Proc. SPIE **3034**, pp.276-286 (1997)
Medical Imaging: Image Processing, K. M. Hanson, ed.

Uncertainties in tomographic reconstructions based on deformable models

Kenneth M. Hanson, Gregory S. Cunningham, and Robert J. McKee

Los Alamos National Laboratory, MS P940
Los Alamos, New Mexico 87545 USA

ABSTRACT

Deformable geometric models fit very naturally into the context of Bayesian analysis. The prior probability of boundary shapes is taken to proportional to the negative exponential of the deformation energy used to control the boundary. This probabilistic interpretation is demonstrated using a Markov-Chain Monte-Carlo (MCMC) technique, which permits one to generate configurations that populate the prior. One of many uses for deformable models is to solve ill-posed tomographic reconstruction problems, which we demonstrate by reconstructing a two-dimensional object from two orthogonal noisy projections. We show how MCMC samples drawn from the posterior can be used to estimate uncertainties in the location of the edge of the reconstructed object.

Keywords: deformable geometric models, Bayesian estimation, tomographic reconstruction, uncertainty estimation, Markov Chain Monte Carlo

1. INTRODUCTION

We have seen in recent years increasing use of deformable geometric models to analyze medical images. For evidence of this trend, one has only to look in the relevant journals or conference proceedings. Many examples of the use of deformable models can be found in this Proceedings for segmentation, registration, and tomographic reconstruction. Deformable models have shown their usefulness by supplying a means of interpolating between noisy or missing data and providing the regularization needed to solve ill-posed problems.

Deformable models allow one to move toward the goal of analyzing medical images in terms of organs. As the field of medical imaging moves toward image interpretation, computer diagnosis, etc., a more automated approach to image analysis and understanding will be essential. Deformable models can provide a framework within which to organize an image. Although individual organs are currently being modeled, we can look forward to analyses involving multiple organs, complete with interactions or constraints between various organs.

The main purpose of this paper is to provide an overview of deformable geometric models. In the authors' view, deformable models fit particularly well into the framework of Bayesian analysis. The smoothness of deformable models is typically controlled by means of a deformation energy function that provides a measure of deviation from smoothness.[1] In the context of Bayesian analysis, the probability of the corresponding geometric configuration is taken to be the negative exponential of this energy function.[2,3] We will emphasize this interpretation by showing a set of random samples of boundary configurations taken from a prior probability based on a typical form for the deformation energy. These samples are generated using the Markov Chain Monte Carlo (MCMC) technique.

A secondary theme of this paper concerns the assessment of uncertainties in estimated models, a capability greatly assisted by the use of Bayesian analysis. The general method we employ here is to generate a sequence of random samples of the posterior probability distribution using MCMC. By fully mimicking the posterior, this sequence of samples can be used to assess the uncertainties in estimated models. As an example of the usefulness of this technique, we consider a problem of reconstructing an object from projections in two directions under the assumption of a known, constant interior density. In the analysis, the boundary of the reconstructed objects are

Further author information –
K.M.H.(correspondence): Email: kmh@lanl.gov; WWW: http://home.lanl.gov/kmh; Telephone: 505-667-1402
G.S.C.: Email: cunning@lanl.gov; Telephone: 505-667-2562

subject to a prior that promotes smoothness. We show how samples from the posterior can be used to estimate uncertainties in the location of the boundary of the reconstructed object.

We deliberately choose our example of tomographic reconstruction from two views because this extremely difficult problem can not be solved with conventional reconstruction algorithms, as shown in Ref. 4, for example. Our ability to obtain an excellent reconstruction emphasizes the advantage of using deformable models in tomography when objects have a relatively simple shape and possess constant density. MCMC provides the means to verify the reliability of our reconstruction. Conventional approaches to uncertainty estimation are inadequate to treat this problem because of the nonlinear relation between the data and the model parameters and, hence, the potential nonGaussian nature of the posterior distribution.

2. DEFORMABLE GEOMETRIC MODELS

2.1. Deformable Boundaries

Researchers began to realize importance of edges in human vision[5] several decades ago. However, it was not until 1987 or so that Kass, Witkin, and Terzopoulos[1] introduced for use in computer vision the use of spline "snakes" to represent the boundaries of objects in images. Deformable curves used to represent object boundaries have also been called active contours. To generalize the terminology, we call them deformable geometric models, or just deformable models, for short, and use this term to cover all types of flexible geometric models.

Clearly, the idea behind the use of snakes is to focus on the geometry or shape of objects as described by their boundaries. This kind of model actually imposes desirable constraints on the result of an analysis, namely that the object's boundary is connected and continuous and that it has a sharply-defined edge.

In the snake approach, the rigidity of the model is restricted through the choice of a restrictive representation of the curve. For example, one could choose polynomial patches, such as B-spline or Bezier patches with continuity constraints applied between them. Each patch has a limited number of degrees of freedom determined by the order of polynomial chosen and the degree of intrapatch continuity constraint. The flexibility of the curve is controlled by the order of polynomial and/or number of patches. This approach, which can be thought of as a finite-element representation, is continues to be used.[6] One has only limited and discrete control in this approach.

We believe there is a significant benefit to using the finely-divided polygon representation instead of finite elements. A polygon can approximate a curve to any desired accuracy by subdividing the sides finely enough. However, by itself, a polygon has no constraint on its shape. Constraints are placed on the curve by introducing a prior. Often the form for the prior is taken to be an exponential of a deformation energy, for example, the integral of the square of the curvature. Constraints other than on the curvature may be used, such as requiring that that curve not intersect itself. The advantage of this overall approach is that alternative forms for the prior can be tried and assessed for any particular imaging situation. The present work uses this approach.

This approach has the major advantage of allowing continuously tunable control over the flexibility of the boundary. The strength of the prior (α in the following Subsection) regulates the effective number of degrees of freedom that the curve has, much in the same way as the number of patches in a Bezier representation. Furthermore, it is possible to multiply the contribution of each vertex in the sum in Eq. (1) by a variable weight. That would allow one to encourage the curve to be smooth in some regions and very flexible in others. In the extreme case of letting the weight go to zero at a vertex, it is possible to let the curve develop a kink there.[8]

Both the above approaches can be easily extended to analyzing three-dimensional objects. Then the boundaries are represented by surfaces instead of curves. The control of smoothness is even greater for boundaries in three dimensions than in two, because of the additional constraint placed on the boundary by the extra dimension. Extensions to even higher dimensions are easily visualized.

A natural way to think about the deformation energy functions used to control boundaries is in terms of the analogy to mechanical systems. Thus, a curve can be thought of as a thin rod. The stiffness of the rod can help control how it is affected by data. Of course, the term "spline" came from the use of bendable strips by draftsmen to draw smooth curves. Just as derivatives of potentials are forces, the derivatives of the energy function can be thought of as forces created by the rod to counteract the pull of data.

2.2. Deformable Interiors

Instead of modeling the boundary of an object in terms of deformable models, it is possible to model its interior. One advantage of this approach is that it is easy to avoid opposite sides of the boundary from crossing each other. The classification of deformable boundary models into representations based on finite elements and those based on a flexible continuous curve hold for deformable interior geometry, as well. A recent example using finite elements to model the interior of the brain is that by Gee et al..[7]

Related to interior deformation models are those based on globally warping the spatial domain of the object. Thin-plate splines falls into this category. Bookstein[9] used these to analyze the deformation of biological shapes, an area of study called morphometrics. Also, Hanson used this kind of model to solve the ill-posed problem of tomographic reconstruction from several projections.[10,11]

Characterizing objects' interiors in terms of deformable models can permit one to more carefully devise models that can mimic the shapes of specific organs. Clearly the shape of a heart does not resemble that of a liver. As the study of matching organs progresses, we can expect to see more attention paid to such organ-specific constraints on the models. Geometric models of object shape may require multiscale representations of objects, for example as used in cores, advanced by Pizer et al.[12]

2.3. Our Boundary Model

Here we use a deformable model to represent the boundary of the object to be reconstructed. The object's boundary is approximated in discrete terms as a finely-divided polygon. The length of the edges of the polygon can be made short enough to adequately describe a curve at any degree of resolution desired.

A smoothness constraint on the boundary is achieved by placing a prior on the curvature of the boundary. The minus-log-prior is taken to be proportional to $\int \kappa^2(s)\, ds$, where $\kappa(s)$ is the curvature of the curve as a function of the arclength along the curve s. This prior serves to keep the curve smooth because large curvature is penalized. This form for a prior has a physical analog in the formula that describes the potential energy created by bending a stiff rod. We note that since the integral has the dimensions of reciprocal length, it depends on the scale of the curve. To achieve a prior that is related to the shape of the curve, not its size, as suggested in Ref. 13, we propose that the integral should be multiplied by the total arclength of the boundary, forming a dimensionless quantity. From the Bayesian point of view, the prior is interpreted probabilistically. A specific probability is assigned to every closed boundary, which ranks boundary shapes according to their plausibility. This probability expresses the uncertainty about the possible shapes of the object that we have before we take data. The probabilistic interpretation of the prior is underscored in Sect. 5.

For our discrete polygon model, we replace the integral by a sum of contributions associated with each vertex in combination with half of each neighboring edge of the polygon. We use for the minus-log-prior a discrete approximation to the expression

$$\frac{\alpha}{(2\pi)^2} \int ds \int \kappa^2(s)\, ds \; . \tag{1}$$

The factor of $(2\pi)^{-2}$ is included to normalize the result to unity for a circle when $\alpha = 1$. The parameter α, called a hyperparameter rather than a parameter because it controls a general aspect of the model, determines the strength of the prior relative to the likelihood. The form of Eq. (1) is such that for equal-sided polygons it is independent of the length of the sides as the length goes to zero. In addition to the smoothness constraint, we find it useful to control the lengths of the sides of the polygon to avoid any side from either getting too small, or too big, relative to the rest. This control is accomplished in this study by adding to the above minus-log-prior an expression that is quadratic in the deviation of each side of the polygon relative to the average side length. Further details can be found in Refs. 8 and 14.

3. BAYESIAN ANALYSIS

Bayesian analysis is a model-based approach to analyzing data with a strong emphasis placed on uncertainty assessment. Every aspect of modeling is assigned a probability that indicates our degree of certainty in its value. At the lowest level of analysis, the estimation of the values of parameters for a specified model, a probability density function (PDF) is associated with each continuous parameter. Loosely speaking, the range of a probability distribution indicates the possible range of its associated parameter. The benefit of Bayesian analysis over traditional methods of uncertainty, or error, analysis is that it permits the use of arbitrary probability distributions, not just Gaussian distributions, and of arbitrary measures of uncertainty, not just rms deviation (or variance). Bayesian analysis reveals the use of prior knowledge and assumptions, which other kinds of analysis incorporate, but do not always make explicit. Furthermore, it extends analysis to higher levels of interpretation, e.g., the choice of the strength of the priors used, the rejection of any particular model, and the selection of appropriate models.[15] For detailed descriptions of Bayesian analysis the reader is referred to the several excellent books that have appeared in the last few years, including Refs. 16–19 that address Bayesian analysis from a practical point of view. Another useful source of background information is the collection of proceedings from the series of workshops on Maximum Entropy and Bayesian Methods, the most recent editions of which are Refs. 20 and 21.

Before conducting an experiment, one starts with some knowledge about the physical object being studied. In addition one often has a model for the object, with associated parameters $\mathbf{x}$, that will be used to analyze the experimental results. In Bayesian analysis, the uncertainties in what is known beforehand are expressed in terms of a PDF on the parameters, $p(\mathbf{x})$, called the prior. This prior knowledge can come from previous measurements, specific information regarding the object itself, or simply general knowledge about the parameters, e.g., that they are nonnegative. With the experimental data in hand, the prior is modified to yield the posterior, which is the PDF $p(\mathbf{x}|\mathbf{d})$ of the parameters given the observed data $\mathbf{d}$, using Bayes law

$$p(\mathbf{x}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{x})\, p(\mathbf{x}) \quad . \tag{2}$$

The probability $p(\mathbf{d}|\mathbf{x})$, called the likelihood, comes from a comparison of the actual data to the data predicted on the basis of the model of the object. The predicted data are generated using a model for how the measurements are related to the object, which we call the measurement model. Under the assumption that the data are degraded by uncorrelated and additive Gaussian noise, it is appropriate to use the exponential of $-\frac{1}{2}\chi^2$ for the likelihood. As usual, χ^2 is the mean squared difference between the actual and the predicted measurements divided by the variance of the measurements. The typical analysis consists of estimating the "best" model to explain the data. Often the model that maximizes the posterior, referred to as the maximum *a posteriori* (MAP) estimate is found, although other estimators can be argued to be more appropriate in some circumstances. While an estimate of the best model is the objective of many analyses, it is only the beginning for the true Bayesian.

Bayesian analysis is open ended. For example, uncertainties may exist in the choice of model to apply to an analysis. One typically has an initial concept of what model to use to analyze the data. However, the preconceived model may fail to account for the data, which can lead to the realization that the model needs to be extended, or even replaced all together. Model checking is an important aspect of any data analysis. Of course, the data may not have enough accuracy to reject a particular model, in which case the model can be considered valid until more discriminating data are available. Bayesian analysis essentially follows the scientific method; a proven model is used until reliable data indicate that it needs to be altered or rejected. Bayes law can be used to choose between models.[15]

The present study is carried out using the Bayes Inference Engine (BIE). We developed the BIE to allow one to easily develop complex models for both the objects under study and the measurement process. Various aspects of the BIE are described elsewhere.[4,22–26] The MCMC technique is a perfect match to the computational approach to Bayesian inference adopted in the BIE.

4. MARKOV CHAIN MONTE CARLO

In Bayesian analysis there is often the need to integrate over the posterior. One way to do that is to use a Monte Carlo technique, i.e. draw random samples from the posterior, The Markov Chain Monte Carlo (MCMC) technique provides a means to sample an arbitrary probability density function (PDF).

A Markov chain is a sequence of states in which the probability of each state depends only on the previous state. In MCMC the objective is to generate a sequence of parameter sets that mimic a specified PDF, let's call it $q(\mathbf{x})$, where $\mathbf{x}$ is a vector of parameters in the relevant parameter space. More precisely, it is desired that the MCMC sequence be in statistical equilibrium with the target PDF $q(\mathbf{x})$, which is achieved when the MCMC sequence is marked by the condition of detailed balance:

$$q(\mathbf{x})\, T(\mathbf{x} \to \mathbf{x}') = q(\mathbf{x}')\, T(\mathbf{x}' \to \mathbf{x}) \quad , \tag{3}$$

where $T(\mathbf{x} \rightarrow \mathbf{x}')$ is the transition probability for stepping from $\mathbf{x}$ to $\mathbf{x}'$. This equation essentially requires that in a very long sequence the number of steps from $\mathbf{x}$ to $\mathbf{x}'$ is identical to the number from $\mathbf{x}'$ to $\mathbf{x}$. For more information about MCMC, the reader is referred to the recent paper by Besag et al.[27] or the excellent book edited by Gilks et al.,[28] which represents the best available compendium on MCMC.

The MCMC technique[27] makes it feasible to perform some of the difficult technical calculations required by probability theory (normalization of PDFs, marginalization, computation of expectation integrals, model selection) in a computer. The MCMC technique has opened up the possibility of applying Bayesian analysis to complex analysis problems. A desirable attribute of MCMC is that there are generally no restrictions on the types of PDFs that can be sampled; no functional form for the PDF is required. In its basic form, MCMC only requires that one be able to calculate $\varphi = -\log(\text{posterior})$, although sometimes the gradient of φ is used.

4.1. Metropolis Algorithm

One of the simplest algorithms used in MCMC calculations is due to Metropolis et al.[29] This algorithm ensures detailed balance (3) for each step in the sequence. One starts at an arbitrary point in the vector space to be sampled, $\mathbf{x}_0$. The general recursion at any point in the sequence $\mathbf{x}_k$ is as follows:

(1) Pick a new trial sample $\mathbf{x}^* = \mathbf{x}_k + \Delta\mathbf{x}$,
where $\Delta\mathbf{x}$ is randomly chosen from a symmetric probability distribution
(2) Calculate the ratio $r = q(\mathbf{x}^*)/q(\mathbf{x}_k)$
(3) Accept the trial sample, that is, set $\mathbf{x}_{k+1} = \mathbf{x}^*$,
if $r \geq 1$,
or with probability r, if $r < 1$,
otherwise start over at step (1).

This algorithm is used in the majority of current MCMC research and works remarkably well.

5. PROBABILISTIC INTERPRETATION OF THE PRIOR

We can now put together the topics discussed in the last three Sections. We demonstrate the probabilistic interpretation of the prior on the deformable models by using MCMC to sample this probability density function. In this example we start with a circle of radius 0.8 mm, represented as a 50-sided polygon. The strength of the prior given in Eq. (1) is set by choosing $\alpha = 5$, consistent with the example of tomographic reconstruction presented in the next Section. Finding that uncorrellated perturbation of the vertex positions does not lead to a very productive MCMC sequence, we have instituted correlations between neighboring vertices by using for the x and y steps, the average over seven uncorrelated x and y perturbations chosen from a Gaussian distribution with an rms width of 0.015 mm. This sampling scheme results in a 35% acceptance of the MCMC steps. Representative samples from the sequence are shown in Fig. 1.

During the generation of the MCMC sequence, the configuration of the boundary is saved at regular intervals, every 100 accepted steps in the present exmaple. After the full sequence is generated, the saved configurations can be played back as a video loop. We observe that the MCMC sequence seems to come into equilibrium with the posterior distribution after about 2000 accepted steps. Visual observation also indicates that there is a strong correlation in the configurations over several frames, that is, over several hundred steps in the sequence. This correlation might be expected because of the small step that each vertex can take in each iteration of the Metropolis algorithm.

6. UNCERTAINTY IN A TOMOGRAPHIC RECONSTRUCTION

6.1. Problem Statement

We demonstrate the usefulness of deformable models and the versatility of the MCMC technique with an example of tomographic reconstruction from just two views. This problem is an extraordinarily difficult inverse problem. Its solution is made feasible by employing the prior information that the object being reconstructed has constant density and consists of a fairly simple shape with smooth boundaries. Figure 2 displays the object devised for this example. It is fashioned to be representative of a lumen, the cross section of an artery, possessing a sizable occlusion. The calculations here are based on images with a full width of 128×128 pixels, arbitrarily set to a size of 4 mm × 4 mm.

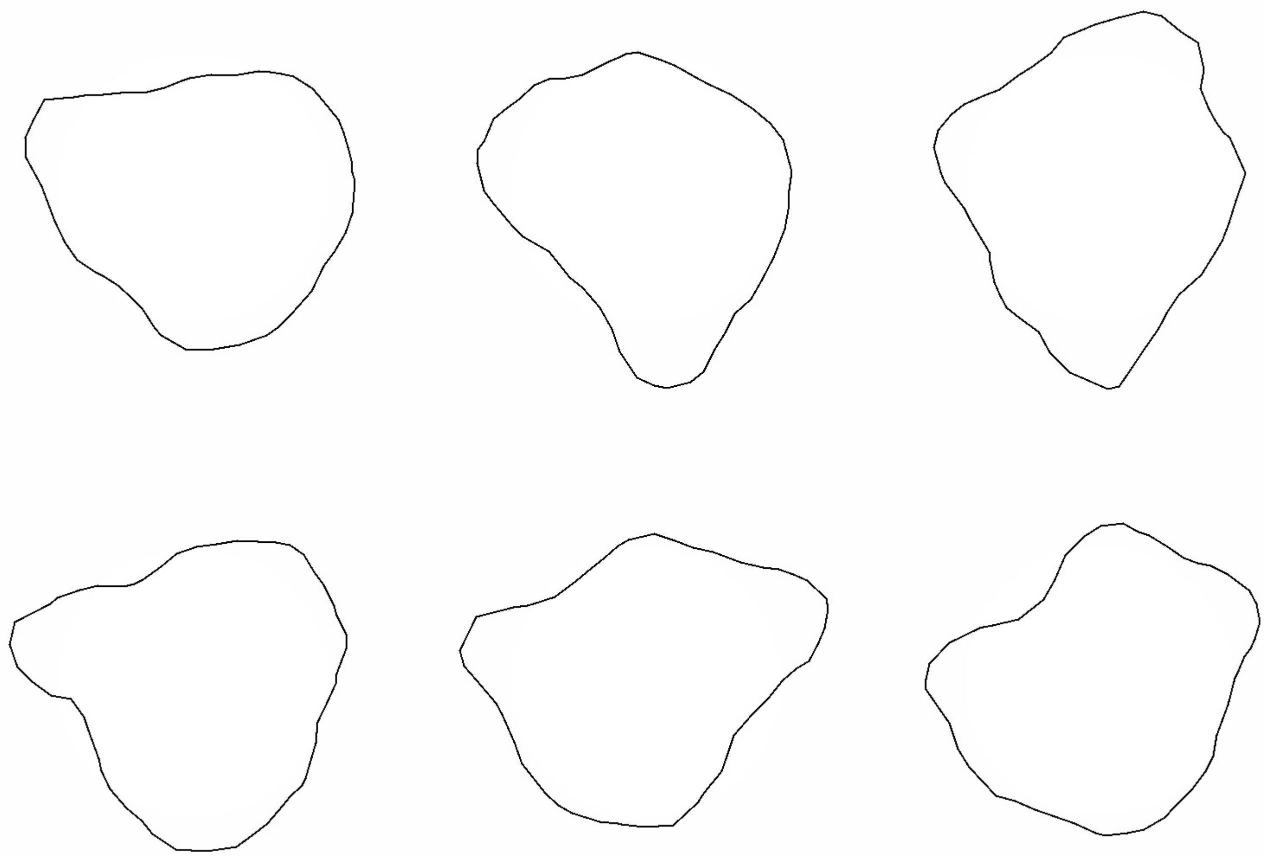

Figure 1. Configurations of the boundary drawn from the prior for $\alpha = 5$ using the MCMC technique. These realizations, taken at intervals of 1000 steps in an MCMC sequence, represent the range of possible shapes allowed by the prior on the curvature.

For better visualization, all the images shown here display just the central 2.5 mm $\times$ 2.5 mm region. To give the scale of the images, the width and height of the object are roughly 64 pixels, or about 2 mm.

Two orthogonal views of the object shown in Fig. 2 are generated, one consisting of the vertical projection and the other of the horizontal. Each projection consists of 128 samples over a distance of 4 mm. Gaussian noise is added to these projections with an rms deviation of 5% of the peak projection amplitude. For this simulation, we ignore blur in the measurement system.

6.2. MAP Reconstruction

For reconstruction the object is modeled in terms of a finely-divided polygon filled with a constant density, which we assume is known beforehand. The polygon has 50 sides to approximate a continuous curve. The parameters in the model consist of the x and y values of the 50 vertices. The –log(posterior) for this problem is the sum of –log(likelihood) $= \frac{1}{2}\chi^2$ and the –log(prior) contributions given in Eq. (1) and the prior of side length. However, the strength of the prior, i.e. α in Eq. (1), must be specified. Ideally, the hyperparameter α that is consistent with the data would be calculated utilizing the next higher level of Bayesian inference.[15] As we are not yet equipped to do that in the BIE, we tried several values for α and selected what seemed to be an appropriate value, $\alpha = 5.0$. Figure 1 shows several realizations of the object shapes allowable with this prior. See below for further justification.

The MAP reconstruction is obtained by using the BIE to find the minimum in $\varphi = -\log$ (posterior) with respect to the 100 variables specifying the polygon model. The BIE accomplishes this in an efficient manner through the use of the Adjoint Differentiation In Code Technique (ADICT)[30] to calculate the gradient of φ with respect to the

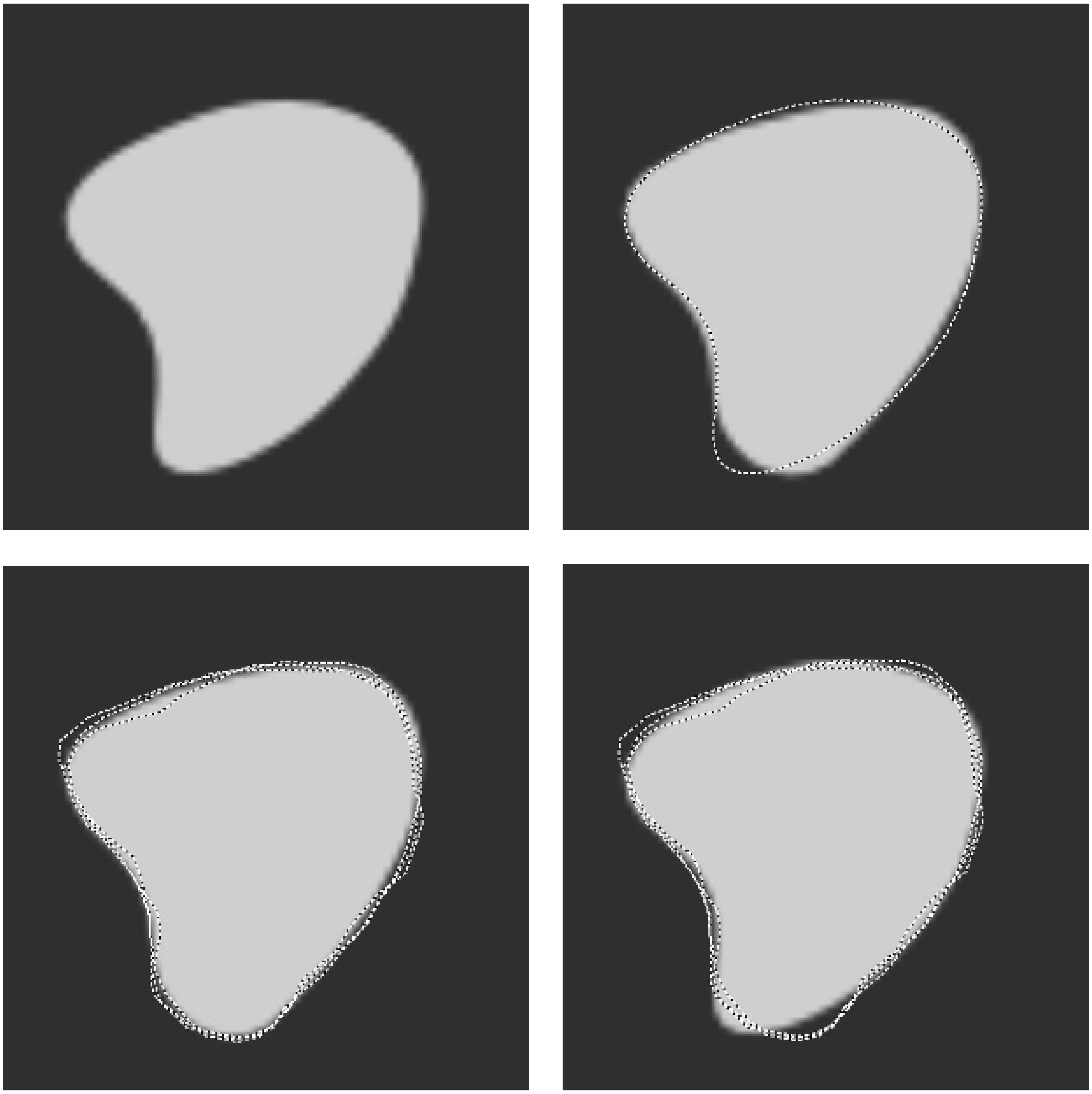

Figure 2. (UL) The original object. (UR) The MAP reconstruction from two orthogonal noisy projections, shown as a grayscale image with the boundary of the original object superimposed. Three representative samples from the posterior shown as curves on top of the grayscale images of (LL) the MAP reconstruction and (LR) the original object.

variables. The initial object is taken to be a circle of diameter 1.6 mm = 51 pixels, for which $\frac{1}{2}\chi^2 = 396.15$ and –log(prior) = 7.88. At the minimum in φ, these values are $\frac{1}{2}\chi^2 = 119.16$, and –log(prior) = 18.24. The resulting reconstructed object compares reasonably well with the original, as shown in Fig. 2. The maximum discrepancy in the position of the two boundaries is about 3.3 pixels, which occurs in the lower lobe. Over the vast majority of the perimeter, the reconstructed boundary lies at most one pixel away from the original.

The reconstruction shown in Fig. 2 is vastly superior to one that would be obtained using conventional reconstruction algorithms. For example, see Ref. 4 for a reconstruction of a similar object from two views obtained using the multipicative algebraic reconstruction technique,[31] which yields an image that maximizes entropy and incorporates a nonnegativity constraint. We will now prove the reliablity of this reconstruction by using MCMC to assess its uncertainty.

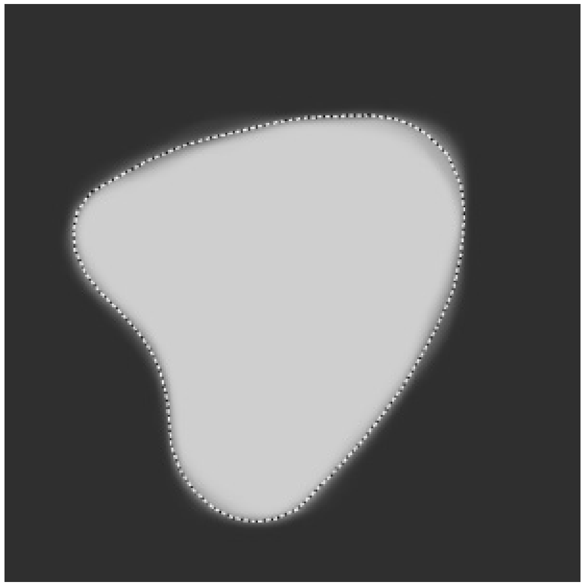

Figure 3. The average of the grayscale images for full MCMC sequence of samples from the posterior with the contour for the MAP reconstruction shown as a dotted line.

6.3. MCMC Results

The MCMC algorithm described above was used to generate samples from the posterior of this reconstruction problem. In our present example, for each MCMC trial step, the increments in the x and y position of each of the vertices are independently chosen from a Gaussian distribution with an rms step size of 0.06 pixels. Since we are drawing vertex step samples from the space of x and y and the priors outlined in Sect. 2.1 are stated in the θ, L space, we need to transform from $p(\theta, L)$ to $p(x, y)$ using the Jacobian of the transform to properly evaluate the priors for the MCMC algorithm. This Jacobian alters the approximately quadratic form of Eqs. (1) by adding terms that are approximately constant for small θs. We believe these terms amount to relatively minor adjustments in the values of the priors and so ignore them.

In all, 150,000 trial steps were calculated for a total computation time of about 16 hours on a DECstation 250 with a DEC Alpha processor running at 266 MHz; 42049 steps were accepted, yielding an acceptance rate of about 28%. Three widely-separated samples from the full MCMC sequence are shown in Fig. 2. While it is not possible to get a quantitative estimate of the uncertainty from these three samples, they provide some indication of the amount of variation in the shapes that occupy the posterior. The amount of waviness observed in the boundary is moderate, as can be observed in Fig. 2. The superfluous waviness compared to the original object is evidence that $\alpha = 5$ is a safe choice, i.e. is weak enough that it does not exert an undue influence on the shape of the MAP solution. Comparison to the configurations from the prior shown in Fig. 1 confirms qualitatively that the posterior is much narrower that the prior.

Visual observation of the replayed MCMC sequence indicates that it takes several hundred steps in the sequence for the boundary to move from plus to minus one standard deviation about the mean position, a distance of a few pixels. Roughly speaking, one might expect that it would take on the order of $[2/(0.06\sqrt{2})]^2 \approx 500$ random steps of rms radial distance $0.06\sqrt{2}$ pixels to move a total distance of two pixels.

A quantitative estimate of the characteristics of the posterior is obtained by averaging over the MCMC sequence. Such an average of the grayscale image of the object is shown in Fig. 3, which is calculated as an overlap of the boundary interior with the pixels of a 512×512 image for increased resolution. Of course, it does not make sense to average the positions of the vertices, because there is nothing to keep the polygon from slipping around the boundary of the object, which has no bearing on the actual object shape. The average MCMC image in Fig. 3 represents the posterior mean image, which may be interpreted as a probability image; the value of each pixel is the posterior probability that the pixel lies inside the boundary of the object. The amount of blur in the edge of the object indicates the variability in the position of that edge allowed by the posterior, i.e. the uncertainty in edge location.

From the measured distance between the 10% and 90% points of the blurred edge of this posterior average, we deduce that the rms uncertainty in edge location varies from about 0.5 pixels to about 1.0 pixels at various positions around the periphery. The smallest rms deviations occur at the limiting edges on the top, bottom, and right sides of the object. The smaller uncertainty at these positions appears to be a consequence of the fact that they are effectively located by the tangential rays of the two projections. The horizontal position of the left edge is not quite as well determined. Of course, it is possible to measure correlations between the uncertainties at different positions.

Another way to summarize the uncertainty in boundary position is to display those pixels in the MCMC average image whose value lies between 0.025 and 0.975. When we do this for the present example, we find that indeed 92% of the original boundary lies inside the 95% credible interval.[14]

Figure 3 also shows that the MAP reconstruction (the model that maximizes the posterior) appears to be consistent with the contour at half the amplitude of the posterior mean image. This result suggests that the posterior probability distribution is symmetric about its maximum. From the shape of the edge profile of the posterior average we tentatively conclude that the posterior may approximately be a multivariate Gaussian distribution, despite the nonlinear relation between the vertex parameters and the measurements.

An important feature of the MCMC technique is that any feature that one wishes to characterize, e.g., the average edge position and its uncertainty in the above example, is not conditional on the other parameters in the model. In the terms of probability theory, MCMC provides marginalized results, which means that the dependence on the uncertainties in "nuisance variables" is integrated out. In the context of the above example, the uncertainty in the edge position deduced for any particular location of the boundary is independent of the uncertainties in edge positions for the rest of the boundary.

7. DISCUSSION

We have presented a brief overview of deformable geometric models. We have demonstrated their usefulness for solving a limit-angle tomographic problem of reconstructing a simple object from two views and have used MCMC to assess the uncertainties associated with the reconstruction.

We have not addressed in much detail the subject of selection of the hyperparameter α. We have chosen α on the basis of the degree of variation observed in the MCMC samples of the posterior, which indicates that the prior is not restricting the smoothness of the boundary very much. Since the value of α affects the width of the prior probability distribution, it influences the posterior uncertainty. Therefore, it is imperative to place more emphasis on understanding how to choose hyperparameters.[15,32] We note that this problem is not confined to Bayesian analysis. As we argued in Sect. 3, choosing α is equivalent to selecting the number of degrees of freedom in the representation of the boundary description. This problem is a common feature of all methods of solving inverse problems, particuarly when they involve finding more model parameters than measurements.

Another limitation of the present study is that α is fixed. A more appropriate approach would be to consider α as a parameter that should be determined from the data, which goes by the name of empirical Bayesian analysis. By including α in the list of variables to be sampled in the MCMC process, the uncertainty in α would be taken into account in the overall uncertainty analysis.

In regard to our use of a polygon to represent a smooth curve, one would ideally like to reduce the length of the polygon sides to a size where this discretization does not affect the results. However, in the present formalism, shorter sides can lead to an increasing frustration of the MCMC algorithm, for example, because of the prior on side length. The result is that one is forced to use increasingly smaller MCMC steps, which leads to a reduced efficiency of the algorithm. One approach to overcoming this problem is to limit the direction of trial steps of each vertex to be along the bisector of angle between neighboring sides, somewhat like the suggestion of Lobregt and Viergever.[13] Another approach avoids the use of the prior on polygon side length by using adaptable discretization, either on an edge-by-edge basis[13] or by routinely resampling the full polygon boundary to create equal-length sides. Another means that warrants investigation is to use a multiresolution representation of the boundary. The advantage would be that one could adjust the size of the MCMC steps at each resolution to improve the overall efficiency of the posterior sampling algorithm.

ACKNOWLEDGEMENTS

We thank John Skilling for helping us realize the need for a Jacobian to transform probabilities on curvature to those on the polygon vertex positions. Many other people have helped us understand MCMC and its potential usefulness, including Julian Besag, James Gubernatis, Richard Silver, and Roger Bilisoly. This work supported by the United States Department of Energy under contract W-7405-ENG-36.

REFERENCES

1. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Inter. J. Comp. Vision* **1**, pp. 321–331, 1988.
2. R. Szeliski, "Probabilistic modeling of surfaces," *Proc. SPIE* **1570**, pp. 154–165, 1991.
3. R. Szeliski and D. Terzopoulos, "Physically-based and probabilistic models for computer vision," *Proc. SPIE* **1570**, pp. 140–152, 1991.
4. K. M. Hanson, G. S. Cunningham, G. R. Jennings, Jr., and D. R. Wolf, "Tomographic reconstruction based on flexible geometric models," in *Proc. IEEE Int. Conf. Image Processing, vol. II*, pp. 145–147, IEEE, 1994.
5. D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, W.H. Freeman, San Francisco, 1982.
6. D. Rueckert and P. Burger, "A multiscale approach to contour fitting for MR images," in *Image Processing*, M. H. Loew and K. M. Hanson, eds., *Proc. SPIE* **2710**, pp. 289–300, 1996.
7. J. C. Gee and D. R. Haynor, "Rapid coarse-to-fine matching using scale-specific priors," in *Image Processing*, M. H. Loew and K. M. Hanson, eds., *Proc. SPIE* **2710**, pp. 416–427, 1996.
8. K. M. Hanson, R. L. Bilisoly, and G. S. Cunningham, "Kinky tomographic reconstruction," in *Image Processing*, M. H. Loew and K. M. Hanson, eds., *Proc. SPIE* **2710**, pp. 156–166, 1996.
9. K. M. Hanson, "The application of protons to computed tomography," *IEEE Trans. Patt. Anal. Mach. Intell.* **PAMI-11**, pp. 567–585, 1989.
10. K. M. Hanson, "Reconstruction based on flexible prior models," in *Image Processing*, M. H. Loew, ed., *Proc. SPIE* **1652**, pp. 183–191, 1992.
11. K. M. Hanson, "Bayesian reconstruction based on flexible prior models," *J. Opt. Soc. Amer. A* **10**, pp. 997–1004, 1993.
12. C. A. Burbeck and S. M. Pizer, "Object representation by cores: indentifying and representing primitive spatial regions," *Vision Res.* **35**, pp. 1917–1930, 1995.
13. S. Lobregt and M. A. Viergever, "A discrete dynamic contour model," *IEEE Trans. Med. Imaging* **MI-14**, pp. 12–24, 1995.
14. K. M. Hanson, G. S. Cunningham, and R. J. McKee, "Uncertainty assessment for reconstructions based on deformable models," submitted to *Int. J. Imaging Systems and Technology* **-**, 1997.
15. D. J. C. MacKay, "Bayesian interpolation," *Neural Computation* **4**, pp. 415–447, 1992.
16. D. S. Sivia, *Data Analysis: A Bayesian Tutorial*, Clarendon, Oxford Univ. Press, New York, 1996.
17. P. H. Garthwaite, I. T. Jolliffe, and B. Jones, *Statistical Inference*, Prentice Hall, London, 1996.
18. J. J. K. Ó Ruanaidh and W. J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*, Springer, New York, 1996.
19. A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, London, 1995.
20. K. M. Hanson and R. N. Silver, eds., *Maximum Entropy and Bayesian Methods*, Kluwer Academic, Dordrecht, 1996.
21. J. Skilling and S. Sibisi, eds., *Maximum Entropy and Bayesian Methods*, Kluwer Academic, Dordrecht, 1996.
22. K. M. Hanson and G. S. Cunningham, "Exploring the reliability of Bayesian reconstructions," in *Image Processing*, M. H. Loew, ed., *Proc. SPIE* **2434**, pp. 416–423, 1995.
23. K. M. Hanson and G. S. Cunningham, "The Bayes inference engine," in *Maximum Entropy and Bayesian Methods*, K. M. Hanson and R. N. Silver, eds., pp. 125–134, Kluwer Academic, Dordrecht, 1996.
24. G. S. Cunningham, K. M. Hanson, G. R. Jennings, Jr., and D. R. Wolf, "An object-oriented optimization system," in *Proc. IEEE Int. Conf. Image Processing, vol. III*, pp. 826–830, IEEE, 1994.

25. G. S. Cunningham, K. M. Hanson, G. R. Jennings, Jr., and D. R. Wolf, "An interactive tool for Bayesian inference," in *Review of Progress in Quantitative Nondestructive Evaluation*, D. O. Thompson and D. E. Chimenti, eds., vol. 14A, pp. 747–754, Plenum, New York, 1995.
26. G. S. Cunningham, K. M. Hanson, G. R. Jennings, Jr., and D. R. Wolf, "An object-oriented implementation of a graphical-programming system," in *Image Processing*, M. H. Loew, ed., *Proc. SPIE* **2167**, pp. 914–923, 1994.
27. J. Besag, P. Green, D. Higdon, and K. Mengersen, "Bayesian computation and stochastic systems," *Stat. Sci.* **10**, pp. 3–66, 1995.
28. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London, 1996.
29. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equations of state calculations by fast computing machine," *J. Chem. Phys.* **21**, pp. 1087–1091, 1953.
30. K. M. Hanson and G. S. Cunningham, "A computational approach to Bayesian inference," in *Computing Science and Statistics* **27**, M. M. Meyer and J. L. Rosenberger, eds., pp. 202–211, Interface Foundation, Fairfax Station, VA 22039-7460, 1996.
31. R. Gordon, R. Bender, and G. Herman, "Algebraic reconstruction techniques for three-dimensional electron microscopy and x-ray photography," *J. Theor. Biol.* **29**, pp. 471–481, 1970.
32. V. Chalana, D. R. Haynor, P. D. Sampson, and Y. Kim, "Parameter estimation in deformable models using Markov Chain Monte Carlo," in *Image Processing*, K. M. Hanson, ed., *Proc. SPIE* **3034**, 1997, to be published.