Estimating Local and Global Image Structure using a Gaussian Intensity Model

Abhir Bhalerao and Roland Wilson

Department of Computer Science
University of Warwick
Coventry
{abhir,rgw}@dcs.warwick.ac.uk

Abstract. This paper describes an estimation and representation method for object structure in 2D and 3D imagery. Local image features are modelled with Gaussian intensity profiles and estimated by a combination of a multiresolution, windowed Fourier approach followed by an iterative, minimum mean-square estimation. The method is computationally efficient and robust, giving accurate estimates of feature position, orientation and size. A structure classification and inference scheme is proposed, which operates over a range of window sizes using the local parametric estimates to find the 'simplest' global object description. Results are presented of estimating vascular (tree-like) structures from 2D and 3D imagery (retinal images and MR angiograms) and preliminary results of a generalised structure classification on 3D synthetic data.

1 Introduction

The problem of inferring object structure from two and three dimensional medical imagery is an important one. Having a concise representation that can capture 'sufficient' shape variation of objects within large data sets is desirable for image segmentation, matching and visualisation, where computational efficiency is important. Previous work in this area has ranged from modelling specific features using deformable templates (e.g for vasculature [1]), filtering and adaptive thresholding [2], and using scale-space approaches [3,4], to recently reported work using level-sets [5]. Statistical methods for image analysis typically use likelihood techniques like Expectation Maximisation to model the global intensity variation, e.g. [6], from which a segmentation can be derived. Here, we use local likelihood maximisation using a Gaussian model of the spatial intensity profile, followed by a structure classification and inference scheme based on the goodness of fit of the model to the data over a range of spatial scales (or window sizes). As well as parameterising the data, the resulting feature likelihood lends itself to the use of Bayesian inference schemes in which prior knowledge about local structure can be explicitly modelled.

2 A Gaussian Intensity Model for Local Features

The local shape of a feature is approximated as being linear, planer or spherical. Examples of these from medical images are: blood vessels and nerve fibres, bone surfaces and interfaces between different 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 a line or a plane is windowed by a smooth function $w()$, then it can be approximated by a n-dimensional Gaussian function:

$$g(\vec{x}|\Theta) = A \exp(-(\vec{x} - \vec{\mu})^T C^{-1}(\vec{x} - \vec{\mu})/2) \tag{1}$$

parameterised by $\Theta = \{A, \vec{\mu}, C\}$ with amplitude A, centred on $\vec{\mu}$ and the covariance matrix $C = R^T C' R$, 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, to the x-axis.

To obtain maximum likelihood (ML) estimates of the parameters, the windowed image data, $f_w(\vec{x}) = w(\vec{x})f(\vec{x})$, are modelled as conditionally normal, given the local model $f_w(\vec{x}) \sim N(g(\vec{x}|\Theta), \sigma^2)$. To maximise the likelihood, we minimise the sum of squared residuals between the windowed data in an image block of size B^n and the model:

$$\chi^2 = \frac{1}{B^n} \sum_{\vec{x}}^{B^n} (g(\vec{x}|\Theta) - f_w(\vec{x}))^2 \tag{2}$$

From an initial estimate of $\Theta_0 = \{A_0, \vec{\mu}_0, C_0\}$ at iteration $t = 0$, we calculate the sample estimates for Θ weighted

by the inner product of f_w and g using the iterative scheme (dropping the position subscript $\vec{x}$ for clarity):

$$A_{t+1} = \sum_{\vec{x}} f_w g(\Theta_t) / \sum_{\vec{x}} g(\Theta_t) g(\Theta_t) \tag{3}$$

$$\vec{\mu}_{t+1} = \sum_{\vec{x}} \vec{x} f_w g(\Theta_t) / \sum_{\vec{x}} f_w g(\Theta_t) \tag{4}$$

$$C_{t+1} = 2 \sum_{\vec{x}} (\vec{x} - \vec{\mu}_t)(\vec{x} - \vec{\mu}_t)^T f_w g(\Theta_t) / \sum_{\vec{x}} f_w g(\Theta_t) \tag{5}$$

The initial estimate Θ_0 is obtained by using the multiresolution Fourier Transform (MFT) [7]. The windowed Fourier transform of $f_w(\vec{x}) \leftrightarrow \hat{f}_w(\vec{u})$ is also a Gaussian with the spectral energy distribution dependent on the type of feature. The principal components of the moment of inertia tensor I of the spectral energy:

$$I = \frac{1}{B^n} \sum_{\vec{u}} \vec{u}\vec{u}^T |\hat{f}_w(\vec{u})|^2 \tag{6}$$

where B is the block size, gives n eigenvalues $\lambda_1 \geq .. \geq \lambda_n$, which are inversely related to the covariance, C. The spatial orientation of a feature will depend on how the results of the PCA analysis of I is interpreted: for linear features, the eigenvector associated with the smallest eigenvalue is the direction of the line; for planes the eigenvector associated with the largest eigenvalue is the normal to the plane. A measure of the local linearity, planarity or isotropy of the feature can be calculated by taking ratios of the eigenvalues [2], table 1. Together with χ^2, these indices can be used to classify a block at any spatial size. Note that $c_p + c_l + c_s = 1$.

Interpretation	Eigenvalues	Index
planes	$\lambda_1 \gg \lambda_2 \approx \lambda_3$	$c_p = \frac{\lambda_1 - \lambda_2}{\lambda_1} \to 1$
lines	$\lambda_1 \approx \lambda_2 \gg \lambda_3$	$c_l = \frac{\lambda_2 - \lambda_3}{\lambda_1} \to 1$
spheres	$\lambda_1 \approx \lambda_2 \approx \lambda_3$	$c_s = \frac{\lambda_3}{\lambda_1} \to 1$

Table 1. Measures of local feature linearity/planarity/isotropy used for feature classification from a PCA of co-variance matrix C.

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

$$\phi'(\vec{u}) = arg(\hat{f}'_w(\vec{u})) = -\vec{\mu}.\vec{u} \tag{7}$$

The feature centroid $\vec{\mu}$ can therefore be estimated by taking average pairwise correlations between neighbouring coefficients along each of the n axes

$$\mu_{0i} = \frac{1}{2\pi B^n} \sum_{\vec{u}} \hat{f}_w(u_i - 1)\hat{f}_w(u_i)^* \tag{8}$$

If $f_w(\vec{x})$ is locally Gaussian and noise free, then Θ_0 will be close to Θ_{ML}. In general, the iterative estimation greatly improves the MFT estimate, converging rapidly (5-10 iterations) to a stable solution. Furthermore, the resulting χ^2 provides a useful measure of goodness of fit of the model to the data.

3   Towards a Global Object Representation

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 because multiple features of multiple types may be present. Therefore, the data will be better explained at the *next* smaller window size. 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. If the estimated feature has an acceptable χ^2 then successive window sizes are compared:

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. The parent is favoured for reasons of greater robustness of the estimate at the larger (parent) window size. By rejecting the hypothesis in (2), degenerate cases can be handled and allowing the hypothesis testing to be repeated at the next finer scale.

4   Experiments

Figure 1 illustrates results of the ML model estimation on part of a 2D retinal angiographic image size 404×404 pixels (fig. 1(a)). In figure 1(b)-(c), data reconstructions are shown of the 2D Gaussians models in each block (at overlapping window sizes of $B = 64$ and $B = 16$) after the iterative ML estimation. The orientation, position and thickness of the vessels (the standard deviation of the model orthogonal to the feature) are accurately modelled at both large and small block sizes. Because the model cannot describe the presence of multiple vessels within the window (e.g. bifurcations), at lower spatial resolutions low-amplitude, isotropic Gaussians are locally the 'best' description of these regions. These blocks can be identified by their higher residual errors, χ^2 (equation (2)).
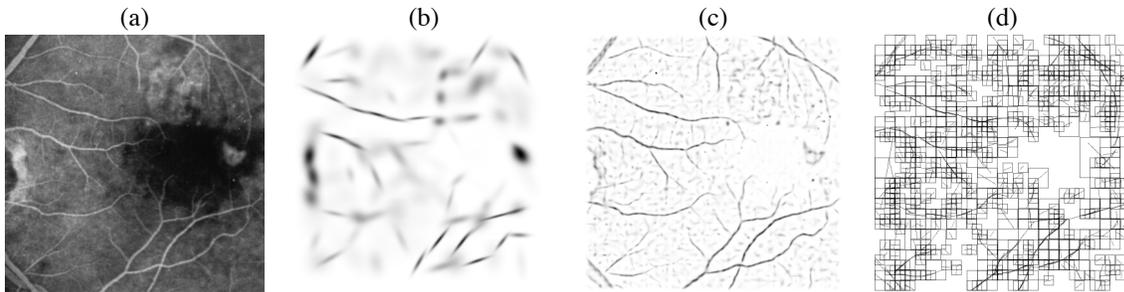
(a)   (b)   (c)   (d)

Figure 1. (a) 2D retinal image (flourescein angiogram) size 404×404 pixels. Reconstruction of data from ML model parameters estimates Θ_{ML} for block sizes (b) 64 and, (c) 16. (d) Multiresolution scale selection based on thresholding goodness of fit between model and data across levels with overlapping block sizes $64, 32, 16$ and 8.

Results from a 3D implementation of the ML model estimator are illustrated in figure 2. The local structure estimator was run on the *speed* image of part of a phase contrast MRA with cerebral blood vessels size $88 \times 58 \times 44$ voxels (fig. 2(a)). The model reconstructions at two window sizes, $B = 32$ and $B = 8$ are shown in (b) and (c) respectively. The major vessels are captured at the larger window size while the finer vessels can be seen in figure 2(c). At both scales, as with the 2D example, the vessel diameters are correctly estimated.
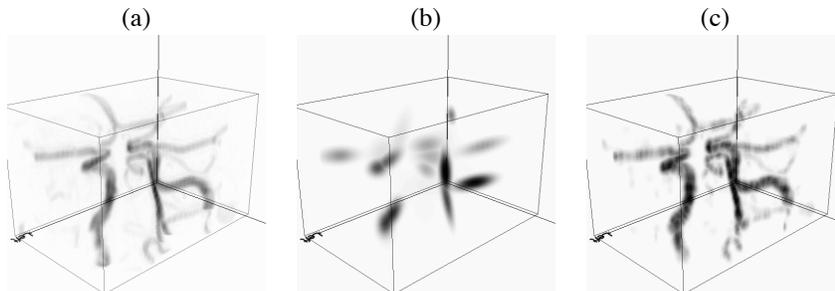
(a)   (b)   (c)

Figure 2. (a) Maximum intensity projection of part of a MR angiogram depicting cerebral blood vessels size $88 \times 58 \times 44$ voxels. Data model based on block-by-block summation of local Gaussian models Θ_{ML} at window sizes (b) $B = 32$ and, (c) $B = 8$. The Gaussian amplitude has been made proportional to the goodness of fit between the data and the model.

To demonstrate the utility of the use multiple window sizes, we generated a multi-level data reconstruction after a simple top-down scale selection based on the normalised block residual errors, figure 1(d). In this reconstruction,

4.4% of the parameters sets were used, Θ_{ML}, from a total of 13596 across 4 block sizes. This reconstruction captures both large and small local structures in the data.

A synthetic 3D image, size 64×64×64 (figure 3(a)), was used to test the local feature-type discrimination indices $\{c_p, c_l, c_s\}$ (table 1). Results of feature estimates with overlapping windows of size $B = 16$ interpreted as planar patches (drawn as circles of radius $B/2$) are shown in figure 3(b). Similarly, in figures 3(c), the estimates are classified as lines. 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 (c), the feature estimates have greater validity as lines using c_l along the linear features in the test data (the two cylinders: one intersecting the sphere and the vertical cylinder on the right hand side). In figure 3(d), the most likely feature in any block is depicted comparing c_p with c_l. As expected, the surface of the sphere is represented by planar patches, while the two cylinders are best modelled as lines. Such a classification process could also operate over a range of window sizes.
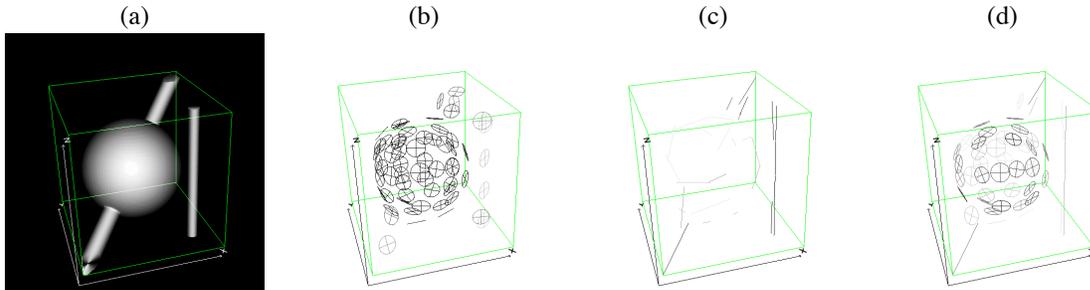
(a) (b) (c) (d)

Figure 3. (a) Synthetic 3D test image containing a sphere and two cylinders with different diameters. (b) Each block interpreted as a plane for window size $B = 16$. (c) Each block is interpreted as a line for window size $B = 16$. (d) Classification of blocks into planar over linear if $c_p > c_l$ (feature intensity proportional classification index) at window size $B = 16$.

5   Conclusions

A maximum likelihood feature estimation method that models local image structure by a Gaussian intensity model was described. It is fast and robust at finding the position, orientation and size of linear and planar features in 2D and 3D data. Additionally, the residual error from the fit can be used to accept or reject the local estimate. By diagonalising the covariance of the Gaussian model, metrics can be applied on the eigenvalues to determine the 'best' structural interpretation of the feature as being locally planar or linear. The model estimation operates over a range of window sizes and we have proposed a scale-selection strategy that tries to find the largest window size in which the single feature hypothesis fits. The illustrative results of the image modelling show it to work well for linear structures like blood vessels in 2D and 3D images. We are investigating generalised Bayesian inference techniques that can explicitly model prior knowledge of the geometry of local structures [8].

References

1. A. F. Frangi, W. J. Niessen, R. M. Hoogeveen et al. "Model-based Quantitation of 3D Magnetic Resonance Angiographic Image." *IEEE Trans. Medical Imaging* **18(10)**, pp. 946–956, 1999.
2. C.-F. Westin, A. Bhalerao, H. Knutsson et al. "Using Local 3D Structure for Segmentation of Bone from Computer Tomography Images." In *Proc. of Computer Vision and Pattern Recognition '97*. Puerto Rico, 1997.
3. G. Gerig, T. Koller, G. Székely et al. "Symbolic description of 3-D structures applied to cerebral vessel tree obtained from MR angiography volume data." In H. H. Barrett & A. F. Gmitro (editors), *Information Processing in Medical Imaging IPMI'93, Lecture Notes in Computer Science*, volume 687, pp. 94–111. 1993.
4. M. E. Martínez-Pérez, A. D. Hughes, A. V. Stanton et al. "Scale-Space Analysis for the Characterisation of Retinal Blood Vessel." In *Proceedings on Medical Image Understanding and Analysis (MIUA)*, pp. 57–60. Oxford, 1999.
5. L. M. Lorigo, O. Faugeras, W. E. L. Grimson et al. "Codimension-Two Geodesic Active Contours for MRA Segmentation." In *Proc. of Intl. Conf. on Information Processing in Medical Imaging*. 1999.
6. D. L. Wilson & J. A. Noble. "An adaptive segmentation algorithm for extracting arteries and aneurysms from time-of-flight MRA data." *IEEE Trans. Medical Imaging* **18(10)**, pp. 938–945, 1999.
7. R. Wilson, A. D. Calway & 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)**, pp. 674–690, 1992.
8. W. R. Gilks, S. Richardson & D. J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, 1996.