Medical Image Registration with Partial Data

Senthil Periaswamy [a] Hany Farid [b]

[a] *Siemens Medical Solutions USA, Inc., Malvern, PA 19355, USA*
[b] *Dartmouth College, Hanover, NH 03755, USA*

Abstract

We have developed a general-purpose registration algorithm for medical images and volumes. The transformation between images is modeled as locally affine but globally smooth, and explicitly accounts for local and global variations in image intensities. An explicit model of missing data is also incorporated, allowing us to simultaneously segment and register images with partial or missing data. The algorithm is built upon a differential multiscale framework and incorporates the expectation maximization algorithm. We show that this approach is highly effective in registering a range of synthetic and clinical medical images.

1 Introduction

The goal of image registration is to find a transformation that aligns one image to another. Medical image registration has emerged from this broad area of research as a particularly active field (see [1–6] for general surveys). This activity is due in part to the many clinical applications including diagnosis, longitudinal studies, and surgical planning, and to the need for registration across different imaging modalities (e.g., MRI, CT, PET, X-RAY, etc.). Medical image registration, however, still presents many challenges. Several notable difficulties are 1) the transformation between images can vary widely and be highly non-rigid in nature; 2) images acquired from different modalities may differ significantly in overall appearance and resolution; 3) there may not be a one-to-one correspondence between the images (missing/partial data); and

This work was supported by an Alfred P. Sloan Fellowship, a NSF CAREER Award (IIS-99-83806), and a department NSF infrastructure grant (EIA-98-02068). This work was done while SP was at Dartmouth College. The authors can be reached at senthil.periaswamy@siemens.com and farid@cs.dartmouth.edu.

4) each imaging modality introduces its own unique challenges, making it difficult to develop a single generic registration algorithm.

In estimating the transformation that aligns two images we must choose: 1) to estimate the transformation between a small number of extracted features, or between the complete unprocessed intensity images; 2) a model that describes the geometric transformation; 3) whether to and how to explicitly model intensity changes; 4) an error metric that incorporates the previous three choices; and 5) a minimization technique for minimizing the error metric, yielding the desired transformation.

Feature-based approaches extract a (typically small) number of corresponding landmarks or features between the pair of images to be registered. The overall transformation is estimated from these features. Common features include corresponding points [7–9], edges [10,11], contours [12,13] or surfaces [14–16]. These features may be specified manually or extracted automatically. Fiducial markers may also be used as features; these markers are usually selected to be visible in different modalities. Feature-based approaches have the advantage of greatly reducing computational complexity. Depending on the feature extraction process, these approaches may also be more robust to intensity variations that arise during, for example, cross modality registration. Also, features may be chosen to help reduce sensor noise. These approaches can be, however, highly sensitive to the accuracy of the feature extraction. Intensity-based approaches, on the other hand, estimate the transformation between the entire intensity images. Such an approach is typically more computationally demanding, but avoids the difficulties of a feature extraction stage.

Independent of the choice of a feature- or intensity-based technique, a model describing the geometric transform is required. A common and straightforward choice is a model that embodies a single global transformation. The problem of estimating a global translation and rotation parameter has been studied in detail, and a closed form solution was proposed by Schönemann [7] in 1966. Other closed-form solutions include methods based on singular value decomposition (SVD) [17], eigenvalue-eigenvector decomposition [18] and unit quaternions [19]. One idea for a global transformation model is to use polynomials. For example, a zeroth-order polynomial limits the transformation to simple translations, a first-order polynomial allows for an affine transformation, and, of course, higher-order polynomials can be employed yielding progressively more flexible transformations. For example, the registration package Automated Image Registration (AIR) can employ (as an option) a fifth-order polynomial consisting of 168 parameters (for 3-D registration) [20,21]. The global approach has the advantage that the model consists of a relatively small number of parameters to be estimated, and the global nature of the model ensures a consistent transformation across the entire image. The disadvantage of this approach is that estimation of higher-order polynomials can lead to an

unstable transformation, especially near the image boundaries. In addition, a relatively small and local perturbation can cause disproportionate and unpredictable changes in the overall transformation. An alternative to these global approaches are techniques that model the global transformation as a piecewise collection of local transformations. For example, the transformation between each local region may be modeled with a low-order polynomial, and global consistency is enforced via some form of a smoothness constraint. The advantage of such an approach is that it is capable of modeling highly nonlinear transformations without the numerical instability of high-order global models. The disadvantage is one of computational inefficiency due to the significantly larger number of model parameters that need to be estimated, and the need to guarantee global consistency. Low-order polynomials are, of course, only one of many possible local models that may be employed. Other local models include B-splines [22–25], thin-plate splines [26,9], and a multitude of related techniques. The package Statistical Parametric Mapping (SPM) uses the low-frequency discrete cosine basis functions [27,28], where a bending-energy function is used to ensure global consistency. Physics-based techniques that compute a local geometric transform include those based on the Navier-Stokes equilibrium equations for linear elasticity [29–31] and those based on viscous fluid approaches [31–33].

Under certain conditions a purely geometric transformation is sufficient to model the transformation between a pair of images. Under many real-world conditions, however, the images undergo changes in both geometry and intensity (e.g., brightness and contrast). Many registration techniques attempt to remove these intensity differences with a pre-processing stage, such as histogram matching [34,35] or homomorphic filtering [36]. The issues involved with modeling intensity differences are similar to those involved in choosing a geometric model. Because the simultaneous estimation of geometric and intensity changes can be difficult, few techniques build explicit models of intensity differences. A few notable exceptions include AIR [20,21], in which global intensity differences are modeled with a single multiplicative contrast term, and SPM [27,28] in which local intensity differences are modeled with a basis function approach.

Having decided upon a transformation model, the task of estimating the model parameters begins. As a first step, an error function in the model parameters must be chosen. This error function should embody some notion of what is meant for a pair of images to be registered. Perhaps the most common choice is a mean square error (MSE), defined as the mean of the square of the differences (in either feature distance or intensity) between the pair of images. This metric is easy to compute and often affords simple minimization techniques. A variation of this metric is the unnormalized correlation coefficient applicable to intensity-based techniques. This error metric is defined as the sum of the point-wise products of the image intensities, and can be efficiently com-

puted using Fourier techniques [37,38]. A disadvantage of these error metrics is that images that would qualitatively be considered to be in good registration may still have large errors due to, for example, intensity variations, or slight misalignments. Another error metric (included in AIR) is the ratio of image uniformity (RIU) defined as the normalized standard deviation of the ratio of image intensities. Such a metric is invariant to overall intensity scale differences, but typically leads to nonlinear minimization schemes. Mutual information [39–42], entropy [43,44] and the Pearson product moment cross correlation [45] are just a few examples of other possible error functions. Such error metrics are often adopted to deal with the lack of an explicit model of intensity transformations [46].

In the final step of registration, the chosen error function is minimized yielding the desired model parameters. In the most straightforward case, least-squares estimation is used when the error function is linear in the unknown model parameters. This closed-form solution is attractive as it avoids the pitfalls of iterative minimization schemes such as gradient-descent or simulated annealing. Such nonlinear minimization schemes are, however, necessary due to an often nonlinear error function. A reasonable compromise between these approaches is to begin with a linear error function, solve using least-squares, and use this solution as a starting point for a nonlinear minimization.

In developing our general-purpose registration algorithm, we have tried to find a compromise between a flexible and robust technique and computational efficiency. To begin, we have chosen an intensity-based approach so as to avoid the various pitfalls involved in feature selection. Geometrically, we model the transformation with a local affine model and a global smoothness constraint. Intensity variations are explicitly modeled with local changes in brightness and contrast and a global smoothness constraint. These model parameters are simultaneously estimated at each pixel in the image, allowing us to capture nonlinear distortions (in both geometry and intensity). We employ a standard MSE error metric on the intensity values. The minimization involves two steps. First an error function that is linear in the model parameters is minimized using least-squares. This error function is then augmented with a nonlinear smoothness constraint, and the least-squares solution is used to bootstrap an iterative nonlinear minimization. This entire procedure is built upon a differential multiscale framework, allowing us to capture both large- and small-scale transformations, see also [46–48] for related techniques.

In addition to this framework we describe an extension that allows us to explicitly contend with missing or partial data. Shown in Fig. 4 are examples of the challenges posed by missing data. In these examples there are large portions of the source image that have no corresponding match in the target image. Without an explicit segmentation or localization of these missing regions, most registration algorithms are unlikely to correctly register these

images. Of course, if the registration between these images were known, then it would be straight-forward to perform the segmentation. Similarly, if the segmentation were known, the registration could proceed. Without a known segmentation or registration, however, we are faced with a bit of a chicken and egg problem - which step should be performed first? In order to contend with this problem we have employed the expectation maximization algorithm that simultaneously segments and registers a pair of images or volumes (see also [49]).

For purposes of completeness we will briefly review our previous registration algorithm [50,51], and then describe the extension that allows us to contend with missing data. We then show the efficacy of this approach on several synthetic and clinical cases.

2 Registration

We formulate the problem of image registration within a differential (non feature-based) framework. This formulation borrows from various areas of motion estimation (e.g., [52,53]). In order to contend with partial or missing data, the expectation maximization algorithm [54] is incorporated into this framework, allowing for simultaneous segmentation and registration. We first outline the basic computational framework, and then discuss several implementation details that are critical for a successful implementation.

2.1 Local affine

Denote $f(x, y, t)$ and $f(\hat{x}, \hat{y}, t-1)$ as the source and target images, respectively. [1] We begin by assuming that the image intensities between images are conserved (this assumption will be relaxed later), and that the geometric transformation between images can be modeled locally by an affine transform:

$$f(x, y, t) = f(m_1 x + m_2 y + m_5, m_3 x + m_4 y + m_6, t-1), \tag{1}$$

where m_1, m_2, m_3, m_4 are the linear affine parameters, and m_5, m_6 are the translation parameters. These parameters are estimated locally for each small spatial neighborhood, but for notational convenience their spatial parameters

[1] We adopt the slightly unconventional notation of denoting the source and target image with a temporal parameter t. This is done for consistency within our differential formulation.

are dropped. In order to estimate these parameters, we define the following quadratic error function to be minimized:

$$E(\vec{m}) = \sum_{x,y \in \Omega} \left[f(x,y,t) - f(m_1 x + m_2 y + m_5, m_3 x + m_4 y + m_6, t-1) \right]^2, \quad (2)$$

where $\vec{m} = \begin{pmatrix} m_1 & \dots & m_6 \end{pmatrix}^T$, and Ω denotes a small spatial neighborhood. Since this error function is nonlinear in its unknowns, it cannot be minimized analytically. To simplify the minimization, we approximate this error function using a first-order truncated Taylor series expansion:

$$\begin{aligned} E(\vec{m}) \approx \sum_{x,y \in \Omega} \left[f_t(x,y,t) - (m_1 x + m_2 y + m_5 - x) f_x(x,y,t) \right. \\ \left. -(m_3 x + m_4 y + m_6 - y) f_y(x,y,t) \right]^2, \end{aligned} \quad (3)$$

where $f_x(\cdot)$, $f_y(\cdot)$, $f_t(\cdot)$ are the spatial/temporal derivatives of $f(\cdot)$. Note that this quadratic error function is now linear in its unknowns, $\vec{m}$. This error function may be expressed more compactly in vector form as:

$$E(\vec{m}) = \sum_{x,y \in \Omega} \left[k - \vec{c}^T \vec{m} \right]^2, \quad (4)$$

where the scalar k and vector $\vec{c}$ are given as: $k = f_t + x f_x + y f_y$ and $\vec{c} = \begin{pmatrix} x f_x & y f_x & x f_y & y f_y & f_x & f_y \end{pmatrix}^T$. This error function can now be minimized analytically by differentiating with respect to the unknowns:

$$\frac{dE(\vec{m})}{d\vec{m}} = \sum_{x,y \in \Omega} -2\vec{c} \left[k - \vec{c}^T \vec{m} \right], \quad (5)$$

setting the result equal to zero, and solving for $\vec{m}$, yielding:

$$\vec{m} = \left[\sum_{x,y \in \Omega} \vec{c}\, \vec{c}^T \right]^{-1} \left[\sum_{x,y \in \Omega} \vec{c}\, k \right]. \quad (6)$$

This solution assumes that the first term, a 6×6 matrix, is invertible. This can usually be guaranteed by integrating over a large enough spatial neighborhood Ω with sufficient image content. When an estimate cannot be made, the local parameters are interpolated from nearby regions. With this approach a dense locally affine mapping can be found between a source and target image.

2.2 Intensity variations

Inherent to the model outlined in the previous section is the assumption that the image intensities between the source and target are unchanged (brightness constancy). This assumption is likely to fail under a number of circumstances. To account for intensity variations, we incorporate into our model an explicit change of local contrast and brightness [55]. Specifically, our initial model, Equation (1), now takes the form:

$$m_7 f(x, y, t) + m_8 = f(m_1 x + m_2 y + m_5, m_3 x + m_4 y + m_6, t - 1), \tag{7}$$

where m_7 and m_8 are two new (also spatially varying) parameters that embody a change in contrast and brightness, respectively. Note that these parameters have been introduced in a linear fashion. As before, the error function is approximated with a first-order truncated Taylor series expansion to yield:

$$E(\vec{m}) = \sum_{x,y \in \Omega} \left[k - \vec{c}^T \vec{m} \right]^2, \tag{8}$$

where the scalar k and vector $\vec{c}$ are now given as:

$$k = f_t - f + x f_x + y f_y \tag{9}$$

$$\vec{c} = \begin{pmatrix} x f_x & y f_x & x f_y & y f_y & f_x & f_y & -f & -1 \end{pmatrix}^T. \tag{10}$$

Minimizing this error function is accomplished as before by differentiating $E(\vec{m})$, setting the result equal to zero and solving for $\vec{m}$. The solution takes the same form as in Equation (6), with k and $\vec{c}$ as defined in Equations (9) and (10).

Intensity variations are typically a significant source of error in differential motion estimation. The addition of the contrast and brightness terms allows us to accurately register images in the presence of these variations.

2.3 Smoothness

Until now, we have assumed that the local affine and contrast/brightness parameters are constant within a small spatial neighborhood, Equation (8). There is a natural trade-off in choosing the size of this neighborhood. A larger area makes it more likely that the matrix in Equation (6) will be invertible. A smaller area, however, makes it more likely that the assumption of constant motion will hold. We can avoid balancing these two issues by replacing

the assumption of constancy with a smoothness assumption [52]. That is, it is assumed that the model parameters $\vec{m}$ vary smoothly across space. A smoothness constraint on the contrast/brightness parameters has the added benefit of avoiding a degenerate solution where a pure intensity-based modulation is used to describe the mapping between images.

We begin with an error function:

$$E(\vec{m}) = E_s(\vec{m}) + E_b(\vec{m}), \tag{11}$$

that combines a smoothness constraint, $E_s(\vec{m})$, with the previous geometric and intensity transformation constraint, $E_b(\vec{m})$. The term $E_b(\vec{m})$ is defined as in Equation (8) without the summation:

$$E_b(\vec{m}) = \left[k - \vec{c}^T \vec{m}\right]^2, \tag{12}$$

with k and $\vec{c}$ given by Equations (9) and (10). The new term $E_s(\vec{m})$ embodies a smoothness constraint:

$$E_s(\vec{m}) = \sum_{i=1}^{8} \lambda_i \left[\left(\frac{\partial m_i}{\partial x} \right)^2 + \left(\frac{\partial m_i}{\partial y} \right)^2 \right], \tag{13}$$

where λ_i is a positive constant that controls the relative weight given to the smoothness constraint on parameter m_i. This error term penalizes solutions proportional to the local change in each parameter across a small spatial neighborhood. In so doing, we allow for a locally smooth, but globally non-rigid transformation. The full error function $E(\vec{m})$ is minimized, as before, by differentiating, setting the result equal to zero and solving for $\vec{m}$. The derivative of $E_b(\vec{m})$ is:

$$dE_b\left(\vec{m}\right)/d\vec{m} = -2\vec{c}\left[k - \vec{c}^T \vec{m}\right]. \tag{14}$$

The derivative of $E_s(\vec{m})$ is computed by first expressing the partials, $\partial m_i/\partial x$ and $\partial m_i/\partial y$ with discrete approximations [52], and then differentiating, to yield:

$$dE_s\left(\vec{m}\right)/d\vec{m} = 2L(\overline{\vec{m}} - \vec{m}), \tag{15}$$

where $\overline{\vec{m}}$ is the component-wise average of $\vec{m}$ over a small spatial neighborhood, and L is an 8×8 diagonal matrix with diagonal elements λ_i, and zero off the diagonal. Setting $dE_b\left(\vec{m}\right)/d\vec{m} + dE_s\left(\vec{m}\right)/d\vec{m} = 0$, and solving for $\vec{m}$

at each pixel location yields an enormous linear system which is intractable to solve. Instead $\vec{m}$ is estimated in the following iterative manner [52]:

$$\vec{m}^{(j+1)} = \left(\vec{c}\,\vec{c}^{T} + L\right)^{-1}\left(\vec{c}\,k + L\overline{\vec{m}}^{(j)}\right). \tag{16}$$

The initial estimate $\vec{m}^{(0)}$ is determined from the closed-form solution of Section 2.2. On the $j+1^{st}$ iteration $\overline{\vec{m}}^{(j)}$ is estimated from the previous estimate, $\vec{m}^{(j)}$.

The use of a smoothness constraint has the benefit that it yields a dense locally affine and smooth transformation. The drawback is that the minimization is no longer analytic. We have found, nevertheless, that the iterative minimization is quite stable and converges relatively quickly (see Section 2.5).

Shown in Fig. 1 are a source and target with both geometric and intensity differences. Also shown in this figure is the registered source and the estimated contrast, brightness, and geometric maps. Shown in Fig. 2 are several more registration examples.

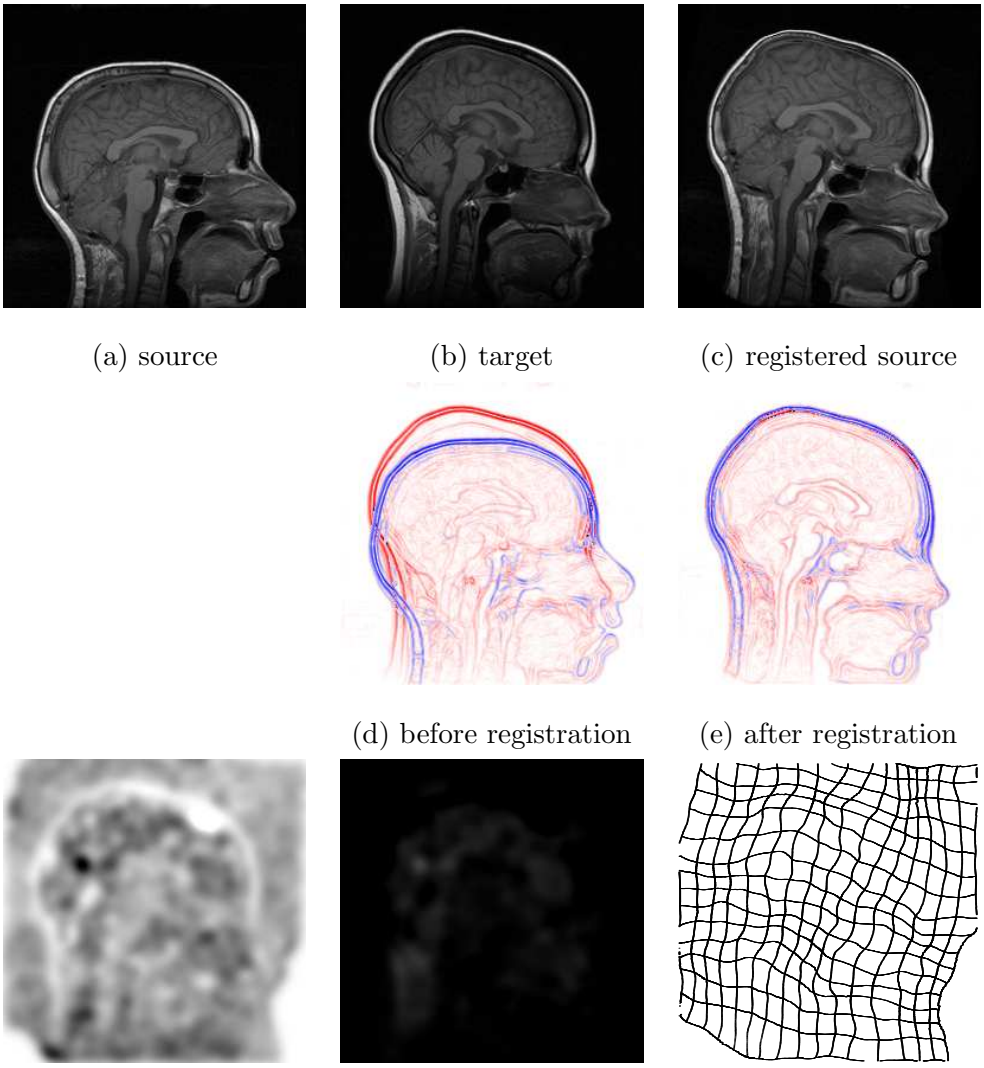
2.4 Partial Data

Inherent to the registration algorithm described above is the assumption that each region in the source image has a corresponding match in the target image. As illustrated in Fig. 4, this need not always be the case. Under such situations, our registration algorithm typically fails. One way to contend with partial or missing data is to employ a pre-processing segmentation step. We propose, however, a more unified approach in which the registration and segmentation are performed simultaneously.

We begin by assuming that each pixel in the source and target are either related through the intensity and geometric model of Equation (7), denoted as model M_1, or cannot be explained by this transformation and therefore belongs to an "outlier" model M_2. Pixels belonging to the outlier model are those that do not have a corresponding match between the source and target images. Assuming that the pixels are spatially independent and identically distributed (iid), the likelihood of observing a pair of images is given by:

$$L(\vec{m}) = \prod_{x,y\in\Omega} P(\vec{q}(x,y)), \tag{17}$$

where, $\vec{q}(x,y)$ denotes the tuple of source, $m_7 f(x,y,t) + m_8$, and target, $f(m_1x + m_2y + m_5, m_3x + m_4y + m_6, t-1)$, image intensities, Equation (7).



(a) source (b) target (c) registered source

(d) before registration (e) after registration

(f) estimated contrast (g) estimated brightness (h) estimated map

Fig. 1. Complete results for a pair of sagittal images from two different patients. Shown along the top are the source, target and registered source. Shown in panels (d) and (e) are the overlaid edges before and after registration. Shown in panels (f), (g) and (h) are the estimated contrast, brightness and geometric maps.

To simplify the optimization of the likelihood function, we consider the log-likelihood function:

$$\begin{aligned}\log[L(\vec{m})] &= \log\left[\prod_{x,y\in\Omega} P(\vec{q}(x,y))\right] \\ &= \sum_{x,y\in\Omega} \log\left[P(\vec{q}(x,y)|M_1)P(M_1) + P(\vec{q}(x,y)|M_2)P(M_2)\right]. \quad (18)\end{aligned}$$

Assuming that the priors on the models, $P(M_1)$ and $P(M_2)$, are equal, the log-likelihood function simplifies to:

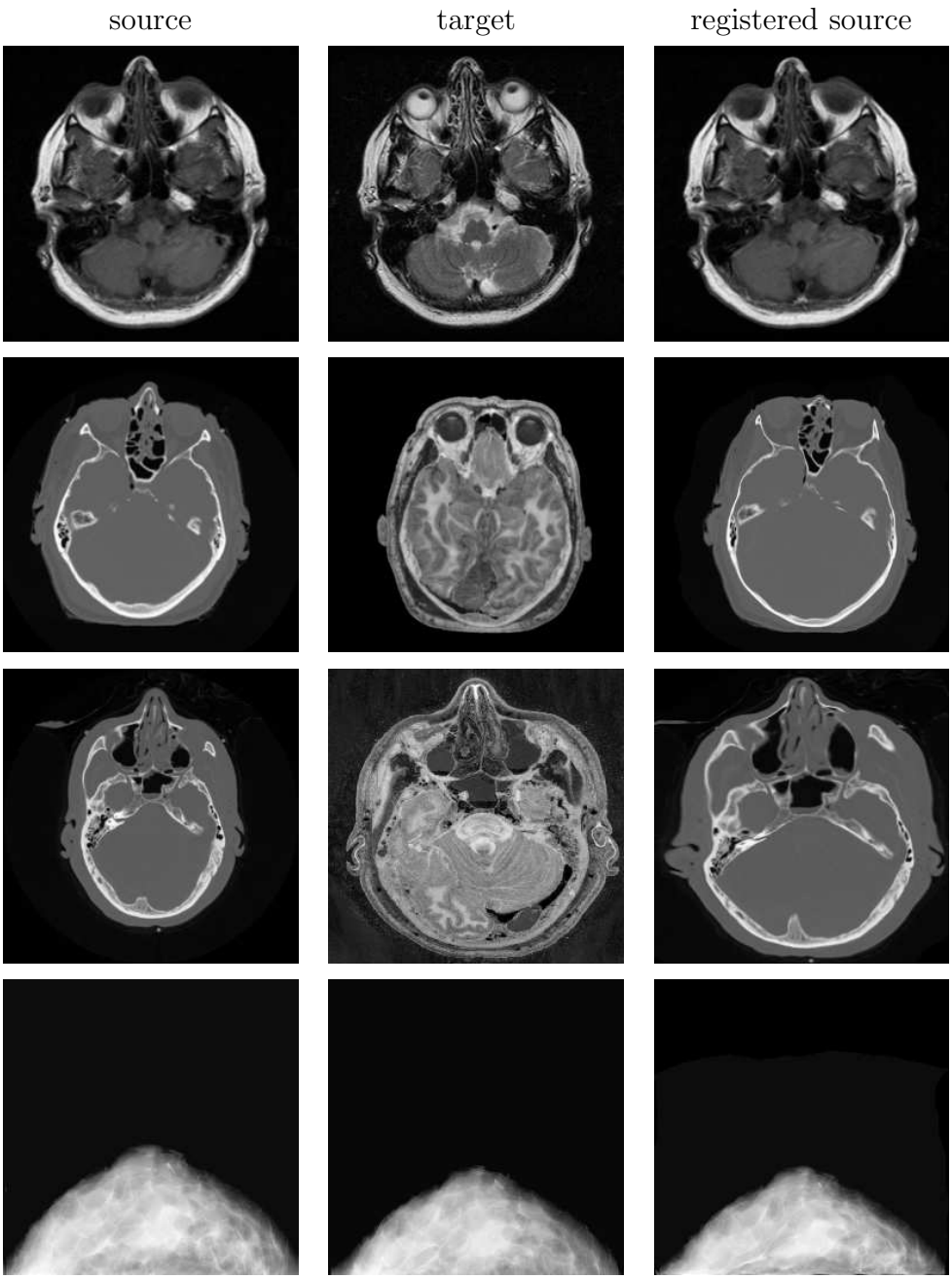


Fig. 2. Shown are examples of registration in the presence of significant geometric and intensity variations. Registration is also successful in the absence of high spatial frequency features, as shown in the mammogram example (bottom row).

$$\log[L(\vec{m})] = \sum_{x,y\in\Omega} \log\left[P(\vec{q}(x,y)|M_1) + P(\vec{q}(x,y)|M_2)\right], \tag{19}$$

where the factored additive constant is ignored for purposes of maximization.

We assume next that the conditional probabilities take the following form:

$$\log[L(\vec{m})] = \sum_{x,y\in\Omega} \log\left[e^{-r^2(x,y)/\sigma^2} + e^{-c^2}\right]. \tag{20}$$

For model M_1 we assume a Gaussian distribution (with variance σ^2), where $r(x,y)$ is the residual error between the source and target defined as:

$$r(x,y) = [m_7 f(x,y,t) + m_8 - f(m_1 x + m_2 y + m_5, m_3 x + m_4 y + m_6, t-1)]. \tag{21}$$

For model M_2 we assume a uniform distribution (i.e., c is a constant). The log-likelihood function is maximized by differentiating, setting the result equal to zero and solving for $\vec{m}$:

$$\frac{d\log[L(\vec{m})]}{d\vec{m}} = \sum_{x,y\in\Omega} \frac{\frac{dr^2(x,y)}{d\vec{m}} e^{-r^2(x,y)/\sigma^2}}{e^{-r^2(x,y)/\sigma^2} + e^{-c^2}} = \sum_{x,y\in\Omega} \frac{dr^2(x,y)}{d\vec{m}} w(x,y) = 0, \tag{22}$$

where $w(\cdot)$ is defined to be the ratio of the exponential distributions. As in the previous sections, the residual $r(\cdot)$ is linearized with respect to the model parameters $\vec{m}$. The derivative of the residual, $dr^2(x,y)/d\vec{m}$, is then substituted into the above to yield:

$$\sum_{x,y\in\Omega} -2\vec{c}[k - \vec{c}^T\vec{m}]w = 0, \tag{23}$$

with $\vec{c}$ and k given by Equations (9) and (10), and, as before, all spatial parameters are dropped for notational convenience. Solving for the model parameters then yields the maximum likelihood estimator:

$$\vec{m} = \left[\sum_{x,y\in\Omega} (\vec{c}\vec{c}^T)w\right]^{-1} \left[\sum_{x,y\in\Omega} (\vec{c}k)w\right]. \tag{24}$$

Note that this solution is a weighted version of the earlier least-squares solution, Equation (6), where the weighting, w, is proportional to the likelihood that each pixel belongs to model M_1. As before, a smoothness constraint can be imposed to yield the following iterative estimator:

$$\vec{m}^{(j+1)} = \left((\vec{c}\,\vec{c}^T)w + L\right)^{-1} \left((\vec{c}\,k)w + L\overline{\vec{m}}^{(j)}\right). \tag{25}$$

This estimator for $\vec{m}$, however, requires an estimate of the weight w which itself requires an estimate of $\vec{m}$. The expectation/maximization algorithm (EM) [54] is used to resolve this circular estimator, and proceeds as follows:

(1) E-step: compute the weights w (with an initial estimate of $\vec{m}$ from the solution of Section 2.3).
(2) M-step: estimate the model parameters $\vec{m}$, Equation (25).
(3) Repeat steps 1 and 2 until the difference between successive estimates of $\vec{m}$ is below a specified threshold.

The E-step is the segmentation stage, where pixels that do not have a corresponding match between source and target images have a close to zero weight w. These pixels are therefore given less consideration in the M-step which estimates the registration parameters $\vec{m}$. The EM algorithm allows for simultaneous segmentation and registration, and hence allows us to contend with missing data.

2.5 Implementation details

While the formulation given in the previous sections is relatively straightforward there are a number of implementation details that are critical for a successful implementation. First, in order to simplify the minimization, the error function of Equation (8) was derived through a Taylor-series expansion. A more accurate estimate of the actual error function can be determined using a Newton-Raphson style iterative scheme [56]. In particular, on each iteration, the estimated geometric transformation is applied to the source image, and a new transformation is estimated between the newly warped source and target image. As few as five iterations greatly improves the final estimate. Second, calculation of the spatial/temporal derivatives in Equations (9) and (10) is a crucial step. These derivatives are often computed using finite differences which typically yield poor approximations. We employ a set of derivative filters, specifically designed for multi-dimensional differentiation [57], that significantly improve the registration results. And third, a coarse-to-fine scheme is adopted in order to contend with larger motions [58,59]. A Gaussian pyramid is first built for both source and target images, and the full registration is estimated at the coarsest level. This estimate is used to warp the source image in the next level of the pyramid. A new estimate is computed at this level, and the process repeated throughout each level of the pyramid. The transformations at each level of the pyramid are accumulated yielding a single final transformation.

The generalization of the algorithm from 2-D images to 3-D volumes is relatively straight-forward. Briefly, to accommodate a 3-D affine transformation, an additional six affine parameters are added to the geometric and intensity transformation model of Equation (7). Linearization and minimization of this constraint proceeds as in the 2-D case. The smoothness constraint of Equation (13) takes on an additional $(\partial m_i/\partial z)^2$ term, and the iterative estimator

of Equation (16) is of the same form, with k and $\vec{c}$ accommodating a different set of, now 3-D, spatial/temporal derivatives. The solution of Section (2.4) proceeds in a similar manner, with the initial constraint of Equation (17) updated to accommodate the 3-D geometric and intensity transformation model. See [51] for more details.

In the current MatLab implementation, running on a 2.8 GHz Linux machine, a pair of 256 $\times$ 256 images requires 4 minutes to register. A pair of 64 $\times$ 64 $\times$ 64 volumes requires 30 minutes.

3 Results

We have tested the efficacy of our registration technique on both synthetic and clinical data in both 2-D and 3-D. Shown in each row of Fig. 3 are source images where a region was removed or replaced with noise. Shown in Fig. 4 are three clinical examples of registration in the presence of varying amounts of missing data. And, shown in Fig. 5 is a synthetic 3-D example. In all cases, the registration is successful even with significant amounts of missing data. This registration would have failed without an explicit model of missing data incorporated directly into the registration algorithm. In all of these results, all system parameters were held fixed. These results show qualitatively the effectiveness of our technique. We next describe a series of results that quantify the robustness and sensitivity of our registration algorithm.

3.1 Analysis

In this section we quantify the accuracy of our registration algorithm with respect to the various design assumptions. Synthetic images and registration maps were used for all the simulations in this analysis. While these images/maps capture only certain aspects of medical images, they allow for a large testbed, and avoid biasing the results towards specific images/maps. We begin with a description of these synthetic images and registration maps. In all of the subsequent simulations, we report the registration accuracy over 100 independently generated images and registration maps. The errors are reported as the root mean square (RMS) error in intensity between the registered source and the target image, and as the RMS error (in pixels) between the estimated and actual registration map.

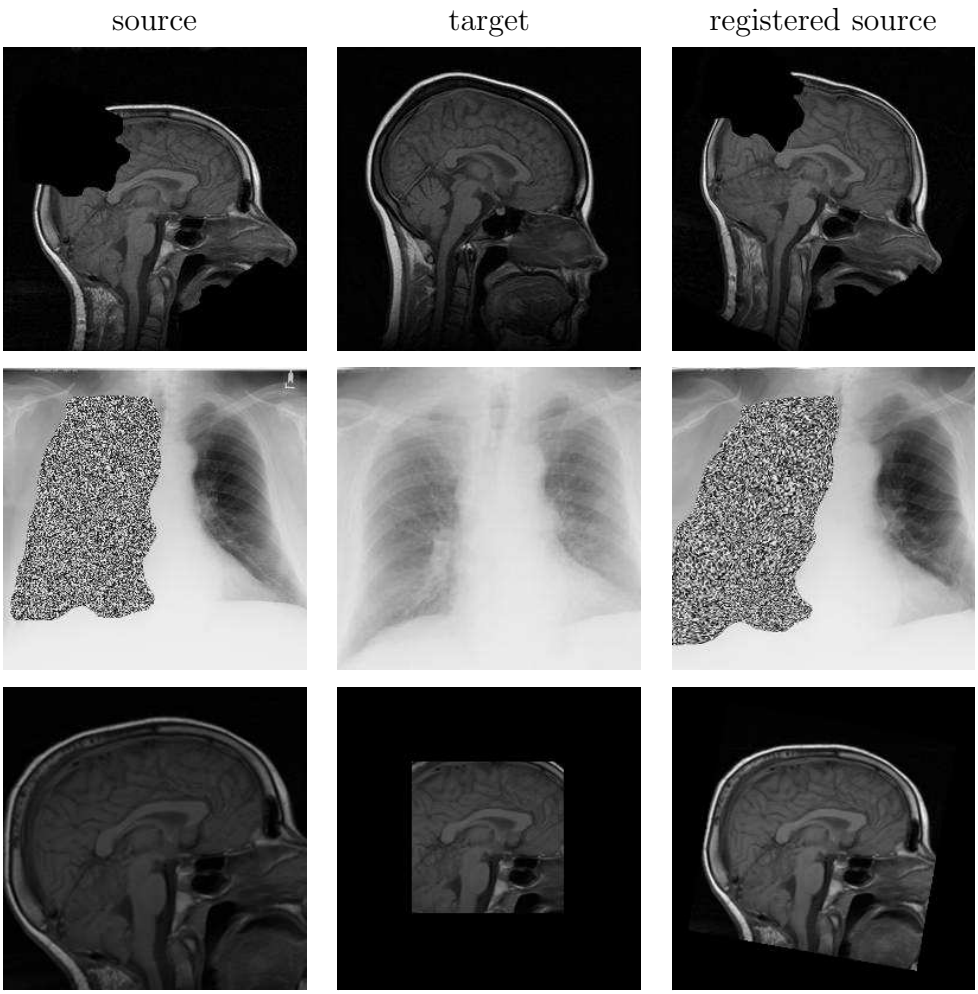


Fig. 3. Shown are synthetic examples of registration with missing data.

3.1.1 Synthetic Image

Consistent with a simple model for natural images [60], the images synthesized for the simulations are fractal in nature and modeled with a $(1/\omega)^{\alpha}$ power spectrum and random phase. The value of α is determined as follows. For a set of sample images, the 2-D power spectrum is first computed. For each discrete frequency ω in the 2-D power spectrum, the median value is computed, thus reducing the 2-D spectrum to 1-D. This 1-D spectrum is then modeled as $(1/\omega)^{\alpha}$, and α is estimated using a simple least-squares estimation. Note that in this process, directional information is lost (since we use the median value at each frequency), along with the phase correlations (since the model uses only the power spectrum). The estimated value of α is 1.4, averaged over a set of ten clinical images (MRI, chest X-rays, mammograms, and CT).

Given the value of α, each synthetic image $s(x, y)$ is generated in the Fourier domain as follows. Let $\Im\{\cdot\}$ denote the Fourier operator, and $\Im^{-1}\{\cdot\}$ the

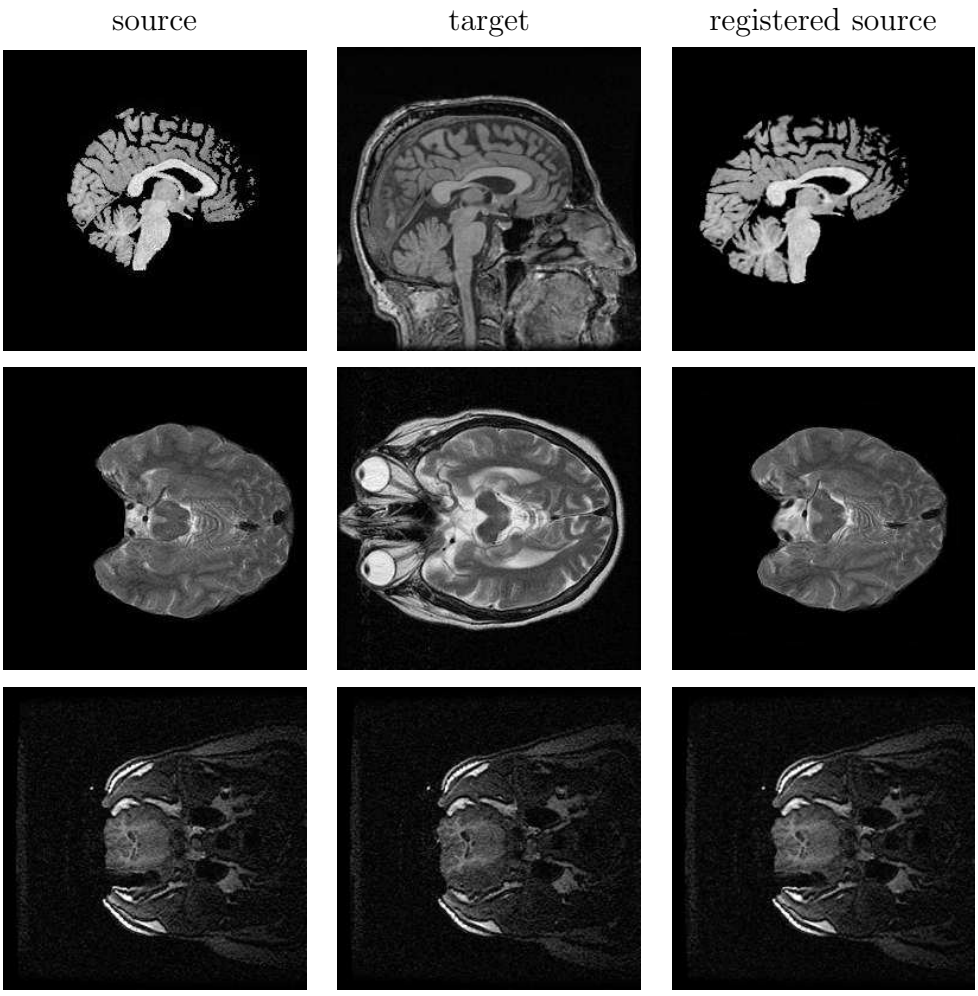


Fig. 4. Shown are clinical examples of registration with missing data.

inverse Fourier operator. Then:

$$s\left(x,y\right)=\Im^{-1}\left\{\Im\left\{r\left(x,y\right)\right\}\cdot H\left(\omega_x,\omega_y\right)\right\},\tag{26}$$

where $r(x,y)$ is a random image of size 128×128, with pixel values chosen from a normal distribution with zero mean and unit variance. The response $H\left(\omega_x,\omega_y\right)$ corresponds to the model of the desired power spectrum in 2-D, given by:

$$H\left(\omega_x,\omega_y\right)=\left(\frac{1}{\sqrt{\omega_x^2+\omega_y^2}}\right)^{\alpha}.\tag{27}$$

The intensity values of the image $s\left(x,y\right)$ are scaled into the range $[0,1]$. A sample synthetic image is shown in Figure 6(a) – the fractal image is padded with a black border to avoid edge artifacts.

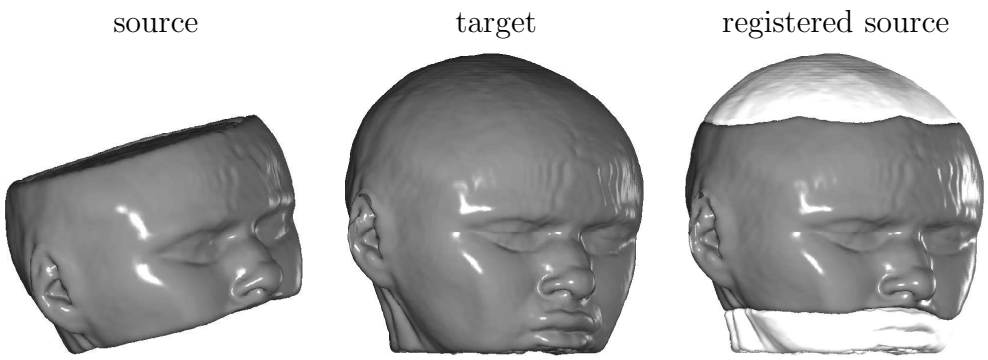


Fig. 5. Shown is an example of 3-D registration with partial data. The brighter regions shown with the registered source are the portions of the target that are missing in the source - these regions are superimposed to show the accuracy of the registration.

3.1.2 Synthetic Registration Map

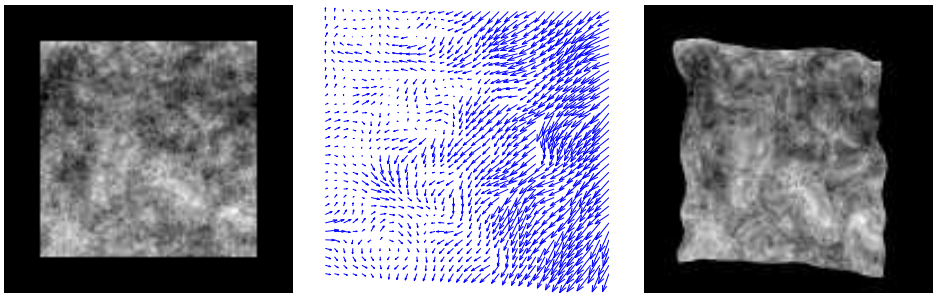
A registration map consists of a displacement vector $\vec{v}(x,y) = (v_x(x,y), v_y(x,y))$ at each pixel location (x,y). The displacement fields $v_x(\cdot)$ and $v_y(\cdot)$ are modeled independently as two Gaussian distributions, and are synthesized independently as follows. All parameter values used were estimated from known registration maps. First, two random images of size 32×32 are generated, with pixel values drawn from a normal distribution with zero mean and a standard deviation of 5.0 pixels. Each image is then up-sampled to a size of 128×128, after which a 3×3 low-pass filter with coefficients $[1\,2\,1; 2\,4\,2; 1\,2\,1]/16$ is repeatedly applied, until a smoothness value of approximately 0.3 is obtained. Smoothness is computed as follows:

$$\left(\frac{1}{N^2}\right)\left(\sqrt{\sum_{x,y\in\Omega}\left(\left(\frac{\partial v_x}{\partial x}\right)^2+\left(\frac{\partial v_x}{\partial y}\right)^2\right)}+\sqrt{\sum_{x,y\in\Omega}\left(\left(\frac{\partial v_y}{\partial x}\right)^2+\left(\frac{\partial v_y}{\partial y}\right)^2\right)}\right), \tag{28}$$

where Ω is the entire image, and N is the total number of pixels. Finally, a global affine $[m_1\ m_2;\ m_3\ m_4]$ and translation $[m_5;\ m_6]$ map is applied to each of the images. The affine parameters vary with a uniform distribution; parameters m_1 and m_4 (the scale parameters) vary from 0.9 to 1.1, parameters m_2 and m_3 (the shear parameters) vary from 0 to 0.2, and parameters m_5 and m_6 (the translation parameters) vary from -10 to $+10$. A sample synthetic registration map is shown in Figure 6(b).

3.1.3 Geometric Distortions

The ability of differential techniques to handle large distortions is generally a concern. Therefore, the extent to which global translations, rotations and scale changes can be recovered was analyzed. In these simulations, the target



(a) (b) (c)

Fig. 6. (a) Sample random synthetic image; (b) sample random synthetic registration map ; (c) result of applying the synthetic registration map to the synthetic image.

image consisted of a random synthetic image, while the source image consisted of the target image with a random global distortion applied.

In the first simulation, a global distortion consisting of only translations, ranging from 0 to 30 pixels in steps of 1 pixel was applied. The algorithm was successful up to a translation of 24 pixels, with an average intensity RMS error of 0.002, and an average registration RMS error of 0.18 pixels. In the second set of simulations, the global distortions consisted of only rotations, applied from 0 to 90 degrees in steps of 1 degree. The algorithm was successful up to a rotation of 45 degrees, with an average intensity RMS error of 0.02, and an average registration RMS error of 0.2 pixels. And, in the third set of simulations, the global distortions consisted of only scale changes, varying from a factor of 1.0 to 2.0 in steps of 0.1. The algorithm was successful up to a scale factor of 1.6, with an average intensity RMS error of 0.04, and an average registration RMS error of 0.3 pixels.

3.1.4 Intensity Distortions

In this simulation, the sensitivity of the registration algorithm to brightness variations was analyzed. The source image consisted of a synthetic fractal image (with intensities in the range $[0, 1]$), with a synthetic registration map applied. The target image consisted of the same synthetic image used in the source, with a random fractal brightness map added to it, so that the intensities ranged from $[0, 1 + b]$. The algorithm was successful with brightness value up to a value of $b = 0.5$, with an average intensity RMS error of 0.05, and an average registration RMS error of 0.5 pixels.

Sensitivity to contrast variations was also analyzed. The source image consisted of a synthetic fractal image (with intensities in the range $[0, 1]$), with a synthetic registration map applied. The target image consisted of the same synthetic image used in the source, with a random fractal contrast map multiplied to it, so that the intensities ranged from $[c, 1]$. The algorithm was

successful with contrast values $c \geq 0.5$, with an average intensity RMS error of 0.05, and an average registration RMS error of 0.5 pixels.

3.1.5 Geometric Smoothness

In this simulation, the ability of the registration algorithm to recover local geometric distortions of varying smoothness was analyzed. The target image consisted of a synthetic fractal image, while the source image consisted of the target image, with a known synthetic geometric distortion applied to it. The smoothness of this distortion field, Equation (28), was varied from 0.1 to 0.6 and was generated similar to the synthetic registration map, as described above; the only two differences being that 1) there was no global distortion applied to the registration map, and 2) the smoothness of the registration map was varied. The algorithm was successful with smoothness values less than 0.3, with an average intensity RMS error of 0.03, and an average registration RMS error of 0.4 pixels.

3.1.6 Noise

In this simulation, the ability of the registration algorithm to perform in the presence of additive uniform noise was analyzed. The target image consisted of a synthetic fractal image, and the source image consisted of the target image with uniformly distributed noise added to it. The intensity of the noise was modulated within the range [0, n], where n varies from 0 (no noise) to 1.0 (100% noise). The algorithm was successful with a PSNR value greater than 3db with an average intensity RMS error of 0.05, and an average registration RMS error of 0.4 pixels. We hypothesize that the registration is successful at very low signal to noise ratios because the estimated contrast and brightness terms absorb the additive noise. This hypothesis was tested by repeating the simulation, estimating only the geometric terms in the registration map. In this case, the average intensity RMS error was 0.13, and the average registration RMS error was 8 pixels, thus confirming that the brightness and contrast terms help contend with large amounts of additive noise.

3.1.7 Missing Data

In this simulation, the ability of the registration algorithm to recover a global geometric distortion when a portion of the data is missing was analyzed. In these simulations, the target image consisted of a random synthetic image, while the source image consisted of the target image with a random global distortion. The global distortion consisted of a translation in the range of 0 to 12 pixels, a rotation in the range 1 to 12 degrees and a scale change between 1 to 1.2. A square of varying size was removed from the source image at random

positions, starting with a size of 16×16 pixels to a size of 128×128 pixels (the entire image). The algorithm was able to successfully recover global distortions with missing data as large as 96×96 pixels, with an average intensity RMS error of 0.01, and an average registration RMS error of 0.2 pixels.

4 Discussion

We have presented a general-purpose registration algorithm. The geometric transformation is modeled as locally affine but globally smooth, and explicitly accounts for local and global variations in image intensities. An explicit model of missing data is also incorporated, allowing us to simultaneously segment and register images with partial or missing data. All of the components are combined within an integrated framework yielding a robust and effective registration algorithm within and across different imaging modalities. We have shown the efficacy of our algorithm on synthetic images/volumes and on clinical images/volumes. We have analyzed the robustness and sensitivity of our algorithm with respect to the various design assumptions. We find that accurate registration is possible under a significant amount of geometric and intensity variations, and missing data.

Our current implementation suffers from one primary shortcoming. On a 2.8 GHz Intel processor with 2 GB memory, the registration of a 256×256 image requires approximately 4 minutes, and the registration of a $64 \times 64 \times 64$ volume requires approximately 30 minutes. We hope that optimization of our algorithm and a C-based implementation, of the current MatLab code, will reduce this run-time.

In designing our registration algorithm we have tried to contend with what we consider to be two of the major obstacles facing general-purpose registration algorithms: intensity variations and missing data between the source and target images/volumes. By contending with these pressing issues within a unified framework, we believe that our registration algorithm will prove useful in a number of clinical settings.

References

[1] L. Brown, A survey of image registration techniques, ACM Computing Surveys 24 (4) (1992) 325–376.

[2] P. Van den Elsen, E.-J. Pol, M. Viergever, Medical image matching - a review with classification, IEEE Engineering in Medicine and Biology 12 (1) (1993) 26–39.

[3] C. Maurer Jr., J. Fitzpatrick, Interactive Image-Guided Neurosurgery, American Association of Neurological Surgeons, Park Ridge, IL., 1993, Ch. A Review of Medical Image Registration.

[4] J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer, R. Kessler, R. Maciunas, Comparison and evaluation of retrospective intermodality image registration techniques, in: Proceedings of the SPIE - The International Society for Optical Engineering, Newport Beach, CA, 1996, pp. 332–347.

[5] J. Maintz, M. Viergever, A survey of medical image registration, Medical Image Analysis 2 (1) (1998) 1–36.

[6] H. Lester, S. Arridge, A survey of hierarchical non-linear medical image registration, Pattern Recognition 32 (1) (1999) 129–149.

[7] P. H. Schonemann, A generalized solution of the orthogonal procrustes problem, Pschometrika 31 (1) (1966) 1–10.

[8] A. Evans, S. Marrett, D. Collins, T. Peters, Anatomical-functional correlative analysis of the human brain using three-dimensional imaging systems, in: Proceedings of the SPIE - The International Society for Optical Engineering, Vol. 1092, 1989, pp. 264–274.

[9] F. Bookstein, Thin-plate splines and the atlas problem for biomedical images, Information Processing in Medical Imaging (1991) 326–342.

[10] M. Nack, Rectification and registration of digital images and the effect of cloud detection, in: Machine Processing of Remotely Sensed Data, West Lafayette, IN, 1977, pp. 12–23.

[11] W. Kerwin, C. Yuan, Active edge maps for medical image registration, in: Proceedings of the SPIE - The International Society for Optical Engineering, 2001, pp. 516–526.

[12] G. Medioni, R. Nevatia, Matching images using linear features, IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (6) (1984) 675–685.

[13] W. Shih, W. Lin, C. Chen, Contour-model-guided nonlinear deformation model for intersubject image registration, in: Proceedings of the SPIE - The International Society for Optical Engineering, Vol. 3034, 1997, pp. 611–620.

[14] C. Pelizarri, G. Chen, D. Spelbring, R. Weichselbaum, C. Chen, Accurate three-dimensional registration of CT, PET and/or MR images of the brain, Journal of Computer Assisted Tomography 13 (1) (1989) 20–26.

[15] B. Fischl, M. Sereno, A. Dale, Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system, Neuroimage 9 (2) (1998) 195–207.

[16] A. Dale, B. Fischl, M. Sereno, Cortical surface-based analysis. I. Segmentation and surface reconstruction, Neuroimage 9 (2) (1999) 179–194.

[17] K. Arun, T. Huang, S. Blostein, Least-squares fitting of two 3-d point sets, IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (1997) 698–700.

[18] B. Horn, H. Hilden, S. Negahdaripour, Closed-form solution of absolute orientation using orthonormal matrices, Journal of the Optical Society of America 5 (1988) 1127–1135.

[19] B. Horn, Closed-form solution of absolute orientation using unit quaternions, Journal of the Optical Society of America 4 (1987) 629–642.

[20] R. Woods, S. Cherry, J. Mazziotta, Rapid automated algorithm for aligning and reslicing PET images, Journal of Computer Assisted Tomography 16 (1992) 620–633.

[21] R. Woods, S. Grafton, C. Holmes, S. Cherry, J. Mazziotta, Automated image registration: I. General methods and intrasubject, intramodality validation, Journal of Computer Assisted Tomography 22 (1998) 141–154.

[22] M. Unser, A. Aldroubi, C. Gerfen, A multiresolution image registration procedure using spline pyramids, in: Proceedings of the SPIE - Mathematical Imaging: Wavelets and Applications in Signal and Image Processing, Vol. 2034, 1993, pp. 160–170.

[23] P.J. Kostelec, J.B. Weaver, D.M. Healy, Jr., Multiresolution elastic image registration, Medical Physics 25 (9) (1998) 1593–604.

[24] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, D. Hawkes, Nonrigid registration using free-form deformations: application to breast MR images, IEEE Transactions on Medical Imaging 18 (8) (1999) 712–721.

[25] J. Kybic, P. Thévenaz, A. Nirkko, M. Unser, Unwarping of unidirectionally distorted epi images, IEEE Transactions on Medical Imaging 19 (2) (2000) 80–93.

[26] F. Bookstein, Principal Warps: Thin-plate splines and the decomposition of deformations, IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (6) (1989) 567–585.

[27] J. Ashburner, K. Friston, Multimodal Image Coregistration and Partitioning - a Unified Framework, NeuroImage 6 (3) (1997) 209–217.

[28] J. Ashburner, P. Neelin, D. Collins, A. Evans, K. Friston, Incorporating prior knowledge into image registration, NeuroImage 6 (1997) 344–352.

[29] C. Broit, Optimal registration of deformed images, Ph.D. thesis, University of Pennsylvania, Pennsylvania, USA (1981).

[30] R. Bajcsy, S. Kovacic, Multiresolution elastic matching, Computer Vision, Graphics and Image Processing 46 (1989) 1–21.

[31] J. Gee, M. Reivich, R. Bajcsy, Elastically deforming an atlas to match anatomical brain images, Journal of Computer Assisted Tomography 17 (2) (1994) 1–21.

[32] G. Christensen, R. Rabbit, M. Miller, 3-D brain mapping using a deformable neuroanatomy, Physics in Medicine and Biology 39 (1994) 609–618.

[33] G. Christensen, R. Rabbit, M. Miller, A deformable neuroanatomy textbook based on viscous fluid mechanics, in: Proceedings of the 1993 Conference on Information Sciences and Systems, Johns Hopkins University, 1995, pp. 211–216.

[34] D. Mishra, A. Chan, C. Chui, Histogram equalization, image registration, and data fusion for multispectral images, in: Proceedings of the SPIE - The International Society for Optical Engineering, Vol. 2496, 1995, pp. 1025–1031.

[35] J. Vlad, S. Eulenstein, W. Wlodarczyk, P. Wust, R. Felix, Registration of portal images for online correction of positioning errors during radiation therapy of prostate cancer, in: Proceedings of the SPIE - The International Society for Optical Engineering, Vol. 4319, 1995, pp. 61–68.

[36] B. Nutter, S. Mitra, T. Krile, Image registration algorithm for a pc-based system, in: Proceedings of the SPIE - The International Society for Optical Engineering, Vol. 829, 1987, pp. 214–221.

[37] E. D. Castro, C. Morandi, Registration of translated and rotated images using finite Fourier transforms, IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (5) (1987) 700–703.

[38] W. Eddy, M. Fitzgerald, D. Noll, Improved image registration by using Fourier interpolation, Magnetic Resonance in Medicine 36 (1996) 923–931.

[39] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, G. Marchal, Information Processing in Medical Imaging, Kluwer, Dordrecht, 1995, Ch. Automated multimodality image registration using information theory, pp. 263–274.

[40] P. Viola, W. Wells III, Alignment by maximization of mutual information, in: IEEE International Conference on Computer Vision, Cambridge, MA, USA, 1995, pp. 16–23.

[41] C. Studholme, D. Hill, D. Hawkes, Incorporating connected region labelling into automated image registration using mutual information, Mathematical methods in biomedical image analysis (1996) 23–31.

[42] P. Thévenaz, M. Unser, Spline pyramids for intermodal image registration using mutual information, in: Proceedings of the SPIE - The International Society for Optical Engineering, Vol. 3169, 1997, pp. 236–247.

[43] T. Buzug, J. Weese, C. Fassnacht, C. Lorenz, Using an entropy similarity measure to enhance the quality of DSA images with an algorithm based on template matching, in: Visualization in Biomedical Computing, Vol. 1131, Hamburg, Germany, 1996, pp. 235–240.

[44] C. Studholme, D. Hill, D. Hawkes, An overlap invariant entropy measure of 3-D medical image alignment, Pattern Recognition 32 (1) (1999) 71–86.

[45] L. Lemieux, U. Wieshmann, N. Moran, D. Fish, S. Shorvon, The detection and significance of subtle changes in mixed-signal brain lesions by serial MRI scan matching and spatial normalization, Medical Image Analysis 2 (3) (1998) 227–242.

[46] A. Roche, G. Malandain, N. Ayache, S. Prima, Towards a better comprehension of similarity measures used in medical image registration, in: Medical Image Computing and Computer Assisted Intervention, Vol. 1679, 1999, pp. 555–566.

[47] O. Nestares, D.J. Heeger, Robust Multiresolution Alignment of MRI Brain Volumes, Magnetic Resonance in Medicine 43 (2000) 705–715.

[48] P. Hellier, C. Barillot, E. Memin, P. Perez, Hierarchical estimation of a dense deformation field for 3-d robust registration, IEEE Transactions on Medical Imaging 20 (5) (2001) 388–402.

[49] R. Bansal, L. Staib, Z. Chen, A. Rangarajan, J. Knisely, R.Nath, J. Duncan, A novel approach for the registration of 2D and 3D CT images for treatment setup verification in radiotherapy, in: Medical image computing and computer-assisted intervention (MICCAI), 1998, pp. 1075–1086.

[50] S. Periaswamy, H. Farid, Elastic registration in the Presence of Intensity Variations, IEEE Transactions on Medical Imaging 22 (7) (2003) 865–74.

[51] S. Periaswamy, General-purpose medical image registration, Ph.D. thesis, Dartmouth College, Department of Computer Science, Hanover, NH (2003).

[52] B. Horn, Robot Vision, MIT Press, Cambridge, MA, 1986.

[53] J. Barron, D. Fleet, S. Beauchemin, Performance of optical flow techniques, International Journal of Computer Vision 12 (1) (1994) 43–77.

[54] A. Dempster, N. Laird, D. Rubin, Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society 99 (1) (1977) 1–38.

[55] S. Negahdaripour, C.-H. Yu, A generalized brightness change model for computing optical flow, in: International Conference of Computer Vision, Berlin, Germany, 1993, pp. 2–11.

[56] J. Shi, C. Tomasi, Good features to track, in: Computer Vision and Pattern Recognition, Seatle, WA, USA, 1994, pp. 593–600.

[57] H. Farid, E. Simoncelli, Differentiation of multi-dimensional signals, IEEE Transactions on Image Processing 13 (4) (2004) 496–508.

[58] B. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, in: International Joint Conference on Artificial Intelligence, Vancouver, 1981, pp. 674–679.

[59] P. Anandan, A computational framework and an algorithm for the measurement of visual motion, International Journal of Computer Vision 2 (3) (1989) 283–310.

[60] A. P. Pentland, Fractal based description of natural scenes, IEEE Transactions on Pattern and Machine Intelligence 6 (6) (1984) 661–674.