I.J. Wireless and Microwave Technologies **2012, 4, 59-64**
Published Online August 2012 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijwmt.2012. 04.09

Available online at http://www.mecs-press.net/ijwmt

A modified semi-supervised color image segmentation method

Wei Hongru, Chai Fangyong

Mathematics Physics School, University of Science and Technology Beijing, Beijing 100083，china

Abstract

The paper proposed a modified color image segmentation method basing on semi-supervised hidden Markov random fields (HMRF) with constraints. Making use of MeanShift algorithm to get supervision information and, cluster number and initial values for cluster centroids, color images can be segmented effectively with the method in this paper by K-Means algorithm. The experimental results are very encouraging.

Index Terms: HMRF, semi-supervised, MeanShift, clustering, K-Means.

1. Introduction

Random field technique can provide exact area features of image. So when the image feature is complex or it can't be dealed with simple methods, satisfying results can be obtained with random field. Markov random field is one of the most used statistical methods, which considers each point as a random variable with certain probability distribution.On the basis of HMRF, we apply the semi-supervised clustering algorithm to image segmentation. In the process of implementation, we modified the module to propose a new method, i.e., MeanShift[1] being used to deal with initialization and constraints. The semi-supervised clustering algorithm[2-3] ,on basis of HMRF, is briefly introduced in Section 2. Modified contents are presented in Section 3. In Section 4, experimental results are shown. Finally, our conclusion is provided.

2. Semi-Supervised Clustering with HMRF

The HMRF probalistic framework[2-3] consists of the following components :

sample set $X = (x_1, \ldots, x_n)$, label set $Y = (y_1, \ldots, y_n)$,

constraint set $C = (c_{12}, c_{13}, \ldots, c_{n-1,n})$,

model parameters $\Theta = \{A, M\}$.

Corresponding author:
E-mail address:

The objective function is as follows:

$$\mathscr{T}_{obj} = \sum_{x_i \in X} d_A(x_i, \mu_{y_i}) + \sum_{c_{i,j} \in C} v(i,j) - \log P(\cdot) + \log Z + n \log Z_{\Theta} \quad (1)$$

where the potential function has the following form:

$$v(i,j) = \begin{cases} w_{i,j} f_{ML}(i,j) & c_{ij} = 1, y_i \neq y_j \\ w_{i,j} f_{CL}(i,j) & c_{ij} = -1, y_i = y_j \\ 0 & otherwise \end{cases} \quad (2)$$

The functions f_{ML}, f_{CL} are chosen as follows:

$$\begin{aligned} & f_{ML}(x_i, x_j) = d_{eucA}(x_i, x_j) \\ & f_{CL}(x_i, x_j) = \varphi^{\max} - d_{eucA}(i,j), \varphi^{\max} = \max_{(x_i, x_j) \in C_{CL}} d_{eucA}(x_i, x_j) \end{aligned} \quad (3)$$

The model provides three kinds of distortion measure, and the squared Euclidean distance is used in this paper, with a symmetric positive-definite matrix A :

$$d_{eucA}(x_i, x_j) = \left\| x_i - x_j \right\|_A^2 = (x_i - x_j)^T A (x_i - x_j) .$$

2.1. semi-supervised clustering with constraints

Pairwise supervision in [2-3] are provided as must-link and cannot-link constraints on data points: a must-link constraint indicates that both points in the pair should be placed in the same cluster, while a cannot-link constraint indicates that two points in the pair should belong to different clusters. Typically, the constraints are “soft”, that is, clusterings that violate them are undesirable but were not prohibited. In certain applications, supervision in the form of class label may be unavailable, while pairwise constraints are easily obtained, creating the necessity for methods that exploit such supervision. Proposed methods for semi-supervised clustering fall into two general categories that we call constraint-based and distance-based. Constraint-based methods use the provided supervision to guide the algorithm toward a data partitioning that avoids violating the constraints. In the distance-based approaches, an existing clustering algorithm that uses a particular distance function between points is employed; however, the distance function is parameterized and the parameter values are learned, by which must-link points are brought together and can not link points taken further apart.

With a unified probalilistic model, semi-supervised clustering based on HMRF combines the constraint-based and distance-based approaches. The objective function can be optimized using an EM[4]-style algorithm , HMRF-KMeans[2-3] which finds a local minimum of the objective function.

2.2. HMRF-KMeans algorithm

Input: Set of data points $X = (x_1, \ldots, x_n)$, number of clusters K, set of constraints C, constraint violation measure D.

Ouput: Disjoint K-partitioning $(X_1, \ldots, X_K)$ of X such that objective function in (1) is locally minimized.

Method:

1). Initialize the K clusters centroids

$M^{(0)} = (\mu_1^{(0)}, \ldots, \mu_K^{(0)})$, set $t \leftarrow 0$

2). Repeat until convergence

2a).E-step: Given centroids $M^{(t)}$ and distortion parameters $A^{(t)}$,re-assign cluster labels

$$Y^{(t+1)} \leftarrow (y_1^{t+1},\ldots,y_n^{t+1})$$

on X to minimize $\mathcal{J}_{obj}$.

2b).M-step(A):Given cluster labels $Y^{(t+1)}$ and distortion parameters $A^{(t+1)}$,re-calculate centroids

$$M^{(t+1)} \leftarrow (\mu_1^{(t+1)},\ldots,\mu_K^{(t+1)})$$

to minimize $\mathcal{J}_{obj}$.

2c).M-step(B):Given cluster labels $Y^{(t+1)}$ and centroids $M^{(t+1)}$,

Re-estimate parameters $A^{(t+1)}$ of the distortion measure to reduce $\mathcal{J}_{obj}$.

2d). $t \leftarrow t+1$

2.3. EM algorithm

E step

$$\mathcal{J}_{obj}(x_i,\mu_h) = d_A(x_i,\mu_h) + \sum_{\substack{(x_i,x_j)\in C_{ML}^i \\ s.t.\, y_i \neq y_j}} w_{ij}\varphi(x_i,x_j) + \sum_{\substack{(x_i,x_j)\in C_{CL}^i \\ s.t.\, y_i = y_j}} w_{ij}(\varphi^{\max} - \varphi(x_i,x_j)) - \log \mathrm{P}(A) \tag{4}$$

The objective function in (1) is specified in (4). Assignments of data points to clusters are updated using the current estimate of cluster representatives and parameters.

M step

In the first part of the M step, the cluster centroids M are re-estimated from points currently assigned to them, to decrease the objective function $\mathcal{J}_{obj}$ in (4). For Bregman divergences, the cluster representative calculated in the M step of the EM algorithm is equivalent to the expectation value over the points in that cluster, which is equal to their arithmetic mean[6].

In the second part of the M step, the parameter matrix A is updated during gradient descent using the rule:

$$A \leftarrow A - \eta \frac{\partial \mathcal{J}_{eucA}}{\partial A}$$

(where η is the learning rate) (5)

$\frac{\partial \mathcal{J}_{eucA}}{\partial A}$ can be expressed as

$$\frac{\partial \mathcal{J}_{eucA}}{\partial A} = \sum_{x_i\in X} \frac{\partial d_{eucA}(x_i,\mu_i)}{\partial A} + \sum_{\substack{(x_i,x_j)\in C_{ML} \\ s.t.\, y_i \neq y_j}} w_{ij}\frac{\partial d_{eucA}(x_i,x_j)}{\partial A} + \sum_{\substack{(x_i,x_j)\in C_{CL} \\ s.t.\, y_i = y_j}} w_{ij}\left(\frac{\partial \varphi_{eucA}^{\max}}{\partial A} - \frac{\partial d_{eucA}(x_i,x_j)}{\partial A}\right) - \frac{\partial \log \mathrm{P}(A)}{\partial A} - n\frac{\partial \log \det(A)}{\partial A} \tag{6}$$

The gradient of the parameterized squared Euclidean distance is given by

$$\frac{\partial d_{eucA}(x_i, x_j)}{\partial A} = (x_i - x_j)(x_i - x_j)^T \tag{7}$$

$\frac{\partial \log P(A)}{\partial a_m}$ is given by

$$\frac{\partial \log P(A)}{\partial a_m} = \frac{1}{a_m} - \frac{a_m}{s^2} \tag{8}$$

The gradient of the distance normalize $\log\det(A)$ is as follows:

$$\frac{\partial \log\det(A)}{\partial A} = 2A^{-1} - diag(A^{-1}) \tag{9}$$

The parameters of the parameterized distortion measure are updated to decrease the objective function to speed up the convergence.

3. Modified contents

3. 1 Initialization

MeanShift algorithm[1] is applied to implementing initialization and the details are as follows:

1)Initialization of must-link constraints C_{ML} . With MeanShift, a little number of points from the clustering result are selected evenly to construct the constraints. C_{ML} consists of point index sets which indicate that the points should be in the same cluster.

2)Initialization of cluster number and centroids. With MeanShift, the number of clusters K and cluster centroids M could be provided. The result of MeanShift will be a desirable choice for the initialization and the centroids are closer to the optimal ones.

3)Initialization of class labels. the labels of points are initialized by computing the minimal distance to centroids M . In the initial model, initialization of class labels is not mentioned.

3. 2 Modification of EM algorithm

The model includes must-link(C_{ML}) and cannot-link(C_{CL}) constraints. C_{ML} can be easily constructed with MeanShift while C_{CL} is more difficult to obtain. As a result, regardless of the C_{CL} , the objective function in E step is modified as follows:

$$\mathcal{J}_{obj}(x_i, \mu_h) = d_A(x_i, \mu_h) + \sum_{\substack{(x_i, x_j) \in C_{ML}^i \\ s.t.\, y_i \neq y_j}} w_{ij}\varphi(x_i, x_j) - \log \mathrm{P}(A) \tag{10}$$

This relieves not only the initialization work, but also the compution greatly.

In the process of updating the parameter matrix by (5) in the second part of M step, divergence comes out. Then this part is omitted. As a result, the compution in (5) will be relieved while the convergence of objective function is not affected.

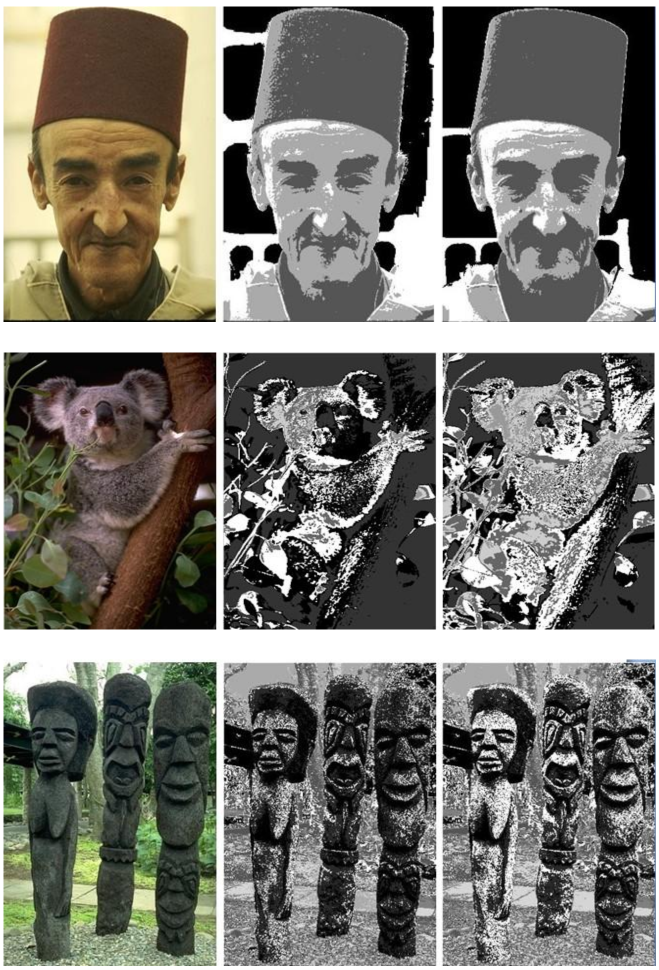
4. EXPERIMENTS

The algorithm is implemented with MATLAB, and the images are downloaded from [5]. The image size is 481×321 or 321×481. Fig. 1 displays the experiment result. The first column is the initial color image. The second column is the clustering result with centroids provided by MeanShift, where the minimal distance to centroids is taken into account. And the third column is the segmentation result of the modified method in this paper.

The experiment suggests that the result of the segmentation method in this paper is desirable. MeanShift algorithm provides valuable initial information, such as the number of clusters, and better segmentation can be achieved. With the objective function, the labels of points can be assigned more exact which could be displayed by Figure1.

5. CONCLUSION

With semi-supervised clustering on basis of HMRF and MeanShift, a new modified semi-supervised color image segmentation method is presented. Experiments demonstrate that color image can be segmented effectively with this algorithm.



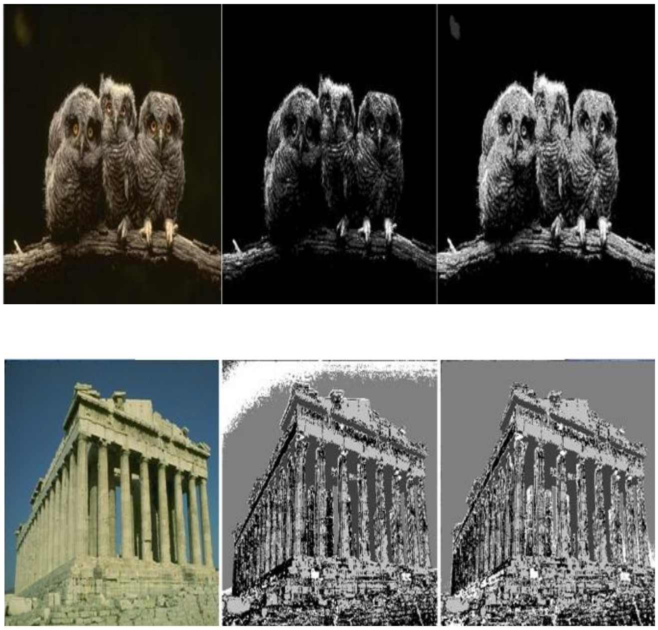


Figure1 experiment results comparison

Acknowledgment

This work is supported by the National Natural Science Foundation of China (No.60873259)and Engineering Research Institute Foundation of USTB (No. 2009-002) and Foundation of Education Teaching Research of USTB (No. 2009-16).

References

[1] D. Comaniciu, P. Meer. Mean Shift: A Robust Approach Toward Feature Space Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 603-619, May 2002, doi:10.1109/34.1000236

[2] S. Basu, M. Bilenko, and R. J. Mooney. A probabilistic framework for semi- supervised clustering, In Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining, pages 59–68, Seattle, WA, 2004b.

[3] Olivier Chapelle, Bernhard Schölkopf, and AlexanderZien. Semi-Supervised Learning, MIT Press, 2006.

[4] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. JRSSB, 39:1–38, 1977.

[5] D. Martin and C. Fowlkes and D. Tal and J. Malik. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics,Vol. 2,pp. 416--423 ,July 2001, ICCV 2001.

[6] Arindam Banerjee, Srujana Merugu, Inderjit S. Dhillon, Joydeep Ghosh. Clustering with Bregman Divergences 6(Oct):1705--1749, 2005.