Image Analysis Using a Generalised Wavelet Transform

Andrew Calway

Introduction

The suitability of wavelet transforms (WT) for use in image analysis is well established: a representation in terms of the frequency content of local regions over a range of scales provides an ideal framework for the analysis of image features, which in general are of different size and can often be characterised by their frequency domain properties [1]. However, the standard form of wavelet decomposition, based on the translation and scaling of a single mother wavelet [2], has its limitations when considering general analysis problems. Apart from its lack of shift invariance [3], it also necessarily links scale and frequency: the size of a given region determines its representative frequencies within the transform. This latter property seems particularly restrictive given that there is no reason in general to assume that the frequency content of an image region should be related to its size. Thus, although the basic advantages of a wavelet approach are well-founded, the need exists for a form which addresses these limitations.

With this in mind, the Image and Signal Processing Group at the University of Warwick has developed a novel form of WT in which the link between scale and frequency is removed, and which provides a degree of shift invariance. Known as the Multiresolution Fourier Transform (MFT) [4][5], this resembles a stack of windowed Fourier transforms (WFT) [6] in which the window size is varied systematically to give a multiresolution representation of the space-frequency plane. As such, it constitutes a superset of the WT and WFT, providing a complete representation of the frequency domain at each scale and hence enabling regions to be analysed over a range of frequencies, yielding a flexibility not possessed by existing WTs. This has allowed the MFT to be used as the basis for tackling a wide range of problems, including linear feature and curve extraction [7]-[10], texture analysis [11] and stereopsis [12], and hence provides a framework for a unified approach to image analysis.

The Multiresolution Fourier Transform

It is simplest to describe the transform in terms of the 1-d continuous case. For a given scale σ, the MFT of the 1-d signal $x(\xi)$ at position ξ and frequency ω is defined as

$$\hat{x}(\xi,\omega,\sigma) \;=\; \sigma^{1/2} \int_{-\infty}^{\infty} w(\sigma(\chi-\xi))\, x(\chi)\; \exp[\,-\jmath\,\omega\,\chi\,]\, d\chi$$

where $\jmath^2 = -1$ and $w(\xi)$ is an appropriate window function typically chosen to be real and even, and to have good localisation both in the signal and frequency domains. For some position ξ, the MFT coefficients therefore correspond to the Fourier transform of the windowed signal $w(\sigma(\chi-\xi))x(\chi)$ and hence represent the frequency content of a region centred at ξ with extent defined by the scaled window $w(\sigma\chi)$. Note that the frequency resolution is inversely proportional to the size of the region and that varying the scale parameter σ alters the region size, thus providing a multiresolution representation. In common with both the WT and WFT, the MFT is invertible if the window is chosen carefully, although unlike the former there are

Department of Computer Science, University of Warwick.

several ways in which the original signal can be recovered, either from the coefficients at a given scale or via some form of multiple scale approach [4][5].

When dealing with discrete signals, the three coordinate axes ξ, ω and σ are sampled at appropriate intervals, the latter usually at powers of 2 in much the same way as the standard WT. The sampling in position and frequency follows that employed in a discrete WFT [6], with the intervals in each domain reflecting the extent of the window function and the requirement for a complete and invertible representation at each scale. For the discrete signal $x(\xi_k)$, its discrete MFT is therefore given by

$$\hat{x}(\xi_i(n), \omega_j(n), n) \; = \; \sum_k \; w_n(\xi_k - \xi_i(n)) \; x(\xi_k) \; \exp\left[-\jmath\, \omega_j(n)\, \xi_k \right]$$

where $\xi_i(n)$ and $\omega_j(n)$ define the sampling points in the signal and frequency domains respectively and $w_n(\xi_k)$ is the discrete window function at scale index n. Note the dependence of the sampling on the latter. Implementation is achieved either via the frequency domain using bandlimited window functions [4] or via the signal domain using truncated windows [6]. In either case, efficient computation is obtained using fast Fourier transform (FFT) techniques.

For the 2-d case, the most straightforward approach is to use a cartesian separable form of the MFT [4][5]. If the spatial sampling is arranged as a quadtree, then the transform as a whole resembles that shown in Figure 1, in which each level consists of a set of 2-d local spectra referring to regions centred according to the relevant node on the tree and with extent defined by the window function used. Note that each of these regions can therefore be described in terms of its associated frequency coefficients, providing a measure of 'local' shift invariance [5] and the capability of modelling region properties using Fourier analysis - an attribute which forms the basis of the applications described below.

Applications

The essential advantage of the MFT over that of a standard WT is that it gives the user greater flexibility in designing models and estimation procedures for analysis problems. In the applications, this is achieved by adopting the following general procedure [5]:

1. The use of local models and hypothesis tests based on frequency domain properties of the features of interest and defined in terms of the local spectral coefficients of the MFT.
2. A subsequent scale selection procedure which determines the optimal scale at which a given local model fits the available data based on the results of the associated hypothesis tests.

In other words, the approach is to identify suitable frequency domain models and hypothesis tests for different classes of feature and to determine the appropriate scale at which the models are valid. This has the advantage of being both efficient and sufficiently general to be applicable in a wide range of problems. Specific examples are described briefly below.

Linear Features and Curves. The initial application of the MFT was in the extraction of linear features and curves [7]-[10]. The scheme follows closely that of the two component procedure detailed above: a frequency domain model and estimator based on the linear phase properties of line and edge segments; and a top-down selection procedure defined within the quadtree framework which identifies the largest regions containing isolated segments. The

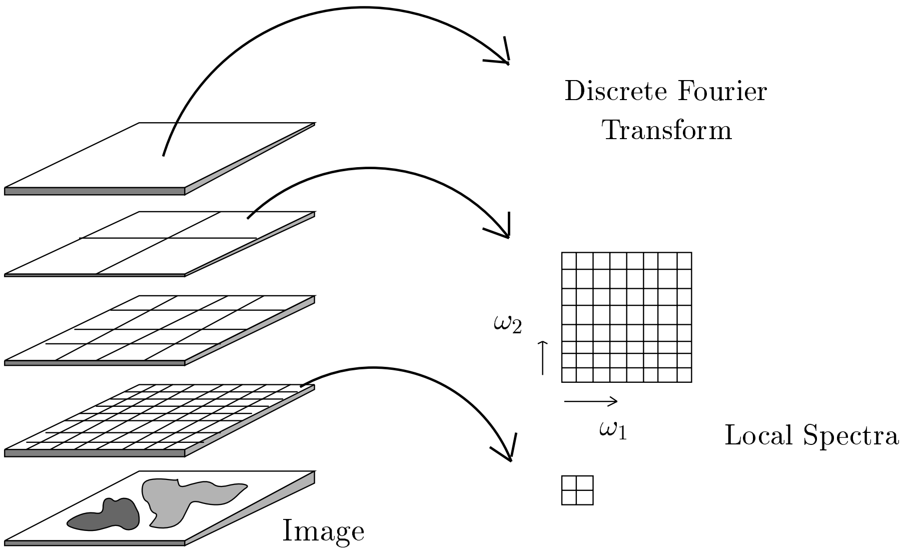

Figure 1: The 2-d cartesian separable MFT forms a quadtree structure in which its local spectra refer to the regions corresponding to the nodes of the tree, culminating at the top with the discrete Fourier transform of the image.

latter are relatively few in number and can be efficiently grouped together to form curves using a simple grouping algorithm based on a measure of continuity. The algorithm is fast, simple to implement and has been shown to be effective for extracting curves from a variety of natural images, comparing well with existing extraction algorithms.

Texture Analysis. A more recent application has been in the area of texture analysis [11]. This is based on the principle of modelling local regions of a homogeneous texture in terms of an affine coordinate transformation of a suitable prototype region - in short, to warp the prototype (or texton) into an approximation of the region of interest. It turns out that the transformation relating two regions can be efficiently computed via the frequency domain using their associated local spectra within the MFT and that the validity of the model can be tested by correlating the transformed prototype with the original region (again via the frequency domain by multiplication of the relevant spectra and inverse FFTs). Applying the procedure over a range of levels within the MFT framework then allows the selection of the optimal scale. Although this work is in its early stages, results obtained from synthesizing various classes of texture have demonstrated the considerable potential of the approach.

Stereopsis. The expediency of the MFT for performing efficient local correlations also forms the basis of a disparity estimation algorithm for use in stereopsis [12]. Using the MFTs of two binocular images, the disparity between given pairs of regions can be calculated via appropriate local correlations. This is incorporated into a coarse-to-fine 'focusing' strategy within the MFT framework, in which disparity estimates between regions at larger scales are used to direct correlations performed at smaller scales, thus yielding a recursive process which updates the disparity as it is tracked through scale. The result is a computationally efficient registration algorithm which is robust and avoids the extensive searching strategies normally encountered.

The Future

As the applications described above illustrate, the MFT has already been used successfully in a number of areas of image analysis. This work is continuing and is moving towards the incorporation of the various techniques into a single framework, thus allowing communication between the different processes and hence the potential for a unified image analysis tool.

References

1. S.Mallat, "Multifrequency channel decompositions of images and wavelet models", *IEEE Trans. Acous. Speech Sig. Proc.*, vol. 37, 12, pp. 2091-2110, 1989.

2. I.Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

3. E.P.Simoncelli *et al.*, "Shiftable multiscale transforms", *IEEE Trans. Information Th.*, vol. 38, 2, pp. 587-607, 1992.

4. A.D. Calway, *The Multiresolution Fourier Transform: A General Purpose Tool for Image Analysis*, Ph.D. Thesis, Warwick Univ., 1989.

5. R.Wilson, A.D.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 Trans. Inform. Th.*, vol. 38, 2, pp. 674-690, 1992.

6. M.R. Portnoff, "Time-frequency representation of digital signals and systems based on short-time Fourier analysis", *IEEE Trans. Acous. Speech Sig. Proc.*, vol. ASSP-28, pp. 55-69, 1980.

7. A.D. Calway and R. Wilson, "A unified approach to feature extraction based on an invertible image transform", *Proc. IEE Int. Conf. on Image Process. and its Appl.*, pp. 651-655, Warwick, England, 1989.

8. A.D. Calway and R. Wilson, "Curve extraction in images using the multiresolution Fourier transform", *Proc. IEEE ICASSP-90*, pp. 2129-2132, Albuquerque, 1990.

9. A.R. Davies and R. Wilson, "Curve and corner extraction using the multiresolution Fourier transform", *Proc. IEE Int. Conf. on Image Process. and its Appl.*, pp. 282-285, Maastricht, Holland, 1992.

10. A.D. Calway and R. Wilson, "Curve extraction in images using a multiresolution framework", Dept. of Computer Science, Warwick Univ., Int. Report RR236, 1992; also submitted for publication.

11. Tao-I. Hsu, A.D. Calway and R. Wilson, "Texture analysis using the multiresolution Fourier transform", submitted for publication.

12. A.D. Calway, H. Knutsson and R. Wilson, "Multiresolution estimation of 2-d disparity using a frequency domain approach", *Proc. British Machine Vision Conf.*, pp. 227-236, Leeds, England, 1992.