

# Binary image restoration at subpixel resolution

D Hitchcock and C A Glasbey  
Biomathematics and Statistics Scotland  
JCMB, King's Buildings, Edinburgh, EH9 3JZ, Scotland

*Abstract* – Recovery of object boundaries to subpixel accuracy is possible from digital images if sufficient assumptions can be made. Subpixel restoration is of most relevance if an object is at the limit of resolution of a sensor, either being only a few pixels in size or having fine structure. A method is proposed for recovering a binary image from multi-level pixel values, which are modelled by a convolution of the true scene with a blurring kernel, together with additive noise. Small objects, such as immunogold-labelled particles in electron micrographs, are parameterised using Fourier descriptors, and a network of splines is used to parameterise the filament structure of fungal hyphae.

*Key words* – Deconvolution, Digital image analysis, Electron micrograph, Fourier descriptor, Fungal hyphae, Normal integral, Splines, Stokes' theorem.

# 1 Introduction

Digital images are imperfect representations of scenes in the real world, because data are affected by blur, subjected to noise and rounding error, and sampled only at a lattice of points (pixels). These factors lead to many problems in digital image analysis. For example, it is very difficult to study objects which are near the resolution limit of an imaging sensor. A fundamental case is where the true scene is binary, appropriate for many simple scenes such as back-illuminated opaque objects (i.e. silhouettes). To illustrate, Fig 1 shows a digitized electron micrograph of an immunogold-labelled section of tulip virus (Roberts, 1994), and Fig 2 shows a digitized photograph of part of a fungal mycelium grown on cellophane (Hitchcock *et al*, 1996). Figs 1 and 2 are blurred and noisy versions of binary scenes, consisting of black blobs (gold particles) and black fibres (fungal hyphae) on grey backgrounds. We wish to determine the shapes and sizes of the gold particles and measure the widths and orientations of the hyphae. In both cases it would clearly be beneficial to recover the binary scenes to subpixel resolution.

The reversal of blur is a numerically unstable procedure which can be made tractable only by including assumptions about the form of the true scene. For example, a linear deconvolution such as the Wiener filter discards the high-frequency components of images, and in maximum entropy restoration the constraint is used that pixels cannot take negative values. (See, for example, Rosenfeld and Kak, 1982, chapter 7.) The vast majority of image processing techniques operate at pixel rather than sub-pixel level. However, several authors have considered sub-pixel restoration. Boulton and Wolberg (1993) proposed a method which depends on an image having a smoothly-varying pixel intensity. If the main features of interest are smooth edges which separate regions of constant intensity, then one approach is to locate an edge to pixel accuracy and then fit a smooth curve (Koplowitz and Raj, 1987; Sriraman *et al*, 1989). Sub-pixel edge detection can also be implemented using a filter (Huertas and Medioni, 1986; Oakley and Shann, 1991) or moments (Tabatabai and Mitchell, 1984; Lyvers *et al*, 1989). Application areas include measuring blood vessel diameters in angiography (Sandor and Spears, 1985), locating industrial parts (Young, 1987), estimation of vegetation cover by remote sensing (Jasinski and Eagleson, 1990), deconvolution of astronomical images (Weir and Djorgovski, 1991) and alignment in lithography (Gatherer and Meng, 1992). Several authors have considered the recovery of binary scenes. Glasbey (1996) showed that inference is very uncertain if data are also binary. However, if pixels are multi-level, more progress can be made. Havelock (1989) found that grey-level quantisation led to greater loss of information than did spatial quantisation. Further, Kiryati and Bruckstein (1991) showed that blur improves the resolving power of digitised images. Jennison and Jubb (1991) and Gavin and Jennison (1997) considered the restoration to subpixel accuracy of binary scenes, using a Bayesian approach and Markov random fields to regularise the problem.

The restoration of binary scenes can be regarded as a statistical problem. In §2 we identify an appropriate statistical model for digital image data, and propose an inferential

procedure. Then, in §3 we apply this methodology to restore the images of blob-like and filamentous objects in Figs 1 and 2. Finally, in §4 we discuss the results.

## 2 Method

The model for image formation used in this paper is as follows. The true binary scene has intensity  $f(x, y)$  at location  $(x, y)$ :

$$f(x, y) = \begin{cases} f_1 & \text{if } (x, y) \in C \\ f_0 & \text{otherwise,} \end{cases}$$

for some set  $C$ , and foreground and background levels of  $f_1$  and  $f_0$  (which may or may not be known). We assume prior knowledge of the shape of  $C$ , the boundary of which may be specified by geometric shapes, Fourier descriptors, splines or other functions, depending on the particular application. The digital image has intensity values,  $z$ , only at integer lattice positions  $(i, j)$  for  $i, j = 1, \dots, n$ , where

$$z_{ij} = \mu_{ij} + e_{ij} \quad \text{and} \quad \mu_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) g(i - x, j - y) dx dy .$$

Here  $\mu_{ij}$  is the expected pixel value at  $(i, j)$ ,  $g$  is a blurring kernel (which integrates to unity) and  $e_{ij}$  is an independently distributed error term with mean zero, variance  $\sigma^2$ . We will assume that  $g$  is of known functional form such as a bivariate Gaussian or Cauchy density, specified by a dispersion parameter  $\tau^2$ . Parameters  $\tau^2$  and  $\sigma^2$  may or may not be known in particular applications. Implicitly, we assume that the signal-to-noise ratio in the data is relatively high, because otherwise sub-pixel restoration is not possible.

We wish to infer  $f$ , given an observed pixel array,  $z$ . We propose to do this by minimising the sum of squares

$$\sum_{i=1}^n \sum_{j=1}^n (z_{ij} - \mu_{ij})^2$$

with respect to parameters specifying  $f$  and  $g$ . This requires a 2-dimensional integral for each pixel position  $(i, j)$ ,

$$\mu_{ij} = f_0 + (f_1 - f_0) \int \int_C g(i - x, j - y) dx dy ,$$

which can be simplified to a 1-dimensional integral for certain functional forms of  $g$ , provided the boundary of  $C$  is piecewise differentiable (see Appendix). Let  $\{(x_t, y_t) : 0 \leq t \leq T\}$  describe the boundary of simply-connected, hole-free set  $C$  in an anticlockwise sense, where  $(x_0, y_0) = (x_T, y_T)$ . Then

$$\mu_{ij} = f_0 + (f_1 - f_0) \int_0^T \frac{h(i - x_t, j - y_t)}{2\pi\{(i - x_t)^2 + (j - y_t)^2\}} \left( - (j - y_t) \frac{dx_t}{dt} + (i - x_t) \frac{dy_t}{dt} \right) dt ,$$

where  $h$  is a function of  $g$ , defined in the Appendix both generally and for Gaussian and Cauchy kernels. We used a standard algorithm for the 1-dimensional integration (routine D01AJF in NAG, 1993) and encountered no problems. In both applications in §3, we used the Nelder-Mead simplex algorithm (routine E04CCF in NAG, 1993) to minimise the sum of squares iteratively. There is the potential for problems with local minima, although we were not aware of encountering any. They can be guarded against by using multiple starting values for the parameters. The computing time in our applications was not great by image analysis standards, because we were only ever considering small regions of images.

There are many ways to parameterise the boundary of  $C$ , and these also serve to regularise the restoration problem, which is otherwise ill-conditioned. In earlier work using a subpixel grid and no further constraints, we found that restoration was only achievable with unrealistically low noise levels (Hitchcock and Glasbey, 1995). Rohlf and Archie (1984) and Mou and Stoermer (1992) reviewed several alternative Fourier descriptors to parameterise the boundaries of mosquito wings and diatoms, respectively. Other options include the use of polygons (Pavlidis, 1977), splines (Hill and Taylor, 1992), conic sections (Bookstein, 1978) and principal curves (Banfield and Raftery, 1992). It is also possible to parameterise other features of  $C$  than its boundary. For example, Blum (1973) used the medial axis transform, a form of skeleton, to describe the shapes of biological objects. A discussion of a number of other morphological approaches to defining and constructing the skeleton in a digital image may be found in Serra (1982, chapter 11). We will return to this method in §3.2.

## 3 Applications

### 3.1 Blob-like objects

If  $C$  is a small object with an approximately circular boundary then a convenient way to parameterise it is as a star-shaped domain, that is, the boundary:

$$\{(x_t, y_t) : 0 \leq t < 2\pi\}$$

is expressible in the form:

$$x_t = x_0 + r_t \cos(t), \quad y_t = y_0 + r_t \sin(t),$$

about an origin  $(x_0, y_0)$ , where  $r_t$  is a positive, periodic function. One advantage of this model is that it automatically ensures the boundary will not self-intersect. However, because the radius function has to be single valued, the model is inappropriate for most non-convex objects. If the boundary is smooth, the radius function can be approximated

by a Fourier descriptor,

$$r_t = \beta_0 + \beta_2 \sin(2t + \theta_2) + \beta_3 \sin(3t + \theta_3) + \dots + \beta_N \sin(Nt + \theta_N) ,$$

with non-negative parameters  $\beta_0, \beta_2, \beta_3, \dots, \beta_N$  and parameters  $\theta_2, \theta_3, \dots, \theta_N$  constrained to the interval  $[0, 2\pi]$ . This parameterisation is rotationally invariant and any star-shaped domain can be approximated using a sufficiently large value of  $N$ . The first-order harmonic,  $\beta_1 \sin(t + \theta_1)$ , has been omitted because it is almost confounded with  $(x_0, y_0)$ : it has approximately the same effect as a change in origin.

Fifty of the gold particles in Fig 1 with no close neighbours were selected for modelling. Initially, a circular model with a Gaussian blurring kernel was fitted to log-transformed pixels in a  $15 \times 15$  square window centred on each particle, such as those within the white frame in Fig 1. The Nelder-Mead simplex algorithm (NAG, 1993) was used in each case to minimise the sum of squares with respect to  $f_0, f_1, \tau^2, x_0, y_0$  and  $\beta_0$ . Then,  $\sigma^2$  was estimated from the residual sum of squares. For illustration, Fig 3a shows, in white, the circular boundary fitted to the region within the white frame in Fig 1, together with log-transformed pixel values. The need for a log-transformation may have arisen from non-linear responses of the photographic film and the digitiser which produced Fig 1. the transformation was used to standardise the residuals in Fig 3c to approximately equal variance. A Gaussian function was found to be adequate to represent the blurring kernel, as can be seen in Fig 4, which shows the predicted and observed log-transformed pixel values for the central row of pixels in Fig 3a.

The model was refitted with a second-order harmonic included. On average, over the 50 particles, the reduction in sum of squares was 29%. Further reductions of 10%, 3.3%, 2.5% and 1.5% were achieved by including third-, fourth-, fifth- and sixth-order harmonics, respectively. These improvements are all highly significant statistically, on the basis of F-tests, under the assumption that the errors are independently distributed. However, there is evidently some autocorrelation in the residuals. Table 1 shows autocorrelations in residuals from a sixth-order model fitted to the data in Fig 3, for a range of shifts in row and column positions. In general, only autocorrelations at shifts of 1 or 2 rows and columns are significantly different from zero.

*Table 1: Percentage autocorrelations of residuals from a sixth-order model fitted to the data in Fig 3.*

column shift	-3	-2	-1	0	1	2	3
row shift							
0	*	*	*	100	50	15	11
1	4	-0	16	53	42	24	12
2	-4	-5	1	23	36	37	24
3	-4	2	1	12	22	36	29

Variances of parameter estimators can be adjusted to take account of error covariances. Consider a single parameter,  $\alpha$  say, whose estimator is linearly dependent on perturbations in  $z$ , by a linearising approximation. Therefore,

$$\tilde{\alpha} \approx \hat{\alpha} + \sum_i \sum_j w_{ij} (\tilde{z}_{ij} - z_{ij}) ,$$

where  $\hat{\alpha}$  is the estimated value of  $\alpha$ , given  $z$ , and  $\tilde{\alpha}$  is the revised estimate if  $z$  is changed to  $\tilde{z}$ . The coefficients,  $w_{ij}$ , can be obtained numerically, and the variance of  $\hat{\alpha}$  can be approximated by

$$\sum_i \sum_j \sum_k \sum_l w_{ij} w_{kl} \text{cov}(e_{ij}, e_{kl}) .$$

Using this formula, and assuming that covariances are zero of  $|i - k| > 3$  or  $|j - l| > 3$ , the adjusted standard errors of  $\beta_2, \beta_3, \dots, \beta_6$  in the sixth-order harmonic model fitted to the data in Fig 3 are given in Table 2. For comparison, also given are the least-squares standard errors, based on the assumed independence of errors.

*Table 2: Estimated parameters and standard errors for sixth-order harmonic model fitted to data in Fig 3. Least-squares-based standard errors, and those adjusted to take account of correlations in errors, are both given.*

	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$
estimate	0.118	0.102	0.073	0.048	0.055
least-squares s.e.	0.018	0.020	0.023	0.027	0.034
adjusted s.e.	0.029	0.027	0.025	0.028	0.035

It can be seen that least-squares-based standard errors of low-order harmonics, in particular, are under-estimated. For this particle, the coefficients  $\beta_5$  and  $\beta_6$  are not significantly different from zero. so a fourth-order model appears to be adequate. Fig 3a shows, in black, the estimated boundary, and Fig 3b shows the predicted pixel values from the fourth-order Fourier model. Figs 3c and 3d show residuals from the fit of the circle and fourth-order Fourier model. Values have been rescaled to cover the full grey-scale range from black to white. The improved fit of the non-circular model is evident.

Although our results show clearly that the gold particles in Fig 1 are not circular, the departures from circularity are only small. The average radius ( $\beta_0$ ) is 3.5 pixels, and the average values of  $\beta_2, \beta_3, \dots, \beta_6$  are 0.173, 0.098, 0.066, 0.063 and 0.067 respectively. No relationship between the phase shifts in different harmonics could be discerned, and therefore no further characterisation of the shape could be achieved. The shapes of the gold particles has some effect on the efficiency of labelling of antigenic sites on the virus section. Glasbey and Roberts (1997) identify a point process model for the spatial distribution of particles in Fig 1.

## 3.2 Fibre network

The fungal hyphae in Fig 2 were modelled as a network of interconnecting curves, each specified by a spline function together with width parameters. The network structure was obtained using standard image analysis techniques. First, the image was thresholded to separate the fibre from the background, with the threshold value selected from the histogram of pixel values. (See, for example, Glasbey and Horgan, 1995, section 3.1.) The precise value of the threshold is not critical provided the connectivity of the thresholded image is not affected. Fig 5a shows the thresholded image. Larger components have been displayed in grey and smaller components in black. We restrict attention to the larger components, which represent the hyphae. A skeleton procedure was then used to reduce each fibre to the width of a single pixel (see, for example, Rosenfeld and Kak, 1982, chapter 11). Fig 5b shows the result. Nodes in the network, that is junctions and end pixels, can be identified from this network. Every connected component of pixels in the skeleton with three or more neighbours constitutes a junction and each pixel with only one neighbour is an end-point. Note that, this pre-processing is only used to identify initial parameter values, for estimation we make use of the original pixel values.

Splines were used to interpolate between nodes. Given two nodes,  $(x_0, y_0)$  and  $(x_1, y_1)$ , with slopes of  $m_0$  and  $m_1$  respectively, the interpolating curve for  $0 \leq t \leq 1$  is given by:

$$\begin{aligned} x_t &= [x_1 t + x_0(1 - t)] + d_y \{ \alpha t(1 - t) + \beta t(1 - t)(1 - 2t) \} , \\ y_t &= [y_1 t + y_0(1 - t)] - d_x \{ \alpha t(1 - t) + \beta t(1 - t)(1 - 2t) \} , \end{aligned}$$

where  $d_x = (x_1 - x_0)$ ,  $d_y = (y_1 - y_0)$ ,

$$\alpha = \frac{(m_1 - m_0)(d_x^2 + d_y^2)}{2(d_x + m_0 d_y)(d_x + m_1 d_y)} \quad \text{and} \quad \beta = \frac{(d_y - m_1 d_x)}{2(d_x + m_1 d_y)} + \frac{(d_y - m_0 d_x)}{2(d_x + m_0 d_y)} .$$

The terms in square brackets are a linear interpolant between nodes, which is modulated by two basis functions,  $t(1 - t)$  and  $t(1 - t)(1 - 2t)$ , to ensure that slopes match at both nodes, and yields a rotationally-invariant solution. The curves are illustrated in Fig 6a. For longer sections of fibre, intermediate nodes were introduced. They were constrained to lie midway between other nodes, as shown in Fig 6b, in order to avoid redundant parameters. A parameter specified the width of each fibre at each node, and the width at any other point on a fibre was found by linear interpolation. The boundary of the fibres was constructed by tracing out a distance of half the width orthogonal to each spline curve. The completion of the boundary of the fibre is shown in Fig 6c. Each fibre was assumed to end in a semi-circle.

Parameters  $f_0$  and  $f_1$  were estimated from the background intensity and a thick portion of fibre, respectively. The blurring kernel, a combination of the effects of photographing the hyphae and using a digitising scanner, was estimated using the techniques in Glasbey *et al* (1994). A Gaussian kernel with  $\tau^2 = 0.8$  was identified. The remaining parameters

in the model, that is the node positions, fibre orientations and widths, were estimated iteratively from starting values extracted from the skeleton. A greedy algorithm was used in which parameters for each node were optimised in turn using the Nelder-Mead simplex algorithm. Nodes were revisited until convergence of parameter values. Because changes in parameter values at a node only affect local pixel values, it was possible to simplify the computations. Pixels were divided among the nodes by initially finding the nearest point on the skeleton and then the nearest node. The boundary of the fibre was described up to adjacent nodes in an anticlockwise direction and the method of §2 used to find the pixel values at the local sites.

To demonstrate the effectiveness of the method, Fig 7 shows a simulated fibre network. In Fig 7b the true boundary is superimposed in black and the estimated boundary in white. The positions of intermediate nodes is also shown. Agreement is good, but is less than perfect at the ends of fibres because of the dependency on a few pixel values. Also, the junction has not been recovered precisely. If noise levels are reduced, the fit becomes almost perfect. However if only one intermediate node is used the parameters cannot exactly match the fibre. Other simulated examples showed similar patterns.

Fig 8a shows the restoration of the fibre within the white frame in Fig 2. The fit is not perfect because an outline of the fibre is visible in the residuals display in Fig 8b, but almost all the variation has been accounted for. We found that problems lay, not so much with large amounts of random noise, but kinks and buds on the fibre which meant that the model did not fit exactly. For this reason we found it best to use reasonably large intervals between the intermediate nodes.

## 4 Discussion

When measurements are needed to greater precision than the inter-pixel distance, it is sometimes possible to increase the magnification of an imaging sensor. However, this can be time consuming, and greatly increases the data handling problems. Subpixel resolution is achievable by the human eye-brain and in this paper we have explored some aspects of what is achievable digitally. Because digital images are corrupted by blur and noise, recovery even at the pixel level is impossible without some assumptions on the form of the image. The binary assumption, that the true image takes only two values, is very strong in the restrictions it places on permissible images and makes problems highly non-linear.

Almost all image processing techniques (filters, segmentation, mathematical morphology, feature extraction) operate at pixel rather than sub-pixel level, and therefore could not have produced the results given in §3. Restoration of binary scenes can be formulated as statistical problem. In §2 we identified an appropriate statistical model and inferential procedure, an approach which is complementary to other sub-pixel methods. The Fourier

descriptor method works well for small blob-like objects. For larger objects, a global parameterisation is inefficient and the optimisation becomes more difficult. However, it is then possible to choose parameters which have only a local effect and optimise subsets of the parameters independently as demonstrated with the fungal hyphae. Long sections of smooth boundary can be recovered very accurately but there is a limit to the amount of boundary detail which can be recovered. Typically a parameterisation needs to be chosen which does not allow a boundary to change direction significantly over the width of a pixel. As noise levels increase, it is important that the number of data points is large compared to the number of parameters. This may force a recovered boundary to be smooth, but that is normally more appropriate than to risk predicting imaginary boundary detail.

## Acknowledgements

We thank Jim Kay for encouraging this work, and Ian Roberts and Karl Ritz of the Scottish Crop Research Institute for permission to use the virus and fungal images. The work was supported by funding from the Scottish Office Agriculture, Environment and Fisheries Department.

## References

- Banfield, J.D. and Raftery, A.E. (1992). Ice-floe identification in satellite images using mathematical morphology and clustering about principal curves. *Journal of the American Statistical Association*, **87**, 7-16.
- Blum, H. (1973). Biological shape and visual science. *Journal of Theoretical Biology*, **38**, 205-287.
- Bookstein, F.L. (1978). *The Measurement of Biological Shape and Shape Change*. Springer-Verlag, Berlin.
- Boult, T.E. and Wolberg, G. (1993). Local image-reconstruction and subpixel restoration algorithms. *CVGIP: Graphical Models and Image Processing* **55**, 63-77.
- Gatherer, A. and Meng, T. (1992). Robust subpixel alignment in lithography. *Journal of Vacuum Science and Technology*, **10**, 2662-2666.
- Gavin, J. and Jennison, C. (1997). A subpixel image restoration algorithm. *Journal of Computational and Graphical Statistics*, **6**, (in press).
- Gilliland, D.C. (1962). Integral of the bivariate normal distribution over an offset circle.

- Journal of the American Statistical Association* **57**, 758-768.
- Glasbey, C.A. (1996). Inference on binary images from binary data. *Journal of Applied Probability*, **28**, 627-640.
- Glasbey, C.A. and Horgan, G.W. (1995). *Image Analysis for the Biological Sciences*. Wiley, Chichester.
- Glasbey, C.A., Horgan, G.W. and Hitchcock, D. (1994). A note on the greyscale response and sampling properties of a desktop scanner. *Pattern Recognition Letters*, **15**, 705-711.
- Glasbey, C.A. and Roberts, I.M. (1997). Statistical analysis of the distribution of gold particles over antigen sites after immunogold labelling. *Journal of Microscopy*, (in press).
- Havelock, D.I. (1989). Geometric precision in noise-free digital images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**, 1065-1075.
- Hill, A. and Taylor, C.J. (1992). Model-based image interpretation using genetic algorithms. *Image and Vision Computing*, **10**, 295-300.
- Hitchcock, D. and Glasbey, C.A. (1995). Binary image restoration at sub-pixel resolution from multi-level data. In *Complex Stochastic Systems and Engineering* (D.M. Titterton, ed.). Clarendon Press: Oxford 177.
- Hitchcock, D., Glasbey, C.A. and Ritz, K. (1996). Image analysis of space-filling by networks: application to a fungal mycelium. *Biotechnology Techniques*, **10**, 205-210.
- Huertas, A. and Medioni, G. (1986) Detection of intensity changes with subpixel accuracy using Laplacian-Gaussian masks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8**, 651-664.
- Jasinski, M.F. and Eagleson, P.S. (1990). Estimation of subpixel vegetation cover using red and infrared scattergrams. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 253-267.
- Jeffreys, H. and Jeffreys, B. (1953) *Methods of Mathematical Physics*. Cambridge University Press, Cambridge.
- Jennison, C. and Jubb, M. (1991) Aggregation and refinement in binary image restoration, *Spatial Statistics and Imaging, IMS Lecture Notes, Hayward*, 150-162.
- Kiryati, N. and Bruckstein, A.M. (1991). Gray levels can improve the performance of binary image digitizers. *CVGIP: Graphical Models and Image Processing* **53**, 31-39.
- Koplowitz, J. and Raj, A.P.S. (1987). A robust filtering algorithm for subpixel reconstruction of chain coded line drawings. *IEEE Transactions on Pattern Analysis and Machine*

*Intelligence* **9**, 451-457.

Lyvers, E.P., Mitchell, O.R., Akey, M.L. and Reeves, A.P. (1989). Subpixel measurements using a moment based edge operator, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 1293-1309.

Mou, D. and Stoermer, E.F. (1992). Separating *Tabellaria* (Bacillariophyceae) shape groups based on Fourier descriptors. *Journal of Phycology*, **28**, 386-395.

Numerical Algorithms Group (1993). *Library Manual Mark 16*. NAG Central Office, 256 Banbury Road, Oxford OX2 7DE, UK.

Oakley, J.P. and Shann, R.T. (1991). Efficient method for finding the position of object boundaries to sub-pixel precision, *Image and Vision Computing*, **9**, 262-272.

Patel, J.K. and Read, C.B. (1982). *Handbook of the Normal Distribution*. Marcel Dekker, New York.

Pavlidis, T. (1977). Polygonal approximations by Newton's method. *IEEE Transactions on Computers*, **26**, 800-807.

Roberts, I.M. (1994). Factors effecting the efficiency of immunogold labelling of plant virus antigens in thin sections. *Journal of Virological Methods* **50**, 155-166.

Rohlf, F.J. and Archie, J.W. (1984). A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Systematic Zoology*, **33**, 302-317.

Rosenfeld, A. and Kak, A.C. (1982). *Digital Picture Processing (2nd edition)*. Academic Press, San Diego.

Roysam, B., Maffitt, D.R., Miller, M.I., Saffitz, J.E. and Thomas, L.J. (1992). A personal computer based implementation of the maximum-likelihood method of analysis of electron microscope autoradiographs. *Microscopy Research and Technique*, **20**, 73-86.

Sandor, T. and Spears, J.R. (1985). Statistical considerations on the precision of assessing blood vessel diameter in cine coronary angiography, *Computers and Biomedical Research*, **18**, 531-543.

Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, London.

Sriraman, R., Koplowitz, J. and Mohan, S. (1989). Tree searched chain coding for subpixel reconstruction of planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**, 95-103.

Tabatabai, A.J. and Mitchell, O.R. (1984). Edge location to subpixel values in digital imagery, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 188-201.

Weir, N. and Djorgovski, S. (1991). A subpixel deconvolution method for astronomical images. *Fundamental Theories of Physics*, **43**, 275-283.

Young, R.A. (1987). Locating industrial parts with subpixel accuracies. *Proceedings of the Society of Photo-optical Instrumentation Engineers*, **728**, 2-9.

## Appendix: Proof of integration result

Define three 2-dimensional functions from the blur function,  $g$ , as follows. Let

$$h(x, y) = 2\pi \int_0^{\sqrt{x^2+y^2}} r g\left(\frac{xr}{\sqrt{x^2+y^2}}, \frac{yr}{\sqrt{x^2+y^2}}\right) dr ,$$

$$u(x, y) = \frac{-y h(x, y)}{2\pi(x^2 + y^2)} \quad \text{and} \quad v(x, y) = \frac{x h(x, y)}{2\pi(x^2 + y^2)} .$$

The 2-dimensional version of Stokes' Theorem (Jeffreys and Jeffreys, 1953, p. 195) states that for any smooth functions  $u(x, y)$  and  $v(x, y)$ :

$$\int \int_C \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) dx dy = \int_0^T \left( u(x_t, y_t) \frac{dx_t}{dt} + v(x_t, y_t) \frac{dy_t}{dt} \right) dt ,$$

where  $\{(x_t, y_t) : 0 \leq t \leq T\}$  defines the piecewise-differentiable closed boundary of region  $C$ , traversed in an anti-clockwise direction. It can be shown from the above definitions of  $u$  and  $v$ , after some algebra, that

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = g ,$$

and hence we have the required result that

$$\int \int_C g(x, y) dx dy = \int_0^T \frac{h(x_t, y_t)}{2\pi(x_t^2 + y_t^2)} \left( -y_t \frac{dx_t}{dt} + x_t \frac{dy_t}{dt} \right) dt .$$

Therefore, for certain functional forms of  $g$ , for which  $h$  is expressible in closed form, we have reduced a 2-dimensional integral to a 1-dimensional one. It will usually be necessary to evaluate this numerically, but provided the integrand is an analytic function, the error decreases exponentially with step size, because the integral is over a complete period of a periodic function. In particular, for integrating a Gaussian density over a circle, the method is competitive with the existing algorithm (Gilliland, 1962, and Patel and Read, 1982, pp 303).

For isotropic Gaussian and Cauchy kernels,  $h$  has an analytic form. In the Gaussian case:

$$g(x, y) = \frac{1}{2\pi\tau^2} \exp\left[\frac{-(x^2 + y^2)}{2\tau^2}\right] \quad \text{and} \quad h(x, y) = 1 - \exp\left[\frac{-(x^2 + y^2)}{2\tau^2}\right] ,$$

and for the bivariate Cauchy density:

$$g(x, y) = \frac{1}{2\pi\tau^2(1 + (x^2 + y^2)/\tau^2)^{\frac{3}{2}}} \quad \text{and} \quad h(x, y) = 1 - \frac{1}{\sqrt{1 + (x^2 + y^2)/\tau^2}} .$$

Roysam *et al* (1992) derived a similar expression for the Cauchy density, and obtained an analytic solution to the 1-dimensional integral for the special case of a piecewise linear boundary.

## Captions for figures

- Fig 1** Electron micrograph of section of pellet of purified tulip virus X which has been immunogold labelled with antiserum. (The square outlines the region shown in Fig 3.)
- Fig 2** A section of fungus mycelium *Trichoderma viride*, that it a network of hyphae from a single fungal organism, grown on cellophane-coated nutrient agar. (The rectangle outlines the region shown in Fig 8.)
- Fig 3** Fourier model fitted to region outlined by white square in Fig 1. **(a)**: log-transformed pixel values, together with a white curve showing the estimated circle and a black curve showing the estimated fourth-order Fourier model. **(b)**: predicted log-transformed pixel values from fourth-order Fourier model. **(c)**: residuals from fit of circle. **(d)**: residuals from fit of fourth-order Fourier model. Residuals in (c) and (d) have been identically rescaled to cover the full grey-scale range from black to white.
- Fig 4** Plot of log-transformed pixel values, and predicted values from the circular model with Gaussian blur, against column number, for the central row of pixels in Fig 3a.
- Fig 5** Extraction of the skeleton on the fungal hyphae in Fig 2. **(a)**: Thresholded image, with large components displayed in grey and small components in black. **(b)**: Skeleton of large components in (a).
- Fig 6** Stages in constructing a section of fibre network from parameter values. **(a)**: Crosses mark positions of end-points and junction points (2 parameters each). Arrows mark directions of fibres at the ends (1 parameter each). Curved lines are determined by position and gradient of their ends. **(b)**: Insertion of intermediate node: position is constrained to lie on fixed straight line (1 parameter). Direction is given by 1 parameter. **(c)**: Completion of outline by tracing boundary. Width is given at every end-point and at every intermediate node by parameter values, and defined at intermediate points by linear interpolation.
- Fig 7** Fit to simulated image of fibres. **(a)**: pixel values, and **(b)**: true outline shown in black and the estimated outline in white superimposed on pixel values. Gaussian blur has standard deviation 0.5, added noise has standard deviation 0.1.
- Fig 8** Model fitted to rectangular section in Fig 2. **(a)**: recovered fibre outline superposed on data values for part of mycelium in. Lines across the fibre show sites of intermediate nodes. **(b)**: residual display.

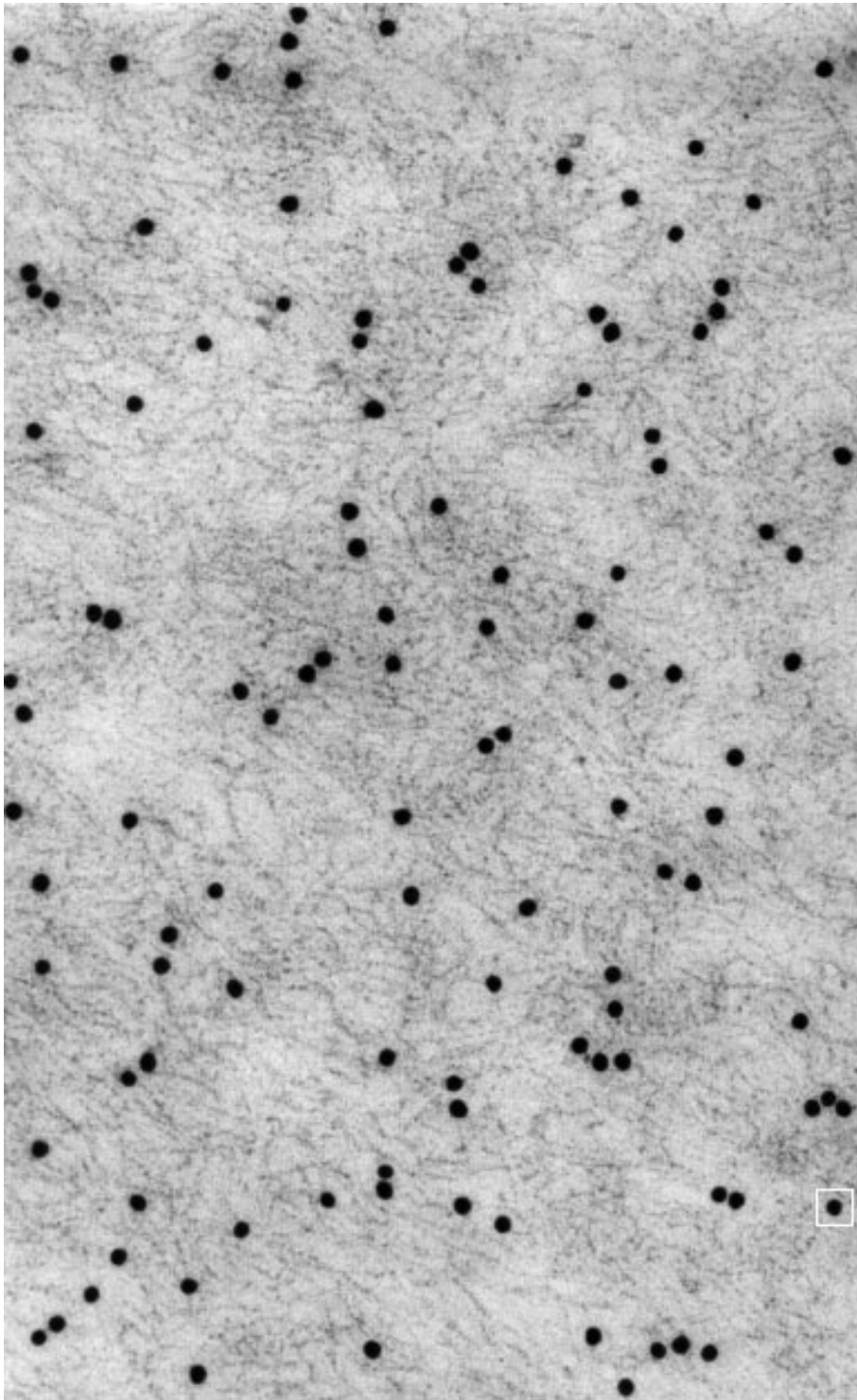


Figure 1:

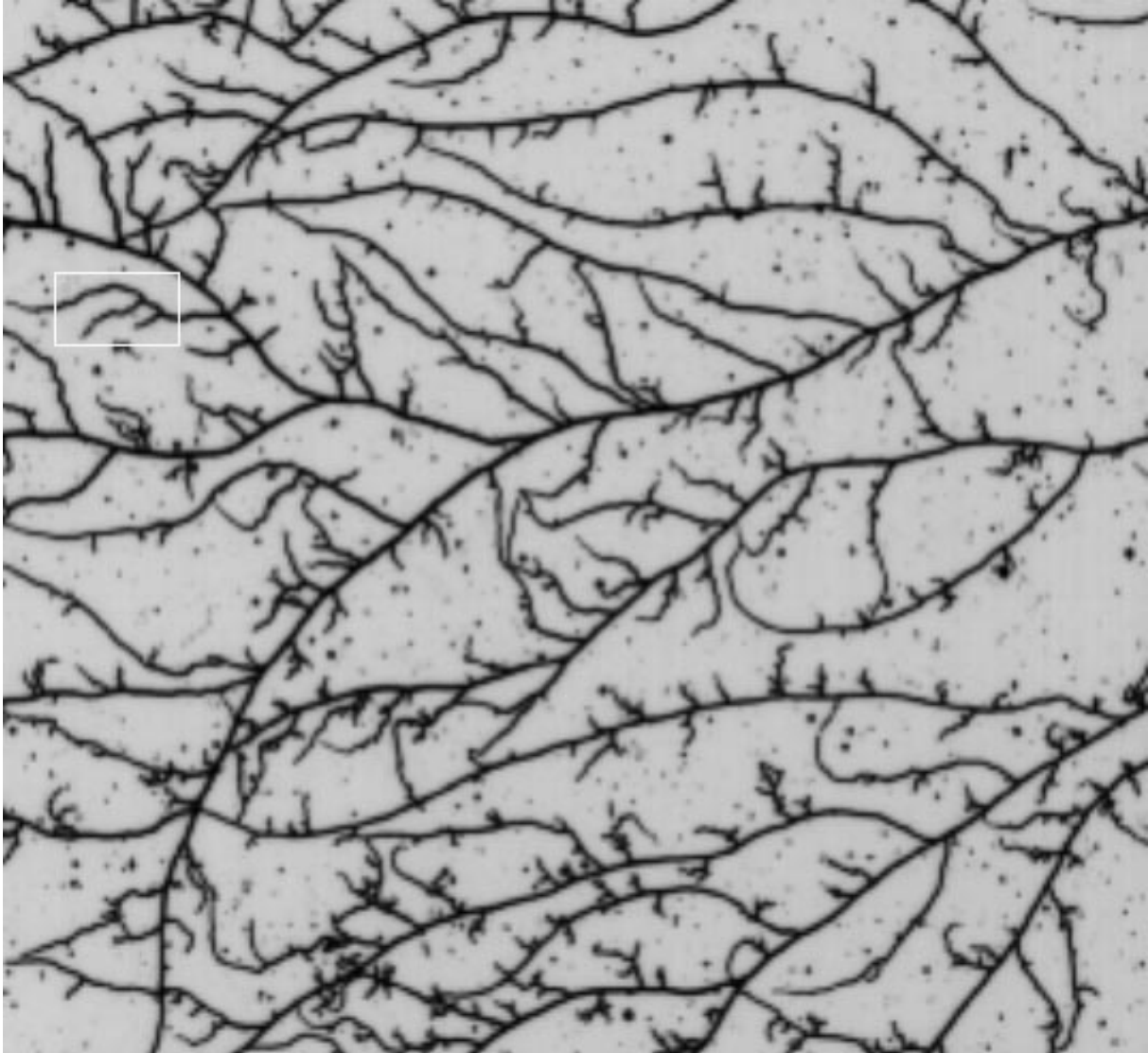
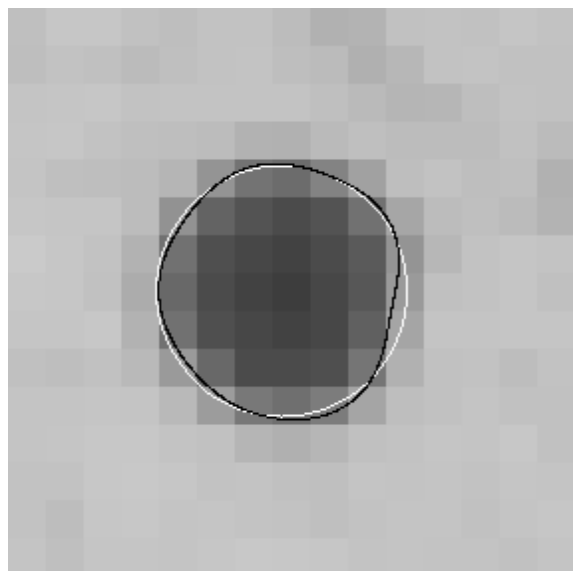
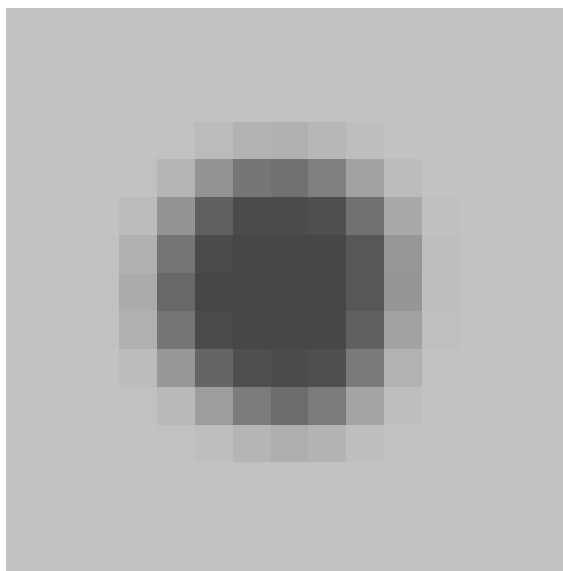


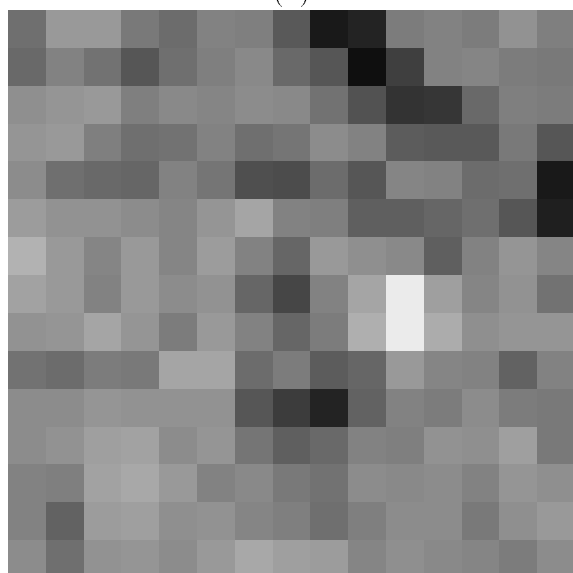
Figure 2:



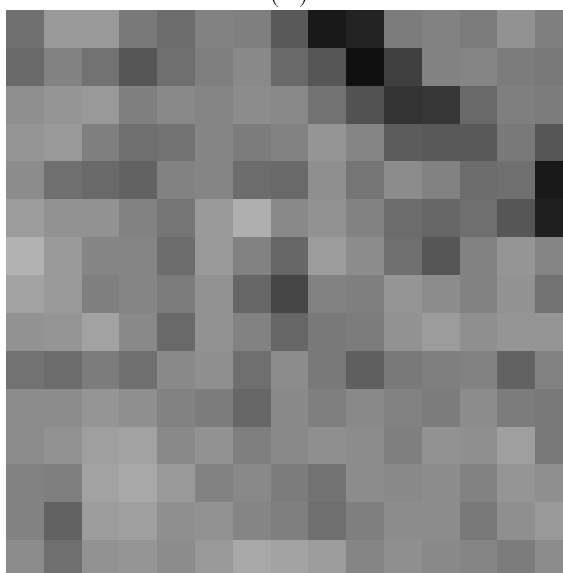
(a)



(b)



(c)



(d)

Figure 3:

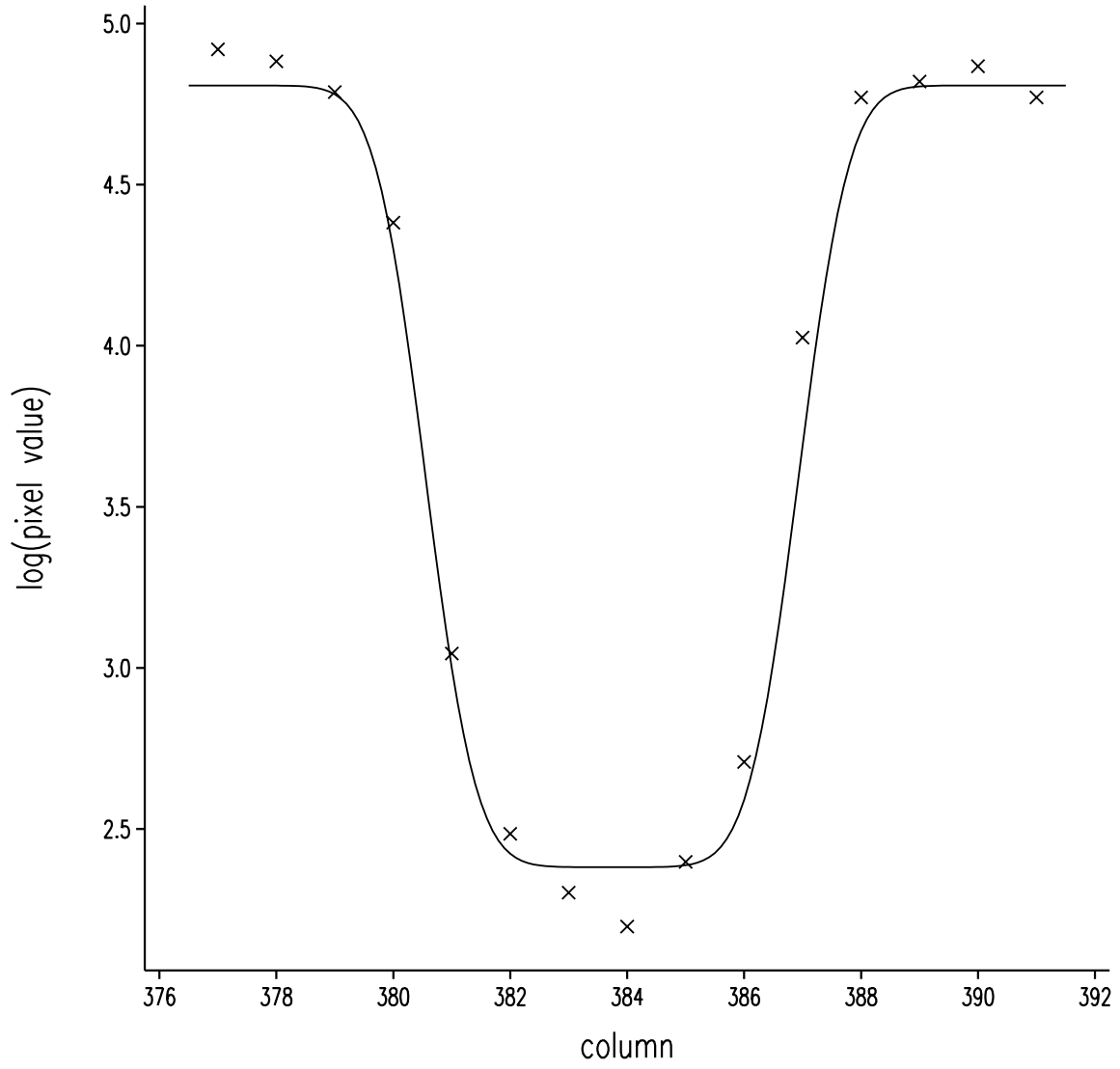
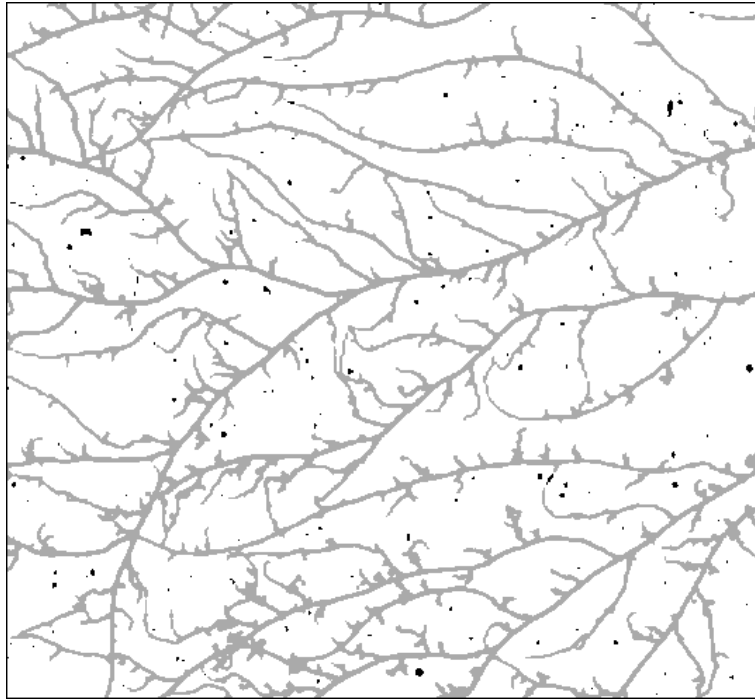
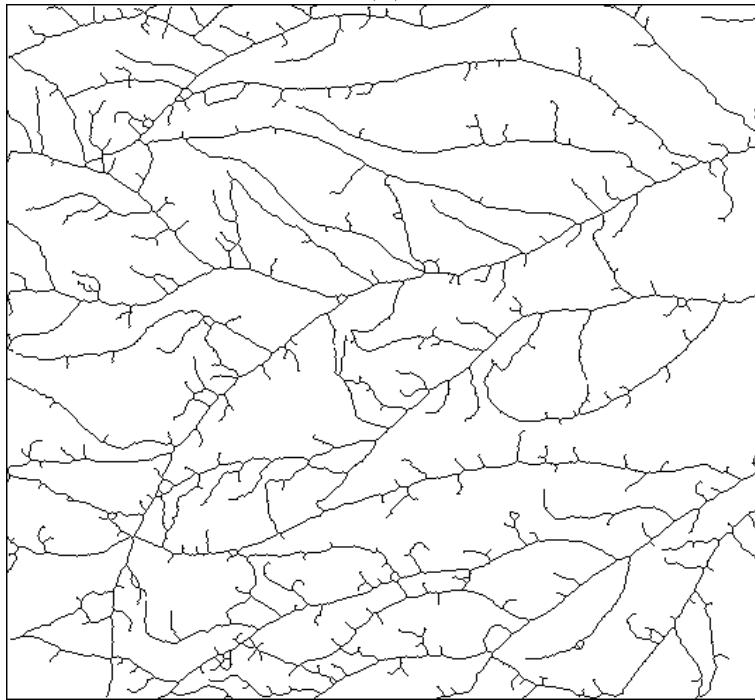


Figure 4:

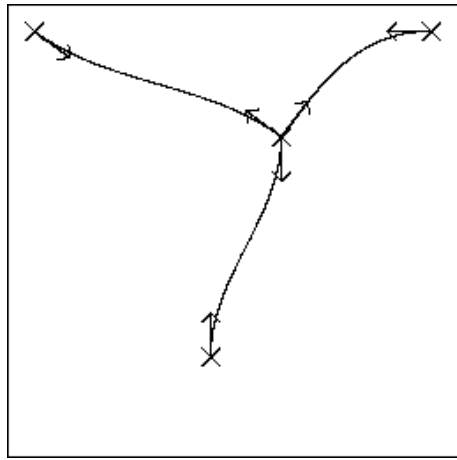


(a)

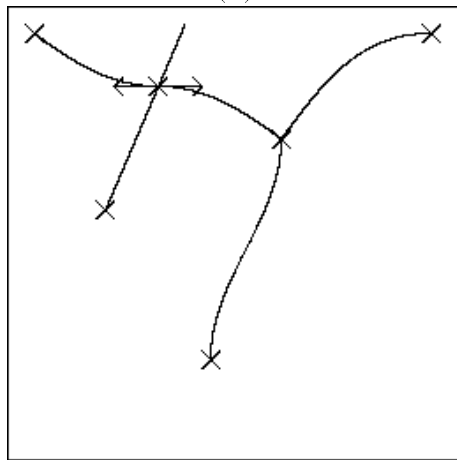


(b)

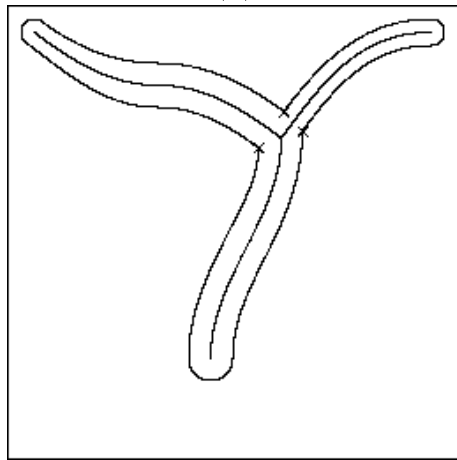
Figure 5:



(a)

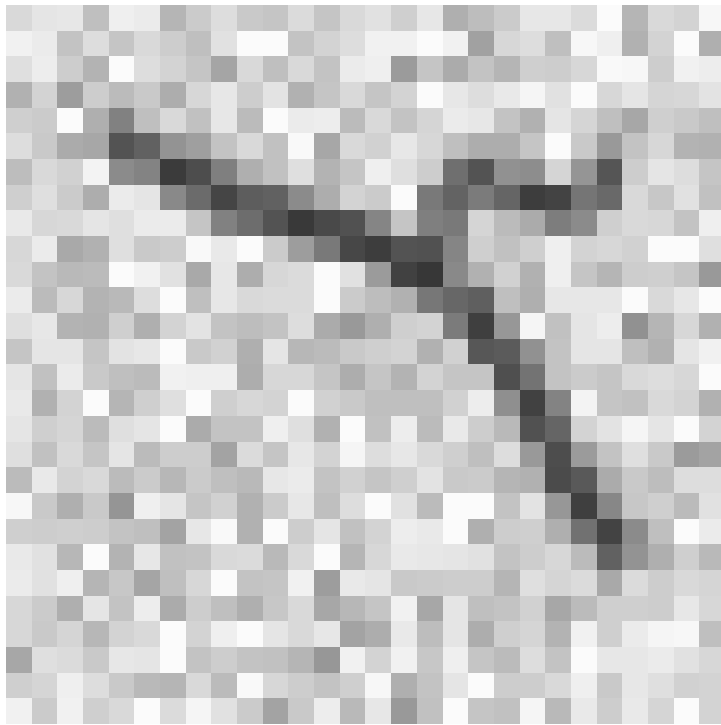


(b)

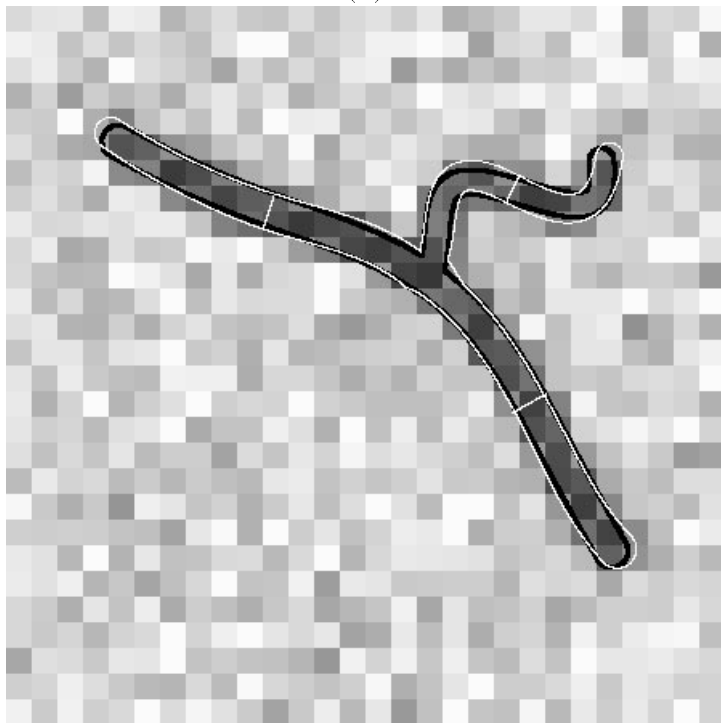


(c)

Figure 6:

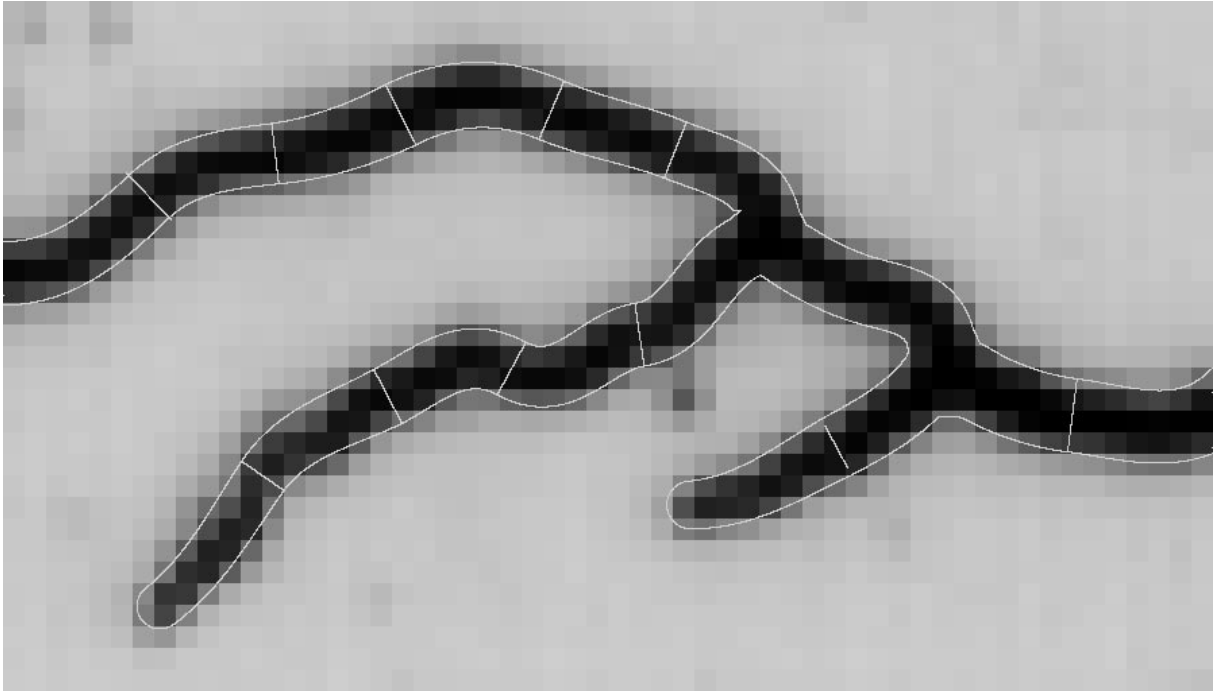


(a)

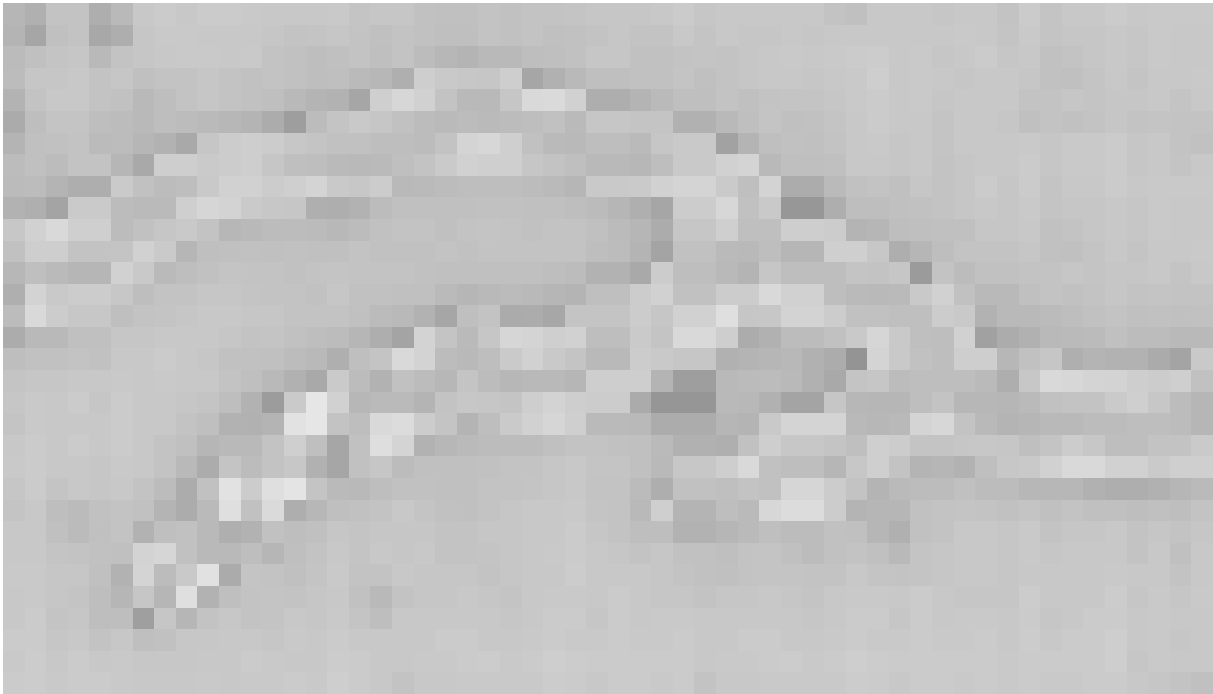


(b)

Figure 7:



(a)



(b)

Figure 8: