Registering a MultiSensor Ensemble of Images

Jeff Orchard, *Member, IEEE*, and Richard Mann

***Abstract*—Many registration scenarios involve aligning more than just two images. These image sets—called *ensembles*—are conventionally registered by choosing one image as a template, and every other image is registered to it. This pairwise approach is problematic because results depend on which image is chosen as the template. The issue is particularly acute for multisensor ensembles because different sensors create images with different features. Also, pairwise methods use only a fraction of the available data at a time. In this paper, we propose a maximum-likelihood clustering method that registers all the images in a multisensor ensemble simultaneously. Experiments involving rigid-body and affine transformations show that the clustering method is more robust and accurate than competing pairwise registration methods. Moreover, the clustering results can be used to form a rudimentary segmentation of the image ensemble.**

***Index Terms*—Gaussian mixture models, multi-image, multisensor, mutual information, registration.**

I. Introduction

IMAGE registration is a fundamental operation in image analysis. It crops up in many realms, including medical imaging, remote sensing, quality control, and super-resolution, to mention only a few. A great deal of research has been devoted to automatically register two images (or volumes).

This paper addresses the question of how to register more than two images. Suppose you have several images—all of the same content—and you want to register them all together. We call this collection of images an *ensemble*. The vast majority of registration methods are designed to register only two images at a time. It is not clear how to use these pairwise methods for ensemble registration.

The problem of registration becomes more difficult when the images come from different sources. For example, a body part could be imaged with different modalities such as magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET), or a region of the earth captured by satellite imagery using a variety of different sensors, or several images of a face acquired with different illumination conditions. In these cases, the image intensities cannot be compared directly because, although the images depict the same content, they do so with different transfer functions. We refer to such registration problems as *multisensor* registration.

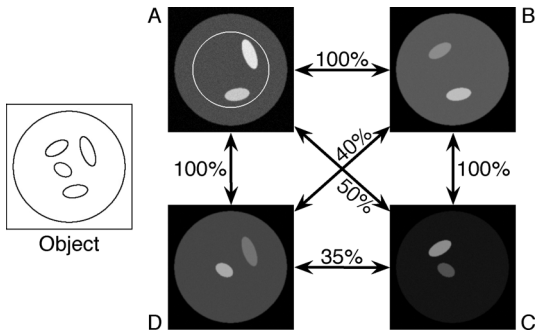


Fig. 1. Success rate for pairwise registration of the multisensor phantom dataset. The edge labels give the percentage of trials for which FLIRT's 64-bin NMI registration had an average pixel displacement of less than 3 pixels. The circular ROI is outlined in image A.

To illustrate the difficulty of multisensor ensemble registration, consider the "phantom" (to borrow a term from medical imaging) shown in Fig. 1. The figure shows a contrived example of multisensor imaging. The "true" object being imaged, shown on the left of Fig. 1, consists of a large circle encapsulating four ellipses. However, only two of the four ellipses are visible in each image, simulating the phenomenon of multisensor image acquisitions. Moreover, a different pair of ellipses is visible in each image, causing some pairs of images to contain disjoint content. The outlined circle in image A shows the region of interest (ROI) in which the images were registered. The success-rate percentages reported in the figure indicate that FLIRT [1], a popular pairwise registration method, had difficulty registering image pairs diagonally across from each other (the term "success" is fully defined in Section IV). It also had difficulty registering the image pair that shared the smallest and roundest ellipse (D and C). Hence, pairwise ensemble registration has the undesirable property that the solution depends on which pairs of images are chosen and registered. We will refer to this issue as *selection dependency*.

In addition, most pairwise registration methods do not offer a way to guarantee that redundancy in the solution is consistent. Consider a pairwise method that registers phantom image A to B and B to C. By composing those two transformations, one can derive a transformation from A to C. However, it is extremely unlikely that registering A to C will yield exactly the same transformation. We refer to this phenomenon as *internal inconsistency*.

We hypothesize that a registration strategy that registers all the images simultaneously can avoid both selection dependency and internal inconsistency. That is, including all the images in a single, global registration problem precludes the need to choose

Manuscript received October 02, 2008; revised November 23, 2009. First published December 28, 2009; current version published April 16, 2010. This work was supported in part by the Natural Science and Engineering Research Council of Canada and in part by the Canada Foundation for Innovation. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Gabriel Marcu.

The authors are with the David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1 Canada (e-mail: jorchard@uwaterloo.ca; mannr@cs.uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2009.2039371

which pairs to register, while generating a solution that is not redundant (and, thus, is internally consistent). Moreover, we hypothesize that the statistical power of using all the images at the same time, rather than just two at a time, will yield more accurate registration solutions.

In this paper, we present a method that employs clustering to simultaneously register an entire ensemble of images. The method computes the registration solution, and at the same time generates a model of the transfer functions among the images of the ensemble.

II. Background

Consider two images, one overlaid on the other. Each pixel corresponds to two intensity values, one from each of the two images. This 2-tuple can be plotted in the joint intensity space, where each axis corresponds to intensity from each of the images. Plotting the points for all the pixels creates a scatter plot in this joint intensity space, and we refer to this scatter plot as the *joint intensity scatter plot*, or JISP.

The idea behind many multisensor registration methods is to reduce the dispersion in the JISP. Why should this be the case? The implicit assumption linking different images of the same object is that they are recognizable as the same object because of some consistency by which intensities are assigned to components in the image. For example, bones show up as black in an MR image, and white in a CT image. Even though bones are rendered with a different intensity in each imaging modality, we still recognize the similarity in global shape because the intensity correspondence is consistent across many pixels. That is, pixels with intensities near x in one image often correspond to pixels with intensities near y in the other image. We call this correspondence an *intensity mapping*. An intensity mapping need not be one-to-one. Indeed, there are lots of examples where two pixels with the same intensity in one image correspond to different objects—and different intensities—in another image. Using our MR/CT example again, white matter and gray matter are virtually indistinguishable in CT, yet yield noticeably different intensities in T1-weighted MRI.

Each object in an image corresponds to a coherent collection of points in the JISP. For many examples, such as bone in MR and CT, the collection of points might be a distribution about a single focal point. For two images of a face—each one illuminated from a different angle—the collection of points corresponding to skin forms a manifold because the gradation of shades in one image corresponds to a different, though consistent, gradation of shades in the other image, creating a curve of points in the JISP.

As two images are moved out of register, the spatial correspondence of objects in the images gets disturbed, causing the coherence of the JISP to be disrupted. The clusters and swaths of scatter points spread out and move around because some bone pixels are now paired with muscle pixels, others with fat pixels, etc. Intensity-based multisensor image registration is based on this observation. The objective is to move the images until the JISP is optimally coherent, or minimally disperse.

One of the most successful applications of this idea uses the entropy of the joint histogram to quantify dispersion. Given the JISP between two images, one forms a joint histogram to reflect the density of points in the scatter plot. One can compute the entropy of this histogram. The lower the entropy, the more compact and tightly clustered the scatter plot and, hence, the more closely registered the two images. Extensions on the same idea include mutual information (MI) [2], [3], and normalized mutual information (NMI) [4].

The same idea can be applied to ensemble registration. The problem with the entropy-based methods is that they do not scale well for registration with more than two images. The joint histogram is an intermediary to those cost functions, and as you add more images to the problem, the number of histogram bins increases exponentially. For example, the joint histogram among five images, with each axis partitioned into 64 bins, has 2^{30} bins (over 1 billion). With 256 intensity bins per image, it gives us 2^{40} bins (over 1 trillion). Hence, these histogram-based methods are infeasible for ensemble registration.

Some registration methods measure the dispersion in the JISP without the need to form the joint histogram. In [5], dispersion is quantified as the length of a minimum-length spanning tree on the joint intensity scatter plot. Roche *et al.* [6] model the clusters in the JISP as a polynomial, thus assuming a functional relation between the intensities in the two images. This is often not the case. That method was expanded in [7] where a two-pass "least trimmed squares" approach allowed for the functional modeling of the outliers from the first least-squares pass. Leventon *et al.* [8] present an iterative method similar to ours, except they must specify the correct JISP in advance, eg., from previously registered images. In contrast, our method derives the JISP and registration solution simultaneously.

The first demonstration of ensemble registration that we are aware of was published in 1998 by Woods *et al.* [9]. Given a set of images, Woods constructs his cost function by adding together the sum of squared differences (SSD) between all possible pairs of images. Minimizing this cost function yields what Woods calls the "reconciled mean transformations". However, since his method uses the SSD cost function, it is only suitable for mono-sensor image registration.

Some other ensemble registration methods have recently emerged in the literature [10]–[14]. However, these methods have not been demonstrated on multisensor image ensembles, but rather focus on the problem of registering a set of images from the same modality to form a template (or so-called *atlas*). A different domain-specific method was designed to simultaneously register sets of brain MR images, but relies on the use of a human brain atlas to perform tissue classification, and then aligns the tissue-classification images [15]. Finally, one method [16] jointly registers and clusters a set of motion-corrupted images, automatically grouping images by similarity. However, their method assumes that the set of images is composed of moved and noisy versions of a set of prototype images, so the registration of the images to their class archetype amounts to mono-modal registration. Hence, these methods are not suitable for general-purpose multisensor ensemble registration.

In this paper, we present an efficient method for multisensor ensemble registration. Our method is based on clustering in the JISP, jointly modeling the distribution of points in the JISP as it estimates the motion parameters. Density estimation of the

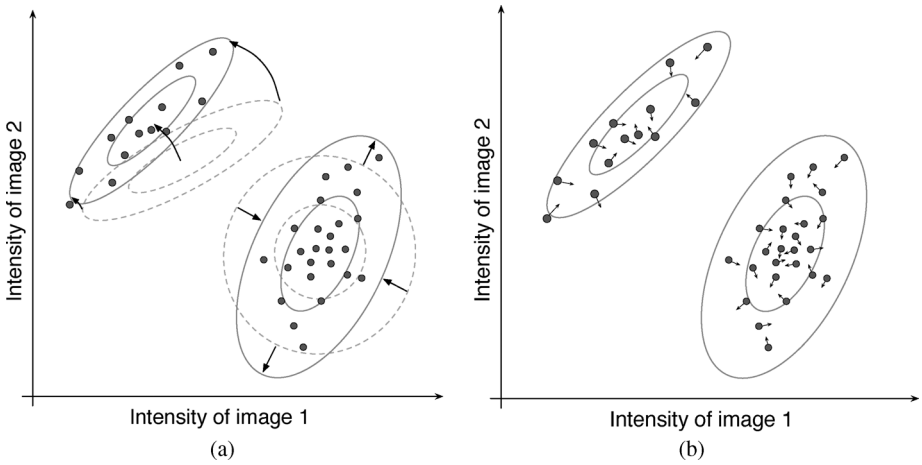


Fig. 2. Two main processes in ensemble registration. In density estimation, the motion parameters are held fixed while a better density estimate is computed by moving and stretching the cluster density components, as shown in (a). In motion adjustment, the density estimate is held fixed and the optimal motion is determined using least-squares. As the images move, the corresponding scatter points move toward the cluster centres (on average), as shown in (b).

clusters is modeled as a Gaussian mixture model (GMM), and is established iteratively using an estimation-maximization (EM) method. The motion parameters are also solved using an iterative Newton-type method. The iterates of these two methods are interleaved, thereby solving the two problems (density estimation and registration) in synchrony.

III. Method

Our approach to minimizing the dispersion in the JISP involves two steps: 1) density estimation of the scatter points in the JISP, followed by 2) moving the images to minimize the dispersion of the scatter plot. We refer to these two processes as *density estimation* and *motion adjustment*, respectively. Fig. 2 depicts density estimation and motion adjustment pictorially.

Suppose we are registering an ensemble of D images. Then, each pixel in our image domain has D values associated with it. We will refer to the vector of intensities for a single pixel as an "intensity vector", and denote the intensity vector for pixel x as $I_x \in \mathbb{R}^D$.

Let us represent our density estimate by ϕ. If we model the pixels as spatially independent variables, the likelihood of observing the images can be written

$$L(\phi) = \prod_x p(I_x|\phi) \tag{1}$$

where p is a probability function (defined later) and x denotes a pixel in our image domain (usually a subset of $\mathbb{R}^2$ or $\mathbb{R}^3$). Thus, $L(\phi)$ is the probability of observing the set of intensity vectors, given the distribution specified by ϕ.

To use this likelihood as a registration cost function, we add the possibility to move the ensemble of images. Let θ represent the set of motion parameters that specify the displacements applied to the image ensemble. Then, our likelihood cost function can be written not only as a function of the pixel density model, ϕ, but also as a function of the motion parameters, θ

$$L(\phi, \theta) = \prod_x p\left(I_x^\theta|\phi\right). \tag{2}$$

The expression I_x^θ represents the intensity vector for pixel x after applying the spatial transformation with parameters θ.

Because of the form of L, it is easier to optimize its logarithm, $\log L$, because the product over x turns into a sum

$$\log L(\phi, \theta) = \sum_x \log p\left(I_x^\theta|\phi\right). \tag{3}$$

Our goal in multisensor ensemble registration is to maximize $\log L(\phi, \theta)$ (and, hence, $L(\phi, \theta)$) by appropriate choice of ϕ and θ. The remainder of this section describes an iterative method that alternately optimizes L with respect to ϕ, and with respect to θ.

A. Gaussian Mixture Model

We will model the density of points in the JISP using a Gaussian Mixture Model (GMM) [17]. The mixture consists of K Gaussian components, each specified by a mean μ and covariance matrix Σ. Then, for a single pixel location x, the likelihood of observing the intensity vector I_x^θ is

$$p\left(I_x^\theta|\phi\right) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(I_x^\theta|\mu_k, \Sigma_k\right) \tag{4}$$

where the kth Gaussian component is specified by μ_k and Σ_k, and π_k are the component weights, with $\sum_k \pi_k = 1$. The function $\mathcal{N}$ denotes the normal (Gaussian) distribution

$$\mathcal{N}\left(I_x^\theta|\mu, \Sigma\right) = \frac{\exp\left(-\frac{1}{2}\left(I_x^\theta - \mu\right)^{\mathrm{T}} \Sigma^{-1}\left(I_x^\theta - \mu\right)\right)}{\sqrt{(2\pi)^D|\Sigma|}}. \tag{5}$$

B. Density Estimation

Taking θ to be the correct motion, we can improve our density estimate by optimizing $\log L(\phi, \theta)$ with respect to the probability density function ϕ.

For the GMM described above, we can find the optimal ϕ iteratively using the *expectation-maximization* (EM) algorithm [17]. The algorithm has an expectation step that maps scatter

points to clusters, followed by a maximization step that re-estimates the optimal clusters. The advantage of using this algorithm with a GMM is that each iteration has a closed-form, least-squares solution.

In the context of ensemble registration, the expectation step divvies-up the membership of each intensity vector among the K clusters. The membership of pixel I_x^θ to cluster k is

$$\tau_{kx} = \frac{\pi_k \mathcal{N}\left(I_x^\theta|\phi_k\right)}{\sum_k \pi_k \mathcal{N}\left(I_x^\theta|\phi_k\right)}. \tag{6}$$

Notice that $\sum_k \tau_{kx} = 1$ for each x.

The maximization step sets the parameters μ_k and Σ_k to their (weighted) maximum likelihood values, given by

$$\mu_k' = \frac{\sum_x \tau_{kx} I_x^\theta}{\sum_x \tau_{kx}}, \tag{7}$$

$$\Sigma_k' = \frac{\sum_x \tau_{kx} \left(I_x^\theta - \mu_k'\right)\left(I_x^\theta - \mu_k'\right)^{\mathrm{T}}}{\sum_x \tau_{kx}}. \tag{8}$$

Then the Gaussian component weights are re-estimated

$$\pi_k' = \frac{\sum_x \tau_{kx}}{\sum_k \sum_x \tau_{kx}}. \tag{9}$$

C. Motion Adjustment

The other half of the method involves holding ϕ fixed, and using it to find a motion increment that moves all the scatter points so that the overall log-likelihood, $\log L(\phi, \theta)$, is increased. We describe here a Newton-type step.

To optimize $\log L(\phi, \theta)$ with respect to the parameters θ, we set its gradient vector to zero

$$\frac{\partial}{\partial \theta} \log L(\phi, \theta) = \mathbf{0}. \tag{10}$$

Note that if each image has M motion parameters, and there are D images in the ensemble, then there are a total of MD motion parameters in θ. Hence, the gradient vector in (10) is an $MD \times 1$ vector.

For notational brevity, we will use $\mathcal{N}_k(I_x^\theta)$ instead of $\mathcal{N}(I_x^\theta|\mu_k, \Sigma_k)$. Recall that I_x^θ denotes the intensity vector at pixel location x after applying the motion parameters θ to the ensemble. The gradient vector of $\log L$ can be written

$$\frac{\partial}{\partial \theta} \log L(\phi, \theta) = \sum_x \frac{\frac{\partial}{\partial \theta} p\left(I_x^\theta|\phi\right)}{p\left(I_x^\theta|\phi\right)}. \tag{11}$$

Substituting in the definition for p [from (4)], and applying the chain rule for differentiation, we can express the numerator of the fraction as

$$\frac{\partial}{\partial \theta} p\left(I_x^\theta|\phi\right) = \sum_{k=1}^{K} \pi_k \frac{\partial I_x^\theta}{\partial \theta} \frac{\partial}{\partial I_x^\theta} \mathcal{N}_k\left(I_x^\theta\right) \tag{12}$$

where $\partial I_x^\theta / \partial \theta$ is an $MD \times D$ matrix holding the derivatives of the D pixel intensities with respect to each of the MD motion parameters, and $(\partial / \partial I_x^\theta)\mathcal{N}_k(I_x^\theta)$ is the $D \times 1$ gradient vector of the normal function. It is worth noting that since each motion parameter affects only one image in the ensemble, the matrix $\partial I_x^\theta / \partial \theta$ is sparse and block-diagonal.

Finally, the derivative $(\partial / \partial I_x^\theta)\mathcal{N}_k(I_x^\theta)$ can be written

$$-\Sigma_k^{-1}\left(I_x^\theta - \mu_k\right)\mathcal{N}_k\left(I_x^\theta\right). \tag{13}$$

Putting it all together, we can write $(\partial / \partial \theta) \log L(\phi, \theta)$ as

$$\sum_x \frac{-1}{p\left(I_x^\theta|\phi\right)} \sum_{k=1}^{K} \pi_k \mathcal{N}_k\left(I_x^\theta\right) \frac{\partial I_x^\theta}{\partial \theta} \Sigma_k^{-1}\left(I_x^\theta - \mu_k\right). \tag{14}$$

We want to find motion parameters θ that set $(\partial / \partial \theta) \log L(\phi, \theta)$ to a zero vector. To accomplish this, we wish to find a small motion increment $\tilde{\theta}$ so that $(\partial / \partial \theta) \log L(\phi, \theta + \tilde{\theta})$ is zero. We replace I_x^θ in the $(I_x^\theta - \mu_k)$ term of (14) with a nudged version of our ensemble, $I_x^{\theta+\tilde{\theta}}$. Approximating the spatial transformation so that it is a linear function of $\tilde{\theta}$ gives us

$$I_x^{\theta+\tilde{\theta}} = I_x^\theta + \frac{\partial I_x^\theta}{\partial \theta}^{\mathrm{T}} \tilde{\theta} \tag{15}$$

where $\tilde{\theta}$ is a small increment to the motion parameters (a "nudge"). Setting (14) to zero then yields a linear equation in $\tilde{\theta}$

$$\sum_x \frac{1}{p\left(I_x^\theta|\phi\right)} \sum_{k=1}^{K} \pi_k \mathcal{N}_k\left(I_x^\theta\right) \frac{\partial I_x^\theta}{\partial \theta} \Sigma_k^{-1} \times \left(I_x^\theta + \frac{\partial I_x^\theta}{\partial \theta}^{\mathrm{T}} \tilde{\theta} - \mu_k\right) = 0. \tag{16}$$

By simply factoring out $\tilde{\theta}$ and collecting the remaining terms, we get

$$\left(\sum_x \frac{1}{p\left(I_x^\theta|\phi\right)} \sum_{k=1}^{K} \pi_k \mathcal{N}_k\left(I_x^\theta\right) \frac{\partial I_x^\theta}{\partial \theta} \Sigma_k^{-1} \frac{\partial I_x^\theta}{\partial \theta}^{\mathrm{T}}\right) \tilde{\theta} = \left(\sum_x \frac{1}{p\left(I_x^\theta|\phi\right)} \sum_{k=1}^{K} \pi_k \mathcal{N}_k\left(I_x^\theta\right) \frac{\partial I_x^\theta}{\partial \theta} \Sigma_k^{-1}\left(I_x^\theta - \mu_k\right)\right) \tag{17}$$

or

$$A\tilde{\theta} = b. \tag{18}$$

where A is the $MD \times MD$ system matrix, and b the $MD \times 1$ vector defined inside the parentheses on the right-hand side of (17). As an example, consider an ensemble with five images (i.e., $D = 5$), each with three motion parameters (i.e., $M = 3$). Then (18) is a linear system of 15 equations involving 15 unknowns. Solving this system for $\tilde{\theta}$ gives us the optimal motion increment, according to the linear approximation. We use this increment to adjust our current estimate for θ.

D. Implementation

We implemented our method in Matlab (Mathworks Inc., Natick, Massachusetts). Algorithm 1 summarizes the code for our method. Here we outline a few details about our implementation.

Algorithm 1 Ensemble clustering registration

input: initial ensemble I_0

input: initial motion parameters θ

input: initial GMM parameters ϕ

for each scale **do**

$I_{\text{scaled}} \leftarrow$ scale ensemble I_0

$I \leftarrow$ apply motion (θ) to ensemble I_{scaled}

repeat

for K_1 iterations **do**

$\phi \leftarrow$ EM step (see Section III-B)

end for

for K_2 iterations **do**

$\tilde{\theta} \leftarrow$ motion adjustment (see Section III-C)

$\theta \leftarrow \theta + \tilde{\theta}$

$I \leftarrow$ apply motion (θ) to ensemble I_{scaled}

end for

until converged ($\tilde{\theta}$ is small)

end for

output: I is registered ensemble at full scale

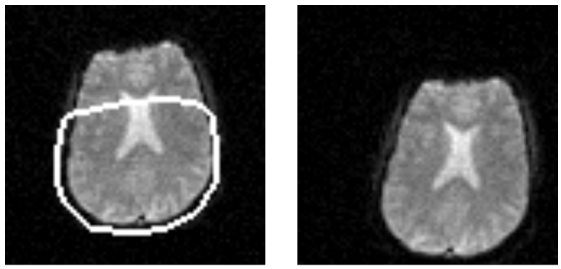
output: θ holds the optimal motion parameters

output: ϕ holds the GMM parameters

As is common practice for image registration (including FLIRT), we use a multiresolution framework in which images are registered first at a low resolution, and then at successively higher resolutions [18], [19]. The solution at each scale is used as an initial guess for the next scale. In general, images were registered at scales 10%, 20%, 50%, and then finally 100%. Exceptions to that schedule are noted in Section IV.

Though not required in our model, each image in the ensemble was subject to the same type of spatial transformation (either rigid-body, or affine). Each image then had the same number of motion parameters, M, associated with its own transformation. Thus, the total number of motion parameters stored in the vector θ is MD.

Our method requires an initial density estimate. Following the method outlined in Section III-B, we implemented a simple EM iteration that attempts to find a GMM (represented by ϕ) to maximize (3). The initialization has three phases, with increasing degrees of freedom in the maximization step. In the first phase, only the means (μ_k) are adjusted using (7). In the second phase, the weights (π_k) are also adjusted using (9). Finally, in the last phase, the covariances (Σ_k) are also adjusted using (8). Ten EM iterations are executed for each of these three phases.



(a) (b)

Fig. 3. Images used in mono-sensor time series experiments (FMRI). (a) Image during the inactive period, as well as the outline of the ROI used for one of the experiments. (b) Image from one of the active periods, displaced randomly. The activity in the occipital region (toward the bottom of the image) is very slightly brighter in (b) than (a).

IV. Experiments

Our new ensemble clustering registration method was compared to two pairwise registration methods, as well as a least-squares ensemble method, to test our hypotheses regarding the advantages of ensemble registration versus pairwise registration. We used our method in its full ensemble registration mode, where all the available images were simultaneously registered. We also used registration methods implemented in FLIRT, a product of the Oxford FMRIB group [1]. Their program is a popular registration tool, and is able to register using a number of different cost functions—we used three of them: normalized mutual information (NMI), correlation ratio (CR), and normalized correlation (NC). We also used our clustering method in a pairwise fashion, where we registered only two images at a time. This pairwise clustering registration method was included to act as a rough experimental control to differentiate between the effects of ensemble registration, and the effects of registration using clustering. Finally, we modified our clustering registration method to implement the sum of squared differences (SSD) cost function, where the cost of a scatter point is measured as the square of its perpendicular distance from the identity line in the JISP. The purpose for including this method was to demonstrate the importance of density estimation in the process of multi-sensor ensemble registration.

The performance of each registration method was gauged by comparing the estimated transformations to the gold standard transformations. This difference was quantified using the average pixel displacement, defined as the distance of each pixel from its true, registered position, averaged over all pixels used in the registration. Hence, an average pixel displacement of 0 indicates perfect registration, and a large average pixel displacement means poor registration. If the average pixel displacement is greater than 3 pixels, then the registration is considered a failure. Hereafter, we will use the word *error* to denote the average pixel displacement.

The registration methods were tested in five different registration applications. For each application, we generated trial ensembles by applying known displacements to the initially-regsitered images. For most of the image ensembles, multiple trials were run. The different experiments are outlined below.

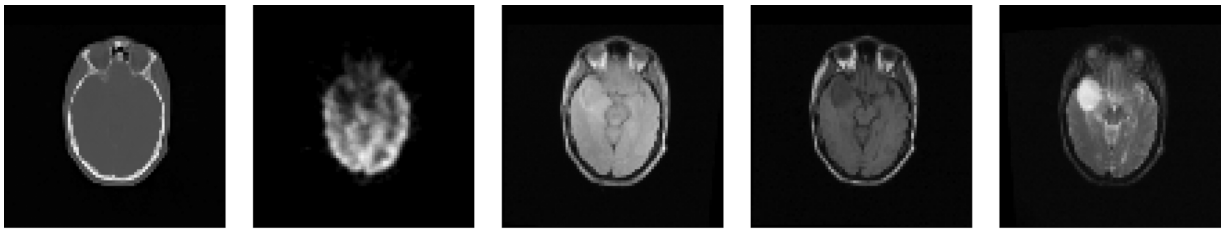


Fig. 4. Slice from the RIRE ensemble used to test 3-D registration.

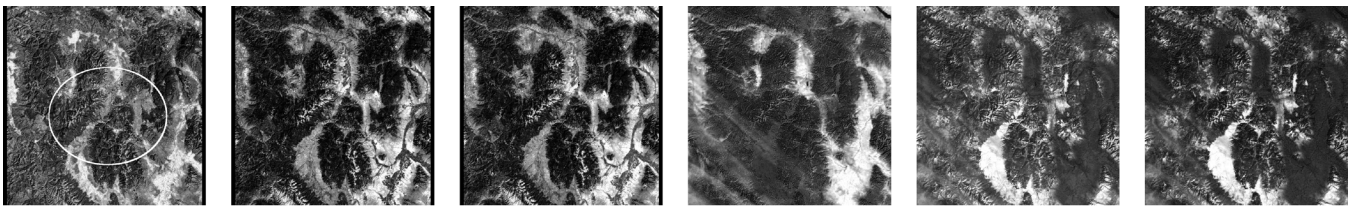


Fig. 5. Satellite images used to test the affine registration. The ROI is outlined in image A.

A. Mono-Sensor Time Series

This image set contains a time series of 56 snapshots of a simulated functional MRI experiment (see Fig. 3). The snapshots were created by duplicating a single 80×80 pixel $T2^*$-weighted MRI 56 times. Artificial activation was added to a large part of the occipital lobe, adding 5% of the average brain intensity to the active region during three 8-frame active periods. Then, randomly-generated rigid-body displacements were applied to all the images, choosing the translations and rotations uniformly from the range $[-10, 10]$ pixels or degrees. Thus, there were 168 motion parameters in total. Finally, Gaussian distributed white noise was added to each image ($\sigma = 4.5\%$ of mean brain instensity).

Because of the time it takes to complete the 3080 pairwise registration runs, we produced just one trial ensemble. The ensemble clustering and the pairwise clustering registration methods were each run with only one Gaussian component, and each used a multiresolution framework with scales 20%, 50%, and 100%. The FLIRT method was run using the normalized correlation cost function. We also ran the ensemble SSD method on this dataset. The entire set of registration scenarios was run twice by each registration method, once using all the pixels in the image, and once using only the pixels in a region of interest that included roughly 75% of the brain [outlined in Fig. 3(a)].

B. Three-Dimensional

We also tested our registration method on an ensemble of medical imaging volumes. The Retrospective Image Registration Evaluation (RIRE) project's training set was used [20] (one slice is shown in Fig. 4). The set consists of five volumes: three MR (magnetic resonance) volumes (T1-weighted, T2-weighted, and PD-weighted), a computed tomography (CT) volume, and a positron emission tomography (PET) volume. The volumes were resampled into register using the true displacement parameters supplied by the RIRE project, then scaled so that each volume was $80 \times 80 \times 26$ isotropic voxels. The volumes were then padded to $80 \times 80 \times 32$ to avoid image content from leaving the field of view after being displaced. Ten trial ensembles were generated using randomly generated 3-D rigid-body transformations, uniformly choosing the three rotations and three translations from the range $[-5, 5]$ (degrees or pixels). The ensemble clustering method was initialized with six Gaussian components, while the pairwise clustering method was initialized with four. Both clustering methods used a multiresolution framework with scales 20%, 50%, and 100%. The FLIRT pairwise registration method was run using the correlation ratio cost function.

C. Satellite

A set of six Landsat satellite images, shown in Fig. 5, were used for testing. Each image was acquired with a different sensor. These six images (761×748 pixels in size) were used to generate ten trial ensembles, each with a different set of randomly-generated 6-degree-of-freedom affine displacements (two scales, one shear, one rotation, and two translations). The motion parameters were chosen uniformly from the following ranges: $[0.95, 1.05]$ for scales, $[-0.2, 0.2]$ for the shear, $[-5, 5]$ degrees for the rotation, and $[-5, 5]$ pixels for the translations. Registration was performed on the region of interest outlined in A of Fig. 5. The ensemble clustering registration method was initialized with six Gaussian components, while the pairwise clustering method was initialized with three components. The FLIRT method was run using the correlation ratio as the cost function.

D. Variable Illumination

Fig. 6 shows an image ensemble of the same face (640×480) with five very different light positions, ranging from far left to far right. The images were taken from the Extended Yale Face Database B [21]. Registration was performed using only the pixels in the region of interest, outlined on F1 of Fig. 6. This set of images poses a very difficult registration scenario, since images F1 and F5 have very little illuminated content in

Fig. 6. Face images used to test the effect of variable illumination. The ROI is outlined in image F1.

common within the region of interest. Ten trial ensembles were generated by applying randomly-generated rigid-body displacements, chosen uniformly from the range [−10, 10] pixels or degrees.

The ensemble clustering registration method was initialized with six Gaussian components, while the pairwise clustering method was initialized with four components. The FLIRT method was run using the NMI cost function with 64 bins. The ensemble SSD method was also run on this dataset.

E. Disjoint Content

Ten trial ensembles of the multisensor phantom (shown in Fig. 1) were generated using randomly-generated rigid-body displacements chosen uniformly from the range [−10, 10] pixels or degrees. The ensemble clustering method was initialized with five clusters, while the pairwise clustering method was initialized with four. Both clustering methods used a multiresolution framework with scales 20%, 50% and 100%. The FLIRT method was run using the NMI cost function with 64 bins. We ran the ensemble SSD method on this dataset as well.

F. Number of Gaussian Components

One of the parameters that needs to be specified in our method is the number of Gaussian components to include in the mixture model. We wanted to know how our the ensemble registration method would behave over a range of different values. We ran only our ensemble registration method (not the pairwise methods) on the multisensor phantom dataset for three different numbers of components: $K = 3$, $K = 5$, and $K = 8$. The trial datasets are the same as those used in the "disjoint content" experiment. Over the ten trial datasets, we recorded the mean error for all trials, the mean error for converged trials (those with an error of less than three pixels), and the number of failed registrations.

V. Results

A. Mono-Sensor Time Series

Among the 56 images in the simulated functional MRI time series, the pairwise methods performed 3080 pairwise registrations. The initial average error for the unregistered images was 10.6 pixels for the trials that involved the ROI. On those trials, the ensemble clustering and ensemble SSD methods successfully registered all the images. The pairwise clustering method failed on 686 (22%) of the registrations, while FLIRT's normalized correlation registration method failed on only one. Of the successful cases, the mean error was 0.028 for the ensemble clustering method, 0.035 for the ensemble SSD method, 0.046 for the pairwise clustering method, and 0.12 for FLIRT. These results suggest a number of things. 1) When compared to the pairwise clustering method, FLIRT is quite a robust pairwise registration method. 2) The ensemble methods (clustering and SSD) avoided all the failures that the pairwise clustering methods encountered. 3) The error was substantially lower for the ensemble methods than for FLIRT.

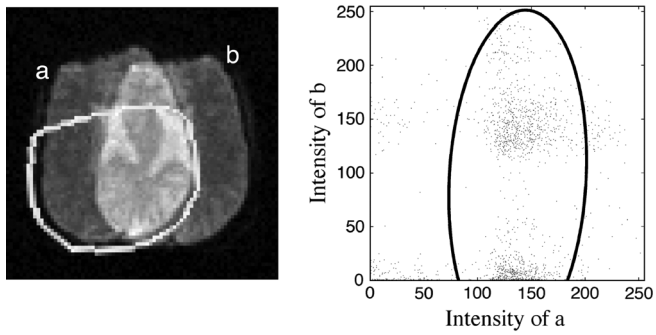


Fig. 7. Initial misregstration for a case on which the pairwise clustering method failed. The large portion of the background of image "b" inside the ROI creates a cluster near the "a"-axis in the JISP that distracts the Gaussian component and causes the registration to diverge.

The cases that failed in the pairwise clustering method were those where the initial misregistration was sufficiently large to cause the solution to diverge. Fig. 7(a) shows one such initial displacement. The image labeled "b" in the figure was displaced to the right far enough that its background filled a significant part of the ROI (which is fixed in relation to image "a"). This region forms a cluster in the JISP along the "a"-axis, which the Gaussian component automatically models [as shown in Fig. 7(b)]. Then, moving image "b" further to the right reinforces that erroneous cluster, and causes the solution to diverge until only the background of "b" is inside the ROI. We do not observe this phenomenon during the ensemble registration because the erroneous clusters are scattered incoherently; they essentially cancel each other out.

For the trials that used the entire image for registration, the initial average error was 11.2 pixels, and all four methods correctly registered all the images. There were differences, however, in the accuracy. The mean error for the ensemble clustering and ensemble SSD registration methods were 0.022 and 0.023 pixels (respectively). The pairwise clustering method was only slightly less accurate, with a mean error of 0.024 pixels. The normalized correlation registration of FLIRT yielded a mean error

TABLE I
AVERAGE ERROR FOR 3-D ENSEMBLE (RIRE). INITIAL AVERAGE ERROR WAS 6.2 PIXELS

Method	CT-PET	CT-PD	CT-T1	CT-T2	PET-PD	PET-T1	PET-T2	PD-T1	PD-T2	T1-T2	Mean
NMI (pairwise)	6.1	7.7	2.6	8.2	7.6	5.9	6.7	6.0	4.9	7.7	6.34
Cluster (pairwise)	3.4	3.8	2.5	2.9	7.4	6.4	5.4	2.8	4.5	3.2	4.23
Cluster (ensemble)	2.3	1.7	1.7	1.8	2.8	2.5	2.9	0.93	0.56	1.1	1.83

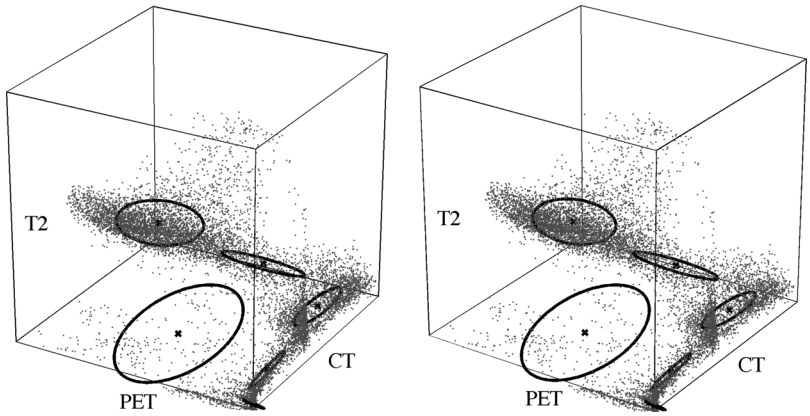


Fig. 8. Three-dimensional joint intensity scatter plot (stereo pair) and corresponding Gaussian components for the 3-D registration test. Note that only three of the five dimensions of the JISP are plotted. To view the stereo pair, hold the page approximately 30 cm from your face and have your left eye focus on the left plot, and your right eye focus on the right plot. You might find it helpful to place a piece of paper vertically between the plots so that your left eye cannot see the right image, and your right eye cannot see the left image.

of 0.087, almost four times larger than the error exhibited by the ensemble clustering registration method.

B. Three-Dimensional

The results for the 3-D registration test, using the RIRE ensemble, are shown in Table I. The initial average error for the unregistered images was 6.2 pixels. Both pairwise methods were less accurate than the ensemble method. The ensemble clustering method successfully registered the entire set of images, with the exception of what appears to be a slight misregistration of the PET image.

Fig. 8 plots three of the five dimensions of the JISP as a stereo pair (see the figure caption for instructions on how to view the stereo pair). That is, only the CT, PET and MRI-T2 intensities are represented in the scatter plot; the other two dimensions of the space—corresponding to MRI-PD and MRI-T1—are not shown in the figure even though they were used in the registration process. The figure also includes a representation of the Gaussian components derived by the clustering registration process, each drawn as an ellipse that indicates its location and covariance. Notice how the six ellipses fit the data, collectively acting as a density estimator. Bear in mind that the Gaussian components are scalar functions in a 5-D space, so our 2-D ellipses drawn in 3-D express only part of the full picture. In particular, each ellipse is drawn to indicate the covariance in the plane of largest variation. The 2-D ellipse is drawn in 5-D and then projected onto the 3-D subspace depicted in the figure.

Each point in the scatter plot maintains a fractional membership to each of the Gaussian components. Using (6), we can compute τ_{kx}, the membership of pixel x to Gaussian component k. Hence, once the registration method is complete and the GMM is stable, we can view the spatial distribution of the influence of a chosen Gaussian component by creating an image of τ_{kx} for a fixed value of k. We call this image the *membership map* for component k. Brighter regions show pixels that are strongly linked to component k. The six membership maps corresponding to Fig. 8 are shown in Fig. 9. In the figure, one can see that the different Gaussian components model coherent structures in the image.

C. Satellite

Out of the 300 pairwise registrations (ten trials, each with 30 registration pairs), the initial average error for the unregistered images was 15.9 pixels. The ensemble clustering registration method failed on 20 of them. The pairwise clustering method failed on 150 pairs (50%), and FLIRT's correlation ratio method failed on 37 of the pairs (12%). It is worth noting that the 20 misregistration cases for the ensemble method were the result of only two registration failures. In each of two trials, one of the six images failed to converge to the other five (and vice versa), and was thus recorded as ten misregistered image pairs. Of the successful cases, the ensemble clustering method had a mean error of 0.31, while the pairwise clustering method and FLIRT's CR method reported 0.65 and 0.41, respectively.

D. Variable Illumination

The results for the variable illumination experiment (involving the face images) are shown in Table II. The initial average error for the unregistered images was 20.3 pixels. As one would expect, the pairwise methods had a great deal of difficulty registering images with vastly different illumination conditions. For example, image F1 is illuminated from the left, while F5 is illuminated from the right. When registering F1

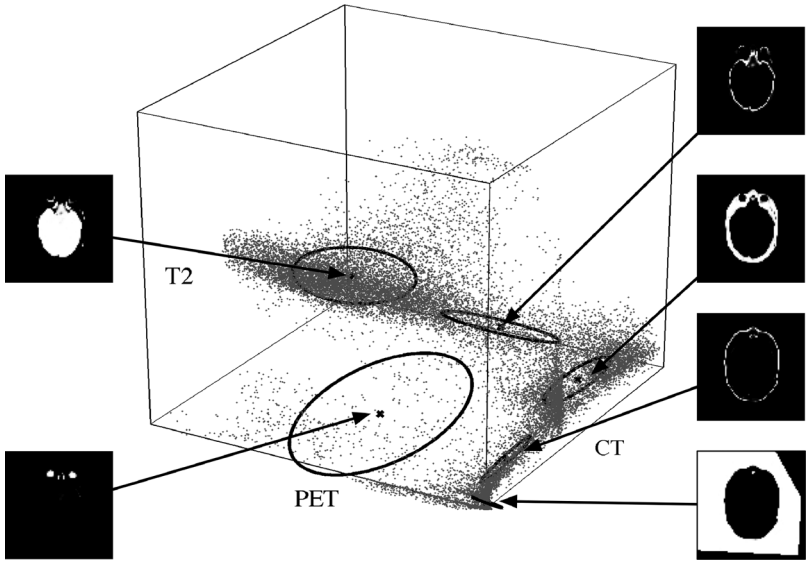


Fig. 9. Membership maps for the 3-D registration test. The scatter plot is the same as shown in Fig. 8. For each Gaussian component, we also show its membership map, an image representation of τ_{kx} for a fixed Gaussian component k. The membership maps show that each Gaussian component models a coherent set of pixels. Starting from the top-right and moving clockwise, the membership maps roughly delineate bone, muscle/fat, skin, background (showing only the part inside the ROI), eyeballs, and brain.

TABLE II
AVERAGE ERROR FOR VARIABLE ILLUMINATION ENSEMBLE (FACE). INITIAL AVERAGE ERROR WAS 20.3 PIXELS

Method	F1-F2	F1-F3	F1-F4	F1-F5	F2-F3	F2-F4	F2-F5	F3-F4	F3-F5	F4-F5	Mean
NMI (pairwise)	33	96	10	41	64	89	47	49	99	15	54.6
Cluster (pairwise)	28	67	13	36	24	27	15	23	44	10	28.6
SSD (ensemble)	122	72	95	29	69	49	141	26	94	117	81.4
Cluster (ensemble)	0.55	1.2	2.8	2.8	0.73	2.5	2.6	2.0	2.1	0.32	1.77

and F5, the mean error was 41 pixels for FLIRT (using NMI with 64 intensity bins), and 36 pixels for the pairwise clustering method. However, the ensemble clustering registration method had a mean average pixel displacement of only 2.8 pixels for those two images. As one would expect, the ensemble SSD method did not perform well on this dataset.

Interestingly, the pairwise methods seemed to do better registering the extreme images (i.e., F1 with F5) than registering images that were two conditions apart (i.e., F1 with F3, F2 with F4, or F3 with F5). This effect might stem from the fact that images with overlapping regions of illumination have more latitude to wander away from the correct solution, while extreme images are more constrained by the sharp shadow border down the centre of the face.

E. Disjoint Content

The initial average error for the unregistered phantom trials was 10.0 pixels. As Table III shows, a number of image pairs in the multisensor phantom proved to be very difficult for the pairwise methods to register accurately. The most notable case is the pairwise registration of A-to-C, and B-to-D. In those image pairs, none of the smaller ellipses are common between the two images. That is, neither of the two ellipses in A are visible in C (and *vice versa*). It should come as no surprise, then, that the pairwise methods exhibited large errors for those image pairs. However, the ensemble clustering registration method successfully registered all the trials. Again, the ensemble SSD method was not successful on this dataset.

TABLE III
AVERAGE ERROR FOR DISJOINT CONTENT ENSEMBLE (MULTIMODAL PHANTOM). INITIAL AVERAGE ERROR WAS 10.0 PIXELS

Method	A-B	A-C	A-D	B-C	B-D	C-D	Mean
NMI (pairwise)	0.59	4.1	0.31	0.83	7.2	4.2	2.87
Cluster (pairwise)	7.0	9.0	6.5	9.8	12	9.7	9.10
SSD (ensemble)	34	76	60	62	50	41	53.9
Cluster (ensemble)	0.89	1.5	0.16	1.0	0.90	1.5	1.00

TABLE IV
EFFECT OF VARYING THE NUMBER OF GAUSSIAN COMPONENTS, K, IN THE ENSEMBLE CLUSTERING METHOD

K	Mean Error (std. dev.)	Mean Error of converged (std. dev.)	Success Rate
3	8.57 ($\pm$11.2)	1.16 ($\pm$0.737)	63.3%
5	1.00 ($\pm$0.681)	1.00 ($\pm$0.681)	100%
8	1.32 ($\pm$0.943)	1.18 ($\pm$0.804)	93.3%

F. Number of Gaussian Components

Table IV shows the results of running the ensemble clustering method using different numbers of Gaussian components. The mean error (over ten trials) suggests that K, the number of Gaussian components, plays a significant role in the success—or failure—of the ensemble clustering registration method. With too few components (3 instead of 5), the registration success rate declined from 100% to 63.3%. Having too many components also reduced the success rate, though only slightly (93.3%). However, the number of components did not seem to have a significant impact on the accuracy of the trials that did converge.

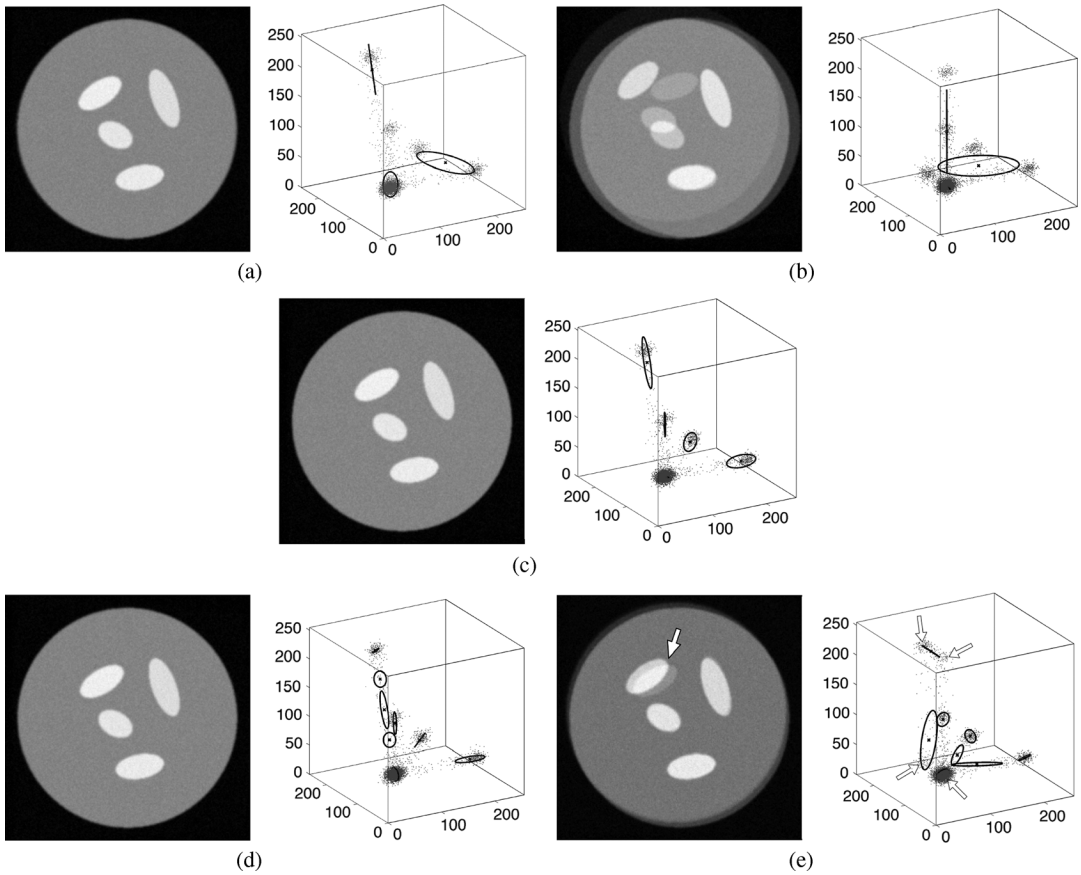


Fig. 10. Sample solutions for varying K, the number of Gaussian components. Overlaid images are shown beside the corresponding JISP with Gaussian components indicated by ellipses. The cases shown in (b) and (e) did not converge to the correct solution, while the others did. (a) $K = 3$, Error $= 2.64$; (b) $K = 3$, Error $= 18.1$; (c) $K = 5$, Error $= 0.678$; (d) $K = 8$, Error $= 1.14$; (e) $K = 8$, Error $= 7.45$.

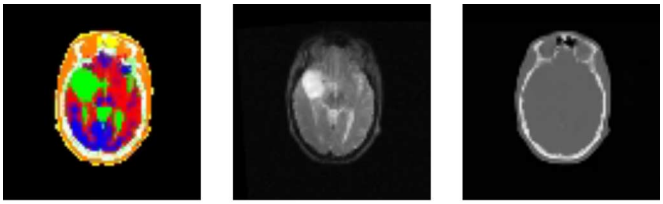
Fig. 10 shows some more results for this experiment. The phantom dataset is particularly informative because we know that there are five different clusters in the 4-D joint intensity space, since there are five regions in total: four small ellipses, and the large embedding circle. Note that the black background is outside the ROI, so does not form a cluster.

The cases shown in Fig. 10(a) and (b) were run with only three Gaussian components, and the scatter plots show that single components were forced to stretch to model multiple clusters. The case shown in (b) did not converge to the correct solution; one component was stretched enough to enable stable modeling of many clusters, some of which correspond to misregistered image content. The case shown in (c) was run with five Gaussian components, each one accurately modeling one of the JISP clusters. Cases (d) and (e) were each run with eight components. While the case in (d) converged to the correct solution, the additional components facilitated the modeling of partial-volume artifacts (caused by interpolation at the high-contrast edges). This accommodation allowed some pixels to deviate from their proper cluster, and instead rest in the corridor modeled by the additional Gaussian component, thereby increasing the error slightly. Case (e) did not converge to the correct solution. One of the Gaussian components (the largest in the JISP) was available to accommodate the misregistered ellipse (indicated by the arrow in the image). As a result, what should have been two clusters in the JISP (one for the ellipse, and one for the embedding circle) was split into four co-planar clusters, as shown by the four arrows in the JISP.

It should be noted that the cluster-modeling hiccups observed in Fig. 10 can be observed with any number of Gaussian components. For example, some $K = 5$ cases showed stretching of a Gaussian component to model two clusters, thereby freeing up one component to model a partial-volume branch [similar to that shown in (d)]. However, these variations do not necessarily devastate the final result.

VI. Discussion

Density estimation of the JISP realizes a combination of the advantages from mutual information (MI) and least-squares (LS). Our method can be viewed as a parametric regression method, with the number of parameters dictated by the number of Gaussian components. At one end of the spectrum, a single Gaussian component with infinite covariance along one direction, and unit covariance along all other directions, is equivalent to least squares. From there, more and more modeling components can be added, each increasing the flexibility of the model. There is a strong connection between our method and entropy-based measures such as MI. Consider, for example, covering the joint intensity space with a grid of Gaussian components. The weights of the components are comparable to the frequencies stored in histogram bins. We plan to fully explore this link in a future paper.



(a) (b) (c)

Fig. 11. Rudimentary segmentation of the RIRE dataset. In (a), the pixels are colour-coded according to their memberships to the ten Gaussian components that were derived during registration. For reference, the corresponding T2 and CT slices are shown in (b) and (c). (a) Segmentation; (b) T2-MRI; (c) CT.

From our results, it is clear that the strength of our method stems from the concept of ensemble registration, and not from the clustering method itself. Indeed, the pairwise clustering method was less successful than FLIRT in many of the experiments, suggesting that there might be room for improvement. In particular, we could either improve the optimization search process for clustering registration, or improve our density estimation, perhaps incorporating a GMM with a "catch all" component to explain outliers.

The clustering registration method scales linearly with the number of Gaussian components (k), and the number of pixels (N). However, the computation time is proportional to the cube of the number of motion parameters (M), and the cube of the number of images (D) because of the matrix products in (17). More precisely, the method has computational complexity $\mathcal{O}(kNM^3D^3)$.

In addition to registering an ensemble, our method inherently offers a rudimentary segmentation. As shown in Fig. 9, membership maps tend to reflect regions in the image that contain consistent intensity combinations across the ensemble of images. In medical imaging, tissue types tend to be manifested as different, but consistent, intensities among the different imaging modalities (e.g., CT and MRI). The notion can even be extended to tissue *function* if a functional modality is included, such as PET. Fig. 11 shows a segmentation that resulted from registering the RIRE ensemble using ten Gaussian components. Each component was given its own colour, and Fig. 11 shows the superposition of all ten tissue classes. This concept was first pointed out by Leventon and Grimson [8], where they generated a segmentation based on a GMM of a 2-D joint histogram. The difference here is that we generate our GMM density estimation *as part of* our registration procedure, while Leventon and Grimson generated their GMM based on images that were already registered. Furthermore, our segmentation has the benefit of a full ensemble of images, while the method proposed in [8] considers only two images.

In this paper, we have not considered strategies for increasing or decreasing the number of Gaussian components during registration. However, we conjecture that the number of Gaussian components can be used in a manner similar to a multiresolution strategy; starting with fewer components will help guide the registration process toward the global optimum, while increasing the number of components as convergence progresses can improve accuracy. We plan to investigate the impact of such techniques.

Prior knowledge could also be applied when choosing the number of Gaussian components. In fact, some applications might even permit a standard GMM as a starting point; such a strategy was demonstrated in [22]. Alternatively, one could employ a graph-theoretic approach such as normalized min-cuts [23] to determine an initial clustering of the scatter plot.

VII. Conclusion

Ensemble registration is the process of registering multiple images together simultaneously within a single optimization problem. This approach for multisensor registration was not previously feasible because the high-dimensional joint histogram was too large to store in memory. Instead, we use a Gaussian mixture model to perform density estimation of the content in the joint intensity space. This GMM model naturally leads to a cost function based on likelihood.

We formulate an optimization problem that has two aspects, developing solutions for the density estimation and motion parameters in synchrony. Within each iteration, we hold the motion parameters fixed and update the density estimation parameters, and then hold the density estimation parameters fixed and update the motion parameters.

Our experiments show that ensemble registration is more robust than pairwise registration. The content shared by one pair of images might be quite different from the content shared by another pair of images. The key is to leverage all these correspondences simultaneously. Ensemble registration does exactly that, implicitly coupling the content of all the images into one optimization problem.

The experiments also show that ensemble registration is more accurate than pairwise registration. Not only does ensemble registration offer more image correspondences (as described above), but it is also less susceptible to noise. This benefit stems from the fact that the estimate of an entity gets more accurate as you include more observations. Adding more images yields greater statistical confidence.

The density estimate generated by our clustering registration method can be viewed as a rudimentary segmentation of the ensemble. This idea offers promising future work in segmentation, labeling pixels and adaptively adjusting the cluters as the alignment is performed.

Our clustering registration method can be used for nonrigid registration as long as the motion transformation can be parameterized (in the variable θ). Examples of nonrigid parametric transformation models include B-splines, Fourier basis functions, and elastic deformation.

Acknowledgment

The authors would like to thank the following organizations for the use of their data: the Retrospective Image Registration Evaluation project, Landsat, and the Yale Face Database.

References

[1] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Med. Image Anal.*, vol. 5, no. 2, pp. 143–156, 2001.

[2] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," in *Proc. Info Proc Med Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, Eds., 1995, pp. 263–274.

[3] W. M. Wells, III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, no. 1, pp. 35–51, 1996.
[4] C. Studholme, D. L. G. Hill, and D. L. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognit.*, vol. 32, pp. 71–86, 1999.
[5] H. Neemuchwala, A. Hero, and P. Carson, "Image matching using alpha-entropy measures and entropic graphs," *Signal Process.*, vol. 85, no. 2, pp. 277–296, 2005.
[6] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Proc. MICCAI Lecture Notes on Computer Science, LNCS*, W. Wells, A. Colchester, and S. Delp, Eds., 1998, vol. 1996, pp. 1115–1124.
[7] A. Guimond, A. Roche, N. Ayache, and J. Meunier, "Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections," *IEEE Trans. Med Imag.*, vol. 20, no. 1, pp. 58–69, Jan. 2001.
[8] M. E. Leventon and W. E. L. Grimson, "Multi-modal volume registration using joint intensity distributions," in *Proc. MICCAI Lecture Notes on Computer Science, LNCS*, W. Wells, Ed., 1998, pp. 1057–1066.
[9] R. P. Woods, S. T. Grafton, C. J. Holmes, S. R. Cherry, and J. C. Mazziotta, "Automated image registration: I. General methods and intrasubject, intramodality validation," *J. Comput. Assist. Tomogr.*, vol. 22, pp. 139–152, 1998.
[10] K. K. Bhatia, J. V. Hajnal, B. K. Puri, A. D. Edwards, and D. Rueckert, "Consistent groupwise non-rigid registration for atlas construction," in *Proc. IEEE Int. Symp. Biomedical Imaging*, Apr. 2004, vol. 1, pp. 908–911.
[11] B. Avants and J. C. Gee, "Geodesic estimation for large deformation anatomical shape averaging and interploation," *NeuroImage*, vol. 23, pp. S139–S150, 2004, supplement.
[12] L. Zöllei, E. Learned-Miller, E. Grimson, and W. Wells, "Efficient population registration of 3D data," in *Proc. 1st Int. Workshop on Computer Vision for Biomedical Image Applications*, Oct. 2005, vol. LNCS 3765, pp. 291–301.
[13] Z. Xue, D. Shen, and C. Davatzikos, "CLASSIC: Consistent longitudinal alignment and segmentation for serial image computing," *NeuroImage*, vol. 30, no. 2, pp. 388–399, April 2006.
[14] S. Baloch, R. Verma, and C. Davatzikos, "An anatomical equivalence class based joint tranformation-residual descriptor for morphological anaysis," *Inf. Process. Med. Imag.*, pp. 594–606, 2007.
[15] P. Lorenzen, M. Prastawa, B. Davis, G. Gerig, E. Bullitt, and S. Joshi, "Multi-modal image set registration and atlas formation," *Med. Image. Anal.*, vol. 10, pp. 440–451, 2006.
[16] M. R. Sabuncu, M. E. Shenton, and P. Golland, "Joint registration and clustering of images," in *Proc. Medical Image Computing and Computer-Assisted Intervention*, 2007, pp. 47–54, ser. LNCS 4791.
[17] G. J. McLachlan and K. E. Basford, *Mixture Models: Inference and Applications to Clustering*. New York: Marcel-Dekker, 1988.
[18] M. Holden, D. L. G. Hill, E. R. E. Denton, J. M. Jarosz, T. C. S. Cox, T. Rohlfing, J. Goodey, and D. J. Hawkes, "Voxel similarity measures for 3-D serial MR brain image registration," *IEEE Trans. Med. Imag.*, vol. 19, no. 2, pp. 94–102, 2000.
[19] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for robust and accurate linear registration and motion correction of brain images," *NeuroImage*, vol. 17, no. 2, pp. 825–841, 2002.
[20] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dewant, J. Calvin, R. Maurer, R. M. Kessler, R. J. Maciunas, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Malandain, X. Pennec, M. E. Noz, J. Gerald, Q. Maguire, M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods, "Comparision and evaluation of retrospective intermodality brain image registration techniques," *J. Comput. Assit. Tomagr.*, vol. 21, pp. 554–566, 1997.
[21] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.
[22] J. Orchard, "Multimodal image registration using floating regressors in the joint intensity scatter plot," *Med. Image Anal.*, vol. 12, no. 4, pp. 385–396, Aug. 2008.
[23] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

Jeff Orchard (M'03) received the B.Math. degree in applied mathematics from the University of Waterloo, Waterloo, ON, Canada, in 1994, the M.Sc. degree in applied mathematics from the University of British Columbia, Canada, in 1996, and the Ph.D. degree in computing science from Simon Fraser University, Canada, in 2003.

Since 2003, he has been with the David R. Cheriton School of Computer Science at the University of Waterloo and was promoted to Associate Professor in 2009. His main research interests include image processing, complex systems, and neural networks.

Richard Mann received the B.Eng. degree in computer engineering and the M.Eng. degrees in electrical engineering from McMaster University, Hamilton, ON, Canada, and the Ph.D. degree in computer science from the University of Toronto, Toronto, ON, Canada.

He worked as a postdoctoral researcher at NEC Research Institute, Princeton, NJ. In 1999, he joined the faculty at the Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada, and is currently an Associate Professor. His research interests are in computational vision, artificial intelligence, and image processing.