– 392 –

Inferring vascular tree structure from 2D and 3D imagery

by

A. Bhalerao, E. Thönnes, W.S. Kendall and R.G. Wilson

January 2002

DEPARTMENT OF STATISTICS

UNIVERSITY OF WARWICK

Inferring Vascular Structure from 2D and 3D Imagery

Abhir Bhalerao[1], Elke Thönnes[2], Wilfrid Kendall[2] and Roland Wilson[1]

Departments of [1]Computer Science and [2]Statistics
University of Warwick, UK
{abhir|elke|wsk|rgw}@{dcs|stats}.warwick.ac.uk

Abstract. We describe a method for inferring vascular (tree-like) structures from 2D and 3D imagery. A Bayesian formulation is used to make effective use of prior knowledge of likely tree structures with the observed being modelled locally with intensity profiles as being Gaussian. The local feature models are estimated by combination of a multiresolution, windowed Fourier approach followed by an iterative, minimum mean-square estimation, which is both computationally efficient and robust. A Markov Chain Monte Carlo (MCMC) algorithm is employed to produce approximate samples from the posterior distribution given the feature model estimates. We present results of the multiresolution parameter estimation on representative 2D and 3D data, and show preliminary results of our implementation of the MCMC algorithm [1].

1 Introduction

The problem of inferring vascular structure from two and three dimensional image data is an important one, especially in the area of surgical planning, which requires a combination of efficient computation and a method of using prior knowledge. Previous work in the area has tended to focus on the modelling of specific vascular features, using deformable templates [7, 3] or to use heuristic approaches such as adaptive thresholding or level sets [4, 6].

The aim of the work described here is to formulate a general method for the inference, which can be applied in two or three dimensions and makes effective use of prior knowledge, yet which is sufficiently general to be applied to a wide range of problems. The common statistical methods for such medical image analysis have typically used likelihood techniques, such as Expectation-Maximisation [11, 9]. Although iterative, EM methods can be efficient computationally, but provide only a limited way of incorporating prior knowledge. A general and powerful way of including further prior information is to use a Bayesian method, such as *maximum a posteriori* (MAP) estimation. The difficulty with Bayesian techniques is essentially a computational one: they typically require the use of Markov chain Monte Carlo (MCMC) algorithms, which may run for hundreds of thousands of iterations to yield reliable results [5]. This has restricted their use in applications involving large data sets such as medical images.

[1] This project is funded by UK EPSRC

The method we have adopted is grounded in statistical inference, combining local likelihood maximisation using a Gaussian model of the spatial intensity profile, and global structure determination using a Bayesian technique derived from a general model of vasculature as a collection of tree structures. By approaching the problem in this way, we can keep the efficiency of likelihood techniques, while exploiting the power and generality of a Bayesian approach. To improve the efficiency and robustness of the computation, we have adopted a multiresolution method, similar to that described in [12]. After a brief description of the estimation algorithms, we present results of two and three dimensional structure inference on real data. The paper is concluded with some observations on the technique.

2 Local structure estimation

We approximate the local shape of a vessel as being linear (lines and cylinders) and employ an iterative fitting technique to minimise the sum of squared residuals between the data f and our model g. 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 cylinder) is windowed by a smooth function $w()$, then it can be approximated by a n-dimensional Gaussian function:

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

parameterised by $\Theta = \{A, \boldsymbol{\mu}, C\}$ with amplitude A, centred on $\boldsymbol{\mu}$ 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, $f_w(\boldsymbol{x}) = w(\boldsymbol{x})f(\boldsymbol{x})$, are modelled as conditionally normal, given the local model $f_w(\boldsymbol{x}) \sim N(g(\boldsymbol{x}|\Theta), \sigma^2)$. 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:

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

From an initial estimate of $\Theta_0 = \{A_0, \boldsymbol{\mu}_0, C_0\}$ 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 $\boldsymbol{x}$ for clarity):

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

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

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

The initial estimate Θ_0 is obtained by using the multiresolution Fourier Transform (MFT) [12]. The windowed Fourier transform of $f_w(\boldsymbol{x}) \leftrightarrow \hat{f}_w(\boldsymbol{u})$ is also a Gaussian with the spectral energy distribution dependent on the type of feature (see [10]). The principal components of the moment of inertia tensor I of the spectral energy:

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

where B is the block size, gives n eigenvalues $\lambda_1 \geq .. \geq \lambda_n$, which are inversely related to the covariance, C, and the spatial orientation of a linear feature is that of the eigenvector associated with the largest eigenvalue λ_1.

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

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

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

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

If $f_w(\boldsymbol{x})$ 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 Inferring the global structure

To draw inferences about the global structure, we employ a Bayesian formalism: the data are modelled as a random tree-like structure and we then use an MCMC algorithm [2] to sample from the posterior distribution, which is conditioned on the data. The sampling distribution is an approximate equilibrium of a random process whose configuration space is the space of tree-like structures and whose equilibrium is designed to be the target conditional distribution. As well as gaining information about the global structure, variation in the posterior samples enables us to quantify uncertainties about the image interpretation.

The prior distribution is that the global structure is a *forest* of a random number of trees. Each such tree is a binary tree: branches divide only into two sub-branches at a time. A physical realisation of such a tree needs each node to be located in space. Unfortunately, the simplistic approach of displacing each node from its parent by a Gaussian displacement of zero mean (a "random-walk" tree) leads to a tangled local structure (left hand figure 1). We therefore introduce a correlation by allowing the mean displacement to be a small linear multiple of the displacement of the parent node from the grandparent node (an "AR(1)" tree) (right hand of figure 1). To each node we then associate a Gaussian kernel that represents the corresponding vessel segment.

The posterior distribution for a random number N of trees $\tau_1, \ldots, \tau_N$ is given by

$$\pi(\tau_1, \ldots, \tau_n|\Theta) \propto \mathbb{P}(N = n) \prod_{i=1}^{n} \prod_{\nu \in \tau_i} p_{v(\nu)} \phi(x_\nu | x_\eta, \eta \in \mathcal{A}(\nu)) \psi(\theta_\nu | \theta_{\text{parent}(\nu)}) \times L(\tau_1, \ldots, \tau_n|\Theta), \tag{9}$$

where p_0, p_1, p_2 are the family size probabilities of the branching process and $v(\nu)$ is the valence of the node ν. Moreover, $\phi(x_\nu|x_\eta, \eta \in \mathcal{A}(\nu))$ is the location distribution of x_ν which depends on the location of its parent and grandparent (if any) given by the ancestor set $\mathcal{A}(\nu)$. The distribution $\psi(\theta_\nu|\theta_{\text{parent}(\nu)})$ of the Gaussian kernel parameters θ_ν associated to each node ν depends on the parameters of the Gaussian kernel associated to its parent node. Finally, $L(\tau_1, \ldots, \tau_n|\Theta)$ denotes the likelihood of the forest given the local structure estimates Θ assuming pixelwise iid white Gaussian noise. The simulation, whose configuration at

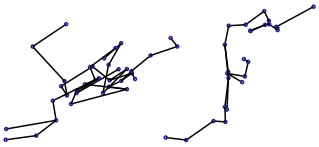

Fig. 1. Illustration of a random-walk tree and an AR(1) tree

any one time is a collection of random trees, is designed to have *moves* which take it from one configuration to the next. These moves, some of which are illustrated in Figure 2, are: adding or deleting a tree, adding or deleting a twig at the end of a branch; displacing a node; splitting a tree into two or grafting two trees together into one; changing the parameters of the Gaussian kernel associated with a node. As long as each move has an 'opposite' (e.g. add versus delete, split

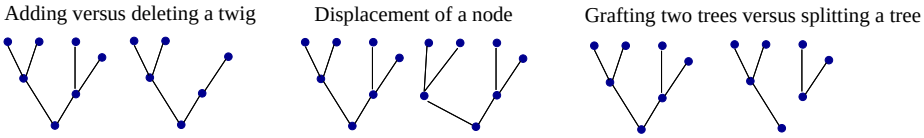

Fig. 2. Illustration of some moves

versus graft) and the chances of each move are balanced against its opposite, it is straightforward to compute the required equilibrium distribution and to design

the move probabilities to give the required posterior as equilibrium, using the Metropolis-Hastings technique (MH) [5]. The MH method iterates in two steps: the first proposes a move from the collection of moves and the second accepts or rejects the move so as to ensure that the equilibrium coincides with the desired posterior. Generally, the decision whether to accept or reject a move depends on whether the resulting new tree will be a more adequate representation from the posterior than the current tree.

The efficiency of the algorithm depends crucially on the proposed moves. As all moves influence the likelihood only locally, the likelihood evaluation can be implemented efficiently. Moreover, to avoid inefficiencies due to a large number of moves being rejected, our moves are guided by likelihood considerations: the locality of a move is dependent on the amount of unexplained data in the surrounding spatial region and the mean direction of a proposed branch segment resembles the direction of the data in the vicinity of the segment. This approach is reminiscent of the Langevin Hastings algorithm, where proposals are influenced by the gradient of the posterior, see [1]. Global structure that can be inferred with high certainty locally will lead to a tree structure that is stable over time, while low local certainty results in a volatile tree-structure that alternates between different explanations for the global structure.

4 Experiments

Figure 3 illustrates results of the ML model estimation. We have used part of a 2D retinal angiographic image size 404 $\times$ 404 pixels (fig. 3(a)) for our 2D experiments. The background is first modelled as locally piece-wise linear and subtracted from the original data prior to the estimating the Gaussian feature parameter estimates. The orientation and position of the MFT feature estimates which form the initial block parameter estimates, Θ_0 are shown in the centre (fig. 3(b) and (c)), where the feature intensity is modulated by the goodness of fit of the model to the data. Despite the noise, there is good correspondence of the large scale arterial structure at block size $B = 64$. The background noise becomes insignificant at block size $B = 16$ (fig. 3(c)). The right hand column (fig. 3(d)-(e)) shows a data reconstruction of the 2D Gaussians in each block (at corresponding block sizes) after the iterative ML estimation. Note that the thickness of the vessels (the standard deviation of the model orthogonal to the feature) are accurately modelled at both large and small block sizes. Clearly, at lower spatial resolutions, the model cannot easily describe the presence of multiple vessels within the window, such as at vessel bifurcations, and the resulting low-amplitude, isotropic Gaussians are locally the 'best' description of these regions. However, these blocks can be identified from higher residual errors, χ^2 in equation (2).

The 3D implementation of the ML model estimator is demonstrated in figure 4(a)-(e). 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. 4(a)). Illustrations of the MFT feature estimates and data reconstructions using 3D Gaussians are shown at two (cubic) window sizes: $B = 32$ and $B = 8$ (fig. 4: top and bottom rows respectively). The major vessels are captured at the

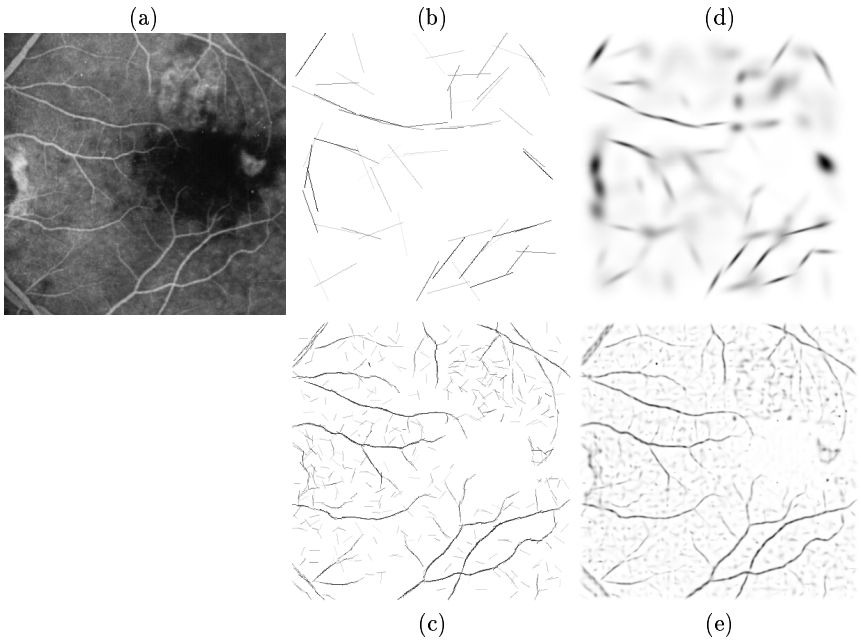

Fig. 3. *Left*: (a) 2D retinal angiogram size 404 × 404 pixels. *Centre*: MFT feature estimates Θ_0 for block sizes (b) 64 and, (c) 16. *Right*: Reconstruction of data from model parameters estimates Θ_{ML} for block sizes (d) 64 and, (e) 16.

lower resolution (fig. 4(b)-(c)) while the finer vessels can be seen in figure 4(d)-(e). Note that at both scales, as with the 2D retinal example, the vessel diameters are correctly estimated.

To demonstrate the utility of the use multiple window sizes, we generated a a multi-level data reconstruction after a simple top-down scale selection based on the normalised block residual errors, figure 5. In this reconstruction, we used 4.4% of the parameters sets, Θ_{ML}, from a total of 13596 across 4 block sizes. This reconstruction is able to capture both large and small local structures in the data.

To infer global structure we have used an MCMC algorithm on the 2D retinal data to sample the posterior distribution in (9) using the Gaussian estimates (figure 6). These show that the method does indeed capture significant global structure.

5 Conclusions

Some encouraging preliminary results have been achieved using the approach described in section 3, demonstrating its potential for modelling vascular structure globally in a computationally efficient way. Fine-tuning the algorithm will lead to significant improvements. These will include, for example, the use of the local estimates to produce initial configurations for the MCMC algorithm, using

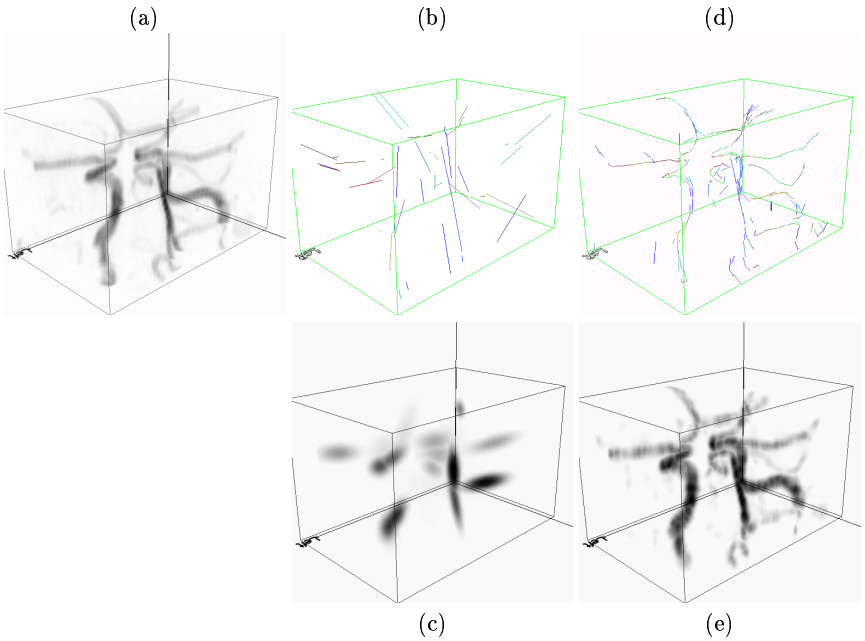

Fig. 4. (a) Maximum intensity projection of part of a MR angiogram depicting cerebral blood vessels size 88 × 58 × 44 voxels. *Top*: MFT feature parameter estimates Θ_0 at window sizes (b) 32 and (d) 8 showing feature orientation and position. *Bottom*: Data model based on block-by-block summation of local Gaussian models Θ_{ML} at window sizes (c) 32 and, (e) 8. The Gaussian amplitude has been made proportional to the goodness of fit between the data and the model.

a posterior based on the multiresolution representation, such as shown in figure 5. Such improvements are currently being implemented.

References

1. J.E. Besag. Contribution to the discussion paper "Representation of knowledge in complex systems" by U. Grenander and M.I. Miller. *Journal of the Royal Statistical Society Series B*, 56:591 – 592, 1994.
2. S.P. Brooks. The Markov Chain Monte Carlo Method and its Application. *The Statistician*, 47:69 – 100, 1998.
3. A. F. Frangi, W. J. Niessen, R. M. Hoogeveen, Th. van Walsum, and M. A. Viergever. Model-based Quantitation of 3D Magnetic Resonance Angiographic Image. *IEEE Trans. Medical Imaging*, 18(10):946–956, 1999.
4. G. Gerig, T. Koller, G. Székely, Ch. Brechbühler, and O. Kübler. Symbolic description of 3-D structures applied to cerebral vessel tree obtained from MR angiography volume data. In H. H. Barrett and A. F. Gmitro, editors, *Information Processing in Medical Imaging IPMI'93, Lecture Notes in Computer Science*, volume 687, pages 94–111, 1993.
5. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, 1996.

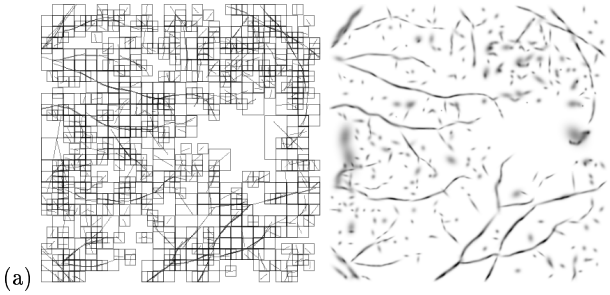

Fig. 5. (a) Multiresolution scale selection based on thresholding goodness of fit between model and data across levels with overlapping block sizes 64, 32, 16 and 8. (b) Multiresolution data reconstruction using parameter set depicted in (a).

Fig. 6. (a) Iteration 2000 of the MCMC simulation based on posterior derived from block sizes $B = 8$ on the 2D retinal image. (b) Iteration 10000. (c) Iteration 50000.

6. 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.
7. T. O'Donnell, A. Gupta, and T. Boult. A new model for the recover of cylindrical structures from medical image data. In J. Troccaz, E. Grimson, and R. Mösges, editors, *Proc. CVRMed-MRCAS'97*, 1997.
8. A. Papoulis. *Signal Analysis.* McGraw-Hill, New York, 1977.
9. W. M. Wells, R. Kikinis, W. E. L. Grimson, and R. Jolesz. Adaptive segmentation of MRI data. *IEEE Trans. Medical Imaging*, 15:429–442, 1996.
10. 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.
11. D. L. Wilson and J. A. Noble. An adaptive segmentation algorithm for extracting arteries and aneurysms from time-of-flight MRA data. *IEEE Trans. Medical Imaging*, 18(10):938–945, 1999.
12. 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.

Other University of Warwick Department of Statistics Research Reports authored or co–authored by W.S. Kendall.

161: The Euclidean diffusion of shape.

162: Probability, convexity, and harmonic maps with small image I: Uniqueness and fine existence.

172: A spatial Markov property for nearest–neighbour Markov point processes.

181: Convexity and the hemisphere.

202: A remark on the proof of Itô's formula for C^2 functions of continuous semimartingales.

203: Computer algebra and stochastic calculus.

212: Convex geometry and nonconfluent Γ-martingales I: Tightness and strict convexity.

213: The Propeller: a counterexample to a conjectured criterion for the existence of certain convex functions.

214: Convex Geometry and nonconfluent Γ-martingales II: Well–posedness and Γ-martingale convergence.

216: *(with E. Hsu)* Limiting angle of Brownian motion in certain two–dimensional Cartan–Hadamard manifolds.

217: Symbolic Itô calculus: an introduction.

218: *(with H. Huang)* Correction note to "Martingales on manifolds and harmonic maps."

222: *(with O.E. Barndorff-Nielsen and P.E. Jupp)* Stochastic calculus, statistical asymptotics, Taylor strings and phyla.

223: Symbolic Itô calculus: an overview.

231: The radial part of a Γ-martingale and a non-implosion theorem.

236: Computer algebra in probability and statistics.

237: Computer algebra and yoke geometry I: When is an expression a tensor?

238: *Itovsn3*: doing stochastic calculus with *Mathematica*.

239: On the empty cells of Poisson histograms.

244: *(with M. Cranston and P. March)* The radial part of Brownian motion II: Its life and times on the cut locus.

247: Brownian motion and computer algebra (Text of talk to BAAS Science Festival '92, Southampton Wednesday 26 August 1992, with screenshots of illustrative animations).

257: Brownian motion and partial differential equations: from the heat equation to harmonic maps (Special invited lecture, 49th session of the ISI, Firenze).

260: Probability, convexity, and harmonic maps II: Smoothness via probabilistic gradient inequalities.

261: *(with G. Ben Arous and M. Cranston)* Coupling constructions for hypoelliptic diffusions: Two examples.
280: *(with M. Cranston and Yu. Kifer)* Gromov's hyperbolicity and Picard's little theorem for harmonic maps.
292: Perfect Simulation for the Area-Interaction Point Process.
293: *(with A.J. Baddeley and M.N.M. van Lieshout)* Quermass-interaction processes.
295: On some weighted Boolean models.
296: A diffusion model for Bookstein triangle shape.
301: COMPUTER ALGEBRA: an encyclopaedia article.
308: Perfect Simulation for Spatial Point Processes.
319: Geometry, statistics, and shape.
321: From Stochastic Parallel Transport to Harmonic Maps.
323: *(with E. Thönnes)* Perfect Simulation in Stochastic Geometry.
325: *(with J.M. Corcuera)* Riemannian barycentres and geodesic convexity.
327: Symbolic Itô calculus in *AXIOM*: an ongoing story.
328: *Itovsn3* in *AXIOM*: modules, algebras and stochastic differentials.
331: *(with K. Burdzy)* Efficient Markovian couplings: examples and counterexamples.
333: Stochastic calculus in *Mathematica*: software and examples.
341: Stationary countable dense random sets.
347: *(with J. Møller)* Perfect Metropolis-Hastings simulation of locally stable point processes.
348: *(with J. Møller)* Perfect implementation of a Metropolis-Hastings simulation of Markov point processes
349: *(with Y. Cai)* Perfect simulation for correlated Poisson random variables conditioned to be positive.
350: *(with Y. Cai)* Perfect implementation of simulation for conditioned Boolean Model via correlated Poisson random variables.
353: *(with C.J. Price)* Zeros of Brownian Polynomials.
371: *(with G. Montana)* Small sets and Markov transition densities.
382: *(with A. Brix)* Simulation of cluster point processes without edge effects.
391: *(with E. Thönnes, A. Bhalerao, R.G. Wilson)* A Bayesian approach to inferring vascular tree structure from 2D imagery.
392: *(with A. Bhalerao, E. Thönnes, R.G. Wilson)* Inferring vascular tree structure from 2D and 3D imagery.

Also see the following related preprints

317: E. Thönnes: Perfect Simulation of some point processes for the impatient user.

334: M.N.M. van Lieshout and E. Thönnes: A Comparative Study on the Power of van Lieshout and Baddeley's J-function.

359: E. Thönnes: A Primer on Perfect Simulation.

366: J. Lund and E. Thönnes: Perfect Simulation for point processes given noisy observations.

If you want copies of any of these reports then please email your requests to the secretary using statistics@warwick.ac.uk (mail address: the Department of Statistics, University of Warwick, Coventry CV4 7AL, UK).