The Institute For Research In Cognitive Science

Bayesian Approach to the Brain Image Matching Problem

by

J.C. Gee
University of Pennsylvania
L. LeBriquer
C. Barillot
Université de Rennes I
D.R. Haynor
University of Washington
R. Bajcsy
IRCS

University of Pennsylvania
3401 Walnut Street, Suite 400C
Philadelphia, PA 19104-6228

April 1995

Site of the NSF Science and Technology Center for Research in Cognitive Science

University of Pennsylvania
Founded by Benjamin Franklin in 1740

IRCS Report 95-08

Bayesian Approach to the Brain Image Matching Problem[1]

J. C. Gee[†,‡], L. LeBriquer[†], C. Barillot[†], D. R. Haynor[§], and R. Bajcsy[‡]

Laboratoire SIM[†], Faculté de Médecine
Université de Rennes I, 35043 Rennes Cedex, France

Department of Computer and Information Science[‡]
University of Pennsylvania, Philadelphia, PA 19104, USA

Department of Radiology[§]
University of Washington, Seattle, WA 98195, USA

ABSTRACT

The application of image matching to the problem of localizing structural anatomy in images of the human brain forms the specific aim of our work. The interpretation of such images is a difficult task for human observers because of the many ways in which the identity of a given structure can be obscured. Our approach is based on the assumption that a common topology underlies the anatomy of normal individuals. To the degree that this assumption holds, the localization problem can be solved by determining the mapping from the anatomy of a given individual to some referential atlas of cerebral anatomy. Previous such approaches have in many cases relied on a physical interpretation of this mapping. In this paper, we examine a more general Bayesian formulation of the image matching problem and demonstrate the approach on two-dimensional magnetic resonance images.

Keywords: Bayesian modeling, image matching, brain atlases, stochastic estimation.

1. INTRODUCTION

The importance of medical imaging in investigating both the structural and functional integrity of the human brain is well established. The various modalities, taken together, are capable of supplying a range of information from richly detailed, three-dimensional (3-D) representations of anatomic structure to maps of physiological activity within the brain. Despite the wealth of data provided by these images, their interpretation can often become a difficult task for human observers because of the many ways in which the identity of a given structure can be obscured. Of particular concern to our work is the uncertainty introduced by the natural anatomic variability among normal individuals. We aim to improve the accuracy and objectivity of anatomical localization within the human brain with the aid of a "computerized" atlas. Our approach is based on the assumption that a common topology underlies the anatomy of normal individuals: differences among individuals are assumed to represent variations on this underlying anatomical plan. To the degree that this assumption holds, the localization problem can be solved by determining the mapping from the anatomy of a given individual to some referential atlas of cerebral anatomy.

A variety of methods exist to estimate the mappings between brain images, and these can be broadly classified into two categories: techniques that only make global corrections versus techniques

[1]This paper is a revised version of one presented at *SPIE Medical Imaging 1995*, San Diego, CA, February 26–March 2, 1995.

that additionally attempt to account for local shape differences between a pair of brain images—see [1] for a comprehensive survey and [2, 3] for classification schemes applicable to more general image domains. Global registration methods are well suited to multimodality studies of the same subject in which sensor distortions within the region of interest can be neglected in the given application. For the atlas problem, however, there are additional differences in the local morphology of the brain that must be addressed. One effective approach has been to involve expert interaction in the definition of homologous landmarks, from which non-rigid mappings can be inferred [4, 5]. Although these mappings can model localized variations, they do so only within those regions which surround the landmarks.

The recovery of local structural differences between two brain images *throughout* the entire volume of the brain was first studied by Broit [6]. In his approach, the anatomy was considered to be linearly elastic. The mappings thus corresponded to deformations of an elastic body. To accomodate large-scale differences in regional morphology, Bajcsy and Kovačič [7] implemented a multiresolution version of the "elastic matching" method of Broit. We have formulated a Bayesian approach to the brain image matching problem, which represents a generalization of many physically-based methods, such as elastic matching. The immediate and potential advantages of adopting Bayesian analysis in image matching are many, and some of these are described in [8], where additional details about our approach can be found. A primary feature of the method is that a probability distribution for the displacement mappings is made available. Using decision-theoretic techniques, the mappings can be chosen in an optimal way in the sense they minimize some loss function tailored to the problem. In this paper, we examine two different kinds of Bayesian solutions that can be derived from the distribution for the mappings, and illustrate our Bayesian approach to image matching with their development.

2. METHODS

To formulate the image matching problem from a decision-theoretic point of view [9], the parameter space Θ is identified with the set of all possible mappings. For three-dimensional images, $\Theta = (\Re^3)^N$, where N is the number of voxels. In estimation problems, such as the present one, the action space comprises the estimates and is therefore equivalent to the parameter space. The choice of a particular estimate is based on its expected loss with respect to the posterior probability for the unknown mapping. In the current work, the performance of the maximum a posteriori (MAP) and minimum mean squared error (MMSE) estimators are examined. The posterior contains all the available information about the unknown mapping. This information is of two kinds. The first is prior information about the mappings, expressed in terms of a probability distribution over the parameter space. The second kind of information is that provided by the sample and is modeled by the likelihood function. The posterior is obtained by combining the prior and likelihood using Bayes's law.

2.1 Likelihood

The likelihood models the degradation process by which the observations or sensed data are obtained and for this reason is also known as the degradation or sensor model. In image matching, however, the measurements are not degraded samples of the unknown mapping. The likelihood of a mapping must be inferred from an observed image pair or their features. This is accomplished by measuring the degree to which one image is made similar to the other through the mapping. These measurements of similarity may be considered as output from a virtual sensor; therefore, like any other measurement device, its error can be modeled [10].

We consider three different similarity measures, two suitable for matching raw intensity data and a third for matching tissue-classified data. Specifically, the normalized cross-correlation is used to define one of the sensor models. In this case, the measurement error at a given point is estimated by approximating with a quadratic the error surface defined by the correlation values computed over a set of possible displacement values; that is, given a point $\mathbf{x}$ in one image, we first determine its cross-correlation value for each point within a 3×3 neighborhood centered at $\mathbf{x} + \boldsymbol{\delta}^*$ in the second image, where $\boldsymbol{\delta}^*$ is the current estimate of the unknown mapping. The quadratic approximation to this collection of error values is then obtained by solving an equal-weighted least squares fitting problem [11]. Finally, the optimum and inverse Hessian of the resultant quadratic are used as the measurement and its variance, respectively. A similar error analysis was performed by Szeliski [10] for optical flow measurements based on the sum of squared differences (SSD) technique of Anandan [12]. The SSD measure is a simplified version of normalized cross-correlation.

Our second intensity-based similarity measure approximates SSD with the windowing function reduced to a single point[2]; its associated likelihood is expressed as follows:

$$p(\{I_T, I_R\}|\boldsymbol{\delta}) \propto \prod_{\mathbf{x}\in\Omega_R} \exp\{-\frac{1}{2\sigma_x^2}|I_T(\mathbf{x}+\boldsymbol{\delta}) - I_R(\mathbf{x})|\}, \tag{1}$$

where Ω_R is the image domain of the brain atlas, and I_R and I_T represent the atlas and the brain image of the subject, respectively. The derivation of the above likelihood assumes that the pair of images represent displaced versions of the same image corrupted with additive white Gaussian noise [10].

To match tissue-classified data, we use again the SSD approximation, except the images now correspond to fields of classification vectors. Because the classification error is normally distributed in the method of Lachmann [13] used in our experiments, the assumption of Gaussian noise is appropriate.

2.2 Prior

Given the likelihood function, which expresses the probability of observing the image pair given a particular value of the mapping, a natural choice for our solution is the value which maximizes the likelihood. The problem of calculating the maximum likelihood estimate (MLE), however, is ill-posed: the matrices which arise in the numerical solution will be singular. In image matching, the information introduced by the prior is essential: it regularizes the MLE problem, producing matrices that are invertible.

For the experiments in Section 3, we chose a Gibbs prior whose energy function is the internal strain energy of a linearly elastic body [14]. The linear elastic strain energy is a kind of first order Tikhonov stabilizer, closely related to the membrane energy [15]. It imposes a smoothness constraint on the solution. The use of this particular prior stems from our previous work in elastic matching [7] and is by no means essential to the present formulation of the matching problem. Nevertheless, the admission of mappings with only C^0 continuity is appealing because the computational complexity is substantially less than that involved in imposing higher order continuity on our mappings. At the same time, the accuracy of the results obtained with these mappings do not appear to be compromised, at least, in the evaluation studies reported in [16, 17, 18]. There is currently no empirical evidence to suggest that higher order smoothness constraints are absolutely necessary in brain image matching. A more important consideration in many applications of matching appears to be the prior modeling of

[2]The intensity difference in Eqn. 1 is not squared as would be the case for the SSD measure.

discontinuities in the mapping. The same concern arises in the solution of a number of inverse visual problems[3] and the modeling techniques that have been developed to address them—see, for example, [15]—are equally applicable here.

2.3 MAP Estimation

The prior and likelihood are combined using Bayes's law to form a posterior distribution for the unknown mapping, which expresses the probability of any mapping $\boldsymbol{\delta}$ given the observed data $\boldsymbol{z}$:

$$\pi(\boldsymbol{\delta}|\boldsymbol{z}) = \frac{f(\boldsymbol{z}|\boldsymbol{\delta})\pi(\boldsymbol{\delta})}{\int_\Theta f(\boldsymbol{z}|\boldsymbol{\delta})\, d\pi(\boldsymbol{\delta})},$$

where f and π are the likelihood and prior, respectively. The prior models our certainty about the mappings in the absence of sample information, whereas the posterior represents our revised beliefs in view of the observed data. We can now consider as our estimate the "most likely" value of the unknown mapping given *both* the prior and sample information.

The problem of determining the maximum a posteriori mapping is made difficult because the posterior distributions are typically multimodal. Our algorithm iteratively searches for the solution by following the currently steepest path over multiple resolution levels. Its implementation has previously been reported in [19] so only a brief overview of the algorithm is provided in the following. To illustrate its operation, consider the following distribution,

$$p(\boldsymbol{\delta}|\{I_T, I_R\}) \propto \exp -U(\boldsymbol{\delta}, \{I_T, I_R\}),$$

with energy equal to

$$U(\boldsymbol{\delta}, \{I_T, I_R\}) = U_{\text{likelihood}}(\boldsymbol{\delta}, \{I_T, I_R\}) + U_{\text{prior}}(\boldsymbol{\delta}),$$

where $U_{\text{likelihood}}$ and U_{prior} are the Gibbs energy functionals corresponding to the likelihood and prior distributions, respectively. The MAP estimate is obtained by minimizing U, a task complicated by the fact that $U_{\text{likelihood}}$ is nonlinear in the unknowns—see Eqn. 1, for example. Our strategy is to replace the nonlinear term at each iteration by either a linearized or quadratic approximation[4]. The optimization algorithm performed at each resolution level can be summarized as follows:

```
Iterate {
    Replace nonlinear likelihood energy by its linearized or quadratic approximation at the current estimate δ*;
    Add approximation to exact quadratic expression for the prior;
    Find the global minimum δ* of the resultant quadratic;
    Find the true minimum of U along the line segment between the previous and current estimates;
    Set the new estimate δ to the line-minimized value;
    Exit when a stable estimate is reached;
}
```

[3]Our image matching problem is conceptually similar to the inverse problems of stereo matching and optical flow estimation, and is related to visible-surface reconstruction.

[4]Using the quadratic approximation is equivalent to modeling with a Gaussian the measurement error at a point, in the same way the likelihood is constructed for the cross-correlation measure.

Note that the solutions to consecutive approximations are used to define a direction along which to search. The subsequent line-minimization determines the new estimate. In certain instances, this step may allow the search to fortuitously escape local minima that would otherwise trap algorithms such as iterated conditional mode.

2.4 MMSE Estimation

Despite its widespread use, the MAP estimator may not be appropriate in situations where some degree of error tolerance is acceptable or desirable. This is due to the fact that the MAP estimator is the Bayes action[5] for the zero-one loss function, which assigns the same increased cost to all estimates other than the correct one. The existence of the posterior enables the use of alternative loss functions, such as the squared-error loss for unknown θ and action a:

$$L(\theta, a) = (\theta - a)^2,$$

whose Bayes action is the mean of the posterior distribution. This action is the minimum mean squared error (MMSE) estimator and is of interest for several reasons. It is a more "robust" estimator than the MAP in the sense that fluctuations in the posterior do not affect its value in the dramatic way in which the MAP estimate may be changed. It is suitable for situations in which the loss is symmetric in $(\theta - a)$. In classical analysis, the squared-error loss arises in the determination of minimum variance unbiased (MVU) estimators and is related to least squares procedures. The generalization of squared-error loss to vector quantities is known as the quadratic loss:

$$L(\theta, a) = (\theta - a)^{\mathrm{T}} \mathbf{Q} (\theta - a),$$

where θ and a are vectors, and $\mathbf{Q}$ a positive definite matrix. The Bayes action in this case is still the posterior mean.

The MMSE estimate can be calculated by Monte Carlo integration. Specifically, the Gibbs sampler was used to sample from the posterior distribution [20]. In [8], we show how to construct an efficient continuous state space Gibbs sampler suitable for sampling distributions with a quadratic energy function. The development capitalizes on our use of a finite element approximation to the mapping, which leads to Markov random field (MRF) models of the Gibbs distributions. Typically, to sample the local conditional distributions of an MRF, a discrete approximation to the state space is made, but this approach introduces quantification errors in the displacement estimates. These errors cannot be made insignificant through refinement of the unit of discretization because the computational complexity grows rapidly with the size of the state space. Given the Gibbs sampler in [8], our algorithm to estimate the posterior mean can be described as follows:

```
Iterate {
    Calculate quadratic approximation to the original posterior energy at the current sample;
    Sample the Gaussian distribution corresponding to this quadratic;
    Collect sample information needed to compute statistics of the original distribution;
}
```

[5] Recall our estimates or actions are chosen to minimize the expected loss with respect to the posterior distribution. Such an action is called a Bayes action.

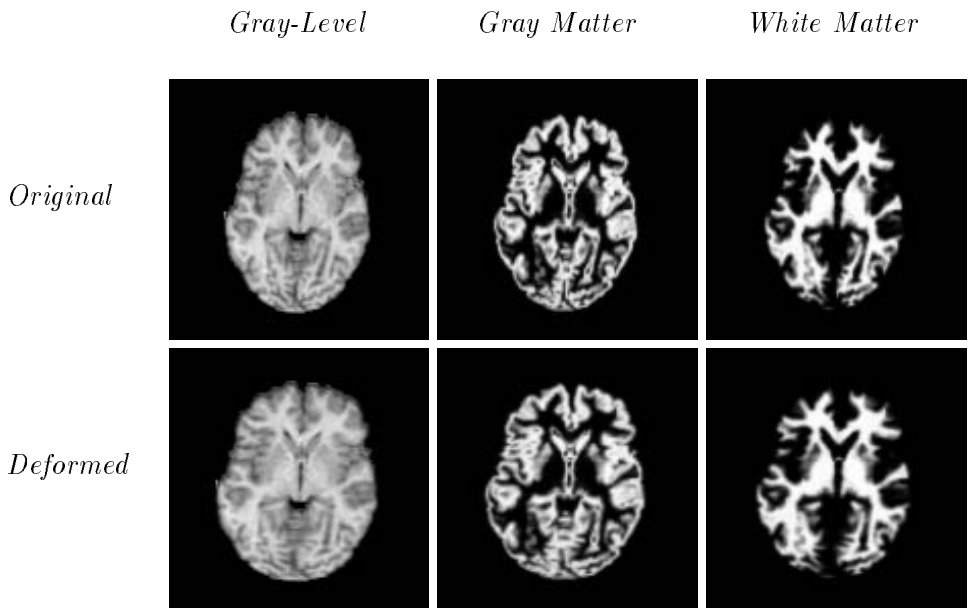

Figure 1: Original tissue-classified data (top row) and its deformed version (bottom row).

3. EXPERIMENTAL RESULTS

Some results from preliminary experiments with two-dimensional data are presented in this section to demonstrate our approach. In every case, the finite element mesh was uniform: 4-noded bilinear quadrilateral elements were used, each being square with sides of length equal to 7 pixels. The 2-point Gauss-Legendre quadrature rule was employed to numerically integrate the various integrals that arise in the finite element implementation. Unlike the linearized approximation used in [19], the Gibbs energy functional for the likelihood was modeled by a quadratic in the following set of experiments. The boundaries of the images were fixed, and matching was performed only at the original resolution level.

The data used in the first set of experiments are shown in Figure 1. The 128×128 sections (1.87 mm $\times$ 1.87 mm) in the top row of the figure were extracted at the level of the anterior commissure-posterior commissure (AC-PC) plane from a brain volume that had previously been segmented into its major tissue components [13]. These images were then deformed using a known thin-plate spline mapping. The spline was inferred from a set of manually defined tie points. The result is shown in the second row of the figure, where the deformed version of each component of the original tissue-classified data is displayed immediately below the image from which it was derived. The objective of the experiment was to determine the accuracy with which the different Bayesian solutions matched each of the original images to its deformed version.

The results of the MAP estimation using the likelihood functions based on cross-correlation and for tissue-classified data are displayed in the middle and rightmost columns of Figure 2, respectively. The images in the leftmost column were obtained by subtracting each of the original images (top row

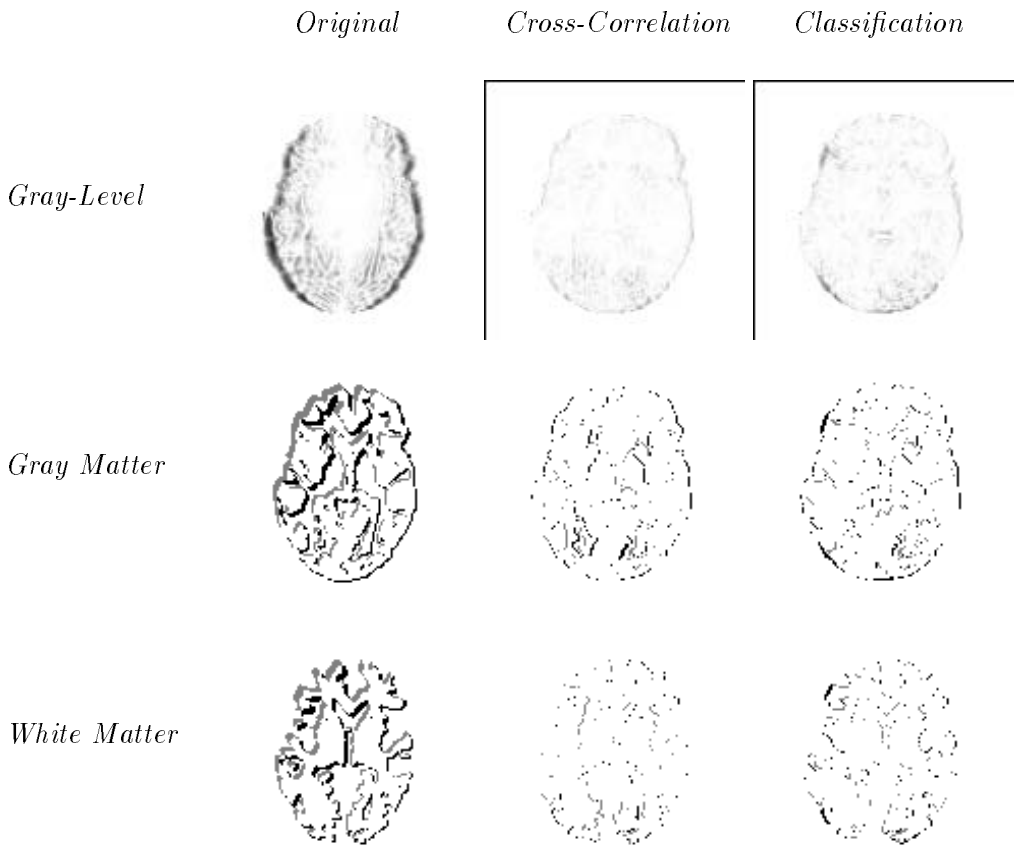

Figure 2: Misalignment before and after matching, by MAP estimation, the data shown in Figure 1. The images in the leftmost column depict the "difference" between each of the original images (top row of Figure 1) and its deformed version (bottom row of Figure 1). The middle and rightmost columns show, respectively, the difference results after matching with cross-correlation and the likelihood for tissue-classified data.

of Figure 1) with its deformed version (bottom row of Figure 1)[6]. These images therefore depict the misalignment in the data set that the matching algorithm has to correct. In the difference images shown in the middle and bottom rows, the gray colored pixels represent points in the original/matched tissue component that do not overlap with any point in its deformed version, whereas the dark pixels correspond to points in the deformed version unmatched by any point in the original/matched tissue component. The aim is to reduce the number of gray colored pixels, that is, to match as much of the original classification image as possible to its deformed version.

To quantify the quality of the results, the relative overlap between the matched and deformed version of each classification component was determined by measuring the ratio between the area of their intersection and the area of their union. This measure assigns an additional penalty for portions of the matched structure which incorrectly label their underlying anatomy in the image as part of the structure of interest. For the data set in Figure 1, the overlap values before matching were 0.66 and

[6]The points of the classification images are given a binary label prior to the subtraction operation

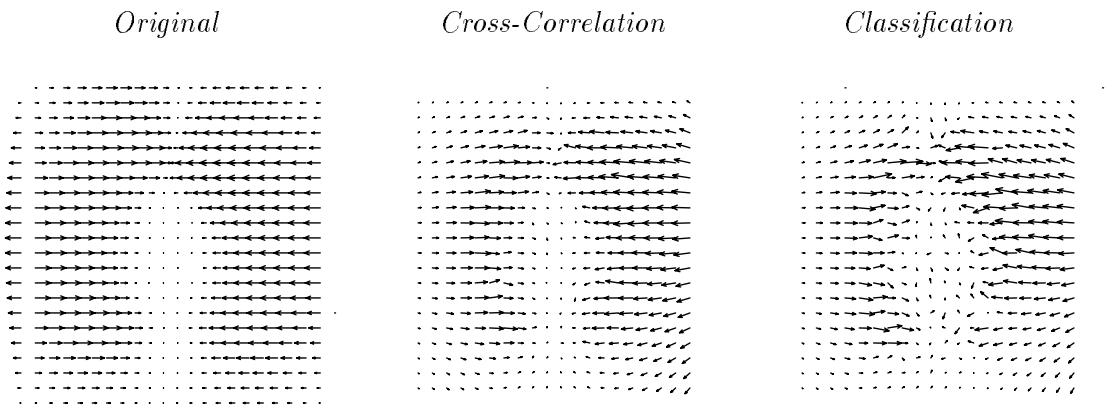

Figure 3: The original thin-plate spline mapping (left) and the MAP estimates obtained by cross-correlation (middle) and classification (right) matching.

0.68 for the gray and white matter components, respectively. After matching using the likelihood for cross-correlation, the overlap for the gray matter was 0.91, whereas the white matter overlap improved to 0.93. Similar results were obtained with the likelihood for tissue-classified data: the gray and white matter overlap were both equal to 0.91 after matching. To further illustrate the quality of the recovered mappings, Figure 3 depicts the estimated mappings together with the original thin-plate spline with which the test data were created.

In addition to the MAP estimate, the posterior mean was stochastically estimated and used to deform the original images. The results are shown in Figures 4 and 5. For cross-correlation, the overlap values of the gray and white matter components were both equal to 0.87. The same overlap values were obtained when the likelihood for tissue-classified data was used.

The comparison above between the MAP and MMSE estimates favors the former because the two likelihood functions produce posterior distributions that are fairly "smooth." This is a consequence of the robustness of the cross-correlation measure to white noise in the data. When matching classification vectors, the data itself is fairly "free" of noise because the very success of any classification algorithm depends on its ability to account for ambiguity in the original data. To distinguish more clearly between the performance of the MAP and MMSE estimators, two additional data sets were created and the likelihood of Eqn. 1 was used.

The first data set was obtained by adding different amounts of white Gaussian noise ($\sigma = 15$ or $\sigma = 30$) to the deformed gray-level image shown in Figure 1. For the second data set, the deformed image was first blurred with a Gaussian filter ($\sigma = 1$ pixel) and then corrupted with additive white Gaussian noise ($\sigma = 15$ or $\sigma = 30$); this data set is shown in Figure 6.

The results for the noise-corrupted ($\sigma = 15$) but blur-free data set are displayed in Figure 7. The overlap values obtained by MAP estimation were 0.84 for both the gray and white matter components. For the MMSE estimate, the overlap values for both classification components were equal to 0.83. In matching the data set where σ of the additive noise was equal to 30, the MMSE estimator (gray and

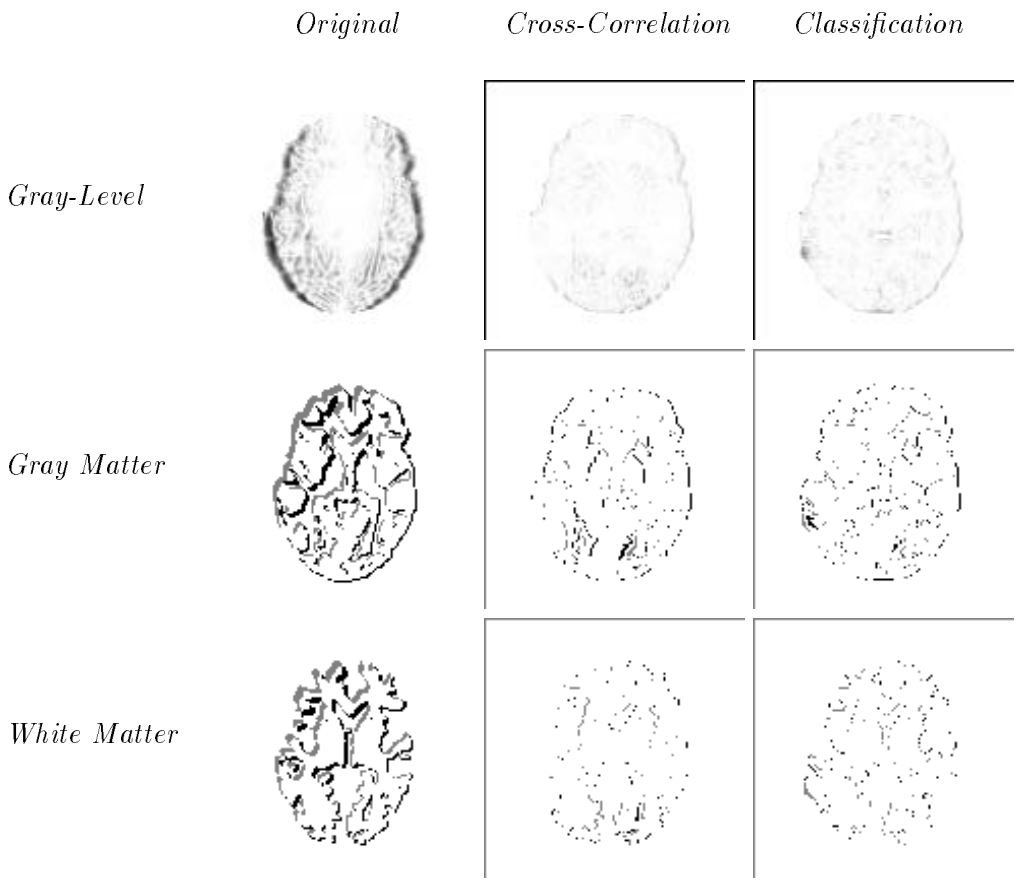

Figure 4: Misalignment before and after matching, by MMSE estimation, the data shown in Figure 1. The images in the leftmost column depict the "difference" between each of the original images and its deformed version. The middle and rightmost columns show, respectively, the difference results after matching with cross-correlation and the likelihood for tissue-classified data.

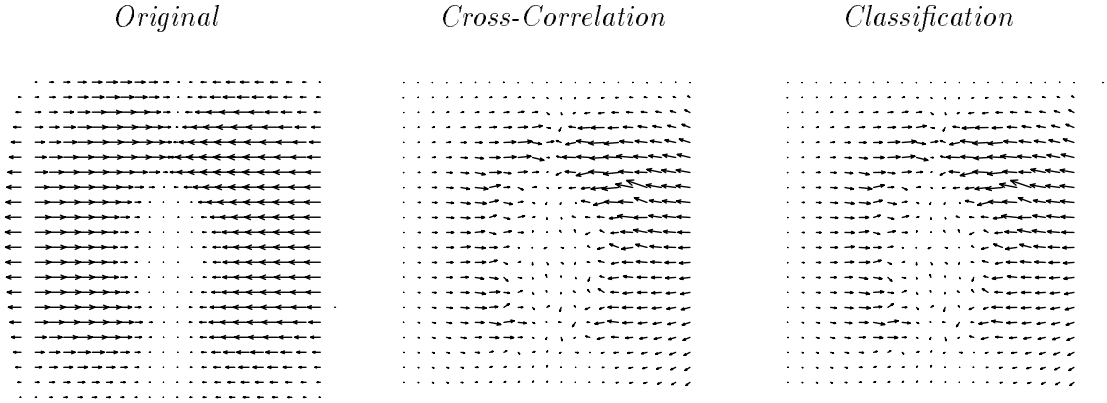

Figure 5: The original thin-plate spline mapping (left) and the MMSE estimates obtained by cross-correlation (middle) and classification (right) matching.

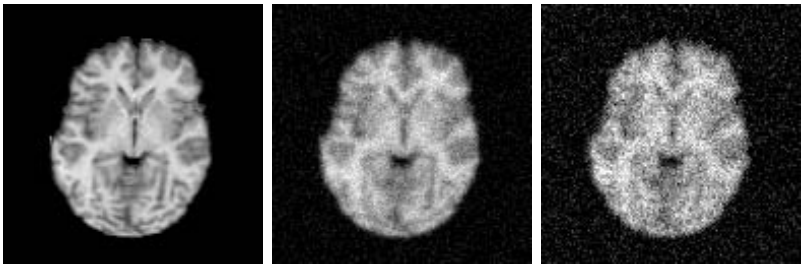

Figure 6: Deformed gray-level image (left) and its degraded versions, obtained by first blurring the image with a Gaussian filter ($\sigma = 1$ pixel) and then adding white Gaussian noise with $\sigma = 15$ (middle) or $\sigma = 30$ (right). The deformed image is the same one shown in Figure 1.

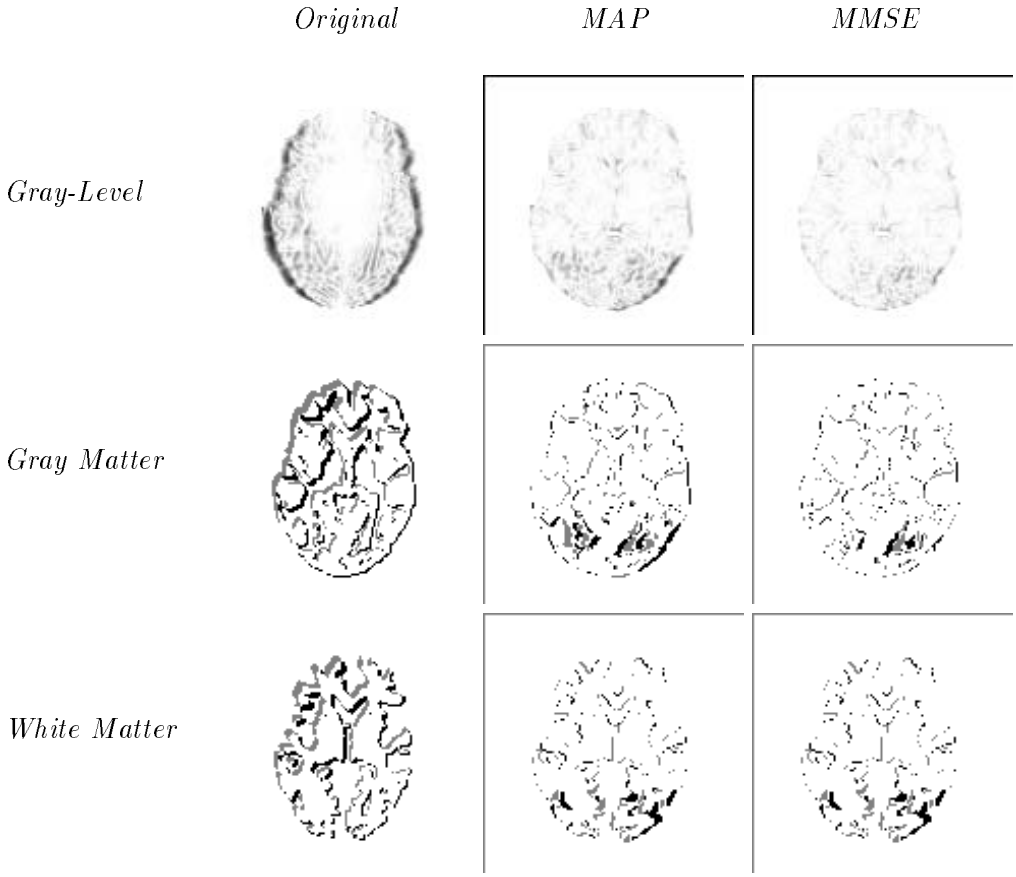

Figure 7: Comparison of the MAP and MMSE estimators for the noise-corrupted ($\sigma = 15$) but blur-free version of the data shown in Figure 1. The images in the leftmost column depict the "difference" between each of the original images and its deformed version. The middle and righmost columns show, respectively, the difference results after matching by MAP and MMSE estimation. Both estimates were obtained using the likelihood corresponding to the SSD-like measure.

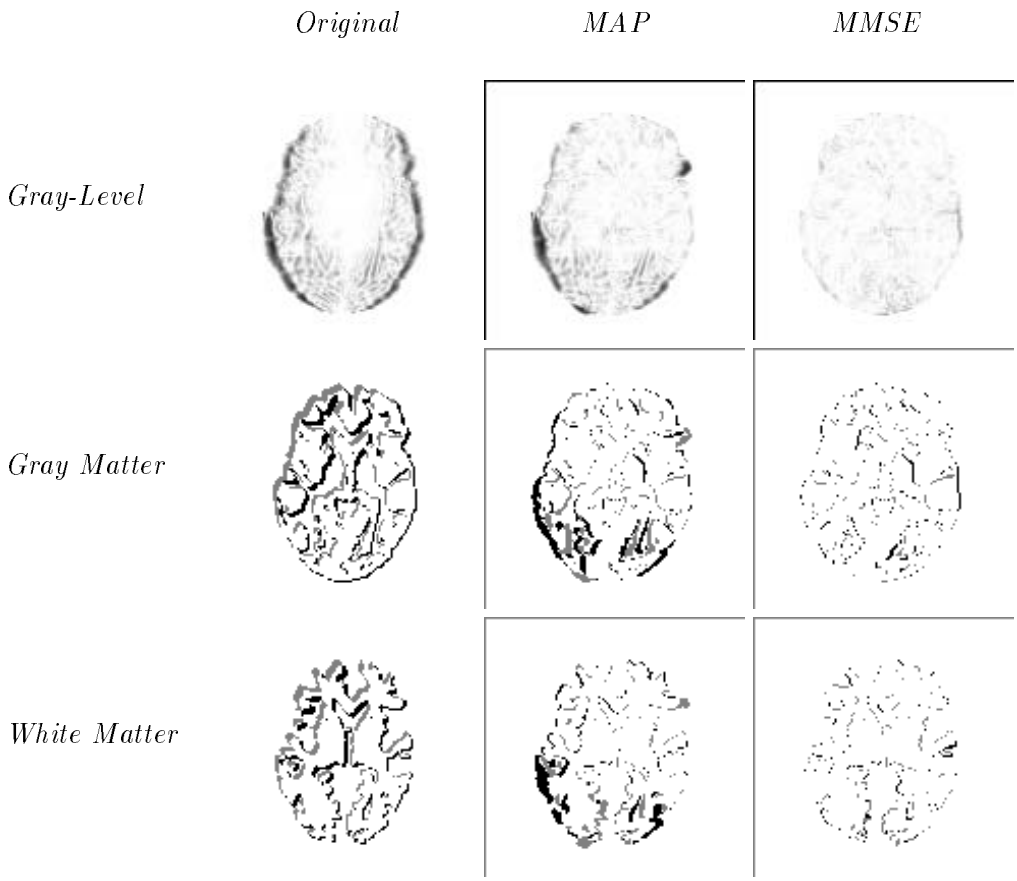

Figure 8: Comparison of the MAP and MMSE estimators for the blurred and noise-corrupted ($\sigma = 15$) data shown in Figure 6. The images in the leftmost column depict the "difference" between each of the original images and its deformed version. The middle and righmost columns show, respectively, the difference results after matching by MAP and MMSE estimation. Both estimates were obtained using the likelihood corresponding to the SSD-like measure.

white matter overlap were 0.80 and 0.79, respectively) performed only slightly better than MAP (gray and white matter overlap were 0.77 and 0.78, respectively). However, the degradation in its performance with the increase in noise, as measured by the overlap values, was less than that of the MAP estimator.

The superiority of the MMSE estimator is more pronounced in one set of the results obtained for the blurred and noise-corrupted data set. Specifically, for the data set where σ of the Gaussian noise was equal to 15, the overlap values that the MMSE estimate obtained for the gray and white matter components were 0.86 and 0.84, respectively. In comparison, the gray and white matter overlap values obtained by the MAP estimate were 0.77 and 0.78, respectively. The corresponding difference images are shown in Figure 8. However, when the data with increased noise ($\sigma = 30$) were processed, the overlap results were similar for both estimators, with the MAP actually producing the best match of the white matter component (0.80 versus 0.78 for MMSE). The gray matter overlap values were equal to 0.79 for both estimators.

4. DISCUSSION

Based on the experimental results reported in the previous section and additional data collected in separate more extensive studies, the performance of the MAP and MMSE estimators was not found to be significantly different. However, our data sets were generated by applying manually defined distortions to MRI images. The results are as yet unavailable for the practical situation in which image volumes derived from different individuals are to be matched. For real data, there are additional issues to consider which may have a significant effect on the results, such as the initial guess used for each of the estimation algorithms. Also, the type of element and configuration of the mesh will have an influence on the performance of the estimators. Nevertheless, some observations of general interest were made and we can comment on aspects related to the implementation of both estimators.

For the SSD-like measure, it was found, as expected, that the MMSE-estimated displacement mappings were in most cases smoother than those produced by MAP estimation. The discrepancy visible near the borders of the motion fields in Figures 3 and 5 is also apparent in the mappings estimated with the SSD-like measure. This effect is due to the zero displacement boundary condition that was imposed on the solutions to simplify the current evaluation. The idea is that it is only important for those parts of the estimated mapping that displace points belonging to the brain to resemble the thin-plate spline mapping. This is indeed the case in many of the results.

Each iteration of our "experimental" version of the MAP estimation algorithm typically required 1 to 3 minutes to complete on a Sun 4 workstation depending on the likelihood function used, with the bulk of the time spent on line-minimization. Estimates based on a linearized approximation to the likelihood term normally converged within 4 or 5 iterations. After the same number of iterations, the estimation using a quadratic approximation to the likelihood is also typically near convergence, but its value usually does continue to change by a very small amount in subsequent iterations. In such cases, the number of iterations was fixed to 10. For the data sets examined in our experiments, neither a decrease in the size of the element nor the use of a higher order interpolation scheme within the element produced significantly better results. Only a small improvement in the results was obtained by increasing the number of Gauss sampling points to perform numerical integration.

The posterior mean was estimated using 300 samples: a complete sweep through the "image"[7] was performed between consecutive samples. The average amount of time for an image visit was approximately 14 seconds on a Sun Sparc 2 workstation. We avoided the use of a discrete state space for reasons of computational efficiency and quantitative accuracy. However, by approximating the local distributions with a continuous quadratic, the samples are no longer taken from the true distribution. The effect of this approximation on our stochastic estimates—at least, in comparison with the values obtained with a discrete state space—needs to be studied further.

In summary, we have posed the problem of matching brain images in terms of Bayesian estimation. The existence of a posterior distribution for the mappings makes possible a range of analyses, including the use of alternative loss functions as was demonstrated with our comparison of the MAP and MMSE estimators. But the flexibility of designing loss functions suited to a specific problem domain is only one "standard" feature of the Bayesian approach. More generally, Bayesian modeling has enabled the development of novel algorithms to account for intensity distortions effected by inhomogeneity in the

[7] The sites of our MRF lattice correspond to the nodes in the finite element discretization of the image domain and their number is only a fraction of the total number of pixels in the image. For example, the number of nodal unknowns that had to be estimated in the experiments of Section 3 was 484.

radio frequency field over the MR imaging volume [21, 22]. Correction for the shading effect has led to improved segmentation of MR data and can similarly benefit the matching of MR images [19]. Statistical information collected from past samples about morphological variation is naturally modeled as prior information in the Bayesian method and can thus be used to guide the matching process. As with loss functions, priors can be constructed to meet the requirements of a given application; for example, it may be important to allow discontinuities in the mappings for studies involving structural pathology. Beyond the estimation of our mappings, Bayesian analysis can provide answers to inquiries about the estimates themselves, such as their uncertainty or reliability [8, 23].

5. ACKNOWLEDGMENTS

The authors are grateful to the Region Council of Brittany for their financial support of this work. We thank Pr. J. M. Scarabin and the Radiology Department of the University Hospital of Rennes for their assistance in obtaining the MRI data sets.

6. REFERENCES

[1] C. R. Maurer Jr. and J. M. Fitzpatrick, "A review of medical image registration" in *Interactive Image-Guided Neurosurgery*, R. J. Maciunas, ed., Park Ridge, IL: American Association of Neurological Surgeons, pp. 17–44, 1993.

[2] P. A. Van den Elsen, E.-J. D. Pol, and M. A. Viergever, "Medical image matching—A review with classification," *IEEE EMBS*, vol. 12, pp. 26–39, 1993.

[3] L. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, pp. 325–376, 1992.

[4] A. C. Evans, C. Beil, S. Marrett, C. J. Thompson, and A. Hakim, "Anatomical-functional correlation using an adjustable MRI-based region of interest atlas with positron emission tomography," *J Cereb Blood Flow Metab*, vol. 8, pp. 513–530, 1988.

[5] F. L. Bookstein, "Thin-plate splines and the atlas problem for biomedical images," in *Proc Information Processing in Medical Imaging*, A. C. F. Colchester and D. J. Hawkes, eds., Berlin:Springer-Verlag, LNCS 511, pp. 326–342, 1991.

[6] C. Broit, *Optimal registration of deformed images*. Doctoral dissertation, Department of Computer and Information Science, University of Pennsylvania, PA, 1981.

[7] R. Bajcsy and S. Kovačič, "Multiresolution elastic matching," *Comput Vision, Graphics, Image Process*, vol. 46, pp. 1–21, 1989.

[8] J. C. Gee, L. Le Briquer, C. Barillot, and D. R. Haynor, "Probabilistic matching of brain images," to appear in *Proc Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, eds., Dordrecht:Kluwer Academic Publishers, 1995.

[9] Berger JO. *Statistical Decision Theory and Bayesian Analysis*. New York:Springer-Verlag, 1985.

[10] R. Szeliski, *Bayesian Modeling of Uncertainty in Low-level Vision*. Norwell, MA:Kluwer Academic Publishers, 1989.

[11] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision: Volume I.* Reading, MA:Addison-Wesley, 1992.

[12] P. Anandan, "Computing dense displacement fields with confidence measures in scenes containing occlusion," in *Proc Image Understanding Workshop*, Miami Beach, FL:Science Applications International Corporation, pp. 186–196, 1984.

[13] F. Lachmann and C. Barillot, "Brain tissue classification from MRI data by means of texture analysis," in *Proc Medical Imaging VI: Image Processing*, M. H. Loew, ed., Bellingham, WA:SPIE, vol. 1652, pp. 72–83, 1992.

[14] S. Timoshenko and J. N. Goodier, *Theory of elasticity.* New York:McGraw-Hill, 1951.

[15] D. Terzopoulos, "Regularization of inverse visual problems involving discontinuities," *IEEE PAMI*, vol. 8, no. 4, pp. 413–424, 1986.

[16] R. Dann, J. Hoford, S. Kovačič, M. Reivich, and R. Bajcsy, "Evaluation of elastic matching system for anatomic (CT, MR) and functional (PET) cerebral images," *J Comput Assist Tomogr*, vol. 13, pp. 603–611, 1989.

[17] J. C. Gee, M. Reivich, L. Bilaniuk, D. Hackney, R. Zimmerman, S. Kovačič, and R. Bajcsy, "Evaluation of multiresolution elastic matching using MRI data," in *Proc Medical Imaging V: Image Processing*, M. H. Loew, ed., Bellingham, WA:SPIE, vol. 1445, pp. 226–234, 1991.

[18] D. L. Collins, A. C. Evans, C. Holmes, and T. M. Peters, "Automatic 3D segmentation of neuroanatomical structures from MRI," to appear in *Proc Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, eds., Dordrecht:Kluwer Academic Publishers, 1995.

[19] J. C. Gee, D. R. Haynor, M. Reivich, and R. Bajcsy, "Finite element approach to warping of brain images," in *Proc Medical Imaging 1994: Image Processing*, M. H. Loew, ed., Bellingham, WA:SPIE, vol. 2167, 1994.

[20] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," *IEEE PAMI*, vol. 6, no. 6, pp. 721–741, 1984.

[21] M. X. H. Yan and J. S. Karp, "An adaptive Bayesian approach to three-dimensional MR brain segmentation," to appear in *Proc Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, eds., Dordrecht:Kluwer Academic Publishers, 1995.

[22] W. M. Wells III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Statistical intensity correction and segmentation of MRI data," in *Visualization in Biomedical Computing 1994*, R. A. Robb, ed., Bellingham, WA: SPIE, vol. 2359, pp. 13–24, 1994.

[23] K. M. Hanson and G. S. Cunningham, "Exploring the reliability of Bayesian reconstructions," to appear in *Medical Imaging 1995: Image Processing*, M. H. Loew, ed., Bellingham, WA:SPIE, vol. 2434, 1995.