

## A MODIFIED MEAN-SHIFT TECHNIQUE FOR SEGMENTATION OF HIGH RESOLUTION SATELLITE IMAGES

Biplab Banerjee<sup>a</sup>, Krishna Mohan Buddhiraju<sup>b</sup> and Surender Varma Gadhira<sup>c</sup>

<sup>a,c</sup> Doctoral Student, Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay  
CSRE, IIT Bombay, Powai, Mumbai – 400076, INDIA

Tel: +91-(022)-2576-4676

Email: [biplab.banerjee/gsvarma@iitb.ac.in](mailto:biplab.banerjee/gsvarma@iitb.ac.in)

<sup>b</sup> Associate Professor, Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay

Tel: +91-(022)-2576-7684

Email: [bkmohan@csre.iitb.ac.in](mailto:bkmohan@csre.iitb.ac.in)

**KEY WORDS:** Mean-Shift, Satellite image segmentation, clustering.

**Abstract:** Image segmentation is a process of sub-dividing a given image into its constituent objects. Segmentation algorithms are inherently unsupervised as there is no available a priori knowledge regarding the approximate number of objects actually present in the image. In this respect, unsupervised clustering techniques are of particular importance. This paper focuses on a particular type of spatial clustering method that is a variant of the well-known method called the mean shift. An adaptive version of the original mean shift is proposed to tune the hidden parameters. The width of the Parzen window is one such variable that controls the degree of over- or under-clustering to a great extent. A density based approach is employed to estimate the optimal window width. A new termination criterion is also developed to control the segmentation. This termination criterion is based on the statistical framework incorporating the variance measure across the windows. We have applied the proposed clustering for satellite image segmentation purpose.

Experimental results showed that this method has performed very well in segmenting satellite images and is comparable to other state of the art clustering based segmentation techniques. It is being extended by embedding the segmentation algorithm as one of the modules in a fully-fledged object based image classification system including pre-processing, segmentation, connected component labeling, feature computation, and classification.

### 1. INTRODUCTION

Image segmentation refers to the process of partitioning an image into groups of pixels where pixels belonging to each group are homogeneous with respect to some measures. The result of segmentation is the division of the image into a set of connected components. Each component represents some region of interest.

Until now, most of the satellite image segmentation methods have used the underlying statistics in the image space in the parametric form. Some of them are supervised and need prior information to achieve a successful segmentation result. But sometimes the required ground truths are not available beforehand. The other classes of segmentation methods are unsupervised in nature which employs watershed transformation combined with region based or clustering based segmentation approaches [1]. The main problem with the parametric methods is that they are not robust as the output gets severely degraded if the model parameters are not tuned properly.

Some segmentation methods based on the spectral signatures of the image such as thresholding and histogram-based finite mixture models are well-studied in the literature. However they usually fail to produce a moderate segmentation result for images with low contrast or noisy images with varying backgrounds. It is noted that these methods don't use the spatial morphological images information [6]. On the other hand, some other methods such as morphological segmentation, region growing and deformable curves, mainly focus on spatial information such as local structures or regions. Unfortunately, the majority of these techniques are not suitable for satellite image segmentation since such type of image presents a non-homogenous texture [7]. In [8] a simulated annealing (SA) based fuzzy clustering method is developed and combined with popular support vector machine (SVM) classifier to fine tune the clustering produced by SA for obtaining an improved clustering performance.

Conventional clustering techniques have been used successfully in the image segmentation literatures. Strategies like k-means, fuzzy c-means, density based clustering [2][3] have already been exercised for this purpose. But most of these approaches are parametric and require the approximate initial number of clusters to proceed further. Mean-

shift clustering [4] has been explored in recent literature as a promising image segmentation method [5]. It does not require any initial cluster count and is based on the concept of kernel density estimation.

In this paper, two modifications to the original mean-shift clustering have been proposed. A  $k$ -dist based method for Parzen window width estimation technique has been developed here.  $k$ -dist calculates for each point its distance to the  $k^{\text{th}}$  neighbor. The distance measure at the sharpest transition point of the distance plot is usually considered as the window size for density based clustering algorithms. The terminating criterion proposed here guarantees fast convergence of the algorithm.

## 2. MEAN-SHIFT ALGORITHM AND THE MODIFICATIONS PROPOSED

This section provides an intuitive idea of Mean-Shift and the proposed modifications. It is a powerful and versatile non parametric iterative algorithm that can be used for purposes like finding modes, clustering etc. It has also been used to analyze satellite remotely sensed images. It has been used to extract objects like roads, water bodies etc [5] from satellite images of various resolutions.

Density estimation refers to the process of describing the underlying probability distribution function of a data set. It can be parametric or non-parametric. In the first case, some assumptions regarding the data points are used to build the distribution function whereas in the latter case, no prior information regarding the data points is known. Histogram based density estimator has been used effectively in many statistical applications but it has some drawbacks [8]. Though it is considered as a non-parametric estimator, the bin width and the starting position of the bin are not fixed. These disadvantages are overcome in the case of kernel density estimator.

Given a set of  $n$  data points  $x_i, i=1,2,3, \dots, n$  in the  $d$  dimensional feature space  $R^d$ , the kernel density estimator at a given location point  $x$  can be calculated as

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right)$$

Where  $K$  is the kernel which is symmetric but not necessarily positive function that integrates to one — and  $h > 0$  is the bandwidth parameter. A kernel with subscript  $h$  is called the scaled kernel and defined as  $K_h(x) = 1/h K(x/h)$ . Intuitively one wants to choose  $h$  as small as the data allow however there is always a trade-off between the bias of the estimator and its variance.

Mean shift treats the points the feature space as a probability density function. Dense regions in feature space correspond to local maxima or modes. So for each data point, we perform gradient ascent on the local estimated density until convergence. The stationary points obtained via gradient ascent represent the modes of the density function. All points associated with the same stationary point belong to the same cluster.

Assuming  $g(x) = -K'(x)$  and considering the gradient ascent property of mean-shift, we have

$$m(x) = \frac{\sum_{i=1}^n g\left(\frac{x - x_i}{h}\right) x_i}{\sum_{i=1}^n g\left(\frac{x - x_i}{h}\right)} - x$$

The quantity  $m(x)$  is called the mean-shift vector. So mean shift procedure can be summarized as,

For each point  $x_i$ ,

1. Compute mean shift vector  $m(x_i')$ .
2. Move the density estimation window by  $m(x_i')$
3. Repeat till convergence.

The mean-shift algorithm is a powerful technique for image segmentation. The algorithm recursively moves to the kernel smoothed centroid for every data point. Mean-Shift based image segmentation is achieved using the following steps,

- Use kernel density estimator to shift the means of pixels in the image.
- Stop when each mean sequence has converged.
- Use the converged means to delineate segments.
- Basically a pair of pixels belong to segment if their
- Convergence color and spatial components are within a threshold.
- For segments with fewer pixels than a given threshold, place them into neighboring segments.

Now the proposed modified adaptive mean-shift algorithm is highlighted. The width or radius of the Parzen window is determined using the well-known  $k$ -dist method, mainly used in DBSCAN. Given a set of  $n$  points  $x_i \in R^d$ ,  $i=1,2, n$  the  $k$ -dist method works as follows

- Select  $k$ , which defines a particular neighbor of a data point ( $k^{\text{th}}$  neighbor).
- For each data point, find the distance of its  $k^{\text{th}}$  neighbor.
- Sort those distances.
- Find the point in the distance curve where the sharpest transition occurs.
- The corresponding distance value  $k_{\text{dist}}$  is recorded.

We have extended this concept and used several values of  $k$  within a given range. The range is purely based on the statistics of the data space like number of points etc. The particular value of  $k$  for which this largest transition in the distance curve is achieved is used as the threshold for minimum number of points needed within the window for further propagation. The  $k_{\text{dist}}$  associated with that particular  $k$  is used as the width of the Parzen window.

The issue regarding the convergence is taken care of now. Starting from any randomly selected point  $x$ , according to the mean-shift concept, the window moves towards the direction of the dense most local region in the data space. Along with the window size, the termination condition is also one of the major factors to be considered to guarantee that the algorithm does not fall into local optima. The original mean-shift uses a threshold and check whether the difference between the means of the data points within two consecutive windows is below or above the threshold. The process is terminated if the difference is above the threshold. We have automated this check by proposing a termination condition.

Let  $x_1, x_2, \dots, x_n$  are the points within the window at a given time at iteration  $i$ . Then after moving to the mean (at iteration  $i+1$ ), the new set of points within the window are  $y_1, y_2, \dots, y_m$  in the  $d$  dimensional feature space. At iteration  $i$ , let us consider  $v_i$  denotes the trace of the covariance matrix of the data points within the corresponding Parzen window. Let us also consider that  $v_{i+1}$  denotes the trace of the covariance matrix of the data points within the window in the next iteration, i.e. iteration  $i+1$ . Now, termination condition is defined as

$$|v_{i+1}-v_i| \times \text{count}(x_i \cap y_j), \{i \in n, j \in m\}$$

Data points belonging to the same class have less variance than the data points belonging to several different classes. Hence, with the movement of the window, once the termination condition attains the minimum value, it suggests that we are at the mode for that single cluster. So, data points belonging to all the windows in this iteration form a single cluster. Now for a new iteration, if there is any intersection of points with previously labeled points, we just merge those two clusters. That is how clusters of any shape and sizes are obtained using this adaptive mean-shift clustering. One issue to be discussed here is that, the above equation, we have also considered the number of common points in two consecutive windows. Now the termination condition ensures that the overall condition term will attain the minimum possible value only at the midway of the window movement path. This process makes the detection of varied shaped clusters more effective. To avoid misclassification, a condition is checked at iteration  $(i+1)$  which suggests the purity of the cluster. In the  $(i+1)^{\text{th}}$  iteration, the points within the window should satisfy

$$\frac{\text{mean}_{i+1}}{\text{variance}_{i+1} + 1} \geq t$$

$\text{mean}_k^i$  and  $\text{variance}_k^i$  define the mean and variance respectively along the  $k^{\text{th}}$  feature dimension of the data points belonging to window  $i$ . We are interested only in the minimum value of the ratio along any dimension. The constraint has been thresholded by experimenting on a number of synthetic and real world data sets.  $t$  is a user defined threshold. The value of  $t$  is dependent primarily on the variance of the data items. It has been observed that for a set of data items belonging to the same cluster, the variance is relatively low than that of data items belonging to several different clusters. In our case, we have fixed the value of  $t$  to 0.18.

### 3. RESULTS

This section presents results of experiments performed based on the proposed algorithms. A set of multi-spectral images have been considered for this purpose. Figure (1) denotes a true color composite of the first experimental image used here. It is a high resolution QuickBird satellite image of Mumbai, India area with 3 bands depicting red, green and blue components. The area is an urban fringe consisting of a few buildings, a quarry site, ponds, road, vegetation, open ground and foot-paths. Figure (2) shows the segmentation output produced after the proposed MS clustering algorithm.



Figure 1: The QuickBird input image



Figure 2: Segmentation result of image in figure 1

The second study area considered here is depicted in Figure (3). It is a true color composite Ikonos satellite image with 4m spatial resolution. Broadly this image contains fields and greenery, river, residential areas with different spectral and textural properties. The result of the image segmented by the proposed MS algorithm has been depicted in Figure (4).



Figure 3: Ikonos image



Figure 4: Segmented result of image in Figure 3

The third study area considered here is a high resolution aerial image acquired by the Geo-Eye sensor. Apparently it is the image with a large building with surrounding road and greenery. Figure (5) shows the image in true color composite and the segmentation result is depicted in Figure (6).



Figure 5 The Geo-Eye image



Figure 6 The segmented output of image in Figure 5

#### 4. DISCUSSION

The results depicted above indicate that the modified MS algorithm has comparable accuracy with the standard conventional MS clustering. The results of some well-known cluster validity indices establish this fact. The Davies-Bouldin (DB), Silhouette indices have been used here in this case. Small values of Davies-Bouldin index suggests good clustering whereas Silhouette index obtains larger value in the optimal scenario. The clustering results have also been validated with the well-known K-means, Fuzzy c-means algorithms. The initial estimates of clusters have been considered with visual inspection for those cases. Figure (7) depicts the results of the DB index for the proposed method and the original MS based clustering for above images.

	Proposed modified MS	Original MS
Image 1 (Figure 1)	0.3145	0.3210
Image 2 (Figure 3)	0.4216	0.4187
Image3 (Figure 3)	0.3614	0.3859

Figure 7: The DB index for the proposed method and original MS based clustering

$k$ -dist method is pre-computed hence it does not have any effect on the time complexity of the proposed clustering. Simulation for a number of  $k$  helps in finding the average width of the clusters present. Hence, determination of the near-optimal window width becomes easy.

#### 5. CONCLUSIONS & RECOMMENDATIONS

A modified mean-shift clustering for satellite image segmentation has been proposed here. First, a pre-processing step has been proposed to determine the optimal Parzen window width. An iterative  $k$ -dist based method has been considered for this purpose. Then a terminating criterion has been proposed which is based solely in the underlying data distribution. Experiments on several multi-spectral images have established the efficiency of the proposed method. It is a step for the object based satellite image analysis model and work is being carried out to perform object based high level analysis of satellite images based on the proposed segmentation.

#### 6. REFERENCES:

Q. Chen, C. Zhou, J. Luo et. al., "Fast Segmentation of High-Resolution Satellite Images Using Watershed Transform Combined with an Efficient Region Merging Approach", Lecture Notes Comput. Sci., vol 33, no 22, pp. 621-630,2004.

G. Castilla, G. J. Hay, and J. R. Ruiz-Gallardo, "Size-constrained region merging (SCRM): an automated delineation tool for assisted photointerpretation", Photogrammetric Engineering and Remote Sensing, vol. 74, no. 4, pp. 409-419, 2008.

J. Yang, Shu. Hao, Pau. Chung. "Color image segmentation using fuzzy C-means and eigenspace projections", Signal Processing vol. 82 no. 3, pp. 461-472, 2002.

D. Comaniciu, P. Meer, "Mean-shift: A Robust Approach Toward Feature Space Analysis", IEEE Trans. Pattern Anal. Mach. Intell. vol. 24 no. 5, pp. 603-619, 2002.

W. Tao, H. Jin, Y. Zhang, "Color Image Segmentation Based on Mean Shift and Normalized Cuts", IEEE Transactions on Systems, Man, and Cybernetics, Part B vol. 37 no. 5, pp. 1382-1389, 2007.

Georgy Kukharev, Adam Nowosielski: Visitor Identification - Elaborating Real Time Face Recognition System. WSCG (Short Papers) 2004: 157-164.

J. Xie, and H. T. Tsui, "Image segmentation based on maximumlikelihood estimation and optimum entropy-distribution(MLE-OED)", Pattern Recognition Letters, Volume 25, Issue 10 , pp. 1133-1141, July 2004.

A. Mukherjee and U. Maulik, Unsupervised Satellite Image Segmentation by Combining SA Based Fuzzy Clustering with Support Vector Machine, International Conference on Advances on Pattern Recognition, PP. 381-384, 2009.

[www.geoeye.com](http://www.geoeye.com)

[www.satimagingcorp.com](http://www.satimagingcorp.com)