Performance Evaluation and Comparative Analysis of Proposed Image Segmentation Algorithm

Rajiv Kumar[1]* and A. M. Arthanariee[2]

[1]Research and Development Centre, Bharathiar University, Coimbatore (Tamil Nadu), India;
dhimanrajiv2000@yahoo.co.in
[2]Department of Science and Humanities, Nehru Institute of Technology, Coimbatore, India;
arthanarimsvc@gmail.com

Abstract

In this paper, a novel approach of K-Region based Clustering image segmentation algorithm has been proposed. The proposed algorithm divides an image of size N × N into K number of regions. The K and N are multiples of 2. The value of K must be less than N. Authors divided the image into 4, 16, 64, 256, 1024, 4096 and 16384 regions, based on the value of K. The adjacent pixels having similar intensity value in each region are grouped into same clusters. Further, the clusters of similar values in each adjacent region are grouped together to form the bigger clusters. The different segmented images have been obtained based on the K number of regions. The four parameters, namely, Probabilistic Rand Index (PRI), Variation of Information (VOI), Global Consistency Error (GCE) and Boundary Displacement Error (BDE) have been used to evaluate the performance of the proposed algorithm. The performance of proposed algorithm was evaluated using 100 images taken from Berkeley image database. The time-complexity of the proposed algorithm has also been calculated. The comparative analysis of proposed algorithm was made with existing image segmentation algorithm, namely, K-mean clustering and Region-growing algorithm. Significant results were obtained in case of proposed algorithm when\the PRI, VOI, GCE and BDE values were compared with those of existing algorithms. MATLAB 7.4 has been used to implement the proposed algorithm.

Keywords: Image Segmentation, Clusters, Regions, K-mean Clustering, Region-growing, MATLAB 7.4.

1. Introduction

Image segmentation is the process of identifying and delineating objects in images. Image segmentation may consist of two related processes—recognition and delineation. Recognition is the high-level process, which is used to find the rough idea about the object of interest present in the given image. On the other hand, delineation is the low-level process, which is used to find the point-to-point composition, and spatial extent of the object present in the image[1, 2]. It is difficult to find the performance evaluation of image segmentation algorithm.

The evaluation of segmentation is difficult to embed in specific type of applications[3, 4]. Sathya et al.[5] suggested that the purpose of clustering technique is to divide an image into K number of disjoint clusters. The members who are similar in some sense become the part of same cluster; otherwise they belong to different clusters. Jaskirat Kaur et al.[3] found that there are many different clustering algorithms to decide the criteria of this similarity. The unsupervised image segmentation is used to separate the image into number of disjoint regions. Each region consists of the finite number of classes. Each class is characterized by distinct parameters. The main goal of

---

*Author for correspondence:

clustering is to maximize the homogeneity within each cluster, and the heterogeneity among different clusters. In other words, objects that belong to the same cluster should be more similar to each other than objects that belong to different clusters. K-mean clustering algorithm partitions the image data set into k number of sets. It finds the Euclidean k-median distance. It minimizes the sum of distance to the nearest center[6]. The advantage of K-mean clustering algorithms is that the classification is simple and easy to implement. But, it is difficult to determine the number of clusters and decrease the numbers of iterations[7]. The number of iterations can be decreased by taking the scale value as large as possible among the disjoint regions in the image. It can also be reduced by differentiating the colours as much as possible[8]. Region growing algorithm uses edge as a growth-stopping condition, and the growing seeds are selected manually. The growth occurs in the homogenous intensity regions and stops at the edges. If there are broken edges, then the segmentation is incorrect.

There are various algorithms to segment the image. But, these algorithms have their pros and cons. For example, the performance of Region-growing algorithm depends on the number of seed points in the image. But, it is very difficult to choose the exact number of seed points to segment the image. On the other hand, the performance of K-mean clustering algorithm highly depends on the number of clusters in the image.

The proposed algorithm is very flexible and simple to design. Authors have combined both cluster-based and region-based approaches. First, an $N \times N$ image is divided into K number of regions. Both N and K are multiples of 2 such that K < N. Subsequently, clustering technique has been applied in each region. Finally, clusters in each adjacent region are merged together to obtainthe larger clusters. In this way, the segmented image was obtained. The number of regions can be chosen easily. It depends on the value of K. Clusters can be formed easily by comparing the intensity value of two adjacent pixels. Formation of clusters depends on some predefined criteria. The obtained segmented image is very useful for image understanding. It can be used in application like face recognition, finger recognition etc. In this paper, the authors have taken 100 images from Berkeley image database[9] to evaluate and perform the comparative analysis of proposed, K-mean clustering and Region-growing algorithm.

This paper is presented in different sections: Section 2 proposes the K-Region based clustering algorithm and has shown the working for image of size 4×4. In section 3, authors have taken four parameters, namely, PRI, GCE, VOI and BDE to evaluate the performance of proposed algorithm. Section 4 performs the comparative analysis of proposed algorithm with K-mean clustering and Region-growing algorithm. Finally, paper concludes in section 5.

2. Proposed Algorithm

Two approaches namely, region-based and cluster-based have been combined together. First, an image of size $N \times N$ is divided into K number of regions. Then, clustering technique is applied in each region. Finally, clusters in each adjacent region are merged together to form the larger clusters based on chosen threshold value. In this way, the segmented image has been obtained. The proposed algorithm divides an image of size $N \times N$ into K number of regions. In each region, the adjacent pixels have similar values belonging to the same cluster. Otherwise, they belong to different clusters. Following this, clusters of similar value in adjacent regions form the larger clusters. Otherwise, they will be separate clusters. This is shown in Figure 1.
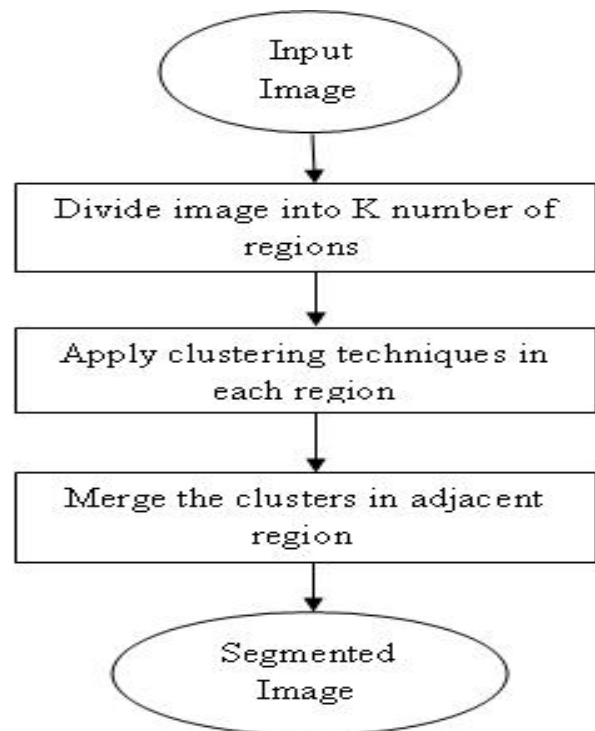

Figure 1.   Proposed algorithm.

2.1 Proposed Algorithm

The following are the steps of algorithm:

1. Select an image of size N × N, where $N = 2^i$, i = 2, 3, ...
2. Divide the image into K number of regions, where $K = 2^j$, j = 1, 2, ... and K < N.
3. Scan the pixels of first row of each region in the image.
   3.1 Compare the first pixel with second pixel in the row. If the difference between these two pixels is less than or equal to the specified threshold, take the average of these two pixels and put them into same cluster. Otherwise, they are part of different clusters.
   3.2 Repeat step 3.1 for the remaining pixels of the first row in the region.
4. Repeat step 3 for the remaining rows of each region in the image.
5. Compare the second row with the first row in each region in the image.
   5.1 The first pixel of second row is compared with the first pixel of first row. As condition specified in step 3.1, clusters are formed.
   5.2 Compare second pixel of the second row with the pixel above and left to it. Find out the difference between second pixel and pixel above as well as pixel left to it. Based on the given threshold, form the clusters.
   5.3 Repeat step 5.2 for the remaining rows.
   After step 5, the clusters in each region are formed.
6. Scan the boundary pixels vertically in each adjacent region of the image.
   6.1 If the difference between first boundary pixel of each adjacent regions is less than or equal to given threshold, then they will merge together and form the clusters. Otherwise, they are treated as different clusters.
   6.2 For the remaining boundary pixels, compare a pixel with the pixel above and left to it. The clusters are formed in similar way as condition specified in step 5.2.
   After this step, the clusters in vertically adjacent region are merged together and form the bigger clusters.
7. Repeat step 6 for the boundary pixel horizontally in each adjacent region of the image. Afterwards, the clusters in horizontally adjacent region are grouped together and form the larger cluster.

8. Exit.

2.2 Working of Proposed Algorithm

The working of proposed algorithm of image of size 4 × 4 is shown as in Figure 2:

1. Here, the image of size 4 × 4 is chosen. This is shown in Figure 2(a).
2. The image is divided into K number of regions such that K < N. In this example, K=4 and N=4 × 4. This is shown in Figure 2(b).
3. Apply the clustering technique in each region. This step is further divided into two parts.
   3.1 Perform the clustering technique row-wise in each region. If the value of adjacent nodes is less than or equal to specified threshold, then the two nodes are merged, and form the clusters. Otherwise, they are treated as separate cluster. Here, 45 is the threshold value. The total number of clusters is 10 as shown in Figure 2(c).
   3.2 Merge the clusters of second row with that of first row and form the larger clusters. After this step, the total number of clusters is 7. This is shown in Figure 2(d).
4. In this step, clusters are grouped region-by-region. This step is further divided into two parts.
   4.1 Merge the clusters which are at the boundary of the vertically adjacent regions and form the larger clusters. Now, the number of clusters is 4 as shown in Figure 2(e).
   4.2 Now, perform the clustering at the boundary in the horizontally adjacent regions. Merge them together and form the bigger clusters. Finally, the total number of clusters is 2 which are shown in Figure 2(f).

3. Performance Evaluation of Proposed Algorithm

The following image segmentation parameters are used to compare the segmentation results for the same set of images. (i) Probabilistic Rand Index (PRI): It counts the fraction of pairs of pixels whose labeling are consistent between the computed segmentation and the ground truth, averaging across multiple ground truth segmentations to account for scale variation in human perception. The Rand index or Rand measure is a measure

of the similarity between two data clusters. (ii) Global Consistency Error (GCE): It measures the extent to which one segmentation can be viewed as a refinement of the other. If one segment is a proper subset of the other, then the pixel lies in an area of refinement, and the error should be zero. If there is no subset relationship, then the two regions overlap in an inconsistent manner. (iii) Variation of Information (VOI): It defines the distance between two segmentations as the average conditional entropy of one segmentation given the other, and thus measures the amount of randomness

in one segmentation which cannot be explained by the other. (iv) Boundary Displacement Error (BDE): It measures the average displacement error of boundary pixels between two segmented images. Particularly, it defines the error of one boundary pixel as the distance between the pixel and the closest pixel in the other boundary image[5].

An image of size 256×256 was divided into 4, 16, 64, 256, 1024, 4096 and 16384 regions based on the value of K. Then, clustering technique is applied on each region. Finally, clusters in the adjacent regions are merged together to form the larger clusters. Segmented images are obtained from the given image based on the K number of regions. Four parameters viz. PRI, GCE, VOI and BDE were considered to evaluate the performance of these segmented images. The experiment was conducted on more than 100 images taken from Berkeley image dataset. The following results were obtained: PRI value has good results when image is divided into 64 regions. Good GCE value was obtained when image is divided into 16384 regions. VOI has good results in case of 1024 and 16384 regions and BDE value is very good in case of 16384 regions. Overall, good results can be obtained when a given image is divided into 16384 regions. These segmented images are quite useful for image understanding purpose. The value of PRI, GCE, VOI and BDE must be low for good segmented images. The results of proposed algorithm on various images are shown in Table 1–4.
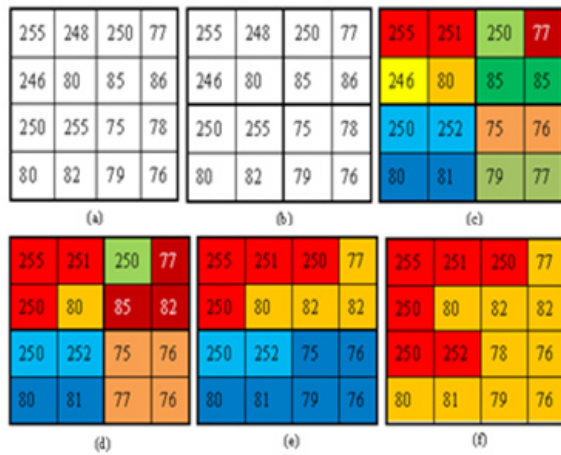

Figure 2. (a) Original Image, (b) Division of Image into 4 Regions, (c) Row-wise Clusters in each Region, (d) Clusters in each Region, (e) Clusters in each adjacent Region of Image, (f) Clusters of Image.

Table 1. Performance evaluation of proposed algorithm on image 1

	R = 4 C = 16384	R = 16 C = 4096	R = 64 C = 1024	R = 256 C = 256	R = 1024 C = 64	R = 4096 C = 16	R = 16384 C = 4
PRI	0.9778	0.9778	0.9778	0.9779	0.9783	0.9788	0.9794
GCE	0.8667	0.8643	0.8645	0.8605	0.8434	0.8202	0.7894
VOI	8.0846	8.0337	8.0236	7.9236	7.6137	7.1893	6.7116
BDE	0.0960	0.0928	0.0951	0.0931	0.0912	0.0841	0.0746

Table 2. Performance evaluation of proposed algorithm on image 2

	R = 4 C = 16384	R = 16 C = 4096	R = 64 C = 1024	R = 256 C = 256	R = 1024 C = 64	R = 4096 C = 16	R = 16384 C = 4
PRI	0.9796	0.9796	0.9796	0.9797	0.9799	0.9802	0.9807
GCE	0.8870	0.8870	0.8871	0.8829	0.8740	0.8590	0.8311
VOI	8.1812	8.1714	8.1790	8.0816	7.8544	7.573	7.0627
BDE	0.0350	0.0356	0.3300	0.0331	0.0328	0.0304	0.0251

Here, R represents the region and C represents the cells within the region.

The comparison of PRI, GCE, VOI and BDE values have shown from Figures 3–6.

The results of proposed algorithm have been shown as in Figure 7.

The 16384 Region segmented image has edge over other segmented images. The percentage difference between 16384 Region and other regions was calculated by taking the formula as shown in equation (1):

$$\text{Result} = \left| \frac{X - Y}{(X + Y)/2} \right| \times 100 \qquad (1)$$

where, X and Y represents the value of parameters of segment image.

Table 5 has been shown the percentage difference between 16384 region and other K regions of proposed algorithm. It is proved that 16384 region segmented image have better results as compared to other segmented images.

Authors have calculated the time-complexity of the proposed algorithm. The total number of comparison of the cells in the image has been calculated. This has been shown in Table 6.

Here, $N = 2^n$, n = 1, 2, 3, …

The total number of the cells in the image is N^2 depending on the value of n. C is the total number of comparison of the cells occurred in the given image. The recurrence relation of the proposed algorithm has been given as:

$$T(n) = \begin{cases} -1 & \text{if } n=0 \\ 4*T(n-1)+8*2^{n-1} & \text{otherwise} \end{cases}$$

It has been given as:

$$T(n) = 4*T(n-1)+8*2^{n-1}$$

$$= 4*[4*T(n-2)+8*2^{n-2}]+8*2^{n-1} \quad \text{(By induction hypothesis)}$$

$$= \dots$$

$$\approx 8*2^{n-1}$$

$$\approx 2^{n-1}$$

$$\approx 2^n$$

$$\approx O(2^n)$$

Table 3. Performance evaluation of proposed algorithm on image 3

	R = 4 C = 16384	R = 16 C = 4096	R = 64 C = 1024	R = 256 C = 256	R = 1024 C = 64	R = 4096 C = 16	R = 16384 C = 4
PRI	0.9686	0.9696	0.9692	0.9681	0.9704	0.9726	0.9747
GCE	0.7119	0.7116	0.7119	0.7135	0.7039	0.6928	0.6763
VOI	6.8931	6.8854	6.8807	6.8923	6.7096	6.5012	6.2096
BDE	0.1192	0.1313	0.1229	0.1117	0.0859	0.0770	0.0726

Table 4. Performance evaluation of proposed algorithm on image 4

	R = 4 C = 16384	R = 16 C = 4096	R = 64 C = 1024	R = 256 C = 256	R = 1024 C = 64	R = 4096 C = 16	R = 16384 C = 4
PRI	0.9864	0.9864	0.9864	0.9864	0.9864	0.9867	0.9870
GCE	0.8971	0.8964	0.8956	0.8937	0.8867	0.8733	0.8545
VOI	8.9476	8.9257	8.9148	8.8164	8.5811	8.1996	7.7593
BDE	0.0352	0.0349	0.0332	0.0324	0.0329	0.0303	0.0256

Table 5. Percentage difference between 16384 Region and other K regions

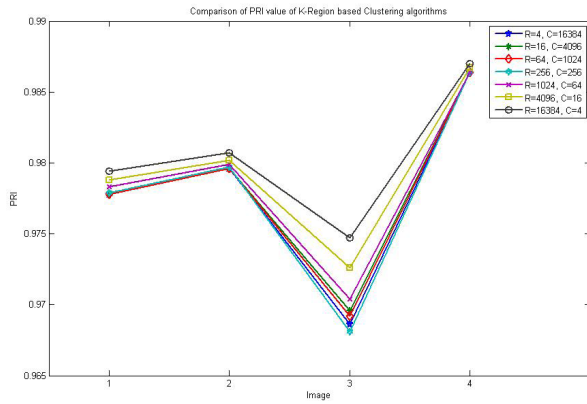|	R = 16384 and R = 4	R = 16384 and R = 16	R = 16384 and R = 64	R = 16384 and R = 256	R = 16384 and R = 1024	R = 16384 and R = 4096
PRI	0.0712	0.0712	0.0814	0.0712	0.0712	0.0509
GCE	2.3895	2.3895	2.3577	2.2412	2.082	2.5587
VOI	9.7378	9.7132	9.6759	9.1222	8.1539	5.5153
BDE	41.0256	42.0382	40	33.557	27.7778	25.3521

Figure 3. Comparative analysis of proposed algorithm on PRI value.
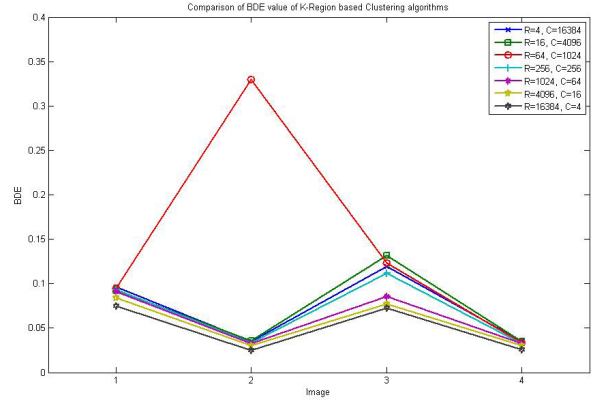

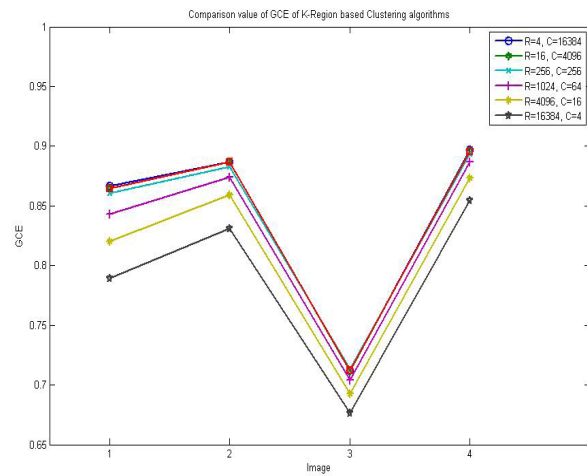**Figure 6.** Comparative analysis of proposed algorithm on BDE value.

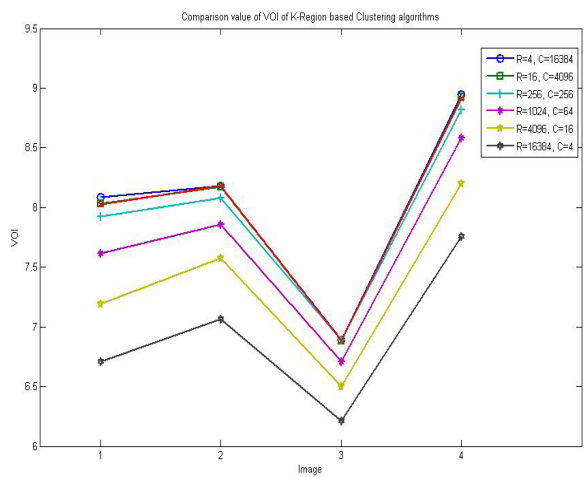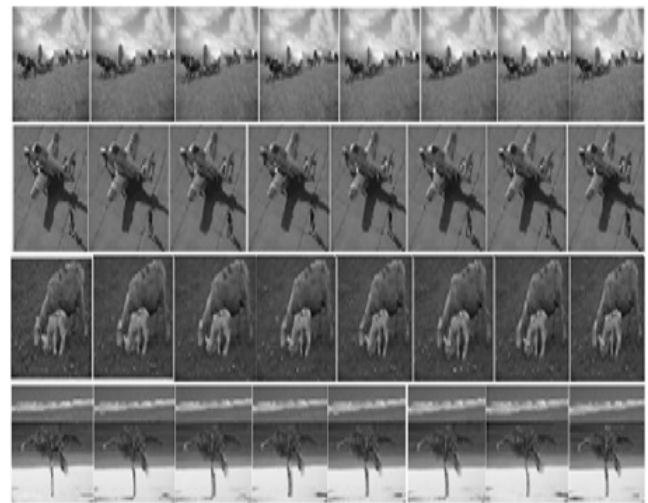**Figure 4.** Comparative analysis of proposed algorithm on GCE value.

Figure 7. i) Original Image, ii) 4-Region Image, iii) 16-Region Image, iv) 64-Region Image, v) 256-Region Image, vi) 1024-Region Image, vii) 4096-Region Image and viii) 16384-Region Image.

Figure 5. Comparative analysis of proposed algorithm on VOI value.

Table 6. Number of Comparison of the cells in image

n	N	C
1	2	4
2	4	24
3	8	112
4	16	480
5	32	1984
6	64	8064
7	128	32512
8	256	130560

The number of comparisons of the cells in the given image has been increased exponentially for the larger value of n. The time complexity of the proposed algorithm has been found $O(2^n)$.

4. Comparative Analysis of Proposed Algorithm with Existing Algorithms

Authors have taken two existing algorithms namely, k-mean clustering and region-growing algorithm. K-mean clustering generates a specific number of disjoint, flat (non-hierarchical) clusters. It is well suited to generating globular clusters. The k-mean method is numerical, unsupervised, non-deterministic and iterative. In k-mean clustering approach, the dataset is partitioned into k clusters and the data points are randomly assigned to the clusters resulting in clusters that have roughly the same number of data points. The time complexity of k-mean clustering algorithm is $O(n*k*l*d)$, where n, k, l and d is the number of points, number of clusters, number of iterations and number of attributes respectively[6].

Region-growing is a procedure that group pixels or sub-regions into larger regions based on predefined criteria for growth. The basic approach is to start with a set of 'seed' points, and from these to grow regions by appending to each seed those neighboring pixels that have predefined properties similar to the seed (such as specific range of gray level or colour). The most serious problem with region growing is power and time consuming.

Authors have compared the proposed 16384 region segmented image algorithm with existing algorithms. The PRI, GCE, VOI and BDE value were calculated to perform the comparative analysis among these algorithms. Table 7–12 has shown the comparative analysis of these algorithms on various images, taken from Berkeley image database[8].

The graph represents the comparative analysis of PRI, GCE, VOI and BDE of proposed, K-mean clustering and Region-growing algorithm. This has been shown from Figures 8–11.

Hindered images from Berkeley image database[2] were taken. The resultant images of proposed, K-mean clustering and Region-growing algorithm have been shown in the Figure 12.

5. Conclusion

In this paper, we obtained different segmented images from the given images, based on the value of K. The segmentation parameters PRI, GCE, VOI and BDE are used to evaluate the performance of the proposed algorithm. Authors h obtained 4, 16, 64, 256, 1024, 4096

Table 7. Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on airplane image

	Proposed algorithm	K-mean clustering	Region-growing
PRI	0.9762	0.1405	0.0997
GCE	0.8609	0.9786	0.9789
VOI	5.8212	5.7939	5.8886
BDE	0.0499	21.2415	24.5727

Table 8. Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on flower image

	Proposed algorithm	K-mean clustering	Region-growing
PRI	0.9855	0.2709	0.3474
GCE	0.9608	0.9870	0.9756
VOI	10.2961	6.4517	6.3213
BDE	0.0159	4.1393	2.9594

Table 9. Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on bird image

	Proposed algorithm	K-mean clustering	Region-growing
PRI	0.9721	0.3357	0.1501
GCE	0.8747	0.9701	0.9731
VOI	6.7815	5.6763	6.0267
BDE	0.0992	7.7076	10.4199

Table 10. Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on tree image

	Proposed algorithm	K-mean clustering	Region-growing
PRI	0.9838	0.4743	0.0914
GCE	0.9564	0.9804	0.9872
VOI	9.7897	5.9191	6.6138
BDE	0.0189	4.2716	10.2072

Table 11. Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on pig image

	Proposed algorithm	K-mean clustering	Region-growing
PRI	0.9868	0.3047	0.4228
GCE	0.9682	0.9879	0.9815
VOI	10.5748	6.4867	6.2959
BDE	0.0096	3.0164	1.8052

Table 12. Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on building image

	Proposed algorithm	K-mean clustering	Region-growing
PRI	0.9869	0.3565	0.1690
GCE	0.9540	0.9867	0.9855
VOI	10.1100	6.3905	6.7290
BDE	0.0287	4.6359	12.4248
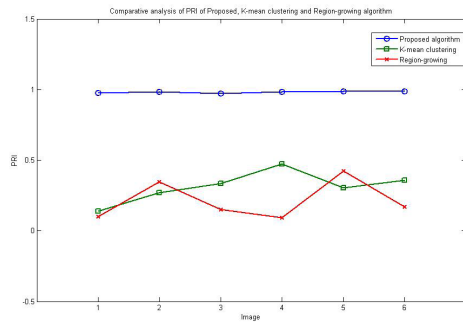
Figure 8. Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on PRI value.
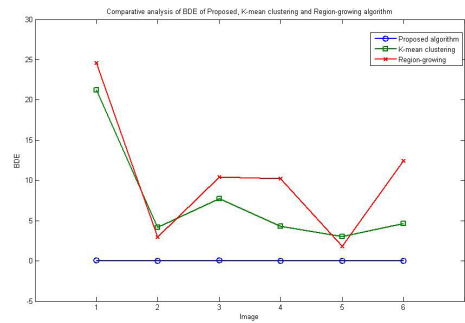

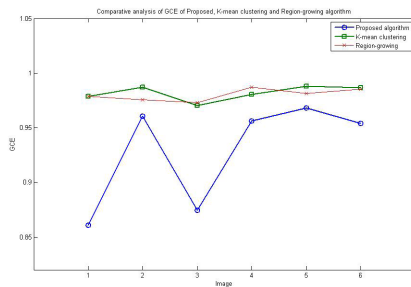**Figure 11.** Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on BDE value.

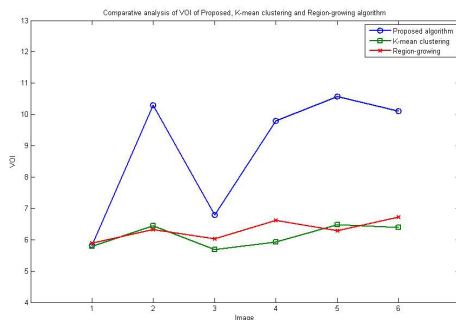**Figure 9.** Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on GCE value.

Figure 10. Comparative analysis of proposed, K-mean clustering and Region-growing algorithm on VOI value.
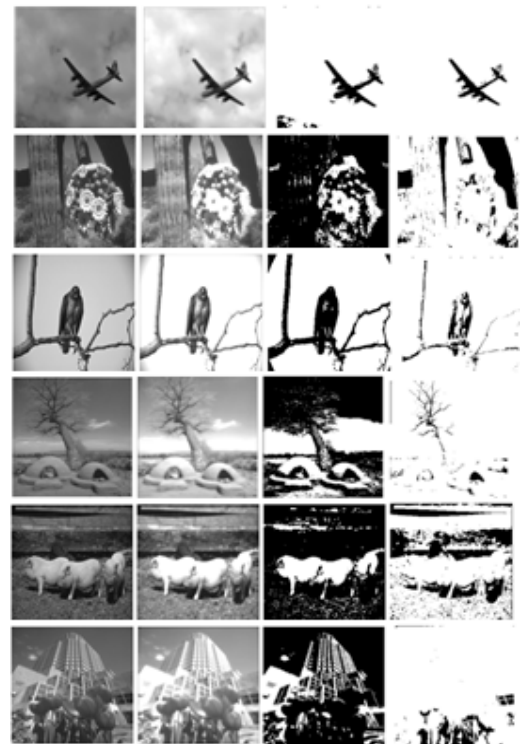

Figure 12. (a) Original image, (b) Proposed algorithm, (c) K-mean clustering, (d) Region-growing.

and 16384 region segmented images from the original image. The PRI value of 16384 region-segmented image is marginally higher than that of other segmented images. Lower GCE value was obtained in case of 16384 region-segmented image. The VOI and BDE value of 16384 region-segmented image is low as compared to other segmented images. The experiment has been conducted on more than 100 images, taken from Berkeley image database. The percentage difference between 16384 region-segmented image and other segmented images were calculated; 16384 region segmented image has shown better results as compared to other segmented images. Authors also calculated the time-complexity of the proposed algorithm which has been obtained $O(2^n)$. Further, the proposed algorithm has been compared with existing K-mean clustering and Region-growing algorithm. K-mean clustering does not work well with non-globular clusters. The problem with region-growing is the formulation of a stopping rule. The PRI value of region-growing algorithm is 29% and 73.5% better compared to K-mean clustering and proposed algorithm respectively. In case of VOI value, K-mean clustering algorithm produced 33.37% and 2.97% better results compared to proposed and Region-growing algorithm respectively. Proposed algorithm showed 4.29% and 2.29% better results in case of GCE value when compared with K-mean clustering and Region-growing algorithm respectively. In case of BDE value, authors obtained 59.43% and 67.76% better results when compared with K-mean clustering and Region-growing algorithm respectively. Thus, the proposed algorithm is proved to be better comparing the existing K-mean clustering and Region-growing algorithm.

6. References

1. Tianhua C. digital image processing. Beijing: Tsinghua University Press; 2007.
2. Udupa JK, Leblanc VR, Zhuge Y, Imielinska C, Schmidt H, Currie LM, Hirsch BE, Woodburn J. A framework for evaluating image segmentation algorithms. Comput Med Imag Grap. 2006; 30(2):75–87.
3. Kaur J, Agrawal S, Vig R. A methodology for the performance analysis of cluster based image segmentation. Int J Eng Res Appl. 2012; 2(2):664–67.
4. Raja SVK, Abdul Khadir AS, Ahamed S SR. Moving toward Region-Based Image Segmentation Techniques: A Study. J Theor Appl Inf Technol. 2009; 5(1):81–7.
5. Sathya B, Manavalan R. Image Segmentation by Clustering Methods: Performance Analysis. Int J Comput Appl. 2011; 29(11):27–32.
6. Kanungo T, Mount, DM, Piatko CD, Silverman, R, Wu, AY. An Efficient k-Means Clustering Algorithm: Analysis and Implementation. IEEE Trans Pattern Anal Mach Intell. 2002; 24(7):881–92.
7. Bin Samma AS, Abdul Salam R. Adaptation of K-means algorithm for image segmentation. Int J Eng Res Appl. 2009; 5(4):58–62.
8. Chitade AZ, Katiyar SK. Colour based image segmentation using K-means clustering. Int J Eng Res Appl. 2010; 2(10):5319–25.
9. Arbelaez P, Maire M, Fowlkes C, Malik J. Contour Detection and Hierarchical Image Segmentation. IEEE TPAMI, 2011; 33(5):898–916.