, we present implementation details and evaluate the algorithm's performance in terms of speed and segmentation quality with different parameter settings.

Overall, this report serves as a comprehensive guide to understanding and implementing the Mean Shift algorithm for image segmentation, highlighting its advantages and potential applications.

2 Methodology

2.1 Find Peaks

The find peaks function is responsible for identifying the peaks in the arrays. The function takes as input the data points and a specified window size, and it iteratively computes the mean shift until convergence. It initializes a starting point and constructs a window around it. It then calculates the distances between the window points and the current point using euclidean distance. Points within the window that satisfy a distance threshold are considered part of the same cluster or peak. The function updates the current point by computing the mean of the points within the window, and the process continues until convergence.

2.2 Mean Shift

The Mean Shift algorithm is an iterative clustering technique that aims to discover clusters in a dataset by iteratively shifting the data points towards the centroid or peak of their local density. This iterative process involves calculating the mean of the data points within the window (radius), and updating the current point to the computed mean. This process continues until convergence is achieved. Find peaks function finds the peak for each point, which then combine with all the other points to create distinct number of clusters.

To improve the Mean Shift algorithm, several optimizations can be applied like the basin of attraction, using points of a particular search path, numpy accelaration.

2.3 Basin of Attraction

The basin of attraction is a region surrounding a peak where data points are likely to converge during the iterative process. By associating nearby points with a peak, the basin of attraction speeds up the algorithm by considering only relevant points within a certain distance, reducing the number of iterations required. This optimization improves the efficiency and accuracy of image segmentation, as it focuses on points that are more likely to belong to the same cluster, resulting in faster convergence and more accurate clustering results.

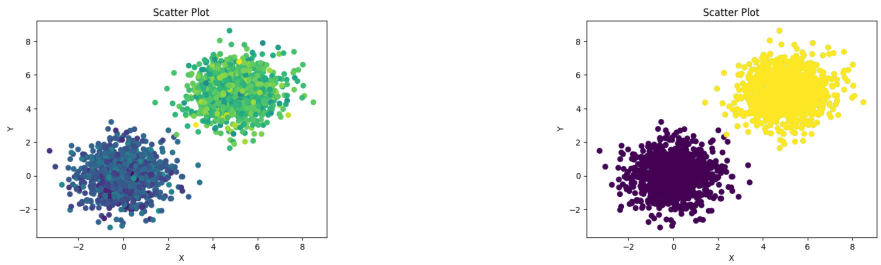
2.4 Search Path Optimisation(r/c)

Another optimization involves considering points within a certain fraction (r/c), where r is the radius and c is a constant, of the search path as part of the converged peak. This optimization helps in associating points that are close to the path but not necessarily within the window size, further improving the clustering accuracy.

By incorporating these optimizations, the Mean Shift algorithm became very fast. Without them, it took at least 7 hours per image segmentation, but after this, an image segmentation took a maximum of 15 minutes. Without any optimisation, the algorithm took more than 60 seconds for the given small sample data to test, but after basin of attraction test, it took less than 17 seconds to run even without any numpy accelarations.

2.5 Pre-processing - Image Segmentation

In the pre-processing step, we resized the image to a smaller size and converted it from the RGB color space to the CIELAB color space. Resizing the image helps improve computational efficiency without significant loss of quality, while



(a) Test Data

(b) Segmented Test Data

Figure 1: Test algorithm with sample data

the conversion to CIELAB enhances color representation for more accurate clustering. These pre-processing steps enable effective preparation of the image data for the subsequent Mean Shift clustering algorithm.

The converted image was then flattened and converted to a 3-by-p matrix of the image's colour information, or to a 5-by-p matrix of the image's colour and it's x,y coordinates information. This array was then given to the mean shift algorithm.

2.6 Post processing - Assign Segment Colours

After the mean shift algorithm, the resulting output consists of two arrays - the peak centroid values and the labels that assign each data point to a specific cluster or segment in the image. However, to visually represent and interpret the segmentation results, an additional step is done. By associating representative colors with each segment, the assign segment colors function enhances the visual representation of the segmented image. It effectively maps the numerical labels to corresponding color values. This post-processing step facilitates easier interpretation, analysis, and understanding of the segmentation output, enabling insights into the distinct regions or objects present in the image.

3 Results

The first test that was done was to establish that the find peak and mean shift algorithm work as expected. The sample test data given to us has 2 clusters as we can clearly see in figure 1. And the algorithm also gave accurate results in a very short time.

3.1 Experiment with r and c

The images used for testing mean shift algorithm for image segmentation are shown in figures 2, 3 and 4. All the following tests are done with image being resized to half the original image size for faster computations. The threshold

Figure 2: Big Ben

Figure 3: Lady on tree

Figure 4: Masked man

used in the find peaks function for combining peaks for the experiments has been used as 0.1 to make computations faster.

3.2 Experiment - 1: r=25, c=4

As seen in figure 5, the segmentation has worked very well and the 5 features with x and y coordinates has done a better job at segmenting the cloud and the shadows more clearly. For this size of image, r=25 and c=4 has worked very well.

3.3 Experiment - 2: r=25, c=1

In figure 6, it is seen that even with a reduction of the constant c, the segmentation has been very good, although there seem to be some artefacts in the segmented place where there should be water. The 5 features segmentation is more cleaner for this case.

3.4 Experiment - 3: r=40, c=4

In figure 7 the 3-by-p segmented more smoothly. Although there are not many artefacts in segmentation with 5 features, it still has two different segmentation for the same pants at the back, or different segments for the grass behind, etc.

3.5 Experiment 4: r=40, c=4, scale=1, threshold=0.01

For this experiment, I ran the masked man image without resizing the image, and with only 3 features, not 5. The threshold in find peaks function is also 0.01 here.

(a) 3 features

(b) 5 features

Figure 5: Big Ben with r=25, c=4

(a) 3 features

(b) 5 features

Figure 6: Experiment 2: Lady on tree with r=25, c=1

(a) 3 features

(b) 5 features

Figure 7: Experiment 3: Masked man with r=40, c=4

The result in figure 8 is clearly a very smooth segmented image without extra segments and clearly covering all different important features.

In all cases, the 3-by-p matrix with only colour data for segmentation has a more smoother segmented image while the 5-by-p matrix has more segments in the image with a more crisp output segmentation. The 5-by-p matrix also takes more than double the time for the same image and parameters.

4 Conclusion

The mean shift algorithm is a powerful clustering technique used for image segmentation. By iteratively shifting data points towards the local mean, it identifies dense regions and assigns points to clusters. The algorithm offers simplicity in implementation and flexibility. However, it can be computationally intensive and its performance depends on the choice of window radius and convergence criteria. By applying optimizations and post-processing steps, the mean shift algorithm produces accurate and visually appealing segmentation results, making it valuable in computer vision and image processing applications.

References

[1] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 24(5):603–619, 2002.

Figure 8: Image segmentation without resizing