Research Report 321

Textured Image Segmentation Using Multiresolution Markov Random Fields and a Two-component Texture Model

Chang-Tsun Li and Roland Wilson

RR321

In this paper we propose a multiresolution Markov Random Field (MMRF) model for segmenting textured images. The Multiresolution Fourier Transform (MFT) is used to provide a set of spatially localised texture descriptors, which are based on a two-component model of texture, in which one component is a deformation, representing the structural or deterministic elements and the other is a stochastic one. Stochastic relaxation labelling is adopted to maximise the likelihood and assign the class label with highest probability to the block (site) being visited. Class information is propagated from low spatial resolution to high spatial resolution, via appropriate modifications to the interaction energies defining the field, to minimise class-position uncertainty. Experiments on the segmentation of natural textures are used to show the potential of the method.

Department of Computer Science
University of Warwick
Coventry CV4 7AL
United Kingdom

March 1997

Textured Image Segmentation Using Multiresolution Markov Random Fields and a Two-component Texture Model

Chang-Tsun Li and Roland Wilson
Department of Computer Science,
University of Warwick,
Coventry CV4 7AL, UK
ctli@dcs.warwick.ac.uk, rgw@dcs.warwick.ac.uk

November 21, 1996

Abstract

In this paper we propose a multiresolution Markov Random Field (MMRF) model for segmenting textured images. The Multiresolution Fourier Transform (MFT) is used to provide a set of spatially localised texture descriptors, which are based on a two-component model of texture, in which one component is a deformation, representing the structural or deterministic elements and the other is a stochastic one. Stochastic relaxation labelling is adopted to maximise the likelihood and assign the class label with highest probability to the block (site) being visited. Class information is propagated from low spatial resolution to high spatial resolution, via appropriate modifications to the interaction energies defining the field, to minimise class-position uncertainty. Experiments on the segmentation of natural textures are used to show the potential of the method.

Keyword: Markov random fields, texture segmentation, deformable templates

1 Introduction

Most attempts to segment or classify textures are based on either statistical or structural descriptions [5]. Statistical approaches such as co-occurrence matrices and auto-regressive models represent texture by statistics extracted from local image measurements. Generally, they are good for textures with random spatial arrangements. Structural approaches consider a textured image as composed of repeating texture elements like a tiled wall. Among the statistical approaches, Markov Random Fields (MRF's) have gained much attention in recent years [1] [2] [3] [7] [8] [9] [10]. Bouman and Shapiro used sequential maximum a posteriori (MAP) estimation in conjunction with a multi-scale random field (MSRF) [1] , which is a sequence of random fields at different scales. Geman et al. [3] used the Kolmogorov-Smirnov non-parametric measure of difference between the distributions of spatial features extracted from pairs of blocks of pixel gray levels, with MAP estimation of the boundary. Panjwani et al. [9] adopted an MRF model to characterise textured colour images in terms of spatial interaction within and between colour planes. Structural methods, by contrast, have received scant attention in recent years, largely because of their relative inflexibility.

In this paper, we describe an attempt to marry the generality and power of multiresolution MRF methods with a description which combines a structural element based on deformable templates [11] and a statistical element. The overall framework is one of MAP classification on a multigrid, using simulated annealing. The descriptors used for classification are based on a two-component model of texture: a 'deterministic' component based on the affine deformation of a patch of texture and a 'statistical' component based on the local Fourier energy spectrum. By combining these approaches, with the multiresolution Fourier transform (MFT) as the analysis tool, we obtain a computationally efficient yet general approach to the segmentation of arbitrary textured fields. After a brief outline of the theory, we present results showing the effectiveness of the method in segmenting natural textures.

2 Multiresolution Markov Random Fields

The feature of a Markov Random Field which makes it attractive in applications is that the state of a given site depends explicitly only on interactions with its neighbours [4]. We model an image as a sequence of MRF's, conforming to a quadtree structure, with a nominal top level l having $2^l \times 2^l$ *sites* and a number of levels $k, l < k \leq N$, each of which has four times more sites than its immediate ancestor. The neighbourhood structure we impose consists of five pixels: the 4-neighbours on the same level and the *father* on the level above:

$$\mathcal{N}_{i,j,k} = \{(i-1,j,k),(i+1,j,k),(i,j-1,k),(i,j+1,k),(\lfloor i/2 \rfloor,\lfloor j/2 \rfloor,k-1)\} \quad (1)$$

where $\lfloor . \rfloor$ denotes the floor of a real number. Each site on level k represents a square region of nominal size $2^{N-k} \times 2^{N-k}$ pixels, from which texture measurements are taken. (In fact, windows with a 50% overlap are used to ensure that the measurements vary smoothly across the image) . The interaction energies defining the MRF at level k in the tree are based on pairwise interactions:

$$U_{i,j,k}(\lambda) = \sum_{m=1}^{M} \alpha_m \sum_{(p,q,r)\in\mathcal{N}_{i,j,k}} U_m(\lambda,\lambda_{p,q,r},X_{i,j,k},X_{p,q,r}) \quad (2)$$

where $\lambda_{i,j,k}, 1 \leq \lambda \leq L$ is the (class) label at (i, j, k) and the sum over m allows for M texture descriptors to be used. The pairwise interactions U are functions of suitably defined *differences* between the measurements $X_{i,j,k}$ at the two sites and the labels at the two. Sampling is then based on the corresponding Gibbs distribution

$$P(\lambda) \propto e^{-\frac{U(\lambda)}{T}} \tag{3}$$

where the position indices have been suppressed and T is the scale parameter, or *temperature*, which is varied using the annealing schedule [4]

$$T_k(i) = \frac{C_k}{\log(1+i)} \tag{4}$$

where C_k is a constant and i is the iteration number.

Although this appears as a simple adaptation of the conventional sampling procedure for MRF's, one significant difference is the *causal* processing across scale implied by the inclusion of the father in the neighbourhood. In effect, the configuration at level k acts to condition the sampler at level $k+1$, giving the multiresolution algorithm:

1. Initialise at top level (l) with a random labelling of sites (i, j, l).
2. For level $k \geq l, k \leq N$
   (a) Copy intitial labels from fathers on level $k-1$.
   (b) Sample at every site on level k using measurements $X_{i,j,k}$ in annealing schedule given in (4) until there is no change over I_k iterations over the whole image at that scale.

3 Texture Measures Based on a Two-component Model

While the above algorithm provides a general framework for segmentation, its effectiveness depends critically on the texture descriptors used. We have four local measurements, which are based on the 'deterministic+stochastic' decomposition, which is a generalisation of the Wold decomposition of signals. The four components are:

1. Differences between the average gray level in the blocks. In effect, this uses the local 'd.c.' component in the MFT, which is computed using a gray level pyramid.

2. Two measures associated with the deterministic component, based on an affine deformation model

$$f_s(\vec{\xi}) = f_{s'}(\boldsymbol{A}^{-1}(\vec{\xi} - \vec{\chi})) + \nu_s(\vec{\xi}) \tag{5}$$

where site $s' = (l, m, k)$ is a 4-neighbour of site $s = (i, j, k)$. $\boldsymbol{A}$ is that 2×2 non-singular linear co-ordinate transform and $\vec{\chi}$ that translation which together give the best fit in terms of total deformation energy between the two patches. These are identified using the method described in [6], which makes use of local Fourier spectra calculated at the appropriate scale using the MFT. The deformation energy consists of:

   (a) The deformation term $||\boldsymbol{A} - \boldsymbol{I}||^2$ represents the amount of 'warping' required to match the given patch using its neighbour.

   (b) The error term $||\nu_s(\vec{\xi})||^2$ is the average residual error in the approximation.

3. A measure for the stochastic component, based on differences in the spectral energy densities estimated at each site via the MFT, $|\hat{f}(\vec{\xi}, \vec{\omega}, \sigma)|$. This is similar to many texture classification methods based on local spectra, Gabor filters or autocovariance estimates.

Each of these measures is scaled by the corresponding (within-class or between-class) sample variance and the four are added with appropriately chosen weights α_m to give the final interaction energy. Only the gray level difference is used for the father interaction, however. In order to obtain a MAP estimate, it is necessary to weight the probabilities in (3) by *priors* based on the likelihood of a pair of neighbours belonging to the same class. These probabilities are estimated directly from the data during the sampling process, as are the within-class and between-class variances. At levels $k > l$, the priors take into account the classification on the previous level, $k - 1$: the probability that a child has the same class as its father is given by

$$P(\lambda_s = \lambda_{f(s)}) = 1 - \rho^{d_s}, \tag{6}$$

where $\rho < 1$ is a constant and d_s is the shortest distance between site s and a site having a different class, ie. it represents distance to the boundary. This amounts to assuming that sites far from the boundary are likely to belong to the same region as their fathers. By varying ρ, it is possible to accommodate the appearance of regions too small to register at the largest scales. In the experiments reported below, $\rho = 0.5$.

4 Experiments

Two images, *Image I* and *Image II*, of size 256×256, each consisting of two textures, were used to test the method (Figure 1 (a) and Figure 2 (a)). The algorithm started

Table 1: Constants C_k and I_k used in experiment and iterations to segm ent images of Figure1(a) and 2(a)

image	*Image I*			*Image II*		
level (k)	C_k	I_k	**iteration**	C_k	I_k	**it eration**
3	8	3	424	8	3	806
4	6	1	30	6	1	38
5	4	1	47	4	1	51
6	4	1	64	4	1	62

at level 3 (8×8 blocks) and stopped at level 6 (64×64 blocks). The segmentation results at different scales are shown as Figure 1 (b) to (e) and Figure 2 (b) to (e). We used different grey values to represent different classes. For example, in Figure 1(b). three grey values are used to represent three classified regions. The white overlays on the original images (Figure 1 (a) and Figure 2 (a)) are the boundaries estimated at level 6. Table 1 summarises the two tests. showing the number of iterations per level and the parameters T_k, I_k controlling the annealing. Note that one iteration is a scan across the whole image at the given level, so that a child level involves four times as many updates as its father. However, many fewer iterations are needed at the smaller scales, whose labelling is constrained by that at the higher levels. Correspondingly, the sampling process runs at lower temperatures at these scales.

The textures in *Image I* are somewhat more structured than those in *Image II*. Consequently, in the first image the energy of the deformation between blocks belonging to the same texture is small relative to that between textures, facilitating segmentation using the deformation component. Although the degree of randomness of the two textures in the second image is higher than that in the first, the algorithm still copes well. However, the total iteration count to convergence is higher at larger scales, as can be seen from Table 1. Note that in this table, although it appears that the bulk of the computation is performed at the higher resolutions (k larger), most of the pixels at these resolutions assume stable states quickly: it is only in the boundary region that changes occur over several iterations. Figure 1 shows clearly the benefit multiresolution approach. Due to the low position resolution, three regions emerge along the boundary at the top level (see Figure 1(b)), although the classification of the blocks away from the boundary is satisfactory. In effect, the boundary blocks are treated as a separate class - not an unreasonable or unfamiliar problem, but none the less undesirable. The degree of misclassification is reduced at level 4 and the the boundary is further refined at lower levels.

5 Conclusions

In this paper, we have presented a novel approach to texture segmentation combining two important ideas: multiresolution MRF's to control the segmentation process and a two-component texture model, in which a deformable template is used to model the structural element of the texture and the energy spectrum is used to capture the stochastic element. In effect, this separation of the classification model from the texture model creates a highly flexible and general segmentation tool. Moreover, the effectiveness of the method has been shown with results from examples using natural textures.

Nevertheless, much remains to be done before the method can show its full potential. For example, no attempt has been made to model the boundary explicitly, through a line process, although that is clearly feasible within the general MMRF framework [4]. The choices of number of classes, L and the parameters C_k, I_k are currently made empirically. Moreover, we have only tested the method on a comparatively small set of images. Work is currently in progress to address these issues.

References

[1] C. A. Bouman and M. Shapiro. A Multiscale Random Field Model for Bayesian Image Segmentation. *IEEE Transactions on Image Processing*, 3:162–176, 1994.

[2] F. S. Cohen and Z. Fan. Maximum Likelihood Unsupervised Textured Image Segmentation. *Computer Vision, Graphics, and Image Processing*, 54:239–251, 1992.

[3] D. Geman, C. Graffigne S. Geman, and P Dong. Boundary Detection by Constrained Optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:609–628, 1990.

[4] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[5] R. Haralick. Statistical and Structural Approaches to Texture. *IEEE*, 67(5):610–621, 1979.

[6] T. I. Hsu. *Texture Analysis and Synthesis using Multiresolution Fourier Transform*. PhD thesis, Department of Computer Science, The University of Warwick, UK, 1994.

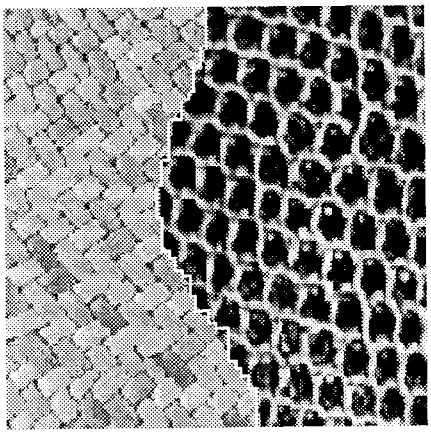
[7] S. M. Lavalle and S. A. Hutchinson. A Bayesian Segmentation Methodology For Parametric Image Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:211–217, 1995.

[8] B. S. Manjunath and R. Chellappa. Unsupervised Texture Segmentation Using Markov Random Field Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:478–482, 1991.

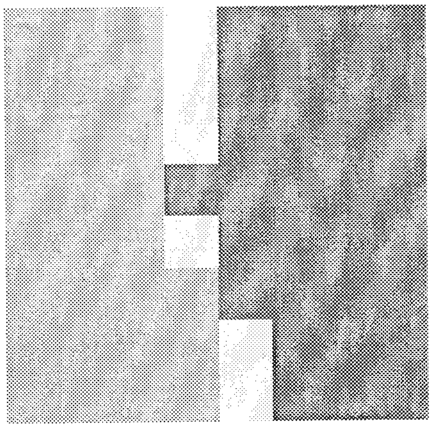
[9] D. K. Panjwani and G. Healey. Markov Random Field Models For Unsupervisd Segmentation Of Textured color Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:939–954, 1995.

[10] R. Szeliski. *Bayesian Modeling of Uncertainty in Low-level Vision*. Kluwer Academic Publishers, 1989.

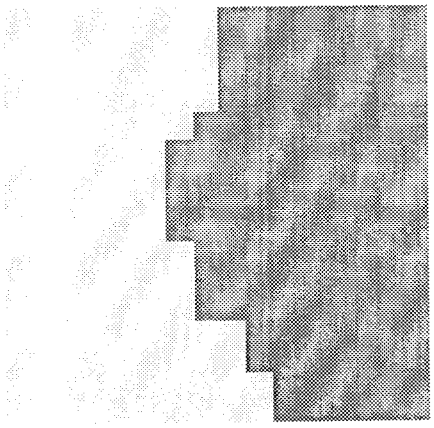
[11] A.J. Yuille and J.J. Clark. Bayesian Models, Deformable Templates and Competitive Priors. In *Spatial Vision in Humans and Robots*, ed. by L Harris and M Jenkin, Cambridge UP, 1993.



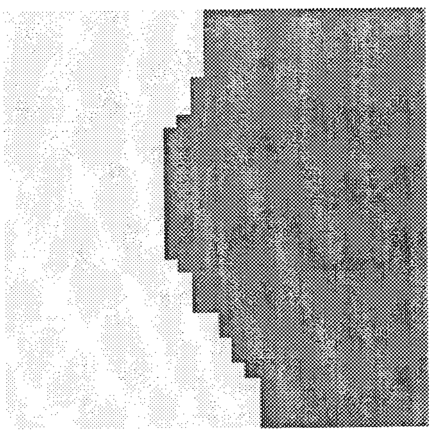
(a) Image I (size 256 X 256)



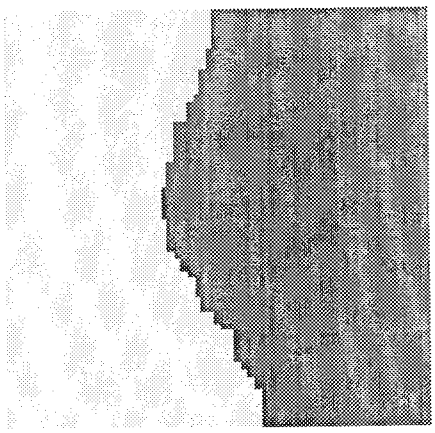
(b) Segmentation at level 3 (16 times enlarged)



(c) Segmentation at level 4 (8 times enlarged)

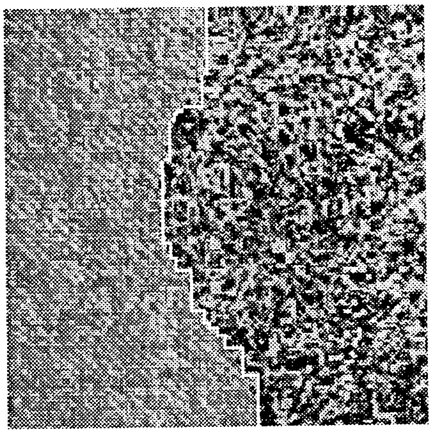


(b) Segmentation at level 5 (4 times enlarged)

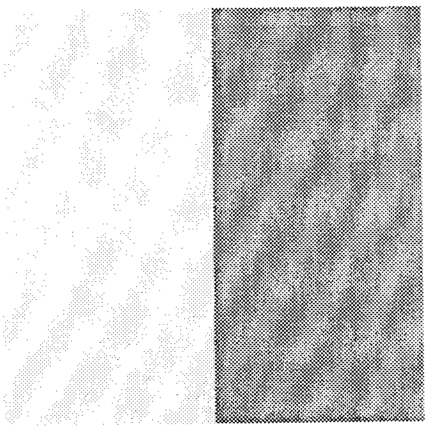


(c) Segmentation at level 6 (2 times enlarged)

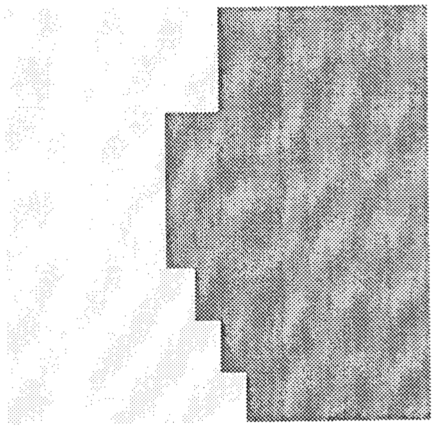
Figure 1: *Image I* and the segmentation results at two different levels. Estimated boundary shown in white in (a).



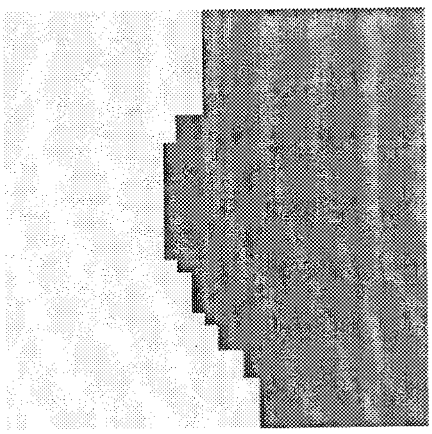
(a) Image II (size 256 X 256)



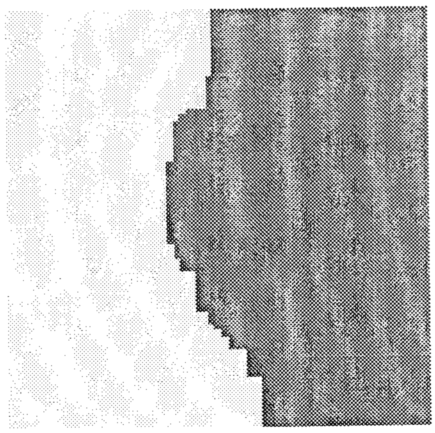
(b) Segmentation at level 3 (16 times enlarged)



(c) Segmentation at level 4 (8 times enlarged)



(b) Segmentation at level 5 (4 times enlarged)



(e) Segmentation at level 6 (2 times enlarged)

Figure 2: *Image II* and the segmentation results at two different levels. Estimated boundary shown in (a).