Image Segmentation Based On The Multiresolution Fourier Transform and Markov Random Fields

Guo-Huei Chen, Roland Wilson
Department of Computer Science,
University of Warwick,
Coventry CV4 7AL

September 21, 1998

Abstract

In this work, the Multiresolution Fourier Transform (MFT) and Markov Random Fields (MRFs) are combined to produce as a tool for image segmentation. Firstly, a Laplacian Pyramid is used as a high-pass filter. Then, the MFT is applied in order to segment images based on the analysis of local properties in the spatial frequency domain. A methodology for edge detection in image segmentation in the Bayesian framework using Markov random field models is then developed. Stochastic Relaxation is also adopted to maximise the likelihood and find the globally minimum energy states using simulated annealing.

Contents

List of Figures

1 Introduction

Image segmentation is a critical technique for image analysis and it has been deeply studied. Generally, there are two main approaches to image segmentation: those using boundary information and those using regional information. The boundary-based approaches involve the detection of luminance discontinuities such as lines and edges, and attempt to estimate their orientation and position. Noise or random fluctuations interfere with the detection of the actual features. Various approaches have been tried in an attempt to achieve noise immunity, such as averaging over a larger region. However increasing the averaging region size causes loss of positional resolution.

On the other hand, region-based approaches include thresholding the grey level image or classifying various attributes, such as granularity, directionality and regularity for texture images [1]. Generally, they all attempt to group together the pixels or image blocks of similar characteristics. The main difficulty in region-based image segmentation is that the classification requires global information.

The implication of this is that the two types of feature refer to different aspects of the information being sought about given primitives. The features used in a boundary-based approach are confined to a small area in order to provide accurate positional information. Conversely, features used in the region-based approach provide information about the class of a region based upon some 'global' property of the pixels. This would inevitably lead to a contradiction. A possible way of avoiding this is to seek a representation of the image (via a suitable transformation) that provides information about all perceptually important features. This leads to multiresolution techniques [1] [2] [3], which are used to allow a trade-off between resolution in class space and position for image segmentation. A particular multiresolution method, the MFT, enables the analysis to be carried out over a range of different levels with kernels/windowing functions of various sizes. By varying the resolution in both the spatial and spatial frequency domains, the uncertainty is confined to a reasonable extent while the computational efficiency is maintained. Markov Random Field models provide a general and natural model for the interaction between spatially related random variables, and there is a relatively flexible optimization algorithm, simulated annealing, that can be used to find the globally optimal realization that, in this case, corresponds to the maximum a posteriori(MAP) interpretation [4]. The work presented in this paper, based broadly on the work of Li [5], incorporates regional and boundary information into the interaction energy function of the multiresolution Markov Random Field (MMRF) [4] and performs the segmentation in a unified process. Detailed descriptions of MFT and MRFs are given in Sections 2, 3 and 4.

2 Multiresolution Fourier Transform

The MFT is based upon the Short Time Fourier Transform (STFT) and is a generalisation of multiscale methods. The basic idea is to combine a set of STFT's into a single hierarchical transform[1][2][6]. Figure 1 shows the structure of the spatial and spatial frequency diagram for 2-D MFT at different levels. In the spatial domain the bottom level is the original image. In the spatial frequency domain the top level is the DFT of the original image, while intermediate levels are supersets of STFT's with differing resolutions. The resolution changes by a factor of two between levels in both domains. This enables the MFT to embody local Fourier transforms over a range of

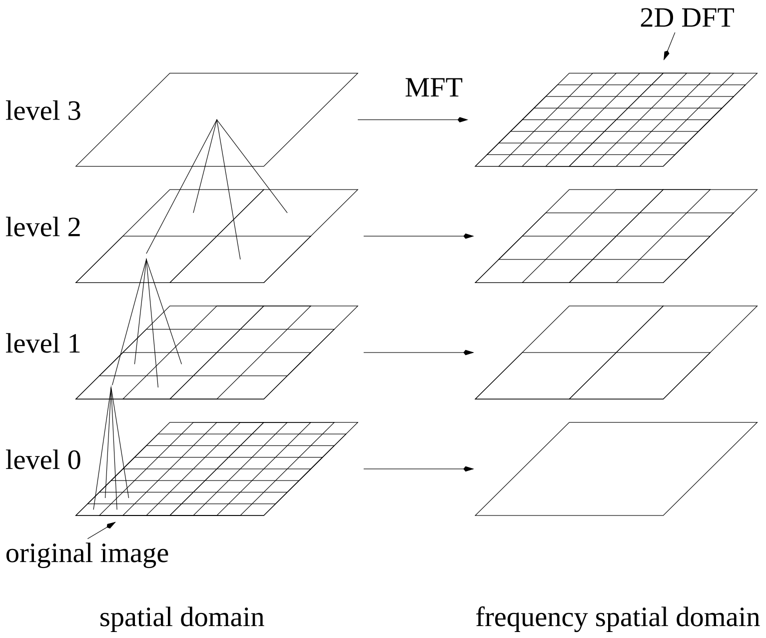


Figure 1: Structure of 2-D MFT in spatial/spatial frequency diagram

scales. Analogous to the continuous transform, the 2-D MFT component at position $\vec{\xi}(l)$ has frequency $\vec{\omega}(l)$ and scale $\sigma(l)$ where l is the transform level. A 2-D MFT coefficient can be represented by

$$\hat{x}(\vec{\xi}(l), \vec{\omega}(l), \sigma(l)) \; = \sum \; w_n(\vec{\xi'}(l) - \vec{\xi}(l)) \; x(\vec{\xi'}(l)) \; e^{[-j\vec{\xi'}(l)\vec{\omega}(l)]} \tag{1}$$

where $w_n(\vec{\xi'}(l) - \vec{\xi}(l))$ is a window function. Therefore each level of MFT resembles a 2-D Short Time Fourier Transform (STFT). $x(\vec{\xi'}(l))$ is the original N $\times$ N image sampled at the point $\vec{\xi'}$ at level l in the spatial domain. The MFT has following important properties for image analysis[1][7][8]:

(1) Linearity: MFT is a hierarchical set of STFT's, and the STFT is linear by definition. Hence, the MFT is linear.

(2) Locality: a level of the MFT is a combined spatial and spatial frequency representation of the image. By choosing analysis vectors that are optimally localised, each level is an optimal representation at its prescribed resolution in each domain.[1] [8].

(3) Invertibility: since the STFT can be defined to be invertible by judicious choice of window function, the MFT has a similar property [8].

(4) Resolution: the MFT contains a multiplicity of resolutions in both domains, from the original image to its DFT. The different resolutions consist of coefficients that are uniformly distributed across the whole domain. These sets of coefficients can represent an uncertain degree of locality of features in each domain. [1]

3 Markov Random Fields

Markov Random Fields (MRFs) have been used as the basis of an evidential approach to many computer vision and image processing tasks in recent years. Coupled Markov random fields can unify the segmentation and reconstruction process [4] and are presented as a mechanism for combining several sources of *a priori* and observational knowledge in a Bayesian framework. The MRFs encode the assignment of labels to image sites. Knowledge is encoded by the neighbourhood structure of the MRF and by the assignment of 'goodness' potentials to local structures (cliques) in the MRF. The potentials then determine the prior probability distribution of labels in the MRF, whose *a posteriori* probability distribution is derived by combining the pooled external observations and *a priori* distribution. The main advantage of the MRF model is that it provides a general and natural model for the interaction between spatially related random variables. There are flexible optimization algorithms, for example simulated annealing or a time evolution state [9], that can be used to find the globally optimal realization which corresponds to the maximum a posteriori (MAP) segmentation [4]. Some of the relevant aspects of MRF theory and its application to image labelling are briefly described in the following pages. Consider a set, X, of discrete-valued random variables. Associate with the random variables is a graph, G, defined as a finite set of vertices, T, and a set of edges of the graph, E. The set of all points which are neighbours of a point t will be denoted by N_t. An assignment of values to all the variables in the field is called a configuration, and is denoted ϖ. ϖ_t is the value given to the point t by the configuration ϖ. A probability measure P will be said to define a Markov Random Field if the local characteristics depend only on the knowledge of the outcomes at neighbouring points [5] [10], ie. if for every ϖ

$$P\,(\varpi) \;>\; 0, \qquad \forall \;\; \varpi \;\in\; X \tag{2}$$

$$P\,(\varpi_t\,|\,\varpi_{T-t}) \;=\; P(\varpi_t\,|\,\varpi_{N_t}) \tag{3}$$

where $P(\varpi)$ and $P(\varpi_t|\varpi_{T-t})$ are the joint and conditional pdf's, respectively. This states, roughly, that the state of a site is dependent only upon the state of its neighbours (N_t). MRFs can also be characterized in terms of an energy function, U, with a Gibbs distribution:

$$P\,(\varpi) \;=\; \frac{e^{-U(\varpi)/T}}{Z} \tag{4}$$

where T is the temperature, and Z is a normalizing constant. If we are interested only in the pdf, $P(\varpi)$, the Gibbs energy function U is defined as:

$$U(\varpi) \;=\; \sum_{c\,\in\,C} V_c(\varpi) \tag{5}$$

where C is the set of cliques defined by the neighbourhood graph $G(T, E)$, and V_c are the clique potentials. Notice that the MRF pdf in Equation (5) is quite general, in that the clique functions can be arbitrary as long as they depend only on the nodes in the corresponding cliques. Due to this unique structure, in which the global and local properties are related through cliques, the MRF model-based approach provides a useful mathematical framework for the study of image segmentation, as will be discussed in more detail later.

4 Image Segmentation

The image segmentation method presented in this work proceeds as follows:

- Representation by the Laplacian Pyramid - This is used as a high-pass filter to create the luminance edges.

- Use of Multiresolution Fourier Transform for Feature Extraction - The MFT is used to transform the luminance edges into a double-sized image consisting of local spectra.

- Extraction of Local Features - Extracting local features helps to determine the boundary structure of the image. This structure relates not only to the magnitude of the spectral components, but also to their relative phases.

- Estimation of Edge Position - These estimate the position of edges and also a certainty measure given by examining the magnitude of the correlation statistic, defined in section 4.4.

- Application of Markov Random Fields - In this section, a method for edge detection in image segmentation in the Bayesian framework using Markov random field models is developed. Stochastic Relaxation is also adopted to maximise the likelihood.

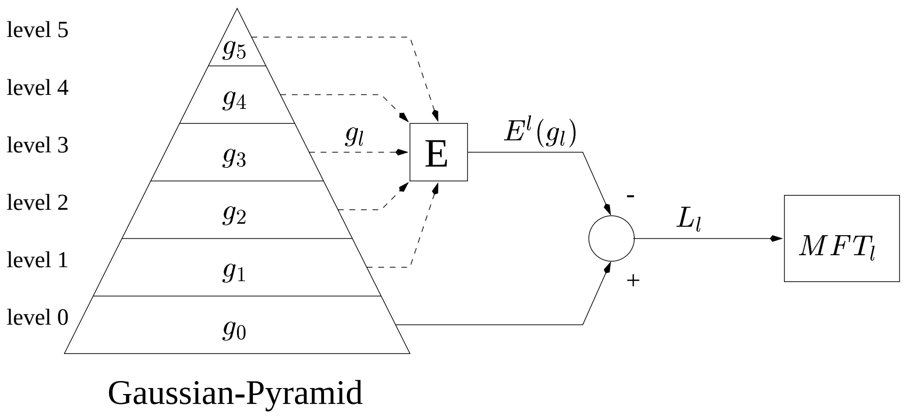


Figure 2: From Gaussian Pyramid to MFT Entry

4.1 Representation by the Laplacian Pyramid

The basic pyramid representation consists of a number of stacked 2-d arrays, each of which represents a different spatial resolution of image. For an image $v(m,n)$, $0 \leq m,n < M$, where $M = 2^N$, $0 < l < N$, the Gaussian Pyramid can be expressed as follow:

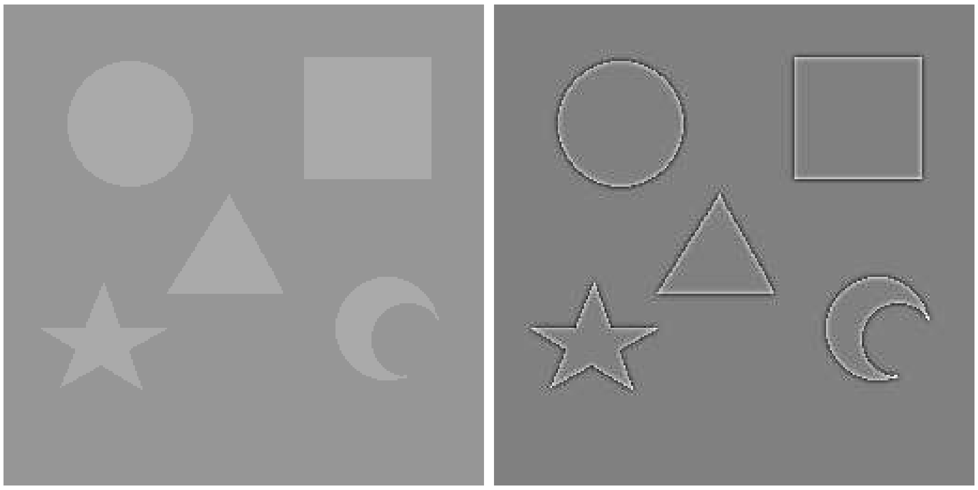
$$g_l(m,n) = \sum_{p=0}^{K-1} \sum_{q=0}^{K-1} w(p,q)\, g_{l-1}(2m+p, 2n+q) \tag{6}$$

The original image, g_0, form is,

$$g_0(m,n) \;=\; v(m,n) \tag{7}$$

where $g_l(m,n)$ are the coefficients or nodes on level l of the representation and $w(p,q)$ is the Gaussian-like weighting function/kernel which is of finite size $K \times K$. The kernel $w(p,q)$ therefore defines the transformation function between the different resolutions. Each node is given by [11]

$$g_l(m,n) \;=\; L_l(m,n) \;+\; E(g_{l+1})(m,n) \tag{8}$$



(a) Original shapes Image   (b) High-pass filtered version of (a)

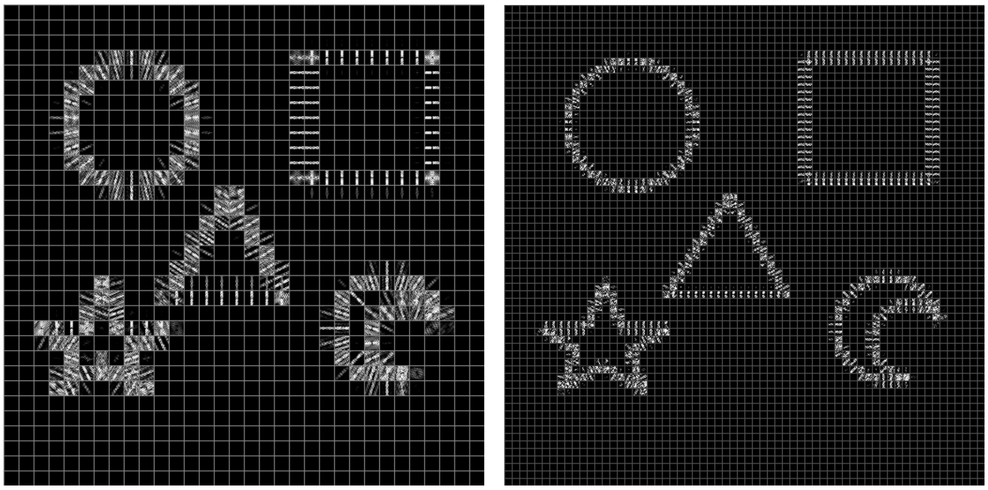
Figure 3: The shapes image and its high-pass filtered version

where $E(I)$ is the expanded version of the image, I, using the kernel, w(p,q), as an interpolator. Formally,

$$E(g_{l+1})(m,n) \;=\; K \sum_{p=0}^{K-1} \sum_{q=0}^{K-1} w(p,q)\, g_{l+1}(\,\lfloor \frac{m-p}{2} \rfloor\,,\,\lfloor \frac{n-q}{2} \rfloor\,) \tag{9}$$

It can be seen that each level of the Laplacian Pyramid is the difference between successive levels of the Gaussian Pyramid. It has been shown that each level of the Gaussian Pyramid, g_{l+1}, is a low-pass filtered version of the previous level, g_l [11]. Conversely, each level of the Laplacian Pyramid, L_l, is the high-pass filtered version of g_l. L_0 is therefore the high-pass filtered version of the original image [8]. Figure 3 shows the 'Shapes' image and its high-pass filtered version, $L_0(m,n)$. To form a suitably pre-whitened input to the MFT on level l, level l of the Gaussian Pyramid is expanded l times and substracted from the original image. Thus the input to the MFT can be written as

$$L_l(m,n) \;=\; g_0(m,n) \;-\; E^l(g_l)(m,n) \tag{10}$$



(a)MFT spectrum at level 3 (b)MFT spectrum level 2

Figure 4: MFT spectrum of shapes image

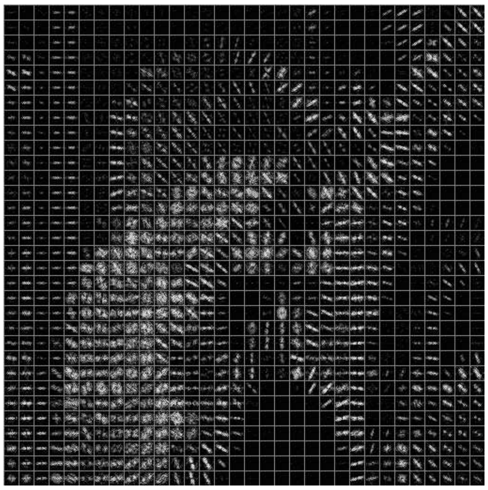
as shown in Figure 2.

4.2 Use of Multiresolution Fourier Transform for Feature Extraction

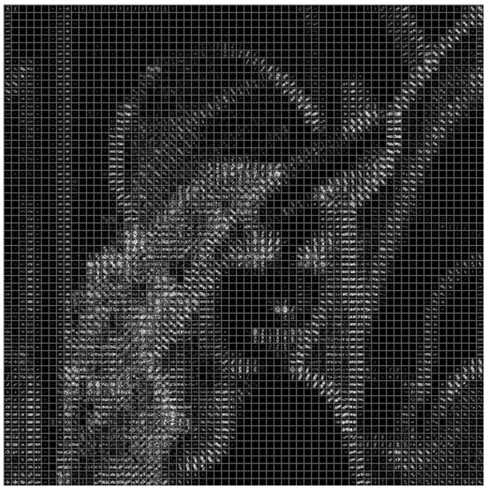
Once the high-pass filtered image of size $N \times N$ is obtained, the MFT is used to transform it into a $2N \times 2N$ spectral image in spatial frequency domain [8]. Figures 4 and 5 show the MFT spectra of the 'Shapes' and 'Lena' images. The local spectra of the MFT are obtained using the FFT. The MFT levels, however, provide both spectral and positional information at two different resolutions, level 3 having greater spatial frequency resolution and level 2 having greater spatial resolution, for example. At levels with high spatial resolution, the boundaries of the shapes are represented by energy concentrated in an orthogonal orientation within the local spectra of MFT. In contrast, on levels with high frequency resolution, a given shape is represented

(a) Original Lena Image

(b) High-pass filtered version of (a)



(a)MFT spectrum at level 3



(b)MFT spectrum at level 2

Figure 5: MFT spectrum of Lena image

within a single spectrum by symmetrically distributed energy.

4.3 Extraction of Local Features

An image can be considered to consist of regions, each with an associated boundary. The boundaries are contiguous to one another and of different size and orientation. These are typically known as lines and edges; a more general term is local image feature. The multiresolution approach to the problem can be based on this characteristic. A hierarchical structure represents an image by local features defined at different spatial resolutions. An ideal local feature in the continuous 2-d spatial domain can be represented by the following oriented region

$$\hat{v}\,(x, y) \;=\; \hat{v}\,(\,x\;cos\,\theta \;+\; y\;sin\,\theta\,) \tag{11}$$

and its Fourier transform is given by[1]

$$\hat{V}\,(\vec{\omega}) \;=\; \hat{U}\,(\,u\;\sin\,\theta \;-\; v\;\cos\,\theta\,)\;\;\hat{V}\,(\,u\;\cos\,\theta \;+\; v\;\sin\,\theta\,) \tag{12}$$

where $\hat{V}\,(u, v)$ is confined to a line which is perpendicular to the orientation of the feature in the spatial domain. Specifically, if $\vec{\eta}_{xy}$ is defined as the centroid of $\hat{v}(x, y)$, then the argument of $\hat{V}\,(\vec{\omega})$ is given by [12]

$$Arg\;[\;\hat{V}\,(\vec{\omega})\;] \;=\; \vec{\omega}\cdot\vec{\eta} \;+\; \varepsilon \tag{13}$$

$$\hat{V}(\vec{\omega}) \;=\; |\;\hat{V}(\vec{\omega})\;|\;e^{-j\;(\vec{\omega}\cdot\vec{\eta}\,+\,\varepsilon)} \tag{14}$$

where ε is a phase constant. In other words, the spatial offset of these features is directly proportional to the phase variation of the spectrum in an orthogonal orientation as illustrated in Figure 6 [1]. In the frequency domain, this correspondonds to convolving the spectrum with the transformed windows, the effect being to smooth out the energy concentration, in the

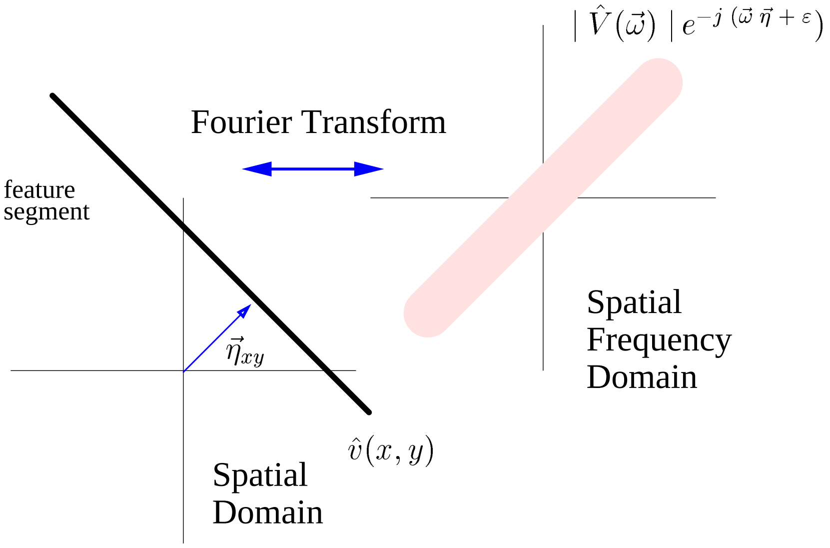
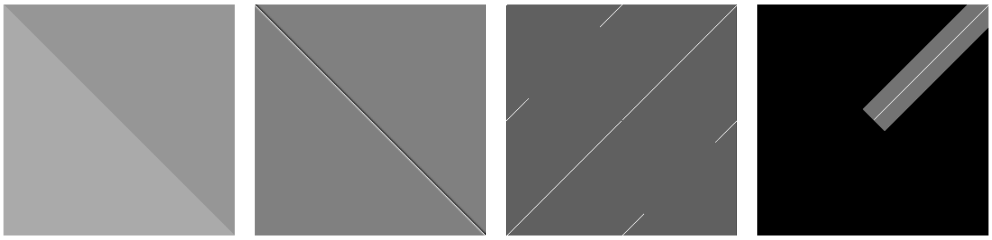


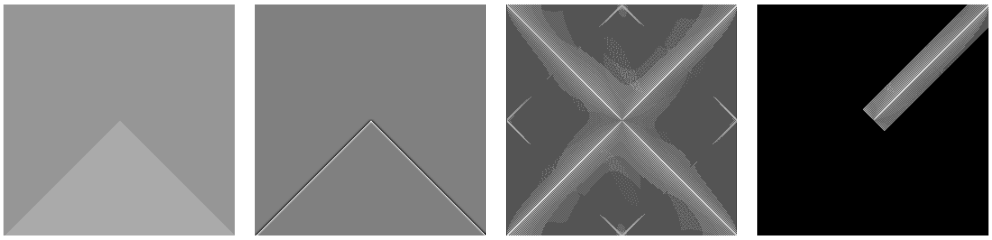
Figure 6: Linear phase property of local feature segment.

form of an oriented and elongated region. Thus, for an edge between two regions with homogeneous gray levels in an image block, the orientation can be estimated by analysing its Fourier spectrum. For example, Figure 7 shows an edge segment between two regions of homogeneous gray level. In order to enhance the edge, Figure 7(b) shows the high-pass filtered version of Figure 7(b). Figure 7(c) shows the Fourier transform of the high-pass image. As expected, the energy is concentrated along the direction orthogonal to that of the edge in spatial domain. However, for a grey level image with a sharp corner as shown in Figure 8 we can see that the Fourier spectrum in Figure 8(c) contains more than one feature inside the region and this will interfere with the analysis. Thus, instead of taking into account all the Fourier coefficients in the whole half plane of the spectrum, a number of orientations are tested and only those coefficients within a strip along the orientation estimated using the method described in section 4.4 are taken into account. This reduces the influence of texture fluctuations and separates the boundary



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

Figure 7: An edge between two adjacent grey level regions and its Fourier spectrum. (a) the original image, (b) the high-passed version of the original image, (c) the Fourier spectrum of the high-passed version, (d) a half-plane of the spectrum containing the Fourier coefficients along a specific orientation.



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

Figure 8: An boundary between two adjacent grey level regions and its Fourier spectrum. (a) the original image, (b) the high-passed version of the original image, (c) the Fourier spectrum of the high-passed version, (d) a half-plane of the spectrum containing the Fourier coefficients along a specific orientation.

from the texture features.

4.4 Estimation of Edge Position

The centroid of the segment contained in each block is calculated using a method similar to Calway [1]. As described in[1][8][12][13] and [14], for an edge or boundary segment between regions of homogeneous gray levels, the spatial information is contained in the phase of the Fourier transform. The spatial position of the centroid of such a linear feature can be estimated by averaging the phase difference over all frequencies. This method uses the *correlation statistic* [1]. Setting $\vec{\omega}$ equal to (u, v), equation (14) can be re-written as [13]

$$\hat{V}(u,v) = \mid \hat{V}(u,v) \mid e^{-j(ux+vy)} \tag{15}$$

where (x, y) is the position of the centroid of the linear feature and $\mid \hat{V}(u,v) \mid$ is the spectrum of a feature at $(0, 0)$. Because of Hermitian symmetry only the coefficients in the half-plane have to be considered.

The autocorrelation coefficients of the image spectrum in the u and v dimensions are

$$\rho_u = \frac{\sum_{u,v \in \Theta_\theta} \hat{V}(u,v)\, \hat{V}^*(u+u',v)}{\sum_{u,v \in \Theta_\theta} |\hat{V}(u,v)|^2} \tag{16}$$

$$\rho_v = \frac{\sum_{u,v \in \Theta_\theta} \hat{V}(u,v)\, \hat{V}^*(u,v+v')}{\sum_{u,v \in \Theta_\theta} |\hat{V}(u,v)|^2} \tag{17}$$

where u' and v' are sampling intervals in the u and v dimensions respectively. By substituting equation (15) into equations (16) and (17), the estimate of the centroid position (x, y) can be given as [8]

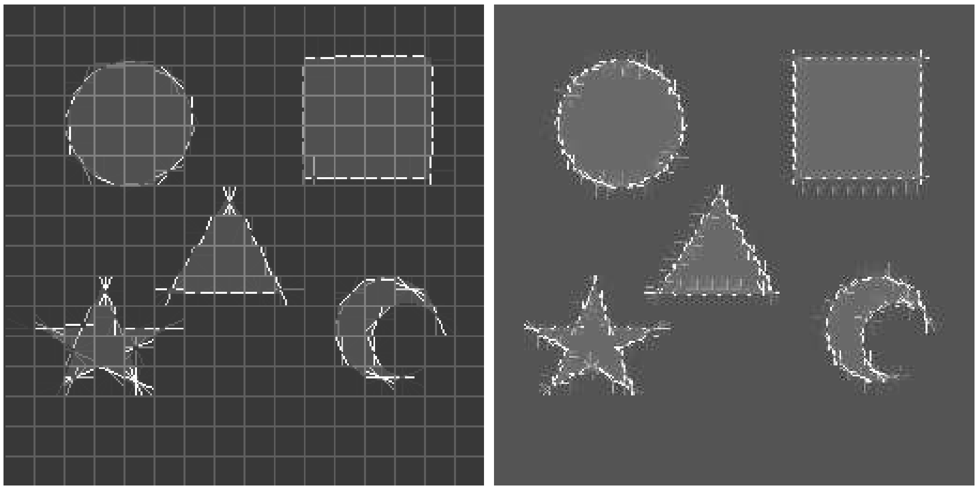
$$x_0 = \frac{N \cdot Arg(\rho_u)}{2\pi} \tag{18}$$

$$y_0 = \frac{N \cdot Arg(\rho_v)}{2\pi}, \tag{19}$$

and the sampling interval is $\frac{2\pi}{N}$. In the multiresolution image model, the parameters obtained refer to local features defined at different spatial resolutions and these are determined from the spatial frequency vectors on different levels of the MFT. For each local feature in a given orientation, the correlation stastistics used in the estimation scheme provide an estimate of the linear phase increment in an orthogonal orientation within the relevant spatial frequency vector. However, the correlation statistics provide a measure of the energy variation over all orientations, given the linear phase model. It is therefore necessary to find other ways to model the magnitude distribution. However, both of Calway's and Li's methods are based on the assumption that each block only contains a single feature. If not, it will be divided into 4 sub-blocks, and each of these will be re-estimated at the higher spatial resolution until a single feature is found or the block is too small to analyse [8]. Here the spatial frequency block is divided into orientation segments instead of dividing the spatial block into sub-blocks. The estimated features are displayed by constructing an image in which each feature is represented by a straight line within the spatial region referred to by the spatial frequency vector. This line is shown at the appropriate orientation and position with a certain length. The luminance value of each line is then set to the magnitude of the MFT coefficients. Figure 9 shows the result of this approach. In order to ensure that interesting features with relatively low energies are not missed, a normalisation process is performed to increase their visibility.

4.5 Application of Markov Random Fields

Based on the result of the feature extraction of MFT, Bayesian probability theory is used to label the image sites. Pixels are labelled corresponding to the number of segments or regions and the label of site is compatible,



(a) image reconstructed at level 4 (b) image reconstructed at level 3

Figure 9: The shapes image reconstructed by MFT coeffecients

in some sense, with the labels of neighbouring sites. Stochastic Relaxation is used to maximise the likelihood and, as in [5], this approach is based on incorporating both regional and boundary information into the interaction energy function of the MMRF and performing the segmentation in a unified process. Li defines [8] the interaction energy between a site and its neighbourhood defining the MRF is a function of the four immediate neighbouring sites and is defined as

$$U(\varpi_t, \varpi_{\mathcal{N}_t}, X_{\mathcal{N}_t}) \; = \; \sum_{t' \in \mathcal{N}_t} V(\varpi_t, \varpi_{t'}, X_{tt'}) \tag{20}$$

where $\varpi_t, 1 \leq \varpi_t \leq L$, is the class label at site t, $X_{\mathcal{N}_t}$ represents the measurements over the sites within the neighbourhood $\mathcal{N}_t$ and the measurement $X_{tt'}$ is the squared grey level difference estimated from these two sites. The pairwise interaction potential V is a suitably defined function of class labels of the two neighbouring sites t and t'. Note that, at present, only the nominal top level in used, with no propagation. Also, there is no interaction between site t and its father in the calculation of V. Define

$$X_{tt'} = (x_t - x'_t)^2 \tag{21}$$

$$\sigma^2_{wx} = \sum_{\varpi_t = \varpi_{t'}} X_{tt'} \quad , \forall t \text{ and } t' \tag{22}$$

$$\sigma^2_{ox} = \sum_{\varpi_t \neq \varpi_{t'}} X_{tt'} \quad , \forall t \text{ and } t' \tag{23}$$

where x_t and x'_t are the grey level at sites t and t', σ^2_{wx} is the average value of the intra-region grey level difference and σ^2_{ox} is the average value of the inter-region grey level difference. These are both re-calculated at the beginning of each iteration because of their dependence on the varying configuration. The probability distribution functions of the two differences roughly match Gaussian distributions with different means and variances. We then define the region and boundary energy function, based on Bayes's theorem as follows: [8]

- The Region Energy Function:

$$V_r(\varpi_t, \varpi_{t'}, X_{tt'}) = \begin{cases} \ln\left(\frac{P(\varpi_t \neq \varpi_{t'})}{\sigma_{ox}}\right) - \frac{X_{tt'}}{2\sigma^2_{ox}} & if\ \varpi_t = \varpi_{t'}\ and\ \varpi_{t'} \neq Bl \\ \ln\left(\frac{P(\varpi_t = \varpi_{t'})}{\sigma_{wx}}\right) - \frac{X_{tt'}}{2\sigma^2_{wx}} & if\ \varpi_t \neq \varpi_{t'}\ and\ \varpi_{t'} \neq Bl \\ 0 & if\ \varpi_{t'} = Bl \end{cases} \tag{24}$$

  where Bl is the boundary label. The physical interpretation of equation (24) is that if sites t and t' belong to the same class, the probability will be higher and assigning t and t' the same label will be encouraged by a lower interaction energy. By contrast, if sites t and t' do *not* belong to the same class, the probability will be lower and it is likely that a larger interaction energy will be produced to discourage the assignment of the same label to sites t and t'. Furthermore, if sites t and t' are assigned different class labels when they *do* belong to the same class, the algorithm will tend to impose a larger interaction energy to penalise the assignment of different labels to sites t and t'. Conversely, if they do not belong to the same class, the algorithm will tend to endorse the assignment of different labels to the sites with a lower interaction energy.

- The Boundary Energy Function:

$$V_b(\varpi_t, \varpi_{t'}, D) = \begin{cases} \ln\left(\frac{P(\varpi_t \neq \varpi_{t'})}{\sigma_{od}}\right) + \frac{\overbrace{(E_t + E_{t'})}^{energy}}{\sigma_{oe}} - \frac{D^2}{2\sigma_{od}^2} & if\ \varpi_t = \varpi_{t'} \\ & and\ \varpi_{t'} = Bl \\ \ln\left(\frac{P(\varpi_t = \varpi_{t'})}{\sigma_{wd}}\right) + \frac{\overbrace{(E_t + E_{t'})}^{energy}}{\sigma_{we}} - \frac{D^2}{2\sigma_{wd}^2} & if\ \varpi_t \neq \varpi_{t'} \\ & and\ \varpi_{t'} = Bl \\ 0 & if\ \varpi_{t'} \neq Bl \end{cases} \tag{25}$$

where $\sigma_{..}$ are appropriate normalisation factors. The potential of the boundary process between sites t and t' is defined, in a fashion similar to equation (24), by combining an energy from MFT coefficients and replacing the grey level difference $X_{tt'}$ by a 'distance' measure, D, between the estimated boundary segments, shown in Figure 10. The distance measure is given by [8]

$$D = \| \vec{l} \| (sin(\theta_1) + sin(\theta_2)), \quad 0 \leq \theta_1, \theta_1 < \pi/2 \tag{26}$$

where $\vec{l}$ is the vector, estimated from MFT coefficients, joining the centroids of the boundary segments within the two blocks, θ_1 is the

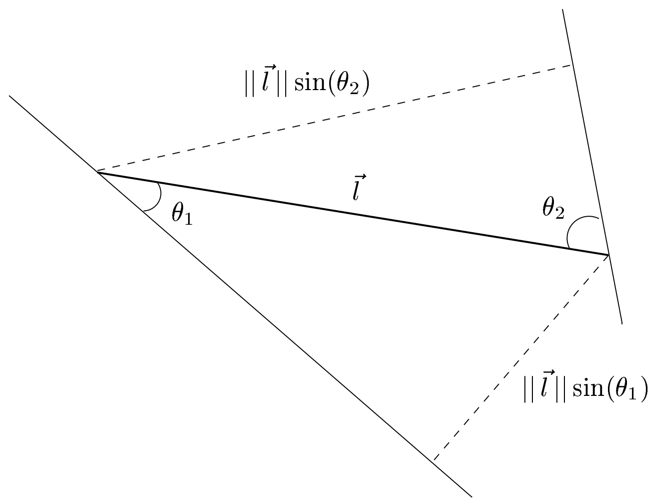
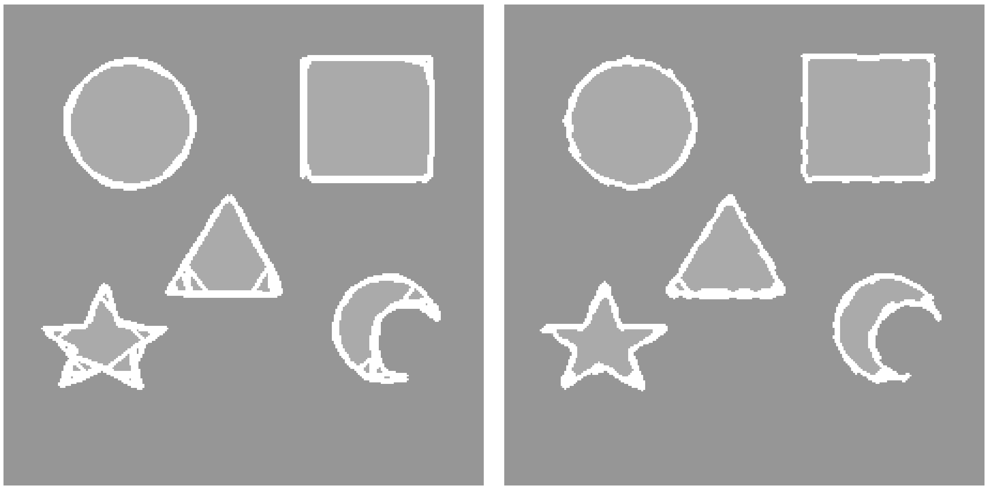


Figure 10: 'Distance' between boundary segments

angle between $\vec{l}$ and one of the boundary segments and θ_2 is the angle between $\vec{l}$ and the other segment. From Figure 10, it is apparent that the better aligned a boundary segment pair, the smaller the θ_1 and θ_2, therefore the shorter the 'distance' between them. The 'energy' term in equation (25) is a negative value added to the energy function encourage the site to be labelled as a boundary. If there is no feature, that is zero MFT coefficients, the block energy will be zero. After the site has been labelled as a boundary, the 'energy' term and 'distance' measure are used to iterate and find the orientation of the site.

5 Experiments

A number of trials have been conducted to test the ability of this approach to refine the boundary connections. To get the appropriate centroid feature from the MFT coefficients, the image scale is limited at Gaussian pyramid level 4 and 5, giving an image size of (32×32) and (64×64). The block



(a) Shape Image at level 3 (b) Shape Image at level 2

(a) Lena image at level 3 (b) Lena image at level 2

Figure 11: The Segmentation Results

sizes are (8×8) and (32×32). The segmentation results for different images are shown in Figure 11. From the these results, it can be seen that while edges are represented reasonably well, there is some over-segmentation at the sharp corners or small regions. This can be improved by adjusting the energy function and the constraints of boundary connection.

6 Conclusions And Future Work

This work is based broadly on that of Li [8], however it is refocused on non-texture image processing. The main contribution of this work is a novel method of improving boundary connections after segmentation. The results, shown in Figure 11, show that edges are generally represented at an appropriate scale. This approach is still in the initial stage and further work could involve multiple levels and refine the boundary connection strategy.

References

[1] A. Calway. *The Multiresolution Fourier Transform: A Genera Purpose Tool for Image Analysis.* PhD thesis, Department of Computer Science, The University of Warwick, UK, September 1989.

[2] R. Wilson, A. Calway, and E.R.S. Pearson. A Generalised Wavelet Transform for Fourier Analysis: The Multiresolution Fourier Transform and Its Application to Image and Audio Signal Analysis. *IEEE Transactions on Information Theory*, 38(2), March 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] Dr. Rama Chellappa and Anil Jain Edited. *Markov Random Fields Theory and Application.* Academic Press, 1993.

[5] Chang-Tsun Li and Roland G. Wilson. Image Segmentation Based on a Multiresolution Bayesian Framework. In *Submitted to IEEE International Conference on Image Processing, Chicago, USA*, 1998.

[6] C. T. Li and R. Wilson. Image Segmentation Using Multiresolution Fourier Transform. Technical report, Department of Computer Science, University of Warwick, 1995.

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

[8] Chang-Tsun Li. *Unsupervise Texture Segmentation Using Multiresolution Markov Random Fields.* PhD thesis, Department of Computer Science, The University of Warwick, UK, 1998.

[9] Yu. A. Rozanov and Translated by Constance M. Elson. *Markov Random Fields.* Springer-Verlag, 1993.

[10] R. Kindermann and J. L. Snell. *Markov Random Fields and Their Applications* . American Math. Society, 1980.

[11] P. J. Burt and E. H. Adelson. The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communication*, COM-31:532–540, 1983.

[12] Athanasios Papoulis. *Signal Analysis.* McGraw Hill, 1981.

[13] Andrew R. Davies. *Image Feature Analysis using the Multiresolution Fourier Transform.* PhD thesis, Department of Computer Science, The University of Warwick, UK, 1993.

[14] Michael Spann. *Texture Description and Segmentation in Image Processing.* PhD thesis, The University of Aston in Birmingham, UK, 1985.