Defence Science Journal, Vol. 60, No. 3, May 2010, pp. 290-301

Split-and-merge Procedure for Image Segmentation using Bimodality Detection Approach

D. Chaudhuri* and A. Agrawal

Defence Electronics Application Laboratory, Dehradun-248 001
**E-mail: deba_chaudhuri@yahoo.co.in*

ABSTRACT

Image segmentation, the division of a multi-dimensional image into groups of associated pixels, is an essential step for many advanced imaging applications. Image segmentation can be performed by recursively splitting the whole image or by merging together a large number of minute regions until a specified condition is satisfied. The split-and-merge procedure of image segmentation takes an intermediate level in an image description as the starting cutest, and thereby achieves a compromise between merging small primitive regions and recursively splitting the whole images to reach the desired final cutest. The proposed segmentation approach is a split-and-merge technique. The conventional split-and-merge algorithm is lacking in adaptability to the image semantics because of its stiff quadtree-based structure. In this paper, an automatic thresholding technique based on bimodality detection approach with non-homogeneity criterion is employed in the splitting phase of the split-and-merge segmentation scheme to directly reflect the image semantics to the image segmentation results. Since the proposed splitting technique depends upon homogeneity factor, some of the split regions may or may not split properly. There should be rechecking through merging technique between the two adjacent regions to overcome the drawback of the splitting technique. A sequential-arrange-based or a minimal spanning-tree based approach, that depends on data dimensionality of the weighted centroids of all split regions for finding the pair wise adjacent regions, is introduced. Finally, to overcome the problems caused by the splitting technique, a novel merging technique based on the density ratio of the adjacent pair regions is proposed. The algorithm has been tested on several synthetic as well as real life data and the results show the efficiency of the segmentation technique.

Keywords: Segmentation, clustering, bimodality, minimal spanning-tree, homogeneity factor, split-and-merge technique, image segmentation

1. INTRODUCTION

Segmentation is a process of partitioning the image into some interesting regions such that each region is homogeneous and the union of two adjacent regions is non-homogeneous. Segmentation sub-divides an image into its constituent parts or objects. The level to which this sub-division is carried, depends upon the problem being solved. That is, the segmentation should stop when the objects of interest in an application have been isolated[1]. Various segmentation algorithms are available in the literature[2-10]. In general, autonomous segmentation is one of the most difficult tasks in image processing. This step in the process determines the eventual success or failure of the analysis. In fact, effective segmentation rarely fails to lead to a successful solution. For this reason, considerable care should be taken to improve the probability of rugged segmentation. This is a useful and important technique in image analysis and pattern recognition, especially in target detection, medical image analysis, and industrial application, etc. For example, in a vision-guided car assembly system, the robot needs to pick up the appropriate components from the bin. For this, segmentation followed by recognition is required. Target extraction may be accomplished by image segmentation so that all the significant objects are detected and further analysed. Its application area varies from the detection of cancerous cells in medical images to the identification of an airport from remote sensing data, etc. In all these areas, the quality of the final output depends largely on the quality of the segmented output. Hundreds of segmentation techniques are available, like threshold-based segmentation, edge-based segmentation, region-based segmentation, etc.[1,2,5], but there is no single method which can be considered good for all images, nor are all methods equally good for a particular type of image. Moreover, algorithms developed for one class of image (say ordinary intensity image) may not always be applied to other classes of images (MRI/RI)[2].

Many authors used the clustering algorithm as a technique for image segmentation[11,12]. Chaudhuri[13], *et al.* presented a new split-and-merge clustering technique based on the directional data density and the method can be used in image segmentation. To capture the data of a non-convex cluster, a novel, multi-seed clustering method was proposed by Chaudhuri and Chaudhuri[14]. Chaudhuri, *et al.*[15] proposed a novel split-and-merge image segmentation technique to overcome the three basic characteristics–

Received 1 June 2009, Revised 19 August 2009

non-parametric, highly overlapping, and non-Gaussian of remote sensing satellite imagery. A voting approach-based multi-seed supervised classification technique[16,17] is used to classify the image pixels into eight land cover types. A variation of the K-means clustering algorithm, called ISODATA clustering, uses splitting and merging clusters method to do the clustering[18]. However, the ISODATA algorithm determines the number of clusters dynamically. To run the ISODATA algorithm, parameters such as initial cluster means, splitting parameters, lumping parameters, the minimum number of pixels in a cluster, and the number of iterations must be specified by the user. Once these parameters are defined, each sample of the feature space is grouped to the nearest cluster centre. The total number of grouped samples in each cluster must meet the minimum required amount. The cluster is eliminated if the minimum amount cannot be reached. The mean of the rest-grouped samples is computed to update each cluster centre. If the splitting condition is met, split that cluster into two clusters. If the lumping condition is met, lump the two clusters. After either splitting or lumping, the new cluster centres have to be recomputed by newly grouped samples. The algorithm terminates until the maximum number of iterations is exceeded or the converged status of cluster centre occurs.

Thresholding is one of the old, simple, and popular techniques for image segmentation. Thresholding can be done based on global information or using local information of the image. Taxt[19], *et al.* refer to the local and global information-based techniques as contextual and non-contextual methods, respectively. Under each of these schemes, if only one threshold is used for the entire image then it is called global thresholding. On the other hand, when the image is partitioned into several sub-regions and a threshold is determined for each of the sub-regions, it is referred to as local thresholding[19]. Nakagawa and Rosenfeld[20], and Yanowitz and Bruckstein[21] call these local thresholding methods as adaptive thresholding schemes.

Thresholding techniques can also be classified as bi-level thresholding and multi-level thresholding. If the image is composed of regions with different gray-level ranges, i.e., the regions are distinct, the histogram of the image usually shows different peaks, each corresponding to one region and adjacent peaks are likely to be separated by a valley. When the histogram has a (or a set of) deep valley(s), selection of threshold(s) becomes easy because it becomes a problem of detecting valley(s). However, normally the situation is not like this and threshold selection is a critical job. There are various methods[1,22-24] available for this. For example, Otsu[25] maximised a measure of class separability. He maximised the ratio of between class variance to the local variance to obtain thresholds.

Typically, a gray-level image histogram is constructed to study the distribution of pixels according to their gray values. In cases where the background covers a big part of the image area, a peak (mode) will be located at the representative gray value of the background. Another peak (mode) will be obtained at gray-value, which represents the pixels belonging to the object(s), where the objects have a nearly constant gray-value[26]. Here, a basic assumption is made, namely that the image histogram is bimodal, and it turns out that this assumption is correct in many practical cases. In particular, when segmentation is based on gray-level distribution, the histogram must first be built and then analysed in terms of its mode distribution. Thresholds may then be computed so that all pixels having a gray-level between two consecutive thresholds will be labelled differently from all the remaining ones.

Bongiovanni[27], *et al.* presented threshold evaluation based on the assumption of a bimodal histogram and on the computation of a statistical parameter, which has been designed to evaluate the best partition of the population for achieving bimodality[26]. Arifin and Asano[12] proposed a new method of image thresholding using cluster organisation from the histogram of an image. A new similarity measure proposed is based on inter-class variance of the clusters to be merged and the intra-class variance of the new cluster. An important issue to be mentioned before performing segmentation on image with different gray-level intervals, a statistical measure about that region is needed to express if that region can be segmented, they called this as separable measure called SP. If the region's $\mathrm{SP} \geq SP_0$ (SP_0 is the predefined separability parameter), then the region can be segmented; if the region's $\mathrm{SP} < SP_0$ this show that the region has uniformity, and need not be segmented.

This paper focuses a new split-and-merge segmentation technique using bimodality detection approach. Initially, bimodality detection approach is used to split the image into different segmented regions until each segmented region is converted into a homogeneous region. The splitting technique is a recursive approach of automatic multi-threshold gray-level image segmentation based on bimodality detection approach, and this approach can automatically select the optimum threshold (one or more) of gray-level image. Since the splitting technique depends upon homogeneity factor, some of the split regions may or may not split properly. It should be rechecking through merging technique between the two adjacent regions to overcome the drawback of the splitting technique. Here, a sequential-arrange-based or a minimal-spanning-tree-based approach is introduced that depends on data dimensionality of the weighted centroids of all split regions for finding the pair-wise adjacent regions. Finally, to overcome the problems caused by the splitting technique, a novel merging technique based on the density ratio of the adjacent pair regions is introduced. Then the approach is evaluated with several synthetic as well as real-life data.

2. BIMODALITY DETECTION AND SPLITTING TECHNIQUE

The bimodality detection approach[27] has a capability to segment the image into two segmented regions towards the optimal partition. Here, the theory of bimodality detection approach is used to segment the gray-level image in recursive way until each segmented region became homogeneous.

2.1 Bimodality Detection

Let *P* be the population. *P* is called a bimodal if it can be divided into two component sub-populations, say $P_{s(k)}, P_{g(k)}$ such that:

(a) $P_{s(k)}$ contains all the pixels with gray-level value $\leq$ some *k*, and $P_{g(k)}$ contains all the pixels with gray-level value > *k*.

(b) The variances of $P_{s(k)}$ and $P_{g(k)}$ are small relative to the variance of *P*.

Let us assume that *n* and σ^2 are the total frequency and variance of *P*, $n_{s(k)}$ and $\sigma^2_{s(k)}$ be the total frequency and variance of $P_{s(k)}$ and similarly, $n_{g(k)}$ and $\sigma^2_{g(k)}$ be the total frequency and variance of $P_{g(k)}$. Now the gray-level value k' will be determined such that the function

$$W(k) = \frac{n_{s(k)}\sigma^2_{s(k)} + n_{g(k)}\sigma^2_{g(k)}}{n\sigma^2}, \quad f_{\min} \leq k \leq f_{\max}$$

is minimised, where $f_{\min}$ and $f_{\max}$ are the minimum and maximum gray-level values of *P*. The gray-level value k' is called bimodality parameter.

$W(k')$ is then taken as a measure of the bimodality of *P*. More specifically, based on the value of $W(k')$, *P* will be regarded as consisting of the two component sub-populations $P_{s(k')}$ and $P_{g(k')}$ (i.e., *P* is taken as bimodal) or as the single population *P* (otherwise). In the proposed splitting technique, for finding the threshold value (bimodality parameter) k' from image *I*, which is a multi-modal, the function $W(k)$ is used in recursive way provided standard deviation of the current image is less than or equal to some specific threshold. The following are the algorithm steps for finding the value bimodality parameter, k':

Algorithm: Bimodality Parameter Detection

Step 1: Sort the pixel population by gray-level value and assign infinity to $W(k')$.

Step 2: For each *k* from $f_{\min}$ to $f_{\max}$ one can:

(a) Consider $P_{s(k)}$ as consisting of the pixels of the sorted sequence having a value $\leq$ *k*: these will be, say, the first i_k pixels of the sorted sequence. The remaining $n - i_k$ pixels of the sorted sequence are taken as $P_{g(k)}$.

(b) Compute the corresponding $W(k)$ value, and

(c) If $W(k)$ is < $W(k')$, then assign the value $W(k)$ to $W(k')$ and store such value along with the values of frequency, mean and variance for each of the two current sub-populations $P_{s(k)}$ and $P_{g(k)}$.

Since the sequence of pixels is sorted by gray-level value, each iteration of *Step 2* is performed after incrementally scanning the sequence and stopping at the next gray-level transition point.

2.2 Splitting Technique

In this study, the splitting technique is based on bimodality detection approach in recursive way until the homogeneous region is detected. The above mentioned two-level threshold method can segment the image of two regions well, the chosen optimum threshold k' (bimodality parameter) can segment the image into two classes, say I_0 and I_1 effectively, therefore the original image $I = I_0 \cup I_1$. But when there is more than one object in the image, the above method will fail to segment the image. In this situation, multi-level threshold technique is needed to segment the image. Based on this automatic threshold method, a recursive automatic multi-level threshold algorithm for splitting the image has been proposed as follows:

Let $f_i, i = 0,1,\ldots,L-1$ be the gray-level value of the image *I* and h_i, $i = 0,1,\ldots,L-1$ be the corresponding frequency of $f_i, i = 0,1,\ldots,L-1$ in the image. Here, the measure of homogeneity is defined as the standard deviation (*Sd*) wrt mode. If $Sd > T_1$ (where, T_1 is the predefined threshold) then the region is a non-homogeneous. Initially, assume that the original image is non-homogeneous. Segment the original image *I* by the above two-level threshold method. The obtained threshold k' of the original image *I* with gray level [0, *L*-1] can segment the original image into two regions I_0 with gray level [0, k'] and I_1 with gray level [$k'+1, L$-1]. So, $I = I_0 \cup I_1$. Again both the regions I_0 and I_1 are recursively segmented by the above two-level bimodality detection threshold approach if these do not satisfy the homogeneity criteria. Let I_0 will be partitioned into I_{00} and I_{01}, and $I_0 = I_{00} \cup I_{01}$. Similarly, I_1 will be partitioned into I_{10} and I_{11}, and $I_1 = I_{10} \cup I_{11}$. In the same way, continue to perform segmentation on I_{00}, I_{01}, I_{10} and I_{11}, so $I = I_0 \cup I_1 = I_{00} \cup I_{01} \cup I_{10} \cup I_{11} = \ldots$. Therefore, the whole segmentation process is to segment gray-level image recursively, until there is no region that can be segmented. Next, the multi-level threshold automatically chosen algorithm is given.

Algorithm: Splitting Technique

Step 1: Initially, the current region is the input image. $N = 0$ (number of region) and L_l and L_h are minimum and maximum gray-level values of the input image.

Step 2: Initially, let $f_{\min} = L_l$, $f_{\max} = L_h$ and $k' = L_h$.

Step 3: Find mode of the current region i.e. $m = f_i$ for which $h_i = \max_{j=f_{\min}}^{f_{\max}} \{h_j\}$.

Step 4: Compute standard deviation (*Sd*) wrt mode (*m*) for the current region, i.e.,

$$Sd = \left[\frac{1}{\sum_{i=0}^{L-1} h_i} \sum_{i=0}^{L-1} (m - f_i)^2 h_i \right]^{1/2}$$

Step 5: If $SD > T_1$ (where T_1 is the predefined threshold) then go to *Step 6*. Otherwise go to *Step 8*. T_1 is called homogeneity factor.
Step 6: Apply bimodality detection approach of the current region and find k'. Split the current region by k'.
Step 7: $f_{\max} = k'$ and go to *Step 3*.
Step 8: Increment number of region i.e. $N = N + 1$.
Step 9: If $k' = L_h$ then go to *Step 11*. Otherwise go to *Step 10*.
Step 10: $f_{\min} = k' + 1$ and $f_{\max} = L_h$. Go to *Step 3*.
Step 11: STOP.

Earlier an approach[15] was suggested for segmentation of remote sensing satellite imagery. Initially, seed point detection technique[15] was suggested. The technique was based on mode (*m*), $\max_d$ (maximum distance between the mode and the other gray-level values of a particular region) and *Sd* wrt mode. According to the algorithm[15], elongatedness measure, $\delta = |\max_d - Sd|$ and if $\delta \le \theta$ (predefined threshold), then don't split, otherwise collect the nearest neighbour gray values of mode (*m*) whose distances from $m \le Sd$ from the present cluster and formed the core clusters. Next, the region was split on the basis of mode and weighted centroid of all clusters by minimum distance approach. But this approach has a drawback that if the standard deviation (*Sd*) and $\max_d$ are very high in the initial stage, and at the same time δ is very small, then the region will not be split but formed a single cluster though the region is not a homogeneous region. In the present study, the proposed splitting technique will overcome this drawback and also give more impressive results than given by the previous algorithm[15] for various images.

3. PAIR OF ADJACENT REGIONS AND MERGING TECHNIQUE

The proposed splitting technique is used for segmenting the regions based on bimodality parameter (k') and the region homogeneity factor (T_1) around the highest frequency point, mode (*m*). Note that, T_1 should not be too big so that it could include the entire image into one segmented region. Also, T_1 should not be too small so that the image may be segmented as many small regions. So T_1 should be taken suitably. Since the splitting technique depends upon homogeneity factor, some of the split regions may or may not split properly. There should be rechecking through merging technique between the two adjacent regions, which are split by the above technique due to homogeneity factor (T_1). Since the real-life data, like remote sensing satellite data, are highly overlapped and only the proposed splitting technique may not be able to segment the data properly, hence the merging technique has been introduced.

3.1 Adjacent Region

Let n_0 be the number of regions, which are detected by splitting technique. This stage contains two steps: (i) a few candidate pairs out of $n_{0C_2} = \frac{n_0(n_0-1)}{2}$ region pairs are chosen, and (ii) only a few of these candidate pairs are merged. To find the candidate pair, first find the weighted centroids for all regions as follows:

Let n_j be the number of distinct gray-level values for j th region, $j = 1,2,\ldots,n_0$. Let f'_{ij} be the gray-level values of i th level and j th region, where, $i = 1,2,\ldots,n_j$ and $j = 1,2,\ldots,n_0$. Let h'_{ij} be the corresponding frequency of f'_{ij}, $i = 1,2,\ldots,n_j$ and $j = 1,2,\ldots,n_0$. Let CS_j represents the j th segmented region, $j = 1,2,\ldots,n_0$. So the gray-value $f'_{ij} \in CS_j$, $i = 1,2,\ldots,n_j$ and $j = 1,2,\ldots,n_0$. Let $M_j, j = 1,2,\ldots,n_0$ be the integer value of the weighted centroid for j th region and is defined by

$$M_j = \left[\frac{\sum_{i=1}^{n_j} f'_{ij} h'_{ij}}{\sum_{i=1}^{n_j} h'_{ij}} \right], j = 1,2,\ldots,n_0$$

where, $[a]$ means the greatest integer value $\le a$.

Since $M_j, j = 1,2,\ldots,n_0$ is single dimensional data, so order the values M_j in increasing order. The lowest and highest values are the end values and form only one pair each. All other values are the intermediate values and form two pairs each. Let $M_1, M_2, \ldots, M_{n_0}$ are the increasing order weighted centroid values, i.e., $M_1 < M_2 < \ldots < M_{n_0}$ and the pairs are (M_1, M_2), $(M_2, M_3), \ldots$, (M_{n_0-2}, M_{n_0-1}), (M_{n_0-1}, M_{n_0}).

Finding the adjacent region pairs of a single-dimensional data is not much difficult and can be done by the above sequential-arrange-based method. But if the data is multi-dimensional, then the above sequential-arrange-based method is not acceptable for finding the adjacent region pairs. For such data, a minimal-spanning-tree (MST) of the weighted centroid data is generated. A pair of regions is considered for merging only if the data points, i.e., weighted centroids form an edge in the MST.

3.2 Merging Technique

To overcome the problems caused by the splitting technique, a novel merging technique based on the density ratio of the adjacent pair regions is proposed here. Some mathematical definitions, which are used in merging technique, are described below before introduction of the merging technique.

3.2.1 Mathematical Formulation

(a) *Neighbouring gray value*:

Let f be a gray value and $f_i, i = 1,2,\ldots,s$ are the neighbouring gray values of f if $|f - f_i| \le D\ \forall i$, where D be the radius (integer) of closed disc and the set A, which is a collection of all neighbouring gray values $f_i, i = 1,2,...,s$, of *f* is called neighbouring gray value set.

(b) *Neighbouring frequency*:

Let $f \in A$ be a gray value and $f_i, i = 1,2,...,s$ are the neighbouring gray values of *f*. The *neighbouring frequency* of *f* for the set *A* is $H(f)_A = \sum_{i=1}^{s} h_i$, where h_i is the frequency of gray value $f_i, i = 1,2,...,s$.

(c) *Neighbouring frequency of a set*:

The neighbouring frequency of a set is the sum of all neighbouring frequency of all gray values for the set. That is, let *A* be a set of gray values and $f_i \in A, i = 1,2,...,s$. Let t_i be the number of distinct gray values for neighbouring gray value set f_i for $i = 1,2,...,s$. Let $g_{ij}, j = 1,2,...,t_i$, $i = 1,2,...,s$ be the neighbouring gray values of $f_i, i = 1,2,...,s$, respectively. Let h_{ij} be the frequency of gray value $g_{ij}, j = 1,2,...,t_i$, $i = 1,2,...,s$. The neighbouring frequency of set *A* is denoted as *H*(*A*) and defined by $H(A) = \sum_{i=1}^{s} H(f_i)_A = \sum_{i=1}^{s}\sum_{j=1}^{t_i} h_{ij}$.

(d) *Overlap neighbouring set*:

Let A and *B* are two sets of gray values. Let $f_i \in A, i = 1,2,...,t$ and $g_i \in B, i = 1,2,...,u$, where $A \cap B \neq \phi$ (Null set). The overlap neighbouring set of *A* and *B* is denoted as $A * B$ and defined by the collection of all gray values which are the common gray values of *A* and *B*. That is, $A * B = \{G_k : G_k = f_k = g_k \forall k, 1 \le k \le t, u\}$.

(e) *Overlap neighbouring set frequency*:

Let $G_i, i = 1,2,...,s$ are the gray values of overlap neighbouring *set* $A * B$. The overlap neighbouring set frequency is $HO(A * B) = \sum_{i}^{s} h_i$, where h_i is the frequency of gray value $G_i, i = 1,2,...,s$.

(f) *N-Overlap neighbouring set*:

Let $A_1, A_2, ..., A_N$ be *N* number sets of gray values, which are pair wise mutually dependent. That is, they are pair wise overlap to each other. Let the gray value $g_{ij} \in A_i, j = 1,2,...,t_i, i = 1,2,...,N$. The N-Overlap neighbouring set of $A_1, A_2, ..., A_N$ is denoted as $A_1 * A_2 * ... * A_N$ and defined by the collection of all gray values, which are the common gray values of any pair of sets among $A_1, A_2, ..., A_N$. That is,

$$A_1 * A_2 * ... * A_N = \{G_k : G_k = g_{1k} = g_{2k}, 1 \le k \le t_1, t_2; \text{or}$$
$$G_k = g_{1k} = g_{3k}, 1 \le k \le t_1, t_3 \text{ or ... or}$$
$$G_k = g_{N-1k} = g_{Nk}, 1 \le k \le t_{N-1}, t_N\}$$

(g) *N-Overlap neighbouring set frequency*:

Let $G_i, i = 1,2,...,s$ are the gray values of N-Overlap neighbouring set $A_1 * A_2 * ... * A_N$. The N-Overlap neighbouring set frequency is $HO(A_1 * A_2 * ... * A_N) = \sum_{i=1}^{s} h_i$, where h_i is the frequency of gray value $G_i, i = 1,2,...,s$.

(h) *Absolute neighbouring frequency*:

Let *A* and *B* are two sets of gray values. The absolute neighbouring frequency of *A* with respect to *B* is denoted as $[HI(A)]_B$ and defined by the difference between neighbouring frequency of set A and overlap neighbouring set frequency of *A* and *B*, i.e., $[HI(A)]_B = H(A) - HO(A * B)$. Onee can extend this definition for *N* sets as below:

Let $A_1, A_2, ..., A_N$ are *N* sets. The absolute neighbouring frequency of A_1 with respect to $A_2, ..., A_N$ is denoted as $[HI(A_1)]_{A_2, A_3, ..., A_N}$ and defined by $[HI(A_1)]_{A_2, A_3, ..., A_N} = H(A_1) - HO(A_1 * A_2 * ... * A_N)$.

3.2.2 *Merging technique*

On the basis of above definitions, the merging technique is described for a pair of regions (S_{k-1}, S_k), which are to be merged and the same procedure is tested for all other pairs, which formed adjacent pairs.

Algorithm: Merging technique

Step 1: Let M_{k-1} and M_k are weighted centroids of the two regions S_{k-1} and S_k, respectively. Compute the standard deviations SD_{k-1} and SD_k of S_{k-1} and S_k wrt M_{k-1} and M_k, respectively. Find $\min_{SD} = \min\{SD_{k-1}, SD_k\}$. Find the mid point of M_{k-1} and M_k. i.e., $M_c = \left[\frac{M_{k-1} + M_k}{2}\right]$, where, $[a]$ means the greatest integer value $\le a$. Find $d_c = [\beta\%(Sd_{k-1} + Sd_K)]$, where $[a]$ means the greatest integer value $\le a$ and β is any integer value.

Step 2: Let *A* be the equidistance region of S_{k-1} and S_k regions. *A* be the collection of all gray-level values which forms a closed region of $M_c - pd_c$ and $M_c + pd_c$, where *p* is an integer multiplier, i.e., $A = \{f_i : f_i \in [M_c - pd_c, M_c + pd_c]\}$. Let $f_i \in A, i = 1,2,...,\alpha$ be the gray values belong to equidistance region *A*. In present case, *p* = 1 is taken.

Step 3: Let h_i be the corresponding frequency of $f_i \in A, i = 1,2,...,\alpha$. Let *t* number gray values called $f_1^{k-1}, f_2^{k-1}, ..., f_t^{k-1}$ belong to S_{k-1} region and remaining $r = \alpha - t$ number gray values called $f_1^k, f_2^k, ..., f_r^k$ belong to S_k region.

Step 4: For each $f_i^{k-1}, i = 1,2,...,t$ find the set $B_i, i = 1,2,...,t$ collection of all gray values from $S_{k-1} \cup S_k$ region whose differences from $f_i^{k-1} \le \min_{Sd}$. Let F_{ij} be the gray value belongs to set B_i and let m_i be the number of gray values of $B_i, i = 1,2,...,t$. The set B_i is defined by

$$B_i = \left\{F_{ij} : \left|F_{ij} - f_i^{k-1}\right| \le \min_{Sd}, F_{ij} \in S_{k-1} \cup S_k, j = 1,2,...,m_i\right\},$$

$i = 1,2,...,t$.

Let h_{ij} be the frequency of $F_{ij} \in B_i, j = 1,2,...,m_i$, $i = 1,2,...,t$. Since all $F_{ij} \in B_i$, $j = 1,2,...,m_i$ $i = 1,2,...,t$ are coming from $S_{k-1} \cup S_k$, so some gray values of the set B_i are coming from S_{k-1} and some are coming from S_k . Let u_i number of gray values are coming from S_{k-1} and remaining $u_i' = m_i - u_i$, $i = 1,2,...,t$ number of gray values are coming from S_k . So for a particular f_i^{k-1} , the neighbouring gray *values* may be from S_{k-1} or S_k . Let $AF_i(S_{k-1})$ and $AF_i(S_k)$ be the absolute neighbouring frequency for $B_i(S_{k-1})$ and $B_i(S_k)$, respectively for a particular gray value f_i^{k-1} , $i = 1,2,...,t$. Let $H_i(S_{k-1})$ and $H_i(S_k)$ be the neighbouring frequency of set $B_i(S_{k-1})$ and $B_i(S_k)$, respectively for a particular gray value f_i^{k-1} , $i = 1,2,...,t$. So for the set S_{k-1} ,

$$AF_1(S_{k-1}) = H_1(S_{k-1}) = \sum_{j=1}^{u_1} h_{1j}$$

$$\begin{aligned} AF_2(S_{k-1}) &= HI[B_2(S_{k-1})]_{B_1(S_{k-1})} \\ &= H_2(S_{k-1}) - HO(B_1(S_{k-1}) * B_2(S_{k-1})) \\ &= \sum_{j=1}^{u_2} h_{2j} - HO(B_1(S_{k-1}) * B_2(S_{k-1})) \end{aligned} \quad \text{,and so on}$$

$$\begin{aligned} AF_t(S_{k-1}) &= HI[B_t(S_{k-1})]_{B_1(S_{k-1}), B_2(S_{k-1}), ..., B_{t-1}(S_{k-1})} \\ &= H_t(S_{k-1}) - HO(B_1(S_{k-1}) * B_2(S_{k-1}) * ... * B_t(S_{k-1})) \\ &= \sum_{j=1}^{u_t} h_{tj} - HO(B_1(S_{k-1}) * B_2(S_{k-1}) * ... * B_t(S_{k-1})) \end{aligned}$$

Similarly for the set S_k , $AF_1(S_k) = H_1(S_k) = \sum_{j=1}^{u_1'} h_{1j}$,

$$\begin{aligned} AF_2(S_k) &= HI[B_2(S_k)]_{B_1(S_k)} \\ &= H_2(S_k) - HO(B_1(S_k) * B_2(S_k)) \\ &= \sum_{j=1}^{u_2'} h_{2j} - HO(B_1(S_k) * B_2(S_k)) \end{aligned} \quad \text{, and so on}$$

$$\begin{aligned} AF_t(S_k) &= HI[B_t(S_k)]_{B_1(S_k), B_2(S_k), ..., B_{t-1}(S_k)} \\ &= H_t(S_k) - HO(B_1(S_k) * B_2(S_k) * ... * B_t(S_k)) \\ &= \sum_{j=1}^{u_t'} h_{tj} - HO(B_1(S_k) * B_2(S_k) * ... * B_t(S_k)) \end{aligned}$$

Let $AF(S_{k-1})$ and $AF(S_k)$ are the total absolute neighbouring frequencies of S_{k-1} and S_k , respectively then compute

$$AF(S_{k-1}) = \sum_{i=1}^{t} AF_i(S_{k-1}) \text{ and } AF(S_k) = \sum_{i=1}^{t} AF_i(S_k) .$$

Step 5: For each $f_i^k, i = 1,2,...,r$ find the set $C_i, i = 1,2,...,r$ collection of all gray values from $S_{k-1} \cup S_k$ region whose differences from f_i^k are less than equal to $\min_{Sd}$. Let F'_{ij} be the gray value belongs to set C_i and let m'_i be the number of gray values of $C_i, i = 1,2,...,r$.

The set C_i is defined by

$$C_i = \left\{F'_{ij} : \left|F'_{ij} - f_i^k\right| \le \min_{Sd}, F'_{ij} \in S_{k-1} \cup S_k, j = 1,2,...,m'_i\right\},$$

$i = 1,2,...,r$.

Let h'_{ij} be the frequency of $F'_{ij} \in C_i$, $j = 1,2,...,m'_i$, $i = 1,2,...,r$. Since all $F'_{ij} \in C_i$, $j = 1,2,...,m'_i$, $i = 1,2,...,r$ are coming from $S_{k-1} \cup S_k$, so some gray values of the set C_i are coming from S_{k-1} and some are coming from S_k . Let v_i number of gray values are coming from S_{k-1} and remaining $v_i' = m'_i - v_i$ number of gray values are coming from S_k . So for a particular f_i^k , the neighbouring gray values may be from S_{k-1} or S_k . Similarly as Step 7, let $AF'_i(S_{k-1})$ and $AF'_i(S_k)$ be the absolute neighbouring frequency for $C_i(S_{k-1})$ and $C_i(S_k)$, respectively for a particular gray value f_i^k , $i = 1,2,...,r$. Let $H'_i(S_{k-1})$ and $H'_i(S_k)$ be the neighbouring frequency of set $C_i(S_{k-1})$ and $C_i(S_k)$, respectively for a particular gray value f_i^k , $i = 1,2,...,r$. It can be found as Step 7, $AF'_i(S_{k-1})$ and $AF'_i(S_k)$ for $i = 1,2,...,r$. Let $AF'(S_{k-1})$ and $AF'(S_k)$ are the total absolute neighbouring frequencies of S_{k-1} and S_k , respectively then one can compute

$$AF'(S_{k-1}) = \sum_{i=1}^{r} AF'_i(S_{k-1}) \text{ and } AF'(S_k) = \sum_{i=1}^{r} AF'_i(S_k)$$

Step 6: Compute $N_{tot}(S_{k-1}) = AF(S_{k-1}) + AF'(S_{k-1})$ and $N_{tot}(S_k) = AF(S_k) + AF'(S_k)$.

Step 7: Find $Max_{tot} = \max\left\{N_{tot}(S_{k-1}), N_{tot}(S_k)\right\}$ and

$$Min_{tot} = \min\left\{N_{tot}(S_{k-1}), N_{tot}(S_k)\right\} .$$

Step 8: Compute $Q(S_{k-1}, S_k) = \dfrac{Min_{tot}}{Max_{tot}}$.

Step 9: If $Q(S_{k-1}, S_k) \ge T_2$, where, T_2 is very small predefined threshold, $(0 < T_2 \le 1)$ then merged S_{k-1} and S_k . Otherwise S_{k-1} and S_k are different segmented regions. *Q* is called the merging factor.

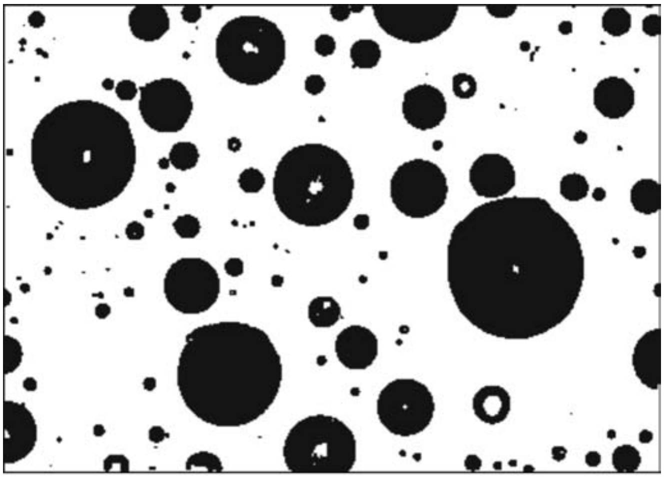
In the above merging technique, to compute d_c , *Step 1* is used for computing the variance tolerance of the equidistance region. Note that β should be specified in this algorithm. As per experience β = 20 is a good choice. The choice of β depends on the number of data and the compactness of the two adjacent regions. *Step 9* represents the density ratio of the two regions with respect to equidistance region. The ratio $Q(S_{k-1}, S_k)$ should be close to 1 if the two regions S_{k-1} and S_k are to be merged. Thus the merging criteria can be decided on the smallness of the quantity $Q(S_{k-1}, S_k)$ and this is always ≤ 1 . So the threshold value T_2 $(0 < T_2 \le 1)$ should be very small (near to 1) and this means, the regions S_{k-1} and S_k are all most equal density.

4. EXPERIMENTAL RESULTS

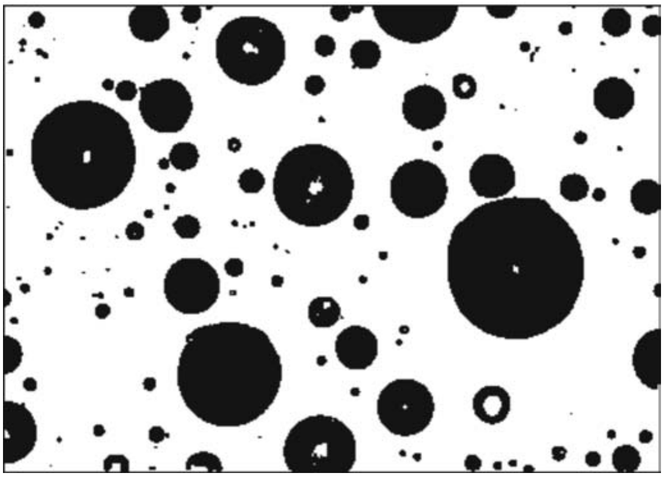
To test the efficiency of the algorithm, several synthetic data as well as real-life data were considered. The algorithm has been tested in binary images, synthetic images of different shapes and grays, Indian Remote Sensing (IRS) satellite panchromatic images, biomedical scanning electron microscope images, and high-resolution satellite panchromatic images. For the present experiments, thresholds T_1 and T_2 are taken as 3.0 and 0.6, respectively, for all images.

Figure 1(a) shows a binary image of several circular shape objects with different sizes, from small dot to big circle. The proposed algorithm is tested on this binary image for validation with thresholds $T_1 = 3.0$ and $T_2 = 0.6$. The segmented image is shown in Fig. 1(b). Each and every object properly segmented irrespective of the size of the object.

Figure 2(a) represents a synthetic image with different shape objects and shades. This image, having limited objects and shades, is taken for testing the capability of the algorithm. The objects marked as 2, 4, and 5 are of the same-gray but different shapes, whereas the objects marked as 1, 3, 6, and 7 are of different grays and shapes. The segmented image by the proposed algorithm is shown in Fig. 2(b). It is noted that the objects 2, 4, and 5 are segmented in the same segmented region with different shapes, whereas objects 1, 3, 6, and 7 are segmented in different segmented regions and these are also different from the segmented regions 2, 4, and 5.

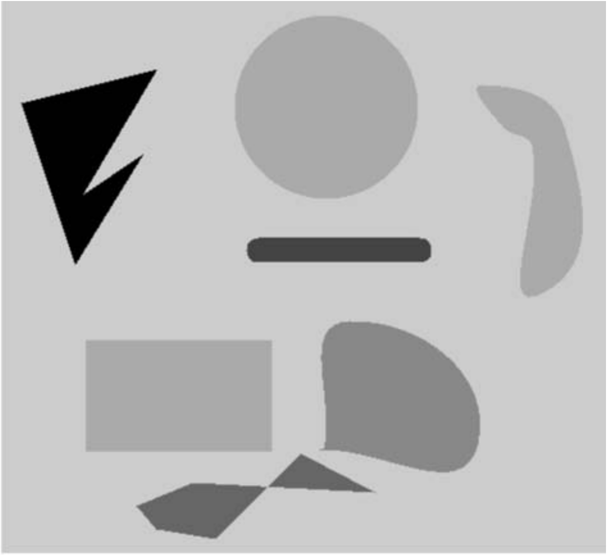


(a)

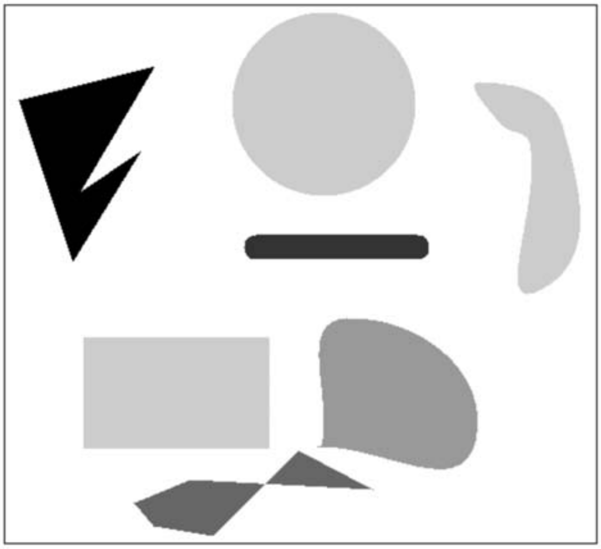


(b)

Figure 1. (a) Binary image and (b) segmented image by the proposed algorithm.



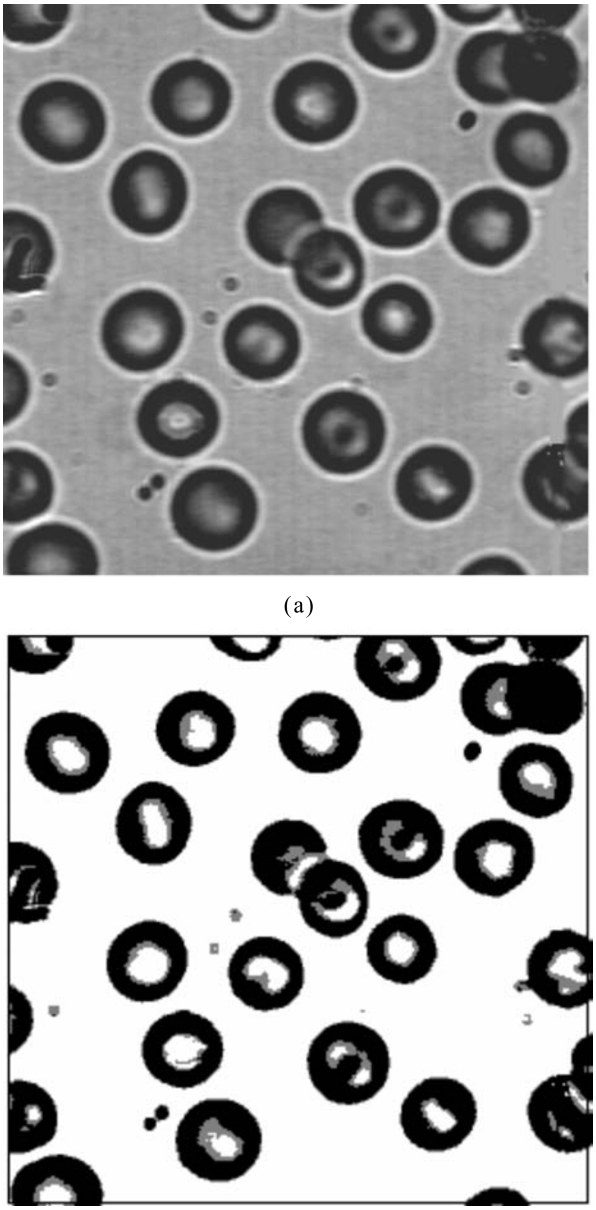
(a)



(b)

Figure 2. (a) Gray-value object image with six different gray and (b) segmented image by the proposed algorithm.

Figure 3(a) shows the original scanning electron microscope gray value image of white blood cells with uniform gray background. Nucleus portions of each cell are different from the outer portions. The segmented image by the proposed algorithm is shown in Fig. 3(b). The nucleus portions of each cell are segmented in different regions (as white region in the centre position of each



(a)

(b)

Figure 3. (a) Original blood cell gray value image and (b) segmented image by the proposed algorithm.

cell) properly from the outer portion of the cells and also each cell is segmented properly from the background.

Figure 4(a) shows the original image of QuickBird satellite in panchromatic mode with 8-bit format. The histogram of the image (Fig. 4(a)) is shown in Fig. 4(b) and the gray-value variation is [0, 255] with multi-modal distribution. It's very difficult to find out the multi-level thresholds for segmentation of the image from the histogram.

There are several objects (buildings) of different shapes and grays (0 – 255) in the original image (Fig. 4(a)). First, the proposed splitting technique is applied for segmentation the regions. Here, initially 16-segmented regions were got with homogeneity factor T_1=3. Figure 4(c) shows the segmented image by the proposed splitting technique. The gray-value ranges of each segmented regions are shown in Fig. 4(d) by alternative colour regions as black and pink. The gray value ranges of all segmented regions from 1 to 16 are (0, 9), (10, 28), (29, 55), (56, 66), (67, 81), (82, 103), (104, 124), (125, 138), (139, 150), (151, 160), (161, 174), (175, 195), (196, 210), (211, 222), (223, 239), and (240, 255), respectively by the proposed splitting technique. This has been reflected in Fig. 4(d).

It can be seen that many regions are not segmented properly by only splitting technique and the regions are segmented into many sub-regions due to choice of homogeneity factor T_1. It is noticed that the modes (m) in some of the segmented regions appear at the boundary of the gray value ranges of the segmented regions. But the central tendency is the best approach for segmentation or clustering technique. To impact the central tendency effect, the weighted centroids of each segmented regions were computed for finding the pair-wise adjacent regions and also for computing M_c and equidistance regions, A. Here, the merging factors, Q for many segmented pair-wise adjacent regions are less than the threshold, T_2 (T_2=0.6). After merging all those adjacent regions the final output of the image is shown in Fig. 4(e) and the corresponding gray value ranges are shown in Fig. 4(f). The final gray-value ranges are (0, 55), (56, 66), (67, 103), (104, 138), (139, 150), (151, 174), (175, 239), and (240, 255) and finally segmented regions wee got after merging. Most of the regions (objects) are segmented properly and the final result (Fig. 4(e)) is much better than the segmented result by only applying splitting technique (Fig. 4(c)). The final output shows the effectiveness of merging technique after splitting technique in the proposed algorithm.

Figure 4(g) shows the segmented image by Revol and Jourlin region growing method[7] and the corresponding gray-value ranges are shown in Fig. 4(h). It can be seen seen closely that many regions are not segmented properly and many objects (regions) are merged with their local backgrounds. It's very difficult to isolate those objects from the background for further high-level processing like object detection/recognition. Also, many regions are segmented into few sub-regions due to intensity variation, which has been reflected in a similar way by the proposed splitting technique (Fig. 4 (c)). Whereas by proposed split-and-merge technique (Fig. 4(e)), the regions are segmented properly and the objects (regions) are isolated from their corresponding local backgrounds, which are useful for further high-level processing. So, the result obtained by the proposed algorithm is better than the result by Revol and Jourlin region growing method.

Figure 5(a) shows an original IRS-1C panchromatic image with spatial resolution 5.8 m × 5.8 m of size 512 × 512. Here, the segmented results of various algorithms and the proposed algorithm are compared. Only the four regions marked as 1, 2, 3, and 4 are highlighted in the original image for comparison of the results. The segmented result by previous algorithm[15] is shown in Fig. 5(b). It can be seen that three black circular-shape objects in the region 2 of the original image (Fig. 5(a)) are almost merged with

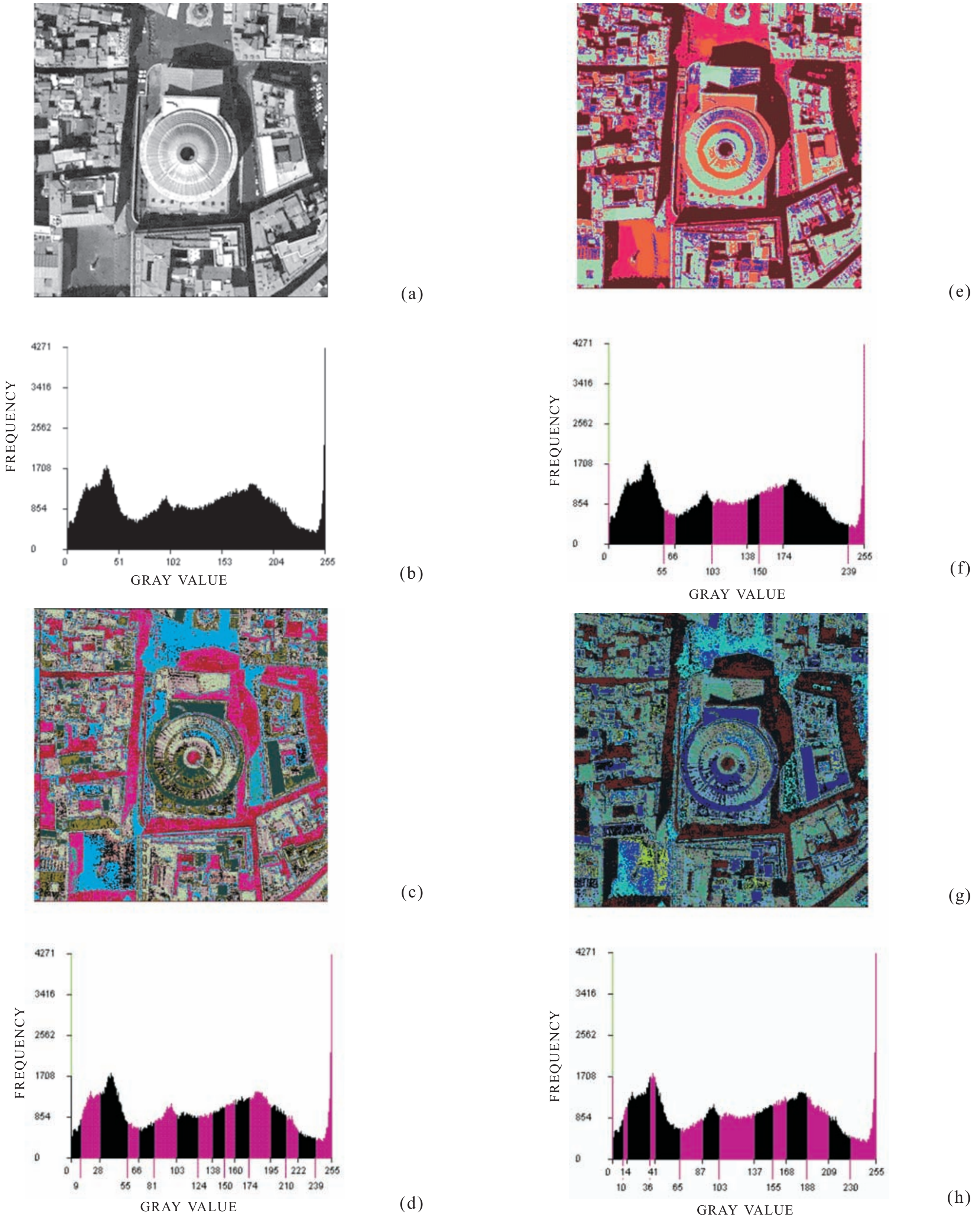
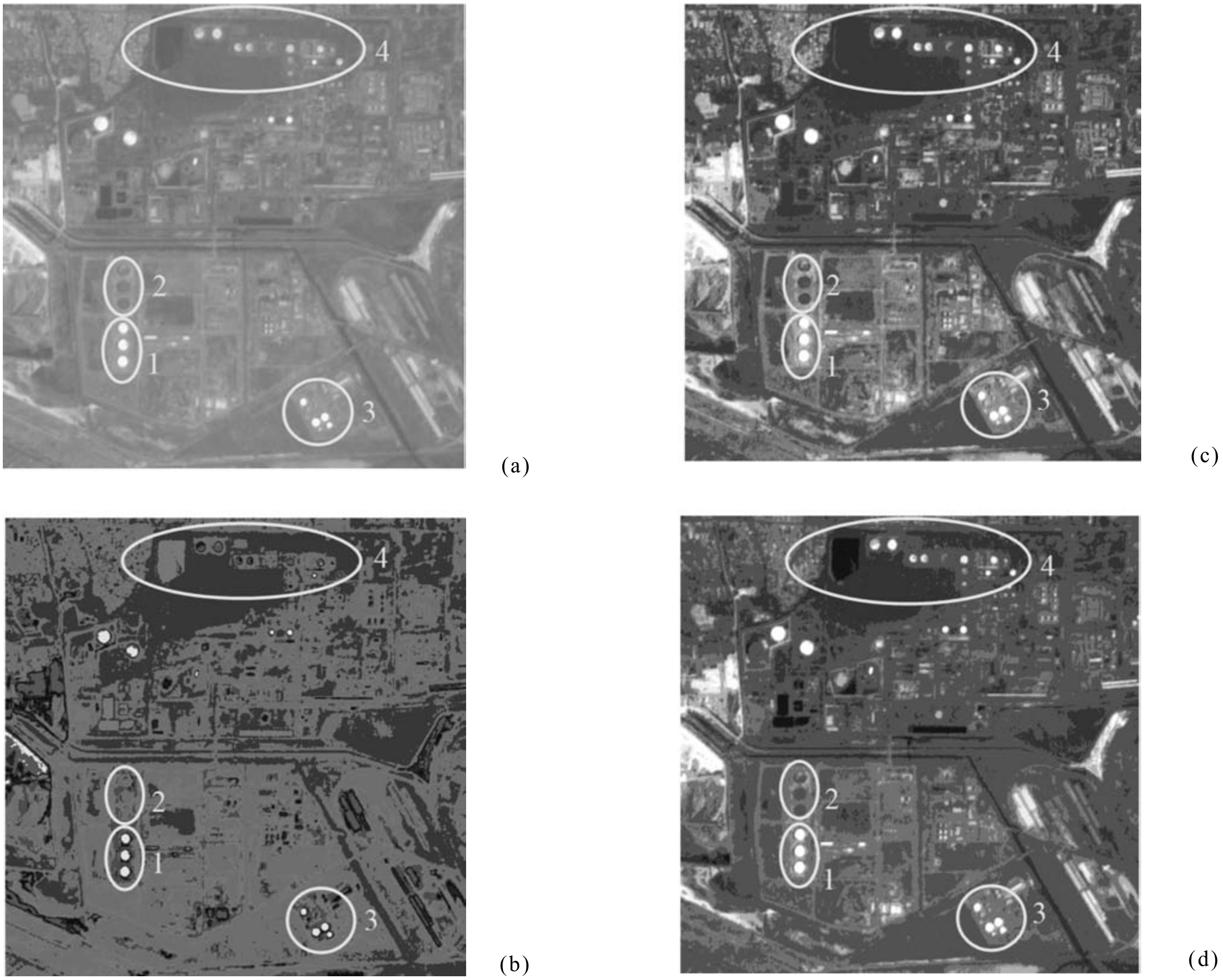


Figure 4. **(a) Original QuickBird satellite image, (b) corresponding histogram of the image, (c) segmented image by the proposed splitting technique, (d) corresponding segmented regions histogram, (e) segmented image by proposed split-and-merge technique, (f) corresponding segmented regions histogram, (g) segmented image by Revol and Jourlin[7] region growing method, and (h) corresponding segmented regions histogram.**



(a) (c) (b) (d)

Figure 5. (a) Original IRS-1C image, (b) segmented image by authors' previous algorithm[15], (c) segmented image by ISODATA algorithm, and (d) segmented image by the proposed algorithm.

the background. These black circular-shape objects are not segmented as separate objects from the background by algorithm[15] (Fig. 5(b)), where these objects are isolated from the background properly by ISODATA algorithm (Fig. 5(c)) and the proposed algorithm (Fig. 5(d)). Also the segmented result of region 4 for bright circular-shape objects in Fig. 5(c) and Fig. 5(d) are much better than the result in Fig. 5(b).

There is a separate black rhombus-type region in the left top corner within region 4 in the original image (Fig. 5(a)). The region is segmented properly by all other reported algorithms except ISODATA algorithm (Fig. 5(c)). This region is merged with the other neighbouring region by ISODATA algorithm. So, in a totality the result by the proposed algorithm is better than the other reported algorithms.

5. SUMMARY

An approach of split-and-merge segmentation technique has been presented. The conventional split-and-merge algorithm is lacking in the adaptability to the image semantics because of its stiff quadtree-based structure. In this study, an automatic thresholding technique based on bimodality detection approach with non-homogeneity criteria is employed in the splitting phase of the split-and-merge segmentation scheme to directly reflect the image semantics to the image segmentation results. Since the splitting technique depends upon homogeneity factor, some of the split regions may or may not split properly. There should be rechecking through merging technique between the two adjacent regions to overcome the drawback of the splitting technique. A sequential-arrange-based or a minimal spanning-tree-based approach has been developed that depends on data dimensionality of the weighted centroids of all split regions for finding the pair-wise adjacent regions. Finally, to overcome the problems caused by the splitting technique, a novel merging technique based on the density ratio of the adjacent pair regions is introduced. The algorithm has been tested on several synthetic as well as real-life data and the results show the efficiency of the segmentation technique.

The complexity of the algorithm is as follows. The algorithm has different modules as: (a) histogram calculation, (b) splitting region based on bimodality detection approach,

(c) pair wise adjacent regions, and (d) merging adjacent region pair. Let n be the size of the image with b-bit data. The order of computation for histogram is $O(n)$. If $\alpha=2^b$, then in the worst case the complexity of splitting technique using bimodality criteria is $O(\alpha^2)$. Since $\alpha=2^b$ and it is constant, so the complexity of splitting technique is constant order. If n_0 be the number of regions after splitting and if the weighted centroids are single dimensional data then for finding the adjacent region pairs the complexity is $O(n_0)$. If the weighted centroids are multi-dimensional data then for finding the adjacent pairs by MST, the complexity is $O(n_0^2)$ (constant). Also the merging technique is constant order. So in totality, the order of the proposed algorithm is $O(n)$.

6. FUTURE RESEACH

The work can be extended in many directions. The segmented technique is based on the gray-level information of the image pixels. So, some mis-segmentation still appears in the output of the experimental results. For better results, this mis-segmentation should be minimised using other higher-level knowledge, i.e., texture, fractal, etc. in the algorithm. Also the thresholds values T_1 and T_2 are heuristic. So there is a scope for further work for automation of these thresholds by data-driven procedure.

ACKNOWLEDGEMENT

The authors are grateful to Mr R.C. Agarwal, Director, Defence Electronics Application Laboratory (DEAL), Dehradun, for granting permission to publish this paper.

REFERENCES

1. Gonzalez, R.C. & Woods, R.E. Digital image processing. Pearson Education, Delhi, 2001.
2. Pal, N.R. & Pal, S.K. A review on image segmentation techniques. *Pattern Recognition*, 1993, **16**(9), 1277-294.
3. Marina, M.; Karl, S. & Hermann, K. Edge and region based segmentation technique for the extraction of large, man-made objects in high-resolution satellite imagery. *Pattern Recognition*, 2004, **37**(8), 1619-628.
4. Serra, J. A lattice approach to image segmentation, *J. Math. Imaging Vision*, 2006, **24**(1), 83-130.
5. Zhang, Y.J. Advances in image and video segmentation. IRM Press, 2006.
6. Boyer, A.L.; Cardenas, C.; Gibou, F.; Liu, P.; Koumrian, T. & Levy, D.; Evaluation of a semi-automated segmentation technique using partial differential equations. *Int. J. Rad. Oncology Bio. Phys.*, 2003, **57**(2), S206-S207.
7. Revol, C. & Jourlin, M. A new minimum variance region-growing algorithm for image segmentation. *Pattern Recog. Lett.*,1997, **18**, 249-58.
8. Zhang, Z.; Chen, C.; Sun, J. & Chan, K.L. EM algorithms for Gaussian mixtures with split-and-merge operation. *Pattern Recognition Lett.*, 2003, **36**(9), 1973-983.
9. Rekik, A.; Zribi, M.; Hamida, A.B. & Benjellounl, M. Review of satellite image segmentation for an optimal fusion system based on the edge and region approaches. *Int. J. Comp. Sci. Network Security*, 2007, **7**(10), 242-50.
10. Zhan, Y.; Wang, W. & Gao, W. A robust split-and-merge text segmentation approach for images. *In* 8th International Conference on Pattern Recognition (ICPR '06), 2006.
11. Dunn, S.M.; Janos, L. & Rosenfeld, A. Bi-mean clustering. *Pattern Recog. Lett.*, 1983, 155-60.
12. Arifin, A.Z. & Asano, A. Image segmentation by histogram thresholding using hierarchical cluster analysis. *Pattern Recog. Lett.*, 2006, **27**(13), 1515-521.
13. Chaudhuri, D.; Chaudhuri, B.B. & Murthy, C.A. A new split-and-merge clustering technique. *Pattern Recog. Lett.*,1992, **13**, 399-09.
14. Chaudhuri, D. & Chaudhuri, B.B. A novel multi-seed non-hierarchical data clustering technique. *IEEE Trans. Syst. Man Cybernetics*, 1997, **27**(5), 871-77.
15. Chaudhuri, D.; Mishra, A.; Anand, S.K. & Gohri, V. Optimal splitting technique for remote sensing satellite imagery data. *SPIE Proc*eedings, 2001, **4388**, 79-88.
16. Chaudhuri, D. & Samal, A. An automatic bridge detection technique for multi-spectral images. *IEEE Trans. Geosci. Remote Sens.*, 2008, **46**(9), 2720-727.
17. Mehere, V. & Chaudhuri, D. Multi-spectral classification technique based on data driven procedure and MDLBC classifier. *In* International Conference on Communication, Devices and Intelligent Systems (CODIS-2004). Vol. 34, 2004. pp. 554-57.
18. Ball, G.H. & Hall, D.J. Some fundamental concepts and synthesis procedures for pattern recognition preprocessors. *In* International Conference on Microwaves, Circuit Theory and Information Theory, September, Tokyo, 1964.
19. Taxt, T.; Flynn, P.J. & Jain, A.K. Segmentation of document images. *IEEE Trans. Pattern Analysis Mach. Intell.*, 1989, **11**(12), 1322-329.
20. Nakagawa, Y. & Rosenfeld, A. Some experiments on variable thresholding. *Pattern Recognition*, 1979, **11**, 191-04.
21. Yanowitz, S.D. & Bruckstein, A.M. A new method for image segmentation. *Comp. Vision Graphics Image Proc.*, 1989, **46**, 82-95.
22. Horowitz, S.L. & Pavlidis, T. Picture segmentation by directed split and merge procedure. *In* Proceedings of 2nd International Joint Conference on Pattern Recognition, 1974. pp. 424-33.
23. Perez, A. & Gonzalez, R.C. An iterative thresholding algorithm for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1987, **9**, 742-51.
24. Jalba, A.C. & Roerdink, J.B.T.M. Automatic segmentation of diatom images. *In* Proceedings of 10th International Conference on Computer Analysis of Images and Patterns, (CAIP–2003), 2003. pp. 329-36.
25. Otsu, N. A threshold selection method from gray level histograms. *IEEE Trans. Syst. Man Cybernetics*, 1979,

9, 62-66.

26. Bongiovanni, G.; Cinque, L.; Levialdi, S. & Rosenfeld, A. Bimodality analysis using pyramids. *In* Multi-computer Vision, edited by S. Levialdi. Academic Press, New York, 1988. pp. 1-7.
27. Bongiovanni, G.; Cinque, L.; Levialdi, S. & Rosenfeld, A. Image segmentation by a multi-resolution approach. *Pattern Recognition*, 1993, **26**(12), 1845-854.

Contributors

Dr D. Chaudhuri obtained his MSc in Applied Mathematics from Jadavpur University, Kolkata. He received his PhD in image processing and pattern recognition from Indian Statistical Institute, Kolkata. He worked at Defence Institute of Advanced Technology, Pune and Integrated Test Range, Chandipur. Presently, he is working at the Defence Electronics Application Laboratory (DEAL), Dehradun. He has published more than 25 papers in reputed international journals and is a reviewer of many international and national journals. His areas of research include: image processing, pattern recognition, computer vision, and remote sensing.

Mr A. Agrawal obtained his BE (Computer Science) from Rajiv Gandhi Technical University, Bhopal, in 2003. Presently, he is working at DEAL, Dehradun. His areas of research include: image processing, computer vision, and remote sensing.