IMAGE SEGMENTATION BY A ROBUST MODIFIED GAUSSIAN MIXTURE MODEL

Hui Zhang[1,2], *Q. M. Jonathan Wu*[2], *Thanh Minh Nguyen*[2]

[1]School of Computer & Software, Nanjing University of Information Science & Technology, China
[2]Department of Electrical and Computer Engineering, University of Windsor, Canada

ABSTRACT

The Gaussian Mixture Model (GMM) with a spatial constraint, e.g. a Hidden Markov Random Field (HMRF), has been proven effective for image segmentation. However, the determination of parameter β in the HMRF model is, in fact, noise dependent to some degree. In this paper, we propose a simple and effective algorithm to make the traditional Gaussian Mixture Model more robust to noise, with consideration of the relationship between the local spatial information and the pixel intensity value information. The conditional probability of an image pixel is influenced by the probabilities of pixels in its immediate neighborhood to incorporate the spatial and intensity information. In this case, the parameter β can be assigned to a small value to preserve image sharpness and detail in non-noise images. At the same time, the neighborhood window is used to tolerate the noise for heavy-noised images. Thus, the parameter β is independent of image noise degree in our model. Finally, our algorithm is not limited to GMM–it is general enough so that it can be applied to other distributions based on the construction of the Finite Mixture Model (FMM) technique.

1. INTRODUCTION

Image segmentation is one of the most important and difficult problems in many applications, such as robot vision, object recognition, and medical image processing. Although different methodologies [1-6] have been proposed for image segmentation, it remains a challenge due to overlapping intensities, low contrast of images, and noise perturbation. One of the most widely used clustering models for image segmentation is the well-known Gaussian Mixture Model (GMM) [7-10]. The main advantage of the standard GMM is that it is easy to implement and the small number of parameters can be efficiently estimated by adopting the Expectation Maximization (EM) algorithm. However, as a histogram-based model, the GMM assumes that each pixel in an image is independent of its neighbors and does not take into account spatial dependencies. Thus, the performance of the GMM is sensitive to noise and image contrast levels. To overcome this shortcoming, a wide variety of approaches have been proposed to incorporate spatial information into the image. A common approach is the use of the Hidden MRF (HMRF) Model [11-13]. In the HMRF model, the spatial information in an image is encoded through the contextual constraints of neighboring pixels, which are characterized by conditional MRF distributions. Parameter estimation in HMRF models usually relies on Maximum Likelihood (ML) or Bayesian methods [14, 15]. Based on this well-known pseudo likelihood approximation [16], various HMRF model estimation approaches have been proposed [17-23].

However, HMRF resorts to the parameter β to control the degree of image smoothness. The chosen parameter has to be both large enough to tolerate the noise, and small enough to preserve image sharpness and detail. Thus, the parameter is selected generally based on experience. In this paper, we propose a robust Modified Gaussian Mixture Model (MGMM) to incorporate local spatial information and pixel intensity value information. The conditional probability of an image pixel is influenced by the probabilities of pixels in its immediate neighborhood. For spatial information, we add weighting for distant pixels in order to distinguish among the contributions of different pixels, as the weighted parameters decrease with increasing distance. In our model, the parameter β can be assigned a small value to preserve image sharpness and detail in non-noise images. At the same time, the conditional probability neighborhood window is used to tolerate the noise for heavy-noised images. The performance of our proposed approach, compared with state-of-the-art technology, demonstrates its improved robustness and effectiveness.

2. FINITE MIXTURE MODEL

Let us first consider two letters: $Q=\{1,2,\ldots,K\}$ and $L=\{1,2,\ldots,D\}$. Let S be a finite index set, $S=\{1,2,\ldots,N\}$. We shall refer to set S as the set of sites or locations. Let X and Y be two random fields, their state space $\mathcal{X}$ and $\mathcal{Y}$ are indexed by the supposed set of sites S (every site $i \in S$), given by $\mathcal{X}_i = \{x_i : x_i \in Q\}$, and $\mathcal{Y}_i = \{y_i : y_i \in L\}$.

Their product space $\mathcal{X} = \prod_{i=1}^{N} x_i$ and $\mathcal{Y} = \prod_{i=1}^{N} y_i$ shall be denoted as the space of the configurations of the state values of the considered site set, $\boldsymbol{x}=(x_i)$ and $\boldsymbol{y}=(y_i)$.

For image segmentation application, an image consisting of N pixels is segmented into K classes. y_i denotes the observation (intensity value) at the i-th pixel of an image and x_i denotes the corresponding class label of the i-th pixel. For every $j \in Q$ and $i \in S$, the probability

$$p(x_i = j) = \pi_j$$

is the prior distribution of the pixel y_i, belonging to the class x_i. y_i follows a conditional probability distribution $p(y_i | \theta_j)$, in which θ_j is the set of parameters. Specific to the GMM, the conditional probability $p(y_i|\theta_j)$ is selected as a Gaussian distribution. Under the independent assumption, the FMM can be calculated as

$$p(y_i \mid \boldsymbol{\pi}, \boldsymbol{\theta}) = \sum_{j=1}^{K} \pi_j p(y_i \mid \theta_j). \tag{1}$$

Although FMM is widely used for its simplicity and effectiveness as a model [1], it only describes the data statistically; no spatial information about the data is utilized. In other words, it does not take into account the spatial correlation between the neighboring pixels in the decision process. Thus, HMRF is selected as a useful tool in order to overcome the problems with the FMM, and to reduce the sensitivity of the segmentation result with respect to noise.

3. PROPOSED METHOD

Let y_i, with dimension d, i=(1,2,…,N), denote the intensity value at the i-th pixel of an image and j (j=1,2,…,K) denote the corresponding class label of the i-th pixel. In HMRF, the neighboring labels influence the class label of a pixel. Inspired by this idea, we calculate the conditional probability of the i-th pixel by its neighborhood probabilities. The modified Gaussian Mixture Model (MGMM) is defined as

$$p(y_i \mid \boldsymbol{\pi}, \boldsymbol{\theta}) = \sum_{j=1}^{K} \pi_{ij} \prod_{m \in \mathcal{N}_i} p(y_m \mid \theta_j)^{\frac{w_m}{R_i}}. \tag{2}$$

where $\mathcal{N}_i$ is the neighborhood of the i-th pixel; including the i-th pixel, this is called the conditional probability window (CPW). The probability $p(y_m \mid \theta_j)$ is the Gaussian distribution and R_i is the normalized factor, defined as

$$R_i = \sum_{m \in \mathcal{N}_i} w_m. \tag{3}$$

Our model degrades to the standard Gaussian Mixture Model when CPW=1×1; thus, the standard GMM can be considered as a special case of our model. A simple choice of w_m is that w_m=1 for all m-th pixels, and R_i equals the number of pixels in the CPW. However, to incorporate the spatial information and pixel intensity value information, the strength of w_m should decrease as the distance between pixel m and i increases. For this reason, we define w_m as the function of d_{mi}, which is the spatial Euclidean distance between pixels m and i.

$$w_m = \frac{1}{\left(2\pi\delta^2\right)^{1/2}} \exp\left(-\frac{d_{mi}^{\;2}}{2\delta^2}\right), \tag{4}$$

$$\delta = \frac{\text{size of CPW} - 1}{4}. \tag{5}$$

The prior probability π_{ij} in (2) represents the prior distribution of pixel y_i belonging to class j, which satisfies the constraint

$$0 \le \pi_{ij} \le 1 \text{ and } \sum_{j=1}^{K} \pi_{ij} = 1. \tag{6}$$

With the intrinsic relationship between the prior probability π_{ij} and the posterior probability, the prior probability can be expressed as

$$\pi_{ij} = \frac{\exp\left(\frac{\beta}{\mathcal{N}_i - 1} \sum_{m \in \partial_i} z_{mj}\right)}{\sum_{k=1}^{K} \exp\left(\frac{\beta}{\mathcal{N}_i - 1} \sum_{m \in \partial_i} z_{mk}\right)}. \tag{7}$$

where z_{ij} is the posterior probability $p(x_i = j \mid y_i)$ and $\partial_i = \mathcal{N}_i - \{i\}$ is the neighborhood of the i-th pixel, called the posterior probability window (PPW). It is noted that (7) is very similar to the HMRF, except with respect to the PPW sizes. For a one-dimensional chain, the Markov property means that the probabilistic behavior of the chain at some time i, given knowledge of its complete past, depends only on its state in the immediate past $i-1$ [24]. For a 2-D image, the definition of neighbors ($i-1$ and $i+1$) extends to horizontal, vertical, and diagonal pixels, which becomes a 3×3 square window. Thus, HMRF has to select only the 3×3 square window due to its Markov property. In our model, 3×3, 5×5, 7×7, etc. square windows can be used, and HMRF is a special case of our model when PPW=3×3.

In general, for heavier-noised images, a larger PPW can be used to obtain better segmentation results. Moreover, parameter β in (7) is set to a small value to preserve image sharpness and detail, and the CPW is used to simultaneously tolerate heavy noise. The details to demonstrate the relationship between CPW, PPW, and parameter β are illustrated in the experimental results presented in Section 4.

We then apply the EM algorithm for parameter learning in our model. First, according to [2], the complete-data log likelihood can be written as

$$Q = \sum_i \sum_j \sum_{m \in \mathcal{N}_i} z_{ij} \left[\log \pi_{ij} + \frac{w_m}{R_i} \log p(y_m \mid \theta_j)\right]. \tag{8}$$

In E-step, the posterior probability can be calculated as

$$z_{ij}^{(k+1)} = \frac{\pi_{ij}^{(k)} \prod_{m \in \mathcal{N}_i} p\left(y_m \mid \theta_j^{(k)}\right)^{\frac{w_m}{R_i}}}{\sum_{h=1}^{K} \pi_{ih}^{(k)} \prod_{m \in \mathcal{N}_i} p\left(y_m \mid \theta_h^{(k)}\right)^{\frac{w_m}{R_i}}}. \tag{9}$$

The M-step evaluates the mean and covariance as follows

$$\mu_j^{(k+1)} = \frac{\sum_{i=1}^{N} \sum_{m \in \mathcal{N}_i} z_{ij}^{(k)} \frac{w_m}{R_i} y_m}{\sum_{i=1}^{N} z_{ij}^{(k)}}. \tag{10}$$

$$\Sigma_j^{(k)} = \frac{\sum_{i=1}^{N} \sum_{m \in \mathcal{N}_i} z_{ij}^{(k)} \frac{w_m}{R_i} \left(y_m - \mu_j^{(k)}\right)\left(y_m - \mu_j^{(k)}\right)^T}{\sum_{i=1}^{N} z_{ij}^{(k)}}. \tag{11}$$

For a deep understanding of our algorithm, we summarize the computation process of our MGMM as follows

Algorithm: The EM Algorithm for MGMM

Step 1. Initialize the algorithm with the *k*-means method to obtain initial values; set CPW and PPW size; assign value of parameter β.

Step 2. In E-step, compute the prior probability $\pi_{ij}^{(k)}$ using (7) and calculate the posterior probability $z_{ij}^{(k)}$ using (9).

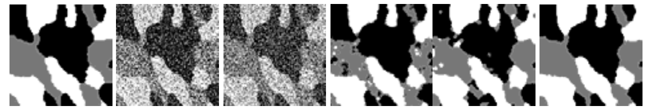
Step 3. Compute the quantities $\mu_j^{(k+1)}$ and $\Sigma_j^{(k+1)}$ using (10) and (11) according to the M-step.

Step 4. Terminate the iterations if the EM algorithm converges; otherwise, increase the iteration (k=k+1) and repeat steps 2-4.

4. EXPERIMENTAL RESULTS

In this section, we experimentally evaluate MGMM in a set of synthetic images and real images. We also evaluate GMM [2], FLICM [3] and HMRF-FCM [10] for comparison. The source codes for the FLICM and HMRF-FCM algorithms can be downloaded from the authors' websites [25-26].

In the first experiment, a three-class synthetic image (128×128, shown in Fig. 1 (a)) is used to compare the performance of the proposed method with others. Fig. 1 (b) shows the same image corrupted by Gaussian noise with zero mean and 0.25 variance. In order to evaluate the segmentation results, we employ the misclassification ratio (MCR) [13] in our experiments. The value of MCR is in the [0%-100%] range, where lower values indicate better segmentation performance. The segmentation results of the noised image (Fig. 1(b)) by GMM, FLICM, HMRF-FCM, MGMM are shown in Fig. 1(c)-(f). We set PPW=5×5, CPW=3×3, and parameter β=2.5 for MGMM. The class number is set to 3. As we observe, GMM does not segment images well. Although FLICM and HMRF-FCM can reduce the effect of noise to some extent, as [3, 10] claim, they are still sensitive to heavy noise and misclassify some portions of pixels, as shown in Fig. 1(d)-(e). However, we observe that the proposed MGMM yields outstanding segmentation results compared to the poor performance of their competitors, as seen in Fig. 1(f). The results obtained by different noise intensities are given in Table 1. As we observe, the proposed MGMM reaches the lowest MCR compared to the other methods.

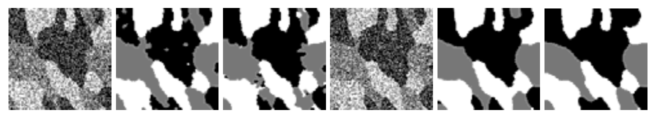


(a) (b) (c) (d) (e) (f)

Fig. 1. (a) Original three-class image (128×128 size); (b) Corrupted by Gaussian noise (zero mean, 0.25 variance); (c) GMM, MCR=51.14%; (d) FLICM, MCR=8.47%; (e) HMRF-FCM, MCR=8.72%; (f) MGMM, MCR=3.37%.

Table 1. The misclassification ratio (MCR %) of synthetic image with additive Gaussian noise for different methods

Methods	var=0.1	var=0.15	var=0.2	var=0.25
GMM	39.31	44.88	48.73	51.14
FLICM	4.14	5.23	7.14	8.47
HMRF-FCM	3.03	6.62	7.74	8.72
MGMM	1.44	2.08	2.72	3.37



(a) (b) (c) (d) (e) (f)

Fig. 2. Methods for different parameter β values. The MGMM is evaluated without CPW affect. (a) HMRF-FCM, β=5, MCR=43.62%; (b) HMRF-FCM, β=10, MCR=4.68%; (c) HMRF-FCM, β=15, MCR=5.27%. (d) MGMM, β=5, MCR=38.89%; (e) MGMM, β=10, MCR=4.86%; (f) MGMM, β=15, MCR=6.82%.

(a) (b)

Fig. 3. MGMM with PPW=5×5 and CPW=3×3 for different parameter β values. (a) MGMM, β=2.5, MCR=3.37%; (b) MGMM, β=3, MCR=2.86%.

For deep investigation of the influence by parameter β, CPW and PPW, we re-evaluate the performance of different methods for heavy-noised image segmentation (shown in Fig. 1(b), corrupted by Gaussian noise with zero mean and 0.25 variance). For fair evaluation of parameter β in HMRF and in our models, we adopt the implementation in [1] instead of mean-field-like algorithm for the prior probability approximation of the HMRF-FCM method. Image segmentation by HMRF-FCM with different parameter β values is shown in Fig. 2(a)-(c). It is observed that the small β value (β=5 in Fig. 2(a)) works poorly for heavy-noised images and the large β values (β=10 and 15 in Fig. 2(b)-(c)) are available for image segmentation. This result is expected, and has been proven by previous research: parameter β has to be large enough to tolerate the noise. In this case, β is 10 to strike a good balance between robustness and noise, and to be effective in preserving the details of the image. To evaluate the impact of PPW, we use MGMM with PPW=5×5, without considering CPW (CPW=1×1), by different values of parameter β. The segmentation results are shown in Fig. 2(d)-(f). Comparing Fig. 2(b)-(c) with Fig.

2(d)-(f), we can observe that a larger PPW can improve the result, but not significantly. However, without the help of CPW, the segmentation result in Fig. 2(d) is still pooled for the small parameter value β=5. Finally, we evaluate the proposed MGMM with PPW=5×5 and CPW=3×3 by different values of parameter β. The segmentation results are shown in Fig. 3(a)-(b). It can be seen in Fig. 3 that CPW improves significantly with the small parameters (β=2.5 and β=3) in MGMM to preserve more image sharpness and detail. From Fig. 2-3, it can be seen that the proposed model yields a nearly perfect result, outperforming the traditional HMRF method with the lowest MCR.

We then evaluate the performance of the proposed MGMM based on a subset of the Berkeley image dataset [27], which is comprised of a set of real-world color images along with segmentation maps provided by different individuals. We employ the Probabilistic Rand (PR) index [28] to evaluate the performance of the proposed method, with the multiple ground truths available for each image within the dataset. The PR index takes values between 0 and 1, with values closer to 0 (indicating an inferior segmentation result) and values closer to 1 (indicating a better result).

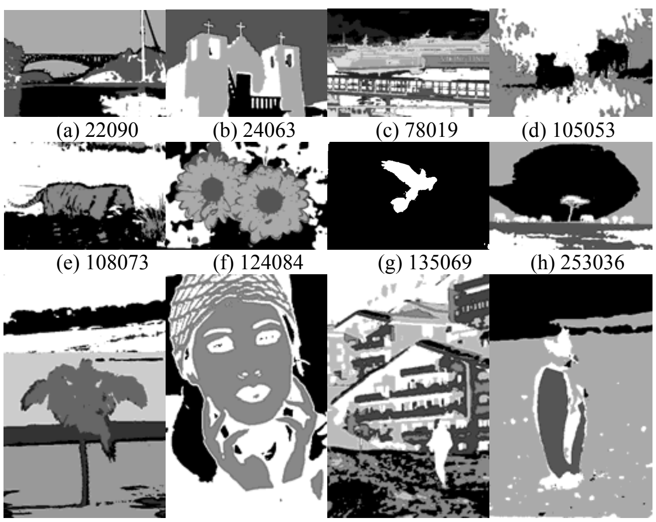


Fig. 4. Berkeley Image segmentation results by MGMM.

Fig. 4 shows the image segmentation results of Berkeley image data. Table 2 presents the average PR values for all methods. Compared to other methods, the proposed MGMM yields the best segmentation results with the highest PR values. We also evaluate the computation time for all methods in the previous experiment. The average computation time of the different methods is presented on the last line of Table 2. It is noted that the computation of our MGMM is slower than GMM, but is still much faster than other methods. Compared to other methods, our models can be calculated more quickly and achieve the best segmentation results.

Table 2. Comparison of different methods for Berkeley image dataset, Probabilistic Rand (PR) Index.

Image #	Class	GMM	FLICM	HMRF-FCM	MGMM
108073	4	0.5792	0.5824	0.5855	0.6305
124084	4	0.6971	0.6	0.7203	0.7310
135069	2	0.979	0.9834	0.9873	0.9733
302003	3	0.7018	0.7172	0.7169	0.7021
105053	3	0.5389	0.51	0.5546	0.6234
22090	4	0.7644	0.7675	0.7777	0.8157
46076	6	0.8226	0.8381	0.8602	0.8984
61086	5	0.6880	0.7213	0.7220	0.7397
106025	4	0.8198	0.7833	0.8241	0.8528
253036	4	0.6856	0.704	0.7181	0.7257
78019	7	0.8153	0.8162	0.8231	0.8378
24063	4	0.7790	0.8139	0.8559	0.8676
Mean		**0.7392**	**0.7364**	**0.7621**	**0.7832**
Computation time		13.96s	49.51s	113.09s	24.80s

5. CONCLUSION

In this paper, we propose a simple and effective algorithm to make the traditional Gaussian Mixture Model more robust to noise, with consideration of the relationship between the local spatial information and pixel intensity value information. The conditional probability of an image pixel is replaced by the calculation of the probabilities of pixels in its immediate neighborhood. We also investigate the influences of PPW, CPW, and parameter β for image segmentation. PPW can improve segmentation results, but not significantly. The parameter β in traditional HMRF is noise dependent to some degree. It has to keep a balance between robustness to noise and effectiveness in preserving the details of the image. In our method, CPW is used to erase the image noise and β is set at 2.5 for MGMM to keep image detail. Thus, our model eliminates the shortcoming of choosing a proper β, and can tolerate noise and preserve image detail simultaneously. Finally, our model is simple and easy to implement, and it can be quickly applied to image segmentation.

6. ACKOWLEDGEMENTS

This work was supported in part by the Canada Chair Research Program and the Natural Sciences and Engineering Research Council of Canada. This work was supported in part by the National Natural Science Foundation of China under Grant 61105007 and 61103141.

7. RELATION TO PRIOR WORK

The work presented here has focused on the image segmentation by incorporating some spatial constraints. Though the basic concept of adding spatial information to GMM is not new [2, 10], the originality of the proposed algorithm is encoding local spatial information by using a conditional probability window.

8. REFERENCES

[1] G. McLachlan and D. Peel, Finite Mixture Models. John Willey and Sons, New York, 2000.

[2] C.M. Bishop, Pattern Recognition and Machine Learning. Springer, 2006.

[3] Krinidis S., and Chatzis V., “A Robust Fuzzy Local Information C-means Clustering Algorithm,” IEEE Trans. Image Process., no. 19, vol. 5, pp. 1328-1337, 2010.

[4] M. Ahmed, S. Yamany, N. Mohamed, A. Farag, and T. Moriarty, “A Modified Fuzzy C-means Algorithm for Bias Field Estimation and Segmentation of MRI Data,” IEEE Transactions on Medical Imaging, vol. 21, pp. 193-199, 2002.

[5] S. Chen and D. Zhang, “Robust Image Segmentation using FCM with Spatial Constraints based on new Kernel-induced Distance Measure,” IEEE Transactions on Systems, Man and Cybernetics, vol. 34, no. 4, pp. 1907-1916, 2004.

[6] W. Cai, S. Chen, and D. Zhang, “Fast and Robust Fuzzy C-means Clustering Algorithms Incorporating Local Information for Image Segmentation,” Pattern Recognition, vol. 40, no. 3, pp. 825-838, March 2007.

[7] C. Carson, S. Belongie, H. Greenspan, and J. Malik, “Blobworld: Image Segmentation Using Expectation-Maximization and its Application to Image Querying,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 8, pp. 1026-1038, Aug. 2002.

[8] M.N. Thanh, Q.M. J. Wu and S. Ahuja, “An Extension of the Standard Mixture Model for Image Segmentation,” IEEE Trans. Neural Netw., vol. 21, no. 8, pp. 1326–1338, 2010.

[9] M.N. Thanh and Q.M. J. Wu, “Gaussian-Mixture-Model-Based Spatial Neighborhood Relationships for Pixel Labeling Problem,” IEEE Trans. Syst. Man Cybern. B, vol. PP, no. 99, pp. 1–10, 2011.

[10] S. P. Chatzis and T. A. Varvarigou, “A Fuzzy Clustering Approach Toward Hidden Markov Random Field Models for Enhanced Spatially Constrained Image Segmentation,” IEEE Trans. Fuzzy Syst., vol. 16, no. 5, pp. 1351–1361, Oct. 2008.

[11] P. Clifford, “Markov Random Fields in Statistics,” in Disorder in Physical Systems. A Volume in Honour of John M. Hammersley on the Occasion of His 70th Birthday, G. Grimmett and D. Welsh, Eds. Oxford, U.K.: Clarendon Press, Oxford Science Publication, 1990.

[12] L.R. Rabiner, “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition,” Proc. IEEE, vol. 77, no. 2, pp. 257-286, 1989.

[13] Y. Zhang, M. Brady, and S. Smith, “Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation Maximization Algorithm,” IEEE Trans. Med. Imag., vol. 20, no. 1, pp. 45–57, Jan. 2001.

[14] Dempster P., Laird N. M., and Rubin D. B., “Maximum Likelihood from Incomplete Data via EM Algorithm,” J. Roy. Stat. Soc. B, vol. 39, no. 1, pp. 1–38, 1977.

[15] G.E.B. Archer and D.M. Titterington, “Parameter Estimation for Hidden Markov Chains,” J. Statistical Planning Inference, 2002.

[16] J. Besag, “Statistical Analysis of Non-Lattice Data,” The Statistician, vol. 24, pp. 179-195, 1975.

[17] J. Zhang, J. W. Modestino, and D. Langan, “Maximum-likelihood Parameter Estimation for Unsupervised Stochastic Model-based Image Segmentation,” IEEE Trans. Image Process., vol. 3, pp. 404–420, July 1994.

[18] G. J. McLachlan and T. Krishnan, “The EM Algorithm and Extensions,” in Series in Probability and Statistics. New York: Wiley, 1997.

[19] Z. Zhou, R. Leahy, and J. Qi, “Approximate Maximum Likelihood Hyperparameter Estimation for Gibbs Priors,” IEEE Trans. Image Process., vol. 6, no. 6, pp. 844-861, 1997.

[20] J. Besag, “On the Statistical Analysis of Dirty Pictures,” J. Roy. Stat. Soc. B, vol. 48, pp. 259–302, 1986.

[21] W. Qian and D. Titterington, “Estimation of Parameters in Hidden Markov Models,” Philos. Trans. Roy. Soc. London A, vol. 337, pp. 407–428, 1991.

[22] J. Zhang, “The Mean Field Theory in EM Procedures for Blind Markov Random Field Image Restoration,” IEEE Trans. Image Process., vol. 2, no. 1, pp. 27–40, Jan. 1993.

[23] F. Forbes and N. Peyrard, “Hidden Markov Random Field Model Selection Criteria Based on Mean Field-like Approximations,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 9, pp. 1089–1101, 2003.

[24] M.N.M. van Lieshout, “Markovianity in Space and Time,” Dynamics and Stochastics: Festschrift in Honor of Michael Keane, pp. 154-167, 2006.

[25] http://www.cs.uoi.gr/~kblekas/sw/MAPsegmentation.html

[26] http://web.mac.com/soteri0s/Sotirios_Chatzis/Software.html

[27] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics,” in Proc. 8th IEEE Int. Conf. Comput. Vis., Vancouver, BC, Canada, vol. 2, pp. 416–423, 2001.

[28] R. Unnikrishnan and M. Hebert, “Measures of Similarity,” in Proc. IEEE Workshop Comput. Vis. Appl., pp. 394–400, 2005.