Unsupervised Multi-Spectral Satellite Image Segmentation Combining Modified Mean-Shift and a New Minimum Spanning Tree Based Clustering Technique

Biplab Banerjee, Surender Varma, Krishna Mohan Buddhiraju, and Laxmi Narayana Eeti

***Abstract*—An unsupervised object based segmentation, combining a modified mean-shift (MS) and a novel minimum spanning tree (MST) based clustering approach of remotely sensed satellite images has been proposed in this correspondence. The image is first pre-processed by a modified version of the standard MS based segmentation which preserves the desirable discontinuities present in the image and guarantees oversegmentation in the output. A nearest neighbor based method for estimating the bandwidth of the kernel density estimator (KDE) and a novel termination condition have been incorporated into the standard MS. Considering the segmented regions as nodes in a low level feature space, an MST is constructed. An unsupervised technique to cluster a given MST has also been devised here. This type of hybrid segmentation technique which clusters the regions instead of image pixels reduces greatly the sensitivity to noise and enhances the overall segmentation performance. The superiority of the proposed method has been experimented on a large set of multi-spectral images and compared with some well-known hybrid segmentation models.**

***Index Terms*—Graph based clustering, image segmentation, mean-shift, minimum spanning tree.**

I. Introduction

IMAGE segmentation is one of the basic and foremost steps required for high level image understanding. It bridges the semantic gap between low level image processing and high level intelligent image analysis. Object recognition and tracking [1], land-cover, land-use classification [2] etc. are some of the applications where segmentation is one of the integral parts.

Data clustering [3] is one of the well-established similarity based methods which has widely been applied in the domain of image segmentation. The main concept of clustering based segmentation is to group image features into different clusters so that the intra cluster variations are minimized and inter cluster variations are maximized to the possible extent.

Several publications report the successful application of various conventional clustering techniques (k-means, fuzzy c-means, density based clustering [4]) for segmentation of satellite images [5], [6]. But most of them are parametric and require the approximate initial number of clusters to proceed further. Mean-shift [7] clustering technique has been explored in recent literature as a promising image segmentation technique [8].

Although the clustering based segmentation approaches are efficient in finding underlying image features, they impose some serious drawbacks too. The spatial structure and the edge information of the image are not preserved and pixels from different regions are difficult to distinguish in case of overlapping feature domains. Spatial segmentation techniques have been explored in recent literature as an alternative segmentation strategy to preserve image discontinuity and spatial relationship between pixels. But the main disadvantage with these algorithms is that they undesirably produce a large number of small quasi-homogeneous regions, i.e. as in the case of watershed transformation. Hence a proper merging technique is needed afterward. [9] has proposed an object based natural image segmentation methodology using super-pixels as a grouping cue. The super-pixels are fused in a bipartite graph partitioning framework. [10] has developed a hybrid segmentation method using watershed transform followed by a modified MS based merging. The framework proposed in [10] has been updated here in this paper in an effective fashion. The image is first oversegmented by the proposed MS based clustering followed by a merging strategy using a new MST based clustering algorithm.

MS algorithm is usually applied to perform discontinuity preserving smoothing followed by image segmentation. Due to its edge preserving filtering property the salient features of the overall image are retained. This property is important for segmenting remotely sensed images in which several distinct regions are used to represent the whole scene. However, it is difficult to partition a remotely sensed image into different land-covers solely based on the MS algorithm as MS is an unsupervised technique where the number and shape of the data clusters are unknown *a priori* though comparatively smaller values of the bandwidth parameter used for the KDE and the smallest region size (M_{min}) allowed can assure oversegmentation in the output. Hence, MS based method is a good choice to perform the initial oversegmentation.

Manuscript received December 31, 2012; revised April 02, 2013 and May 06, 2013; accepted May 29, 2013. Date of publication July 01, 2013; date of current version March 14, 2014.

B. Banerjee is with the Indian Institute of Technology (IIT) Bombay, CSRE, Powai, Mumbai 400076, India.

S. Varma is with the Indian Institute of Technology (IIT) Bombay, CSRE, Powai, Mumbai, Maharashtra 400076, India.

K. M. Buddhiraju is with the Indian Institute of Technology (IIT) Bombay, Centre of Studies in Resources Engineering, Mumbai, Maharashtra 400076, India.

L. N. Eeti is with the Indian Institute of Technology (IIT) Bombay, CSRE, Mumbai, Maharashtra, India.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2013.2266572

Typical graph based segmentation methods are difficult to apply directly to image pixels, as the storage and computational costs become huge. Hence, for hybrid segmentation algorithms, graph based clustering is the usual choice for the final region merging purpose where the nodes of the graph represent the objects extracted from the initially oversegmented image (produced by some other algorithms like MS based segmentation) which are far less in number than the pixels present in the original image. Algorithms like C-means or Fuzzy c-means are generally not encouraged to be used for this final merging because they tend to produce hyperspherical clusters and they have the inherent parametric nature.

This paper proposes a satellite image segmentation method incorporating two of the clustering approaches mentioned above, i.e., mean-shift and minimum spanning tree based clustering. The novelties of the proposed work are as follows,

- A modification to the conventional mean-shift clustering has been proposed. A k-dist based method for Parzen window width estimation technique has been developed here. For each point, k-dist calculates, its distance to the kth nearest 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 [10].
- A non-parametric clustering algorithm using MST has also been proposed. The tree is divided hierarchically and a single cluster is generated in each iteration. A separation predicate has been proposed which guides the clustering process.

The paper is structured as follows. In Section II, an overview of the proposed segmentation algorithm is presented. The modifications proposed for the standard MS based clustering are mentioned in Section III. Feature extraction from the segmented objects is dealt with in this section. Section IV details the proposed MST based clustering technique. Experimental results are discussed in Section V. Section VI concludes the article along with the possible future extension of the proposed segmentation framework.

II. The Proposed Image Segmentation Algorithm

Fig. 1 depicts the flowchart of the proposed segmentation method. Broadly, it is a three-step process.

- Apply the proposed modified MS based clustering algorithm to cluster the pixels of the input satellite image in the spectral domain.
- A feature extraction step is carried out to calculate some low level features from each of the objects found in the previous state.
- Considering each region as a node in the newly defined feature space, the proposed MST based non-parametric clustering method is applied for final region merging.

The steps are described in detail in subsequent sections.

III. Modifications Proposed to the Mean-Shift Technique

The mathematical details of the standard mean-shift technique can be obtained in [7]. Though mean-shift does not require the initial number of clusters actually present in the data space, the width of the Parzen window is to be mentioned beforehand. If the window size is larger, there is possibility of overmerging whereas smaller window leads to the generation of several small sub-clusters in the output. Here a k-dist based method for estimating the (near)-optimal bandwidth has been proposed. The traditional mean-shift process stops when the density gradient attains a value close to zero. It is an iterative process and may generate unnecessary larger number of clusters. Here an adaptive termination criterion has been developed which has been demonstrated to accelerate the clustering process. This terminating condition detects the sub-clusters present in a given cluster first and then a postprocessing step combines them together into a single group.

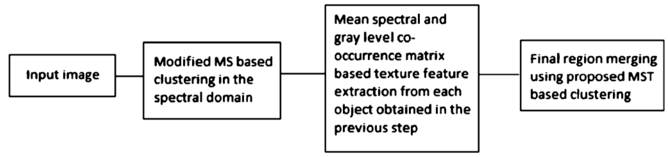

Fig. 1. Flowchart of the proposed segmentation algorithm.

A. *k-Dist Based Bandwidth Selection*

Nearest Neighbor (NN) based bandwidth estimation [11] is well known in the clustering literature. Here, an extension of the popular k-dist technique (used mainly along with density based clustering like DBSCAN [12]) has been carried out to estimate the bandwidth parameter (h). Algorithm 1 presents the scheme.

Algorithm 1 Calculate h given $\{x_i\}_{i=1}^{n} \epsilon R^d$, $\forall k \epsilon [k_{min}, k_{max}]$

for $k = k_{min}$ to k_{max} **do**
  for $i = 1$ to n **do**
    $dist[i]$=$KNN_{dist}(x_i)$. $\{KNN_{dist}(x)$ calculates the distance between x and its k^{th} NN using kd tree [13].$\}$
  end for
  Plot $dist$ after sorting it in ascending order.
  Find the distance measure corresponding to the sharpest transition (k_{dist}) found in the plot and store it.
end for
Calculate h using Equation (1).

The k_{dist} corresponding to the particular k for which the **smallest of the largest** transitions in the $dist$ plot is obtained, is considered as h:

$$h = min\left\{k_{dist}^{i}\right\}, \ \forall i \epsilon \left[2, \frac{n}{4}\right] \tag{1}$$

A bandwidth is favorable if a high proportion of data points falls within circles of radius h centered at a certain given object. The mode seeking process becomes faster if the points within the window at a time belong to the same class. To ensure this, the average cluster width needs to be approximated. Hence given a point x_i and M_{min}, it has been assumed that any of its kth neighbor in the above mentioned range will eventually cross the boundary of the class where x_i is likely to be present. The range is empirically set and exhibits good performance for remotely sensed images with medium and high spatial resolutions.

B. Proposed Termination Condition

Starting from any randomly selected point x, according to the mean-shift concept, the window moves towards the direction of the mostly populated 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 traditional MS clustering method stops when the density gradient attains a value close to 0. The proposed criterion instead combines the local modes to generate a global uniform smooth density estimator. The proposed technique is far more useful when a cluster is sparsely scattered in the feature space.

Let $s_1, s_2, \ldots, s_n$ be 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 $t_1, t_2, \ldots, t_m$. 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, the termination condition is defined as

$$|v_{i+1} - v_i| \times count\left(\{s_k\} \cap \{t_j\}\right)_{k \epsilon n, j \epsilon m} \tag{2}$$

Data points belonging to the same class have less variance than the data points belonging to different classes. Hence, with the movement of the window, once the termination condition of (2) attains the minimum value, it suggests that the local mode for that cluster is reached. So data points belonging to all the windows in this iteration form a single sub-cluster. Now for another iteration, if there is any intersection of points with previously labeled points, those two clusters are merged. That is how clusters of any shape and size are obtained using this adaptive mean-shift clustering. One issue to be addressed here is that, in (2), the number of common points in two consecutive windows has also been considered. The termination function ensures that the overall condition term will attain the minimum possible value only at the midway of the window movement path. These local modes (detected sub-clusters) can be grouped to obtain a single larger cluster.

C. Feature Extraction From the Segmented Objects

The objects found after this step undergo a feature extraction stage. Mean spectral value of the region and some statistical texture features have been considered for this purpose. Shape feature has not been considered here as a given land use class may have different size and shape at different geographical location.

Gray level co-occurrence matrix (GLCM) [14] based texture features have been considered as these kinds of features can classify images with micro-textural elements. This concept is key in classifying different land cover classes of remotely sensed satellite images. GLCM is defined as the distribution of the co-occurring values within a given offset in the image plane. It represents the distance and angular spatial relationship over an image sub-region of specific size. The GLCM used here is direction invariant. Hence the average of all four spatial arrangements depicting 0°, 45°, 90° and 135° has been used. A pixel offset of 1 and a quantization level of 256 have also been considered. From the co-occurrence matrix, many texture parameters [14] can be calculated. Entropy, energy, homogeneity and contrast have been selected for this case to capture the relative gray level randomness and brightness of a given region. Hence for a p-band image, the regions found after the previous step are represented in a $(p+4)$-dimensional feature space which are further clustered using the proposed MST based clustering.

IV. Proposed MST Based Clustering

This section describes the proposed data clustering technique using minimum spanning tree. MST based clustering [15] technique is known to be capable of detecting clusters with irregular boundaries.

A. Proposed MST Based Clustering Algorithm

Given a set of points $x_1, \ldots, x_N (x_i \epsilon R^{(p+4)})$ with unknown underlying probability distribution, a graph $G(V, E)$ is constructed with vertex set V and edge set E where each of the x_i is represented by a node in the graph. The edge between a given pair of vertices x_i and x_j is weighted (w) by the Euclidean distance:

$$w(x_i, x_j) = 1 - \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{3}$$

σ has been fixed by grid search [16] technique using 50% of the total data for cross-validation. The value of w will be smaller if x_i and x_j are similar to each other in some sense whereas w will get increased in case the data points differ. A minimum spanning tree $T(V, E_T)$ of G is constructed henceforth using the Kruskal's approach [17].

The proposed clustering method checks iteratively whether a given edge of the tree can be deleted to form two compact sub-clusters. The edges are considered in the descending order of the corresponding edge weights because it has been assumed that edges with larger weights are the ones which span different clusters. Algorithm 2 describes the proposed clustering process.

A is a measure of the cluster compactness. If a given edge $E_{W(i)}$ has drastically larger weight value than the component edge weights of its two adjoining connected components (CC's), then it can be presumed that $E_{W(i)}$ spans two different clusters and hence can be deleted. The parameter α (Algorithm 3) of (4) defines a tolerance per CC. Given a CC, the corresponding α controls the gap allowed in the feature space between this CC and any other CC.

B. Efficiency Assessment of the MST Based Clustering

The goodness of the proposed clustering has been measured in terms of the Davies-Bouldin (DB) index and Silhouette (Sil) index. The value of DB index is minimum at the optimal clustering point whereas Sil index attains the maximum value at this point. The value of these two indices for the proposed clustering have been compared with those of some of the well-known clustering (parametric and non-parametric) techniques.

Table I lists all the important variables used in the algorithm.

V. Experimental Results

A. Study Areas and Experimental Setup

Four study areas have been presented here for analysis. The images considered here are primarily of the Maharashtra Area,

TABLE I
IMPORTANT PARAMETERS USED IN THE ALGORITHM

Parameter Name	Usage
k_{dist}	Distance at sharpest transition of the distance plot for a given k
v_i	Variance of data points inside the Parzen window at iteration i (Equation (2))
M	Mean edge weight of a connected component (Centroid) (Equation (4))
α	An adaptive parameter used for cluster separation (Equation (4))
σ	Standard deviation of the edge weights of a given connected component (Equation (4))

TABLE II
COMPARISON OF THE BANDWIDTH MEASURES (FOR IMAGES IN FIG. 2(a)–2(d))

No of data points	Dimension	Actual no of classes	h by proposed method	No of classes	Bandwidth by MISE	No of classes
120000	5	7	6.75	7	9.80	4
688492	3	5	31.14	6	37.54	3
743400	4	9	27.46	11	43.55	2
1048576	4	7	21.94	10	32.81	5

Algorithm 2 Input: $T(V, E_T)$, A list W which stores the edge weights in descending order
Output: Array L of labels of the vertices of T

for $i = 1$ to $size(W)$ **do**
Consider the (sub)tree where the $E_{W(i)}$ belongs to. $\{E_{W(i)}$ is the edge in E_T having weight $W(i).\}$
Let T_1 and T_2 be the two CC adjacent to $E_{W(i)}$.
Calculate A_1 and A_2 corresponding to T_1 and T_2 where $A_j, (j \epsilon \{1, 2\})$ is defined as:

$$A_j = M_j + (\alpha_j \times \sigma_j) \tag{4}$$

$\{M_j$ and σ_j define the mean and standard deviation of the weights of the edges contained in T_j. α_j is calculated using Algorithm 3$\}$
if $W(i) > min(A_1, A_2)$ **then**
Delete $E_{W(i)}$ to generate two separate CC's T_1 and T_2.
Label the points of the two CC's and store the labels in L.
else
Break loop.
end if
end for

TABLE III
EXECUTION SPEED COMPARISON BETWEEN PROPOSED MS AND CLASSICAL MS BASED CLUSTERING IN TERMS OF AVERAGE CPU TIME REQUIRED (MEASURED IN SECONDS)

No of data points	proposed MS	classical MS
120000	11.48	14.56
688492	88.37	96.82
743400	100.47	108.61
1048576	155.02	165.10

India acquired by Indian Remote Sensing Satellite P6 Linear Imaging Self-Scanner IV (IRS P6 LISS (IV) (Fig. 2(b)), Quick-Bird (Fig. 2(c)), IRS 1B LISS II (Fig. 2(a)) and IRS 1C LISS III (Fig. 2(d)). Spatial resolution of Fig. 2(a) and (d) are 23.8 m × 23.8 m each and for Fig. 2(b) the resolution is 5.8 m × 5.8 m. Fig. 2(c) is pan-sharpened and has a spatial resolution of 0.61 m.

GLCM calculation requires the generation of the corresponding gray level image given a multi-spectral image. Principal Component Analysis (PCA) based and uniform band

Algorithm 3 Calculate α for a given T_i

1: Use BORDER algorithm [18] to find out the set of data points (S_i) of T_i which resides along the boundary in the feature space. {BORDER employees Gorder KNN join and reverse KNN to find out the boundary points of a data set.}
2: D=average width of the points of T_i in the feature space. $\{D$ is calculated by averaging the pairwise distances of all the points of $(S_i).\}$
3: $\alpha = \frac{D}{M}$. $\{M$ has been defined in the same way as that of Algorithm 2$\}$

averaging based methods have been explored for this purpose. The intensity image (mean of all the available bands) has been selected for texture analysis as PCA based technique has tendency to discard noise, hence, some of the texture components remain unnoticed.

B. *Effectiveness of the Proposed k-Dist Based Bandwidth Selection Technique*

Table II compares the bandwidth measures as obtained by the proposed k-dist based technique with that of the popular Mean Integrated Squared Error (MISE) [19] based technique.

It can be observed from Table II that the bandwidths measured by the MISE algorithm produce over-merged clusters for comparatively large data sets whereas the bandwidth calculated by the proposed k-dist based technique produce near-optimum clustering. As in the proposed case, $h_{\text{used}} = h/2$ (h is calculated from (1)) and $M_{min} = 30$ have been considered, the output of the MS based clustering integrated with the proposed bandwidth selection method guarantees over-segmentation. Gaussian kernel has been used with MS algorithm in all the cases.

C. *Time Analysis of the Proposed MS Based Clustering With the Traditional MS Based Clustering*

The convergence analysis of the proposed MS based clustering with the standard MS clustering in terms of the mean CPU time needed for the execution of the algorithm has been shown in Table III. 20 realizations have been performed in each case and the average CPU time requirement is calculated. The proposed k dist based bandwidth selection strategy has been

TABLE IV
CLUSTERING EFFICIENCY COMPARISON BASED ON THE MEAN SILHOUETTE INDEX (FOR IMAGES IN FIG. 2(a)–2(d))

No of objects	Proposed MST based clustering	Normalized cut	Mean-shift	C-means++	Fuzzy c-means
642	**0.7345** ($N_{MST}=7$)	0.6607 ($N_c=7$)	0.6422	**0.7405** ($N_c=7$)	0.7293 ($N_c=7$)
784	**0.5190** ($N_{MST}=5$)	0.4698 ($N_c=5$)	0.4687	**0.5325** ($N_c=5$)	0.5082($N_c=5$)
1588	**0.6114** ($N_{MST}=9$)	**0.6128** ($N_c=9$)	0.4821	0.6059 ($N_c=9$)	0.5657 ($N_c=9$)
2574	**0.6463** ($N_{MST}=7$)	0.5911 ($N_c=7$)	0.6332	0.6337 ($N_c=7$)	0.6442 ($N_c=7$)

TABLE V
CLUSTERING EFFICIENCY COMPARISON BASED ON THE DB INDEX (FOR IMAGES IN FIG. 2(a)–2(d))

No of objects	Proposed MST based clustering	Normalized cut	Mean-shift	C-means++	Fuzzy c-means
642	**0.5041** ($N_{MST}=7$)	0.5134 ($N_c=7$)	0.6492	0.5186 ($N_c=7$)	0.5213 ($N_c=7$)
784	**0.4301** ($N_{MST}=5$)	0.4319 ($N_c=5$)	0.5183	0.4518 ($N_c=5$)	0.4482($N_c=5$)
1588	**0.3614** ($N_{MST}=9$)	0.4171 ($N_c=9$)	0.4011	**0.3597** ($N_c=9$)	0.3691 ($N_c=9$)
2574	**0.3679** ($N_{MST}=7$)	0.3412 ($N_c=7$)	0.4702	0.3770 ($N_c=7$)	**0.3602** ($N_c=7$)

adopted in both cases. It can be observed that a time reduction in the range of 7%–20% can be achieved using the proposed termination condition for sufficiently large data-sets. The algorithm has been implemented in MATLAB in a machine with 2.66 Ghz Core 2 Duo processor and 4 GB RAM.

D. Validity Comparison of the Proposed MST Based Clustering

The objects extracted after the proposed MS based segmentation undergo a feature extraction stage. The objects are further merged using the proposed MST based clustering technique. The efficiency of this clustering has been compared with some well-known clustering techniques from the literature, i.e., normalized cut based graph clustering [20], adaptive mean-shift based clustering with data centric bandwidth estimation, C-means++ [21] and Fuzzy c-means in terms of Silhouette index and DB index. The initial cluster centers selected by C-means++ have been used for Fuzzy c-means also which is based upon the intuition of spreading the C initial clusters away from each other. Tables IV and V show the results of the comparison based on cluster validity indices. For normalized cut based clustering, C-means++ and Fuzzy c-means, the number of clusters (N_c) has been set manually to the approximate number of classes present in the image based on expert opinion. The values mentioned here are the average of 25 realizations.

N_{MST} denotes the number of clusters **generated** by the proposed MST based clustering. It can be observed from both Tables IV and V that the proposed clustering technique performs near-optimal clustering in almost all the cases and is comparable to other popular clustering techniques. The problem with techniques like Normalized cut, C-means++ or Fuzzy c-means is that the approximate number of clusters needs to be supplied beforehand and even under-merging occurs with those methods. It is difficult to ascertain this information from remotely sensed satellite images *a priori*. The N_{MST} measures exhibit that the proposed method can detect the actual number of clusters. Problems due to cluster shape and size are well taken care of by the underlying tree structure.

E. Comparison of the Segmentation Results

The segmentation results of the proposed hybrid clustering based method have been compared with two recent object based clustering techniques [10] and [8]. Like [10], [8] is also a two step process where the objects extracted from the over-segmented output of the MS based clustering are merged using Normalized cut based clustering. The accuracy assessment for the segmentation technique is performed by comparing the results to the reference classified images obtained by expert annotation.

TABLE VI
COMPARISON IN TERMS OF OVERALL ACCURACY (IN %)
(FOR IMAGES IN FIG. 2(a)–2(d))

Proposed method	Watershed + MS based	MS + Normalized cut based
85.34	73.68	77.24
84.62	81.11	82.33
79.41	76.29	80.70
81.67	79.32	76.61

From the segmentation output of Fig. 2(i), it is evident that the proposed segmentation is able to detect the underlying land-cover classes with maximum possible merging. The outputs in Fig. 2(m) and (q) are largely over-segmented and a certain amount of mis-classification can be observed in both cases. The proposed method also preserves the local texture properties (Fig. 2(j) and (l)). The corresponding results obtained by the other algorithms (Fig. 2(n), (p), (r) and (t)) are over-merged to some extent, hence the local texture details are difficult to ascertain.

These three segmentation methods have further been compared with respect to the reference manually classified image in terms of overall classification accuracy. The reference images have been labeled manually using the domain knowledge. In a classification matrix of size $N_c \times N_c$ where N_c denotes the number of clusters, a typical entry q_{ij} defines how many samples of class i have been misclassified to class j. The overall classification accuracy is defined as

$$\zeta_{\text{overall}} = \frac{\sum_{i=1}^{N_c} q_{ii}}{N} \tag{5}$$

N is the number of samples in the data set. Table VI shows the accuracy assessment of the proposed segmentation technique with respect to the other ones used in terms of ζ_{overall}.

Overall, the proposed segmentation algorithm is capable of segmenting a given satellite image into almost accurate number of land cover classes without any user intervention. Use of MS algorithm with properly tuned parameters guarantees an over-

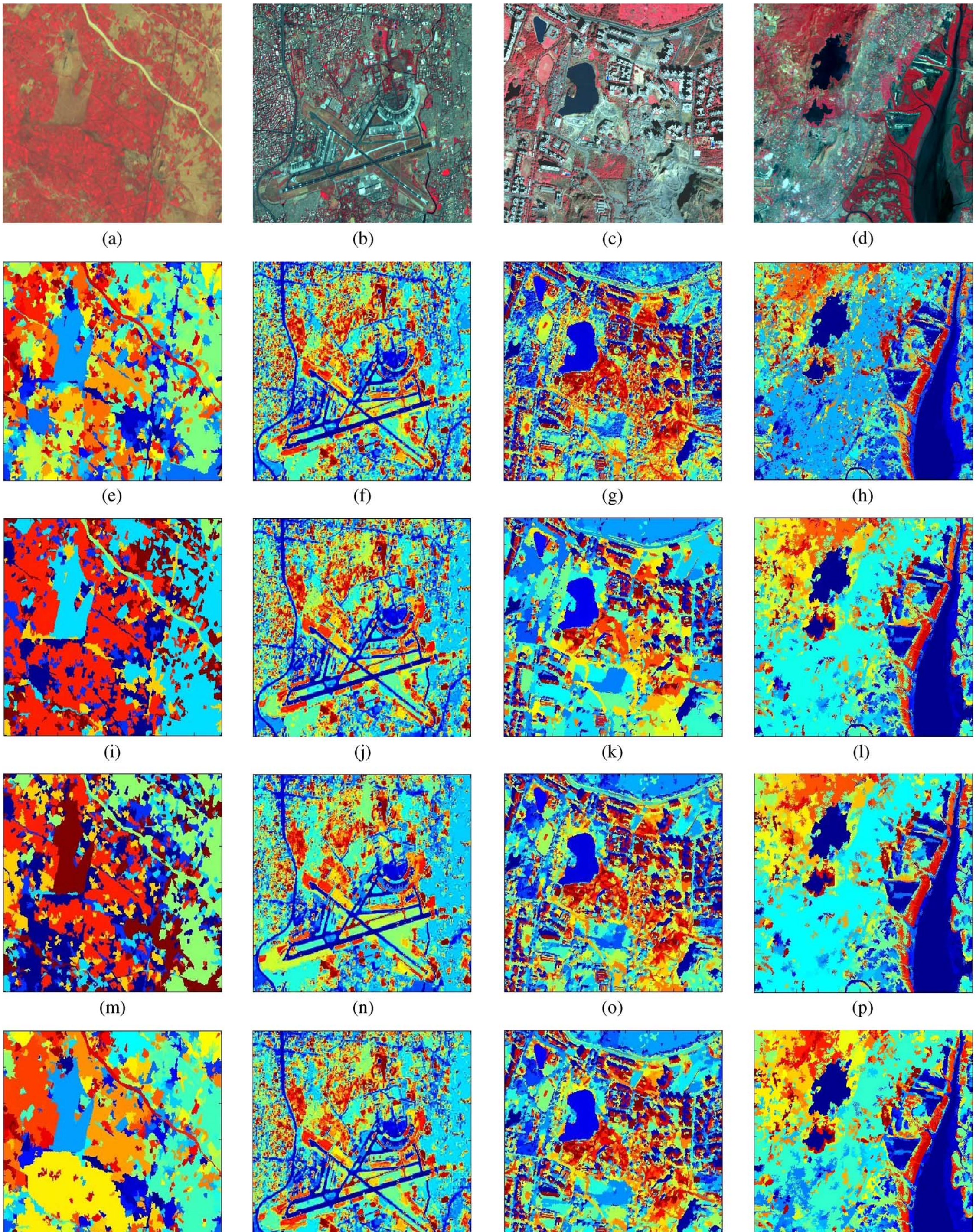

Fig. 2. Comparison of segmentation outputs. (a) FCC of input image 1 (400 × 300), 5 bands; (b) FCC of input image 2 (938 × 734), 3 bands; (c) FCC of input image 3 (900 × 826), 4 bands; (d) FCC of input image 1 (1024 × 1024), 4 bands; (e) output of proposed MS segmentation of (a); (f) output of proposed MS segmentation of (b); (g) output of proposed MS segmentation of (c); (h) output of proposed MS segmentation of (d); (i) output of proposed MST merging on (e); (j) output of proposed MST merging on (f); (k) output of proposed MST merging on (g); (l) output of proposed MST merging on (h); (m) output of watershed+MS clustering based segmentation on (a); (n) output of watershed+MS clustering based segmentation on (b); (o) output of watershed+MS clustering based segmentation on (c); (p) output of watershed+MS clustering based segmentation on (d); (q) output of MS+ Normalized cut based segmentation on (a); (r) output of MS+ Normalized cut based segmentation on (b); (s) output of MS+ Normalized cut based segmentation on (c); (t) output of MS+ Normalized cut based segmentation on (d).

segmented image while keeping fine image details. The edge preserving smoothing property of MS based segmentation helps in removing unnecessary noise elements that may be present in the image due to improper sensing or environmental hazards. Hence, MS based clustering is a better option than the traditional watershed based segmentation in the initial stage. Now to merge these initially found objects, it is always better to adopt any pure non-parametric method due to lack of domain knowledge available for a given remotely sensed image. Non-parametric algorithms like MS are not encouraged to use for this merging stage as these algorithms are often stuck in local modes present in the feature space. The proposed MST based clustering does not suffer from such kind of problems and furthermore, the proposed separation predicate allows sub-division up to a level when the clusters under consideration are sufficiently far from each other.

VI. Conclusion

A novel unsupervised satellite image segmentation algorithm has been proposed here. It consists of two parts. First the image is segmented in the spectral domain by a modified mean-shift based clustering technique. A novel terminating criteria and a novel method for estimating the Parzen window width have been added to the traditional MS clustering algorithm which guarantees fast convergence while performing near-optimal clustering by the MS algorithm. The output of this step is an oversegmented version of the original image. A feature extraction is performed to extract color and texture features from each object thus extracted. The objects are clustered in the feature space using a non-parametric minimum spanning tree based clustering method. This MST based clustering does not require the number of clusters to be mentioned *a priori*. The algorithm preserves the notion of object based image analysis and is comparable to other object based segmentation techniques. The experiments performed show that, the method exhibits better classification accuracy than methods like Watershed + MS based segmentation or MS + Normalized cut based segmentation methods. This segmentation can be used further to extract specific objects like roads, buildings from the image. Object based image understanding which requires clean and efficient extraction of image regions is another domain where the proposed method can enhance accuracy to some extent.

References

[1] Q. Hao, R. Cai, Z. Li, L. Zhang, Y. Pang, and F. Wu, "3D visual phrases for landmark recognition," in *Proc. 2012 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2012, pp. 3594–3601.

[2] W. Long, III and S. Srihann, "Land cover classification of SSC image: Unsupervised and supervised classification using Erdas Imagine," in *Proc. 2004 IEEE Geoscience and Remote Sensing Symp.*, 2004, vol. 4, pp. 2707–2712.

[3] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Computing Surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.

[4] J.-F. Yang, S.-S. Hao, and P.-C. Chung, "Color image segmentation using fuzzy C-means and eigenspace projections," *Signal Process.*, vol. 82, no. 3, pp. 461–472, 2002.

[5] S. Saha and S. Bandyopadhyay, "Application of a new symmetry-based cluster validity index for satellite image segmentation," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 166–170, 2008.

[6] C. Wemmert, A. Puissant, G. Forestier, and P. Gancarski, "Multiresolution remote sensing image clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 533–537, 2009.

[7] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[8] W. Tao, H. Jin, and Y. Zhang, "Color image segmentation based on mean shift and normalized cuts," *IEEE Trans. Syst., Man, Cybern., B: Cybernetics*, vol. 37, no. 5, pp. 1382–1389, 2007.

[9] Z. Li, X.-M. Wu, and S.-F. Chang, "Segmentation using superpixels: A bipartite graph partitioning approach," in *Proc. 2012 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2012, pp. 789–796.

[10] B. Banerjee, V. Surender, and K. Buddhiraju, "Satellite image segmentation: A novel adaptive mean-shift clustering based approach," in *Proc. 2012 IEEE Geoscience and Remote Sensing Symp. (IGARSS)*, Jul. 2012, pp. 4319–4322.

[11] B. Georgescu, I. Shimshoni, and P. Meer, "Mean shift based clustering in high dimensions: A texture classification example," in *Proc. 9th IEEE Int. Conf. Computer Vision*, 2003, pp. 456–463.

[12] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining*, 1996, vol. 1996, pp. 226–231, AAAI Press.

[13] R. Panigrahy, "An improved algorithm finding nearest neighbor using kd-trees," *LATIN 2008: Theoretical Informatics*, pp. 387–398, 2008.

[14] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, 1979.

[15] O. Grygorash, Y. Zhou, and Z. Jorgensen, "Minimum spanning tree based clustering algorithms," in *Proc. 18th IEEE Int. Conf. Tools With Artificial Intelligence, ICTAI'06*, 2006, pp. 73–81.

[16] C.-W. Hsu, C.-C. Chang, and C.-J. Lin *et al.*, "A Practical Guide to Support Vector Classification," 2003 [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

[17] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proc. American Mathematical Society*, vol. 7, no. 1, pp. 48–50, 1956.

[18] C. Xia, W. Hsu, M. L. Lee, and B. C. Ooi, "Border: Efficient computation of boundary points," *IEEE Trans. Knowledge Data Eng.*, vol. 18, no. 3, pp. 289–303, 2006.

[19] D. Comaniciu, "An algorithm for data-driven bandwidth selection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 281–288, 2003.

[20] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[21] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proc. 18th Annual ACM-SIAM Symp. Discrete Algorithms*, 2007, pp. 1027–1035.

Biplab Banerjee is currently a research scholar at the Satellite Image Analysis lab, CSRE, IIT Bombay, India. His research interests include computer vision, image processing, machine learning, algorithm analysis and design.

Surender Varma is currently a research scholar at the Satellite Image Analysis lab, CSRE, IIT Bombay, India. His current research interests are image processing, machine learning, and remote sensing.

Krishna Mohan Buddhiraju has been an Associate Professor at CSRE, IIT Bombay, India, since 1989. His current research interests include high resolution image segmentation and classification, hyper-spectral image analysis, and geographic information systems and science.

Laxmi Narayana Eeti is currently pursuing the Ph.D. degree at CSRE, IIT Bombay, India. His current research interests are image processing and remote sensing.