A NEW MULTIRESOLUTION ALGORITHM FOR IMAGE SEGMENTATION

M. Saeed[1], W. C. Karl[2,3], T. Q. Nguyen[2], H. R. Rabiee[4]

[1]*Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology*
[2]*Dept. of Electrical and Computer Engineering, Boston University*
[3]*Dept. of Biomedical Engineering, Boston University*
[4]*Media & Interconnect Technology Lab, Intel Corporation*

ABSTRACT

We present here a novel multiresolution-based image segmentation algorithm. The proposed method extends and improves the Gaussian mixture model (GMM) paradigm by incorporating a multiscale correlation model of pixel dependence into the standard approach. In particular, the standard GMM is modified by introducing a multiscale neighborhood clique that incorporates the correlation between pixels in space and scale. We modify the log likelihood function of the image field by a penalization term that is derived from a multiscale neighborhood clique. Maximum Likelihood (ML) estimation via the Expectation Maximization (EM) algorithm is used to estimate the parameters of the new model. Then, utilizing the parameter estimates, the image field is segmented with a MAP classifier. It is demonstrated that the proposed algorithm provides superior segmentations of synthetic images, yet is computationally efficient.

1. INTRODUCTION

The overall goal of image segmentation is to identify regions that display similar characteristics in some sense. Image segmentation algorithms accomplish this by assigning each pixel to be a member of one of K classes or homogeneous regions. Robust segmentation algorithms often utilize a parametric model of the image field [1]. A statistical model of the image field is in the form of a probability density function (*pdf*) of the pixel intensities. Often the parameters of the *pdf* are not known *a priori*, thus parameter estimation theory can be utilized to achieve an efficient and consistent estimate of the model parameters [2]. In this section, the Gaussian Mixture Model (GMM) is formally described. The objective is to provide a framework for modeling the image field. In subsequent sections, it is demonstrated how images can be segmented using ML parameter estimation of the GMM.

By using thee GMM. we assume the image field, $Y(i,j)$, consists of intensities from K different classes. As an example, in MR brain images these classes represent different tissues (i.e. White Matter, Gray Matter, and CSF). In the GMM, pixel intensities are assumed to be independent and identically distributed (*i.i.d*). The *i.i.d* assumption allows for simple computation with the well-known Gaussian density functions. The intensity of each class is characterized by one of K different Gaussian density functions. Therefore, the model for the data is given by

$$p(y_i) = \sum_{j=1}^{K} p(y_i = Y|k_i = j)p(k_i = j) \tag{1.1}$$

where $p(y_i = Y|k_i = j)$ is the conditional probability of each pixel and is defined by the Gaussian

$$p(y_i = Y|k_i = j) = N(\mu_j, \sigma_j^2) \tag{1.2}$$

and $p(k_i = j)$ is the prior probability that the class of pixel i is class j. It is pointed out that the notation emphasizes individual pixel statistics rather than the entire image. Thus (1.1) defines a Gaussian mixture.

To characterize the GMM, we define the parameter vector $\Phi = [\mu_1\ \sigma_1^2\ \mu_2\ \sigma_2^2\ ..\ \mu_k\ \sigma_k^2]^T$. Given the data and with knowledge of Φ, the *Maximum a posteriori* (MAP) estimate of the class, $\hat{k}_i$ at pixel i can easily be computed. The MAP estimate, $\hat{k}_i$, of the class of pixel i is defined as:

$$\hat{k}_i = \underset{j}{\operatorname{argmax}}\ p(k_i = j|y_i = Y) \tag{1.3}$$

We can proceed to segment an image by assigning class memberships to each pixel individually using the above MAP estimate of the pixel class.

In practice, the conditional density parameters, Φ_i (e.g. μ_i and σ_i^2), and prior probabilities, $p(k_i = j)$, are not known *a priori*. For these reasons, the Maximum Likelihood (ML) estimation technique is used to find the estimated value of Φ_i based upon data in the image field. Since the class correspondence of each pixel in the image field is not known *a priori*, ML estimation for the conditional density parameters is a challenging nonlinear optimization problem. An attractive iterative technique to solve this problem is the Expectation Maximization (EM) algorithm [1].

The GMM-based segmentation algorithm assumes neighboring pixels are independent and identically distributed (*i.i.d*). However, pixels in homogenous regions of most natural images are correlated with one another which leads the GMM-based algorithm to yield poor segmentations [3]. Markov random fields (MRF) have been used to model this correlation. MRF models are not computationally tractable, thus we propose a simplified multiresolution-based algorithm which incorporates neighboring pixel correlation to yield improved segmentations. The proposed scheme is an improvement to the method of Ambroise et al [4]. The next section shall formally describe the novel multiresolution algorithm.

2. THE MR EM ALGORITHM: A NEW MULTIRESOLUTION LIKELOHOOD FUNCTION

The point of departure for the proposed new multiresolution EM algorithm is a particular convenient form of the log likelihood equation arising in the standard GMM segmentation approach. In particular, it is demonstrated in [4] that the log likelihood function can be written as

$$L(Z,\Phi) = \sum_{k=1}^{K}\sum_{i=1}^{n} z_{ik}\log(p(k_i = k)p\langle y_i|\Phi_k,(y_i \in k)\rangle)$$

$$-\sum_{k=1}^{K}\sum_{i=1}^{n} z_{ik}\log(z_{ik}) \qquad (2.1)$$

Z is a matrix whose elements are all the z_{ik} of the image, where at the p^{th} iteration of the EM algorithm, z_{ik} is defined by

$$z_{ik}^{(p)} = \frac{p\langle k_i = k|\Phi^{(p)}\rangle p\langle y_i|k_i = k, \Phi^{(p)}\rangle}{\sum_{j=1}^{K} p\langle k_i = j|\Phi^{(p)}\rangle p\langle y_i|k_i = j, \Phi^{(p)}\rangle} \qquad (2.2)$$

Recall, that the EM algorithm iterates until the parameter matrix, Φ, converges to a local maxima of the log likelihood function. In the E-step of the EM algorithm, z_{ik} is the probability that pixel i belongs to class k. Thus, the outputs of the EM algorithm, Φ and Z, also maximize $L(Z,\Phi)$. The proof that the EM algorithm maximizes $L(Z,\Phi)$ can be found in [4].

As is apparent, $L(Z,\Phi)$ does not account for the spatial correlation of the data. We propose to modify the log likelihood equation (2.1), by the addition of a penalization term, $V(Z)$. The penalization term will bias the log likelihood of a pixel, i, belonging to the same class, k_a, of its neighbors. Thus, it is observed that $V(Z)$ is modifying the pdf to incorporate desirable correlation properties. This prior probability on the pixel class probability is of a Gibbs form and thus like an MRF on the class probabilities. The new log likelihood expression is given by the following equation

$$U(Z,\Phi) = L(Z,\Phi) + V(Z) \qquad (2.3)$$

The penalization term, $V(Z)$, will incorporate the quadtree data structure illustrated in Figure 1 as well as a simple "clique" or pixel neighborhood system. Within the same resolution, we define the neighborhood of a pixel, i, to be all pixels which are adjacent to pixel i (top, down, right, left, and diagonal). Furthermore, a pixel at resolution J-2, will be defined to have a neighborhood at resolution J-1 which consists of the parent of i as well

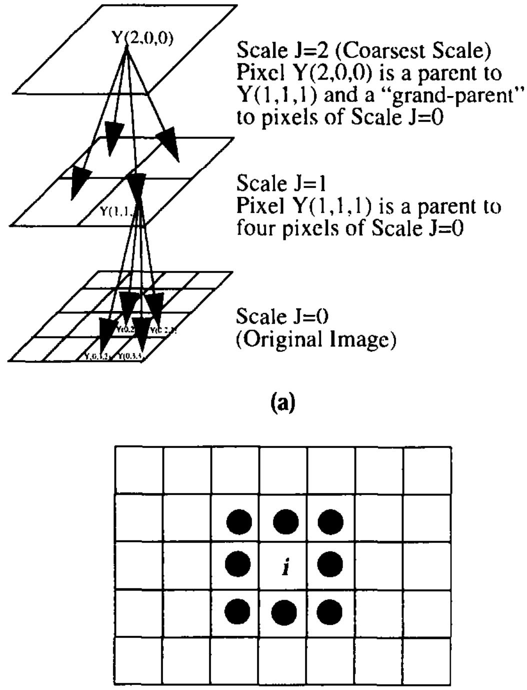


Figure 1 - Multiresolution Neighborhood Clique. (a) Multiresolution scales illustrating the parent-child structure. (b) Neighborhood of pixel ***i*** within the same scale is shown as small circles. Note, the total neighborhood of pixel ***i*** also includes its parents and grandparents as well as their respective neighborhoods.

as the parent's neighborhood. This neighborhood can be extended further across resolution to include the "grand-parents" of i at resolution J. In practice, we only incorporate the information from scales J = 0, 1, and 2. The neighborhoods at each resolution will have different weights in the neighborhood interaction weights of the penalization term. Let us define the following neighborhood interaction weight (NIW).

$$v_{irj} = \begin{cases} \alpha & \text{if pixel i and r are neighbors at resolution 0} \\ \beta & \text{if pixel i and r are neighbors at resolution 0 and 1} \\ \gamma & \text{if pixel i and r are neighbors at resolution 0 and 2} \\ 0 & \text{else} \end{cases} \qquad (2.4)$$

Using the NIW of (2.4), we propose the following penalization term

$$V(Z) = \sum_{j=0}^{J}\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{r=1}^{N} z_{jik}z_{jrk}v_{irj} \qquad (2.5)$$

Where z_{jik} is the probability of pixel, i, from resolution j being a member of class k. Note, that this changes the E-step of the EM algorithm such that z_{ik} is now defined by

$$z_{0ik} = z_{ik}^{(p)} =$$

$$\frac{p(k_i = k|\Phi^{(p)})p(y_i|k_i = k, \Phi^{(p)})\exp\left(\sum_{j=0}^{J}\sum_{r=1}^{n} z_{jrk}v_{irj}\right)}{\sum_{s=1}^{K} p(k_i = s|\Phi^{(p)})p(y_i|k_i = s, \Phi^{(p)})\exp\left(\sum_{j=0}^{J}\sum_{r=1}^{n} z_{jrs}v_{irs}\right)} \quad (2.6)$$

(Z) weights neighborhoods which have pixels that are members of the same class more than heterogeneous neighborhoods. Furthermore, it is observed that $V(Z)$ is only dependent on the probability matrix, Z, whose elements are the individual pixel probabilities, z_{jik}.

A modified version of the EM algorithm can be used to maximize the new penalized log likelihood equation, $U(Z,\Phi)$. We shall call the modified EM algorithm the *Multiresolution EM (MEM) algorithm*. The attractiveness of the MEM algorithm is in the approach of utilizing a multiresolution neighborhood. The coarser resolutions will allow for the segmentation of the more prominent features in the image. However, the information at the finer levels is important for accurate segmentation along boundaries and for segmenting highly detailed regions. Thus this has three advantages: 1) the MEM algorithm has desirable correlation properties and avoids blurring, 2) misclassifications are reduced, and 3) the MEM algorithm is computationally more tractable than MRF models. The subsection below presents an overview of the MEM algorithm.

3. THE MULTIRESOLUTION SEGMENTATION ALGORITHM

Given an image Y, a three-level Discrete Wavelet Transform (DWT) using the Haar basis is computed. The DWT will provide a collection of low-pass filtered images, $\{S_1, S_2\}$, where S_2 is the coarsest image as well as the original finest scale image, Y. We derive z_{ik1} and z_{ik2} using the conventional EM algorithm via the monoresolution Gaussian mixture model, and segment S_1 and S_2. Then after this information is provided, we can apply the MEM algorithm to Y. The MEM segmentation algorithm can be summarized as follows:

Step 1: Compute 3 level DWT to obtain S_1, S_2.
Step 2: Run standard EM on S_1 and S_2 to obtain z_{ik1} and z_{ik2}.
Step 3: Run MEM on Y using $U(Z,\Phi)$ as the log likelihood.
Figure 2 provides an illustration of the algorithm.

4. EXPERIMENTAL RESULTS

To demonstrate the robustness of the MEM algorithm and compare its performance against the traditional GMM-based segmentation algorithm, we applied the MEM algorithm to a test image. Specifically, the goal is to demonstrate that a multiresolution segmentation will result in a more accurate segmentation of the image field. The experimental results will demonstrate that while the NEM algorithm of Ambroise et al. performs better than the GMM-based segmentation algorithm, however the NEM's segmentations are not as accurate as the MEM algorithm's segmentations.

In the experiments, we segment the test image (a) of Figure 3. In test image (a), the three classes arise from three different Gaussian noise processes. There are two main challenges in segmenting the test image. The variances of the three classes were allowed to be large such that the pdf's of each Gaussian have a significant overlap. This is significant because once the parameters, Φ, are estimated, the MAP classifier becomes a simple minimum distance classifier. Thus, pixels are labelled to a class whose mean intensity they are nearest to. The GMM-based segmentation of this image results in many errors due to this overlap of pdf's. The other challenge of test image (a) is the 2-pixel wide horizontal line which runs through all the classes. Since the coarser resolutions of the MEM algorithm are in practice low-pass filtered versions of the original image, the line can be blurred and the segmentation map of the final image will have a poor labelling of the horizontal line. However, since the MEM algorithm utilizes information from all resolutions, the horizontal line is preserved in the segmentation.

The segmentations of the test image are shown in Figures 3(b)-3(d). Clearly, the GMM-based segmentation failed to demonstrate the spatial correlation existing between pixels of the same region due to the large variance of each class. The MEM algorithm provides a subjectively and quantitatively superior segmentation. The MEM algorithm is also compared to the Neighborhood EM (NEM) algorithm of Ambroise et al. [4]. While the NEM algorithm performs better than the conventional EM algorithm, the new MEM algorithm yields a more accurate segmentation map.

Figure 4 demonstrates an application of the MEM algorithm to the MR Brain image segmentation. The challenge in segmenting MR Brain images is to accurately label tissues such as white matter, gray matter, and cerebrospinal fluid. Figure 4(b) shows the segmentation map of the brain into the different tissue types.

Thus, the proposed MEM algorithm has been shown to segment both synthetic and real images accurately. Fine structure is preserved and the MEM algorithm incorporates the pixel correlation across space and scale which allows for better segmentations than the standard GMM. Moreover, the MEM algorithm is computationally efficient which is an important advantage when compared to MRF-based segmentation algorithms.

REFERENCES

[1] J. Zhang, J.W. Modestino, and D.A. Langan, "Maximum-Likelihood Parameter Estimation for Unsupervised Stochastic Model-Based Image Segmentation," IEEE Transactions on Image Processing, Vol. 3, No. 4, July 1994, pp. 404-419.

[2] A. Papoulis, "Probability, Random Variables, and Stochastic Processes," McGraw-Hill, third ed., 1991.

[3] M. Saeed, "ML Parameter Estimation of Mixture Models and its Application to Image Segmentation and Restoration," MIT S.M. Thesis, 1997

[4] C. Ambroise, and G. Govaert, "Spatial Clustering and the EM Algorithm," Conference Proceedings of International Workshop on Mixtures, Aussois, France, Sept. 17-21, 1995.

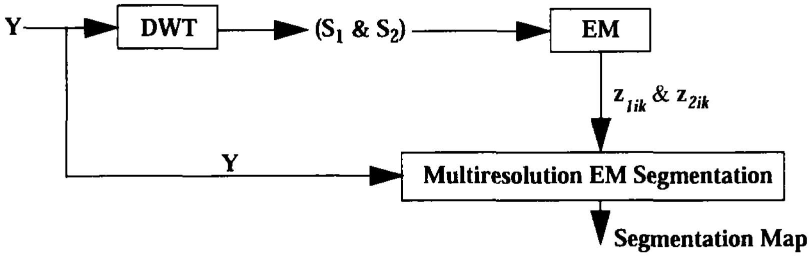
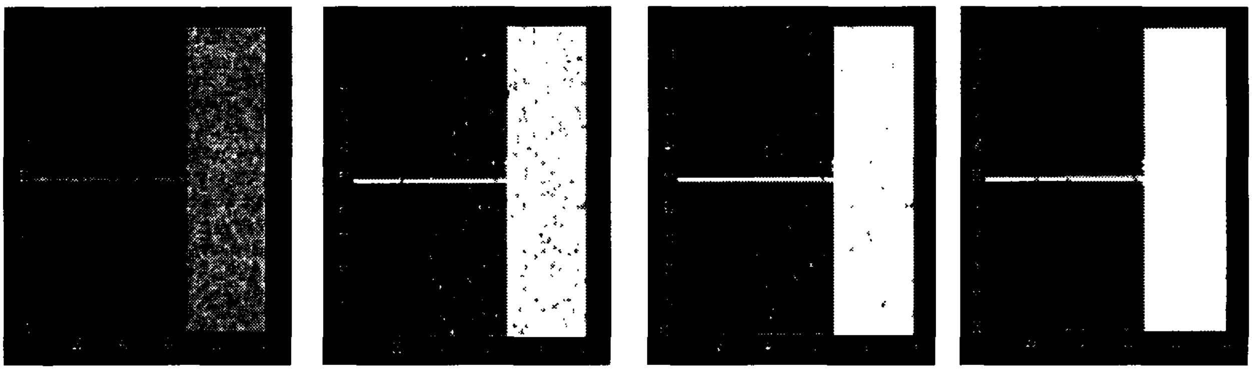
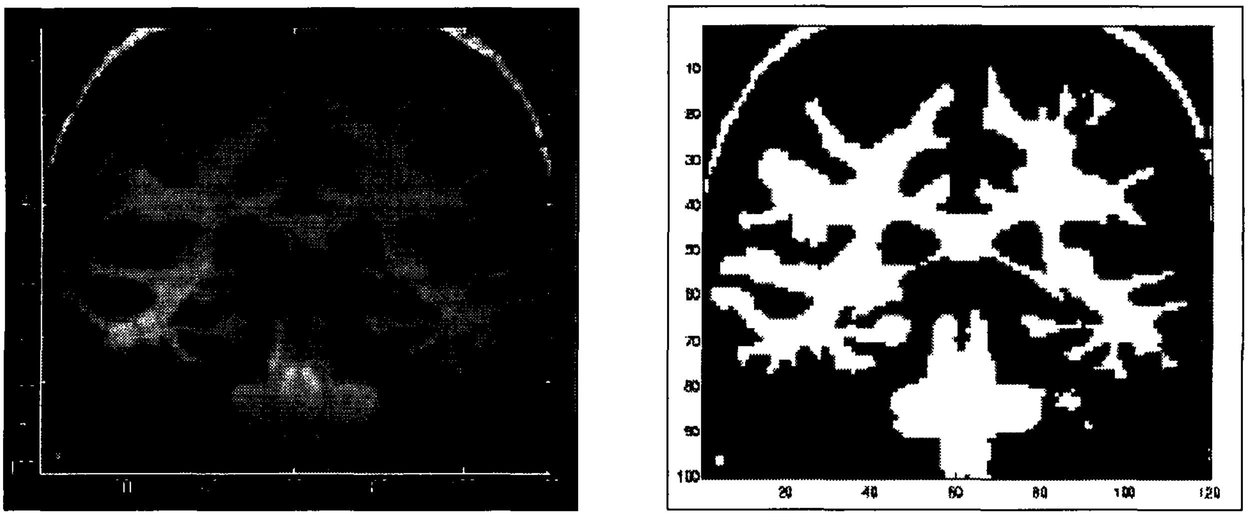


Figure 2 - Schematic of the MEM segmentation algorithm.



(a) (b) (c) (d)

Figure 3 -(a) Classes 1,2, and 3 and have means of 50, 100, and 150, respectively. All classes have a variance of 225 with a 2-pixel wide horizontal strip from class 3 running through the other classes. (b) Segmentation using conventional Gaussian Mixture Model. (c) Segmentation using modified Neighboring EM (NEM) Algorithm of Ambroise et al. (d) Segmentation using novel MEM algorithm.



(a) (b)

Figure 4 - MEM segmentation of MR brain image into white matter, gray matter, air, and cerebrospinal fluid (shown as black background). (a) Original MRI of brain, (b) Segmented brain image.