ELSEVIER

October 1994

Pattern Recognition Letters

Pattern Recognition Letters 15 (1994) 969–976

A systematic way for region-based image segmentation based on Markov Random Field model

Il Y. Kim, Hyun S. Yang *

Center for Artificial Intelligence Research, Department of Computer Science, KAIST, 373-1 Kusung-Dong, Yusung-Ku, Taejon, South Korea 305-701

Received 20 March 1993; revised 11 April 1994

Abstract

In this paper, we propose a Markov Random Field model-based approach as a systematic way for integrating constraints for robust image segmentation. In our approach, the image is first segmented into a set of disjoint regions by one of the region-based segmentation techniques which operates on image pixels, and a Region Adjaceny Graph (RAG) is then constructed from the resulting segmented regions based on the spatial adjacencies between regions. Our approach is then applied by defining an MRF model on the corresponding RAG. Constraints for improving the segmentation results are incorporated into an energy function via clique functions and optimal segmentation is then achieved by finding a labeling configuration which minimizes the energy function through simulated annealing.

Keywords: Image segmentation; Region Adjacency Graph; Markov Random Field; Clique functions; Simulated annealing

1. Introduction

Image segmentation is the process of segmenting an image into a set of disjoint regions whose characteristics such as intensity, color, texture, etc. are similar. Image segmentation is the first stage of the image understanding task and severely affects the later process, image interpretation, by providing useful image structures such as regions and edges. In general, there are two important criteria to be considered in image segmentation; one is the homogeneouity of the region and the other is the discontinuity between adjacent disjoint regions. Although there is a wide variety of image segmentation techniques which are well surveyed in (Haralick and Shapiro, 1985), it is still difficult to satisfy all the properties for the optimal set of segmented regions. For example, in typical region-growing image segmentation techniques (Haralick and Shapiro, 1985), the resulting segmented image depends on the predetermined threshold values and so the algorithm often fails to split the regions that must be separated or fails to merge the regions that need not be separated because the information about the uniformity in a region corresponding to an object surface and the discontinuity between regions corresponding to different objects' surfaces are not well incorporated into the algorithms.

Since the Markov Random Field (MRF) model has powerful capability to integrate various visual information, it has been recently applied to early vision problems such as image restoration and segmentation (Geman and Geman, 1984; Derin et al., 1984; Lakshmanan and Derin, 1989; Jain and Nadabar, 1990; Dubes et al., 1990; Daily, 1989; Gamble et al.,

* Corresponding author.

SSDI 0167-8655(94)00028-2

1989). But due to the large number of pixels on which the MRF is to be defined, it usually takes a lot of computational time to get optimal labels, which makes it difficult to apply the MRF model to the real scene domains. Therefore in our MRF-based image segmentation scheme we define an MRF model on a set of preliminary segmented regions resulted from one of the region-based segmentation methods rather than image pixels, by which the number of sites to be considered is significantly reduced. In measuring region uniformity we consider several features such as average intensity, color, texture and so on. And we also introduce a *boundary process* which reflects average discontinuities in intensity value across the common boundary between two adjacent regions. In finding the optimal segmentation result we use the simulated annealing algorithm which is timely complexive but can be modified to find a sub-optimal solution with practical computational time.

This paper is organized as follows. In Section 2 is described how to apply the MRF model to the image segmentation problem by defining an MRF on the RAG. In Section 3 we explain the proposed image segmentation scheme and its results are presented in Section 4. Finally we conclude in Section 5.

2. Defining MRF on RAG

In this section, we describe how to define an MRF on the corresponding Region Adjacency Graph (RAG) which is constructed from preliminary segmented regions.

2.1. Region Adjacency Graph

Before defining an MRF on RAG we assume that the image is preliminarily segmented into a set of disjoint regions. The corresponding RAG is then constructed from those regions based on spatial adjacencies between regions. In Fig. 1 is shown a typical segmented image and corresponding Region Adjacency Graph.

2.2. MRF definitions on RAG

Let $G=\{R, E\}$ be a RAG, where

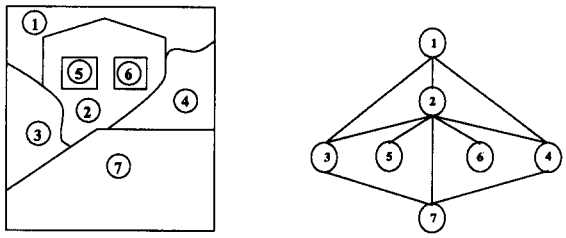

Fig. 1. An example of segmented image (a) and corresponding RAG (b).

$$R=\{R_i: 1\leqslant i\leqslant N\} \tag{1}$$

is a set of disjoint regions; and E is a set of edges connecting them. Suppose that there exists a neighborhood system on G, denoted by

$$\Gamma=\{n(R_i): 1\leqslant i\leqslant N\}\,, \tag{2}$$

where $n(R_i)$ is the set of regions in R that are neighbors of R_i. Let $X=\{X_i, 1\leqslant i\leqslant N\}$ denote any family of random variables each X_i of which is associated with region R_i, and $\Lambda=\{\lambda_1, ..., \lambda_m\}$ be a set of possible labels so that $X_i\in\Lambda$ for all i. Let Ω be the set of all possible *configurations*:

$$\Omega=\{\omega=(x_1, ..., x_N): x_i\in\Lambda, 1\leqslant i\leqslant N\}\,. \tag{3}$$

As usual, the event $\{X_1=x_1, ..., X_N=x_N\}$ is abbreviated to $\{X=\omega\}$. Then X is an MRF with respect to Γ if

$$P(X=\omega)>0 \quad \forall\omega\in\Omega\,; \tag{4}$$

$$P(X_i=x_i \mid X_j=x_j, R_j\neq R_i)$$

$$=P(X_i=x_i \mid X_j=x_j, R_j\in n(R_i))$$

for every $R_i\in R$ and $(x_1, ..., x_N)\in\Omega$. And $P(\cdot)$ and $P(\cdot \mid \cdot)$ are the joint and conditional probability density functions, respectively.

An alternate characterization is given by the Hammersley–Clifford theorem: X is a Markov Random Field if and only if the probability distribution $P(X=\omega)$ is a Gibbs distribution

$$P(\omega)=\frac{1}{Z}\exp-\frac{1}{T}U(\omega) \tag{5}$$

where T is the temperature, U is an energy term, and Z is a normalizing constant. Let c denote a clique which is defined to be a fully connected subgraph of

G. U is then obtained by summing over applicable clique potentials $V_c(\omega)$:

$$U(\omega) = \sum_{c \in C} V_c(\omega) , \tag{6}$$

where C is the set of all cliques contained in the graph G. In the work presented here, we wish to take into account not only the prior probabilities of particular configurations, but also external evidence. The external evidence for a particular label ω_s at a site s after an observation O_s is given by the likelihood $P(O_s \mid \omega_s)$. If we assume that these likelihoods are conditionally independent such that

$$P(O \mid \omega) = \prod_{s \in S} P(O_s \mid \omega_s) \tag{7}$$

and start with Bayes' rule, it can be seen that this posterior probability also has the Gibbs distribution form (see (Regier, 1991) for a proof):

$$P(\omega \mid O) = \frac{1}{Z'} \exp - \frac{1}{T} U^p(\omega \mid O) . \tag{8}$$

Since our goal is to assign each region a region label, the problem configuration can be stated as follows:

$$X = \{X_1, X_2, ..., X_N\} , \tag{9}$$

$$\Lambda = \{1, 2, ..., N\} \quad \text{so that } X_i \in \Lambda \text{ for all } i ,$$

where N is the number of preliminarily segmented regions.

Then the problem is to find the optimal region labeling configuration given the observation O on and between regions corresponding to the maximum of $P(\omega \mid O)$, and $P(\omega \mid O)$ can be maximized by minimizing Gibb's energy function, $U(\omega \mid O)$. Since the energy function is again the summation of $V_c(\omega \mid O)$'s, we can encode the various constraints for acquiring robust segmentation results by defining clique functions appropriately.

3. Image segmentation based on MRF model

The overall scheme of our image segmentation scheme is depicted in Fig. 2.

Before applying our segmentation scheme to the given image we should first segment a given image into a set of disjoint regions. To do this we use one of

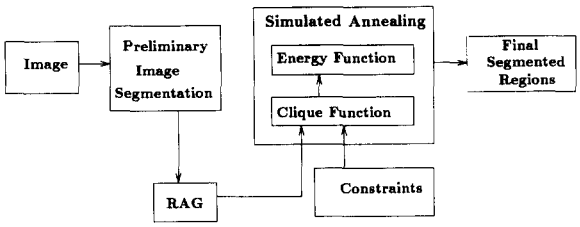

Fig. 2. Overview of the proposed image segmentation scheme.

the region-based segmentation algorithms which is called *Partition-Mode-Test* proposed by Suk and Jung (1983). The PMT algorithm is quite fast because it performs image segmentation by just one scan of the whole image but its result is still dependent on two threshold values. When preliminarily segmenting the image, over-segmentation (splitting the regions that need not be separated) is assumed because we do not consider of the recovery from the under-segmentation (failure to split the regions that must be separated). After segmenting the given image, we construct the corresponding RAG based on the spatial adjacencies between the segmented regions.

3.1. Defining clique functions

Before defining clique functions we impose the following constraints that the optimally segmented regions should have.

- A single segmented region should be uniform in spectral features such as intensity, color, texture and so on.
- On the common boundary between two different segmented regions there should exist strong discontinuities in intensity value.

3.1.1. Region process

Let $F = \{F_1, F_2, ..., F_N\}$ be a set of region processes each F_i of which represents average values on a set of visual data observed on the region R_i. Since we measure several spectral features on the region, a region process on the region R_i can be defined again as $F_i = \{F_i^1, F_i^2, ..., F_i^M\}$, where M is the number of features of the region.

Now if we consider the first constraints of the above, a clique function for a corresponding clique c can be defined as

$$V_c(\omega \mid F) = \sum_{R_i \in c} \sum_{R_j \in c, R_j \neq R_i} \eta_{ij} \max_{1 \leq k \leq M} |F_i^k - F_j^k| , \tag{10}$$

where R_i and R_j are regions belonging to the clique c. η_{ij} is a binary variable which has the value 1 if X_i and X_j have the same region label and value 0 otherwise. Since we only consider the set of adjacent two-region pairs each of which is directly connected in the RAG, the above equation simply becomes

$$V_c(\omega \mid F) = \eta_{ij} \max_{1 \leq k \leq M} |F_i^k - F_j^k| , \tag{11}$$

where the clique c contains two regions, R_i and R_j.

As features on the region itself, we use several spectral feature descriptions which are summarized in Table 1.

However, with only the above definition, the global energy has the minimum when all the segmented regions have distinct region labels different from adjacent ones because η will be 1 only when two adjacent regions have the same region label. Therefore we should introduce the boundary process as in the following section.

3.1.2. Boundary process

Let $B = \{B_{ij}: 1 \leq i, j \leq N\}$ be a set of boundary processes between two adjacent regions, R_i and R_j. By the second constraint we imposed in the above, B_{ij} can be defined as

$$B_{ij} = \frac{1}{n_{ij}} \sum_{p \in D_{ij}} e_p \tag{12}$$

where D_{ij} is the set of pixels composing the common boundary between regions R_i and R_j, n_{ij} the cardinality of D_{ij} and e_p the edge magnitude on the pixel p. That is, the boundary process represents the average discontinuity on the common boundary between two adjacent regions. In the conventional region-based segmentation algorithm, there are often the cases that a single object surface whose intensity values are smoothly changing is splitted into disjoint regions although there is no salient discontinuity between them. So by incorporating the boundary processes into the clique functions we can avoid such an erroneous problem.

Now the clique function with boundary process can be defined again as

$$V_c(\omega \mid F, B) = \eta_{ij} \max_{1 \leq k \leq M} |F_i^k - F_j^k| + \alpha(1 - \eta_{ij}) \frac{1}{B_{ij}} \max_{1 \leq k \leq M} |F_i^k - F_j^k| , \tag{13}$$

where α is a parameter for controlling weights between region and boundary processes. For convenience, let us simplify the above equation as:

$$V_c(\omega \mid F, B) = \eta_{ij} FT_{ij} + \alpha(1 - \eta_{ij}) ST_{ij} . \tag{14}$$

Then the clique function has the value FT_{ij} if the two regions R_i and R_j have been assigned the same region label, and αST_{ij}, otherwise. So if we want to get the optimal set of segmented regions, we should tune the weighting parameter α so that the two terms in the clique function may behave as in Table 2.

For determining the appropriate value of parameter α, we have got samples sets of segmented regions, estimated region and boundary processes, and known that the two constraining terms show desirable be-

Table 1
Descriptions of the features measured on a region R_k: (R, G, B) = average values of three color planes of pixels in the region, f_i = intensity value of a pixel at location i, N_i = a set of four neighboring pixels of pixel at i

Feature	Definition
r	$\frac{R}{R+G+B}$
g	$\frac{G}{R+G+B}$
b	$\frac{B}{R+G+B}$
intensity	$\frac{r+g+b}{3}$
r-b	$r-b$
saturation	$\frac{\max(r,g,b)-\min(r,g,b)}{\max(r,g,b)}$
smoothness	$\frac{1}{\text{area}(R_k)} \sum_{i \in R_k} \sum_{j \in N_i} \frac{\lvert f_i - f_j \rvert}{4}$

Table 2
Desired behaviors of FT_{ij} and αST_{ij}

R_i, R_j	Desirable behavior
should be merged	$FT_{ij} < \alpha ST_{ij}$
should not be merged	$FT_{ij} > \alpha ST_{ij}$

haviors when the parameter α has the value around 0.03.

3.2. Estimation by simulated annealing

Now we should find the optimal region labeling configuration that minimizes the energy function defined as

$$U(\omega \mid F, B) = \sum_{c \in C} V_c(\omega \mid F, B) ,$$

$$V_c(\omega \mid F, B) = \eta_{ij} \max_{1 \le k \le M} |F_i^k - F_j^k| + \alpha(1-\eta_{ij}) \frac{1}{B_{ij}} \max_{1 \le k \le M} |F_i^k - F_j^k| . \quad (15)$$

But a simple exhaustive search procedure will result in an exponential complexity of $O(N^N)$. So we use the simulated annealing algorithm (Kirkpatrick et al., 1983) that has been widely used for solving such a combinatorial optimization problem and are guaranteed to find a global minimum by following the specific temperature schedule (Geman and Geman, 1984) which is not usually taken for its time complexity. The annealing procedure can be described as follows.

- *Stage* 1. Set initial temperature T and set X_i in X to $i \; \forall i$.
- *Stage* 2. For each X_i in X,
  - 2.1. Change its label to any one of Λ.
  - 2.2. Calculate the amount of change in the energy function, ΔU.
  - 2.3 If $\Delta U < 0$ accepted the current ω; otherwise, accept it if $e^{-\Delta U/T} > \xi$, where ξ is a randomly generated number from the uniform distribution over [0, 1].
- *Stage* 3. If the energy becomes stabilized, exit; otherwise, lower the temperature by $0 < \kappa < 1$ such that $T_{k+1} = \kappa T_k$ and go back to Stage 2.

4. Experimental results

In Fig. 3 we show some experimental results on natural outdoor scene images. In the leftmost column of the figure are shown the images exploited to evaluate our proposed image segmentation scheme. As in the image of the topmost row, the upper-right part of the road is very bright and the darker its brightness is the farther it is from the upper-right due to the lighting condition. This style of changes of brightness causes a single object surface to be divided into several disjoint regions although there are no salient discontinuities between those regions as is shown in the middle column of the figure.

Now we improve the preliminary segmentation result by our proposed scheme. First we construct a RAG from the preliminarily segmented regions and initially assign labels to all the nodes (regions) in the RAG. We then find optimal labeling configurations by minimizing the given energy function through simulated annealing. In the annealing procedure, we set the initial temperature T to 1 and κ to 0.96. The boundaries of the segmented regions according to the final labeling configuration are depicted in the rightmost column of the figure.

In Fig. 4a we present the energy height graph showing the estimated value of the energy function from the initial to the final state in the annealing process. And also in Fig. 4b we show the energy values at the *exit* point of the annealing process in several executions under the same initial conditions such as initial temperature and κ except initial labels. Although we did not follow the specific temperature schedule for finding global optimal solutions (Geman and Geman, 1984), it can be seen that we reach the minimum energy state in most cases.

Finally we compare the proposed scheme with the *probabilistic relaxation* method (Hummel and Zucker, 1983). In probabilistic relaxation (PR), a priori probability that a site X_i has a certain label λ is represented as $p_i(\lambda)$, compatibility between labels λ and λ' as $r_{ij}(\lambda, \lambda')$ and influential structures between sites as C_{ij}. Thus if we set $p_i(\lambda)$ to equal value for all sites and labels, evaluate $r_{ij}(\lambda, \lambda')$ as the clique functions we defined and set C_{ij} to 1 whenever site X_i and X_j are directly connected in the RAG, we can implement the given image segmentation problem as the PR method.

In probabilistic relaxation, the final label of the site X_i is determined to be the label λ that maximizes $p_i(\lambda)$. And those a priori probabilities are iteratively updated through the relaxation process. In Fig. 5 we compare the energy values at the final states of segmented regions resulted from the proposed scheme and the PR method. As in the figure, the PR method

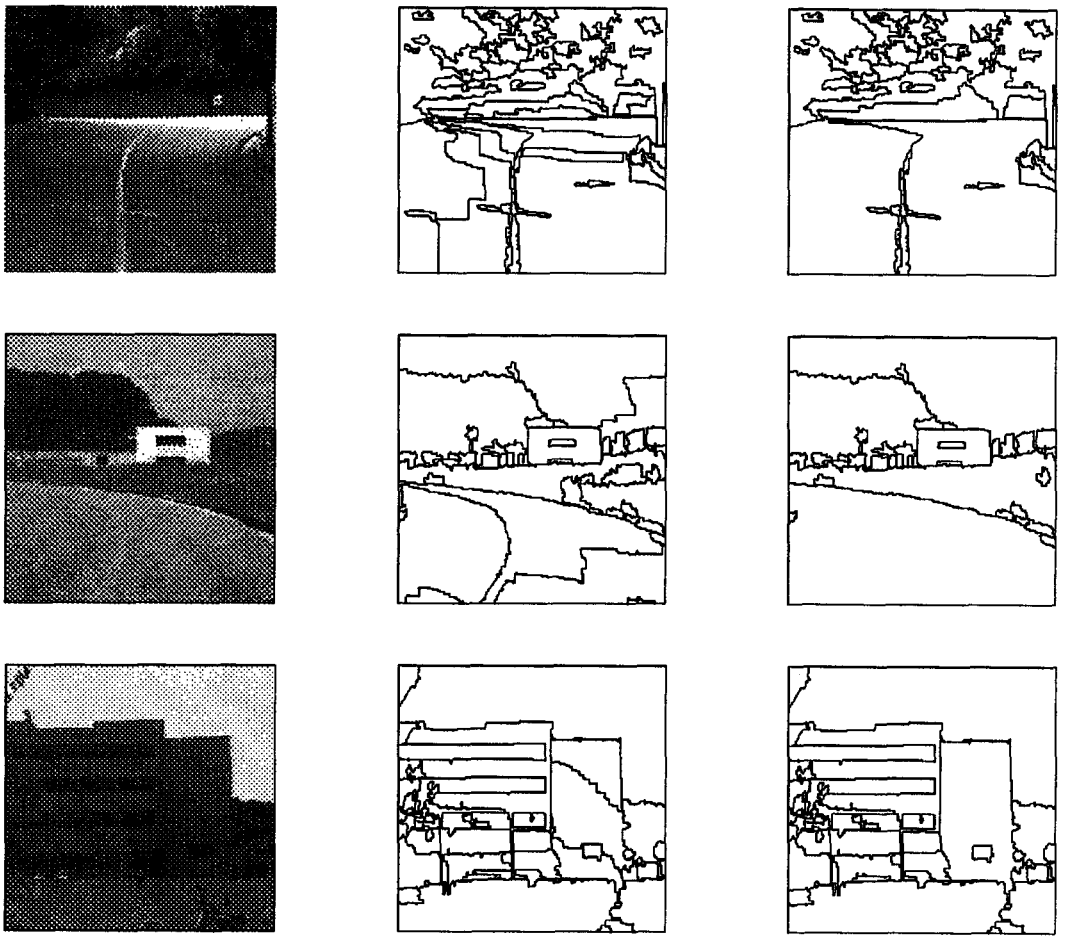

Fig. 3. Experimental results on real outdoor scene images.

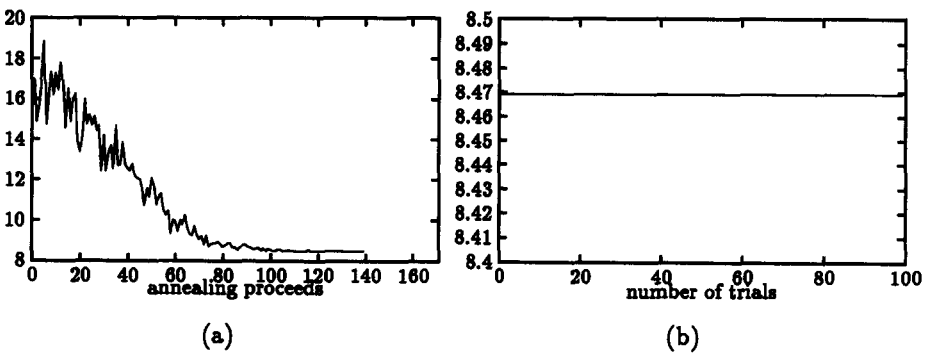

Fig. 4. (a) Energy values from the initial to the final labeling state. (b) Energy values when annealing stops at several experiments under the same initial conditions ($T=1.0$, $\kappa=0.96$).

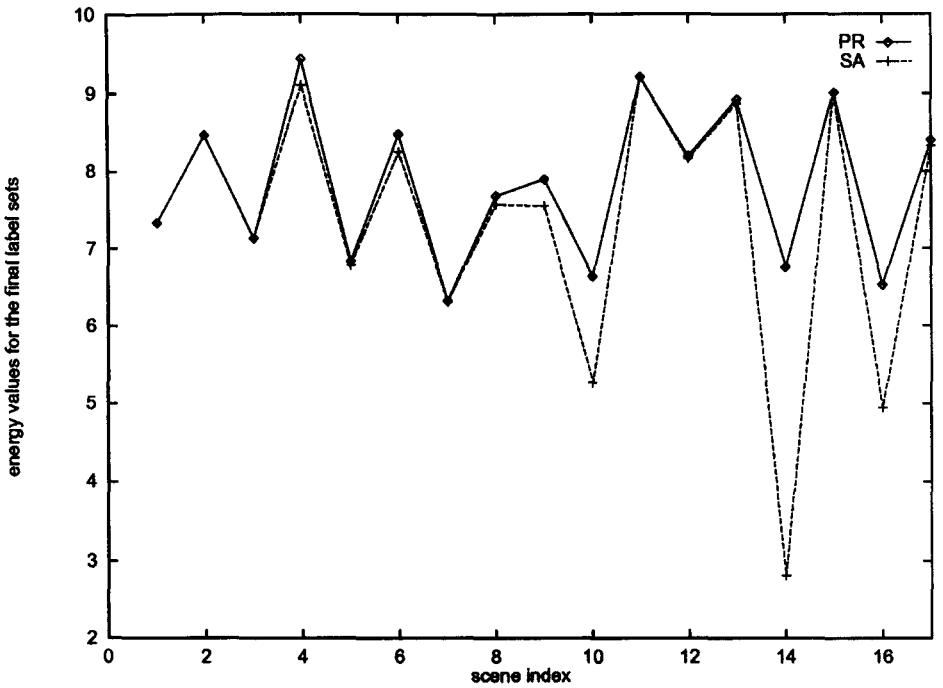

Fig. 5. Simulated annealing vs. probabilistic relaxation. Energy values are estimated for the resulting set of final labels using the clique functions we defined.

sometimes falls in a local minimum and thus produces a set of lables which is not globally optimal with respect to the given clique functions although it is timely efficient rather than simulated annealing.

5. Conclusions

In this paper, we have proposed a robust image segmentation scheme based on the Markov Random Field model. We defined an MRF on the preliminarily segmented regions and introduced region and boundary processes for measuring uniformity and discontinuity information in various spectral features between regions, which has not been well incorporated in region-based segmentation techniques. We then acquired suboptimal segmentation results by minimizing the energy functions through the simulated annealing algorithm.

References

Daily, M.J. (1989). Color image segmentation using Markov Random Field. *Proc. Internat. Conf. on Computer Vision and Pattern Recognition*, 304–312.

Derin, H., H. Elliott and D. Geman (1984). Bayes smoothing algorithms for segmentation of binary images modeled by Markov Random Fields. *IEEE Trans. Pattern Anal. Machine Intell.* 6, 707–720.

Dubes, R.C., A.K. Jain, S.G. Nadabar and C.C. Chen (1990). MRF model based algorithms for image segmentation. *Proc. Internat. Conf. on Pattern Recognition*, 808–814.

Gamble, E.B., D. Geiger and T. Poggio (1989). Integration of vision modules and labeling of surface discontinuities. *IEEE Trans. Syst. Man Cybernet.* 19, 1576–1581.

Geman, S., and D. Geman (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.* 6, 721–741.

Haralick, R.M., and L.G. Shapiro (1985). Image segmentation techniques. *Computer Vision, Graphics, and Image Processing* 29, 100–132.

Hummel, R. and S. Zucker (1983). On the foundations of relaxation labeling processes. *IEEE Trans. Pattern Anal. Machine Intell.* 5, 267–287.

Jain, A.K. and S.G. Nadabar (1990). MRF model-based segmentation of range images. *Proc. Internat. Conf. on Pattern Recognition*, 667–671.

Kirkpatrick, S., C.D. Gelatt and M.P. Vecchi (1983). Optimization by simulated annealing. *Science* 220, 671–680.

Lakshmanan, S. and H. Derin (1989). Simultaneous parameter estimation and segmentation of Gibbs Random Fields using simulated annealing. *IEEE Trans. Pattern Anal. Machine Intell.* 11, 799–813.

Regier, T. (1991). Line labeling using Markov Random Fields. Technical Report, UC Berkeley, TR-91-059.

Suk, M.S. and S.M. Jung (1983). A new image segmentation technique based on Partition Mode Test. *Pattern Recognition* 16, 469–480.