January 1994

Finite Element Approach to Warping of Brain Images

James C. Gee
University of Pennsylvania

D. R. Haynor
University of Washington

Martin Reivich
University of Pennsylvania

Ruzena Bajcsy
University of Pennsylvania

Follow this and additional works at: http://repository.upenn.edu/cis_reports

Recommended Citation
http://repository.upenn.edu/cis_reports/276


This paper is posted at ScholarlyCommons. http://repository.upenn.edu/cis_reports/276
For more information, please contact repository@pobox.upenn.edu.
Finite Element Approach to Warping of Brain Images

Abstract
A probabilistic approach to the brain image matching problem is proposed in which no assumptions are made about the nature of the intensity relationship between the two brain images. Instead the correspondence between the two intensities is represented by a conditional probability, which is iteratively determined as part of the matching problem. This paper presents the theory and describes its finite element implementation. The results of preliminary experiments indicate that there remain several aspects of the algorithm that require further investigation and refinement.

Comments
Finite Element Approach to Warping of Brain Images

MS-CIS-94-11
GRASP LAB 370

J.C. Gee
D.R. Haynor
M. Reivich
R. Bajcsy

University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

March 1994
Finite element approach to warping of brain images

J. C. Gee\textsuperscript{1}, D. R. Haynor\textsuperscript{2}, M. Reivich\textsuperscript{3}, and R. Bajcsy\textsuperscript{1}

Department of Computer and Information Science\textsuperscript{1}
University of Pennsylvania
Philadelphia, PA 19104

Department of Radiology\textsuperscript{2}
University of Washington
Seattle, WA 98195

Cerebrovascular Research Center\textsuperscript{3} of the Department of Neurology
University of Pennsylvania
Philadelphia, PA 19104

ABSTRACT

A probabilistic approach to the brain image matching problem is proposed in which no assumptions are made about the nature of the intensity relationship between the two brain images. Instead the correspondence between the two intensities is represented by a conditional probability, which is iteratively determined as part of the matching problem. This paper presents the theory and describes its finite element implementation. The results of preliminary experiments indicate that there remain several aspects of the algorithm that require further investigation and refinement.

1. INTRODUCTION

The anatomy of normal individuals can be thought of as quantitative variations on a common underlying qualitative plan: for example, we all have two eyes, a nose, a rough symmetry midplane, etc. A natural way of using this prior knowledge about human anatomy is to imagine that a given individual's anatomy (the test anatomy) is a warped version of some universal, or reference, anatomy. If this warping function could be calculated, many useful applications are possible. The warping function yields direct knowledge about organ size and shape. A physiologically defined region of interest (ROI) can be automatically related to a standard anatomic ROI (which need be defined only once) in the reference anatomy, facilitating comparison between individuals. The existence of the reference anatomy also makes possible an approach to tissue classification and segmentation which is global, rather than local, in character. The fundamental problem is to define the class of allowable deformations (which must be broad enough to enable the reference anatomy to fit all test anatomies) and to come up with efficient, automated algorithms for the calculation of the appropriate deformation. In previous meetings we presented a system that "elastically" matches points of two three-dimensional (3-D) image volumes \[1, 2\]. The system uses a multiresolution coarse-to-fine strategy to search for regions with informative features, i.e. steep intensity gradients [3]. The goal of our current research is to generalize the above approach by considering the brain image warping problem within a probabilistic framework.

To illustrate the warping problem, consider two datasets obtained via either magnetic resonance imaging (MRI) or x-ray computed tomography (CT). Let one represent the test anatomy and the other, the reference anatomy. Suppose that for each point \( r \) of the reference domain \( \Omega_R \), we know the reference intensity (possibly a vector), \( I_R(r) \), and that for each point \( t \) of the test domain \( \Omega_T \), we know the test intensity \( I_T(t) \). The matching problem is to find the displacement field \( \Phi : \Omega_R \rightarrow \Omega_T \) that maps each point \( r \) into a corresponding point \( \Phi(r) \) in \( \Omega_T \). The problem is made more difficult because, due to the fact that \( I_R \) and \( I_T \) are acquired at different times, possibly by different modalities, etc., the exact correspondence between \( I_R \) and \( I_T \) values may not be known; in fact, the determination of this grayscale mapping function is an additional part of the problem. We represent this mapping by a conditional probability density \( P(I_T|I_R) \), which is determined simultaneously with the mapping \( \Phi \). We provide the user with the option of specifying a small set of tie points, i.e. points where \( \Phi(r) \) is prescribed. The problem of finding \( \Phi \) may then be put in a probabilistic framework: find \( \Phi \) and \( P \) to minimize
where

\[ x \text{ is a generic point in } \mathbb{R}^3 \text{ or } \mathbb{R}^2, \text{ where in the latter case the integrals are taken over areas instead of volumes,} \]

\[ \delta(x_1, x_2, x_3) = \delta(x) \text{ is a vector-valued deformation or displacement,} \]

\[ x' = \Phi(x) = R(x + \delta(x)) \text{ is the correspondence induced by the local deformation } \delta \text{ and the global affine map } R; \]

\[ R(x)_i = \sum_j R_{ij} x_j + C_i, \]

\[ (p_i, p'_i) \text{ is a set of user-defined (approximately) corresponding points, where the true location of } \Phi(p_i) \text{ is assumed to have a distribution centered at } p'_i, \text{ with uncertainty measured by the Gaussian with width } \eta_i, \text{ and} \]

\[ \sigma \text{ and } \epsilon \text{ are the stress and strain vectors respectively—for simplicity we assume the deformations are linearly elastic.} \]

The first term in the expression above represents the logarithm of the likelihood of observing the given test intensity at a point, given the reference intensity at the corresponding point, integrated over the points in the reference image. The second term represents the sum over the user-specified tie points \( p_i \), which must be mapped by \( \Phi \) into the points \( p'_i \); the value of \( \alpha \) is adjusted so that \( ||\Phi(p_i) - p'_i|| \) is of order \( \eta_i \). Note that this term also expresses a likelihood, where we have now made an explicit assumption about the joint probability density relating the tie points. To match surfaces, the same expression can be used once the corresponding surface points have been identified. The last term represents the internal strain energy of the deformation \( \delta \) [4]. It penalizes deviations from a smooth (globally affine) map \( \Phi \). We assume that the reference anatomy, in addition to being linear elastic, is isotropic and homogeneous; thus, only Lamé's elastic constants \( \lambda \) and \( \mu \) remain to define its material properties. The values of the elastic constants may be related to a prior distribution for \( \delta \); therefore, their values can be fixed for a given domain.

2. METHOD

The expression in Eq. (1) is highly nonlinear, so iterative techniques must be used for its solution. We begin by computing a spline approximation to \( \Phi \) if tie points are available. Otherwise, an initial approximation to \( \Phi \) is inferred by aligning the centroids and principal axes of the brain volumes. The joint histogram between \( I_R(\Phi) \) and \( I_T(\Phi(r)) \) then provides us with an estimate of \( P \), which is smoothed. For simplicity we use a \( C^0 \) finite element approximation to \( \delta \). In the finite element method [5] the domain of the problem is first divided into regions called elements. The elements are connected at discrete nodal points along their periphery. The only unknowns are those at the element nodes. We define \( N \) to be the set of all nodes and \( E \) the set of all elements. The \( j \)th component of the displacement \( \delta \) at any other point within the element is interpolated from its nodal values:

\[ \delta^j = \sum_{n \in N(e)} \delta^j_n N_n, \quad (3) \]

where \( N(e) \) represents the set of nodes in element \( e \) and we write \( \delta^j_n \) for the value of \( \delta^j \) at node \( n \). In our preliminary investigations the interpolating or shape functions \( N_i \) that we have used correspond to those for the two-dimensional (2-D) parabolic isoparametric element [6].

Of the three terms in Eq. (1), the second and third are immediately seen to be quadratic in \( \delta \) (because of the local character of finite element basis functions, the middle term in Eq. (1) involves only a few of the \( \delta^j_n \)'s). If the
first term in Eq. (1) were also quadratic, then the problem of minimizing $\mathbf{II}$ would reduce to solving a large but sparse system of linear equations. Our strategy is to replace the first term by a linearized approximation and find the minimum $\delta^*$ of this quadratic approximation. This is used as the next estimate of $\delta$ and $P$ is recomputed again. We then develop a new linearized approximation valid near $\delta^*$ and repeat the process until a stable solution is reached.

The algorithm is illustrated in Figure 1. It is has an outer loop over the different resolution levels and an inner optimization algorithm. The inner algorithm, given a starting approximation to $\delta$, fixed values for $\alpha$, elastic constants $\lambda$ and $\mu$, and $P(I_T|I_R)$, finds a local optimum for $\mathbf{II}$ given by Eq. (1). The outer loop defines a series of image sets $I^k_R$, $I^k_T$, Lagrange multipliers $\alpha^k$, and elastic constants $\lambda^k$ and $\mu^k$ designed so that if $\delta^k$ is the local minimum for $\mathbf{II}^k(\delta) = \mathbf{II}(\delta; I^k_R, I^k_T, \alpha^k, \lambda^k, \mu^k)$, then:

(i) $\delta^k \rightarrow \delta^*$, the global minimum of $\mathbf{II}(\delta; I_R, I_T, \alpha, \lambda, \mu)$;
(ii) $\alpha^k \rightarrow \alpha^*$, which is such that $\Phi(p_i) \approx p_i'$ to within $\eta$;
(iii) $P^k \rightarrow P^*$, which is an estimate for the joint probability density of $I_R$ and $I_T$;
(iv) $I^k_R \rightarrow I_R$ and $I^k_T \rightarrow I_T$.

$I^k_R$ and $I^k_T$ represent smoothed versions of $I_R$ and $I_T$ that have been reduced in resolution under some fixed filtering schedule. For example, the Gaussian pyramid is a suitable representation for multiscale applications [7, 8] and one that we have successfully applied in our previous work. We wish in general the elastic constants $\lambda^k$ and $\mu^k$ to be decreasing but experimental investigation will be required to determine their schedules. Initially, assuming $\lambda^k \equiv \lambda$, $\mu^k \equiv \mu$ for all $k$ is reasonable. The element and node sets $\mathcal{E}^k$, $\mathcal{N}^k$ are normally different at each iteration $k$, with the size of the elements decreasing as the value of $k$ increases. $\alpha$ is adjusted within the optimization loop so that $\Phi(p_i)$ agrees with $p_i'$ to within approximately $\eta$. Thus, $\alpha$ is decreased if $\sum_{i=1}^N (\Phi(p_i) - p_i')^2/2\eta_i^2$ is too small and increased if it is too large. As with the elastic constants, an initial estimate of $\alpha$ will be obtained from prior experience. In order to calculate $P^{(k,i)}$ and $\partial P^{(k,i)}/\partial I_T$ we first construct the joint empirical probability histogram of $I_R$ and $I_T$ by determining $(\mathbf{I}(x), \mathbf{I}^{(\delta)}(x))$ for a large number of points $x$. This empirical estimate is then smoothed. We can calculate $\partial P(I_T|I_R)/\partial I_T$ either by a filtering method or by the method used to derive the forces in our previous elastic matching algorithm, where $P$ now becomes the similarity function $S$. That is, given a point $x$ in $I_R$ we determine the values of $P(I_T|I_R)$ for a neighborhood centered at $x + \delta$ in $I_T$, where $\delta$ is the current estimate of $\delta^*$. We then model this neighborhood of values with a quadratic and take its derivative to obtain $\partial P(I_T|I_R)/\partial I_T$ at $x$. Following the density estimation step the algorithm attempts to find a new local minimum of $\mathbf{II}(\delta)$ given the current estimate of $\delta$. We proceed by differentiating each of the three terms in Eq. (1) separately:

$$\mathbf{II}^{(i)}: \text{For each } i, p_i \in e_i \in \mathcal{E}, \text{ where } e_i \text{ is some uniquely determined element of } \mathcal{E}. \text{ For } x \in e_i,$$

$$\Phi(x) = R(x + \sum_{n \in \mathcal{N}(e_i)} \delta_n N_n(x)).$$

(4)

It follows that

$$\frac{\partial}{\partial \delta_n} \frac{(\Phi(p_i) - p_i')^2}{2\eta_i^2} = \left\{ \begin{array}{ll} \sum_k \frac{1}{\eta_i} [\Phi(p_i) - p_i']^T R_{jk} N_n(p_i) & \text{if } n \in \mathcal{N}(e_i), \\
0 & \text{otherwise.} \end{array} \right.$$ (5)

In the finite element implementation it is convenient to develop an equivalent matrix form of the above derivation. Let

$$\delta^e = [\delta_1, \delta_2, \ldots, \delta_n]^T,$$

(6)

the vector of nodal displacements of the element $e$ and

$$N = [N_1, N_2, \ldots, N_n],$$

(7)
for \( k = 0; k < N; k++ \) {  
  Loop over each of \( N \) resolution levels.

  if \( k = 0 \) {
    Calculate \( R^{(0,0)} \) global affine map so that \( R^{(0,0)}(p_i) \approx p_i' \) if tie points, \((p_i, p_i')\), are available; otherwise, \( R^{(0,0)} \) is obtained by aligning centroids and principal axes of the brain volumes;
    Calculate \( \Phi^{(0,0)} \) initial spline approximation to \( \Phi \) if tie points, \((p_i, p_i')\), are available;
    Calculate initial displacement field \( \delta^{(0,0)}(x) = (R^{(0,0)})^{-1}\Phi^{(0,0)}(x) - x \) for \( x \in \Omega_R \);
  }

  else {
    \( R^{(k,0)} = R^{(k-1,m_k-1)} \);
    \( \delta^{(k,0)} = \delta^{(k-1,m_k-1)} \);
    \( \alpha^{(k,0)} = \alpha^{(k-1,m_k-1)} \);
    If \( N_k = l \), \( \mathcal{E}_k \) are different from \( N_k, \mathcal{E}_k, \delta^{(k,0)} \) is interpolated from \( \delta^{(k-1,m_k-1)} \);
  }

  for \( l = 0; l < m_k; l++ \) {  
    Optimize \( \Phi \) over \( m_k \) iterations, updating \( \alpha, P \) at intermediate steps.

    if \( l > 0 \) {
      \( R^{(k,l)} = R^{(k,l-1)} \);
      \( \delta^{(k,l)} = \delta^{(k,l-1)} \);
      \( \alpha^{(k,l)} = \alpha^{(k,l-1)} \);
    }

    Calculate \( P^{(k,l)} \), smoothed estimate of \( P \), based on the map \( \Phi^{(k,l)}(x) = R^{(k,l)}(x + \delta^{(k,l)}(x)) \);
    Optimize local deformation \( \delta^{(k,l)}(I^{k,l}, J^{k,l}, \{(p_i, p_i')\}, \alpha^{(k,l)}, P^{(k,l)}, R^{(k,l)}, \lambda_k, \mu_k) \);
    Optimize global affine map \( R^{(k,l)}(I^{k,l}, J^{k,l}, \{(p_i, p_i')\}, \alpha^{(k,l)}, P^{(k,l)}, R^{(k,l)}, \lambda_k, \mu_k) \);  
    Optional.
    Update \( \alpha^{(k,l)} \);
  }

Figure 1: Algorithm to iteratively calculate the mapping \( \Phi \) and the conditional probability density \( P \) by minimizing \( \Pi \) (see Eq. (4)) over multiple resolution levels.
where \( N_i = N_i \mathbf{I} \) and \( \mathbf{I} \) is the \( r \times r \) identity matrix with \( r \) the number of unknowns per node. Then Eq. (3), which defines the displacement at a point within the element, can be written as

\[
\delta = N \delta^e. \tag{8}
\]

For simplicity we will ignore the affine component \( R \) in the matrix derivations. The minimization of \( \Pi^I \) with respect to the nodal displacements \( \delta^e \) for each \( i \) becomes in matrix notation:

\[
\frac{\partial}{\partial \delta^e} \left( \frac{\Phi(p_i) - p_i'}{2\eta^2} \right) = \frac{1}{2\eta^2} \frac{\partial}{\partial \delta^e} \left( (p_i + N \delta^e) - p_i' \right)^2 \tag{9}
\]

\[
= \frac{1}{2\eta^2} \frac{\partial}{\partial \delta^e} \left( (p_i - p_i')^2 + 2[\delta^e]^T[N]^T(p_i - p_i') + [\delta^e]^T[N]^TN\delta^e \right) \tag{10}
\]

\[
= \frac{1}{\eta^2}[N]^T(p_i - p_i') + \frac{1}{\eta^2}[N]^TN\delta^e \tag{11}
\]

\[
= F_i^e + K_i^e \delta^e, \tag{12}
\]

where

\[
K_i^e = \frac{1}{\eta^2}[N]^TN \tag{13}
\]

is the contribution of the point matching term to the element “stiffness” matrix \( K^e \) of the element \( e \) and

\[
F_i^e = \frac{1}{\eta^2}[N]^T(p_i - p_i') \tag{14}
\]

is the contribution to the element “load” vector \( F^e \).

\( \Pi^{III} \): In the finite element method integrals such as the one representing the internal strain energy in \( \Pi^{III} \) are calculated on an element-by-element basis and then summed. Thus, we write

\[
\Pi^{III} = \sum_e \Pi_e^{III} = \sum_e \int_{V_e} [\sigma]^T \varepsilon \, dV, \tag{15}
\]

where it is understood that the summation is over the elements comprising \( \Omega_R \). In most elastic problems the strains at a point within the element can be interpolated from the nodal displacements as follows:

\[
\varepsilon = B \delta^e, \tag{16}
\]

where \( B \) is the strain matrix. For example, in plane problems the strain-displacement relationship can be written as

\[
\varepsilon = \begin{bmatrix} \varepsilon_{x1} \\ \varepsilon_{x2} \\ \gamma_{x1x2} \end{bmatrix} = \begin{bmatrix} \frac{\partial \delta^1}{\partial x_1} \\ \frac{\partial \delta^2}{\partial x_2} \\ \frac{\partial \delta^1}{\partial x_2} + \frac{\partial \delta^2}{\partial x_1} \end{bmatrix} = B \delta^e, \tag{17}
\]

where

\[
B = [B_1, B_2, \ldots, B_n] \tag{18}
\]

and

\[
B_i = \begin{bmatrix} \frac{\partial N_i}{\partial x_1} & 0 \\ 0 & \frac{\partial N_i}{\partial x_2} \end{bmatrix}. \tag{19}
\]

The stresses in turn can be expressed in terms of the strains as

\[
\sigma = D \varepsilon, \tag{20}
\]
where \( \mathbf{D} \) is the elasticity matrix. Again, considering the plane example and in particular isotropic plane stress, the elasticity matrix becomes

\[
\mathbf{D} = \frac{E}{1 - \nu^2} \begin{bmatrix}
1 & \nu & 0 \\
\nu & 1 & 0 \\
0 & 0 & \frac{1 - \nu}{2}
\end{bmatrix},
\]

where \( E \) and \( \nu \) are Young’s modulus and Poisson’s ratio, respectively—these can be directly related to Lame’s elastic constants \( \lambda \) and \( \mu \). Substituting Eqs. (16) and (20) into Eq. (15), we obtain for an arbitrary element \( e \):

\[
\Pi_{e}^{III} = -\frac{1}{2} \int_{V_e} [\delta_e]^T [\mathbf{B}]^T \mathbf{D} \delta_e \, dV.
\]

Minimizing with respect to \( \delta_e \), as was done in \( \Pi^I \), we have:

\[
\frac{\partial \Pi_{e}^{III}}{\partial \delta_e} = \int_{V_e} [\mathbf{B}]^T \mathbf{D} \delta_e \, dV = \mathbf{K}_{e}^{III} \delta_e,
\]

where

\[
\mathbf{K}_{e}^{III} = \int_{V_e} [\mathbf{B}]^T \mathbf{D} \, dV
\]

is the contribution of the deformation term to the element stiffness matrix \( \mathbf{K}^e \).

Alternatively, we may choose to define the deformation energy \( \Pi^{III} \) as follows:

\[
\Pi^{III} = \lambda \int_{V_e} \| \mathbf{\nabla} \delta \|^2 \, dV
\]

where it follows that

\[
\int_{V_e} \lambda \| \mathbf{\nabla} \delta \|^2 \, dV = \sum_j \lambda \int_{V_e} (\sum_{n \in \mathcal{N}(e)} \delta_{ij} \nabla N_n)^2 \, dV
\]

\[
= \sum_{j_{nn'}} \lambda \delta_{ij} \delta_{ij'} \int_{V_e} \nabla N_n \nabla N_{n'} \, dV.
\]

Thus,

\[
\frac{\partial}{\partial \delta_{ij}} \int_{V_e} \lambda \| \mathbf{\nabla} \delta \|^2 \, dV = 2\lambda \sum_{n' \in \mathcal{N}(e)} \delta_{ij} \int_{V_e} \nabla N_n \nabla N_{n'} \, dV \quad \text{for } n \in \mathcal{N}(e)
\]

In matrix form, we write

\[
\int_{V_e} \lambda \| \mathbf{\nabla} \delta \|^2 \, dV = \int_{V_e} \lambda [\delta_e]^T \mathbf{\nabla N} \mathbf{\nabla N}^T \delta_e \, dV,
\]

where \([\mathbf{\nabla N}]_{ij} = \partial N_j / \partial x_i \mathbf{I}\) and \( \mathbf{I} \) is the \( r \times r \) identity matrix with \( r \) the number of coordinate components. The derivative with respect to \( \delta_e \) can then be expressed as

\[
\frac{\partial \Pi_{e}^{III}}{\partial \delta_e} = \int_{V_e} 2\lambda [\mathbf{\nabla N}]^T \mathbf{\nabla N} \delta_e \, dV = \mathbf{K}_{e}^{III} \delta_e,
\]

where

\[
\mathbf{K}_{e}^{III} = \int_{V_e} 2\lambda [\mathbf{\nabla N}]^T \mathbf{\nabla N} \, dV
\]
Calculate $K$, $f_0$; \[ \text{See Eq. (36)} \]

for ($i = 0; i < N; i++$) \{

Calculate $f'_0$ at $\delta$, the current estimate of $\delta^*$;

Solve $K\delta^* = f_0 + f'_0$;

Find $\eta \in [0, 1]$ such that $\Pi(\eta\delta + (1 - \eta)\delta^*)$ is minimal;

Set $\delta = \eta\delta + (1 - \eta)\delta^*$;

If $\delta$ agrees with its previous value to within a predefined tolerance, then exit;

\}

Figure 2: Algorithm to minimize $\Pi$ (see Eq. (4)) with respect to the local deformation.

is the element stiffness contribution of the alternative deformation term.

We can express the sum of $\Pi^{II}$ and $\Pi^{III}$ as follows:

\[ \frac{\partial (\Pi^{II} + \Pi^{III})}{\partial \delta} = K\delta + f_0 \] \[ (36) \]

with $K$ a constant (global stiffness) matrix and $f_0$ a constant (global load) vector.

$\Pi^I$: To differentiate $\Pi^I$ with respect to $\delta$, the integral is again divided into the contributions of individual elements:

\[ \Pi^I = -\sum_e \int_{V_e} \ln P(I_T(R(x + \delta(x))) \mid I_R(x)) \, dV \]

\[ = -\sum_e \int_{V_e} \ln P(I_T(R(x + \sum_{n \in N(e)} \delta_n N_n(x))) \mid I_R(x)) \, dV \] \[ (37) \]

\[ (38) \]

The derivative of $\Pi^I$ with respect to $\delta^I_n$ arises only from the terms in Eq. (38) for which $n \in N(e)$, i.e. for $e \in E(n)$ where $E(n)$ is the set of elements containing $n$. We have, for such an $e$:

\[ \frac{\partial \Pi^I}{\partial \delta^I_n} = \int_{V_e} -\ln P(I_T(R(x + \sum_{n \in N(e)} \delta_n N_n(x))) \mid I_R(x)) \, dV \]

\[ = \int_{V_e} \left[ \frac{-1}{P(I_T(\Phi(x)) \mid I_R(x))} \cdot \frac{\partial P(I_T(\Phi(x)) \mid I_R(x))}{\partial I_T} \right] \cdot \sum_k (\nabla I_T(\Phi(x)) R_{kj} N_k(x)) \, dV, \] \[ (39) \]

\[ (40) \]

where $k$ is the number of coordinate components. The key approximation is to treat this as a constant, say $f'_0$: we then have, with $f = f_0 + f'_0$:

\[ \frac{\partial \Pi}{\partial \delta} \approx K\delta + f; \] \[ (41) \]

a minimum is sought for $\Pi$ by solving Eq. (41). This is a large sparse matrix equation and $K$ is symmetric positive definite for the deformation energies that we have considered. We solve Eq. (41) using the ITPACK iterative linear system solver [9] and call the solution $\delta^*$. We then seek the true minimum for $\Pi$ along the line segment between $\delta$ and $\delta^*$. The algorithm is presented in Figure 2.

The procedure to optimize the global affine map is similar to, but simpler than, the one to optimize $\delta$. First, we can ignore $\Pi^{III}$ since it does not depend on $R$. The gradient of $\Pi^{III}$ with respect to $R$ is easily calculated. The gradient of $\Pi^I$ is computationally intensive, and is similar to the calculations for optimizing $\delta$. However, $R$ can be updated much less frequently than $\delta$ since the initial solution is likely to be quite close to the truth.
3. PRELIMINARY EXPERIMENTS

We are currently investigating the method by evaluating an implementation that works with 2-D images. The objective is to examine every aspect of the algorithm in two dimensions in order to guide the design and implementation of its 3-D version. Figures 3 and 4 illustrate the types of reference images that have been used in our preliminary experiments. In one set of evaluations test images were created by applying a second order polynomial warp to the reference images with small values for the higher order coefficients. The initial displacements are computed by fitting a thin-plate spline to the pairs of tie points we define: 19 for the simulated brain sections. Because the spline thus computed will recover exactly the polynomial warp, a small error is introduced into the location of the tie points so that the calculated displacement field now only approximates the true mapping between the reference and test images. In the simulated brain images, for example, the rms error in the fit of the landmarks is around 0.8 pixels. This value is within the range of distances that each point can be displaced at a given resolution level. Larger displacements are discouraged because of the uncertainty that arises from the computation of the likelihood gradients. In the current implementation \( \frac{\partial P(I_l|I_R)}{\partial I_l} \) is obtained at a point by first approximating the conditional probability with a quadratic function. The approximation is valid only over a small local neighborhood about the point; therefore, we bound the displacements to decrease the likelihood of false matches.

To build the joint empirical probability histogram, the intensity value of the nearest pixel to the mapped point \( \Phi(x) \) in the test image is used. The histogram is then convolved with a Gaussian filter \( (\sigma = 0.9) \) a number of times depending on the current iteration \( k \). In general, the smoothing is reduced with increasing \( k \) but a fixed filtering schedule was empirically determined to be satisfactory for our experiments.

We employ the Gauss-Legendre quadrature method to numerically evaluate the various integrals that make up the element stiffness matrices and load vectors. The same code is used within the line minimization procedure to evaluate Eq. (1). The \( n \)-point Gauss rule integrates exactly all polynomials of degree \( 2^{n-1} \), or less. In spite of its accuracy the method may not be ideal for the current application because within any element only 4 points are sampled to compute the integral—we have used the 2-point rule for square elements that correspond to a 5 \( \times \) 5 region within the image. For the intensity value at each Gauss point, again that of the nearest pixel is used in order to simplify the conditions of these initial experiments.

Figure 3 shows the results for a pair of images which contain only two graylevel values. The original and final misalignment between the images are fairly close: the overlap, defined by the intersection of the gray matter and ventricles in both images divided by the union of the area of these structures, improved from 0.755 to 0.706. Without matching the landmarks explicitly, the rms error of their fit was 0.70 pixels, with the original error equal to 0.85 pixels. This error can be made arbitrarily small by adjusting the width of the uncertainty \( \eta \). Because the term representing each landmark involves only a few of the unknowns its solution does not significantly affect the displacement field elsewhere. However, by heavily weighting the fit of the landmarks the solution can be biased to such an extent that only the landmarks are matched without consideration of the remaining points within the images.

In the preceding example, the structures were colored the same in both the test and reference images. To determine the efficacy of performing the histogram estimation, the same images were recolored as shown in Figure 4. The very different intensity relationship between the images in this case did not affect the outcome of the results although the difference between the final and initial displacement field is negligible.

For the current evaluation the values of the elastic constants \( \mu \) and \( \lambda \) and the Lagrange multiplier \( \alpha \) were adjusted through experimentation. In fact much of the results thus far have been qualitative in nature to simply gain insight into the behavior of the algorithm. It is our objective to carry out a rigorous analysis of the method in the near future and provide along with the algorithm a set of techniques for systematically determining the appropriate values and updating schedules of its various free parameters.

4. DISCUSSION

We have presented a probabilistic formulation of the brain image matching problem and described its finite element implementation. The finite element method is well established as an efficient and accurate technique for
Figure 3: Example of the simulated brain sections (top row) and their misalignment after matching (bottom row). The difference image in the bottom left depicts the remaining misalignment after initial displacements have been applied. The final misalignment is shown in the bottom right.
Figure 4: Example of the simulated brain sections with which the histogram estimation and probabilistic matching is being evaluated. The pair of images shown are identical to the ones in Figure 3 except that these have been recolored to exhibit a very different intensity relationship.

solving systems of partial differential equations subject to appropriate boundary and initial conditions. The method is very general; it facilitates well structured implementations; and the accuracy of its solutions can be controlled by adjusting the grid size or by using higher order elements. In addition, the method is especially suitable for problems with complex geometrical configurations. In addition to exploring these variables of the method, an important component, particularly in our application, that will require further investigation is the numerical algorithm with which the integrals are evaluated. For image matching it will most likely be advantageous to apply a method that requires more sampling points than that is needed by Gaussian quadrature.

We have only begun to examine the method's ability to accommodate arbitrary intensity distortion models but it is evident from our early evaluations that the procedures to construct and smooth the histogram will require careful consideration and additional experimentation. For example, if a nearest neighbor approach is not used to obtain the intensity value at a point not coincident with a pixel then the effect of intensity interpolation on the estimation procedure must be studied. Once the conditional probability is determined its gradient must also be computed. This is where the accuracy of the numerical integration method is especially important. It remains to be seen whether the local quadratic approximation to the conditional probability is sufficiently accurate for the derivative of the probability to be correctly determined.

Finally, we expect the method to be used primarily to match structural images of the human brain. In the case of MRI, precise information about local anatomical landmarks and more global features such as the gray-white interface may be available [10]. These features, which further constrain the mapping, can be introduced into our formulation in a manner similar to how tie points are treated.

5. ACKNOWLEDGMENTS

We thank Dr. Yongmin Kim for generously lending us access to the resources at IPSL. The first author is grateful to Kevin Atteson and Pedro Peralta for many fruitful discussions. This work was supported by the U.S.P.H.S. under Program Project Grant NS-14867-13.

6. REFERENCES


