Curve Extraction in Images Using the Multiresolution Fourier Transform*

A.Calway, R.Wilson
University of Warwick
Coventry CV4 7AL
England

Abstract

In this paper, a curve extraction algorithm which is defined within a multiresolution framework is described. An image model based on local features such as lines and edges is assumed and the parameters of this model are estimated using the multiresolution Fourier transform (MFT). The estimated features, which exist at different spatial resolutions, are then combined into curves using an appropriate curvature measure. The scheme is generally applicable and efficient in terms of computation. Furthermore, results presented for a natural image illustrate that the scheme compares favourably with existing approaches.

1 Introduction

There is often a requirement in image analysis for the identification of curves in an image. Such features are useful for providing an intermediate step towards the identification of higher order features such as shape and object boundaries. The traditional approach to the problem has been to apply some form of edge detection scheme and then to combine the edge points into suitable curves. There have been a number of techniques proposed, ranging from line following [1] to parameter space methods [2]. However, although these methods have found some success, it is often the case that they either lack generality or are computationally expensive.

The purpose of this work is to present a new approach to curve extraction based on a multiresolution framework. It has the dual advantage of being both generally applicable and efficient in terms of computation. Starting from a general image model, a feature estimation scheme is used which assumes the image to consist of a set of line and edge features defined at different spatial resolutions. The estimator is implemented by making use of the multiresolution Fourier transform (MFT) [3], an invertible transform that has been shown to be useful for a variety of image analysis problems [4, 5]. A curve representation is defined within the image model and a hierarchical extraction process is used to identify the curves. Results are presented for a natural image.

*Presented at the IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Albuquerque, NM, 1990

2 A Multiresolution Feature Model

The ideas described in this work are based upon a multiresolution image model. In its general form this has a linear recursive structure [6]

$$\mathbf{v}(n) = \mathbf{A}(n)\mathbf{v}(n-1) + \mathbf{B}(n)\mathbf{w}(n) \qquad 0 < n \leq N \tag{1}$$

where $\mathbf{A}(n)$ and $\mathbf{B}(n)$ are linear operators and $\mathbf{v}(n)$ and $\mathbf{w}(n)$ are vectors representing 2-d arrays of size $M \times M$, $M = 2^N$. These arrays are partitioned into $\Omega_n \times \Omega_n$ contiguous regions each of size $,_n \times ,_n$ and this structure is indicated by the appropriate number of vector indices, ie

$$\mathbf{v}(n) = [v_{xykl}(n)] \qquad \mathbf{w}(n) = [w_{xykl}(n)] \tag{2}$$

$$0 \leq x, y < \Omega_k \quad 0 \leq k, l < ,_k \quad ,_n \Omega_n = M \quad \Omega_n = 2^n$$

where xy defines a given region and kl a position within that region. In terms of the model, the vectors $\mathbf{w}(n)$ represent feature innovations and the region vectors $\mathbf{w}_{xy}(n)$ contain a finite number of features. Furthermore, subject to the initial condition $\mathbf{v}(0) = \mathbf{B}(0)\mathbf{w}(0)$ the image is given by $v(i, j) = v_{ij00}(N)$.

In the present work, two additional constraints are imposed upon the model. First, the innovation vectors $\mathbf{w}_{xy}(n)$ are restricted to a class that contain single local features

$$w_{xykl}(n) = h_{nxy}(n) + g(x,_n + k, y,_n + l) \tag{3}$$

where $h_{nxy}(n)$ is some locally defined real function which has an associated orientation θ_{nxy} and perpendicular offset η_{nxy}. This function represents the local feature. The function $g(x, y)$ is a lowpass image, which is assumed here to be the result of convolving a Gaussian kernel with the original image. Secondly, the model is limited such that a given image region is represented exclusively by a single innovation vector at an appropriate scale, ie the vector $\mathbf{v}_{xy}(n)$ is either equal to a new feature vector $\mathbf{w}_{xy}(n)$ or is equal to a quadrant of the relevant vector on the previous level. The structure of this model for a simple line drawing is illustrated in fig. 1.

3 Estimation of Model Parameters

The image model of eqns (1)-(3) requires the estimation of three parameters for each feature: its scale index n, its orientation θ_{nxy} and its offset η_{nxy}. These are obtained from the original image by using a combined estimation/decision procedure. A brief outline of the scheme is provided here, a more detailed description can be found in [4, 5].

3.1 Local Feature Modelling and Estimation

The estimation at a given scale index n is based upon a frequency domain model of oriented, localised features such as lines and edges. These features are characterised by a concentration of energy in the Fourier transform along a line passing through the origin and orthogonal to the orientation of the feature. Moreover, there exists a linear phase relationship amongst the coefficients such that the linear component is directly related to the position of the feature. This latter property enables a distinction to be made between 'localised' features and oriented texture fields [5].

It is assumed that the Fourier transform of the image (region), or an estimate of it, can be sampled at intervals $\rho_s = s\pi/S$ along any radius at an orientation θ to give a sequence of complex samples $u(\rho_s, \theta)$, $1 \leq s \leq S$. The model then takes the form of a 1-d complex normal Markov process

$$u(\rho_s, \theta) = \alpha u(\rho_{s-1}, \theta) + \beta v(\rho_s, \theta) \qquad s > 1 \tag{4}$$

where $u(\rho_1, \theta) = v(\rho_1, \theta)$ and $\alpha = |\alpha| e^{-j\phi(\theta)}$, $0 \leq |\alpha| < 1$. The complex innovations $v(\rho_s, \theta)$ are normally distributed with zero mean and unit variance. This model incorporates the linear phase requirement, since if a feature is present at an orientation θ and an offset η, then $\phi(\theta) = (\rho_s - \rho_{s-1})\eta$. Given this model, a maximum likelihood (ML) estimate for α can be obtained from the correlation statistic

$$R_\theta = \frac{1}{S-1} \sum_{s=1}^{S-1} u^*(\rho_s, \theta) u(\rho_{s+1}, \theta) \tag{5}$$

where the resulting estimates are unbiased, since [5]

$$\mathrm{E}\left[R_\theta \mid \ |\alpha| = 0\right] = 0 \qquad\qquad \mathrm{E}\left[R_\theta \mid \ |\alpha| > 0\right] = \alpha \tag{6}$$

However, in practice the above procedure must be approximated, since the statistic must be computed for a finite number of orientations and the scale index implies the use of a finite resolution spectral estimate. This means that it is necessary to use a set of correlation statistics $R_i(n)$, defined at scale index n and corresponding to orientations $\psi_i = i\pi/L_n$, $0 \leq i < L_n$. Although in general this will introduce bias due to interference and feature misalignment, it can be shown that by appropriate choice of window functions, acceptable estimates can be obtained given the presence of a single feature or several features in different orientations [5].

3.2 Single Feature Regions

The image model described in section 2 is based upon regions that contain a single local feature. Thus, having calculated the statistics $R_i(n)$ for a given image region, it is then necessary to determine whether these satisfy the single feature hypothesis. This is done in two ways. First, the distribution of magnitudes $|R_i(n)|$ over all orientations is considered and a measure of the anisotropy or 'orientedness' of the distribution is sought. The scheme is based upon the tensor analysis described in [7], in which the eigenvalues of the average inertia tensor provide a measure of the anisotropy. The principal axis of the tensor then defines the dominant orientation. Secondly, acceptance of the single feature hypothesis is also dependent upon a *scale consistency* criterion. This relates the feature estimate obtained at a given resolution to those estimates at an increased resolution and seeks a consistency of information measure [5].

The estimation/decision procedure is therefore defined as:

(i) From an estimate of the local spectrum at scale index n and position (x, y), compute the statistics $R_{ixy}(n)$ corresponding to the orientations $\psi_i = i\pi/L_n$, $0 \leq i < L_n$.

(ii) From the values $R_{ixy}(n)$, $0 \leq i < L_n$, determine the average inertia tensor and calculate its eigenvalues.

(iii) If the anisotropy measure exceeds the threshold and the scale consistency is acceptable, then the block is classified as containing a single feature and its orientation is given by the principal axis. The feature position is then derived from the argument of a correlation statistic calculated in the principal orientation using an interpolation procedure [5].

(iii) Otherwise, the single feature hypothesis is rejected and the block is split, the procedure being repeated for each of its 4 child subregions.

Once the above procedure is complete, the image will be represented by a structure which resembles that in fig. 1.

4 Estimator Implementation

The above estimation scheme is based upon the availability of local spectral estimates across a full range of scales. These are obtained from the MFT of the original image. In its cartesian separable form and for the image vector $\mathbf{v}$, $v_{ij} = v(i,j)$, $0 \leq i,j < M$, the transform is defined by an invertible linear operator $\mathbf{G}$ [3]

$$\mathbf{u} = \mathbf{G}\mathbf{v} \qquad \mathbf{v} = \mathbf{G}^{-\mathbf{1}}\mathbf{u} \tag{7}$$

which is defined by a set of level operators $\mathbf{G}(n)$, where n is the scale parameter, and the individual coefficients are given by

$$u_{xykl}(n) = \sum_{ij} G_{xyklij}(n) \;\; v_{ij} \qquad 0 \leq n \leq N \tag{8}$$

$$0 \leq x,y < \Omega_n \quad 0 \leq k,l < ,_n \quad \Omega_n = 2^n \quad \Omega_n ,_n = M$$

The vector $\mathbf{u}(n)$ is therefore given by the inner product of the vectors $\mathbf{G}_{xypq}(n)$ and the image $\mathbf{v}$. These vectors are finite prolate spheroidal sequences (fpss) [8]. They are bandlimited and maximally concentrated into separate and contiguous regions of size $,_n \times ,_n$ and $\Omega_n \times \Omega_n$ in the spatial and spatial frequency domains respectively. Readers familiar with combined representations will recognise that a given vector $\mathbf{u}(n)$ resembles a 2-d short-time Fourier transform, where the resolution in each domain is defined by parameters $,_n$ and Ω_n. The MFT is a whole can be considered as containing a set of such levels in which the resolution varies uniformly from the image at one extreme to its 2-d DFT at the other. A level can therefore be considered to consist of a set of local spectral estimates $\mathbf{u}_{xy}(n)$ which refer to contiguous image regions, the region size and spectral resolution being dependent upon the level. The use of the fpss basis functions means that the bias associated with such estimates is minimised [5].

The transform is used in the estimation procedure detailed in the previous section. Specifically, the coefficients $\mathbf{u} = \mathbf{G}\mathbf{v}_h$ are derived, where $\mathbf{v}_h = \mathbf{v} - \hat{\mathbf{g}}$ and $\hat{\mathbf{g}}$ is a vector representing the estimated lowpass image in eqn (3). Referring to step (i) and eqn (5), the correlation estimates are then obtained according to

$$R_{ixy}(n) = \mathbf{u}^{+}_{xy}(n)\mathbf{A}\mathbf{u}_{xy}(n) \tag{9}$$

where $^+$ indicates conjugate transpose. The operator $\mathbf{A}$ has a dual purpose. First, it ensures that the estimates are made along the orientations ψ_i within the vector $\mathbf{u}_{xy}(n)$ using a (bilinear) interpolation formula and, secondly, it incorporates the shift indicated in eqn (5). Using these estimates, the remaining steps in the estimation procedure are implemented.

5 Curve Representation and Extraction

The model of section 2 provides a straightforward way of representing curves in an image. It is a piecewise representation in which a curve is defined by a set of local features, where each feature corresponds to a section of the curve and the change in orientation between adjacent features defines a local curvature measure. An example is shown in fig. 2. Note that the regions associated with the local features vary in size according to the curvature: small regions corresponding to high curvature and large regions to low curvature.

The curve extraction scheme is based upon the above representation. It is a recursive scheme that operates on the truncated quadtree that represents structures of the type in fig. 1, ie the root node corresponds to the whole image and subsequent child nodes refer to quadrants as the image is recursively split. Leaf nodes of this tree then correspond to single feature regions. The idea is to use an upward directed process within this tree to form a curve by combining features at successively lower spatial resolutions, until at some node the curve is identified. The following steps are applied at each node in the tree, starting at the root node:

(i) For child nodes that are not leaf nodes do steps (i)-(iii).

(ii) Form curve sections by combining features or curve sections defined at child nodes. This is based upon a local curvature measure and a best fit search process.

(iii) Assign these sections to the current node.

The scheme is therefore a fine-to-coarse analysis that begins by descending to the nodes whose child nodes are all leaf nodes. Curve sections are formed and these are propagated up through the tree to enable them to be combined with other similarly formed sections or features defined at lower spatial resolutions. An example is illustrated in fig. 3.

6 Results

The curve extraction scheme was implemented on the (512×512) 8-bit grey level 'girl' image shown in fig. 4a. Results of the feature estimation are shown in fig. 4b, where a single feature region is shown as a block containing a line in the relevant orientation and position. The grey level regions indicate regions that have been classified as containing predominantly lowpass energy based on the magnitude values of the correlation statistics and the lowpass image $g(x, y)$. Fig. 4c. shows the result of the curve extraction algorithm. This is a binary image in which curves are represented by single lines which have been derived by fitting a B-spline representation to the set of features identified for a curve [5]. From these results, it can be seen that line and edge features have been detected at an appropriate scale and that the curves identified correspond closely to those in the image.

7 Conclusions

A curve extraction scheme based within a multiresolution framework has been described. The scheme is defined in terms of a general image model and incorporates the inherent multiresolution structure of curves. It is both generally applicable and can be implemented in an efficient manner within a hierarchical data structure. Results presented demonstrate that it is capable of identifying curves in natural images. Current work is aimed at extending the scheme to include the use of for example local curvature measures from the original image data.

Acknowledgement

This work has been performed with the support of the UK SERC. The authors would also like to thank the Computer Vision Laboratory at Linköping University for their help in preparing this paper.

References

[1] A.Martelli, *An Application of Heuristic Search Methods to Edge and Contour Detection*, Com. ACM 19, 73-83, 1976.

[2] J.Illingworth, J.Kittler, *Survey of the Hough Transform*, Comput. Vision Graph. and Image Process. 44, 87-116, 1988.

[3] R.Wilson, A.D.Calway,*A General Multiresolution Signal Descriptor and Its Application to Image Analysis*, Proc. EUSIPCO, Grenoble, 1988.

[4] A.Calway, R.Wilson, *A Unified Approach to Feature Extraction Based on an Invertible Image Transform*, Proc. 3rd IEE Int. Conf. Image Processing, 651-655, Warwick, 1989.

[5] A.Calway, *The Multiresolution Fourier Transform: A General Purpose Tool for Image Analysis*, University of Warwick Ph.D Thesis, 1989.

[6] R.Wilson, S.C.Clippingdale, *A Class of Nonstationary Image Models and Their Applications*, Proc. IMA Conf. Math. in Sig. Process., Warwick, 1988.

[7] H.Knutsson, *Representing Local Structure Using Tensors*, Proc. 6th Scan. Conf. on Image Analysis, Oulu, 1989.

[8] R.Wilson, *Finite Prolate Spheroidal Sequences and Their Applications I: Generation and Properties*, IEEE Trans. vol. PAMI-9, 787-795, 1987.

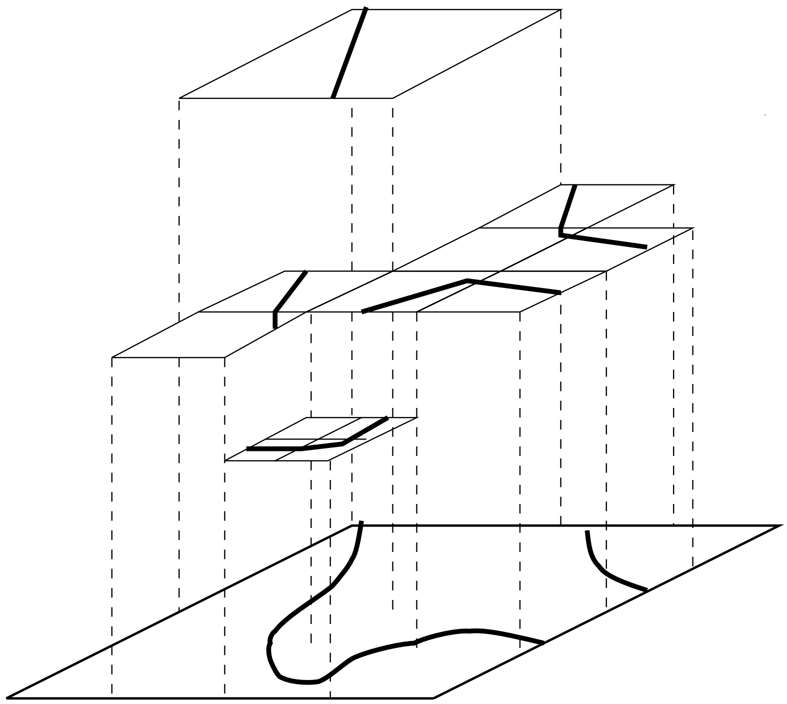


Figure 1: Example of multiresolution image model structure.

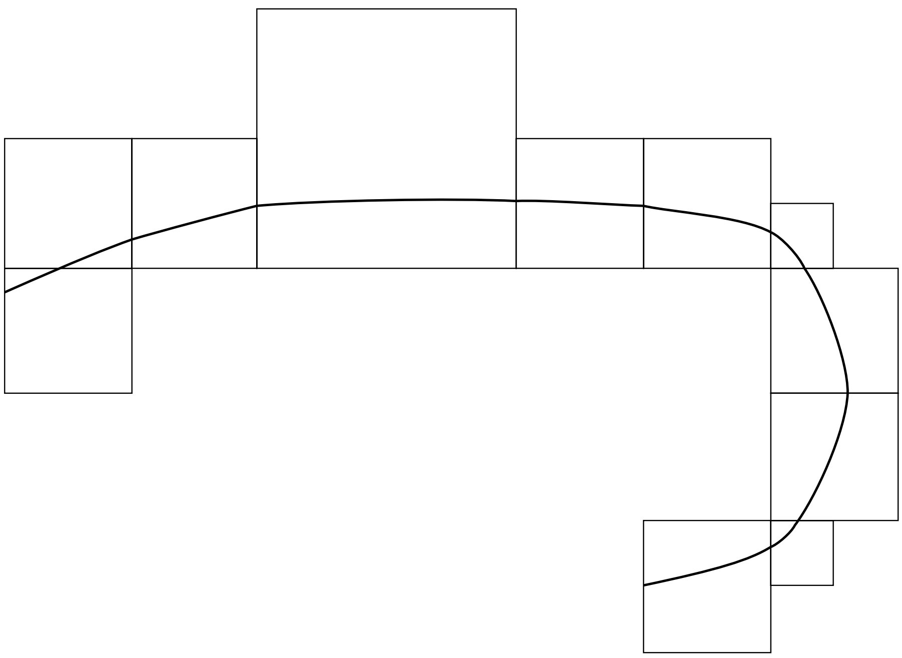


Figure 2: Curve representation.

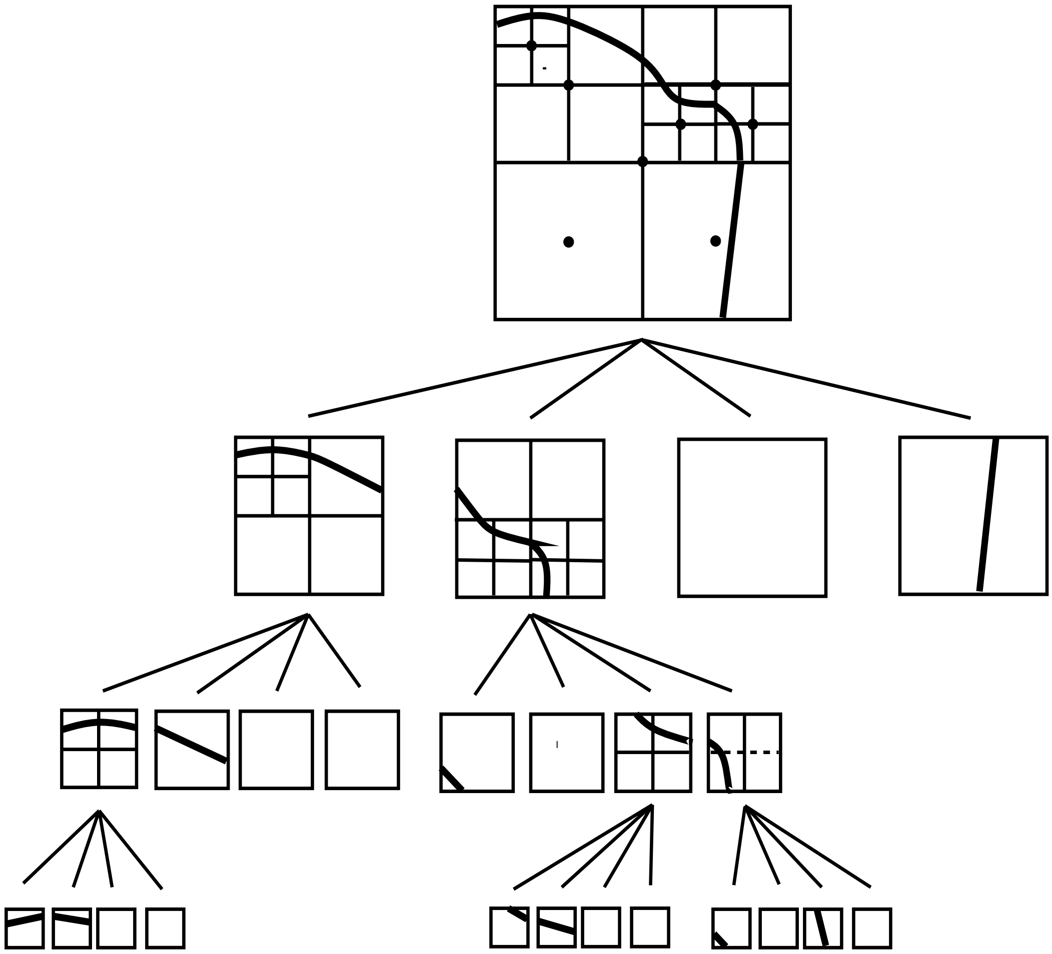
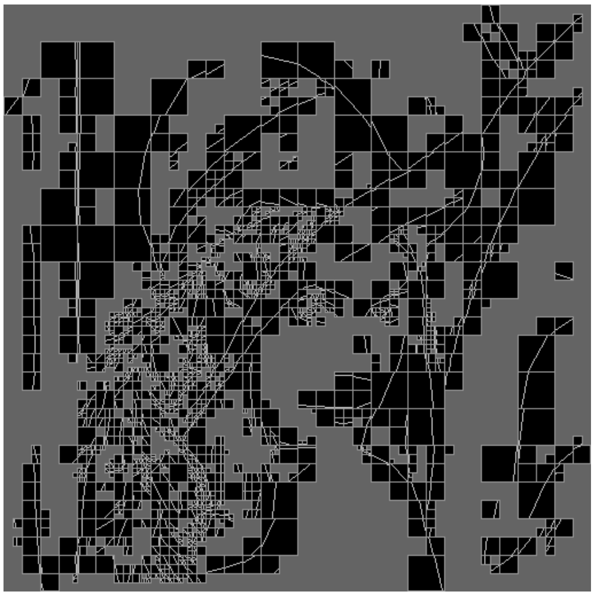


Figure 3: Hierarchical curve extraction.

(a)



(b)

(c)

Figure 4: (a) Original image (b) Local feature estimation (c) Curve extraction.