

A Fourier Approach to 3D Local Feature Estimation from Volume Data

Abhir Bhalerao and Roland Wilson
Department of Computer Science
University of Warwick
Coventry, UK
{abhir|rgw}@dcs.warwick.ac.uk

Abstract

This paper describes an estimation and representation method for object structure in 3D image data. A windowed Fourier transform is used to estimate the parameters of 3D local structures (lines, planes and spheres) across a number of window sizes or scales. We present the estimation method and describe a decision process which aims to find the simplest object description. Results of the method on test data and blood vessels from magnetic resonance imaging are presented.

1 Introduction

The modelling and estimation of structure in 3D imagery has many important applications, such as segmentation [1, 2]. Also, a concise representation that can capture ‘sufficient’ shape variation of objects within large data sets is desirable for image matching and visualisation, where computational efficiency is important [3]. In this paper, we apply the Multiresolution Fourier Transform (MFT) [4], extending ideas that have been successfully applied in 2D image analysis and motion estimation (e.g. [5]) to develop a scheme for 3D data.

The MFT is closely related to a short-time Fourier Transform (SFT) and more akin to an image pyramid or a Wavelet Transform (WT), employing a localised Fourier bases at varying window sizes, than a scale-space or feature space (e.g.[6, 7]). An MFT is an orthonormal and invertible representation but unlike a WT, its multiresolution decomposition is over-complete: a complete image representation exists at any given scale. Image feature modelling and estimation using the MFT, like the WT, have a number of advantages over single resolution methods and compares favourably with spatial filtering by a set of gradient kernels. The Fourier transform is fast to compute for any image block and only a single forward transform is required to make all parameter estimates. A continuous estimate of feature orientation is obtained, and the goodness of fit of the feature estimate can be checked directly by correlation. There is flexibility in feature modelling, such as modelling regional properties like texture as well as gradient changes. Local features are modelled at different window sizes, which reflects the nature of biological 3D imagery including medical imaging data that often exhibit self-similarity. The MFT naturally inherits all Fourier theorems and properties, such as shift-invariance, rotational equivariance, and invertibility, which are exploited by the parameter estimation described

below. Perhaps the greatest benefit of using a windowed scheme like the MFT, is decision schemes that follow feature estimation and operate through scale in a coarse-to-fine manner, which leads to simplified hypothesis testing and a recursive implementation.

2 3D Feature Estimation

The types of local structure which we wish to model are: linear structures (lines and cylinders), planar structures and spherical or isotropic structures. Examples of these from medical images are respectively: blood vessels and nerve fibres, bone surface and interfaces between tissue types. The global shape of an object, in general, cannot be modelled by a single such primitive structure, hence the need to localise the model to a small neighbourhood. In the continuous spatial domain, if a feature such as line/cylinder, plane, or sphere is windowed by a smooth function, then it can be approximated by a 3D Gaussian:

$$f(\vec{x}) = (2\pi)^{-\frac{3}{2}} |C|^{-\frac{1}{2}} \exp(-(\vec{x} - \vec{\rho})^T C^{-1} (\vec{x} - \vec{\rho})/2) \quad (1)$$

centred on $\vec{\rho}$ with $C = R^T C' R$, where C' is the diagonal matrix of variances representing the extent of the function in the major axes and R is the matrix of rotation from the feature orientation vector $\vec{\theta}$ to the x-axis. For lines, the feature orientation vector is aligned along the feature; for planes it is the normal vector to the plane. For spheres, $\vec{\theta}$ has no meaning and R can be the identity with C' being diagonal.

The Fourier transform of this function $f(\vec{x}) \leftrightarrow \hat{f}(\vec{u})$ is also a 3D Gaussian with the spectral energy distribution dependent on the type of feature. For linear features, the spectral distribution will lie in a plane orthogonal to $\vec{\theta}$. For planar features, it is along the planar normal $\vec{\theta}$ and for spherical features, it will be evenly distributed and the compactness will be inversely proportional to the spatial size of the sphere.

2.1 Feature orientation estimation

Because the feature orientation lies along one of the axes of the ellipsoidal mass in the Fourier domain, it can be estimated by finding the principal components of the moment of inertia tensor I of the spectral energy:

$$I = \frac{1}{B^3} \sum_{\vec{u}} \vec{u} \vec{u}^T |\hat{f}(\vec{u})|^2 \quad (2)$$

where B is the block or window size. Performing PCA analysis on the symmetric tensor I yields 3 eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$ and corresponding eigenvectors \vec{e}_i , $1 \leq i \leq 3$.

The estimate of the local orientation of the feature, $\vec{\theta}$, will depend on how the results of the PCA analysis are interpreted. For linear features, $\vec{\theta} = \vec{e}_1$ and for planar features, the normal vector, $\vec{\theta} = \vec{e}_3$. For isotropic regions, all three eigenvectors can be used to estimate the spatial extent.

2.2 Local structure classification

Measures of local linearity of co-planarity of the image block can be formed by considering ratios of the eigenvalues λ_i [1]:

1. **planes:** If the feature is planar then $\lambda_1 \gg \lambda_2 \approx \lambda_3$ and the index

$$c_p = \frac{\lambda_1 - \lambda_2}{\lambda_1} \rightarrow 1 \quad (3)$$

2. **lines:** If the feature is linear then $\lambda_1 \approx \lambda_2 \gg \lambda_3$ and the index

$$c_l = \frac{\lambda_2 - \lambda_3}{\lambda_1} \rightarrow 1 \quad (4)$$

3. **spheres:** If the feature is isotropic then $\lambda_1 \approx \lambda_2 \approx \lambda_3$ and the index

$$c_s = \frac{\lambda_3}{\lambda_1} \rightarrow 1 \quad (5)$$

Together with the average spectral energy, these indices can be used to classify a block at any spatial size. Note that $c_p + c_l + c_s = 1$.

2.3 Feature centroid estimation

The phase spectrum, $\phi(\vec{u})$, will be related to the position or centroid $\vec{\rho}$ of the spatial function if the window function for the image block is real and even, via the Fourier shift theorem [8]:

$$\phi(\vec{u}) = \arg(\hat{f}(\vec{u})) = -\vec{\rho} \cdot \vec{u} \quad (6)$$

This expresses the linear-phase relationship between the position of the feature within the window and the gradient of the phase across the spectrum. The feature centroid $\vec{\rho}$ is then estimated by taking average pairwise correlations between neighbouring coefficients in each of the three axes directions, i . Since $\hat{f}(u_i - 1)\hat{f}(u_i)^* = \phi(u_i - 1) - \phi(u_i)$, then

$$\rho_i = \frac{1}{2\pi B^3} \sum_{\vec{u}} \hat{f}(u_i - 1)\hat{f}(u_i)^* \quad (7)$$

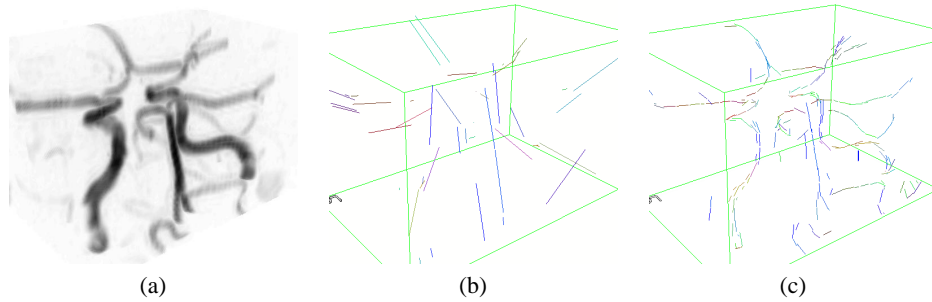


Figure 1: (a) Maximum intensity projection of part of a MR angiogram depicting cerebral blood vessels. (b) and (c): linear feature estimates at window sizes $B = 16, 4$.

3 Object Representation Across Scales

The goal of the feature estimation and local structure type classification is to produce a succinct global object representation. At large window sizes, the single feature hypothesis (which simplifies the estimation scheme) is generally false and multiple features of multiple types may be present. Therefore, the data are likely to be better explained at the *next* smaller window size. If the window sizes are chosen to decrease by a factor of two, then there will be 8 sub-blocks (or children) of each block (or parent) forming an oct-tree feature representation of the data. The desired compact representation will form the leaves of a pruned sub-tree within the original feature oct-tree.

A recursive coarse-to-fine process scale selection process can be used to move from each parent block with its children, either accepting the parent or repeating the process in a pre-order traversal. The scale selection criteria are:

1. **accept** the feature estimate at the parent over the children by confirming that one or more children features form *all* parts of the feature described in the parent window.
2. **reject** the feature estimate (and type classification) at the parent because the child features disagree with the parent. This may be because of the introduction of new features additional to that at the parent or because the feature types differ.

An example of (1) will be if a single linear feature is seen equally at both scales or within some tolerance such as the degree of curvature or co-linearity. An example of (2) would be if multiple, possibly separate linear features are present at the two scales. Note that in (1), the parent is favoured for reasons of greater robustness of the estimate at the larger (parent) window size. Also, rejecting the hypothesis in (2) will catch any degenerate cases and allow the hypothesis to be tested at the next finer scale.

4 Experiments

Figure 1 shows estimates of linear features in part of a 3D magnetic resonance angiogram (MRA) (size $88 \times 58 \times 44$) at window sizes $B = 16, 4$ in figures (b) and (c) respectively. A maximum intensity projection of the vessel data at the same view orientation is given in (a). The position and orientation of larger vessels are captured in (b) but, unsurprisingly, in some parts a *single* feature estimate is invalid. At the smaller window size, finer blood vessels and bifurcations become apparent.

A 3D synthetic image size $64 \times 64 \times 64$ was used to test the local feature-type discrimination indices $\{c_p, c_l, c_s\}$ (figure 2(a)). The top row (2(b)-(d)) depicts the feature estimates at $B = 16, 8, 4$ classified as planar patches (shown as a circle radius $B/2$). Similarly, in figures 2(e)-(g), the estimates are classified as lines and the index c_l is used to represent the estimate magnitude. In (b), the estimates on the sphere have the highest c_p values, whereas the patches along the cylinders have a low planar index. Conversely, in (e), the feature estimates have greater validity as lines using c_l along the linear features in the test data (the diagonal ends of the intersecting cylinder and the vertical cylinder in the right foreground). This pattern is partially repeated at the next window size ($B = 8$) shown in (c) and (f), but the curvature of the cylinders is sufficiently low for it to be seen locally equally well as a plane. For the same reasons, at the smallest window size ($B = 4$), the planar index c_p appears to be the same across all blocks.

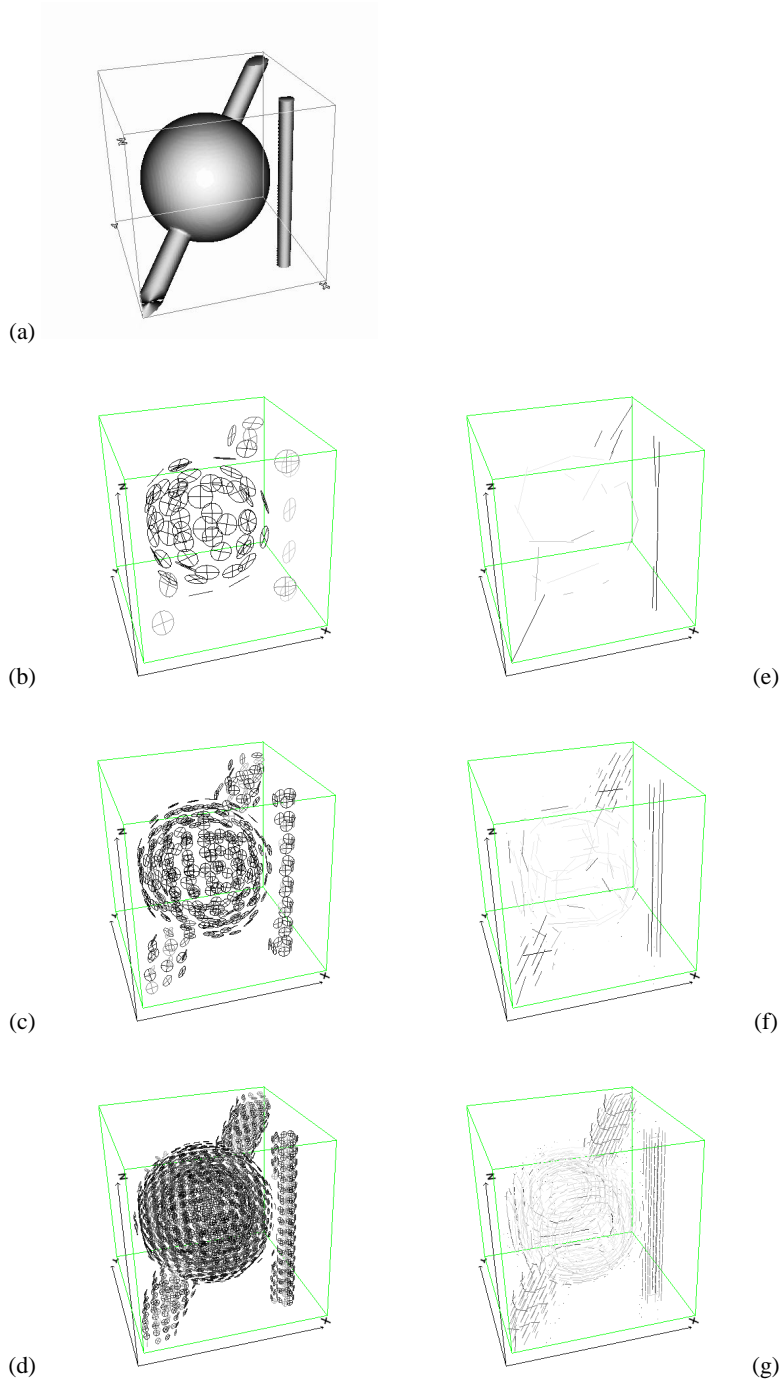


Figure 2: (a) Synthetic 3D test image containing a sphere and two cylinders with different diameters. (b)-(d): each block interpreted as plane where intensity proportional to c_p . (e)-(g): each block interpreted as line where intensity proportional to c_l .

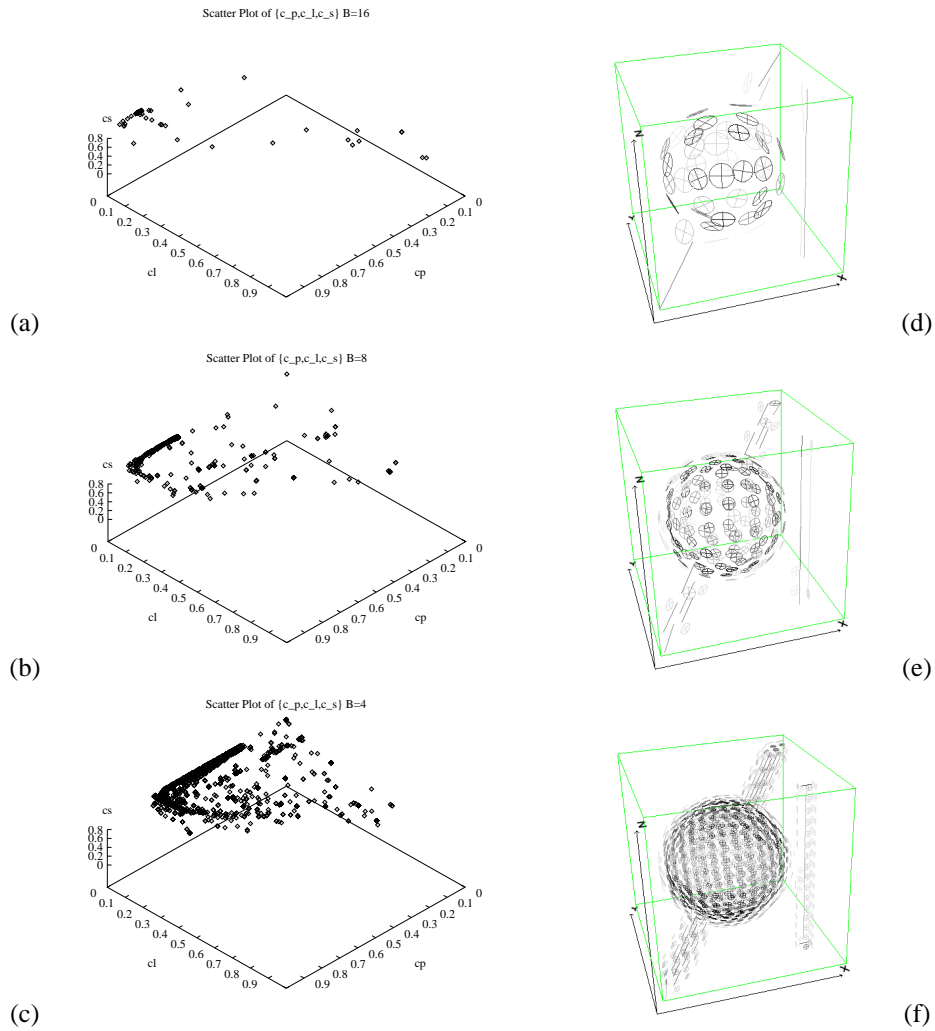


Figure 3: (a)-(c) Scatter plots of local feature-type index classification space $\{c_p, c_l, c_s\}$ across scales $B = 16, 8, 4$ for the mixed object test images (figure 2(a)). (d)-(f) Classification of blocks into linear over planar if $c_l > c_p$ with feature intensity proportional to average spectral energy at block sizes $B = 16, 8, 4$ respectively.

To visualise the $\{c_p, c_l, c_s\}$ classification space across scale, we produced scatter plots for all blocks at a given size (figures 3(a)-(c)). In (a), there are clusters around $c_p \approx 1$ and $c_l \approx 1$. In (b), a line of points begins to form across the $c_l = 0$ plane, spreading toward $c_s \approx 1$. For $B = 4$ in (c), the points begin to become evenly spread in the plane $c_p + c_l + c_s = 1$. In figure 3(d)-(f), we classified the feature estimates independently at each window size into lines or planes by testing whether $c_l > c_p$. What is encouraging is that at all scales the surfaces of the sphere is seen locally as a plane and is depicted as a planar-patch, while those blocks containing parts of the cylinders are correctly classified as being locally linear.

To test out the scale selection based on the scheme given above (section 3), we used two MRA data sets: one containing a giant aneurysm, which is a balloon like structure caused by the distending of an artery (figure 4(a)), and a complete MRA scan of the head-neck arteries and veins from a normal subject (data size $256 \times 256 \times 60$) (figure 5(a)) that has vascular structure at coarse and fine scales. In both cases because of the relatively low SNR, it is impossible to achieve a satisfactory segmentation containing both large and fine structures by manually setting an object/background threshold. We used a coarse-to-fine traversal of 3 levels of an oct-tree with overlapping (by 50%) window sizes $B = 32, 16, 8$ to produce a combined classification and scale-selection. The selection criteria were:

1. At each scale, only consider feature blocks for which the linear/planar feature certainty was greater than a half, i.e. $c_l > 0.5$ or $c_p > 0.5$.
2. **accept** a feature estimate at a parent over a child if the feature type classification (i.e. $\max(c_l, c_p)$) at the parent was greater than *any* child.
3. **reject** a parent if the feature classification at any child disagrees with a parent.

Test 1 will discard most features and only consider those that look like lines or planes. Test 2 accepts parent features if they are better fit to the data. Test 3 over rules test 2, favouring a diversity of feature types at the next finer scale.

Figure 4(b) shows a scale-selection on the aneurysm data set with the arrow labelled 1 marking the wall of the giant aneurysm correctly represented by planar patches, whereas the other arteries in the image are classified as lines (e.g. arrow labelled 2). The presence of multiple estimates at some points is a result of the overlapping window structure used in these experiments.

Results of feature estimation and scale selection for the head-neck data set are given in figures 5(b)-(f). Figures 5(b)-(d) show a linear *only* classification of blocks at three separate scales: $B = 32, 16, 8$ respectively. The large vein (sagittal sinus denoted by the arrow labelled 1) is seen at all levels. The vessels which feed the head from the neck (arrow labelled 2), are unbundled from the coarse to fine representation with most of the finer vessels clearly indentifiable in (d). This is an example of *new* features being introduced at the next finer scale which we do not deal with explicitly in the current scale-selection. In figure 5(e) and (f), we have used the scale-selection on the feature estimates shown in (b)-(d), to produce a combined object representation. We have restricted the classification to only pick and show linear features in (e), i.e. only blocks for which $c_l > c_p$. This results is less noisy than (d) and contains 14, 123 and 1943 blocks from the three scales respectively. In (f), we allow both lines and planes to coexist in the process using $\max(c_l, c_p)$ to make the feature type classification at any block. As expected, most of the image is dominated by linear features but, other than the a few planes found along

the large vein (arrow labelled 3) in the centre of the image, what is encouraging is that some of the artifactual fat signal from the patient’s skin surface is correctly seen as planar patches by the tessellation (arrow labelled 4 in (f)). In this image the scale selection picked 14, 148 and 2335 blocks across the three scales.

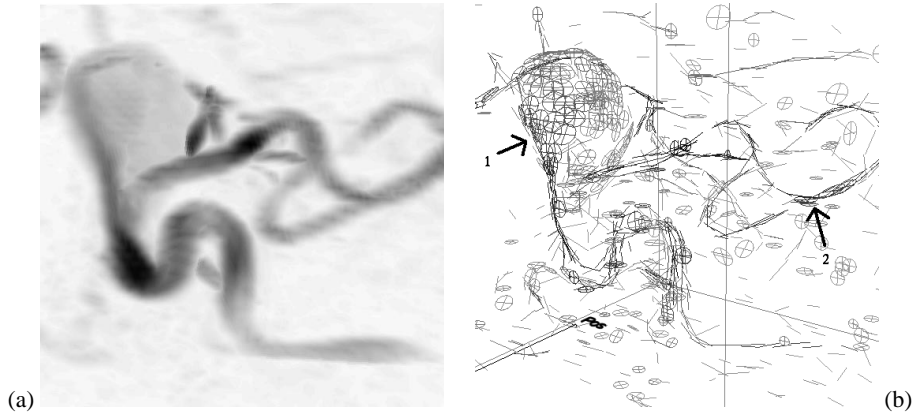


Figure 4: (a) MIP of part of an MRA data set depicting a giant aneurysm. (b) A multi-level object representation using 3 oct-tree scales: $B = 32, 16, 8$ from which 4, 34 and 1108 (overlapping) blocks of linear or planar classification have been selected using $\max(c_l, c_p)$ and a parent-child decision scheme (see main text). The walls of the aneurysm are classified as planes (arrow marked 1) while the other vessels are classified as lines (e.g. vessel at arrow marked 2).

5 Conclusions

We have described a windowed Fourier based method for describing 3D local structure data which could be employed for feature representation in volume data, such as medical imaging data, for a variety of applications. The Multiresolution Fourier Transform (MFT) is used to estimate the parameters of local features classified into lines, planes or spheres. As well as estimating the feature orientation and position, we outlined a way of classifying features by considering the principal components of the second-order moment of the energy spectrum. An important aspect of this method is that we perform the estimation and hypothesis testing across a number of window sizes or scales, which enables the local structure to be fixed and its extent relate to the window size.

The example results demonstrate that the simplest of classification methods can effectively find a ‘descriptive’ local structure for any window size. The object representation scheme proposed in section 3 is an extension to that used in [4] for curve extraction in 2D. The scale selection is complicated in 3D because there are multiple feature types possible at each block such as at vessel bifurcations in angiographic data. In such instances, a superposition of local Gaussian features can be used to model the region with the individual features being estimated separately using the local Fourier spectra [9]. Also, in the presence of noise, the feature estimate is less robust and some form of local regression can be usefully improve on the initial estimate. To achieve a segmentation it is necessary

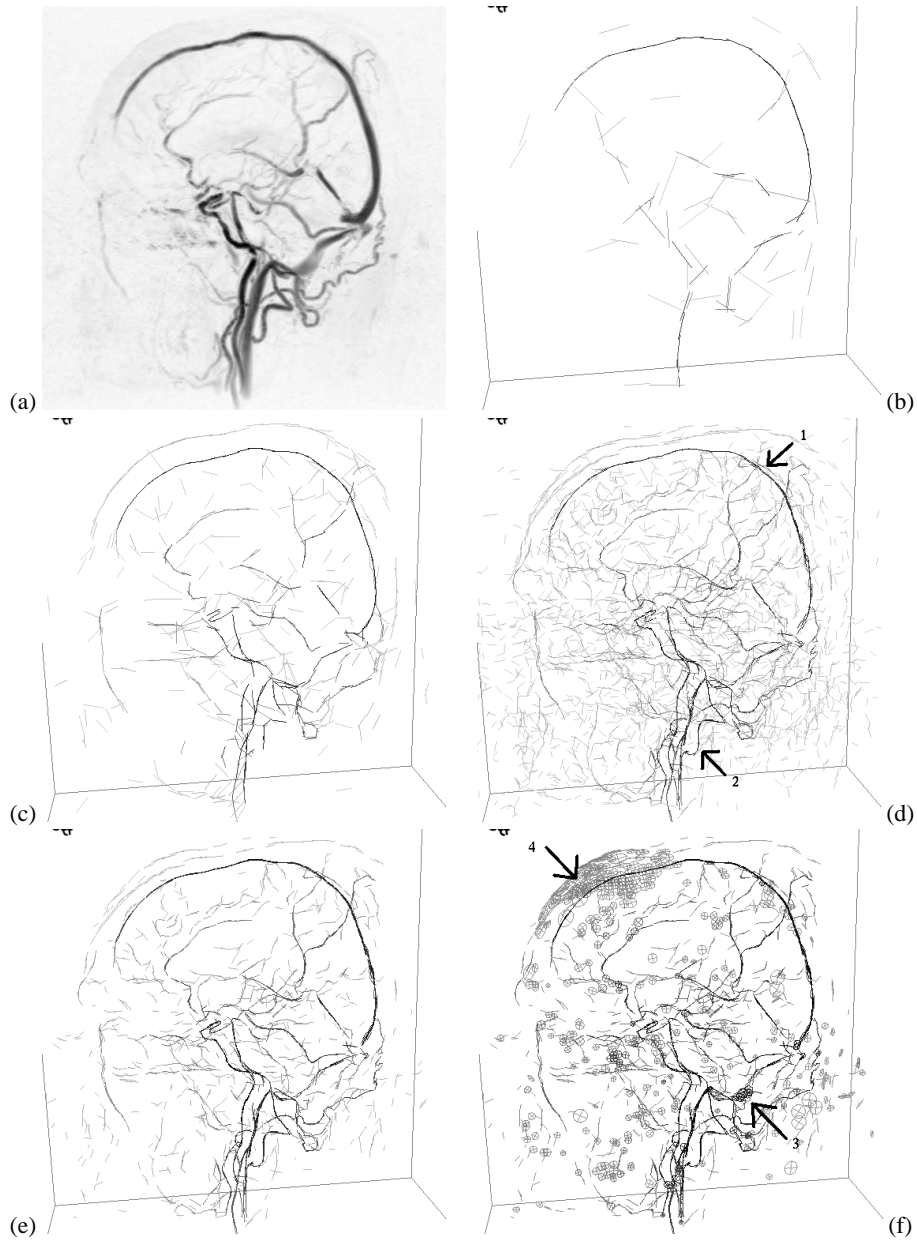


Figure 5: (a) MIP of a normal MRA data set size $256 \times 256 \times 60$ depicting major arteries and veins of head and neck. (b)-(d): Linear feature estimates at 3 block sizes: $B = 32, 16, 8$ respectively where feature intensity is proportional to estimated amplitude. Only linear estimates for which $c_l > 0.5$ are shown. Large structures such as the vein (sagittal sinus) marked by arrow labelled 1, appears across all three levels. The bundle of vessels in the neck (marked by arrow 2) are properly discriminated at the finest scale. (e) A multi-level tessellation of (overlapping) blocks from the three scales shown in (b)-(d) selecting only linear features using the parent-child decision scheme (see main text). (f) Same as (e) with this time taking $\max(c_l, c_p)$ allowing planar features to be incorporated (e.g. surface features marked 3 and 4).

to then connect together the elements of the object representation based on continuity and labelling of features across neighbouring blocks. We are currently investigating both a maximum-likelihood iterative estimation method to improve the initial feature estimates obtained from the Fourier approach described here and, a stochastic scheme to achieve a global description of the objects [10].

References

- [1] C-F. Westin, A. Bhalerao, H. Knutsson, and R. Kikinis. Using Local 3D Structure for Segmentation of Bone from Computer Tomography Images. In *Proc. of Computer Vision and Pattern Recognition '97*, Puerto Rico, 1997.
- [2] L. M. Lorigo, O. Faugeras, W. E. L. Grimson, R. Keriven, R. Kikinis, A. Nabavia, and C-F. Westin. Codimension-Two Geodesic Active Contours for MRA Segmentation. In *Proc. of Intl. Conf. on Information Processing in Medical Imaging*, 1999.
- [3] S. Warfield, A. Robatino, J. Dengler, F. Jolesz, and R. Kikinis. Nonlinear Registration and Template Driven Segmentation. In A. W. Toga, editor, *Brain Warping*, pages 67–84. Academic Press, 1999.
- [4] 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*, 38(2):674–690, 1992.
- [5] S.A. Kruger and A.D. Calway. A Multiresolution Frequency Domain Method For Estimating Affine Motion Parameters. In *Proceedings of the IEEE International Conference on Image Processing*, pages 113–116, 1996.
- [6] T. Lindeberg. Scale-Space: A framework for handling image structures at multiple scales. In *Proc. CERN School of Computing*, Egmond ann Zee, The Netherlands, 1996.
- [7] M. E. Martínez-Pérez, A. D. Hughes, A. V. Stanton, S. A. Thom, A. A. Bharath, and K. H Parker. Scale-Space Analysis for the Characterisation of Retinal Blood Vessel. In *Proceedings on Medical Image Understanding and Analysis (MIUA)*, pages 57–60, Oxford, 1999.
- [8] A. Papoulis. *Signal Analysis*. McGraw-Hill, New York, 1977.
- [9] A. R. Davies and R. Wilson. Curve and Corner Extraction using the Multiresolution Fourier Transform. In *Proc IEE Conf. Image Processing and its Applications*, pages 282–285, 1992.
- [10] A. Bhalerao, E. Thönnies, W. Kendall, and R. Wilson. Inferring Vascular Structure from 2D and 3D Imagery. In *Proc. MICCAI'01 (to appear)*, 2001.