Multi-modal image set registration and atlas formation

Peter Lorenzen [a,*], Marcel Prastawa [a], Brad Davis [a], Guido Gerig [a,b],
Elizabeth Bullitt [a,c,d], Sarang Joshi [a,e]

[a] Department of Computer Science, University of North Carolina, Chapel Hill, NC 27599, USA
[b] Department of Psychiatry, University of North Carolina, Chapel Hill, NC 27599, USA
[c] Department of Surgery, University of North Carolina, Chapel Hill, NC 27599, USA
[d] Department of Radiology, University of North Carolina, Chapel Hill, NC 27599, USA
[e] Department of Radiation Oncology, University of North Carolina, Chapel Hill, NC 27599, USA

Abstract

In this paper, we present a Bayesian framework for both generating inter-subject large deformation transformations between two multi-modal image sets of the brain and for forming multi-class brain atlases. In this framework, the estimated transformations are generated using maximal information about the underlying neuroanatomy present in each of the different modalities. This modality independent registration framework is achieved by jointly estimating the posterior probabilities associated with the multi-modal image sets and the high-dimensional registration transformations mapping these posteriors. To maximally use the information present in all the modalities for registration, Kullback–Leibler divergence between the estimated posteriors is minimized. Registration results for image sets composed of multi-modal MR images of healthy adult human brains are presented. Atlas formation results are presented for a population of five infant human brains.
© 2005 Elsevier B.V. All rights reserved.

1. Introduction

With the increasing number of imaging techniques and imaging sensors, multi-modal image registration has become an active area of research in medical image analysis. An increasingly important area of medical image analysis is computational anatomy, the study of anatomical variation. Understanding anatomical variability requires robust high-dimensional image registration methods where the number of parameters used to describe the mappings between images is on the order of the number of voxels describing the space of the images.

Modern imaging techniques provide an array of imaging modalities which enable the acquisition of complementary information representing an underlying anatomy. Most image registration algorithms find a mapping between two scalar images. To utilize multi-modal images of a single anatomy, we define a multi-modal image set, $\bar{I}$, as a collection of m co-registered multi-modal images where, for a given spatial location $x, \bar{I}(x) \in \mathbb{R}^m$. For example, $\bar{I}$ might represent a CT image, a T1-weighted MR image, and a PET image of a single

* Corresponding author. Address: Medical Image Display Analysis Group, The University of North Carolina at Chapel Hill, CB# 3175, Sitterson Hall, Chapel Hill, NC 27599-3175, USA. Tel.: +1 919 962 1879; fax: +1 919 962 1699.

E-mail addresses: lorenzen@cs.unc.edu (P. Lorenzen), prastawa@cs.unc.edu (M. Prastawa), davisb@cs.unc.edu (B. Davis), gerig@cs.unc.edu (G. Gerig), bullitt@med.unc.edu (E. Bullitt), joshi@cs.unc.edu (S. Joshi).

anatomy. We assume that the underlying neuroanatomy, represented in two acquired sets of multi-modal images, consists of a set, C, of separate anatomical structure classes, c_j. Throughout this paper, we assume that, for a given subject, the multi-modal images of that subject are co-registered.

Many image registration techniques that register probability functions involve information theoretic distance measures. Mutual information is one such method that is typically used to register multi-modal images. High-dimensional image registration in the context of mutual information and other dissimilarity measures has been studied extensively. A thorough investigation of these dissimilarity measures in high-dimensional image registration is presented in (Hermosillo, 2002). Mutual information is, equivalently, the Kullback–Leibler divergence between the joint distribution and product of marginal distributions of two random variables. A multi-modal free-form registration algorithm that matches voxel class labels, rather than intensities, via minimizing Kullback–Leibler divergence, another information theoretic distance measure, is presented in (D'Agostino et al., 2003a,b). This method only finds correspondences between two scalar images. A method that minimizes Kullback–Leibler divergence between expected and observed joint class histograms is presented in (Chan et al., 2003). This technique, however, estimates class labels as a preprocessing step and is used only for rigid registration between scalar images. The method presented in this paper is more general in that registration is performed on sets of images, of arbitrary number, and is not constrained by an initial class labeling. Although inter-subject high-dimensional image registration has received much attention (Rueckert et al., 1998; Miller et al., 1999; Gaens et al., 1998; Studholme et al., 1999), to our knowledge, little attention has been given to the use of multi-modal image sets in image registration. The foundation for the work presented in this paper have been proposed in (Lorenzen and Joshi, 2003; Lorenzen et al., 2004a; Davis et al., 2004; Lorenzen et al., 2004b).

1.1. Model-based multi-modal image set registration

Across image sets, the number of constituent images may vary, thus registration based on an intensity similarity measure is not possible in this setting. While mutual information can be extended to multiple random variables, its extension to registration involving three or more images is problematic in that it requires maintaining an impractical number of histogram bins (Bhatia et al., 2004). Consider a multi-modal image set registration involving four twelve-bit DICOM images, an example of which is provided in Section 5.2. Using mutual information would require the construction of a $2^{4.12} \approx 2.8 \times 10^{14}$ bin joint histogram, which is impractical.

Given these difficulties we move to a model-based approach where the registration is performed using underlying anatomical structures. We incorporate these anatomical structures as a prior in a Bayesian framework.

This framework is based on the assumption that human brain anatomy consists of finitely enumerable structures such as grey matter, white matter, and cerebrospinal fluid. These structures present with varying radiometric intensity values across disparate image modalities. Given two multi-modal image sets, we capture the underlying structures by estimating, for each image set, the class posterior mass functions associated with each of the structures. These class posteriors are then used to produce a mean posterior atlas by estimating high-dimensional diffeomorphic registration maps relating the coordinate spaces of the class posteriors. The Kullback–Leibler divergence is used as a distance function on the space of probability mass functions to estimate the transformations. The use of the class posteriors provides an image intensity independent approach to image registration.

1.2. Multi-class atlas formation

An important problem in computational anatomy is the construction of an exemplar atlas from a population of medical images. This atlas represents the anatomical variation present in the population (Miller et al., 1997; Grenander and Miller, 1998; Thompson et al., 2000). Many images are mapped into a common coordinate system to study intra-population variability and inter-population differences, to provide voxel-wise mapping of functional sites, and facilitate tissue and object segmentation via registration of anatomical labels. Common techniques for creating atlases often include the choice of a template image, which inherently produces a bias. Motivated by the atlas construction framework presented in (Joshi et al., 2004), we propose the construction of an unbiased multi-class atlas from a population of anatomical class posteriors using large deformation diffeomorphic registration. When applied to two image sets, this atlas formation method yields inverse-consistent image registration.

1.3. Inverse consistent registration

Many registration algorithms are not inverse consistent since their dissimilarity metrics are computed in the coordinate system of either one of the images involved in the registration. This leads to order non-preservation of optimization energy cost functions. In traditional techniques for image registration, solutions may be systematically biased with respect to expanding and contracting regions in the estimated transformation (Cachier and Rey, 2000). Inverse consistent registration

is desired when there is no preference, or believability, for one image over another. Existing methods for generating inverse consistent registration approximate inverse consistency by adding an inverse consistency penalty to the optimization cost function. The registration frameworks formulated in these methods are not intrinsically symmetric. A method for approaching this problem, involving an algorithm that estimates incremental transformations while approximating inverse consistency constraints on each incremental transformation, is presented in (He and Christensen, 2003). The approach presented in this paper is intrinsically inverse consistent as the registration problem is formulated symmetrically. Therefore, no correction penalty for consistency is required.

The remainder of this paper is organized as follows. Sections 2 and 3 cover the methodology of the multi-class atlas formation and the specialization to two image set registration, results for these methods are presented in Sections 4 and 5, respectively, and finally, a discussion and conclusion are presented in Sections 6 and 7.

2. Methodology: atlas formation

In this section, we present the atlas formation framework and the motivation for the use of Kullback–Leibler divergence to drive the registration. We begin by describing the Bayesian framework for representing the anatomical class structures. The specialization to image set registration involving two image sets is presented in Section 3.

2.1. Bayesian framework

From a population of N multi-modal image sets $\{\bar{I}_i\}_{i=1}^N$, for each class $c_j \in C$, we first estimate the class posterior mass functions $p_i(c_j(x)|\bar{I}_i)$ for each image set i where $c_j(x)$ is the class associated with the voxel at spatial position $x \in \Omega \subset \mathbb{R}^3$. This method is independent of the number of images comprising each image set. These class posteriors are produced using the expectation maximization method described in (Prastawa et al., 2003) and (van Leemput et al., 1999a). Following van Leemput et al. (1999a), for each class c_j the associated data likelihood, $p(\bar{I}_i(x)|c_j(x), \mu_j, \Sigma_j)$, is modeled as a normal distribution with mean, μ_j, and covariance, Σ_j.

2.2. Large deformation multi-class atlas formation

We now consider the problem of estimating an atlas class posterior $\hat{p}$ that is the best representative for a population of N class posteriors, $\{p_i\}_{i=1}^N$, representing the N individual image sets $\{\bar{I}_i\}_{i=1}^N$. The atlas $\hat{p}$ is not a member of the set $\{p_i\}$. To this end, we consider the problem of constructing a mapping between $\hat{p}$ and each class poster-

ior in the set $\{p_i\}$. That is, we estimate the mappings $h_i: \Omega \to \Omega_i$ where $\Omega \subset \mathbb{R}^3$ and $\Omega_i \subset \mathbb{R}^3$ are the coordinate systems of the class posteriors $\hat{p}$ and p_i, respectively. The coordinate system Ω is chosen to be independent of the individual population class posterior coordinate systems, Ω_i. This framework is depicted in Fig. 1.

We seek the representative atlas class posterior $\hat{p}$ that requires the minimum amount of energy to deform into every population class posterior p_i. More precisely, given a transformation group S with associated metric $D: S^2 \to \mathbb{R}$, along with a probability density dissimilarity measure $E(p, q)$, we wish to find the class posterior density $\hat{p}$ such that

$$\{\hat{h}_i, \hat{p}\} = \arg\min_{h_i \in S, p} \sum_{i=1}^N E(p_i \circ h_i, p) + D^2(e, h_i), \tag{1}$$

where e is the identity transformation.

In this paper, we focus on the infinite dimensional group of diffeomorphisms $\mathcal{H}$ where we apply the theory of large deformation diffeomorphisms (Miller et al., 1999; Joshi and Miller, 2000) to generate deformations h_i that are solutions to the Lagrangian ODEs $\frac{d}{dt} h_i(x, t) = v_i(h(x, t), t)$, $t \in [0, 1]$. The transformations h_i are generated by integrating velocity fields forward in time and h_i^{-1} are generated by integrating velocity fields backward in time. The relationship between spatial locality, velocity fields, and time is shown in
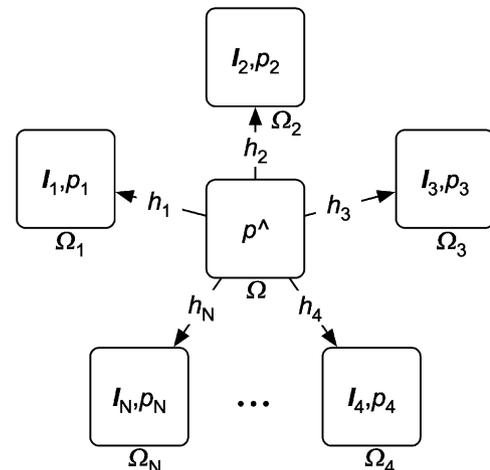
Fig. 1. Atlas formation. Unbiased atlas constructed as the intrinsic mean of a population of class posteriors.
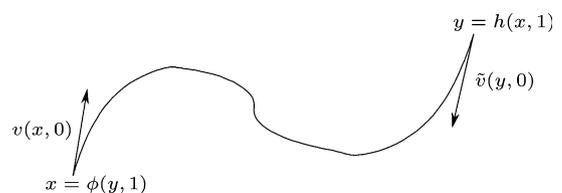
Fig. 2. Velocity fields.

Fig. 2. The spatial location y is described in terms of the forward integration of the velocity field v starting from spatial location x, that is, $y = h(x,1) = x + \int_0^1 v(h(x,\tau),\tau)d\tau$. Similarly, x can be described in terms of integrating the reverse velocity field $\tilde{v}$ starting at y, that is, $x = \phi(y,1) = y + \int_0^1 \tilde{v}(\phi(y,\tau),\tau)d\tau$. From this figure we note that $v(h(x,t),t) = -\tilde{v}(\phi(y,1-t),1-t)$ and, hence, $\|Lv(x,t)\|^2 = \|L\tilde{v}(y,1-t)\|^2$, where $L = \alpha\nabla^2 + \beta\nabla(\nabla\cdot) + \gamma$ is the Navier–Stokes operator.

We induce a metric on the space of diffeomorphisms by using a Sobolev norm (a norm that involves derivatives of a function) via the partial differential operator L on the velocity fields v. Let h be a diffeomorphism isotopic to the identity transformation e. We define the squared distance $D^2(e,h)$ as

$$D^2(e,h) = \min_v \int_0^1 \int_\Omega \|Lv(x,t)\|^2 \, dx \, dt$$

subject to

$$h(x) = x + \int_0^1 v(h(x,t),t) \, dt.$$

The distance between any two diffeomorphisms is defined by

$$D(h_1,h_2) = D(e, h_1^{-1} \circ h_2).$$

The construction of h and h^{-1}, as well as the properties of D, are described in (Miller and Younes, 2001) and (Miller et al., 2002).

Having defined a metric on the space of diffeomorphism and a regularization operator L, the energy minimization problem described in Eq. (1) is formulated as

$$\{\hat{h}_i, \hat{p}\} = \arg\min_{h_i,p} \sum_{i=1}^N E(p_i \circ h_i, p) + \int_0^1 \int_\Omega \|Lv_i(x,t)\|^2 \, dx \, dt \tag{2}$$

subject to

$$h_i(x) = x + \int_0^1 v_i(h_i(x,t),t) \, dt.$$

2.3. Kullback–Leibler divergence

As a measure of dissimilarity between two class posteriors, we use Kullback–Leibler divergence (relative entropy).

Definition 1. Let p and q be probability mass functions on a set C. The Kullback–Leibler divergence (Kullback and Leibler, 1951) between p and q is defined as

$$D_{KL}(p\|q) = \sum_{c_j \in C} p(c_j) \log \frac{p(c_j)}{q(c_j)}.$$

The Information Inequality theorem provides the basic properties of $D_{KL}(p\|q)$:

Theorem (Information Inequality). *Let $p(c_j)$, $q(c_j)$ be two probability mass functions on the set C. Then $D_{KL}(p\|q) \geqslant 0$ with equality if and only if $p(c_j) = q(c_j)$ for all $c_j \in C$.*

Proof. By an application of Jensen's inequality, see (Cover and Thomas, 1991). $\square$

In our setting, we use the Kullback–Leibler divergence as a measure of dissimilarity between the two probability mass functions p_Ω and p_i at spatial location $x \in \Omega$,

$$D_{KL}(p_\Omega(x)\|p_i(x)) = \sum_{c_j \in C} p_\Omega(c_j(x)) \log \frac{p_\Omega(c_j(x))}{p_i(c_j(x))}.$$

From an information theoretic viewpoint (Cover and Thomas, 1991), this dissimilarity can be interpreted as the inefficiency of assuming that $p_i(x)$ is true when $p_\Omega(x)$ is true. That is, if we have a model expressed as a probability mass function $p_i(x)$, we can then measure how far an observation, also expressed as a probability mass function, $p_\Omega(x)$, deviates from $p_i(x)$ using Kullback–Leibler divergence.

2.3.1. Bayes probability of error

The use of Kullback–Leibler divergence as a distance function is appropriate in this setting as it provides a lower bound on the Bayes probability of error, $P(error)$, between two probability mass functions. Specifically, reducing $D_{KL}(\cdot)$ increases a lower bound on $P(error)$.

Motivated by Kailath (1967), we can consider the two-hypothesis decision-theory problem of classifying an observation as coming from one of two possible hypotheses H_1, the average class posterior $p_\Omega(x)$, and H_2, the class posterior of one of the deformed image sets $\bar{I}_i$, $p_i(h_i(x))$. Let $q(H_1)$ and $q(H_2)$ denote the a prior probabilities on the two hypotheses, and let $q(c|H_1)$ and $q(c|H_2)$ denote the class conditional probability density functions given the true hypotheses. For an observed class structure, c, (e.g., white matter), the posterior probability of H_k is

$$q(H_k|c) = \frac{q(c|H_k)q(H_k)}{q(c|H_1)q(H_1) + q(c|H_2)q(H_2)}$$

for $k = 1,2$. To minimize the probability of error, choose the hypothesis with the larger posterior probability. Therefore,

$$P(error|c) = \min\left\{\frac{q(c|H_1)q(H_1)}{q(c|H_1)q(H_1) + q(c|H_2)q(H_2)}, \frac{q(c|H_2)q(H_2)}{q(c|H_1)q(H_1) + q(c|H_2)q(H_2)}\right\}$$

and, hence,

$$P(error) = \mathscr{E}_c[P(error|c)]$$
$$= \sum_{c \in C} \min\{q(c|H_1)q(H_1), q(c|H_2)q(H_2)\},$$
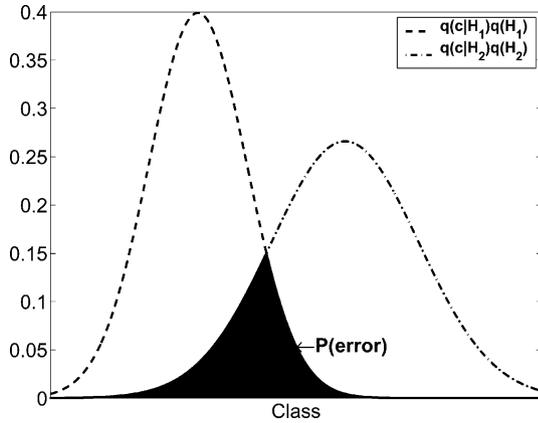
Fig. 3. Bayes probability of error for two distributions. Let $q(H_1)$ and $q(H_2)$ represent the a priori probabilities of hypothesis H_1 and H_2 being true, respectively, and $q(c|H_1)$ and $q(c|H_2)$ be the hypothesis-conditioned likelihoods for class c. In the context of neuroanatomical matching, c could be grey matter, white matter, or cerebrospinal fluid. The densities $q(c|H_1)q(H_1)$ and $q(c|H_2)q(H_2)$ correspond to the atlas class posterior $p_\Omega(x)$ and the class posterior of the deformed image set $\bar{I}_i$, p_i, at a given spatial location x.

where $\mathscr{E}_c$ denotes the expectation with respect to c. The probability of error is graphically depicted in Fig. 3. In our registration problem we want to *maximize P(error)*. That is, the closer these posterior distributions become, the larger the probability of mistaking one for the other, and, hence, an increase in indistinguishability.

In practice, real distributions are not known so the $P(error)$ cannot be directly computed. Given this reality, we look for a lower-bound on $P(error)$ which is maximized when $D_{\mathrm{KL}}(\cdot)$ is minimized.

2.3.2. Bounds on P(error) for our two-hypothesis problem

One of the first divergence measures involving density ratios is the Jeffreys divergence (Jeffreys, 1946),

$$J(q_1\|q_2) = D_{\mathrm{KL}}(q_1\|q_2) + D_{\mathrm{KL}}(q_2\|q_1).$$

As will be described in Section 3.2, this symmetric form of Kullback–Leibler divergence will be used as the distance measure to drive the registration. Using inequalities derived in (Hoeffding and Wolfowitz, 1958) a lower bound on $P(error)$ in terms of $J(q_1\|q_2)$ is presented in (Kailath, 1967) and (Toussaint, 1972), $\frac{1}{4}e^{-\frac{J}{2}} \leqslant P(error)$. Thus, a reduction in $D_{\mathrm{KL}}(\cdot)$ leads to an increase in the lower bound of $P(error)$. It should be noted, that while we have defined Jeffreys divergence in terms of symmetric Kullback–Leibler divergence, Sir Harold Jeffreys published his divergence measure in 1946 (Jeffreys, 1946) while Solomon Kullback and Richard Leibler published theirs in 1951 (Kullback and Leibler, 1951).

2.4. Registration

Under the Kullback–Leibler divergence measure, Section 2.3, the atlas estimation problem in Eq. (2) becomes

$$\hat{h}_i, \hat{p} = \arg\min_{h_i, p} \sum_{i=1}^{N} \int_\Omega D_{\mathrm{KL}}(p(x)\|p_i(h_i(x)))\,\mathrm{d}x$$
$$+ \int_0^1 \int_\Omega \|Lv_i(x,t)\|^2\,\mathrm{d}x\,\mathrm{d}t. \tag{3}$$

This minimization problem can be simplified by noticing that for fixed transformations h_i, the $\hat{p}$ that minimizes Eq. (3) is given by normalized geometric mean of the deformed class posteriors, $p_i(h_i(x))$,

$$\hat{p}(c_j(x)) = \frac{\left(\prod_{i=1}^{N} p_i(h_i(c_j(x)))\right)^{\frac{1}{N}}}{\sum_{c_k \in C}\left(\prod_{i=1}^{N} p_i(h_i(c_k(x)))\right)^{\frac{1}{N}}}. \tag{4}$$

Combining Eqs. (3) and (4) results in the following minimization problem

$$\hat{h}_i = \arg\min_{h_i} \sum_{i=1}^{N} \int_\Omega D_{\mathrm{KL}}(\hat{p}(x)\|p_i(h_i(x)))\,\mathrm{d}x$$
$$+ \int_0^1 \int_\Omega \|Lv_i(x,t)\|^2\,\mathrm{d}x\,\mathrm{d}t. \tag{5}$$

Note that the solution to this minimization problem is independent of the ordering of the N image sets and increases linearly as image sets are added, thus, making the algorithm scalable.

2.5. Implementation

We use Christensen's greedy algorithm for propagating templates (Christensen et al., 1996). In the atlas formation setting, the velocity v_i^n for each iteration n is computed as follows. First, compute the updated atlas estimate (i.e. the normalized geometric mean)

$$\hat{p}^n(c_j(x)) = \frac{\left(\prod_{i=1}^{N} p_i(h_i^n(c_j(x)))\right)^{\frac{1}{N}}}{\sum_{c_k \in C}\left(\prod_{i=1}^{N} p_i(h_i^n(c_k(x)))\right)^{\frac{1}{N}}}$$

for each class component c_j. Next, using the second order approximation to Kullback–Leibler divergence, we define the body force functions

$$F_i^n(x) = \sum_{c_k \in C}\left[\frac{p_i(h_i(c_k(x)))}{\hat{p}(c_k(x))} - 2\right]\nabla p_i\bigg|_{c_k(h_i(x))}.$$

This is the variation of the class posterior dissimilarity term in Eq. (5) with respect to the transformation h_i. The velocity field v_i^n is computed at each iteration by applying the inverse of the differential operator L to the body force function, i.e. $v_i^n(x) = L^{-1}F_i^n(x)$. This computation is performed in the Fourier domain (Joshi et al., 2003).

The forward and inverse integration is described as follows. At time t the transformations h_i are described as

$$h_i(x, t+\delta) = h_i(x,t) + \int_t^{t+\delta} v_i(h_i(x,\tau),\tau)\,\mathrm{d}\tau$$
$$\approx h_i(x,t) + \delta v_i(h_i(x,t),t)$$

for small δ. At iteration k of the algorithm, the transformations h_i become the telescoping compositions $h_i = h_i^1 \circ h_i^2 \circ \cdots \circ h_i^k$. At time t the inverse transformations h_i^{-1} are described as

$$h_i^{-1}(y,t) = h_i^{-1}\left(y - \int_t^{t-\delta} v_i(y,\tau)\,\mathrm{d}\tau, t-\delta\right)$$
$$\approx h_i^{-1}(y - \delta v_i(y,t), t-\delta)$$

for small δ. At iteration k of the algorithm, the transformations h_i^{-1} become the telescoping compositions $h_i^{-1} = h_i^{-1,k} \circ h_i^{-1,k-1} \circ \cdots \circ h_i^{-1,1}$.

3. Methodology: image set registration

In this section, we present the specialization of the atlas formation problem of Section 2 to the problem of image set registration involving two image sets. That is, we consider the problem of finding a mapping between image sets $\bar{I}_1$ and $\bar{I}_2$ (Fig. 4). We would like to find the mappings $f:\Omega_1 \to \Omega_2$ and $g:\Omega_2 \to \Omega_1$ where Ω_1 and Ω_2 are the coordinate systems of image sets $\bar{I}_1$ and $\bar{I}_2$, respectively. Again, we introduce a new coordinate system Ω, independent of Ω_1 and Ω_2. Let transformations h_1 and h_2 map Ω to Ω_1 and Ω_2, respectively. By construction, $f = h_2 \circ h_1^{-1}$ and $g = h_1 \circ h_2^{-1}$. This registration method is inverse consistent as $f \circ g = g \circ f = e$, the identity map.

3.1. Bayesian framework

From the multi-modal image sets $\bar{I}_1$ and $\bar{I}_2$, for each class $c_j \in C$ we jointly estimate the posterior mass functions $p_1(x) = p(c_j(h_1(x))|\bar{I}_1)$ and $p_2(x) = p(c_j(h_2(x))|\bar{I}_2)$ along with the registration maps $h_1(x)$ and $h_2(x)$, that map the independent coordinate space $\Omega \subset \mathbb{R}^3$, into the space of $\bar{I}_1$, $\Omega_1 \subset \mathbb{R}^3$, and $\bar{I}_2$, $\Omega_2 \subset \mathbb{R}^3$, respectively. This method is independent of the choice of the number of images comprising each image set. Optimal inter-subject multi-modal image registration is estimated by
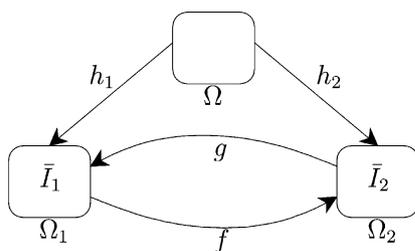
Fig. 4. Registration framework. Registration of image sets $\bar{I}_1$ and $\bar{I}_2$ through the unbiased coordinate space Ω.

an alternating iterative algorithm which is motivated by an expectation maximization method used in (Prastawa et al., 2003; van Leemput et al., 1999a). Our algorithm interleaves the estimation of the posteriors associated with $\bar{I}_1$ and $\bar{I}_2$ and the estimation of the registration maps $h_1:\Omega \to \Omega_1$ and $h_2:\Omega \to \Omega_2$.

As defined in Section 2.1, for each class c_j the associated data likelihoods, $p(\bar{I}_{\{1,2\}}(x)|c_j(x),\mu_j,\Sigma_j)$, are modeled as a normal distributions with means, μ_j, and covariances, Σ_j. Given the transformations h_1 and h_2 and the current estimates μ_j and Σ_j for both image sets, the posteriors of $\bar{I}_1$ and $\bar{I}_2$ are associated with the independent coordinate probability mass function p_Ω by using Bayes's Rule with p_Ω as the prior for both posteriors $p_1(x)$ and $p_2(x)$. Having defined the posteriors, the parameters μ_j and Σ_j are updated by their expected values. We are currently investigating the use of kernel density estimation as a replacement for the Gaussian models as described in (Prastawa et al., 2004).

3.2. Registration

At a given spatial location $x \in \Omega$, the dissimilarity between image sets $\bar{I}_1(x)$ and $\bar{I}_2(x)$ is measured by the dissimilarity between the posterior mass functions modeling them, $p_1(x)$ and $p_2(x)$. As we have seen in Section 2.3.2, minimizing the Jeffrey's divergence between two probability mass functions increases the lower bound on the Bayes' probability of error in the two-hypothesis decision problem, and, thus, renders the probability mass functions more indistinguishable. That is, brings them closer together. The following distance is used to drive the registration

$$D(p_1(x),p_2(x))$$
$$= \underset{p_\Omega(x)}{\arg\min}\left[D_{\mathrm{KL}}(p_\Omega(x)\|p_1(x)) + D_{\mathrm{KL}}(p_\Omega(x)\|p_2(x))\right].$$
(6)

From Eq. (6), $D(p_1(x),p_2(x)) = D(p_2(x),p_1(x))$. For known transformations h_1 and h_2 the probability mass function in the independent coordinate system, $p_\Omega(x)$, that minimizes the distance above is the normalized geometric mean

$$p_\Omega(c_j(x)) = \frac{(p_1(c_j(x))p_2(c_j(x)))^{\frac{1}{2}}}{\sum_{c_k \in C}(p_1(c_k(x))p_2(c_k(x)))^{\frac{1}{2}}}.$$

Thus, the dissimilarity measure can be expressed wholly in terms not involving the independent coordinate system. After substituting the above expression for p_Ω into Eq. (6) we obtain the following distance at position $x \in \Omega$

$$D(p_1(x),p_2(x))$$
$$= -2\log\sum_{c_j \in C}\left(p_1(h_1(c_j(x))|\bar{I}_1)p_2(h_2(c_j(x))|\bar{I}_2)\right)^{\frac{1}{2}}.$$

With this result we re-write the minimization problem stated in Eq. (1) as follows

$$\hat{v}_1, \hat{v}_2 = \arg\min_{v_1, v_2} \int_\Omega \log \sum_{c_j \in C} \left(p_1(h_1(c_j(x))|\bar{I}_1) p_2(h_2(c_j(x))|\bar{I}_2) \right)^{\frac{1}{2}} dx$$
$$+ \int_0^1 \int_\Omega \|Lv_1(x)\|^2 \, dx \, dt + \int_0^1 \int_\Omega \|Lv_2(x)\|^2 \, dx \, dt.$$

3.3. Implementation

As described in Section 2.5, we compute the variation for h_1 of the average $D(p_1(x), p_2(x))$ term

$$\frac{\partial}{\partial h_1} \frac{1}{|\Omega|} \int_\Omega D(p_1(x), p_2(x)) \, dx$$
$$= -\frac{1}{|\Omega|} \int_\Omega \frac{\sum_{c_j \in C} \left(\frac{p_2(c_j(x))}{p_1(c_j(x))}\right)^{\frac{1}{2}} \nabla p_1 \big|_{c_j(h_1(x))}^{\mathrm{T}}}{\sum_{c_k \in C} (p_1(c_k(x)) p_2(c_k(x)))^{\frac{1}{2}}} \, dx.$$

In a similar manner the variation for h_2 is computed. The velocity fields $v_{\{1,2\}}$ at each iteration are updated by solving the partial differential equations,

$$Lv_{\{1,2\}}(x, t) = \frac{\partial}{\partial h_{\{1,2\}}} \frac{1}{|\Omega|} \int_\Omega D(p_1(x), p_2(x)) \, dx.$$

4. Results: atlas formation

To evaluate the performance of the atlas formation method, we applied the algorithm to a set of five class posterior mass functions that where derived from a pop- ulation of T1-weighted, T2-weighted, and proton density 3D MR images of brains of health two-year-old children, acquired at UNC Chapel Hill, using an expectation maximization segmentation method (van Leemput et al., 1999a,b). As a preprocessing step, these images were aligned using affine registration. An axial slice from each derived class posterior mass function is shown in Fig. 5. There is noticeable variation between these anatomies, especially in the ventricular region.

Fig. 6 shows the normalized geometric mean of the five class posterior mass functions and the final estimate of the atlas. The normalized geometric mean is blurry since it is an "average" of the varying individual neuroanatomies. Ghosting is evident around the lateral ventricles and near the boundary of the brain. In the final
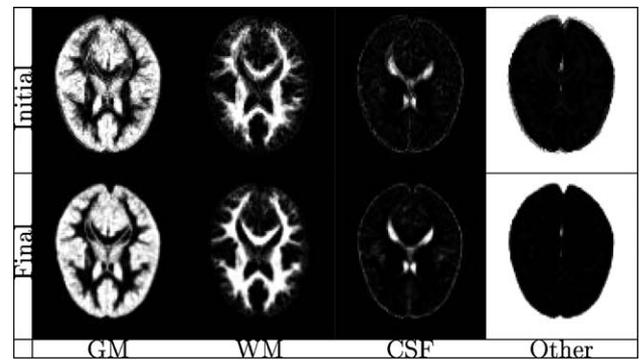
Fig. 6. Atlas construction. The top row shows the normalized geometric mean class posterior density following an affine registration of all five subjects. The bottom row represents the estimated atlas after the final iteration of the algorithm.
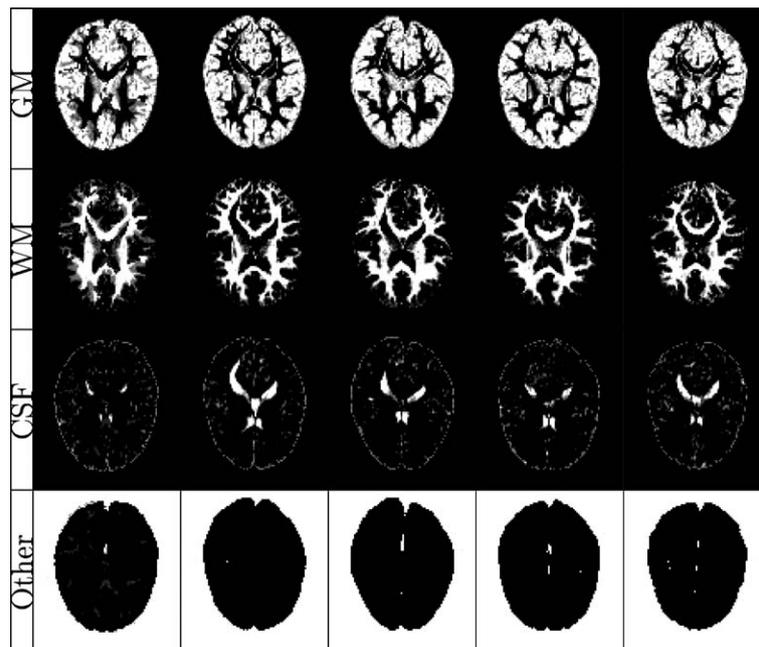
Fig. 5. Population of class posteriors. Five class posteriors (columns) each with four classes (rows). These images clearly show the large inter-subject variability, especially in the ventricular system.

estimate of the atlas these variations have been accommodated by the high-dimensional registration.

5. Results: image set registration

To evaluate the image set registration method, we defined a collection of image sets of increasing complexity. For each image subject, the image sets were created using 3D scalar images from a population of four imaging modalities: MRA, T1-FLASH MR, T1-MPRAGE MR, and T2 MR. The composition of these image sets is described in Section 5.2.1. The individual MR images were acquired at UNC Chapel Hill using a Siemens head-only 3Tesla system (Allegra, Siemens Medical Systems Inc.) and a Siemens 1.5Tesla system (Sonata, Siemens Medical Systems Inc.) with a head coil. Imaging parameters, TR/TE/TH/in-plane resolution, for the T1 and T2 image acquisitions are as follows, 15 ms/7 ms/1 mm/1×1 mm^2 and 7730 ms/80 ms/1 mm/1×1 mm^2, respectively. Additionally, a 3D time-of-flight MRA sequence was acquired. Velocity compensation along both frequency and phase encoding direction was used to maximize signal de-phasing induced by the flowing spins. A magnetization transfer pulse was used to suppress signal from brain parenchyma while maintaining signal from flowing spins. The acquired spatial resolution for the MRA images was $0.5134 \times 0.5134 \times 0.78$ mm^3 and $1 \times 1 \times 1$ mm^3 for the T1 and T2 images. The MRA images were resampled to the $1 \times 1 \times 1$ mm^3 resolution.

5.1. Data preprocessing

The tissue exterior to the brain was removed using a mask generated by a brain segmentation tool based on the statistical classification method described in (Prastawa et al., 2004). The geometric prior used to initialize the algorithm was also produced using this tool. Mid-axial, mid-coronal, and mid-sagittal slice views for subjects one and two are presented in Figs. 7 and 8, respectively. These four modalities provide complementary information. For example, the T1-FLASH and T1-MPRAGE images have contrast differences and the MRA images exhibit missing information due to grey matter/white matter wash out and axial slab effect. In these examples, the set of structural classes is taken to be

$$C = \{c_1 = \text{grey matter}, c_2 = \text{white matter},$$
$$c_3 = \text{cerebrospinal fluid}, c_4 = \text{other}\}.$$

5.2. Registration experiments

To evaluate this image set registration framework, we first estimated transformations, f_1 and g_1, relating the coordinate spaces of subject one and subject two by
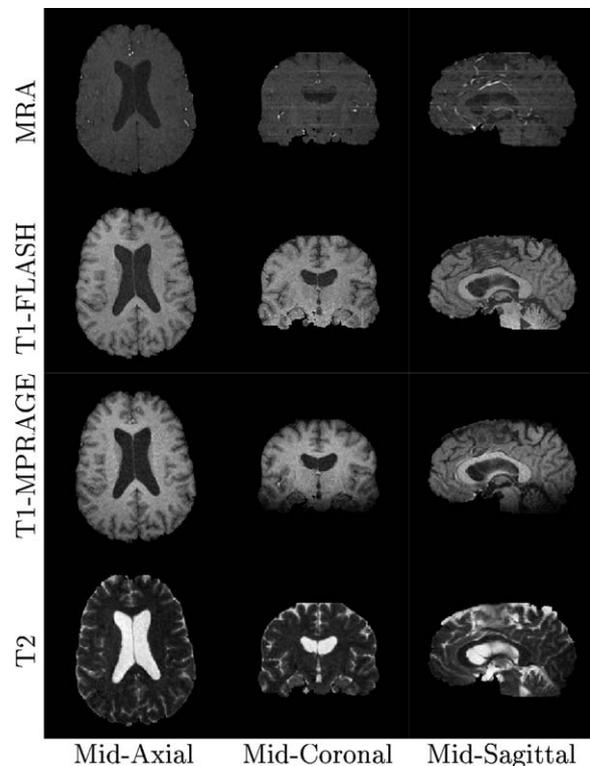
Fig. 7. Subject one.
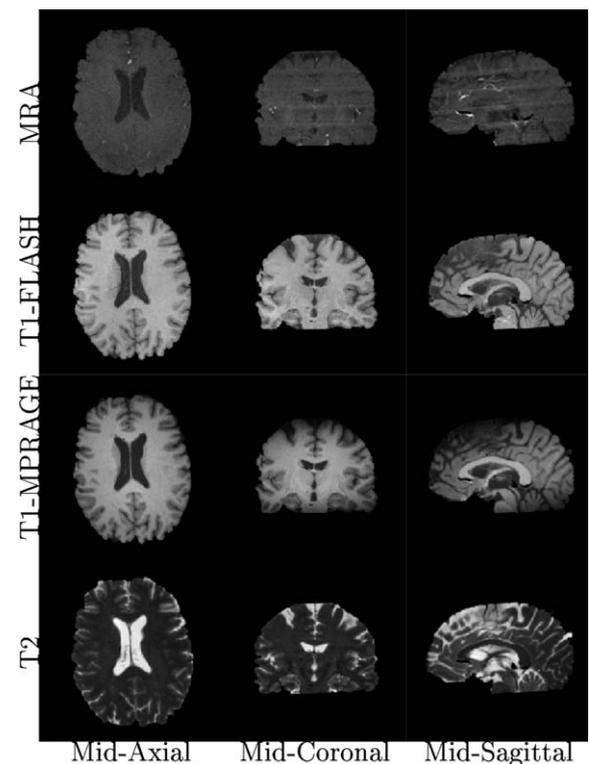
Fig. 8. Subject two.

applying our method to a mono-modal registration. These two transformations were then used as "ground truth" for the purpose of evaluating an increasingly

complex collection of image set registrations. Quantitative analysis of these results involves computing the inverse-consistency error.

5.2.1. Set-up

We performed the following eight registration experiments

(1) Mono-modal/Mono-modal (common): $\bar{I}_1 = $ T1−FLASH of subject one and $\bar{I}_2 = $ T1 − FLASH of subject two.
(2) Mono-modal/Mono-modal (mutually exclusive): $\bar{I}_1 = $ T1 − FLASH of subject one and $\bar{I}_2 = $ T2 of the subject two.
(3) Bi-modal/Bi-modal (fully common): $\bar{I}_1 = $ T1−FLASH and T2 of subject one and $\bar{I}_2 = $ T1−FLASH and T2 of subject two.
(4) Bi-modal/Bi-modal (single common): $\bar{I}_1 = $ T1−FLASH and T2 of subject one and $\bar{I}_2 = $ T1−MPRAGE and T2 of subject two.
(5) Bi-modal/Bi-modal (mutually exclusive): $\bar{I}_1 = $ T1 − FLASH and T2 of subject one and $\bar{I}_2 = $ T1 − MPRAGE and MRA of subject two.
(6) Bi-modal/Mono-modal (mutually exclusive): $\bar{I}_1 = $ T1 − FLASH and T2 of subject one and $\bar{I}_2 = $ MRA of subject two.
(7) Tri-modal/Tri-modal (fully common): $\bar{I}_1 = $ T1−FLASH, T1-MPRAGE, and T2 of subject one and $\bar{I}_2 = $ T1 − FLASH, T1-MPRAGE, and T2 of subject two.
(8) Quad-modal/Quad-modal (full common): $\bar{I}_1 = $ T1 − FLASH, T1-MPRAGE, T2, and MRA of subject one and $\bar{I}_2 = $ T1 − FLASH, T1-MPRAGE, T2, and MRA of subject two.

From each of these experiments, transformations f_i and g_i are obtained. The first experiment provides the "ground truth" transformations, f_1 and g_2. The T1-FLASH modality was chosen for the first experiment due to its relatively good white matter/grey matter contrast.

5.2.2. Bi-modal/bi-modal (mutually exclusive) registration

For the purposes of brevity we present qualitative results for the most interesting of these experiments, the bi-modal/bi-modal mutually exclusive registration. In this experiment, $\bar{I}_1$ represents the T1-FLASH and T2 images acquired from subject one and $\bar{I}_2$ represents the T1-MPRAGE and MRA images acquired from subject two. The estimated forward, f, and inverse, g, transformations are depicted, in three orthogonal views, in
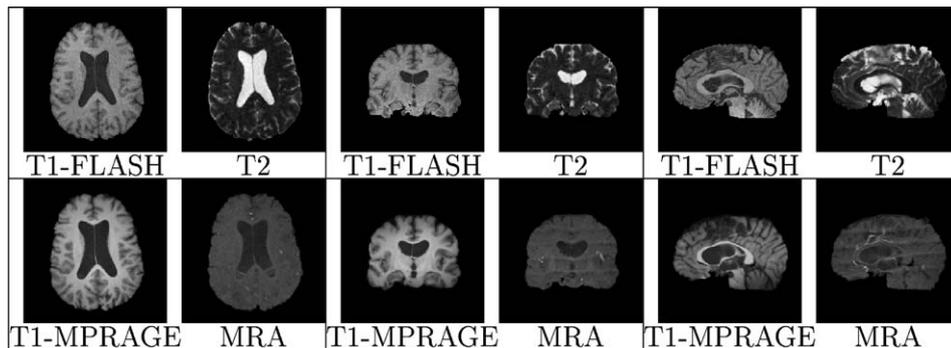
Fig. 9. Forward mapping. The top row shows mid-axial, mid-coronal, and mid-sagittal views of image set $\bar{I}_1$ and the bottom row shows the same views for the deformed image set $\bar{I}_2(f)$.
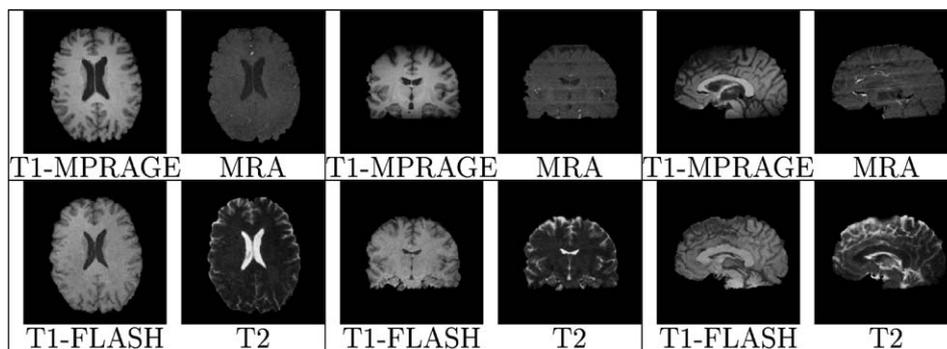
Fig. 10. Inverse mapping. The top row shows mid-axial, mid-coronal, and mid-sagittal views of image set $\bar{I}_2$ and the bottom row shows the same views for the deformed image set $\bar{I}_1(g)$.
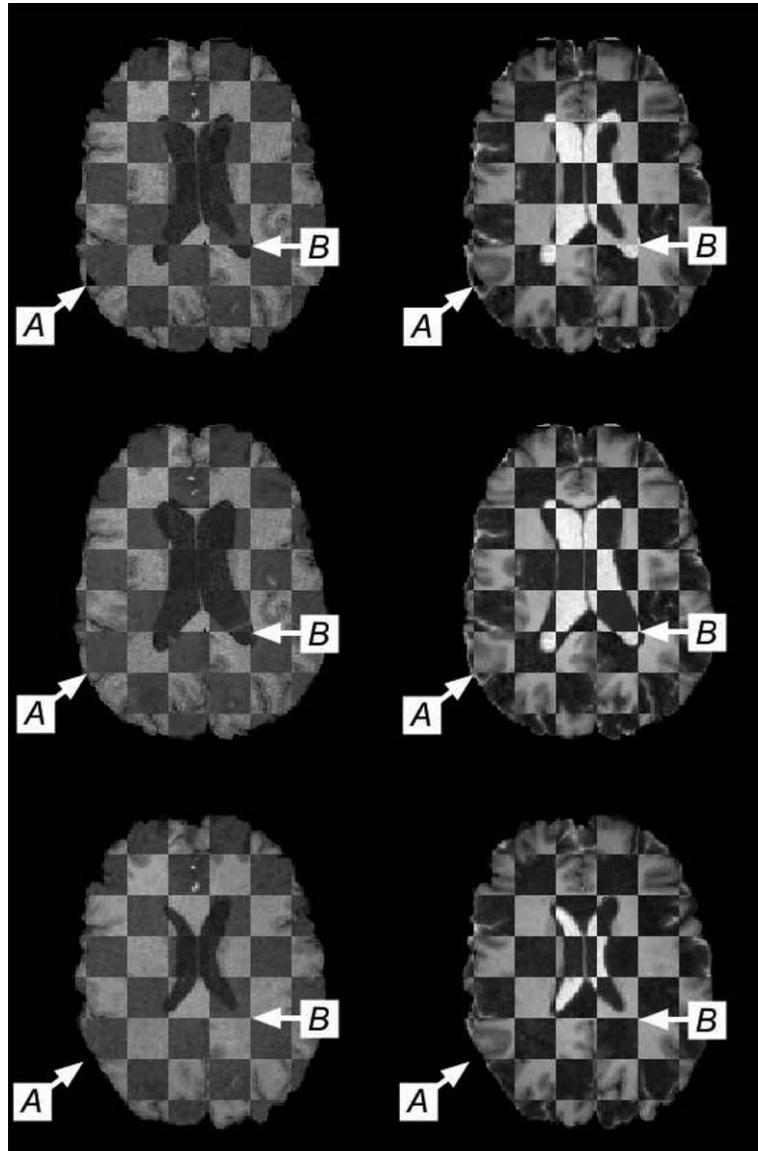
Fig. 11. Qualitative assessment. The top row shows a checker board blending of an axial slice of the image sets $\bar{I}_1$ and $\bar{I}_2$. The middle row shows a blending between $\bar{I}_1$ and $\bar{I}_2(f)$, the forward mapping and the bottom row a blending between $\bar{I}_2$ and $\bar{I}_1(g)$, the inverse mapping. The left column is composed of the T1-FLASH of $\bar{I}_1$ and the MRA of $\bar{I}_2$ and the right column is composed of the T2 of $\bar{I}_1$ and the T1-MPRAGE of $\bar{I}_2$. The registration has accommodated the local variability between the two image sets, especially in the cortical region, e.g. point "A", and in the ventricular region, e.g. point "B".

Figs. 9 and 10, respectively. In Fig. 11, a qualitative assessment of the registration can be made by examining axial slices of $\bar{I}_1$ and $\bar{I}_2$ in greater detail through a checkerboard pattern. The mismatch between the image sets is clearly evident in the top row of the figure. The second and third rows illustrate the effectiveness of the registration under the estimate forward and inverse transformations, respectively.

5.2.3. Inverse-consistency

We used the L_2 difference norm, $\|f_1(g_i(x)) - x\|_2$, to evaluate inverse-consistency between each experiment, i, and the first experiment for each spatial location $x \in \Omega$. For numerical stability these inverse-consistency

errors were computed via telescoping compositions as described in Section 2.5. Over all eight experiments, the maximum computed inverse consistency error was 3.12×10^{-4} voxels with an average of 5.04×10^{-5} voxels.

6. Discussion

The multi-modal image set registration as presented here might be potentially significant in various applications which rely on the measurement of image sets. For example, multi-modal imaging is standard in the imaging of pathologies such as tumors and lesions. Registration between images presenting pathology and

images of healthy subjects is a challenging task since space-occupying lesions have to be treated differently from infiltrating lesions. Specifically, the registration needs to accommodate both local spatial deformation and local change of image intensity. Existing registration method involving scalar images based on image brightness do not accommodate pathologies. In the formation of the class posteriors, we can explicitly assign classes to the various healthy and pathological tissues. This allows us to potentially model the behaviors of the different tissues during the registration process.

Another potential application for this method is the registration of images acquired from scanners of different field-strength. Image set registration across different scanners becomes an increasingly important component in multi-center studies. For example, in studies of developmental changes covering multiple years, and in follow-up studies of diseases with change of scanner technology. Images acquired from different scanners potentially have different contrasts and different spatial distortions. Our new method may help address these problems as the registration would be based on underlying anatomical structures rather than simply image intensities.

7. Conclusion

Image set registration is a generalization of pair-wise scalar image registration where each image set is composed of an arbitrary number of scalar images. In this paper, we have presented a novel method for multimodal image set registration and multi-class atlas formation. The method is model-based and finds correspondences between underlying class structures using posterior mass functions. We use Kullback–Leibler divergence on the space of posteriors as the minimization of this distance maximizes Bayes probability of error and, hence, indistinguishability between two posteriors. This method extends to the problem of multi-class atlas formation via computing an intrinsic, and hence, unbiased, mean. The introduction of this unbiased mean makes this registration approach intrinsically inverse-consistent. We have presented results, for image set configurations of varying complexity, which suggest that this method produces image set registrations with low inverse-consistency error.

Acknowledgements

References

Bhatia, K.K., Hajnal, J.V., Puri, B.K., Edwards, A.D., Ruekert, D., 2004. Consistent groupwise non-rigid registration for atlas construction. In: Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI), pp. 908–911.

Cachier, P., Rey, D., 2000. Symmetrization of the non-rigid registration problem using inversion-invariant energies: application to multiple sclerosis. In: Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI), Lecture Notes in Computer Science (LNCS). Springer-Verlag, Berlin, pp. 472–481.

Chan, H.-M., Chung, A.C.S., Yu, S.C.H., Norbash, A.,W.M.W. III, 2003. Multi-modal image registration by minimizing Kullback–Leibler distance between expected and observed joint class histograms. In: Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR), vol. 2, IEEE, pp. 570–576.

Christensen, G.E., Rabbitt, R.D., Miller, M.I., 1996. Deformable templates using large deformation kinematics. IEEE Transactions on Image Processing (TIP) 5 (10), 1435–1447.

Cover, R., Thomas, J., 1991. Elements of Information Theory. John Wiley and Sons, Inc., New York.

D'Agostino, E., Maes, F., Vandermeulen, D., Suetens, P., 2003a. An information theoretic approach for non-rigid image registration using voxel class probabilities. In: International Workshop on Biomedical Image Registration (WBIR), Lecture Notes in Computer Science (LNCS), vol. 2717. Springer-Verlag, Berlin, pp. 122–131.

D'Agostino, E., Maes, F., Vandermeulen, D., Suetens, P., 2003b. A viscous fluid model for multimodal non-rigid image registration using mutual information. Medical Image Analysis 7 (4), 565–575.

Davis, B., Lorenzen, P., Joshi, S., 2004. Large deformation minimum mean squared error template estimation for computation anatomy. In: Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI), pp. 173–176.

Gaens, T., Maes, F., Vandermeulen, D., Suetens, P., 1998. Non-rigid multimodal image registration using mutual information. In: Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI), Lecture Notes in Computer Science (LNCS), Springer-Verlag, pp. 1099–1106.

Grenander, U., Miller, M., 1998. Computational anatomy: an emerging discipline. Quarterly of Applied Mathematics 56, 617–694.

He, J., Christensen, G.E., 2003. Large deformation inverse consistent elastic image registration. In: Information Processing in Medical Imaging (IPMI), Lecture Notes in Computer Science (LNCS). Springer-Verlag, Berlin, pp. 438–449.

Hermosillo, G., 2002. Variational methods for multimodal image matching, Ph.D. thesis, Universite de Nice – Sophia Antipolis (May 2002).

Hoeffding, W., Wolfowitz, J., 1958. Distinguishability of sets of distributions. The Annals of Mathematical Statistics 29, 700–718.

Jeffreys, H., 1946. An invariant form for the prior probability in estimation problems. Proceedings of the Royal Society A 186, 453–461.

Joshi, S., Miller, M.I., 2000. Landmark matching via large deformation diffeomorphisms. IEEE Transactions on Image Processing (TIP) 9 (8), 1357–1370.

Joshi, S., Lorenzen, P., Gerig, G., Bullitt, E., 2003. Structural and radiometric asymmetry in brain images. Medical Image Analysis 7 (2), 155–170.

Joshi, S., Davis, B., Jomier, M., Gerig, G., 2004. Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage: Supplement issue on Mathematics in Brain Imaging 23 (Suppl. 1), S151–S160.

Kailath, T., 1967. The divergence and Bhattacharyya distance measures in signal selection. IEEE Transactions on Communication Theory (1), 52–60.

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. The Annals of Mathematical Statistics 22 (1), 79–86.

Lorenzen, P., Joshi, S., 2003. High-dimensional multi-modal image registration. In: International Workshop on Biomedical Image Registration (WBIR), Lecture Notes in Computer Science (LNCS), vol. 2717. Springer-Verlag, Berlin, pp. 234–243.

Lorenzen, P., Davis, B., Joshi, S., 2004a. Model based symmetric information theorectic large deformation multi-modal image registsration. In: Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI), pp. 720–723.

Lorenzen, P., Davis, B., Gerig, G., Bullitt, E., Joshi, S., 2004b. Multiclass posterior atlas formation via unbiased Kullback–Leibler template estimation. In: Proceedings of Medical Image Computing and Computer-assisted Intervention (MICCAI), Lecture Notes in Computer Science (LNCS). Springer-Verlag, Berlin, pp. 95–102.

Miller, M.I., Younes, L., 2001. Group actions, homeomorphisms, and matching: A general framework. Internation Journal of Computer Vision (IJCV) 41 (1–2), 61–84.

Miller, M., Banerjee, A., Christensen, G., Joshi, S., Khaneja, N., Grenander, U., Matejic, L., 1997. Statistical methods in computational anatomy. Statistical Methods in Medical Research 6, 267–299.

Miller, M.I., Joshi, S.C., Christensen, G., 1999. Large deformation fluid diffeomorphism for landmark and image matching. In: Brain Warping. Academic Press, San Diego, pp. 115–131 (Chapter 7).

Miller, M.I., Trouve, A., Younes, L., 2002. On the metrics and euler-lagrange equations of computational anatomy. Annual Review of Biomedical Engineering 4, 375–405.

Prastawa, M., Moon, N., Bullitt, E., van Leemput, K., Gerig, G., 2003. Automatic brain tumor segmentation by subject specific modification of atlas priors. Academic Radiology 10, 1341–1348.

Prastawa, M., Bullitt, E., Ho, S., Gerig, G., 2004. A brain tumor segmentation framework based on outlier dection. Medical Image Analysis (MedIA), 275–283.

Rueckert, D., Hayes, C., Studholme, C., Summers, P., Leach, M., Hawkes, D.J., 1998. Non-rigid registration of breast MR images using mutual information. In: Proceedings of Medical Image Computing and Computer-assisted Intervention (MICCAI), Lecture Notes in Computer Science (LNCS). Springer-Verlag, Berlin, pp. 1144–1152.

Studholme, C., Hill, D.L.G., Hawkes, D.J., 1999. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition 32 (1), 71–86.

Thompson, P., Woods, R., Mega, M., Toga, A., 2000. Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. Human Brain Mapping 9 (2), 81–92.

Toussaint, G.T., 1972. Comments on the divergence and Bhattacharyya distance measures in signal selection. IEEE Transactions on Communications 20 (3), 485.

van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P., 1999a. Automated model-based tissue classification of MR images of the brain. IEEE Transactions on Medical Imaging (TMI) 18 (10), 897–908.

van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P., 1999b. Automated model-based bias field correction of MR images of the brain. IEEE Transactions on Medical Imaging (TMI) 18 (10), 885–896.