ISSN:0975-9646

Garima Singh Rawat et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (5) , 2012,5128 - 5132

Proposed Method for Image Segmentation Using Similarity Based Region Merging Techniques

Garima Singh Rawat[1] ,Joy Bhattacharjee[2], Roopali Soni[2]

[1]M. Tech. IV SEM, [2]Asst. Prof. Dept. of C.S.E.
Dept. of C.S.E.
Oriental College of Technology
Bhopal (M.P), India

Abstract- Image Segmentation is a technique that partitioned the digital image into many number of homogeneous regions or sets of homogeneous pixels is called image segmentation. Efficient and effective image segmentation is an important task in computer vision and object recognition. Since fully automatic image segmentation is usually very hard for natural images, interactive schemes with a few simple user inputs are good solutions. In this paper we presents image segmentation using similar region merging and comparison among different interactive image segmentation techniques. Segmentation should stop as object of interest in an application is isolated. The ultimate goal is to make the image more simplified and that to get more meaningful to analyze. This paper deals with the several segmentation techniques using human interactions models like PGM, Bayesian Networks Unified Graphical Models, Similar region merging etc., are taken from the literature are reviewed and a new interactive method of image segmentation is proposed using similar region merging.

Keywords:- Region Merging, Image Segmentation, Probabilistic Graphical Model (PGM), Region Merging, Unified Graphical Model (UGM).

1. INTRODUCTION:

Image segmentation is the process of partitioning an image into multiple regions or sets of homogeneous pixels. Image segmentation is to separate the desired objects from the background. These partitions are different objects in image which have the same features of image. The result of image segmentation is a number of regions that collectively cover the entire image.

In general, the color and texture features in a natural image are very complex so that the fully automatic segmentation of the object from the background is very hard. Therefore, semi-automatic segmentation methods incorporating user interactions have been proposed [1,2,3] and are becoming more and more popular. For instance, in the graph cut algorithm [4, 5, 6], the prior information obtained by the users is critical to the segmentation performance. The low level image segmentation methods, such as mean shift [7, 8], watershed [2], level set [9], usually divide the image into many small regions. Although may have severe over segmentation, these low level segmentation methods provide a good basis for the subsequent high level operations, such as region merging. Because of less over segmentation, the statistic features of each region, which will be exploited by the proposed region merging method, can be more robustly calculated and then be used in guiding the region merging process. In this paper, we did the comparison among different interactive image segmentation techniques. For example the interactive information is introduced as human interactions, which are input by the users to roughly indicate the position and main features of the object and background. In this paper we focus on region merging of initial image segmentation, it will be our proposed method which calculate the similarity of different regions and merge them based on the proposed similarity rule with the help of human interaction. The object will then be extracted from the background when the merging process ends. The key contribution of the object segmentation method is a similarity region merging (SRM) mechanism, which is adaptive to image content and does not require a preset threshold. With the region merging algorithm, the non-marker background regions will be automatically merged and labeled, while the non-marker object regions will be identified and avoided from being merged with background. Once all the non-marker regions are labeled, the object contour can then be readily extracted from the background.

Image Segmentation is a process which partitioned image into multiple unique regions, where region is set of pixels. If I is set of all image pixels, then by applying segmentation we get different-different unique regions like $\{R_1,R_2,R_3,..,R_n\}$ which when combined formed the image 'I'. Many approaches have been proposed earlier which includes region growing [7], normalized cuts [3], grab-cut method [4], active contours [5], and MRF (Markov Random Field) based approaches [10] etc.

The automatic segmentation approaches may also fail even though they use prior information in segmentation process. The main reason behind this is the complexity of segmentation of image in real applications. The complexity is due to the several reasons like shadow, low contrast areas, occlusion, cluttering and noise in the image. These reasons make the segmentation process quite difficult and challenging. The use of interactive image segmentation process is the solutions of such problems. There are many works on interactive image segmentation [2, 3, 4]. They demonstrate the usefulness of the user's intervention for improving segmentation.

Many methods have been proposed and a dense literature is available for extracting information from an image and to partition it into different regions. Probabilistic Graphical

Models (PGM) [10, 11, 12] are very powerful models for extracting features from the image. There are two basic types of graphical models: the undirected graphical model and the directed acyclic graphical model. The undirected graphical model can represent noncasual relationships among the random variables. The Markov Random Field (MRF) [13] is a type of well-studied undirected graphical model. MRF models have been widely used for image segmentation. They incorporate the spatial relationships among neighbouring labels as a Markovian prior. This prior can encourage the adjacent pixels to be classified into the same group. As an extension to MRF, the Conditional Random Field (CRF) [14] is another type of undirected graphical model that has become increasingly popular.

While both MRF and CRF models can effectively capture noncasual relationships among the random variables (i.e., the nodes in a graphical model), such as the spatial homogeneity, they cannot model some directed relationships (e.g., the causalities) that extensively exist and are also important. Fortunately, this problem can be complementarily solved by another type of graphical model, i.e., the directed acyclic graphical model such as Bayesian Network (BN) [12]. BN can conveniently model the causal relationships between random variables using directed links and conditional probabilities. It has been successfully applied to medical diagnosis systems, expert systems, decision-making systems, etc. For image segmentation, there are some relationships that can be naturally modelled as causal relationships. The existing graphical models for image segmentation are directed, undirected which can effectively capture one type of image relationship and often fail to capture the complex image relationships of different types, and combination of both directed (BN model) and undirected (CRF Model).

2. Previous Work

Lei Zhang, Zhi Zeng, and Qiang Ji [15] proposed a method to extend the Chain Graph (CG) model to with more general topology and the associated methods for learning and inference.CG is a hybrid Probabilistic Graphical Model (PGM) which Contains both directed and undirected links. Its representation is powerful enough to capture heterogeneous relationships among image entities. For CG they first oversegment the image into superpixels and find out different heterogeneous relationships among image entities (superpixels, vertices or junctions, edges, regions etc.) They construct the CG model with parameterization of links with derived Joint Probability Distribution (JPD). They represent these links by either potential function or conditional probabilities.

They first create a Directed Master Graph then create directed sub-graphs for some terms in the JPD of Directed Master Graph. They segment the image into two parts foreground and background. Finally applying the probabilistic inference model applies in the foreground to find out Most Probable Explanation (MPE). Li Zhang and Qiang Ji [16] have proposed a Bayesian Network (BN) Model for both Automatic (Unsupervised) and Interactive (Supervised) image segmentation. They Constructed a Multilayer BN from the oversegmentation of an image, which find object boundaries according to the measurements of regions, edges and vertices formed in the oversegmentation of the image and model the relationships among the superpixel regions, edge segment, vertices, angles and their measurements. For Automatic Image Segmentation after the construction of BN model and belief propagation segmented image is produced. For Interactive Image Segmentation if segmentation results are not satisfactory then by the human intervention active input selection are again carried out for segmentation.

Kittipat Kampa, Duangmanee Putthividhya and Jose C. Principe [17] design a probabilistic unsupervised framework called Irregular Tree Structure Bayesian Network (ITSBN). The ITSBN is made according to the similarity of image regions in an input image. ITSBN is a Directed acyclic graph (DAG) with two disjoint sets of random variables hidden and observed. The original image is oversegmented in multiscale hierarchical manner then they extracted features from the input image corresponding to each superpixel. According to these superpixels ITSBN is built for each level. After applying the learning and inference algorithms the segmented image is produced.

Fei Liu, Dongxiang Xu, Chun Yuan and William Kerwin [18] combined the BN and MRF (Markov Random Field) to form an image segmentation approach. The BN generates a probability map for each pixel in the image and then MRF prior is incorporated to produce the segmentation. It is a supervised image segmentation method. First each pixel will be individually assigned a probability value to be each given class. According to such probability map, BN provides a mechanism to convert the problem from feature space to image domain. Second they consider the prior knowledge on the image model and the spatial relationships between pixels, they used MRF based model to generate the segmentation.

Costas Panagiotakis, Ilias Grinias, and Georgios Tziritas [19] proposed a framework for image segmentation which uses feature extraction and clustering in the feature space followed by flooding and region merging techniques in the spatial domain, based on the computed features of classes. A new block-based unsupervised clustering method is introduced which ensures spatial coherence using an efficient hierarchical tree equipartition algorithm. They divide the image into different-different blocks based on the feature description computation. The image is partitioned using minimum spanning tree relationship and mallows distance. Then they apply K-centroid clustering algorithm and Bhattacharya distance and compute the posteriori distributions and distances and perform initial labelling. Priority multiclass flooding algorithm is applied and in the end regions are merged so that segmented image is produced.

Eric N. Mortensen and Jin Jia [20] proposed a two layer BN model for image segmentation, which captures the relationships between edge segments and their vertices. Given a user input seed path, they use minimum path spanning tree graph search to find the most likely object boundaries. They also encode a statistical similarity measure between the adjacent regions of an edge into its a priori probability therefore implicitly integrating region information.

In the early study [21] they use the similar BN model for both automatic and interactive segmentation. Their approach can find multiple non-overlapping closed contours before any given user's intervention. The intervention will serve as an evidence to help select a single closed contour that covers the object of interest.Lei Zhang, and Qiang Ji [22] develop a unified graphical model in which they combined directed graphical model and undirected graphical model. The combination allows capturing more complex and heterogeneous relationships among image entities. The unified model is more expressive and more powerful. But it only used for automatic segmentation not for interactive segmentation. They first propose to employ Conditional Random Field (CRF) to model the spatial relationships among image superpixel regions and their measurements. They introduce a multilayer Bayesian Network (BN) to model the causal dependencies that naturally exist among different image entities, including image regions, edges, and vertices. The CRF model and the BN model are then systematically and seamlessly combined through the theories of Factor Graph to form a unified probabilistic graphical model that captures the complex relationships among different image entities. Using the unified graphical model, image segmentation can be performed through a principled probabilistic inference.

In Probabilistic Image Modelling with an Extended Chain Graph for Human Activity Recognition and Image Segmentation they proposed a model called Chain Graph which is a hybrid probabilistic graphical model (PGM) capable of modelling heterogeneous relationships among random variables. They apply it to two challenging image and video analysis tasks: human activity recognition and image segmentation. Extended CG models are constructed to capture useful heterogeneous relationships among multiple entities for solving these problems. In [16] experiments show that the CG Models outperform conventional undirected PGMs or directed PGMs. It demonstrates the applicability of the proposed CG model to different image and video analysis problems as well as its potential benefits over standard directed or undirected PGMs in improving classification and recognition performance.

3. Proposed Method

In our method, an initial segmentation is required to partition the image into homogeneous regions for merging. Any existing low level segmentation methods, such as super-pixel, mean shift [6,7], watershed [3] and level set [8], can be used for this step. Mean shift method is best for initial segmentation because it has less over segmentation and can well preserve the object boundaries. In the proposed method, we only focus on the region merging.

3.1 Region representation and similarity measure

After mean shift initial segmentation, we have many small regions available. To guide the following region merging process, we need to represent these regions using some descriptor and define a rule for merging. A region can be described in many aspects, such as the colour, edge, texture, shape and size of the region. Among them the color histogram and gray level co-occurrence matrix (GLCM) for texture is an effective descriptor to represent the object color feature statistics and it is widely used in pattern recognition and object tracking, etc. In the context of region merging based segmentation, color histogram is more robust than the other feature descriptors. This is because the initially segmented small regions of the desired object often vary a lot in size and shape, while the colors of different regions from the same object will have high similarity. Therefore, we use the color histogram or GLCM to represent each region in this our proposed work.

In our proposed method we select the Euclidean descriptor or Bhattacharyya coefficient between region A and region B which provides similarity between region A and region B. If two regions have similar contents, their histograms will be very similar, and hence their Euclidean descriptor or Bhattacharyya coefficient will be very high, i.e. the angle between the two histogram vectors is very small. Certainly, it is possible that two perceptually very different regions may have very similar histograms. Fortunately, such cases are rare because the region histograms are local histograms and they reflect the local features of images. Even in case two perceptually different regions have similar histograms, the similarity between them is rarely the highest one in the neighborhood. Coupling with the "highest similarity rule", the Bhattacharyya similarity works well in the proposed region merging method. However, it should be stressed that other color spaces, such as the HSI color space, and other distance measures, such as the Euclidean distance between histogram vectors, can also be adopted in the proposed region merging scheme. We also present some examples of proposed method by using HSI color space and Euclidean distance, respectively. The results are similar to those by using the RGB color space.

3.2 Selecting of object and background

In the interactive image segmentation, the users need to specify the object and background conceptually. Similar to [10,13,17], the users can input interactive information by supervised techniques, which could be lines or curves on the image. The regions that have pixels inside the object markers are thus called object marker regions, while the regions that have pixels inside the background markers are called background marker regions. We use different color markers to mark the object while using and to represent the background. Actually, the less the required inputs by the users, the more convenient and more robust the interactive uses. After object marking, each region will be labeled as one of three kinds of regions: the marker object region, the marker background region and the non-marker region. To completely extract the object contour, we need to region merging for each non-marker region with a correct label of either object region or background region. Further we use region growing techniques to draw object clearly from background regions.

3.3 Similarity based merging rule

After object and background identification, it is still a challenging problem to extract accurately the object contour from the background because only a small portion of the object and background features are indicated by the user. The conventional region merging methods merge two adjacent regions whose similarity is above a preset threshold. These methods have difficulties in adaptive threshold selection. A big threshold will lead to incomplete merging of the regions belonging to the object, while a small threshold can easily cause over-merging, i.e. some object regions are merged into the background. Moreover, it is difficult to judge when the region merging process should stop.

Object and background selection can provide some key features of object and background, respectively. Similar to graph cut and marker based watershed, where the marker is the seed and starting point of the algorithm, the proposed region merging method also starts from the seed pixels and all the non-marker regions will be gradually labeled as either object region or background region. In this paper, we present an adaptive similarity based merging mechanism to identify all the non-marker regions under the guidance of object and background markers.

3.4 Expected outcomes and comparison with various existing methods

Table 1, shows that expected outcomes of our proposed method is depends on the size of image and number of initial segmented regions, we measure time in seconds and table 2, shows that the TPR and FPR values of different methods on the test images. Table 3 shows that the comparison of result for proposed method using 10000 super pixels with recent baselines, and Table 4 shows that the Time required for segmenting a normal size image:

Table1

The running time of Proposed Methods on different images

Image	*Bird*	*Horse*	*Bus*	*Dog*
Size of image	163×192	376×425	448×368	335×295
Number of regions after initial segmentation	170	522	1088	196
Running time (s)	6-20	32-46	80-95	12-35

Table 2

Error percentage of several methods with different numbers of initial segmented regions (n_s).

Initial Segmented Regions	*Graph (%)*	*PGM*	*Hyper graph*	*Proposed Method (Approx)*
2000	6.34	5.86	5.71	4.0-5.50
6000	6.14	5.67	5.52	5.0-6.25
10000	5.92	5.50	5.33	5.50-6.00

Table 3

Comparison of result for proposed method using 10000 super pixels with recent baselines

Method	*Error rate (%)*
GM-MRF [48]	7:9
Hypergraph [24]	7:3
s-Laplacian [13]	5:4
Proposed	5.33

Table 4

The Typical Time Required for Segmenting a Normal Size Image

Method	*Normal Image Size*	*Segmentation Time*
ITSBN	300 × 200	2 min/image
UGM	320 × 213	Less than 30 seconds
PMCFA	214 × 320	13.2 seconds
MPST	128 × 128	Approx. 1 second.

4. CONCLUSION

This paper proposed a novel region merging based interactive image segmentation method. The image is initially segmented by using some segmentation and the users only need to roughly indicate the main features of the object and background by using some supervised technique. Since the object regions will have high similarity to the marked object regions and so do the background regions, a similarity based region merging mechanism was proposed to extract the object. The proposed scheme is simple yet powerful and it is image content adaptive. The proposed scheme efficiently exploits the color similarity of the target object so that it is robust to the variations of input markers.

The proposed method provides a general region merging framework. It does not depend essentially using some segmentation and other color image segmentation methods can also be used for initial segmentation. Although some marker based interactive image segmentation methods have been proposed, the proposed algorithm firstly exploits a novel adaptive maximal similarity based region merging mechanism. In the future, we will explore how to introduce pixel classification into the merging process to make the algorithm more intelligent.

REFERENCES

[1]. M. Galun, E.sharon, R. Basri and A. Brandt “Texture Segmentation by Multiscale Aggregation of Filter Responses and Shape Elements” *Proc. IEEE Int’l conf. Computer Vision and Pattern Recognition*, pp. 716-723, 2003.

[2]. E. Sharon, M. Galun, D. Sharon, R. Basri and A. Brandt, “Hierarchy and Adaptivity in Segmenting Visual Scenes,” Nature vol. 442, no. 7104, pp. 810-813, june 2006.

[3]. J. Shi and J. Malik, “Normalized Cuts and Image Segmentation,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, August 2000.

[4]. D. Martin, C.Folwkes, D. Tal, J.Malik, “A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics,” *Internatinal Conference on Computer Vision,* pp. 416-423, 2001.

[5]. M. Kass, A. Witkin, D. Terzopoulos, "Snake: Active Contour Models," *International Journal of Computer Vision,* vol.1, no.4, pp. 321-331, 1987.
[6]. F. Meyer, S. Beucher, "Morphological Segmentation," *Journal of Visual Communication and Image representation*, vol.1, no. 1, pp. 21-46, 1990.
[7]. D.A. Forsyth, J.ponce, Computer Vision: A Morden Approch, Prentice-Hall Englewood Cliffs, NJ, 2002.
[8]. N.Bonnet, J.Cutrona, M.Herbin, A no Threshold Histogrm Based Image Segmentation Method, Pattern Recognition vol. 35, no.10, pp. 2319-2322, 2002.
[9]. Y. Boykov, G. Funka-Lei, "Graph cuts and efficient n-d image segmentation" *International Journal of Computer vision, vol 70. no.2*, pp.109-131,2006.
[10]. J. Lafferty, A. McCallum, and F. Pereira,"Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," Proc. Int'l Conf. Machine Learning,pp. 282-289, 2001.
[11]. S.L. Lauritzen, Graphical Models. Oxford Univ. Press, 1996.
[12]. F.V. Jensen, Bayesian Networks and Decision Graphs. Springer-Verlag, 2001.
[13]. S.Z. Li, Markov random field modeling in computer vision, Springer-Verlag New York, Inc. Secaucus, NJ, USA, 1995.
[14]. J. Lafferty, A. McCallum, and F. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," Proc. Int'l Conf. Machine Learning, pp. 282-289, 2001.
[15]. Lei Zhang, Zhi Zeng, and Qiang Ji "Probabilistic Image Modeling With an Extended Chain Graph for Human Activity Recognition and Image Segmentation", IEEE Transaction on Image Processing, VOL. 20, NO. 9, September 2011.
[16]. Lei Zhang and Qiang Ji, "A Bayesian Network Model for Automatic and Interactive Image Segmentation", IEEE Transaction on Image Processing, VOL. 20, NO. 9, September 2011.
[17]. Kittipat, Kampa, Duangmanee Putthividhya and Jose C. Principe, "Irregular Tree-Structured Bayesian Network for Image Segmentation.", IEEE International Workshop on Machine Learning for Signal Processing September 18-21, 2011, Beijing, China.
[18]. F. Liu, D. Xu, C. Yuan, and W. Kerwin, "Image segmentation based on Bayesian network-Markov random field model and its application on in vivo plaque composition," in *Int. Symp. Biomed. Imag.*, 2006, pp. 141–144.
[19]. Costas Panagiotakis, Ilias Grinias, and Georgeios Tziritas "Natural Image Segmentaion Based on Tree Equipartition, Bayesian Flooding and Region Merging", IEEE Transactions on Image Processing, Vol. 20, No. 8, August 2011.
[20]. E. N. Mortensen and J. Jia, "Real-time semi-automatic segmentation using a Bayesian network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 1007–1014.
[21]. E. Mortensen and J. Jia, "A Bayesian network framework for real-time object selection," in *Workshop Percept. Org. Comput. Vis.*, 2004, pp. 44-49.
[22]. L. Zhang and Qiang Ji, "Image Segmentation with a Unified Graphical Model,"IEEE Transactions
Pattern Anal. Mach. Intell., vol. 32, no. 8, pp. 1406–1425, August 2010

.