PROBABILISTIC MODEL BASED IMAGE SEGMENTATION

Mrudula Karande[1] and Prof. D. B. Kshirsagar[2]

[1]Department of Information Technology, K. K. Wagh Polytechnic, Nashik, India
[2] Department of Computer Engg, S.R. E.S. COE, Kopargaon, India

ABSTRACT

There exists a plethora of algorithms to perform image segmentation and there are several issues related to execution time of these algorithms. Image Segmentation is nothing but label relabeling problem under probability framework. To estimate the label configuration, an iterative optimization scheme is implemented to alternately carry out the maximum a posteriori (MAP) estimation and the maximum likelihood (ML) estimations. In this paper this technique is modified in such a way so that it performs segmentation within stipulated time period. The extensive experiments shows that the results obtained are comparable with existing algorithms. This algorithm performs faster execution than the existing algorithm to give automatic segmentation without any human intervention. Its result match image edges very closer to human perception.

KEYWORDS

Maximum a Posteriori, Maximum Likelihood, graphcut

1. INTRODUCTION

Image Segmentation is part of Image analysis which leads us to automated comprehension of the image by the computer. There has been tremendous work done in the field of Image analysis. Many researchers have developed numerous algorithms to achieve segmentation but till this date no algorithm has surpassed the segmentation performed by the humans. Also there are issues regarding the execution time of these algorithms. Since tremendous amount of time is spent in performing the various complex tasks it takes more time. There has been always a quest for segmentation algorithm which will work with all types of images and give good performance. In this paper we had modified the pixel relabeling algorithm in such a way it ultimately leads to faster execution which gives comparable results with the original existing algorithm [1].

1.1. Related Work

Available image segmentation algorithms can be classified into two groups: contour-based approaches and region-based approaches. Contour-based approaches try to find the boundaries of objects in an image, while region-based approaches attempt to split an image into connected regions. In contour-based approach we generally start with some spline curve and we refine it by shrink and expansion operations minimizing energy function. One problem existing in these algorithms is that they are easy to get trapped in local minima. In addition, they need manually specified initial curves close to the objects of interest. Region-based approaches try to classify an image into multiple consistent regions or classes. Thresholding is the simplest segmentation method but its performance is usually far from satisfactory.

Watershed segmentation [2] is one of the traditional region-based approaches. It is used for images containing touching objects. It finds high intensity regions and low intensity regions. It suffers from over-segmentation. The various morphological operations are used to handle this problem. Usually, watershed is used for the segmentation of foreground and background (two-class) of an image. For a general color image with many different regions, it often gives a bad result. Hence it is not used widely.

The K-means algorithm [3] is the basic one. However, the K-means is not good enough because it does not take account of the spatial proximity of pixels. It is, thus, often used in the initialization step for other approaches.

Expectation-maximization (EM) [4] performs segmentation by finding a Gaussian mixture model in an image feature space. EM is not suitable for images containing different number of regions. The disadvantage of EM is that it does not change the number of regions during the segmentation, which leads to wrong segmentation. Theoretically, the minimum description length (MDL) principle [4] can be used to alleviate this problem, but the segmentation has to be carried out many times with different region numbers to find the best result. This takes a large amount of computation, and the theoretically best result may not accord with this perception.

In [5], a mean shift algorithm is proposed for image segmentation. Mean shift is a nonparametric clustering technique which neither requires to know the number of clusters in advance nor constrains the shapes of the clusters However, it often obtains over-segmented results for many natural images.

Recently, a number of graph-based approaches are developed for image segmentation. Shi and Malik's [6] normalized cuts are able to capture intuitively salient parts in an image. Normalized cuts are one of the popular spectral clustering algorithms. Normalized cuts are not suitable for image segmentation because adhoc approximations are to be considered to relax the NP-hard computational problem. These vague approximations are ambiguous leading to unsatisfactory results. Also, due to this, spectral clustering algorithms suffer from the expensive computational cost.

Another popular segmentation approach based upon MRFs is graphcut algorithm [7]. This algorithm relies on human interaction, and solves the two-class segmentation problem only, i.e., separating an image into only background and object regions, with some manually given seed points.

In [9], authors have used Fuzzy Rule based graphcut to achieve perfect segmentation. This method definitely gives better results but is time consuming for segmenting large number of images.

All of the above techniques have their advantages and disadvantages. Some techniques suffer from over-segmentation while some of the techniques suffer from under-segmentation. The MAP-ML [1] algorithm overcomes the disadvantages in above algorithms and gives result more closely to human perception.

We are going to implement the MAP-ML algorithm on the Berkeley database containing 500 natural images of size 321 x 481 (or 481x321), with ground truth segmentation results obtained from human subjects for evaluating segmentation algorithm and we will compare the results with those obtained by state-of-the-art image segmentation algorithms such as Mean Shift and Normalized Cuts. Section 2 introduces the probability framework used in the algorithm. Section 3 discusses the proposed modified MAP-ML Algorithm. Section 4 discusses the results obtained. Section 5 concludes our work.

2. PROBABILISTIC MODEL

For a given image P, the features of every pixel p are expressed by a 4-D vector

$$I(p)=(I_L(p), I_a(p), I_b(p), I_t(p))^T \tag{1}$$

where $I_L(p), I_a(p), I_b(p)$ are the components of p in the L*a*b*color space, and $I_t(p)$ denotes the texture feature of p. In this seminar, the texture contrast defined in [2] (scaled from [0, 1] to [0,255]) is chosen as the texture descriptor. Fig. 3.4 shows an example of the features.

The task of image segmentation is to group the pixels of an image into relevant regions. If the problem is formulated as a labeling problem, the objective is then to find a label configuration $f = \{f_p | p\}$ where f_p is the label of pixel p denoting which region this pixel is grouped into. Generally speaking, a "good" segmentation means that the pixels within a region i should share homogeneous features represented by a vector $\varphi(i)$ that does not change rapidly except on the region boundaries. The introduction of $\varphi(i)$ allows the description of a region, with which high level knowledge or learned information can be incorporated into the segmentation. Suppose that there are k possible region labels.

A 4-D vector

$$\varphi(i) = (\bar{I}_L(i), \bar{I}_a(i), \bar{I}_b(i), \bar{I}_t(i))^T \tag{2}$$

is used to describe the properties of label (region), where the four components of $\varphi(i)$ have the similar meanings to those of the corresponding four components of I(p).

Let $\varphi - \{\varphi(i)\}$ be the union of the region features. If P and φ are known, the segmentation is to find an optimal label configuration $\hat{f}$, which maximizes the posterior possibility of the label configuration.

$$\hat{f} = \underset{f}{arg\ max}\ Pr\ (f\,|\,\varphi, P) \tag{3}$$

where φ can be obtained by either a learning process or an initialized estimation. However, due to the existence of noise and diverse objects in different images, it is difficult to obtain φ that is precise enough. Thus, an iterative method is used to solve the segmentation problem.

Suppose that φ^n and f^n are the estimation results in the nth iteration. Then the iterative formulas for optimization are defined as

$$f^{n+1} = \underset{f}{arg\ max}\ Pr\ (f\,|\,\varphi^n, P) \tag{4}$$

$$\varphi^{n+1} = \underset{\varphi}{arg\ max}\ Pr\ (f^{n+1}\,|\,\varphi, P) \tag{5}$$

This iterative optimization is preferred because (4) can be solved by the MAP estimation, and (5) by the ML estimation.

2.1. MAP Estimation

Given an image P and the potential region features φ, f is inferred by the Bayesian law, i.e., $Pr(f|\varphi, P)$ can be obtained by

$$\Pr(f|\varphi, P) = \frac{\Pr(\varphi, P|f)\Pr(f)}{\Pr(\psi, P)} \propto Pr(\varphi, P|f)\,Pr(f) \tag{6}$$

which is a MAP estimation problem and can be modeled using MRFs.

Assuming that the observation of the image follows an independent identical distribution, $\Pr(\varphi, P|f)$ is defined as

$$\Pr(\varphi, P|f) \propto \prod_{p \in P} \exp\left(-D(p, f_p, \varphi)\right) \tag{7}$$

where $D(p, f_p, \varphi)$ is the data penalty function which imposes the penalty of a pixel p with a label f_p for given φ. The data penalty function is defined as

$$D(p, f_p, \varphi) = \|I(p) - \varphi(f_p)\|^2$$
$$= (I_L(p) - \bar{I}_L(f_p))^2 + (I_a(p) - \bar{I}_a(f_p))^2 + (I_b(p) - \bar{I}_b(f_p))^2 + (I_t(p) - \bar{I}_t(f_p))^2 \tag{8}$$

MRF's whose clique potentials involve pairs of neighboring pixels only is considered. Thus

$$\Pr(f) \propto \exp\left(-\sum_{p \in P}\sum_{q \in N(p)} V_{p,q}(f_p, f_q)\right) \tag{9}$$

where $N(p)$ is the neighborhood of pixel p. $V_{p,q}(f_p, f_q)$, called the smoothness penalty function, is a clique potential function, which describes the prior probability of a particular label configuration with the elements of the clique(p, q). The smoothness penalty function is defined as follows using a generalized Potts model [7]:

$$V_{p,q}(f_p, f_q) = c \cdot \exp\left(\frac{-\Delta(p,q)}{\sigma}\right) \cdot T(f_p \neq f_q) = c \cdot \exp\left(\frac{-|I_L(p) - I_L(q)|}{\sigma}\right) \cdot T(f_p \neq f_q) \tag{10}$$

where $\Delta(p, q) = -|I_L(p) - I_L(q)|$, called brightness contrast, denotes how different the brightnesses of p and q, c>0 is a smoothness factor, $\sigma > 0$ is used to control the contribution of $\Delta(p, q)$ to the penalty, and T(.) is 1 if its argument is true and 0 otherwise. $V_{p,q}(f_p, f_q)$, depicts two kinds of constraints. The first enforces the spatial smoothness; if two neighboring pixels are labeled differently, a penalty is imposed. The second considers a possible edge between p and q; if two neighboring pixels cause a larger Δ, then they have greater likelihood to be partitioned into two regions.

In this algorithm, the boundaries of the segmentation result are pulled to match the darker pixels which are more likely to be edge pixels.

From (6), (7), and (9) the equation can be written as,

$$\Pr(f|\varphi, P) \propto \left(\prod_{p \in P} \exp\left(-D(p, f_p, \varphi)\right)\right) \cdot \exp\left(-\sum_{p \in P}\sum_{q \in N(p)} V_{p,q}(f_p, f_q)\right) \tag{11}$$

Taking the logarithm of (11), the following energy function is as:

$$E(f, \varphi) = \sum_{p \in P} D(p, f_p, \varphi) + \sum_{p \in P}\sum_{q \in N(p)} V_{p,q}(f_p, f_q) \tag{12}$$

where $E(f, \varphi) \propto -\log \Pr(f|\varphi, P)$. It includes two parts: the data term

$$E_{data} = \sum_{p \in P} D(p, f_p, \varphi) \tag{13}$$

and the smoothness term

$$E_{smooth} = \sum_{p \in P}\sum_{q \in N(p)} V_{p,q}(f_p, f_q) \tag{14}$$

From (12), it is clear that maximizing $\Pr(f|\varphi, P)$ is equivalent to minimizing the Markov energy $E(f, \varphi)$ for a given φ. In this paper, graphcut algorithm is used to solve this minimization problem.

2.3. ML Estimation

A 4-D vector $\varphi(i)$ given by equation 2 is used to describe the properties of label (region). The ML estimation $\varphi = \varphi(i)$ is obtained, where

$$\varphi(i) = \frac{1}{num_i} \sum_{f_p = i} I(p) \tag{15}$$

with num_i being the number of pixels within region i. Here (15) is exactly the equation to obtain $\bar{I}_L(i)$, $\bar{I}_a(i)$, $\bar{I}_b(i)$ and $\bar{I}_t(i)$ and in (2).

3. PROPOSED MODIFIED MAP-ML ALGORITHM

The MAP-ML [1] is used to segment the image by each object in the same image. The algorithm starts with finding the texture and contrast feature of every pixel present in the image. The texture and contrast feature is used to segment the outline of the each object in the image and labelling is used to delete the unwanted portion of the image and segment each object by each color. The K-means Algorithm is used for initializations of the regions. The MAP estimation is used to detect the edges of the image and the color space is used to segment the images by colors. The graph cut algorithm is an unsupervised algorithm used for over segmentation and computation problem. We had modified the existing MAP-ML [1] algorithm and the modified algorithm is given below:

Algorithm: Modified MAP-ML Image Segmentation:
Input: an RGB color image.
Step 1: Convert the image into L*a*b* space and calculate the texture contrast [4].
Step 2: Use the K-means algorithm to initialize features of region φ
Step 3: Find the standard deviation σ (camera noise) for the image
Step 4: Iterative optimization.
4.1: MAP estimation-
Estimate the label configuration f based upon current φ using the graph cut algorithm [8].
4.2: Relabeling-
Remove small regions (same label configuration) having less than 100 pixels. Reorder the labels in proper sequence, obtaining a new f
4.3: ML estimation-
Refine φ based upon current f
Step 5: If f do not change between two successive iterations or the maximum number of iterations is reached, go to the output step; otherwise, go to step 4.
Output: Multiple segmented regions of the image.

After step 4.1, it is possible that two non adjacent regions are given the same label. The MAP estimation is an NP-hard problem. Boykov et al. [8] proposed to obtain an approximate solution via finding the minimum cuts in a graph model. Minimum cuts can be obtained by computing the maximum flow between the terminals of the graph. In [8], an efficient Maxflow algorithm is given for solving the binary labelling problem. In addition, an algorithm, called α expansion with the Maxflow algorithm embedded, is presented to carry out multiple labelling iteratively. In this algorithm, the α expansion algorithm is used to perform step 4.1. To increase the speed of the algorithm we had used Maxflow 3.01 algorithm.

In the original MAP-ML Algorithm [1], the authors had initiated the MAP-ML algorithm with default 10 labels and then in the iteration each region is labelled uniquely. Since the number of labels is unique and increases with each iteration, the time to execute the MAP Estimation goes up. So instead of that we had kept the initial number of labels=10 by default but we had not uniquely labelled the regions so thereby the image will have utmost 10 or less than 10 labels

hence the time to take the MAP Estimation is less comparative to original MAP-ML Algorithm. To achieve the equivalent result as the original we had calculated the standard deviation (camera noise) for each image automatically since it will be different for each image. It is calculated by taking expectation of all the pairs of neighbors in an image. So we had obtained results as near as possible to the original algorithm in less amount of time.

Briefly we can say that the modified algorithm has three enhancements over Original MAP-ML:

1) Use of Maxflow 3.01 Algorithm with the reuse trees option
2) Unlike original algorithm the regions are not labelled uniquely
3) For every image sigma (standard deviation) is calculated. Sigma is an important factor used in deciding the smoothness penalty for an image. Here it is calculated based on average value of all pairs of neighbors in an image.

4. EXPERIMENTAL RESULTS

Our algorithm is tested on the Berkeley benchmark for evaluating segmentation algorithms and compares the results with those obtained by state-of-the-art image segmentation algorithms. The Berkeley database contains 500 natural images of size 321 x 481 (or 481 x 321), with ground truth segmentation results obtained from human subjects.

The compared algorithms in these experiments include: Mean Shift (MS) [5] and Normalized cuts (NC) [6]. In this algorithm, the initial cluster number in the K-means algorithm is set to 10 and the smoothness factor c is 100. The region number in NC is set to 20, which is the average number of segments marked by the human subjects in each image.

In the MS algorithm the default parameters used are hs=15, hr=13, and the minimal region=20 pixels are chosen. Since NC cannot handle an image of size 321 x 481(or 481 x 321) due to the overflow of the memory, all the input images for them are shrunk into a size 214 x 320 (or 320 x 214), and the segmentation results are enlarged to their original sizes.

All the above experiments had been conducted on Intel Core 2 Duo 2.2 GHz 4GB RAM Windows 7 platform. The code has been developed in JAVA which makes it portable.

4.1. Qualitative Comparison Results

The part of the images in the Berkeley benchmark is classified into 7 sets ("Landscape", "Grassplot and Sky", "Craft", "Human", "Bird", "Felid" and "Buildings"), and show the segmentation results obtained by the three algorithms in Figure 1-7.
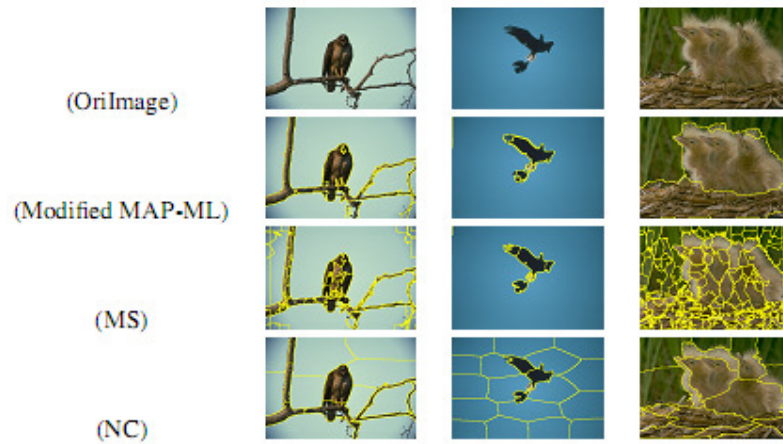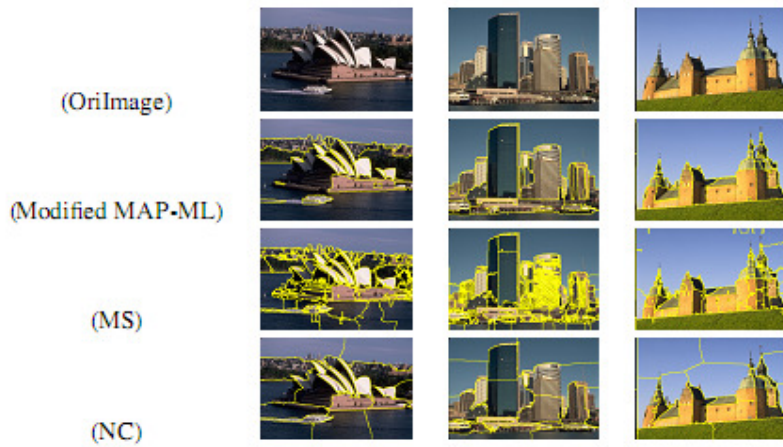
(OriImage)

(Modified MAP-ML)

(MS)

(NC)

Figure 1. Results obtained on "Bird" images

(OriImage)

(Modified MAP-ML)

(MS)

(NC)

Figure 2. Results obtained on "Buildings" images

(OriImage)

(Modified MAP-ML)

(MS)

(NC)

Figure 3. Results obtained on "Feline" images

17

Figure 4. Results obtained on "Craft" images

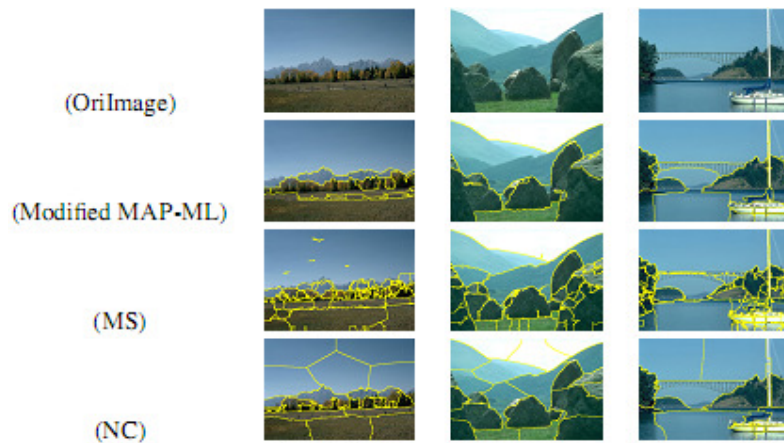Figure 5. Results obtained on "GrassPlot and Sky" images
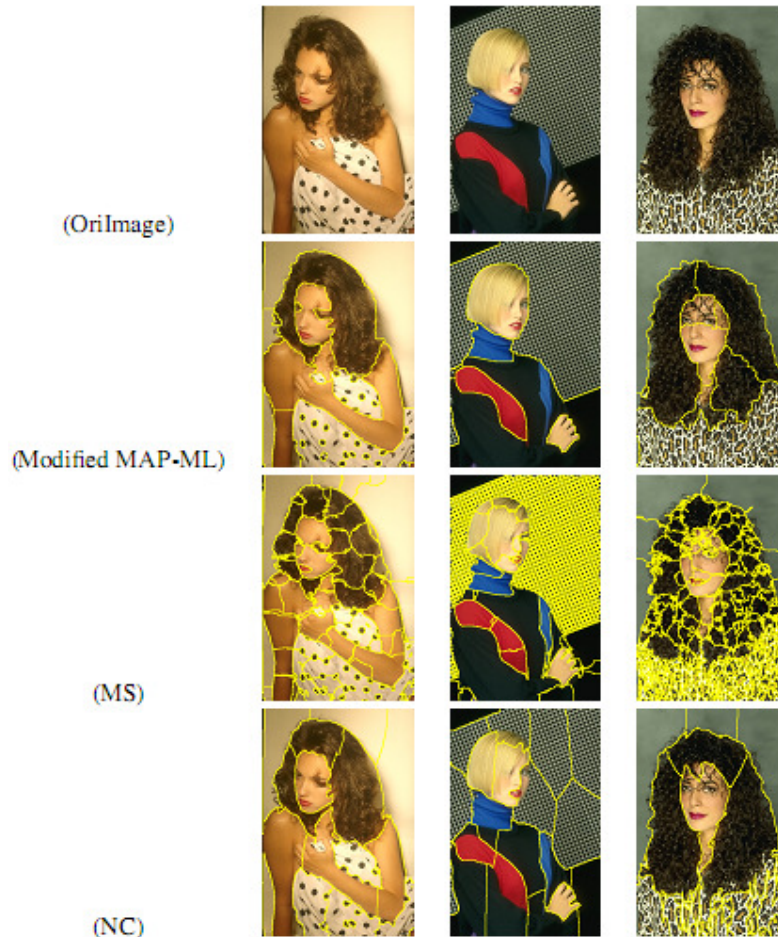

Figure 6. Results obtained on "Landscape" images

Figure 7. Results obtained on "Humans" images

From these examples, the following observations are seen:

NC tends to partition an image into regions of similar sizes, resulting in the region boundaries different from the real edges. MS give strongly over-segmented results. Compared with these other algorithms, it is easy to see that this algorithm obtains the best results, in which the generated boundaries match the real edges well and the segmented regions are in accordance with human perception.

4.2. Quantitative Comparisons Results

Quantitative comparisons are also important for objectively evaluating the performance of the algorithms. There have been several measures proposed for this purpose. Region differencing and boundary matching are two of them. Region differencing measures the extent to which one segmentation can be viewed as a refinement of the other. Boundary matching measures the average displacement error of boundary pixels between the results obtained by an algorithm and the results obtained from human subjects. However, these two measures are not good enough for segmentation evaluation. For example, a segmentation result with each pixel being one region obtains the best score using these two measures. A strongly over-segmented result, which does not make sense to human visual perception, may be ranked good.

In these experiments, two more stable and significant measures, variation of information (VoI) and probabilistic rand index (PRI) are used to compare the performances of the three algorithms, to objectively evaluate image segmentation algorithms.

VoI is an information-based measure which computes a measure of information content in each of the segmentations and how much information one segmentation gives about the other. It is related to the conditional entropies between the region label distributions of the segmentations. PRI compares an obtained segmentation result with multiple ground truth images through soft non uniform weighting of pixel pairs as a function of the variability in the ground truth set. The value of VoI falls in $[0,\infty]$, and the smaller, the better. The value of PRI is in $[0,1]$, and the larger, the better.

The average values of PRI and VoI for the three algorithms are given in Table 1. In this table, the second column shows the average PRI and VoI values between different human subjects, which are the best scores. From these results, one can see that this algorithm outperforms the other algorithms because it obtains the smallest VoI value and the largest PRI value. Among other algorithms, MS gives sometimes better PRI values to this algorithm. However, their VoI values are much larger than algorithm.

Table 1. Average Values of PRI and VOI on the images.

	Human	Modified MAP-ML	Modified MAP-ML	NC	MS
		10 Labels	20 Labels		
PRI	0.8961	0.7876	0.7954	0.7501	0.7769
VOI	0.9219	2.220	2.215	2.8327	3.747

To demonstrate the performances of these algorithms on each image, the PRI and VOI curves are shown in Figure 8 (default 10 labels) and Figure 9 (default 20 labels).   It is clearly observed that modified algorithm performs the best. There is slight trade off between speed and accuracy in the modified MAP-ML Algorithm. The elapsed time calculated between original MAP-ML and modified MAP-ML Algorithm is shown in Figure 10.
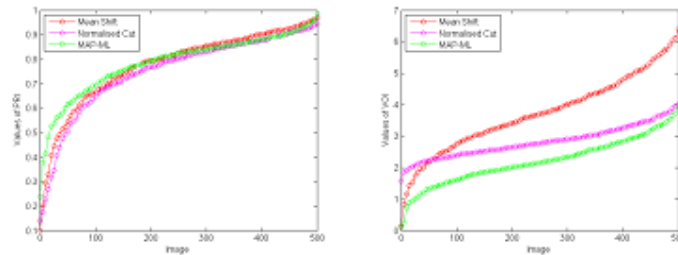

Figure 8. PRI and VOI values achieved on individual images by the three algorithms when default labels are 10. The values are plotted in increasing order.
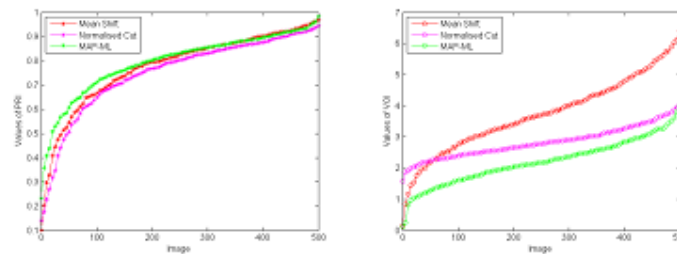
Figure 9. PRI and VOI values achieved on individual images by the three algorithms when default labels are 20. The values are plotted in increasing order.
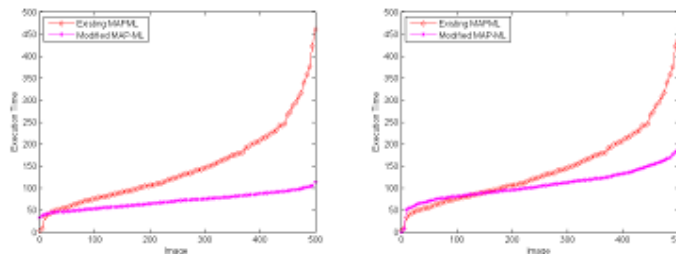

Figure 10. Elapsed time Comparison between Original MAP-ML and Modified MAP-ML Algorithm when default labels are a) 10 and b) 20. The values are plotted in increasing order.

5. CONCLUSION

We had implemented our modified MAP-ML algorithm which gives comparable results with the original MAP-ML algorithm performing the image segmentation. Thus from the experimental results we had successfully shown that the modified MAP-ML algorithm takes less time to execute as compared to the original MAP-ML algorithm giving nearly same results as the original algorithm.

REFERENCES

[1]   Shifeng Chen, Liangliang Cao, Yueming Wang, Jianzhuang Liu (September 2010), "Image Segmentation by MAP-ML Estimations", IEEE Trans. on Image Processing, vol. 19, No. 9, pp. 2254-2264.

[2]   L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations", IEEE Trans. Pattern Anal. Mach. Intell., vol.13, no. 6, pp. 583-598, Jun.1991.

[3]   R. Duda, P. Hart, and D. Stork, "Pattern Classification", 2nd ed. Hoboken, NJ: Wiley, 2001.

[4]   C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image segmentation using expectation-maximization and its application to image querying", IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 8, pp. 1026-1038, Aug. 2002.

[5]   D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis", IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603-619, May 2002.

[6]   J. Shi and J. Malik, "Normalized cuts and image segmentation", IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 888-905, Aug. 2000.

[7]   V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?", IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 2, pp. 147-159, Feb. 2004.

[8] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts", IEEE Trans. Pattern Anal. Mach. Intell., vol.23, no. 11, pp. 1222-1239, Nov.2001.

[9] Muhammad Rizwan, Khokher, Abdul Ghafoor, Adil Masood Siddiqui, "GraphCuts based Image Segmentation using Fuzzy Rule Based System", IEEE Radio Engineering, vol . 21, no. 4, Dec 2012, pp. 1236-1245.

AUTHORS

Mrudula Karande received the B.E. (Comp) degree from the Nagpur University of India and the M.E.(Comp. Engg) degree from the Pune University, in first class. She is working as the Head of the Department of Information Technology in K. K. Wagh Polytechnic, Nashik India. Her research interests include image processing.

Prof. D. B. Kshirsagar received the B.E. (CSE), Computer Engineering, from the Walchand College of Engineering, Sangli, M.E. (CSE), Engineering degree from Shivaji University in first class with distinction and is currently pursuing Phd. He is working as the Prof. and the Head of the Department of Computer Engg in S. R. E. S. COE Kopargaon, India. His research interests include image processing.