Research Report 202

Curve and Corner Extraction using the Multiresolution Fourier Transform

Andrew Davies, Roland Wilson

RR202

A novel method of segmenting images into regions which contain linear or circular arc features is presented. The feature models are based upon spectral properties, local estimates of which are provided, over a range of scales, by the Multiresolution Fourier Transform (MFT). The algorithm ensures that detected features both accurately model the data and are consistent across scale. A method of combining the primitive line and arc segments into more complex features is also considered. Some results from application of the algorithm to a number of natural and synthetic images are presented to illustrate its effectiveness in practical applications.

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

November 1991

Curve and Corner Extraction using the Multiresolution Fourier Transform

Andrew Davies, Roland Wilson
Department of Computer Science,
University of Warwick,
Coventry

November 28, 1991

Abstract

A novel method of segmenting images into regions which contain linear or circular arc features is presented. The feature models are based upon spectral properties, local estimates of which are provided, over a range of scales, by the Multiresolution Fourier Transform (MFT). The algorithm ensures that detected features both accurately model the data and are consistent across scale. A method of combining the primitive line and arc segments into more complex features is also considered. Some results from application of the algorithm to a number of natural and synthetic images are presented to illustrate its effectiveness in practical applications.

1 Introduction

The problem of identifying boundary contours or line structures is widely recognised as an important component in many applications of image analysis and computer vision. Typical solutions to the problem employ some form of edge detection and line following or, more commonly in recent years, Hough transforms (Ballard and Brown [1], Illingworth and Kittler [2]). Because of the processing requirements of such methods and to try to improve the robustness of the algorithms, a number of authors have explored the use of multiresolution approaches to the problem (Princen et al [3]). Non-parametric, iterative approaches such as relaxation labelling (Zucker et al [4]) and "Snakes" (Kass et al [5]) have also been used.

In an attempt to overcome some of the limitations of these, a multiresolution feature model based on a type of wavelet transform, the Multiresolution Fourier Transform (MFT), was described in [6],[7],[8],[9]. This 'local spectral' model of line

and edge features uses a Markov process defined in the Fourier domain and gives both estimates of the feature parameters and a measure of how well the feature models the data. Efficient maximum likelihood estimation and decision procedures give a fast segmentation of the image into a set of disjoint 'single feature' regions, of varying scales, each region containing a single straight edge or line feature. Use of the MFT allows any arbitrarily compact region to be represented independently of what is present outside of it [6], and thus an appropriate scale can be chosen for any feature in the image.

The work described here is an extension of these ideas, in which a given region can contain multiple straight line or edge features or a circular curve segment. The general framework of frequency domain modelling is still employed and the estimation and decision procedures are broadly comparable, both theoretically and computationally, to the previous work. The scope of the model and consequently its utility have been increased, without greatly increasing the computational requirements. The processing takes place entirely within the framework of the MFT and does not require separate edge detection processes, while the resolution of feature parameters is primarily limited by noise and does not affect the computational complexity, as happens with the Hough Transform. Because the modelling is performed using an invertible transform it is possible to measure the model error and, if required, produce an approximation of the image based upon the estimated parameters [6]. A method of combining these simple features into more complex curves is also described. The new method is shown to be effective in segmenting a variety of synthetic and natural images.

2 Feature Detection Using The MFT

2.1 The Multiresolution Fourier Transform

The feature models are based upon spectral properties, local estimates of which are provided by the MFT [6], in which the size of the 'locality' is dependent upon a scale parameter. The continuous MFT for an image $x(\vec{\xi})$ is defined by

$$\hat{x}(\vec{\xi},\vec{\omega},\sigma) = \sigma \int_{-\infty}^{\infty} w(\sigma(\vec{\chi}-\vec{\xi})x(\vec{\chi}) \exp\left[-\jmath\vec{\chi}.\vec{\omega}\right] d\vec{\chi} \tag{1}$$

where $\jmath = \sqrt{-1}$, $\vec{\xi} = (\xi_1\ \xi_2)^T$ is a spatial position vector, $\vec{\omega}$ is the frequency vector,σ is a scale parameter and $w(\vec{\xi})$ a window function.

The discrete MFT, which is used in the feature extraction process, is similarly given by

$$\begin{aligned} \hat{x}(i,j,n) &\triangleq \hat{x}(\vec{\xi}_i(n),\vec{\omega}_j(n),\sigma(n)) \\ &= \sum_k w_n(\sigma(\vec{\xi}_k - \vec{\xi}_i(n)))x(\vec{\xi}_k) \exp\left[-\jmath\vec{\xi}_k.\vec{\omega}_j(n)\right] \end{aligned} \tag{2}$$

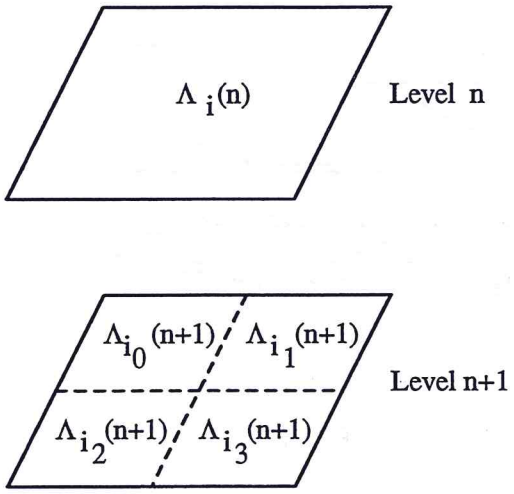

Figure 1: Parent - Child Relationship for Regions

where '.' denotes scalar product, n is the MFT level and specifies the discrete scale parameter,$\sigma(n)$, $\vec{\xi}_i(n)$ and $\vec{\omega}_j(n)$ are the sampling points in the spatial and frequency domains respectively. Each level of the MFT may be considered as a form of windowed Fourier transform [10] or Gabor representation [11]. As such, the MFT may be considered as a 'stack' of windowed Fourier Transforms, each level representing the image at a different spatial resolution.

A region $\Lambda_i(n)$ is associated with each sample point $\vec{\xi}_i = (\xi_{i1}\ \xi_{i2})^T$, such that

$$\vec{\xi} \in \Lambda_i(n) \iff |\xi_l - \xi_{il}| < \Xi(n)/2, l = 1,2 \tag{3}$$

where $\Xi(n)$ is the spatial sampling interval at level n. For a region $\Lambda_i(n)$ the set of coefficients $\{\hat{x}(i,j,n), 0 \leq j < 4\Xi^2(n)\}$, is a discrete estimate of its local spectrum. The spatial blocks $\Lambda_i(n)$ conform to a quad-tree structure: to each block $\Lambda_i(n)$ there correspond via (3) 4 blocks $\Lambda_{i_k}(n+1)$, $0 \leq k \leq 3$, which are disjoint and whose union is equal to $\Lambda_i(n)$ (see Fig. 1),

$$\Lambda_i(n) = \bigcup_{k=0}^{3} \Lambda_{i_k}(n+1) \tag{4}$$

Figure 2 shows the levels of the MFT for a small image. In the spatial domain, each region at level n is split into 4 at level $n+1$ (its children); thus spatial resolution increases with n. The block of local Fourier coefficients for each of the children, however, is a quarter of the size of that for the parent, i.e. the frequency resolution decreases with n.

In the feature detection process, each block $\Lambda_i(n_u)$, for some maximum scale index n_u, is first checked against the feature model. This gives an estimate of the feature

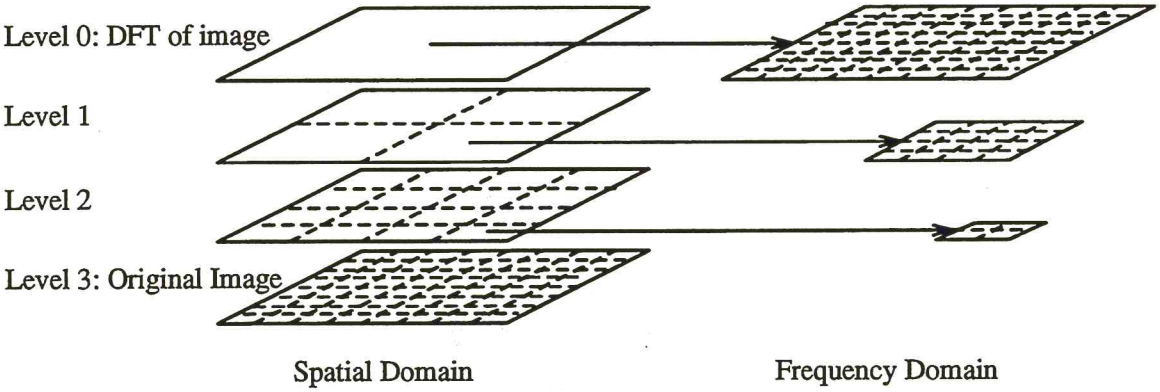

Figure 2: MFT Structure

parameters and a measure of how well the feature models the actual data, which can be used to decide whether such a feature is present in the region. At the next level, the relevant spectral properties for each block are estimated using the features detected in its children, in such a way that if these correlate well with the data in the block, consistency between the blocks features at the two levels may be assumed. Each block that passes the consistency test is then checked against the model and parameters estimated. This is repeated for all levels down to some lowest index n_l. Once level n_l has been processed, an attempt is made to link the detected primitive features into more complex features, using a continuity criterion based upon local correlation. Figure 3 gives an overview of the algorithm.

2.2 Multiple Linear Features

The set of coefficients $\{\hat{x}(i,j,n), 0 \leq j < 4\Xi^2(n)\}$ provides an estimate of the spectrum in a region $\Lambda_i(n)$ centered on the sampling point $\vec{\xi}_i(n)$. In the case where the region contains a single linear feature such as an edge, the spectrum will be oriented orthogonally to the feature orientation, and the phase will have a linear characteristic, i.e. its partial derivatives, with respect to the two Fourier co-ordinates, will be constant [8],[6]. The orientation is calculated by using the energy spectrum to construct a moment of inertia tensor, as described in [12]. The gradient of the transform phase is the vector $\vec{\phi}(\vec{\xi},\vec{\omega},\sigma)$ defined by

$$\vec{\phi}(\vec{\xi},\vec{\omega},\sigma) = \partial \arg\left(\hat{x}(\vec{\xi},\vec{\omega},\sigma)\right)/\partial\vec{\omega} \tag{5}$$

In the discrete case the corresponding vector is written as $\vec{\phi}(i,j,n)$ and is estimated using a fifth order FIR approximation to the partial derivative (Appendix A). From the gradient, the feature centroid is estimated as the magnitude weighted average

$$\eta_{ik} = \arg\left(\sum_j |\hat{x}(i,j,n)| \exp\left[\jmath\phi_k(i,j,n)\right]\right), k = 1,2 \tag{6}$$

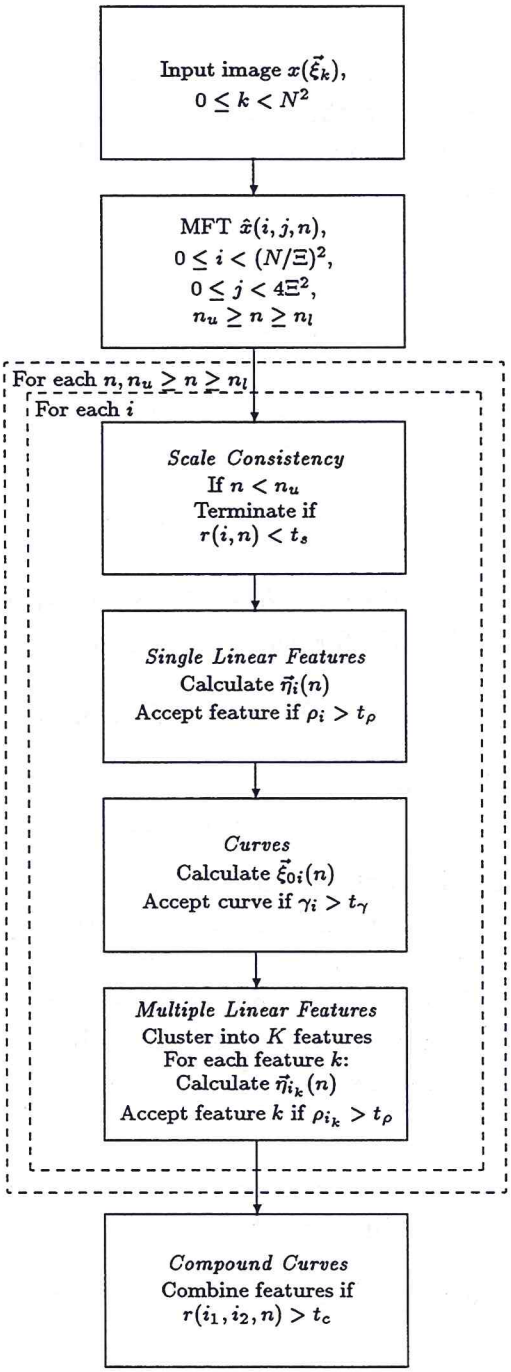

Figure 3: Structure of Algorithm

The ratio ρ_i defined by

$$\rho_i = \sum_k |\sum_j |\hat{x}(i,j,n)| \exp[\jmath\phi_k(i,j,n)]|/2 \sum_j |\hat{x}(i,j,n)| \tag{7}$$

is then used to provide a measure of the consistency of the model with the data in $\Lambda_i(n)$, and checked against a threshold t_ρ.

If more than one linear feature is present, then it is necessary to segment the spectrum into components due to each feature. If there are K features in $\Lambda_i(n)$, then the general MFT coefficient may be written as

$$\hat{x}(i,j,n) = \sum_{k=1}^{K} \hat{x}_k(i,j,n) \tag{8}$$

where $\hat{x}_k(i,j,n)$ is the contribution due to the k^{th} feature. If these features have well separated orientations, then to a first approximation, for each j,

$$|\hat{x}_k(i,j,n)\hat{x}_l^*(i,j,n)| = \delta_{kl}|\hat{x}_k(i,j,n)|^2 \tag{9}$$

In other words, the contributions from different features can be estimated directly from the sum in (8). Clearly for each j, for some k

$$\hat{x}(i,j,n) = \hat{x}_k(i,j,n) \tag{10}$$

and hence, in the ideal case

$$\vec{\phi}(i,j,n) = \vec{\eta}_{i_k} \tag{11}$$

where $\vec{\eta}_{i_k}$ is the centroid of the k^{th} feature.

Assuming that this condition is met, a clustering approach based upon the phase derivatives can be used to segment the coefficients into K sets - one for each of the features. Of course, if any two or more of the features are in the same orientation, the process will fail: a smaller scale will have to be used.

A version of the K-means algorithm, as described by Therrien [13], is used to perform the clustering task:

1. Choose K arbitrary feature centroids within the spatial region - where K is the number of features being sought;

2. Using the the phase gradient $\vec{\phi}(i,j,n)$, assign coefficient (i,j,n) to the cluster to which it is closest;

3. Calculate the centroid for each cluster;

4. If any coefficient has been reclassified since the last time this step was executed, return to step 2; if not continue;

5. Check each cluster to see whether it fits the 'single feature' criteria previously described. If it represents a single feature, calculate feature parameters, if not then replace by two with centroids displaced from the original and return to step 2.

In practice this algorithm has been found to converge within 4 or 5 iterations.

2.3 Curve Extraction

A curve may be considered as a set of straight line segments, varying in orientation and position from one to the next, with each small segment represented by MFT coefficients in the appropriate orientation. The phase derivatives will give the position of the centroids of these line segments. Thus each coefficient gives an estimate of a point on the curve and the orientation of the tangent at that point. The magnitude $|\hat{x}(i,j,n)|$ gives a measure of how much energy is in the given orientation.

If there is a single circular arc feature in the region, then all the coefficients with significant energy should represent points that are some distance R, the radius of the circle, from some specific point $\vec{\xi}_0$, the centre. In addition, the tangent vector at that point should be orthogonal to the radial vector to that point. Hence, in a block i at level n ,for all j

$$(\phi_1(i,j,n) - \xi_{01})^2 + (\phi_2(i,j,n) - \xi_{02})^2 = R^2 \tag{12}$$

and

$$\cos(\psi_j)(\phi_1(i,j,n) - \xi_{01}) + \sin(\psi_j)(\phi_2(i,j,n) - \xi_{02}) = 0 \tag{13}$$

where ψ_j is the tangent orientation at the point (ξ_{j1}, ξ_{j2}), and is given by

$$\psi_j = \arctan(\omega_{j2}/\omega_{j1}) + \pi/2 \tag{14}$$

Defining the vector $\vec{\Psi}_j^T = (\cos(\psi_j)\ \sin(\psi_j))$ these conditions can be expressed as

$$\| \vec{\phi}(i,j,n) - \vec{\xi}_0 \| = R \tag{15}$$

$$\vec{\Psi}_j.(\vec{\phi}(i,j,n) - \vec{\xi}_0) = 0 \tag{16}$$

It should be noted that the size of the arc within a block must be less than π radians in order to ensure that no two distinct line segments are in the same orientation. This is necessary since coefficient $\hat{x}(i,j,n)$ will be determined by all line segments in the region with the orientation ψ_j.

Estimation of $\vec{\xi}_0$ is based on the error measure

$$E(i,j,n) = |\vec{\Psi}_j.(\vec{\phi}(i,j,n) - \vec{\xi}_0)|^2 \tag{17}$$

which will be minimum (in fact 0) when the two vectors are orthogonal. The total error can then be defined as the weighted sum

$$E(i,n) = \sum_j |\hat{x}(i,j,n)| E(i,j,n) \tag{18}$$

To minimise with respect to both ξ_{01} and ξ_{02},

$$\frac{\partial E(i,n)}{\partial \xi_{01}} = \frac{\partial E(i,n)}{\partial \xi_{02}} = 0 \tag{19}$$

which gives

$$\begin{aligned} &\xi_{01} \sum_j |\hat{x}(i,j,n)| \cos^2(\phi_j) + \xi_{02} \sum_j |\hat{x}(i,j,n)| \cos(\phi_j) \sin(\phi_j) = \\ &\sum_j |\hat{x}(i,j,n)| (\phi_1(i,j,n) \cos^2(\phi_j) + \phi_2(i,j,n) \cos(\phi_j) \sin(\phi_j)) \end{aligned} \tag{20}$$

$$\begin{aligned} &\xi_{01} \sum_j |\hat{x}(i,j,n)| \sin(\psi_j) \cos(\psi_j) + \xi_{02} \sum_j |\hat{x}(i,j,n)| \sin^2(\psi_j) = \\ &\sum_j |\hat{x}(i,j,n)| (\phi_1(i,j,n) \sin(\psi_j) \cos(\psi_j) + \phi_2(i,j,n) \sin^2(\psi_j)) \end{aligned} \tag{21}$$

This can be expressed in matrix form by defining the tensors

$$\begin{aligned} T(i,j,n) &= |\hat{x}(i,j,n)| \begin{bmatrix} \cos(\psi_j) \\ \sin(\psi_j) \end{bmatrix} [\cos(\psi_j) \ \sin(\psi_j)] \\ &= |\hat{x}(i,j,n)| \Psi_j \Psi_j^T \qquad (22) \\ T(i,n) &= \sum_j T(i,j,n) \qquad (23) \end{aligned}$$

to give

$$T(i,n) \begin{bmatrix} \xi_{01} \\ \xi_{02} \end{bmatrix} = \sum_j T(i,j,n) \begin{bmatrix} \phi_1(i,j,n) \\ \phi_2(i,j,n) \end{bmatrix} \tag{24}$$

or

$$\vec{\xi_0} = T(i,n)^{-1} \sum_j T(i,j,n) \vec{\phi}(i,j,n) \tag{25}$$

Having estimated the position of the centre, the radius of the arc can be estimated from

$$R = \frac{\sum_j |\hat{x}(i,j,n)| \parallel \vec{\phi}(i,j,n) - \vec{\xi_0} \parallel}{\sum_j |\hat{x}(i,j,n)|} \tag{26}$$

The correlation function γ_i, defined by

$$\gamma_i = \frac{|\sum_j \hat{x}(i,j,n) \tilde{x}^*(i,j,n)|}{\sqrt{\sum_j |\hat{x}(i,j,n)|^2 \sum_j |\tilde{x}^*(i,j,n)|^2}} \tag{27}$$

where $\tilde{x}(i,j,n)$ are coefficients derived from the feature model, provides a measure of consistency between the model and the data, and is compared with a threshold t_γ. Detection of curve segments is performed before testing for multiple linear features, but after testing for single features, to minimise processing and to find the simplest hypothesis compatible with the data (Fig. 3).

3 Scale Consistency

Because features should be consistent across scale, features in region $\Lambda_i(n)$ should be consistent with those in the children $\Lambda_{i_l}(n+1), 0 \leq l \leq 3$. If $\Lambda_i(n)$ contains K features, the centroid for each feature, $\vec{\eta}_{ik}(n)$, will be the weighted average of the centroids of its components in the children, $\vec{\eta}_{ik_l}(n+1)$, $0 \leq l \leq 3$. If only one feature contributes energy at each frequency in the parent, the phase derivatives $\vec{\phi}(i,j,n)$ can be estimated using the centroids of features detected in the children. If derivatives estimated using this approach correlate highly with those estimated for $\Lambda_i(n)$ directly, it would suggest that each parent frequency component is dependent upon only one feature - a necessary condition for the block to fit the feature models described in sections 2.2 and 2.3.

The coefficients $\vec{\phi}(i,j,n)$ are estimated from the features found in the children $\Lambda_{i_l}(n+1)$, for $m = 1, 2$,

$$\tilde{\phi}_m(i,j,n) = \arg\left(\sum_{l=0}^{3}\sum_{k=0}^{K_l-1} |\hat{x}_k(i_l, j_C, n+1)| \exp\left[\jmath(\eta_{lkm} - \Delta_{i_l m})\right]\right) \tag{28}$$

where $\vec{\Delta}_{i_k} = \vec{\xi}_i - \vec{\xi}_{i_k}$ is the displacement of the child region centre from the centre of its parent, K_l is the number of features in the region $\Lambda_{i_l}(n+1)$, $\vec{\eta}_{lk}$ is the centroid of the k^{th} feature in $\Lambda_{i_l}(n+1)$ and $\hat{x}_k(i_l, j_C, n+1)$ are the coefficients due the k^{th} feature. These derivative estimates are then compared with those obtained directly at level n, using a correlation function which takes into account the energy at level n

$$r(i,n) = \frac{\sum_j \sum_{m=1}^{2} |\hat{x}(i,j,n)|^2 \exp\left[\jmath(\phi_m(i,j,n) - \tilde{\phi}_m(i,j,n))\right]}{\sum_j |\hat{x}(i,j,n)|^2} \tag{29}$$

The correlation will be high only if there are features in some of the children that form one feature in the parent and there is consistency of the features detected in both the father and the children. This implies that the parent may be considered as one of the feature regions previously described. If the correlation is high (above some threshold t_s), then the parent may be used to extract feature parameters with higher resolution than the children. If the correlation is low, then it is assumed that the contents of the parent region cannot be modelled in terms of primitive features using the MFT coefficients of the parent level and the detection process terminates at the child level. This bottom-up (Fig. 4) approach continues until the image is segmented into the largest possible regions from which the primitive features may be extracted.

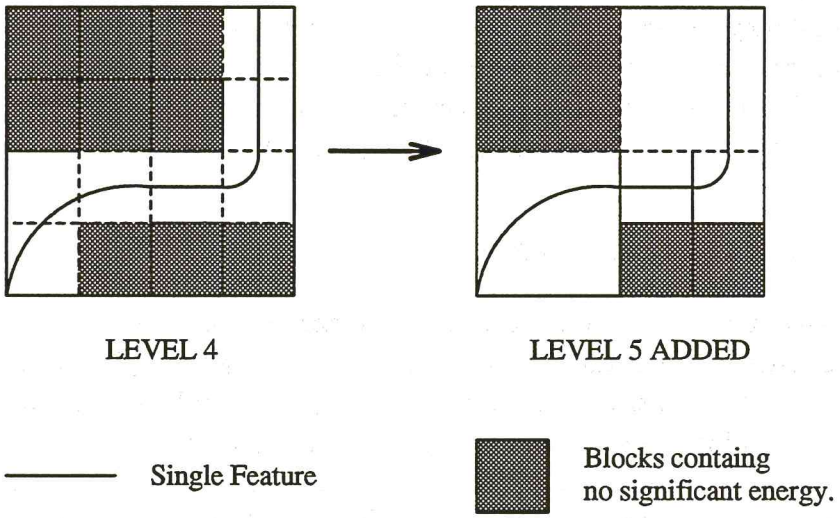

Figure 4: Combining regions

4 Combining Primitive Features

More complex shapes are modelled as concatenations of the above primitive features. A compound feature that exists in a particular region at a high level may be defined recursively in terms of the features that exist in its descendants. Thus, having segmented the image into primitive features, the next stage is to combine these into more complex features where possible [6]. If two primitives in adjacent blocks are part of the same compound feature, then it may be assumed that they

1. Meet at the boundary between the blocks, and
2. They have some continuity of curvature at this point.

To combine two primitives into a compound feature, consider how these primitives act at the ends where they meet. Since any feature may be considered as a set of linear feature segments at some scale, there must be a level at which the ends of both features are representable by such segments. Figure 5 shows this for two adjacent features. Moreover, the orientations must be similar if they are part of the same feature. In other words the magnitude of the correlation $r(i_1, i_2, n)$, defined by

$$r(i_1, i_2, n) = \frac{|\sum_j \hat{x}(i_1, j, n)\hat{x}^*(i_2, j, n)|}{\sqrt{\sum_j |\hat{x}(i_1, j, n)|^2 \sum_j |\hat{x}(i_2, j, n)|^2}} \tag{30}$$

should be high for some scale index n greater than that at which the primitive features were identified. If it is above a threshold t_c, the two features are combined into a single feature; otherwise they are left separate. This process can be used to join

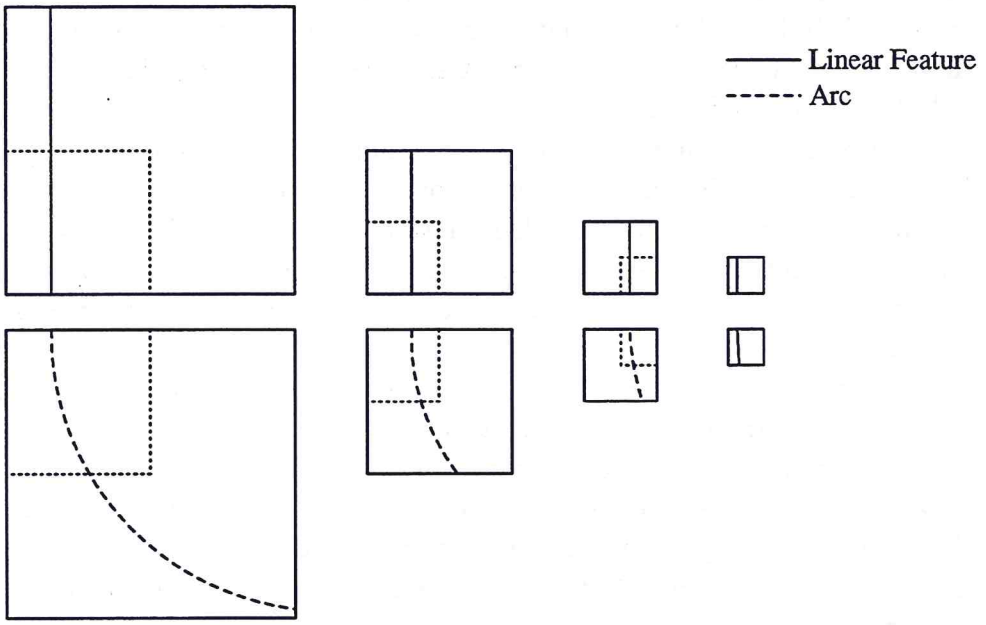

Figure 5: Combining low level features

compound features by determining whether the primitive features at the ends can be combined in this way.

5 Results

The algorithm was run over a number of test images and the results are presented in this section. In each case the levels of the MFT were thresholded so that all blocks containing low energy were ignored. The MFT uses a window that is bandlimited and is therefore not spatially confined [14],[15]. Indeed adjacent spatial regions will overlap and features between two region centers (which are separated by $\Xi(n)$) will be seen by both regions. To over come this, spatial features that are more than $\Xi(n)/2$ away from the region center will be considered to be outside the region. This means however that the spectrum of a region will be influenced by features in adjacent blocks. This will cause no problems if they are linear features in different orientations, but may cause interference for features of similar orientation. In these cases the spectral properties of a region may not fit the model, even if there is no more than a single feature within a distance of $\Xi(n)/2$ from the region center.

Figure 7 shows a simple 256 $\times$ 256 image containing a circle and a square. The algorithm was run over levels 3-6 of the MFT for this image but due to the relatively large scale of the features they were all extracted at level 3. Figures 8 and 9 show the magnitudes of MFT levels 3 and 5. Figure 10 shows the features extracted from the

image, with the blocks indicating the spatial regions in which they were found. The grey levels indicate features that were joined together by the linking process described in section 4. As can be seen the circle is represented as one feature, while the square is linked into four features.

Figures 11 - 13 show the 256 × 256 'Shapes' image and the magnitudes for two level of its MFT. The process was run over levels 3-6 and features were detected at a number of scales (Fig. 15) and then plotted on the original image (Fig. 15). The left and right points of the star are not clearly detected due the closeness in orientation of the upper and lower edges, but most of the major edges are correctly detected and linked.

For the 256 × 256 'Lena' (Fig. 16) image the same algorithm was used over MFT levels 3-6. Figure 17 shows the features detected in the image, and figure 18 shows the features shows these drawn over the original. The result may be compared to that obtained using the single linear feature model [6], (which used a 512 × 512 version of the image). The top of the hat has been missed due to the low contrast between the hat and the background, but most of the important features in the image have been identified correctly.

6 Conclusions

This paper contains a novel method of segmenting an image into regions, each of which contains either one or more linear features (straight lines or edge), or a single section of a circular arc edge or line. The features were modelled using the properties of local spectra, which are provided at different scales by the Multiresolution Fourier Transform (MFT). The image is initially segmented at a low level with small spatial regions, each of which is tested for the presence of one of the feature types. The parameters of these features are estimated for the regions in which they are detected. These are then grouped together into larger features, where the larger features also fit the models. Once this process terminates, more complex features are produced by combining primitive elements based upon the correlation between the linear features at their ends. The algorithm has been tested on some demanding synthetic and natural images containing a variety of polygonal and curved objects. It has been shown to be effective at dealing with complexity of the data, but further work is required to test its robustness in the presence of noise and other degradations. Extension of the model to deal with more general curve models is also under examination.

Acknowledgements

This work was supported by the UK SERC.

A Derivative Estimation

Since the modelling of the simple features is largely based upon the derivatives of the MFT phases, it is necessary to estimate these using the discrete data available. Although the more data used, the higher the accuracy of the derivative estimates, $\hat{\phi}_1(i,j,n)$ and $\hat{\phi}_1(i,j,n)$, for sake of computational simplicity it was decided to use a five sample approximation that fitted a minimum mean square error criterion.

Differentiating in the spatial domain is equivalent to multiplying by $\jmath\omega$ in the frequency domain. Approximating the frequency domain form of the estimator as the sum of two sine functions, the error, between the theoretical differentiator and the approximation at each ω is defined by

$$E(\omega) = (\omega - (a_1 \sin(\omega) + a_2 \sin(2\omega))^2 \tag{31}$$

giving a total error over the interval $[-0.5\pi, +0.5\pi]$

$$E = \int_{-0.5\pi}^{0.5\pi} (\omega - (a_1 \sin(\omega) + a_2 \sin(2\omega))^2 d\omega \tag{32}$$

Minimising this with respect to a_1 and a_2 requires

$$\frac{\partial E}{\partial a_1} = \frac{\partial E}{\partial a_2} = 0 \tag{33}$$

giving

$$\begin{aligned} \frac{\partial E}{\partial a_1} &= 2\int_{-0.5\pi}^{0.5\pi} -a_1(\omega - (a_1 \sin(\omega) + a_2 \sin(2\omega))d\omega \\ &= -4 + a_1\pi + a_2\frac{8}{3} \\ &= 0 \end{aligned} \tag{34}$$

Similarly, from the other partial derivative

$$-\pi + a_1\frac{8}{3} + a_2\pi = 0 \tag{35}$$

Solving these two equations simultaneously for a_1 and a_2 gives

$$a_1 = 1.5185 \tag{36}$$

$$a_2 = -0.2889 \tag{37}$$

Since multiplication in the frequency domain is equivalent to convolution in the spatial domain, this may be transformed to give a spatial sequence that can be convolved with the phase values. This sequence, normalised to give a derivative of one if the value changes by one between samples, is

$$\{-a_2, -a_1, 0, a_1, a_2\} = \{0.1535, -0.8069, 0, 0.8069, -0.1535\} \tag{38}$$

Calculating the derivative of phases using this approach may cause problems of rap around if performed in the complex domain. To get around this the five point linear combination shown above can be decomposed into two separate operations. The first operation is to find the forward difference of the phase in the complex domain from

$$\Delta_1\phi(i,j,n) = \delta_1 \arg\left(\hat{x}(\vec{\xi}_i(n),\vec{\omega}_j(n),\sigma(n))\hat{x}^*(\vec{\xi}_i(n),\vec{\omega}_{j_{(1,0)}}(n),\sigma(n))\right) \tag{39}$$

where

$$\vec{\omega}_{j_{(l_1,l_2)}} = \begin{bmatrix} \omega_1 + l_1 \\ \omega_2 + l_2 \end{bmatrix} \tag{40}$$

and δ_1 is the sampling interval and is required to convert from a phase difference to an estimate of derivative. These differences are then combined using the second operation - a four point integrator:

$$\begin{aligned} \hat{\phi}_1(i,j,n) &= (-a_2)\Delta_u\phi(i,j_{(-2,0)},n) + (-a_1-a_2)\Delta_u\phi(i,j_{(-1,0)},n) + \\ &\qquad (-a_1-a_2)\Delta_u\phi(i,j,n) + (-a_2)\Delta_u\phi(i,j_{(1,0)},n) \qquad (41) \\ &= 0.1535\Delta_u\phi(i,j_{(-2,0)},n) - 0.6535\Delta_u\phi(i,j_{(-1,0)},n) - \\ &\qquad 0.6535\Delta_u\phi(i,j,n) + 0.1535\Delta_u\phi(i,j_{(1,0)},n) \qquad (42) \end{aligned}$$

This derivative estimator has been defined in the frequency domain to approximate the the linear characteristic of the ideal differentiator, in the frequency interval $[-0.5\pi, +0.5\pi]$. As can be seen in figure 6 it fits well for these frequencies, but its performance tails off for higher frequencies.

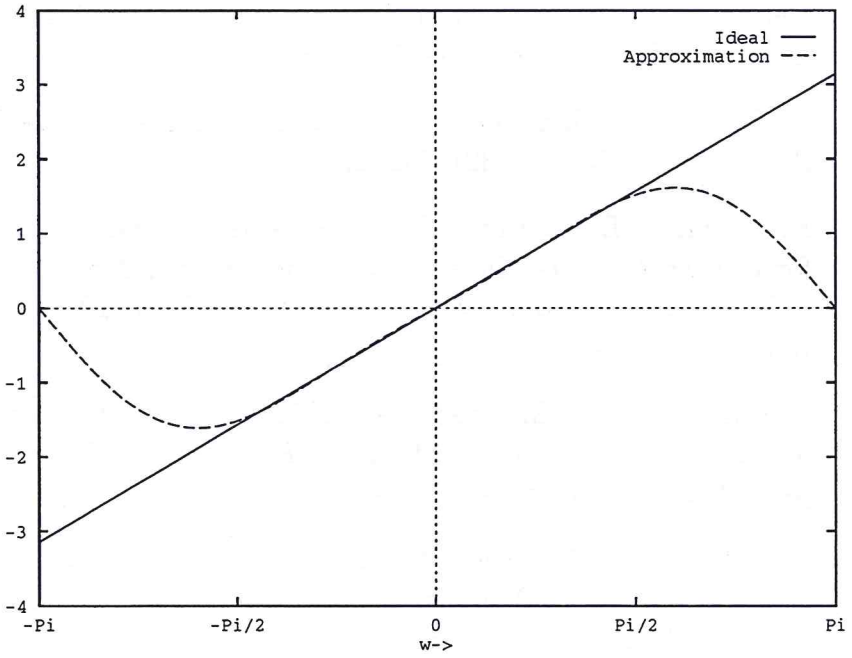

Figure 6: Differentiators in Frequency Domain

References

[1] D. H. Ballard and C. M. Brown. *Computer Vision.* Prentice Hall, New Jersey, 1982.

[2] J. Illingworth and J. Kittler. A Survey of the Hough Transform. *Computer Vision, Graphics and Image Processing*, 44:87–116, 1988.

[3] J. Princen, J. Illingworth, and J. Kitler. A Hierarchical Approach to Line Extraction Based on the Hough Transform. *Computer Vision, Graphics, and Image Processing*, 52(1):57–77, October 1990.

[4] S. W. Zucker, C. David, A. Dobbins, and L. Iverson. The Organization of Curve Detection: Coarse tangent fields and fine spline coverings. In *Proc. of Int. Conf. Computer Vision*, pages 568–577, 1988.

[5] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. *Int. J. Computer Vision*, 1:321–332, 1988.

[6] R. Wilson, A. D. Calway, and E. R. S. Pearson. A Generalized Wavelet Transform for Fourier Analysis: the Multiresolution Fourier Transform and its Application to Image and Audio Signal Analysis. *IEEE Trans. IT, Special Issue on Wavelet Representations*, 1992.

[7] A. Calway and R. G. Wilson. A Unified Approach to Feature Extraction Based on an Invertible Image Transform. In *Proc. 3rd IEE Int. Conf. Image Processing*, pages 651–655, Warwick, U.K, 1989.

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

[9] A. Calway and R. Wilson. Curve extraction in images using the multiresolution Fourier transform. In *Proc. of IEEE ICASSP-90*, pages 2129–2132, 1990.

[10] M. R. Portnoff. Time-frequency Representation of Digital Signals and Systems Based on Short-Time Fourier Analysis. *IEEE Trans. A.S.S.P*, 28(1):55–69, February 1980.

[11] D. Gabor. Theory of Communications. *Proc. IEE*, 93(26):429–441, November 1946.

[12] A. I. Borisenko and I. E. Tarapov. *Vector and Tensor Analysis.* Dover Publications, Inc., New York, 1979.

[13] C. W. Therrien. *Decision, Estimation and Classification.* J. Wiley and Sons, 1989.

[14] A. Papoulis. *Signal Analysis.* McGraw-Hill, New York, 1977.

[15] R. G. Wilson and G. H. Granlund. The Uncertainty Principle in Image Processing. *IEEE Trans. P.A.M.I*, 6(6):758–767, November 1984.

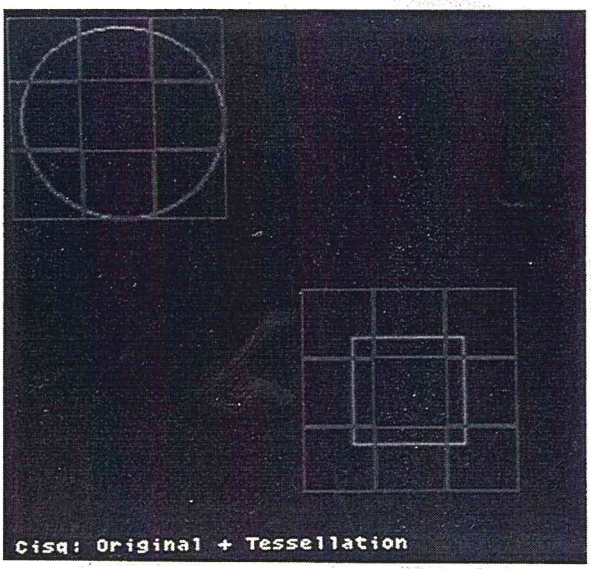

Figure 7: Circle+Square: Original

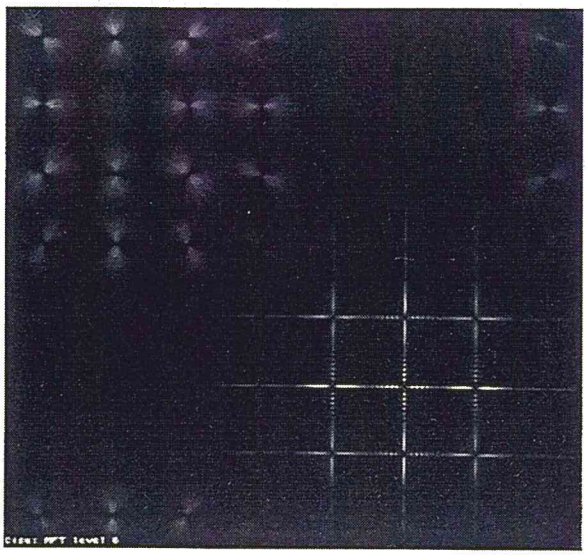

Figure 8: Circle+Square: MFT Level 3

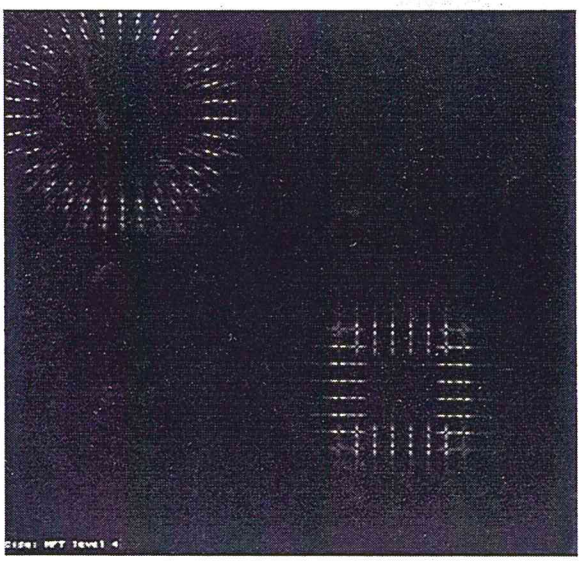

Figure 9: Circle+Square: MFT Level 5

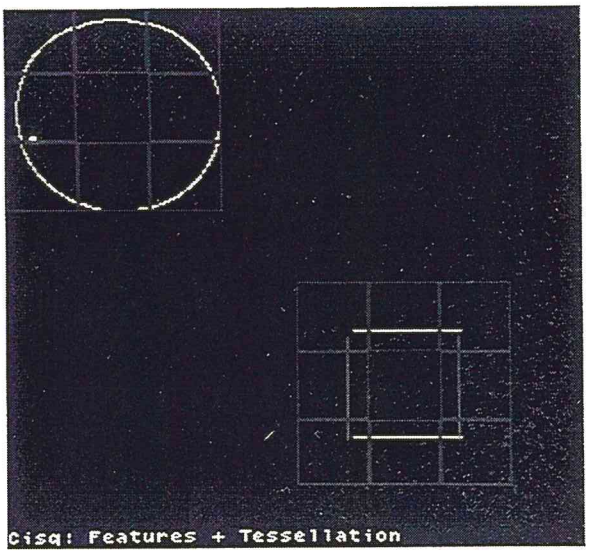

Figure 10: Circle+Square: Features

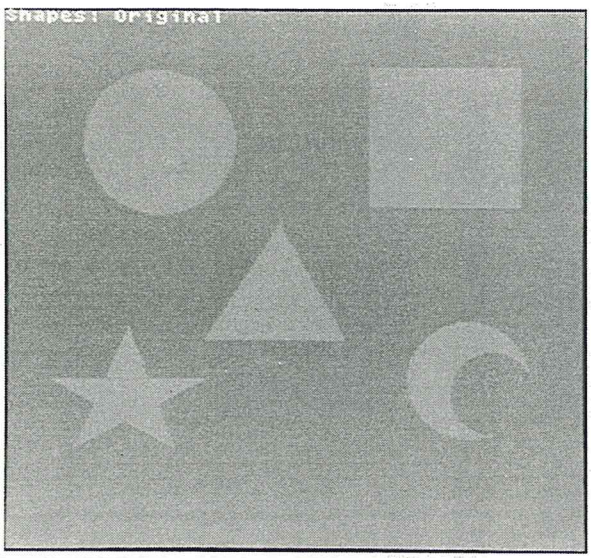

Figure 11: Shapes: Original

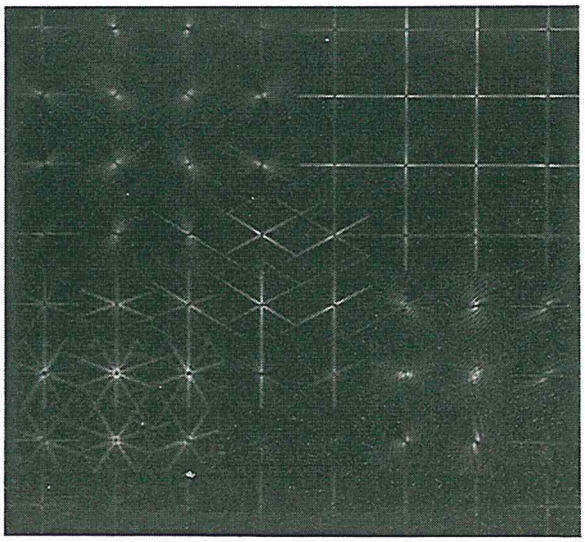

Figure 12: Shapes: MFT Level 3

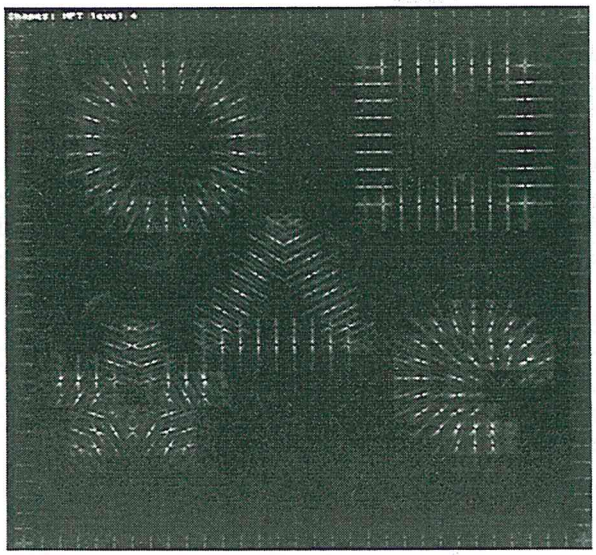

Figure 13: Shapes: MFT Level 5

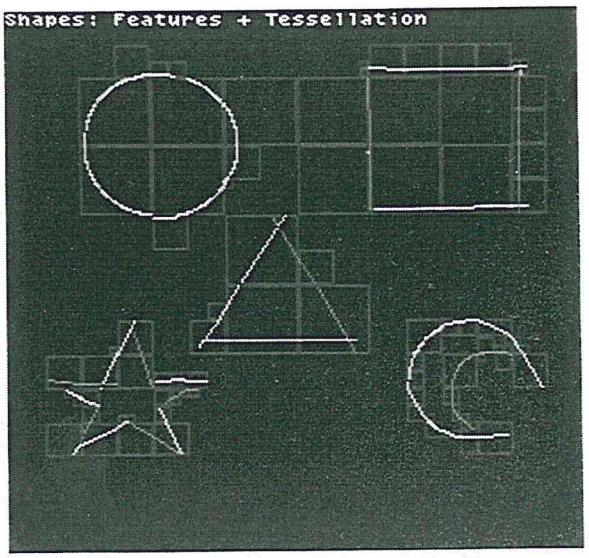

Figure 14: Shapes: Features

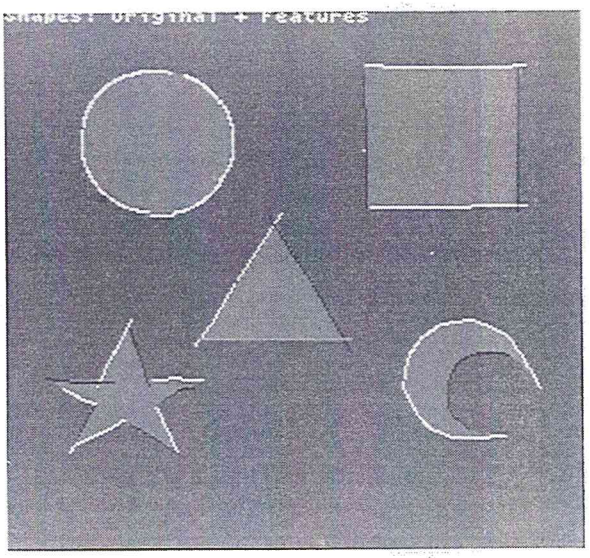

Figure 15: Shapes: Features Overlaid

Figure 16: Lena: Original

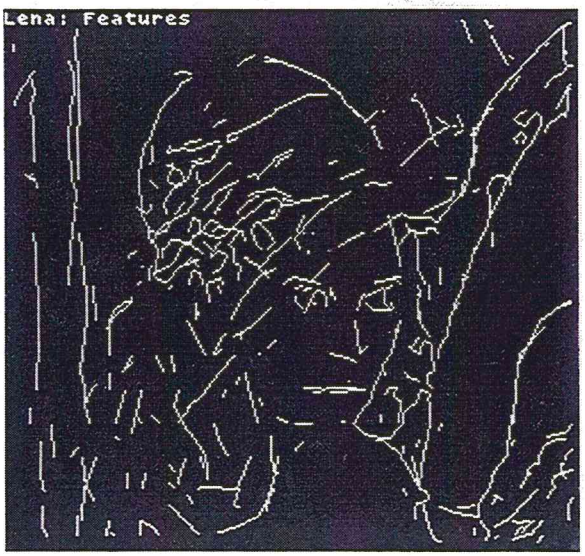

Figure 17: Lena: Features

Figure 18: Lena: Features Overlaid