Modern Education
and Computer Science
PRE∫∫

A New Enhanced semi supervised image segmentation using Marker as Prior information.

L.Sankari [1]
[1].Asst Professor, Department of Computer Science,
Sri Ramakrishna College of Arts and Science for women, Coimbatore, India
E-mail: sankarivnm@gmail.com

Dr.C.Chandrasekar [2]
[2].Associate Professor , Dept of Computer Science, Periyar University, Salem, India
E-mail: ccsekar@gmail.com

Abstract— **In Recent days Semi supervised image segmentation techniques play a noteworthy role in image processing. Semi supervised image segmentation needs both labeled data and unlabeled data. It means that a Small amount of human assistance or Prior information is given during clustering process. This paper discusses an enhanced semi supervised image segmentation method from labeled image. It uses both a background selection marker and fore ground object selection marker separately. The EM (Expectation Maximization) algorithm is used for clustering along with must link constraints. The proposed method is applied for natural images using MATLAB 7. Thus the proposed method extracts Object of Interest (OOI) from OONI (Object of Not Interest) efficiently and the experimental results are compared with Standard K Means and EM Algorithm also. The results show that the proposed system gives better results than the other two methods. It may also be suitable for object extraction from natural images and medical image analysis.**

Index Terms— *Semi supervised image segmentation - prior knowledge – constrained clustering.*

I. INTRODUCTION

Image segmentation is the method of dividing an image into different regions such that each region is homogeneous. By partitioning an image into a set of disjoint segments, image segmentation leads to more compact image representation. As the central step in computer vision and image understanding, image segmentation has been extensively investigated in the past decades, with a large number of image segmentation algorithms. There are number of segmentation techniques exist in the literature. But no single method can be considered best for all kind of images. Most of the techniques are being pretty ad hoc in nature.

A)Need of semi supervised image segmentation:
Semi supervised method is the combination of both supervised (classification) and unsupervised (clustering) and classification concept. Before performing clustering some prior knowledge is given. If the algorithm is purely an unsupervised (clustering) algorithm it will not show good result for all kind of images since an iterative

clustering algorithms commonly do not lead to optimal cluster solutions. Partitions that are generated by these algorithms are known to be sensitive to the initial partitions that are fed as an input parameter. A "good" selection of initial seed is an important clustering problem. Likewise the classification algorithm will not give best solution since the result depends on type of classifier. So this paper discuss about combination of these two methods called 'semi supervised model' for image segmentation.

The following section discuss about semi supervised model[13][14].

b.)Semi supervised clustering:
During clustering process a small amount of prior knowledge is given either as labels or constraints or any other prior information. The following figure explains about the semi supervised clustering model[15][16][17].
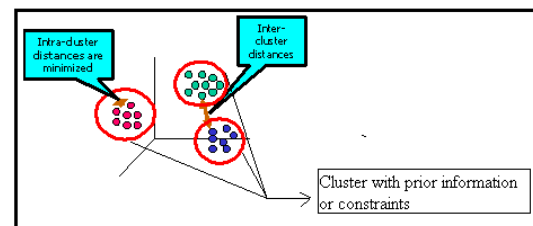

Figure 1. Semi supervised clustering model

In the above figure, the three clusters are formed using certain constraints or prior information. Besides the similarity information which is used as color knowledge, the other kind of knowledge is also available by either pair wise (must-link or cannot-link) constraints between data items or class labels for some items. Instead of simply using this knowledge for the external validation of the results of clustering, one can imagine letting it "guide" or "adjust" the clustering process, i.e. provide a limited form of supervision. There are two ways to provide information for semi supervised clustering.

1. **Search based.**
2. **Similarity based.**

1. Search based :

The clustering algorithm itself is modified so that user-provided constraints or labels can be used to bias the search for an appropriate clustering. This can be done in several ways, such as by performing a transitive closure of the constraints and using them to initialize clusters [4], by including the cost function a penalty for lack of compliance with the specified constraints [10][11], or by requiring constraints to be satisfied during cluster assignment in the clustering process [12].

2. Similarity based:

There are several similarity measures existing in the domain. Any one similarity measure is adapted[7][8][9] so that the given constraints can be easily satisfied.

In this paper semi supervised image segmentation with minimum user label is discussed. Instead of selecting some sample pixels with mouse clicks [1], a group of pixels are selected as a region using mouse. A pixel which has the same color and intensity as in the selected marker region (by mouse selection) will come under one cluster and others pixels will not be. This concept is given in detail in the following sections.

I.   PREVIOUS RELATED WORKS

Recently there are many papers focusing the importance of semi supervised image segmentation. Among them a few papers are analyzed. According to paper [3], the semi-supervised C-Means algorithm is introduced in this paper to solve three problems in the domains like choosing and validating the correct number of clusters, Insuring that algorithmic labels correspond to meaningful physical labels tendency to recommend solutions that equalize cluster populations. The algorithm used MRI brain image for segmentation.

In this [4] paper, how the popular k-means Clustering algorithm can be modified to make use of the available information with some artificial constraints. This method was implemented for six datasets and it has showed good improvement in clustering accuracy. This method was also applied to the real world problem of automatically detecting road lanes from GPS data and observed dramatic increases in performance.

In paper [5] a novel semi-supervised Fuzzy C-means algorithm is proposed. A set called as seed set which contains a small amount of labeled data is used. First, an initial partition in the seed set is done, then use the center of each partition as the cluster center and optimize the objective function of FCM using EM algorithm. Experiments results show that the defect of fuzzy c-means is avoided that is sensitive to the initial centers partly and give much better partition accuracy.

In Paper [6], Semi-supervised clustering uses a small amount of labeled data to aid and bias the clustering of unlabeled data. Here   labeled data is used to generate initial seed clusters along with the constraints generated from labeled data to guide the clustering process. It introduces two semi-supervised variants of KMeans clustering that can be viewed as instances of the EM algorithm, where labeled data provides prior information about the conditional distributions of hidden category labels. Experimental results demonstrate the advantages of these methods over standard random seeding and COP-KMeans, a previously developed semi-supervised clustering algorithm.

This paper [12] focuses on semi-supervised clustering, where the goal is to cluster a set of data-points given a set of similar/dissimilar examples. Along with instance-level equivalence (similar pairs belong to the same cluster) and in-equivalence constraints (dissimilar pairs belong to different clusters) feature space level constraints (how similar are two regions in feature space) are also used for getting final clustering. This task is accomplished by learning distance metrics (i.e., how similar are two regions in the feature space?) over the feature space which that are guided by the instance-level. A bag of words models, which are nothing but code words (or visual-words) are used as building blocks. Our proposed technique learns non-parametric distance metrics over codewords from these equivalence (and optionally, in-equivalence) constraints, which are then able to propagate back to compute a dissimilarity measure between any two points in the feature space. Thus this work is more advanced than previous works. First, unlike past efforts on global distance metric learning which try to transform the entire feature space so that similar pairs are close. This transformation is non-parametric and thus allows arbitrary non-linear deformations of the feature space. Second, while most Mahalanobis metrics are learnt using Semi-Definite Programming (SDP), this paper discuss about a Linear Program (LP) and in practice, is extremely fast. Finally, Corel image datasets (MSRC, Corel) where ground-truth segmentation is available. Over all, this idea gives improved clustering accuracy.

II.   METHODOLOGY

In this paper, ground truth image is taken with proper class labels. The Octree color quantization algorithm is applied to get the reduced colors. This color table is integrated with must link constraints for the given image using EM algorithm.

A group of pixels as a region must be selected separately for object of interest (OOI) and object of not interest (OONI) for back ground and foreground from an input image. Here OOI and OONI refer foreground and back ground objects respectively. Find the smallest

distance for each and every pixel in the marker to its neighboring pixels using Mahalanobi's formula.

$$dist2_{i,j} = (x_i - x_j) * COV^{-1}(x_i - x_j)$$ ---- (1)

Where

x   row vector which contains the pixels inside a marker and the other area.

COV   represents sample covariance matrix.

If this distance is less than the assumed threshold value (0.5) then find the exact color index. If the index values are same then assign to the same region otherwise (same value & different class index) delete any one index and group. The above process is repeated for all pixels in the marker. Finally the object of interest (OOI) is clearly segmented than the other methods.
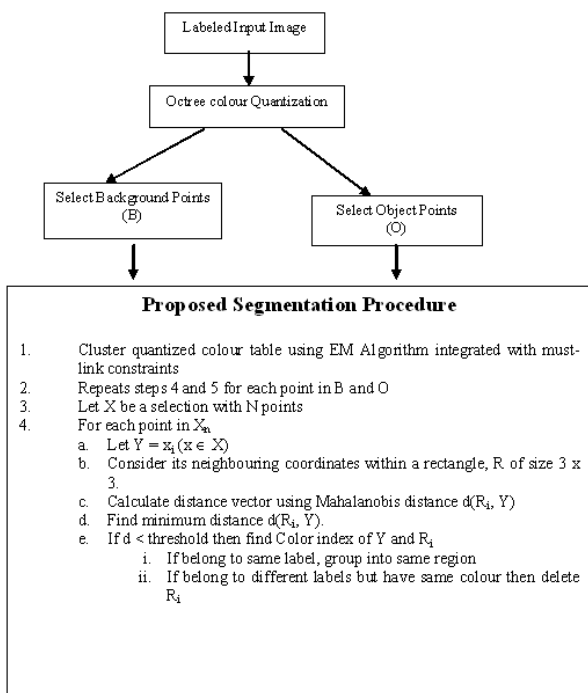

Fig 1. Work Flow Diagram of the proposed work

The algorithm based on proposed idea is given below:

1. Labeled image is taken as input image.
2. Octree color quantization to reduce the colors and store with class label.

3. Mark object of interest (OOI) and object of not interest (OONI) using mouse.
4. Cluster quantized color table using EM Algorithm interated with must-link constraints. Must link constraints means that if any point

belongs to same region, group them into one region otherwise need not group.

5. Repeats steps 4 and 5 for each point in object of interest(OOI) and object of not interest(OONI)
6. Let X be a selection with N points
7. For each point in X_n
   a. Let $Y = x_i$ ($x \in X$)
   b. Consider its neighboring coordinates within a rectangle, R of size 3 x 3.
   c. Calculate distance vector using Mahalanobis distance $d(R_i, Y)$
   d. Find minimum distance $d(R_i, Y)$.
   e. If d < threshold then find Color index of Y and R_i
      i. If belong to same label, group into same region
      ii. If belong to different labels but have same color then delete R_i

III. RESULTS AND DISCUSSIONS

Figure 2:

Here three different labeled images are taken as input . These three images are partially separated images.

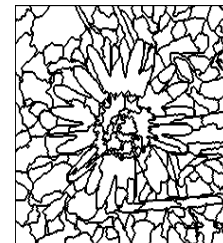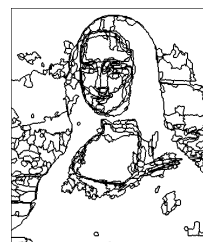

Fig 2.a   Fig 2.b

Fig 2.c

Figure 3:

Marking of OOI (Object of interest) and OONI (Object of not interest) for all the above three figures using blue and green color.

Object of intereste (Green Color mark)
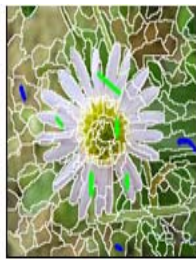
Object of Not interest (Blue color mark)

Fig 3.a   Fig 3.

Fig 3.c

Figure 4:
Resullt of proposed method for the above three figures.

Fig 4.a   Fig 4.b

Fig 4.c

Figure 5:
Labeled image and different marker selection.

Fig 5.a   Fig.5.b

Fig 5.c

Figure 6:
Results using proposed idea for the above figure 5 using different marker selection.

Fig 6.a   Fig 6.b

(From these two pictues the object is segmented properly)

Fig.6.c

The object of interest is not segmented properly

Figure 7:
Results of input images using K Means method.
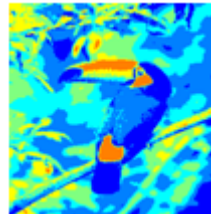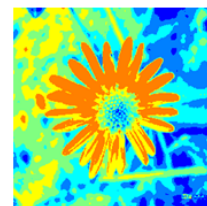

Fig 7.a   Fig 7.b
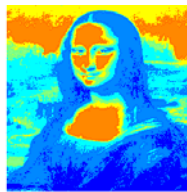

Fig 7.c

Figure 8:
Result of input images using Standard EM method.

Fig 8.a   Fig.8.b

Fig 8.c

The time taken for getting the result of the input **image1** (Bird) using proposed method is noted down in the table.

Table 1. Performance Table for image1

S No	Methods	Time in sec.
1	K Means	6.3214
2	EM	40.22405
3	Proposed	13.05598

It shows that the proposed method has taken more time than K means and less time than EM method. The performance is shown in the chart given below.
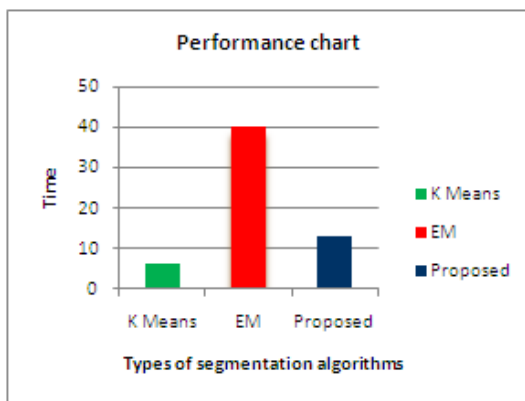

Figure 9. Performance chart

Using proposed idea any labeled image is taken as input. The two markings are given like OOI and OONI(object of interest and object of not interest) with different colors. There are two different marker selections are shown in the figure 3 and 5. According to the marker selection given, the object is extracted from its background. Compared to the results in Figure 4, the segmentation results are better in figure 6 for some pictures. This is because the quality of segmentation depends on marker selection. These proposed results are also compared with K means and Standard EM Algorithm. In Figure7 and 8, the K Means & EM Algorithms do not segment the foreground of the object accurately. But the proposed method results are better in figure 6 and the object of interest is also separated accurately.

IV. CONCLUSION

The above result shows that the proposed semi supervised segmentation extracts the object of interest precisely. But the result of segmentation depends on the marker selection on left and right side of the image. If the marker selection is not given properly, the result of segmentation will not be good. This may be eliminated in future by adding certain other constraints for texture, color etc.

REFERENCES

[1]   Yuntao Qian, Wenwu Si, IEEE, " Semi-supervised Color Image Segmentation Method"-2005

[2]   Yanhua Chen, Manjeet Rege, Ming Dong, JingHua Farshad Fotouhi Department of Computer Science Wayne State UniversityDetroit, MI48202 "Incorporating User Provided Constraints into Document Clustering",2009

[3]   Amine M. Bensaid, Lawrence O. Hall Department of Computer Science and Engineering Universit of South Florida Tampa, Partially Supervised Clustering for Image Segmentation -1994

[4]   Kiri Wagstaff, Claire Cardie ,Seth Rogers &Stefan Schroedl ,"Constrained K-means Clustering with Background Knowledge-2001"

[5]   Kunlun Li; Zheng Cao; Liping Cao; Rui Zhao; Coll. Of Electron. & Inf. Eng., Hebei Univ., Baoding, China ,"A novel semi-supervised fuzzy c-means clustering method"2009,IEEE Explorer

[6]   Sugato Basu , Arindam Banerjee , R. Mooney , In proceedings of 19th international conference on Machine Learning(ICML-2002),Semi-supervised Clustering by Seeding (2002)

[7]   David Cohn, Rich Caruana, and Andrew McCallum. Semi-supervised clustering with user feedback, 2000.

[8]   Dan Klein, Sepandar D. Kamvar, and Christopher D. Manning. From instance-level constraints to space-level constraints: Making the most of prior knowledge in data clustering. In Proceedings of the 19th International Conference on Machine Learning, pages 307–314. Morgan Kaufmann Publishers Inc., 2002.

[9]   Eric P. Xing, Andrew Y. Ng, Michael I. Jordan, and Stuart Russell. Distance metric learning with application to clustering with sideinformation. In S. Thrun S. Becker and K. Obermayer, editors, Advances in Neural Information Processing Systems 15, pages 505–512, Cambridge, MA, 2003. MIT Press.

[10]   A. Demiriz, K. Bennett, and M. Embrechts. Semi-supervised clustering using genetic algorithms. In C. H. Dagli et al., editor, Intelligent Engineering Systems Through Artificial Neural Networks 9, pages 809–814. ASME Press, 1999. [11]

[11]   K.Wagstaff and C. Cardie. Clustering with instance-level constraints. In Proceedings of the 17th International Conference on Machine Learning, pages 1103–1110, 2000.

[12]   Dhruv Batra, Rahul Sukthankar and Tsuhan Chen,"Semi-Supervised Clustering via Learnt Codeword Distances",2008.

[13]   Richard Nock, Frank Nielsen,Semi supervised statistical region refinement for color image segmentation,2005

[14]`   Jan Kohout,Czech Technical University in Prague
Faculty of Electrical Engineering,Supervisor: Ing .
Jan Urban,Semi supervised image segmentation of
biological samples-PPT,July 29, 2010

[15]   Ant´onio R. C. Paiva1 and Tolga Tasdizen,Fast Semi
Supervised image segmentation by novelty selection,2009

[16]   Kwangcheol Shin and Ajith Abraham, Two Phase Semi-
supervised ClusteringUsing Background Knowledge,2006.

[17]   M´ario A. T. Figueiredo, Dong Seon Cheng, Vittorio Murino,
Clustering Under Prior Knowledge with Application to
Image Segmentation,2005.

Biography:

Mrs L.Sankari is currently working as an Assistant Professor in the Department of Computer Science, Sri Ramakrishna College of Arts and Science for women, Coimbatore- 641 044, Tamilnadu, India. She is about 16 years of teaching experience. She has published four national and four international research papers. Her research interest area includes image processing, Data mining , Pattern classification and optimization techniques.

Dr. C. Chandrasekar received his Ph.D. degree from Periyar University, Salem.   He has been working as Associate Professor at Dept. of Computer Science, Periyar University, Salem – 636 011, TamilNadu, India. His research interest includes Wireless networking, Mobile computing, Computer Communication and Networks. He was a Research guide at various University in India.   He has been published more than 50 research papers at various National/ International Journals.