

Nonrigid Image Registration Using an Entropic Similarity

Mohammed Khader and A. Ben Hamza, *Senior Member, IEEE*

Abstract—In this paper, we propose a nonrigid image registration technique by optimizing a generalized information-theoretic similarity measure using the quasi-Newton method as an optimization scheme and cubic B-splines for modeling the nonrigid deformation field between the fixed and moving 3-D image pairs. To achieve a compromise between the nonrigid registration accuracy and the associated computational cost, we implement a three-level hierarchical multiresolution approach such that the image resolution is increased in a coarse to fine fashion. Experimental results are provided to demonstrate the registration accuracy of our approach. The feasibility of the proposed method is demonstrated on a 3-D magnetic resonance data volume and also on clinically acquired 4-D CT image datasets.

Index Terms—Image registration, nonrigid, Tsallis entropy.

I. INTRODUCTION

NONRIGID image registration is of paramount importance in the field of medical imaging and has sparked a flurry of research interest in many other applications of image analysis such as remote sensing, movie editing, and archeology [1]. In recent years, various techniques have been proposed in the literature to tackle the nonrigid image registration problem [1]–[3]. Most of these techniques may be classified into two broad categories: feature-based and intensity-based methods. Feature-based approaches determine the registration at the feature locations, and an interpolation method is required at other locations. The registration accuracy for feature-based algorithms depends on the accuracy of the feature detector and involves detecting surface landmarks, edges and points [4]–[8]. On the other hand, intensity-based approaches employ matching techniques that involve the use of distance measures, such as the normalized cross correlation or the sum of squared differences [1], [2]. A general framework for intensity-based registration methods relies on information-theoretic similarity measures. One such similarity measure is the mutual information (MI), which was proposed independently by Viola and Wells [9] and by Maes *et al.* [10]. The MI measure has been shown to be effective in the development of the intensity-based image registration because of its

ability to register images from different modalities [11], [12]. Moreover, registration algorithms that maximize MI over rigid and affine transformations have reported impressive registration results [12]. Rueckert *et al.* [13] presented MI-based schemes for matching multimodal image pairs using B-splines to represent the deformation field on a regular grid. Most accurate methods for nonrigid registration are inspired by models from physics, either from elasticity [14], [15] or fluid mechanics [16], [17], but they are considered computationally expensive. Hence, several methods have been proposed based on various heuristics to approximate the underlying physical reality by alternative mathematical models [18]. Likar and Pernus [19] proposed a hierarchical image subdivision strategy by decomposing the nonrigid registration problem into an elastic interpolation of various local rigid registrations of subimages of decreasing size. This algorithm is applicable to both intra- and intermodal cases as it maximizes the local MI among subimages. Although MI has been successfully applied to nonrigid image registration, it is worth noting that MI-based registration methods might have a limited performance, once the initial misalignment of the two images is large or equally the overlap region of the two images is relatively small [20]. Moreover, MI is sensitive to the changes that occur in the distributions (overlap statistics) as a result of changes in the region of overlap. To circumvent these limitations, a number of information-theoretic methods have been proposed to improve the robustness of MI-based registration, including the normalized MI (NMI) approach [21]. The NMI-based approach is a robust similarity measure that allows for fully automated intermodal image registration algorithms. Wang and Vemuri [22] introduced the cross-cumulative residual entropy (CCRE), which is a measure of entropy defined in terms of cumulative distribution functions. In this approach, the CCRE between two images to be registered is maximized over the space of smooth and unknown nonrigid transformations. The reported results showed a better performance than MI-based methods. In [23], Loeckx *et al.* proposed the conditional MI (cMI) as a new similarity measure for nonrigid image registration. This measure was calculated as the expected value of the cMI between the image intensities given the spatial distribution. Recently, Myronenko and Song [24], [25] proposed to minimize a residual complexity (RC) instead of MI. This approach deals with complex spatially varying intensity distortions and produces accurate registration results.

In recent years, there has been a concerted research effort in statistical physics to explore the properties of Tsallis entropy, leading to a statistical mechanics that satisfies many of the properties of the standard theory [26]. Wachowiak *et al.* [27] introduced a generalized MI measure based on Tsallis entropy

Manuscript received November 2, 2010; revised March 30, 2011 and May 11, 2011; accepted May 28, 2011. Date of publication June 16, 2011; date of current version September 2, 2011. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Discovery Grant N00929.

The authors are with the Concordia Institute for Information Systems Engineering, Concordia University, Montréal, QC H3G 1M8, Canada (e-mail: ciise@linuxmail.org; hamza@ciise.concordia.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/TITB.2011.2159806

for 2-D–3-D multimodal biomedical image registration, and showed that their metric often produces fewer misregistrations compared to the MI approach. However, no results on nonrigid image registration were reported. Extending entropic methods to nonrigid image registration is the main focus of this paper. In the proposed approach, we model the nonrigid transformation of image coordinates using cubic B-splines [28]. Because of their attractive characteristics, such as inherent control of smoothness, separability in multiple dimensions, and computational efficiency, B-splines are used widely in the literature to model nonrigid deformations [13], [22]–[24], [29].

In this paper, we propose a nonrigid image registration method by optimizing the Jensen–Tsallis (JT) similarity measure using the quasi-Newton L-BFGS-B method [31] as an optimization scheme and cubic B-splines for modeling the nonrigid deformation field between the fixed and moving 3-D image pairs. The analytical gradient of the JT similarity is derived so that we can achieve an efficient and accurate nonrigid registration. In order to achieve a compromise between the nonrigid registration accuracy and the associated computational cost, we implement a three-level hierarchical multiresolution approach such that the image resolution is increased, along with the resolution of the control mesh, in a coarse to fine fashion. Experimental results are provided to demonstrate the registration accuracy of the proposed approach in comparison to RC and NMI approaches. The feasibility of the proposed algorithm is demonstrated on medical images from MRI with different protocols and also on clinically acquired 4-D CT image datasets.

The rest of the paper is organized as follows. In Section II, we describe in detail the proposed method, including the JT divergence and its main properties, the problem formulation, the deformation model, the estimation of the JT similarity and its derivative, as well as the summary of our proposed algorithm. Section III provides experimental results on a medical imaging dataset that demonstrate the effectiveness and superior performance of our method compared to RC and NMI approaches. And finally, in Section IV, we conclude and point out future work directions.

II. PROPOSED METHOD

In this section, we present the details of our proposed nonrigid image registration approach. First, we introduce the JT similarity measure that calculates how well the fixed and deformed image match. Then, we describe the transformation model that defines the space in which the best solution is found. Next, we present the optimization and derivative of the analytic gradient of the JT cost function with respect to nonrigid transformation parameters. Finally, we summarize our nonrigid registration algorithm.

A. JT Similarity

Shannon’s entropy of a discrete probability distribution $\mathbf{p} = (p_1, p_2, \dots, p_k)$ is defined as $H(\mathbf{p}) = -\sum_{j=1}^k p_j \log(p_j)$. A

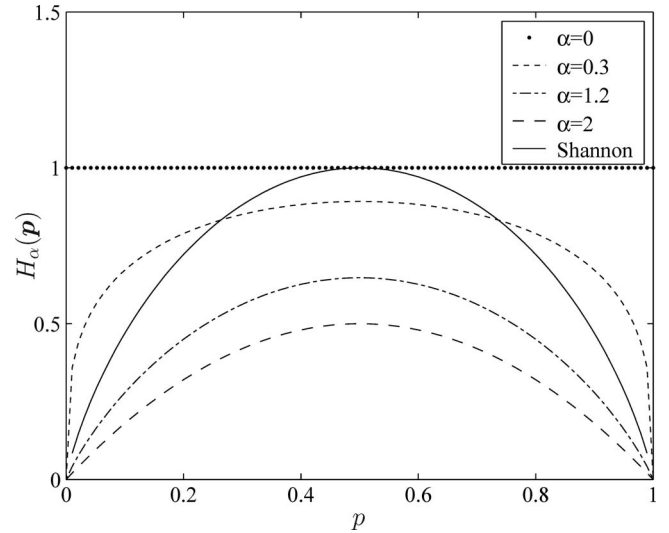


Fig. 1. Tsallis entropy $H_\alpha(\mathbf{p})$ of a Bernoulli distribution $\mathbf{p} = (p, 1-p)$ for different values of α .

generalization of Shannon entropy is Tsallis entropy given by

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \left(\sum_{j=1}^k p_j^\alpha - 1 \right) = -\sum_{j=1}^k p_j^\alpha \log_\alpha(p_j) \quad (1)$$

where \log_α is the α -logarithm function defined as $\log_\alpha(x) = (1-\alpha)^{-1}(x^{1-\alpha} - 1)$ for $x > 0$, and $\alpha \in (0, 1) \cup (1, \infty)$ is an exponential order also referred to as entropic index. This generalized entropy is widely used in statistical physics applications [26].

If we consider that a physical system can be decomposed in two statistical independent subsystems with probability distributions \mathbf{p} and \mathbf{q} , then it can be shown that the joint Tsallis entropy is pseudoadditive

$$H_\alpha(\mathbf{p}, \mathbf{q}) = H_\alpha(\mathbf{p}) + H_\alpha(\mathbf{q}) + (1-\alpha)H_\alpha(\mathbf{p})H_\alpha(\mathbf{q}) \quad (2)$$

whereas the joint Shannon’s entropy is additive: $H(\mathbf{p}, \mathbf{q}) = H(\mathbf{p}) + H(\mathbf{q})$. Pseudoadditivity implies that Tsallis entropy has a nonextensive property for statistical independent systems. Further, standard thermodynamics is extensive because of the short-range nature of the interaction between subsystems of a composite system. In other words, when a system is composed of two statistically independent subsystems, then Shannon entropy of the composite system is just the sum of entropies of the individual systems, and hence the correlations between the subsystems are not accounted for. However, Tsallis entropy does take into account these correlations due to its pseudoadditivity property. Fig. 1 depicts Tsallis entropy of a Bernoulli distribution $\mathbf{p} = (p, 1-p)$, for different values of the entropic index. As illustrated in Fig. 1, the measure of uncertainty is at a maximum when Shannon’s entropy is used, and for $\alpha \geq 1$ it decreases as the parameter α increases. Furthermore, Tsallis entropy attains a maximum uncertainty when its exponential order α is equal to zero.

Definition 1: Let $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ be n probability distributions. The JT divergence is defined as

$$D_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) = H_\alpha \left(\sum_{i=1}^n \omega_i \mathbf{p}_i \right) - \sum_{i=1}^n \omega_i H_\alpha(\mathbf{p}_i) \quad (3)$$

where $H_\alpha(\mathbf{p})$ is Tsallis entropy, and $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ is a weight vector such that $\sum_{i=1}^n \omega_i = 1$ and $\omega_i \geq 0$.

Using the Jensen inequality, it is easy to check that the JT divergence [30] is nonnegative for $\alpha > 0$. It is also symmetric and vanishes if and only if all the probability distributions are equal, for all $\alpha > 0$.

The following result establishes the convexity of the JT divergence of a set of probability distributions [32].

Proposition 1: For $\alpha \in [1, 2]$, the JT divergence D_α^ω is a convex function of $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$.

In the sequel, we will restrict $\alpha \in [1, 2]$, unless specified otherwise. In addition to its convexity property, the JT divergence is an adapted measure of disparity among n probability distributions as shown in the next result [33].

Proposition 2: The JT divergence D_α^ω achieves its maximum value when $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ are degenerate distributions, that is $\mathbf{p}_i = (\delta_{ij})$, where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

Proposition 3: The upper bound of the JT divergence is given by $D_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) \leq H_\alpha(\omega)$.

Proof: Since the JT divergence is a convex function of $\mathbf{p}_1, \dots, \mathbf{p}_n$, it achieves its maximum value when Tsallis entropy of the ω -weighted average of degenerate probability distributions, achieves its maximum value as well. Assigning weights ω_i to the degenerate distributions $\delta_1, \delta_2, \dots, \delta_n$, where $\delta_i = (\delta_{ij})$, the following upper bound

$$D_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) \leq H_\alpha \left(\sum_{j=1}^n \omega_j \delta_j \right) = H_\alpha(\omega)$$

is achieved. \blacksquare

Since $H_\alpha(\omega)$ attains its maximum value when the weights are uniformly distributed (i.e., $\omega_i = 1/n, \forall i$), it follows that a tight upper bound of the JT divergence is given by

$$D_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) \leq H_\alpha(1/n, \dots, 1/n) = \log_\alpha n. \quad (4)$$

If we are measuring the similarity, $S_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)$, between densities, then $S_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)$ should satisfy the conditions

$$\begin{aligned} S_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) &\geq 0 \\ S_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) &= \max \end{aligned}$$

where max is the maximum similarity possible on the scale of measurement being used, and often this will be unity. Therefore, using (4) we may define the JT similarity measure as follows:

$$S_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) = \frac{\varrho(D_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)) - \varrho(\log_\alpha n)}{\varrho(0) - \varrho(\log_\alpha n)} \quad (5)$$

where ϱ is a monotonous decreasing function such that $\varrho(\log_\alpha n) < \varrho(0)$. For simplicity, we choose $\varrho(x) = 1 - x$ as a monotonous decreasing function and by substituting this func-

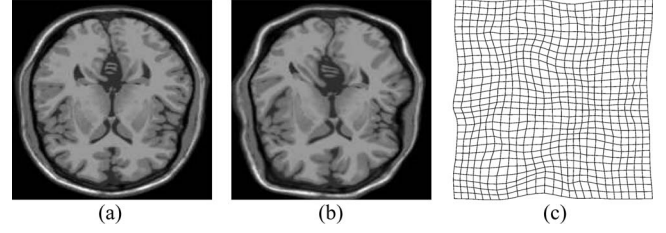


Fig. 2. (a) Fixed image I ; (b) moving image J ; (c) deformation field Φ

tion into (5), the JT similarity measure becomes

$$S_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) = 1 - \frac{D_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)}{\log_\alpha n}. \quad (6)$$

It is worth pointing out that the main purpose of introducing the JT similarity measure S_α^ω is to use the limited memory quasi-Newton minimization method, which efficiently solves nonlinear and large-scale minimization problems.

B. Problem Statement

In the sequel, we will use the JT similarity measure, given by (6), as a matching criterion to solve the image alignment problem. Let I and J be two misaligned images to be registered, where I is the fixed image and J is the moving image. The moving image J is obtained by applying a deformation field Φ to the fixed image I , as depicted in Fig. 2. The deformation field Φ is described by a transformation function $g(\mathbf{x}; \boldsymbol{\mu}) : V_J \rightarrow V_I$, where V_J and V_I are continuous domains on which J and I are defined, and $\boldsymbol{\mu}$ is a set of transformation parameters to be determined. The image alignment or registration problem may be formulated as an optimization problem

$$\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu}} S_\alpha^\omega(I(\mathbf{x}), J(g(\mathbf{x}; \boldsymbol{\mu}))). \quad (7)$$

To align the transformed moving image $J(g(\mathbf{x}; \boldsymbol{\mu}))$ to the fixed image I , we seek the set of transformation parameters $\boldsymbol{\mu}$ that minimize the JT cost function $S_\alpha^\omega(I(\mathbf{x}), J(g(\mathbf{x}; \boldsymbol{\mu})))$.

C. Transformation Model

Several transformation models have been proposed over the years to represent a nonrigid deformation field. In this paper, we model the transformation $g(\mathbf{x}; \boldsymbol{\mu})$ using the free-form deformation (FFD) [29], which is based on cubic B-splines, and $\boldsymbol{\mu}$ represents the parameter vector of deformation coefficients. Let Φ denote a $n_x \times n_y \times n_z$ mesh of control points $\varphi_{i,j,k}$ with a uniform spacing Δ . Then, the 3-D transformation at any point $\mathbf{x} = [x, y, z]^T$ in the moving image is interpolated using a linear combination of cubic B-spline convolution kernels as follows:

$$g(\mathbf{x}; \boldsymbol{\mu}) = \sum_{ijk} \eta_{ijk} \beta^{(3)} \left(\frac{\mathbf{x} - \varphi_{ijk}}{\Delta} \right) \quad (8)$$

where $\beta^{(3)}(\mathbf{x}) = \beta^{(3)}(x)\beta^{(3)}(y)\beta^{(3)}(z)$ is a separable cubic B-spline convolution kernel, and η_{ijk} are the deformation coefficients associated to the control points φ_{ijk} . The degree of nonrigidity can be adopted to a specific registration problem by varying the mesh spacing or the resolution of the mesh Φ

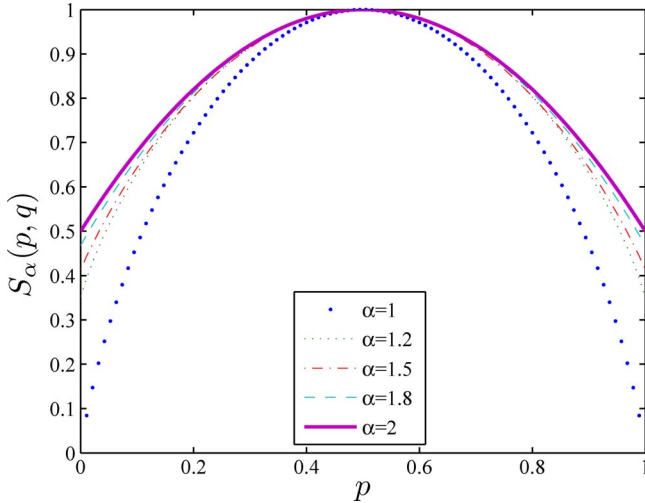


Fig. 3. JT similarity $S_\alpha(\mathbf{p}, \mathbf{q})$ between two Bernoulli distributions $\mathbf{p} = (p, 1 - p)$ and $\mathbf{q} = (1 - p, p)$ for different values of α .

of control points. The parameter vector $\boldsymbol{\mu} = (\eta_{ijk})$ represents the vector of deformation coefficients associated to the control points φ_{ijk} , where the indices i, j, k denote the coordinates of the control points on the mesh grid.

D. Optimization of the JT Similarity

We adopt a limited memory quasi-Newton method for solving the optimization problem given by (7). The calculation of the analytical gradient of the JT cost (similarity) function is necessary to not only avoid discretization errors but also to achieve an efficient and robust minimization scheme. Denote by $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ and $\mathcal{Y} = \{y_1, y_2, \dots, y_n\}$ the sets of pixel intensity values of the fixed image $I(\mathbf{x})$ and the deformed moving image $J(g(\mathbf{x}; \boldsymbol{\mu}))$, respectively. Let X and Y be two random variables taking values in \mathcal{X} and \mathcal{Y} . Then, we define the conditional intensity probability distributions \mathbf{p}_i as follows:

$$\mathbf{p}_i = \mathbf{p}_i(J(g(\mathbf{x}; \boldsymbol{\mu})|I(\mathbf{x})) = (p_{ij})_{j=1, \dots, n}, \quad \forall i = 1, \dots, n$$

where $p_{ij} = P(Y = y_j | X = x_i) = p(j|i; \boldsymbol{\mu})$, $j = 1, \dots, n$. Note that in p_{ij} the parameter vector $\boldsymbol{\mu}$ is omitted for notational simplicity.

It is worth pointing out that if the images I and J are exactly matched, then $\mathbf{p}_i = (\delta_{ij})$ and by Proposition 2, the JT divergence is, therefore, maximized and consequently the JT similarity measure is minimized. By substituting (1) into (3), we obtain

$$D_\alpha^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) = \frac{1}{1 - \alpha} \left[\sum_{j=1}^n \left(\sum_{i=1}^n \omega_i p_{ij} \right)^\alpha - \sum_{i=1}^n \omega_i \sum_{j=1}^n p_{ij}^\alpha \right]. \quad (9)$$

Fig. 3 illustrates the JT similarity between two Bernoulli distributions $\mathbf{p} = (p, 1 - p)$ and $\mathbf{q} = (1 - p, p)$ for different values of the entropic index. As shown in Fig. 3, the highest similarity corresponds to the entropic index $\alpha = 2$. In the sequel, we choose an entropic index $\alpha = 2$. Thus, the JT divergence is

reduced to

$$D_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) = - \sum_{j=1}^n \left(\sum_{i=1}^n \omega_i p_{ij} \right)^2 + \sum_{i=1}^n \omega_i \sum_{j=1}^n p_{ij}^2 \quad (10)$$

and the JT similarity becomes

$$S_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) = 1 - \frac{D_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)}{\log_2 n}. \quad (11)$$

The calculation of the registration function is as follows. First, we calculate the conditional intensity probability distributions \mathbf{p}_i , $i = 1, \dots, n$ between the fixed and deformed moving images. Then, we compute $D_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)$ according to the formula given by (10). Finally, we compute the similarity $S_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)$ given by (11), which represents our registration function. Note that we are using conditional (not marginal) probabilities to compute the JT similarity metric. Moreover, it is important to point out that n denotes the number of pixel values in the fixed image and also in the moving image. Thus, the number of conditional probabilities is also equal to n . In addition, both indices i and j in (10) run from 1 to n .

1) *Conditional Intensity Probability Estimation*: In a typical registration problem, direct access to the marginal and joint probability densities is not available and hence the densities must be estimated from the image data. Parzen windows (also known as kernel density estimators) can be used for this purpose. In this scheme, the densities are constructed by taking intensity samples from the image and super-positioning kernel functions centered on the elements of these samples. A variety of functions can be used as smoothing kernels with the requirement that they are smooth, symmetric, have zero mean and integrate to one. For example, boxcar, Gaussian, and B-spline functions are suitable candidates. We propose to use the B-spline Parzen window to estimate the conditional intensity probability of the interpolated moving image given the fixed image. The advantage of using a B-spline kernel over a Gaussian kernel is that the B-spline kernel has a finite support region which is computationally attractive, as each intensity sample only affects a small number of bins and hence does not require an $N \times N$ loop to compute the metric value. Let $\beta^{(0)}$ be a zero-order spline Parzen window and $\beta^{(3)}$ be a cubic spline Parzen window, then the smoothed conditional probability of $J(g(\mathbf{x}; \boldsymbol{\mu}))$ given $I(\mathbf{x})$ is expressed as follows:

$$p(j|i; \boldsymbol{\mu}) = \frac{p(j, i; \boldsymbol{\mu})}{p_I(i)} \quad (12)$$

where

$$p(j, i; \boldsymbol{\mu}) = \xi \sum_{\mathbf{x} \in V} \beta^{(0)} \left(i - \frac{I(\mathbf{x}) - f_I^0}{\Delta b_I} \right) \cdot \beta^{(3)} \left(j - \frac{J(g(\mathbf{x}; \boldsymbol{\mu})) - f_J^0}{\Delta b_J} \right). \quad (13)$$

The normalization factor ξ ensures sum to one of the probabilities, and $I(\mathbf{x})$ and $J(g(\mathbf{x}; \boldsymbol{\mu}))$ are samples of the fixed and interpolated moving images, respectively. These samples are normalized by the minimum intensity value, f_I^0 , f_J^0 , and intensity range of each bin, Δb_I and Δb_J , respectively.

The marginal probability of the fixed image is independent of the transformation parameters, and can be computed as follows:

$$p_I(i) = \xi \sum_{\mathbf{x} \in V} \beta^{(0)} \left(i - \frac{I(\mathbf{x}) - f_I^0}{\Delta b_I} \right). \quad (14)$$

Since the fixed image probability density function does not contribute to the cost function derivative, it does not need to be smooth. Hence, a zero order B-spline kernel is used.

By taking the derivative of the conditional probability with respect to $\boldsymbol{\mu}$, we get

$$\begin{aligned} \frac{\partial p(j|i; \boldsymbol{\mu})}{\partial \boldsymbol{\mu}} &= \frac{\gamma}{p_I(i) \Delta b_J} \sum_{\mathbf{x} \in V} \beta^{(0)} \left(i - \frac{I(\mathbf{x}) - f_I^0}{\Delta b_I} \right) \\ &\quad \cdot \beta^{(3)} \left(j - \frac{J(g(\mathbf{x}; \boldsymbol{\mu})) - f_J^0}{\Delta b_J} \right) \\ &\quad \cdot \left(\frac{\partial J(t)}{\partial t} \Big|_{t=g(\mathbf{x}; \boldsymbol{\mu})} \right) \frac{\partial g(\mathbf{x}; \boldsymbol{\mu})}{\partial \boldsymbol{\mu}} \end{aligned} \quad (15)$$

where $\partial J(t)/\partial t$ is the image gradient, and $\beta^{(3)}$ is the derivative of the cubic spline kernel

$$\beta^{(3)}(u) = \begin{cases} 0.0 & u \leq -2 \\ 2u + 2 + \frac{1}{2}u^2 & -2 < u \leq -1 \\ -2u - \frac{3}{2}u^2 & -1 < u \leq 0 \\ -2u + \frac{3}{2}u^2 & 0 < u \leq 1 \\ 2u - 2 - \frac{1}{2}u^2 & 1 < u \leq 2 \\ 0.0 & u > 2 \end{cases} \quad (16)$$

E. Derivative of the JT Similarity

Now, taking the derivative of the JT divergence with respect to $\boldsymbol{\mu}$ yields

$$\begin{aligned} \frac{\partial D_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)}{\partial \boldsymbol{\mu}} &= \frac{\partial \left(- \sum_{j=1}^n \left(\sum_{i=1}^n \omega_i p_{ij} \right)^2 \right)}{\partial \boldsymbol{\mu}} \\ &\quad + \frac{\partial \left(\sum_{i=1}^n \omega_i \sum_{j=1}^n p_{ij}^2 \right)}{\partial \boldsymbol{\mu}} \end{aligned} \quad (17)$$

where

$$\begin{aligned} &\frac{\partial \left(- \sum_{j=1}^n \left(\sum_{i=1}^n \omega_i p_{ij} \right)^2 \right)}{\partial \boldsymbol{\mu}} \\ &= -2 \sum_{j=1}^n \left(\sum_{i=1}^n \omega_i p_{ij} \right) \frac{\partial \left(\sum_{i=1}^n \omega_i p_{ij} \right)}{\partial \boldsymbol{\mu}} \\ &= -2 \sum_{j=1}^n \left(\sum_{i=1}^n \omega_i p_{ij} \right) \left(\sum_{i=1}^n \omega_i \frac{\partial p_{ij}}{\partial \boldsymbol{\mu}} \right) \end{aligned} \quad (18)$$

Algorithm 1 Proposed nonrigid registration approach

- 1: Initialize the deformation field Φ
 - 2: **for** *hierarchical_level* = 1 to 3 **do**
 - 3: Calculate the cost function and its gradient as given by Eq. (20)
 - 4: **repeat**
 - 5: Use the quasi-Newton method to solve the optimization problem given by Eq. (7)
 - 6: Update the deformation field
 - 7: Recalculate the cost function and its gradient
 - 8: **until** the difference in consecutive iterates is less than $\epsilon = 0.01$
 - 9: Increase the resolution of both the deformation field and the image.
 - 10: **end for**
-

and

$$\begin{aligned} \frac{\partial \left(\sum_{i=1}^n \omega_i \sum_{j=1}^n p_{ij}^2 \right)}{\partial \boldsymbol{\mu}} &= \sum_{i=1}^n \omega_i \sum_{j=1}^n \frac{\partial p_{ij}^2}{\partial \boldsymbol{\mu}} \\ &= 2 \sum_{i=1}^n \omega_i \sum_{j=1}^n p_{ij} \frac{\partial p_{ij}}{\partial \boldsymbol{\mu}}. \end{aligned}$$

Therefore

$$\begin{aligned} \frac{\partial D_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)}{\partial \boldsymbol{\mu}} &= -2 \sum_{j=1}^n \left(\sum_{i=1}^n \omega_i p_{ij} \right) \left(\sum_{i=1}^n \omega_i \frac{\partial p_{ij}}{\partial \boldsymbol{\mu}} \right) \\ &\quad + 2 \sum_{i=1}^n \omega_i \sum_{j=1}^n p_{ij} \frac{\partial p_{ij}}{\partial \boldsymbol{\mu}}. \end{aligned} \quad (19)$$

Consequently, the JT similarity measure and its derivative are given by

$$\begin{cases} S_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n) = 1 - \frac{D_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)}{\log_2 n} \\ \frac{\partial S_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)}{\partial \boldsymbol{\mu}} = - \frac{\partial D_2^\omega(\mathbf{p}_1, \dots, \mathbf{p}_n)}{\partial \boldsymbol{\mu}} \times \frac{1}{\log_2 n}. \end{cases} \quad (20)$$

F. Summary of the Proposed Algorithm

The proposed algorithm is implemented by changing the deformation in the moving image until the discrepancy between the moving and fixed images is minimized. The main algorithmic steps of our nonrigid image registration approach are summarized in Algorithm 1. First, the algorithm initializes the deformation field Φ by creating a uniform B-spline control grid with predefined spacing knots. Next, a three-level hierarchical multiresolution scheme is used to achieve the best compromise between the registration accuracy and the associated computational cost. As the hierarchical level increases the resolution of the control mesh is increased, along with the image resolution, in a coarse to fine fashion. In each hierarchical level, a limited-memory, quasi-Newton minimization scheme is used to find the optimum set of transformation parameters that reduce the JT cost function until the difference between the cost function values in two consecutive iterations is less than $\epsilon = 0.01$. The

resolution of the optimum set of transformation parameters, at a courser level, is increased to be used as starting point for the next hierarchical level.

III. EXPERIMENTAL RESULTS

We tested the performance of the proposed approach on a medical imaging dataset that was obtained from the brainweb database at the Montreal Neurological Institute [34]. This dataset contains a full 3-D simulated brain magnetic resonance (MR) data volumes from several protocols, including T1-weighted (MR-T1), T2-weighted (MR-T2), and proton density (MR-PD) with a variety of slice thicknesses, noise levels, and levels of intensity nonuniformity. All the corresponding slices from different protocols are originally aligned with each other. The images used in our experiments have $181 \times 217 \times 181$ voxels with a 1-mm voxel size in each dimension. To validate the nonrigid registration accuracy of the proposed method, we first applied both geometric and intensity distortions to the fixed image in order to generate moving image. Then, we aligned the moving image with the fixed image. We also compared the image registration results of our approach to RC and NMI approaches, which are implemented in the Medical Image Registration Toolbox [24] and in the Image Registration Toolkit [13], respectively. In all the experiments, we used an entropic index $\alpha = 2$ and the normalized histogram of the fixed image as the weight vector ω for the JT similarity measure.

For the implementation of RC and NMI methods, we essentially model the nonrigid deformation field as a FFD based on B-splines and then we employ an iterative gradient descent scheme as an optimization algorithm. In all the experiments, the moving image is generated by applying a random perturbation to the corresponding fixed image using a thin-plate spline interpolation (TPS) such that the mean nonrigid displacement of the pixels, caused by the relative displacement between the fixed and generated moving images, is the ground truth deformation field μ_g . Moreover, it is important to mention that none of the three methods (NMI, RC, and proposed approach) uses TPS interpolation as its transformation model. Therefore, using TPS interpolation for generating moving image is not unfairly advantageous to any of these methods.

A. Registration Functions

An ideal registration function that measures the similarity between two images should be smooth and convex with respect to different transformation parameters. Also, the global minimum of the registration function should be close to the correct transformation parameters that align two images perfectly [35]. Moreover, the capture range around the global minimum should be as large as possible, and the number of local minima of the registration function should be as small as possible. These criteria will be used to evaluate the registration functions generated by NMI, RC, and the proposed approach. The registration function of our algorithm can be generated by computing the JT similarity between two images under all possible transformations. Similarly, the registration functions of NMI and RC can be obtained.

The registration functions of NMI, RC, and the proposed approach with respect to different rotation and translation parameters are shown in Fig. 4. We can observe that the registration functions of the proposed approach are much smoother than those of NMI and RC. In addition, the capture range in the registration function of our method is considerably large. In particular, the change of the registration function with respect to rotations is smoothly extended relatively far from the global minimum, indicating a better performance of the proposed approach.

B. Monomodality Test

In the first experiment, we distorted the fixed image MR-T1 with a known nonrigid transformation field, or the so-called ground truth deformation μ_g . Then, we applied the proposed approach, RC and NMI algorithms. And finally, we compared the obtained deformations fields with the ground truth. Fig. 5 shows the results obtained from this experiment. Note that the registered moving images obtained by the proposed method and the RC approach are visually more similar in shape to the moving image than the image produced by the NMI approach. Moreover, the estimated transformation field resulted from our approach is more similar to the ground truth than of those obtained using RC and NMI approaches. To measure the registration accuracy of the proposed method, we computed the mean $\hat{\mu}_{\text{MSE}}$ and standard deviation $\hat{\sigma}_{\text{MSE}}$ of the MSE between the ground truth and estimated displacement vectors. Table I displays the MSE statistics of the estimated nonrigid deformation, when compared to the ground truth. The first column shows the mean ground truth deformation, which represents the magnitude of the displacement vector that is used to generate the moving images in each experiment. For each row twenty different transformation fields with this mean are generated and applied to the fixed image in order to generate the corresponding moving images. The second and third columns display the average and standard deviation of MSE for the generated 20 pairs of fixed-moving images. The results obtained using the proposed approach are considerably small compared to those of RC and NMI methods.

In the second experiment, we used similar steps as in the first experiment, but this time we generated the moving image by distorting the fixed image with both intensity and geometric distortions. The intensity distortion is generated by corrupting the fixed image as follows [24]:

$$1) \quad I(x, y) = I^\gamma(x, y) + v \frac{xy}{MN} + \frac{1}{K} \sum_{k=1}^K \exp\left(-\frac{\| [x; y] - \Psi_K \|^2}{2\sigma^2}\right)$$

2) Rescale I to $[0, 1]$,

where the first term represents the gamma correction on I after geometric distortion, the second term models a smoothly varying global intensity field, and the third term models locally varying intensity field with a mixture of K Gaussian densities. In this experiment, we chose a distortion level 2 with parameters as follows: $v = 0.4$, $K = 1$, Ψ_K were randomly selected from the interval $[1, \nu]$ (ν is the size of the image domain), $\sigma = 30$, and γ is selected randomly from $[0.9, 1.2]$. The registered images

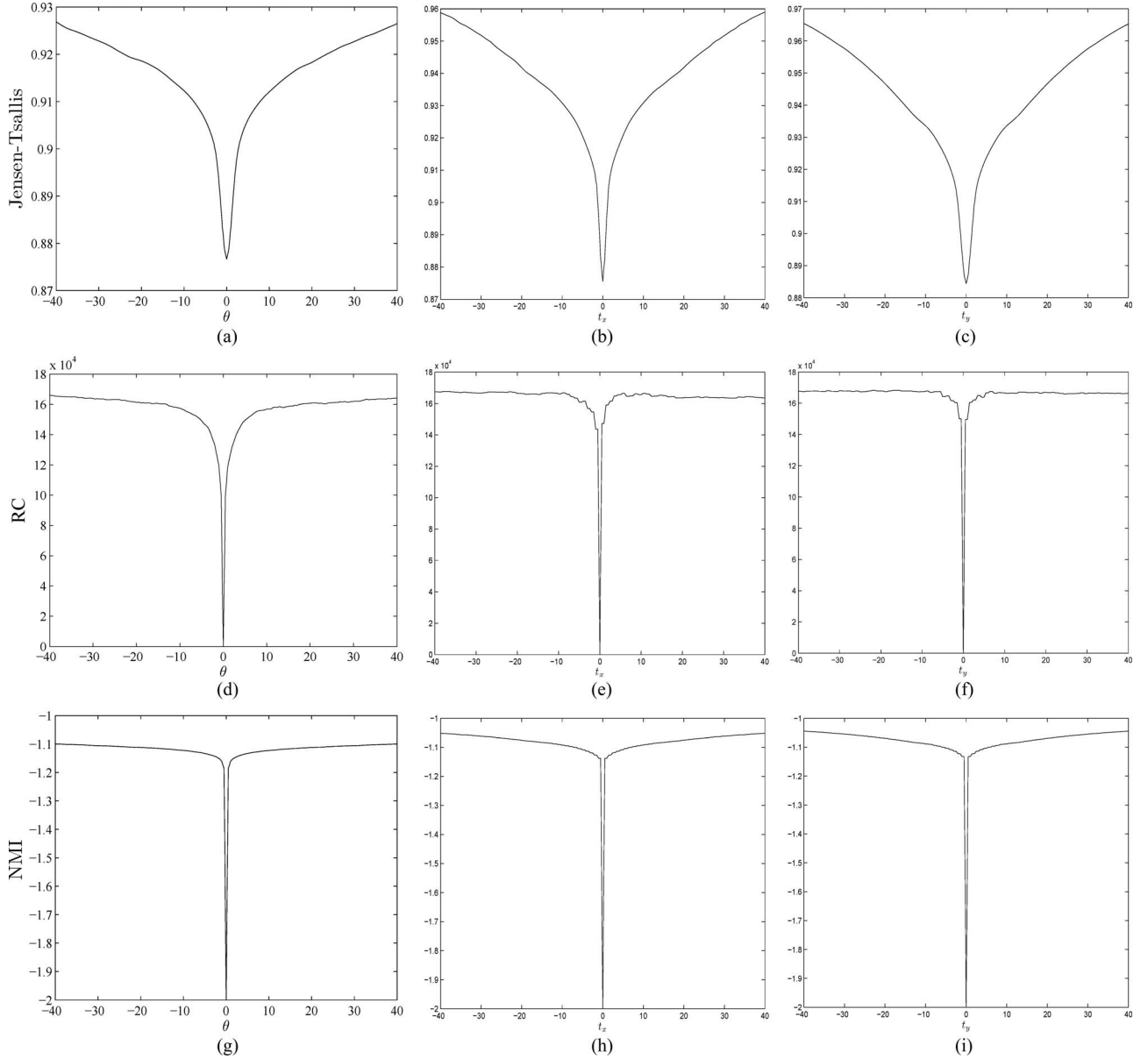


Fig. 4. Registration functions of our approach, RC, and NMI in aligning MRI-T1 images. From top to bottom: (a), (d), and (g) rotation only; (b), (e), and (h) translation along x -axis; (c), (f), and (i) translation along y -axis.

shown in Fig. 6 demonstrate that the proposed algorithm outperforms the RC approach, in the presence of spatially varying intensity distortion. The result obtained by the NMI approach shows a poor performance. Moreover, the estimated deformation field obtained by our approach shows superior accuracy in comparison to RC and NMI methods.

C. Multimodality Test

The images used in this experiment are corresponding slices from MR-T1 and MR-T2 image pair, and they are originally aligned with each other. In this experiment, we registered the geometrically deformed MR-T1 image onto MR-T2 image using our approach and the NMI method. We omitted the result of the RC approach because it is not applicable to multimodal images. Fig. 7 shows the accuracy of our method in registering

images from different modalities. As can be seen, the registered image using the NMI approach still has a considerable amount of misregistration. However, most of the visible amount of misalignment in the moving image has been removed after applying the proposed approach. In addition, the nonrigid transformation estimated by the proposed method looks very similar to the ground truth, indicating a much better performance of our approach.

D. Statistical Significance Test

We used a paired t -test to determine if the difference in MSE for the pairs of registration methods is statistically significant. Table II shows the p -values of the differences of MSE between each pair of registration methods after applying paired t -test. At 95% level of confidence, it is evident from Table II that the

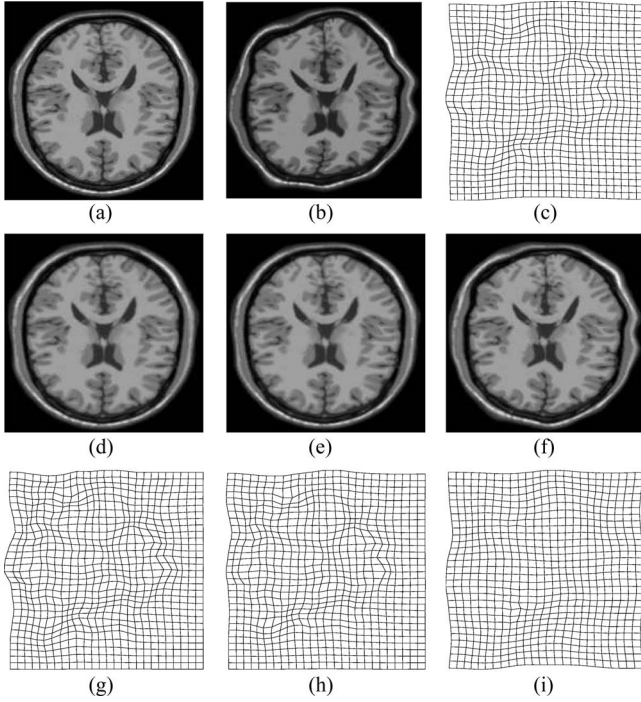


Fig. 5. Geometric distortion experiment : (a) MR-T1 image; (b) distorted MR-T1 image with geometric distortion; (c) ground truth deformation field; (d)–(f) registered images using our approach RC, and NMI, respectively; (g)–(i) estimated transformation using our approach RC, and NMI, respectively.

TABLE I
MSE STATISTICS OF THE ESTIMATED NONRIGID DEFORMATION

μ_g	our method		RC		NMI	
	$\hat{\mu}_{MSE}$	$\hat{\sigma}_{MSE}$	$\hat{\mu}_{MSE}$	$\hat{\sigma}_{MSE}$	$\hat{\mu}_{MSE}$	$\hat{\sigma}_{MSE}$
2.1	0.4982	0.0325	0.5013	0.0251	0.7133	0.0436
3.0	0.5971	0.0843	0.5994	0.0413	0.8936	0.0362
4.6	0.7002	0.0157	0.7342	0.0321	1.0933	0.0733
5.1	0.7621	0.0241	0.7998	0.0379	1.4992	0.0536
5.7	0.7922	0.0252	0.8213	0.0317	1.6916	0.0837

MSE resulting from our approach is significantly lower than the MSEs resulting from RC and NMI methods, indicating the better performance of the proposed approach.

E. Real-Life Data Set: 3-D Thoracic CT Images

We also validated the proposed method, RC and NMI approaches on real-life clinically acquired 4-D CT image datasets [36]. These publicly available datasets consist of thoracic 4-D CT images acquired as part of the standard planning process for the treatment of esophageal cancer at the University of Texas M. D. Anderson Cancer Center in Houston. Each 4-D CT scan consists of a time-varying stack of ten 3-D images of the entire thorax and upper abdomen, and were acquired at 2.5-mm slice spacing with a General Electric Discovery ST PET/CT scanner (GE Medical Systems, Waukesha, WI). The utmost inhale and exhale phases of the 4-D CT sets were used for nonrigid registration evaluation. For each patient dataset, an expert in thoracic imaging manually delineated and registered pulmonary landmark features such as vessel bifurcations between the fixed and moving image pairs as shown in Fig. 8. A total number of

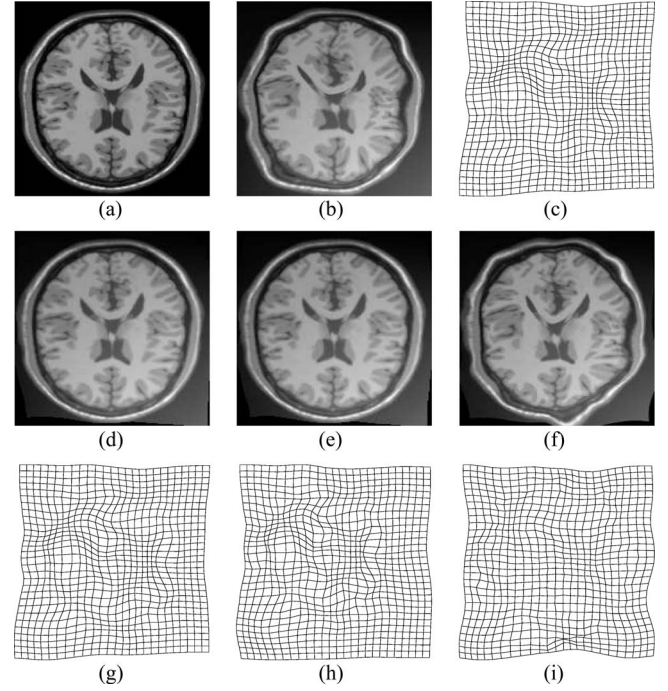


Fig. 6. Geometric and intensity distortion experiment: (a) MR-T1 image; (b) distorted MR-T1 image with geometric and intensity distortion; (c) ground truth deformation field; (d)–(f) registered images using our approach, RC, and NMI, respectively; (g)–(i) estimated transformation using our approach, RC, and NMI, respectively.

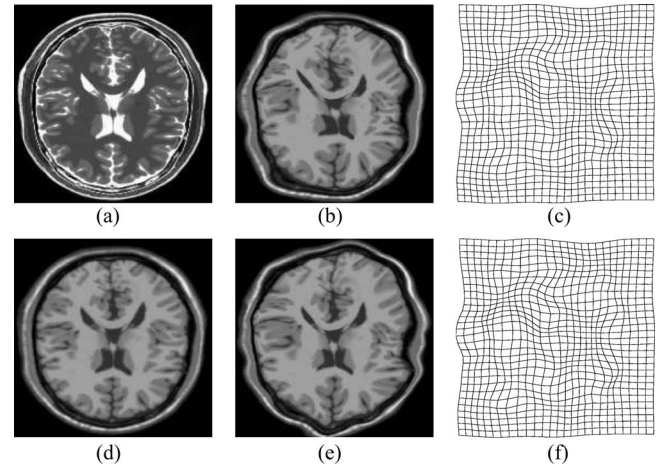


Fig. 7. Multimodality experiment: (a) MR-T2 image; (b) distorted MR-T1 image with geometric distortion; (c) ground truth deformation field; (d) and (e) registered images using our approach, and NMI, respectively; (f) estimated transformation using our approach.

four cases are used in this evaluation with a minimum of 1 166 registered landmark features in each case as shown in Table III.

The deformable registration accuracy was evaluated by calculating the point registration errors as a 3-D Euclidean distance between the manually determined landmark position in the exhale image and that calculated by applying the registration methods to the corresponding feature location in the inhale image. The mean registration errors and corresponding standard deviations were calculated for all cases before

TABLE II
COMPARISON BETWEEN THE REGISTRATION METHODS USING STATISTICAL SIGNIFICANT TEST (p -VALUES)

Registration Methods	μg				
	2.1	3.0	4.6	5.1	5.7
our method vs. RC	0.0493	0.0537	0.0067	0.0013	0.000585
our method vs. NMI	2.5033×10^{-10}	7.544×10^{-12}	3.2419×10^{-14}	0	0
RC vs. NMI	5.1467×10^{-10}	5.3916×10^{-11}	1.6183×10^{-12}	0	0

TABLE III
CT-IMAGE AND REFERENCE LANDMARK PROPERTIES

CT properties	Case 1	Case 2	Case 3	Case 4
Volume dimensions	$256 \times 256 \times 94$	$256 \times 256 \times 112$	$256 \times 256 \times 104$	$256 \times 256 \times 99$
Voxel dimensions(mm)	$0.97 \times 0.97 \times 2.5$	$1.16 \times 1.16 \times 2.5$	$1.15 \times 1.15 \times 2.5$	$1.13 \times 1.13 \times 2.5$
Landmark points distribution				
# Right lung	672	767	637	803
# Left lung	608	720	529	758
Total	1280	1487	1166	1561

The image volumes dimensions in units of voxels and also the voxel dimensions are shown for each patient case used in this experiment. The distribution of manually registered landmark pairs regarding right, left and total lung feature points are shown as well. More details about generation and characterization of landmarks can be found in [37].

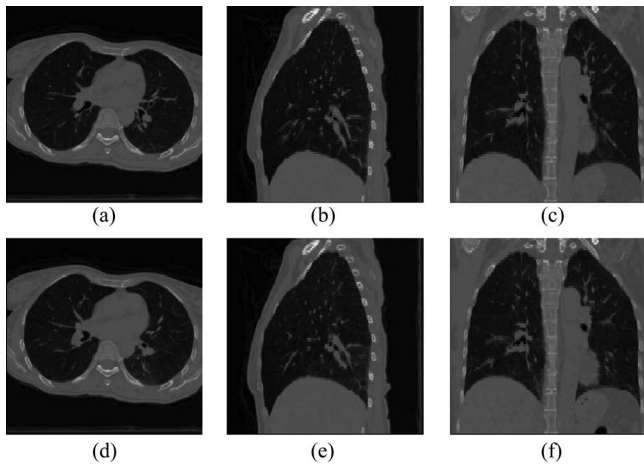


Fig. 8. 4-D CT scan of a lung taken during a single breath cycle: fixed (top row) and moving (bottom row) represent component phase 3-D volumes from a 4-D set for patient case number one. (a)–(c) display a 3-D volume of the maximum inhale phase in transverse, coronal, and sagittal orientation, respectively; (d)–(f) display a 3-D volume of the maximum exhale phase in transverse, coronal, and sagittal orientation, respectively.

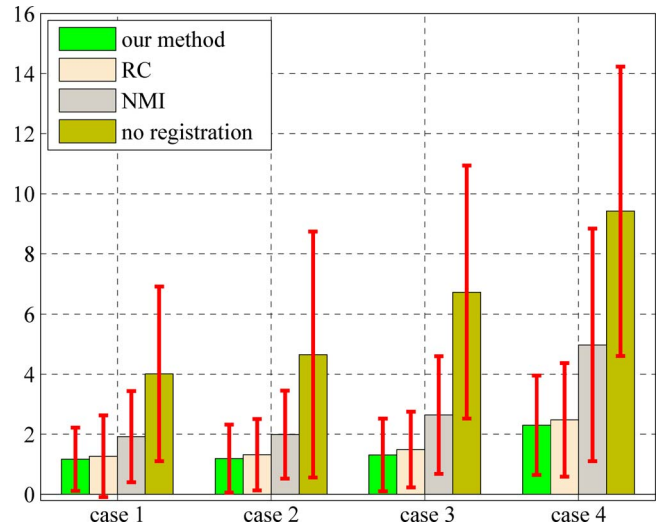


Fig. 9. 3-D volume registration accuracy. Mean and standard deviation of 3-D registration errors are depicted for each case before registration and after applying the proposed method, RC and NMI approaches. All values are in units of millimeters.

registration, to represent the initial misregistration, and after applying our algorithm, RC and NMI methods. As shown in Fig. 9, it is evident that the proposed method achieved lower 3-D registration errors than RC and NMI approaches. For the four patient cases, mean(standard deviation) 3-D errors ranged from 1.17(1.05) to 2.30(1.65) mm for the proposed method, 1.27(1.36) to 2.48(1.89) for the RC method, and 1.92(1.52) to 4.97(3.87) for the NMI approach. For the cumulative validation landmark set of 5494 landmark pairs, the mean(standard deviation) registration errors were 1.49(1.26), 1.64(1.43) and 2.88(2.20) for the proposed method, RC and NMI approaches, respectively.

IV. CONCLUSION

An entropic framework for nonrigid image registration is proposed in this paper. The experimental results on (MR-)T1–T1 and (MR-)T1–T2 registrations indicate the feasibility of the proposed approach and a much better performance compared to RC and NMI methods, not only in terms of registration accuracy in the presence of intensity and geometric distortion but also in terms of nonrigidly registering images of different protocols and modalities. The feasibility of the proposed algorithm is also demonstrated on clinically acquired 4-D CT image datasets. Future work will focus on extending our approach to non-rigid multimodal image registration using MR-CT and MR-PET images.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the associate editor for helpful and very insightful comments.

REFERENCES

- [1] A. A. Goshtasby, *2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications*. New York: Wiley, 2005.
- [2] I. Bankman, *Handbook of Medical Image Processing and Analysis*. New York: Academic, 2008.
- [3] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, *Medical Image Registration*. Boca Raton, FL: CRC Press, 2001.
- [4] H. Chui, L. Wing, R. Schultz, J. Duncan, and A. Rangarajan, "A unified non-rigid feature registration method for brain mapping," *Med. Image Anal.*, vol. 7, no. 2, pp. 112–130, 2003.
- [5] M. Yasein and P. Agathoklis, "A feature-based image registration technique for images of different scale," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2008, pp. 3558–3561.
- [6] N. Paragios, M. Rousson, and V. Ramesh, "Non-rigid registration using distance functions," *Comput. Vis. Image Underst.*, vol. 89, no. 2–3, pp. 142–165, 2003.
- [7] M. Audette, K. Siddiqi, F. Ferrie, and T. Peters, "An integrated range-sensing, segmentation and registration framework for the characterization of intra-surgical brain deformations in image-guided surgery," *Comput. Vis. Image Underst.*, vol. 89, no. 2–3, pp. 226–251, 2003.
- [8] A. Leow, P. Thompson, H. Protas, and S. Huang, "Brain warping with implicit representations," in *Proc. IEEE Int. Symp. Biomed. Imag.: Nano Macro*, 2004, pp. 603–606.
- [9] P. Viola and W. M. Wells, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 173–154, 1997.
- [10] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, 1997.
- [11] W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multimodal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, no. 1, pp. 35–51, 1996.
- [12] F. Maes, D. Vandermeulen, and P. Suetens, "Medical image registration using mutual information," *Proc. IEEE*, vol. 91, no. 10, pp. 1699–1722, Oct. 2003.
- [13] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–721, Aug. 1999.
- [14] C. Davatzikos, "Spatial transformation and registration of brain images using elastically deformable models," *Comput. Vis. Image Underst.*, vol. 66, no. 2, pp. 207–222, 1997.
- [15] J. Gee, M. Reivich, and R. Bajcsy, "Elastically deforming 3d atlas to match anatomical brain images," *J. Comput. Assist. Tomogr.*, vol. 17, no. 2, pp. 225–236, 1993.
- [16] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," in *Proc. Int. Conf. Vis. Biomed. Comput.*, 1993, pp. 267–276.
- [17] G. Christensen, R. Rabbitt, and M. Miller, "Deformable templates using large deformation kinematics," *IEEE Trans. Image Process.*, vol. 5, no. 10, pp. 1435–1447, Oct. 1996.
- [18] A. Andronache, M. Siebenthal, G. Szekely, and P. Catti, "Non-rigid registration of multi-modal images using both mutual information and cross-correlation," *Med. Image Anal.*, vol. 12, no. 1, pp. 3–15, 2008.
- [19] B. Likar and F. Pernus, "A hierarchical approach to elastic registration based on mutual information," *Image Vis. Comput.*, vol. 19, pp. 33–44, 2001.
- [20] J. P. W. Pluim, J. B. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 986–1004, Aug. 2003.
- [21] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognit.*, vol. 32, no. 1, pp. 71–86, 1999.
- [22] F. Wang and B. Vemuri, "Non-rigid multi-modal image registration using cross-cumulative residual entropy," *Int. J. Comput. Vis.*, vol. 74, no. 2, pp. 201–215, 2007.
- [23] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens, "Nonrigid image registration using conditional mutual information," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 19–29, Jan. 2007.
- [24] A. Myronenko and X. Song, "Image registration by minimization of residual complexity," in *Proc. Comput. Vis. Pattern Recognit.*, 2009, pp. 49–56.
- [25] A. Myronenko and X. Song, "Intensity-based image registration by minimizing residual complexity," *IEEE Trans. Med. Imag.*, vol. 29, no. 11, pp. 1882–1891, Nov. 2010.
- [26] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *J. Statist. Phys.*, vol. 52, pp. 479–487, 1988.
- [27] M. P. Wachowiak, R. Smolikova, G. D. Tourassi, and A. S. Elmaghraby, "Similarity metrics based on nonadditive entropies for 2D-3D multimodal biomedical image registration," *Proc. SPIE Med. Imag.*, vol. 5032, pp. 1090–1100, 2003.
- [28] G. Wahba, *Spline models for observational data*. Philadelphia, PA: SIAM, 1990.
- [29] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "PET-CT image registration in the chest using free-form deformations," *IEEE Trans. Med. Imag.*, vol. 22, no. 1, pp. 120–128, Jan. 2003.
- [30] A. Ben Hamza, "A nonextensive information-theoretic measure for image edge detection," *J. Electron. Imag.*, vol. 15, no. 1, pp. 130111–130118, 2006.
- [31] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York: Springer-Verlag, 2000, ch. 8–9.
- [32] J. Burbea and C. R. Rao, "On the convexity of some divergence measures based on entropy functions," *IEEE Trans. Inf. Theory*, vol. 28, no. 3, pp. 489–495, May 1982.
- [33] W. Mohamed, Y. Zhang, A. Ben Hamza, and N. Bouguila, "Stochastic optimization approach for entropic image alignment," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, Canada, 2008, pp. 2126–2130.
- [34] [Online]. Available: <http://www.bic.mni.mcgill.ca/brainweb/>
- [35] H. Luan, F. Qi, Z. Xue, L. Chen, and D. Shen, "Multimodality image registration by maximization of quantitative-qualitative measure of mutual information," *Pattern Recogn.*, vol. 41, pp. 285–298, 2008.
- [36] [Online]. Available: <http://www.dir-lab.com>
- [37] R. Castillo, E. Castillo, R. Guerra, V. Johnson, T. McPhail, A. Garg, and T. Guerrero, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," *Phys. Med. Biol.*, vol. 54, no. 7, pp. 1849–1870, 2009.

Authors' photographs and biographies not available at the time of publication.