

MRF Based Image Segmentation Augmented with Domain Specific Information

Özge Öztimur Karadağ and Fatoş T. Yarman Vural

Department of Computer Engineering,
Middle East Technical University, Ankara, Turkey
{oztimur,vural}@ceng.metu.edu.tr
<http://image.ceng.metu.edu.tr>

Abstract. A Markov Random Field based image segmentation system which combines top-down and bottom-up segmentation approaches is proposed in this study. The system is especially proposed for applications where no labeled training set is available, but some priori general information referred as *domain specific information* about the dataset is available. *Domain specific information* is received from a domain expert and formalized by a mathematical representation. The type of information and its representation depends on the content of the image dataset to be segmented. This information is integrated to the segmentation process in an unsupervised framework. Due to the inclusion of domain specific information, this approach can be considered as a first step to semantic image segmentation under an unsupervised MRF model. The proposed system is compared with the state of the art unsupervised image segmentation methods quantitatively via two evaluation metrics; consistency error and probabilistic rand index and satisfactory results are obtained.

Keywords: image segmentation, Markov Random Fields, domain specific segmentation.

1 Introduction

Markov Random Fields are widely used in image segmentation studies, since they provide an effective framework for modeling spatial relations. A group of MRF based methods employs a supervised approach to image segmentation problem by utilizing labeled datasets. They construct complicated energy functions which include various image features and relations among image parts. A second group of studies, takes an unsupervised approach and construct a relatively simpler energy function. The major difference between these two approaches is the availability of labels. In most of the real life problems, labeling the images is not practical or possible. On the other hand, depending on the application domain one may extract domain specific information which can be employed to guide the segmentation process. Remote sensing applications are good examples of such data sets, where one seeks a group of objects in a highly cluttered

background. For instance, if the goal is to detect the airplanes or airports in a remotely sensed image, the unsupervised segmentation algorithms are quite naive to extract the airport regions or airplanes with a complete segmentation method. Similarly, supervised segmentation approaches require large amount of labeled data. Even if sufficient labeled data is available, due to the large within class variances, supervised segmentation algorithms fail to extract the targeted objects. However, if one can employ a priori domain specific information into the segmentation method, it is more likely that the object of interest is extracted in whole regions. For example, it is well known that runways consists of two parallel lines and airports are constructed in a planar regions. This information is easily formalized by a mathematical model and can be employed in the segmetation algorithm. Similarly, in an image database of animals, we know that zebras have stripes, but we may not have a labeled set of zebras. In this study, this type of information about the problem domain is referred as *domain specific information*, and it can be represented in a wide range of mathematical forms, such as, a set of relations among the certain image parts, or basic image features such as color, texture or shape.

2 Related Work

Markov Random Fields are first proposed as an image processing method by Geman and Geman [7] in 1984. They fomulate image segmentation in Markov Framework and showed that this problem can be modeled as a Gibbs Distribution. Majority of the studies in the literature employ a double clique model by employing *eq. 1*.

$$E(x) = \sum_{i \in S} \psi_i(x_i) + \sum_{i \in S, j \in N_i} \psi_{ij}(x_i, x_j) \quad (1)$$

Here, S is the set of all image pixels corresponding to sites, N is a neighborhood system and x_i denotes the labelling of pixel i . In this equation, the first term, referred as unary potential, is generally defined as the negative log likelihood of a label being assigned to pixel i . Its formula is given in *eq. 2*. It is assumed that the features follow Gaussian distribution and various features are utilized for this term [16,8]. The second term, referred as double clique, models the relations between neighboring sites. It is a smoothness term which is usually modeled via Potts model. It takes zero if same label is assigned to two neighboring sites, otherwise it takes 1 or some positive value β , as in *eq. 3*. Various versions of Potts model are encountered in the literature. One direct modification is contrast sensitive Potts model [8], where the function takes a value, based on the level of difference between two neighboring sites.

$$\psi_i(x_i) = \sum_{i \in S} \ln(\sqrt{2\pi}\sigma_{x_i}) + \frac{(i - \mu_{x_i})^2}{2(\sigma_{x_i})^2} \quad (2)$$

$$\psi_{ij}(x_i, x_j) = \begin{cases} 0 & \text{if } x_i = x_j \\ \beta & \text{otherwise} \end{cases} \quad (3)$$

Energy of MRF system is minimized usually by Simulated Annealing or Iterated Conditional Modes. Due to its stochastic nature, simulated annealing converges slowly. For this reason, Besag [1] proposed Iterated Conditional Modes for minimizing MRF energy, which is a deterministic algorithm based on the idea of optimizing local energy iteratively. On the other hand, graph cut based energy minimization methods propose fast approximate solutions for MRF based image segmentation [2].

3 MRF Based Segmentation Augmented with Domain Specific Information

3.1 Motivation

Our goal is to propose an MRF based image segmentation system which can be employed in applications where labels are not present, but some priori information about the given problem domain is available. This information, referred as domain specific information, may be represented in various forms and may be employed via different techniques. In this study, we consider the application areas where this type of expert information is available and assume that this information can be represented by some mathematical tools, such as low level image features, spatial relationships among the object of interest, general structure of background clutter.

The novelty of this study is the incorporation of high level information into segmentation process in an unsupervised framework. This is possible only if expert knowledge about given problem is available. If this is the case, then the system initially detects certain image parts via object detection, shape detection or obtain an initial labeling for the image using certain low level image features which is specified by domain expert. This initial labeling is later utilized in MRF based segmentation.

3.2 System Architecture and Energy Function

The proposed system consists of two levels of information processing, in the first level a set of bottom-up image segmentation systems are employed; these may be distinct algorithms or they may be same algorithm with different parameters. In the second level, MRF system is constructed whose neighborhood weights are determined by information from the first level. Initially, all the weights are identical, after introduction of information from the first level some weights are weakened while some others are strengthened. System architecture is shown in *Figure 1*.

Level 1: Level 1 consists of two parts, first part is a set of bottom-up segmentations. Any bottom-up segmentation algorithm may be employed in this part. Each bottom-up segmenter obtains a segmentation which is denoted as R_i . In this study, mean shift segmentation with various values of spatial bandwidth

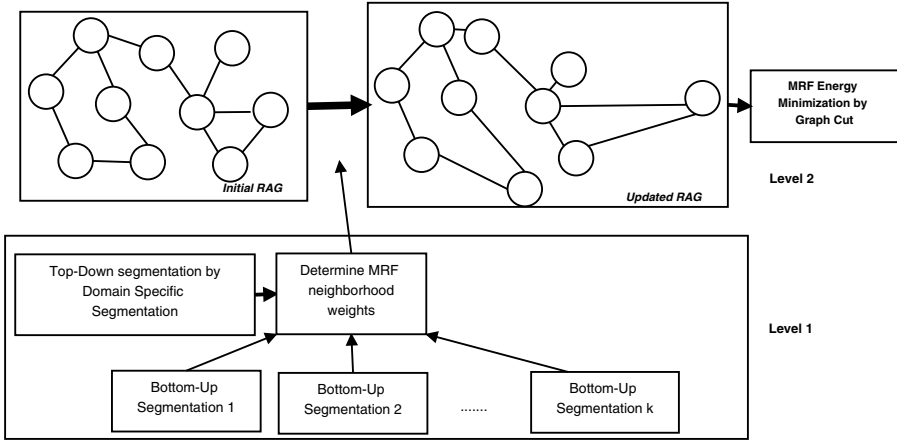


Fig. 1. System Architecture

and range bandwidth are employed. Second part of Level 1 is domain specific segmentation. As the name implies, this segmentation depends on the application domain. This information is provided by a domain expert and it may be in tree forms:

- **DS information related to low level image features:** Domain specific information related to color, texture, intensity information is in this group. This group of information may be referred as mid level information. Expert knowledge regarding illumination conditions, or surface reflectances can be considered in this group. Basic thresholding segmentation methods may be employed to find a domain specific segmentation, which is expected to segment salient image regions or background, or only certain image parts.
- **DS information related to primitive shapes:** Domain expert may state that certain shapes; such as line, circle etc. are expected in a given image. In this case, shape detectors are employed to segment related image parts. This group of information may also be referred as mid level information, but it is one step closer to high level information when compared to the first group domain specific information.
- **DS information related to object(s) in the image:** Domain expert may provide a higher level information regarding which object(s) are expected in the given domain. In this case, object detectors are employed to segment corresponding image parts. This type of domain specific information may be referred as high level information.

Segmentation obtained by domain specific information is referred as top-down segmentation in *Figure 1*, since it is directed by mid or high level domain information. If this domain information is in the first group, that is, it is related to low level image features then all image pixels are segmented. On the other hand,

if domain specific information is in the second or third group, certain image parts are segmented while the rest of the image is left as a single region.

Level 2: Markov Random Fields based image segmentation employed in this study utilizes super pixels as sites, hence image is first over segmented and super pixels obtained are denoted as S . Then region adjacency graph (RAG) for S is obtained. In this graph all neighboring nodes (super pixels) have equal weights. These weights are updated by information from Level 1. Let w_{ij} denote weight between two neighboring super pixels s_i and s_j . This weight is updated by eq. 4. Here, K is the number of bottom-up segmenters at Level 1. D_{s_i} is the region index of s_i that is assigned by domain specific segmenter; each pixel is assigned a region index and region index of the super pixel is determined by majority voting, that is the region index that is assigned to majority of pixels in that super pixel is selected as the super pixel's region index. p is the number of bottom-up segmenters that puts s_i and s_j in the same region. And β is a constant that determines the weight of domain specific information. If domain specific segmentation does not assign same region index to s_i and s_j then weight w_{ij} is determined only by bottom-up segmenters, as given in first condition of eq. 4. If domain specific segmenter assigns them same region index and at least half of the bottom-up segmenters agree with DS segmenter, then weight is increased by β , as in the second condition. If bottom-up segmenters do not agree with DS segmenter then DS segmenter is given same weight as other Level 1 segmenters.

$$w_{ij} = \begin{cases} p/K & \text{if } D_{s_i} \neq D_{s_j} \\ \beta + p/K & \text{if } D_{s_i} = D_{s_j} \wedge p \geq (K/2) \\ (p+1)/(K+1) & \text{otherwise} \end{cases} \quad (4)$$

Once, neighborhood weights are determined and updated RAG is obtained, MRF based segmentation is realized by graph cut based image segmentation of Boykov et al [2]. Their swap algorithm is utilized for minimizing MRF energy in eq. 5. Here, s_i represents super pixels as sites of Markov Rand Field and l_{s_i} is the label assigned to that super pixel. First term of energy function is constructed using CIELab color features and classical unary potential as in eq. 2. Second term is modeled as weighted Potts model.

$$E(x) = \sum_{s_i \in S} \psi_i(l_{s_i}) + \sum_{s_i \in S, s_j \in N_{s_i}} w_{ij} \times \psi_{ij}(l_{s_i}, l_{s_j}) \quad (5)$$

The steps of the algorithm are provided at Algorithm 1. First, Level 1 bottom-up segmentations and DS segmentation are obtained. Output of first Level 1 bottom-up segmenter is employed for estimating mean and covariance parameters of region indexes to be used in unary potential of MRF energy. After that, super pixels are obtained and RAG is constructed. Edge weights (neighborhood weights) are adjusted by information from Level 1. Finally, MRF energy is minimized by graph cut based swap algorithm, which initializes the system with a random labeling L and updates this labeling by swap moves.

Algorithm 1. Steps of DS-MRF

- 1: Obtain bottom-up segmentations R_i
 - 2: Obtain DS by utilizing expert knowledge
 - 3: Obtain super pixels, S , by Mean Shift
 - 4: Construct RAG for S
 - 5: **for** each region index l_i **do**
 - 6: estimate model parameters (μ, Σ) from R_1
 - 7: **end for**
 - 8: Adjust RAG edge weights by eq. 4
 - 9: Initialize a random labeling L
 - 10: $success = 0$
 - 11: **for** each pair of labels $\{\alpha, \beta\}$ **do**
 - 12: find $\hat{L} = \operatorname{argmin} E(\hat{L})$ among \hat{L} within one $\alpha - \beta$ swap of L
 - 13: if $E(\hat{L}) < E(L)$, set $L = \hat{L}$ and $success = 1$
 - 14: **end for**
 - 15: if $success = 1$ go to 10
 - 16: return L
-

4 Experiments

The proposed system is run on a subset of Berkeley Segmentation Dataset [10] which is widely used for segmentation evaluation. Berkeley dataset consists of 300 images each of which have five to seven groundtruth segmentations. The dataset contains images from various domains; people, animals, scene etc. A subset of this dataset consisting of 80 images is selected to construct a domain specific dataset. Selected images are outdoor images with some green area. For this dataset, domain specific information states that green areas can be detected via Normalized Difference Index whose eq. is given in eq. 6. Motohka et al. proposed employing NDI [11], which they refer as Green-Red Vegetation Index, for green land and vegetation detection in remote sensing applications. They propose thresholding at $NDI = 0$ for phenology detection. Similarly, NDI thresholding is employed as DS segmenter in this experiment. As bottom-up segmenters, Mean Shift segmentation is employed for several values of range bandwidth and spatial bandwidth parameters. Both parameters take values as $\{3, 7, 11, 15\}$ and a total of 16 segmentations are obtained. Hence $K = 16$ and β is set as 2 in this experimental setup.

$$NDI = \frac{\rho_{green} - \rho_{red}}{\rho_{green} + \rho_{red}} \quad (6)$$

4.1 Segmentation Evaluation

Consistency Error. Given two segmentations masks of an image, consistency error measures the level of consistency among them. For each pixel p_i , segment S_1 containing this pixel in the first segmentation and the segment S_2 containing this pixel in the second segmentation are compared. Either there is a refinement relation between two segments, or there are overlapping pixels in two segments.

Consistency between two segments are measured with *eq. 7*.

$$E(S_1, S_2, p_i) = \frac{|R(S_1, p_i) \setminus R(S_2, p_i)|}{R(S_1, p_i)} \quad (7)$$

Here " \setminus " is set difference operator which is non-symmetric. Using this local error measure, two error measures are defined first one is Global Consistency Error (GCE), which is provided in *eq. 8*, and second measure is Local Consistency Error (LCE), which is given in *eq. 9*. GCE and LCE take values in the range [0,1], values close to 0 indicates high segmentation performance. Consistency errors are informative if two segmentations have approximately the same number of regions.

$$GCE(S_1, S_2) = \frac{1}{n} \min \left\{ \sum_i E(S_1, S_2, p_i), \sum_i E(S_2, S_1, p_i) \right\}, \quad (8)$$

$$LCE(S_1, S_2) = \frac{1}{n} \sum_i \{ \min \{ E(S_1, S_2, p_i), E(S_2, S_1, p_i) \} \}. \quad (9)$$

Probabilistic Rand Index. Probabilistic Rand Index (PRI) takes pixels in pairs and measures the ratio of compatibly labeled pixels in segmentations S_{test} and ground truth segmentations S_k with *eq. 10*.

$$PR(S_{test}, \{S_k\}) = \frac{1}{\binom{N}{2}} \sum_{i < j} [c_{ij} p_{ij} + (1 - c_{ij})(1 - p_{ij})] \quad (10)$$

Here, N is the number of pixels in image, c_{ij} is the event of a pair of pixels i and j having the same label in image S_{test} and p_{ij} is the ground truth probability of two pixels having the same label estimated over all ground truth segmentations of the image. PRI takes values in the interval [0,1], where values close to 1 indicates high segmentation performance. PRI is meaningful even if two segments have different number of regions.

4.2 Comparison of Methods and Segmentation Examples









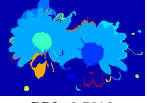










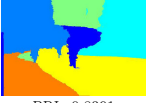








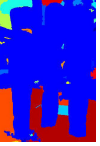

Proposed system is compared with three state of the art image processing methods; Mean Shift Segmentation [6], Multiscale Normalized Cut Segmentation [5] and Efficient Graph Based Segmentation [12]. Publicly available systems

Table 1.

Method	PRI	GCE
Mean Shift [6,3]	0.7272	0.2399
EGS [12,13]	0.6241	0.1293
Multiscale N-Cut [5,14]	0.6892	0.2880
DS-MRF	0.7492	0.2139

[3,14,13] are utilized for implementation of the algorithms. Mean Shift segmentation parameters spatial bandwidth and range bandwidth take values {3,7,11,15}. Segmentation result with highest PRI is provided at *Table 1*. Efficient Graph Based Segmentation parameter k , which defines the merging criterion is set as {5,8,11}, neighborhood size is set as 1 and minimum region size of a region is set as 50. In Multiscale Normalized Cut, number of regions is set as 6. This is determined as the average number of regions in DS-MRF segmentation results. Segmentation performance of these methods are compared via PRI and GCE as provided in *Table 1*. Parameters maximizing PRI and minimizing GCE are selected for each algorithm.

Table 2. Segmentation Examples

Original Image	DS-MRF	Mean shift	EG	NCut
	 PRI=0.9388 GCE=0.0837	 PRI=0.5689 GCE=0.3055	 PRI=0.8791 GCE=0.1011	 PRI=0.5321 GCE=0.2400
	 PRI=0.7207 GCE=0.2843	 PRI=0.7211 GCE=0.2773	 PRI=0.7616 GCE=0.2276	 PRI=0.6392 GCE=0.4810
	 PRI=0.8298 GCE=0.1519	 PRI=0.8377 GCE=0.1208	 PRI=0.7739 GCE=0.2468	 PRI=0.7243 GCE=0.3060
	 PRI=0.8286 GCE=0.1619	 PRI=0.7916 GCE=0.1693	 PRI=0.7364 GCE=0.1120	 PRI=0.8391 GCE=0.1460
	 PRI=0.7796 GCE=0.1242	 PRI=0.7897 GCE=0.1706	 PRI=0.7450 GCE=0.1384	 PRI=0.7435 GCE=0.2134
	 PRI=0.7305 GCE=0.1507	 PRI=0.8224 GCE=0.3072	 PRI=0.5386 GCE=0.1461	 PRI=0.8106 GCE=0.4225

Segmentation results for sample images from the outdoor green area dataset are given at *Table 2*. PRI and GCE values are also provided so that, performance can be examined both visually and quantitatively. In these examples, contribution of domain specific information is realized clearly.

5 Conclusion and Future Work

In this study, an unsupervised MRF based image segmentation system is proposed and competitive results with state of the art segmentation methods is obtained. The contribution of this study is to embed the domain specific information into the MRF energy function in a simple but effective way. High segmentation accuracy is obtained for the selected set of images with provided domain specific property. Throughout the experiments in this study, it is assumed that only one specification is available and it is related to low level image features, hence only one domain specific segmentation is obtained and incorporated into the MRF energy. Nevertheless, there may be more information available and for each piece of information a distinct top-down segmentation may be obtained and integrated to MRF energy. Therefore, due to the design of the energy function, the proposed system can be considered as a consensus segmentation. In this new approach information from various sources are joined under the MRF energy function to improve the segmentation quality.

This study can be considered as a first step to unsupervised semantic segmentation systems where semantic information is introduced via the energy function without a training phase. In this preliminary study, the domain specific information is modeled by NDI thresholding. However, depending on the problem domain more sophisticated models can be incorporated in the energy function.

References

1. Besag, J.: Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society, Series B* 36(2), 192–236 (1974)
2. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(11), 1222–1239 (2001)
3. Christoudias, C., Georgescu, B., Meer, P.: Synergism in low-level vision. In: 16th International Conference on Pattern Recognition, Quebec City, Canada, vol. IV, pp. 150–155 (2002)
4. Pantofaru, C., Hebert, M.: A Comparison of Image Segmentation Algorithms. The Robotics Institute, Carnegie Mellon University, Number CMU-RI-TR-05-40 (2005)
5. Cour, T., Benezit, F., Shi, J.: Spectral segmentation with multiscale graph decomposition. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, vol. 2, pp. 1124–1131, 20–25 (2005)
6. Comaniciu, D., Meer, P.: Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(5), 603–619 (2002)

7. Geman, S., Geman, D.: Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6(6), 721–741 (1984)
8. Kohli, P., Ladicky, L., Torr, P.: Robust higher order potentials for enforcing label consistency. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2008*, pp. 1–8, 23–28 (2008)
9. Ladicky, L., Russell, C., Kohli, P., Torr, P.H.S.: Associative hierarchical CRFs for object class image segmentation. In: *2009 IEEE 12th International Conference on Computer Vision*, pp. 739–746 (2009)
10. Martin, D.R., Fowlkes, C., Tal, D., Malik, J.: A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. Technical Report, EECS Department University of California, Berkeley (2001)
11. Motohka, T., Nasahara, K.N., Oguma, H., Tsuchida, S.: Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. *Remote Sens.* 2, 2369–2387 (2010)
12. Felzenswalb, P.F., Huttenlocher, D.P.: Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision* 59(2) (2004)
13. Dongcai, S.: Efficient Graph Based Image Segmentation Code, <http://www.mathworks.com/matlabcentral/fileexchange/29299-efficient-graph-based-image-segmentation>
14. Cour, T., Yu, S., Shi, J.: Normalized Cut Segmentation Code. Copyright 2004 University of Pennsylvania, Computer and Information Science Department (2004)
15. Unnikrishnan, R., Pantofaru, C., Hebert, M.: Toward Objective Evaluation of Image Segmentation Algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* 6(29), 929–944 (2007)
16. Kato, Z., Pong, T.C., Lee, J.C.M.: Color Image Segmentation and Parameter Estimation in a Markovian Framework. *Pattern Recognition Letters* 22(3-4), 309–321 (2001)