Pattern Recognition Letters
journal homepage: www.elsevier.com

Intensity-Based 3D Local Image Registration

Huajun Song[a], Peihua Qiu[b,**]

[a] College of Information and Control Engineering, China University of Petroleum (East China)
[b] Department of Biostatistics, University of Florida

ABSTRACT

In many imaging applications, including image fusion and image sequence monitoring, we need to geometrically match up one image to another of a same scene, so that information from different images can be compared or combined properly. This is the image registration (IR) problem that has received much attention in recent years due to its broad applications. In the literature, early IR methods are for analyzing 2D images. Because of the rapid progress in image acquisition technologies, 3D images have become increasingly popular in magnetic resonance imaging (MRI) and other applications in recent years. Since the structure of a typical 3D image is substantially more complicated than that of a typical 2D image, 3D image registration is challenging. In this paper, we develop a new 3D image registration method using local smoothing statistical methods. By the flexibility of local smoothing, our method does not require any parametric form or other global regularity conditions on the related geometric transformation. It is shown that this method works well in practice.

In many imaging applications, we need to compare two or more images of a same object so that useful information of the images can be combined and the difference among the images can be detected. One such example is about the comparison of several MRI brain images of a brain tumor patient that were taken at different times, so that the tumor growth can be monitored closely. To make the comparison meaningful, the related images should be geometrically matched up first, because the relative positions between the imaging device and the image object would hardly be the same at different times when the images are taken. Image registration (IR) is for solving this and some other problems, including inter-subject anatomical comparisons, anatomical atlas creation, multi-modal image fusion, and longitudinal studies [1, 2]. It is a fundamental task in many imaging applications, including medical imaging [3], remote sensing [4], finger print or face recognition [5], image compression [6], video enhancement [7], and so forth.

In the literature, early IR methods are for analyzing 2D images. They can be roughly divided into two groups: feature-based and intensity-based methods. To register two images by a feature-based method, we first select two sets of features in the two images under consideration, and then find a geometrical transformation $\mathbf{T}$ to best match the two sets of features [8, 9, 10]. Commonly used features include landmarks or control points that can be selected manually or automatically by a computer [11], edge lines or curves that are often detected by gradient-based methods [12], regions, centroids or templates that are usually determined by ways of thresholding and segmentation [13], and degenerate pixels of the image intensity functions that are defined specifically for IR [10]. Because feature extraction is often a time-consuming and challenging task with much arbitrariness involved, recent IR research focuses more on the search of the transformation $\mathbf{T}$ based directly on the observed image intensities of the two images. Such methods are often called intensity-based image registration (IBIR) methods. Commonly used IBIR methods include those based on parametric transformation families [14], and more flexible ones using nonparametric transformations [15, 16, 17, 18, 19, 20]. IBIR methods are usually computing-intensive and not always robust to intensity biases in images.

In practice, most objects in our real life are 3D. Thanks to the rapid development of image acquisition techniques, 3D images have become more and more popular nowadays in certain applications (e.g., medical imaging). Consequently, 3D image registration has been an active research problem in recent 10-15 years. Generally speaking, 3D image registration is challenging

**Corresponding author;
e-mail: pqiu@ufl.edu (Peihua Qiu)

because 3D images have much more complicated structure than 2D images. For instance, edge locations are surfaces in 3D images which have more complicated structure than edge curves in 2D images. So far, some 2D image registration methods have been generalized to 3D cases. Some of them were incorporated in the software packages 3DSlicer (http://slicer.org/) and ANTS/SyN (http://www.picsl.upenn.edu/ANTS/). Some methods in these packages will be further studied in Section III of the paper. Another 3D IR method is based on the Iterative Closest Point (ICP) algorithm [21, 22]. This is a feature-based IR method based on the assumption that $\mathbf{T}$ has a parametric form. To use such a method, two sets of features need to be extracted first from the two related images. Then, the ICP algorithm alternates between estimating the parameters in $\mathbf{T}$ and searching for the optimal point-wise correspondence between the two images. Some other 3D image registration methods were discussed in papers such as [23, 24, 25, 26].

In this paper, we propose a new 3D IBIR method. This method is designed for cases when the geometric transformation $\mathbf{T}$ from one image to the other is non-rigid-body. Recall that a rigid-body transformation implies that the Euclidean distance between any two voxels in an image will not change after the geometric transformation. Non-rigid-body transformations, on the other hand, are much more flexible and they can cope with articulated objects or soft bodies that change shape over time. Typical applications of non-rigid-body IR methods can be found in biomedical studies, where soft bodies (e.g., human organs) are common [27, 28]. In the proposed new method, we actually do not impose any restrictive assumptions on the geometric transformation $\mathbf{T}$. To allow this great flexibility, we adopt the nonparametric local smoothing approach that was first discussed in [20] in 2D cases. By such an approach, estimation of $\mathbf{T}$ at a given voxel depends only on nearby image intensities of the two 3D images. This *local estimation* nature makes it possible to not impose restrictive assumptions on $\mathbf{T}$. Our proposed method will be described in detail in Section II. Some numerical results are presented in Section III. Several remarks conclude the article in Section IV.

1. Proposed 3D IBIR Method

1.1. Description of the 3D IBIR problem

The 3D image registration problem can be described as follows. Let R and M be two 3D images to register. In the literature, R is often called the reference image, and M the moved image. Their true image intensity functions are denoted as $R(x, y, z)$ and $M(x, y, z)$. It is assumed that they have the following relationship:

$$\begin{aligned} &M(T_1(x, y, z), T_2(x, y, z), T_3(x, y, z)) \\ &\quad = R(x, y, z), \qquad \text{for } (x, y, z) \in \Omega, \end{aligned} \tag{1}$$

where Ω is the design space of the image R, and $\mathbf{T}(x, y, z) = (T_1(x, y, z), T_2(x, y, z), T_3(x, y, z))$ is an unknown geometric transformation to estimate. IBIR methods try to estimate $\mathbf{T}(x, y, z)$ from the observed image intensities of the two images following the statistical model:

$$\begin{aligned} Z_M(x_i, y_j, z_k) &= M(x_i, y_j, z_k) + \varepsilon_M(x_i, y_j, z_k), \\ Z_R(x_i, y_j, z_k) &= R(x_i, y_j, z_k) + \varepsilon_R(x_i, y_j, z_k), \\ &i, j, k = 1, 2, \ldots, n, \end{aligned} \tag{2}$$

where $\{(x_i, y_j, z_k)\}$ are voxel locations, and $\varepsilon_M(x_i, y_j, z_k)$ and $\varepsilon_R(x_i, y_j, z_k)$ are i.i.d. random errors in the two images with mean 0 and unknown variances σ_M^2 and σ_R^2, respectively. Nonparametric IBIR methods try to estimate $\mathbf{T}(x, y, z)$ from the observed image intensities, without imposing any parametric form on $\mathbf{T}(x, y, z)$. In (2), we assume that the two observed images contain pointwise noise only, for simplicity. In practice, they may also contain spatial blur and other degradation [29, 30], and the noise may be non-additive (e.g., follows the Rician noise model) so that a biase-correction procedure should be applied in advance [31]. At a given voxel $(x, y, z) \in \Omega$, we can write

$$\begin{pmatrix} T_1(x, y, z) \\ T_2(x, y, z) \\ T_3(x, y, z) \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} b(x, y, z) \\ c(x, y, z) \\ d(x, y, z) \end{pmatrix},$$

where $b(x, y, z) = T_1(x, y, z) - x$, $c(x, y, z) = T_2(x, y, z) - y$ and $d(x, y, z) = T_3(x, y, z) - z$. Therefore, estimation of $\mathbf{T}(x, y, z)$ is equivalent to estimation of $(b(x, y, z), c(x, y, z), d(x, y, z))$. After the estimators of $(b(x, y, z), c(x, y, z), d(x, y, z))$, denoted as $(\widehat{b}(x, y, z), \widehat{c}(x, y, z), \widehat{d}(x, y, z))$, are obtained, the estimator of $\mathbf{T}(x, y, z)$ can be written as

$$\widehat{\mathbf{T}}(x, y, z) = (x, y, z) + \left(\widehat{b}(x, y, z), \widehat{c}(x, y, z), \widehat{d}(x, y, z)\right) \tag{3}$$

1.2. 3D IR and Non-Degenerate Voxels

It has been pointed out in the literature that the 2D image registration problem is ill-posed at degenerate pixels [20, 17], in the sense that the geometric transformation is not well defined around the degenerate pixels. In 3D cases, we have the similar issue, which will be discussed in this part. At a given voxel $(x, y, z) \in \Omega$, if the magnitude of the geometric transformation $\mathbf{T}(x, y, z)$ is small (i.e., $\mathbf{T}(x, y, z) - (x, y, z) = (b(x, y, z), c(x, y, z), d(x, y, z))$ is small) and M has the first-order partial derivatives at (x, y, z), then by the Taylor's expansion, we have

$$\begin{aligned} &M(T_1(x, y, z), T_2(x, y, z), T_3(x, y, z)) \\ &= M(x, y, z) + M'_x(x, y, z)b(x, y, z) \\ &+ M'_y(x, y, z)c(x, y, z) + M'_z(x, y, z)d(x, y, z) \\ &+ o\left(\|\mathbf{T}(x, y, z) - (x, y, z)\|\right), \end{aligned} \tag{4}$$

where $\|\cdot\|$ is the Euclidean norm. By (1) and (4), it can be seen that $R(x, y, z) = M(T_1(x, y, z), T_2(x, y, z), T_3(x, y, z))$ can be well approximated by $M(x, y, z) + M'_x(x, y, z)b(x, y, z) + M'_y(x, y, z)c(x, y, z) + M'_z(x, y, z)d(x, y, z)$ in such cases. Therefore, $(b(x, y, z), c(x, y, z), d(x, y, z))$ can be chosen such that the approximation error

$$\begin{aligned} &R(x, y, z) - \big[M(x, y, z) + M'_x(x, y, z)b(x, y, z) \\ &+ M'_y(x, y, z)c(x, y, z) + M'_z(x, y, z)d(x, y, z)\big] \end{aligned}$$

is as small as possible. In reality, however, $R(x, y, z)$, $M(x, y, z)$, $M'_x(x, y, z)$, $M'_y(x, y, z)$ and $M'_z(x, y, z)$ are all unobservable.

What observed are the image intensities $\{Z_M(x_i, y_j, z_k)\}$ and $\{Z_R(x_i, y_j, z_k)\}$ defined in (2), which contain random noise. To smooth out noise while estimate $(b(x, y, z), c(x, y, z), d(x, y, z))$, we adopt the idea of local linear kernel (LLK) estimation in statistical nonparametric regression [32] as follows. First, the quantities $M'_x(x, y, z)$, $M'_y(x, y, z)$ and $M'_z(x, y, z)$ can be estimated by their conventional LLK estimators $\widehat{M}'_x(x_i, y_j, z_k)$, $\widehat{M}'_y(x_i, y_j, z_k)$ and $\widehat{M}'_z(x_i, y_j, z_k)$, respectively. Then, in a spherical neighborhood of (x, y, z) with radius h_n, denoted as $O(x, y, z; h_n)$, $b(x, y, z)$, $c(x, y, z)$ and $d(x, y, z)$ can be estimated by the solution of the following minimization problem:

$$\begin{aligned}\min_{b(x,y,z),c(x,y,z),d(x,y,z)} \sum_{i,j=1}^{n} \Big[& Z_M(x_i, y_j, z_k) \\ & - Z_R(x_i, y_j, z_k) + \widehat{M}'_x(x, y, z)b(x, y, z) \\ & + \widehat{M}'_y(x, y, z)c(x, y, z) + \widehat{M}'_z(x, y, z)d(x, y, z)\Big]^2 K_{h_n},\end{aligned} \tag{5}$$

where $K_{h_n} = K((x_i - x)/h_n, (y_j - y)/h_n, (z_k - z)/h_n)$, and K is a trivariate density kernel function with unit circular support. The minimization problem (5) searches for estimators of $b(x, y, z)$, $c(x, y, z)$ and $d(x, y, z)$ such that the weighted sum of squares of the approximation errors reaches the minimum, and the weights are determined by the kernel function K. In the statistical literature, K is often chosen to be the Epanechnikov kernel function $C_K(1 - x^2)(1 - y^2)(1 - z^2)I(x^2 + y^2 + z^2 \le 1)$, where C_K is a normalization constant and $I(\cdot)$ is an indicator function. By (5), if a voxel $(x_i, y_j, z_k) \in O(x, y, z; h_n)$ is farther away from the given voxel (x, y, z), then the corresponding approximation error at (x_i, y_j, z_k) would receive a less weight, which is intuitively reasonable, because the observed image intensity at (x_i, y_j, z_k) would provide less information about $(b(x, y, z), c(x, y, z), d(x, y, z))$ in such cases, compared to the observed image intensities at voxels closer to (x, y, z). It is not difficult to check that the problem (5) has the solution given in (6). In (6), for $s_1, s_2 = x, y, z$,

$$K_{s_1 s_2} = \sum_{i,j,k=1}^{n} \widehat{M}'_{s_1}(x_i, y_j, z_k)\widehat{M}'_{s_2}(x_i, y_j, z_k)K_{h_n},$$

$$K^*_{s_1} = \sum_{i,j,k=1}^{n} \left[Z_R(x_i, y_j, z_k) - Z_M(x_i, y_j, z_k)\right] \widehat{M}'_{s_1}(x_i, y_j, z_k)K_{h_n},$$

and the conventional LLK estimators of $M'_x(x, y, z)$, $M'_y(x, y, z)$ and $M'_z(x, y, z)$ are: for $s = x, y, z$,

$$\widehat{M}'_s = \frac{\sum_{ijk}(s_i - s)Z_M(x_i, y_j, z_k)K_h(x_i - x, y_j - y, z_k - z)}{\sum_{ijk}(s_i - s)^2 K_h(x_i - x, y_j - y, z_k - z)}.$$

From the above description, we know that formula (6) is obtained in the ideal situation when (i) $\|\mathbf{T}(x, y, z) - (x, y, z)\|$ is small such that the first-order approximation to $M(T_1(x, y, z), T_2(x, y, z), T_3(x, y, z))$ in (4) is good, (ii) M has the first-order partial derivatives at (x, y, z), and (iii) the denominator on the right-hand-side of (6) is not zero. The above conditions (i) and (ii) imply that the estimator defined by (3) and (6) may not estimate $\mathbf{T}(x, y, z)$ well at places where the transformation is relatively large or where M is not smooth (e.g., edge locations of M). Condition (iii) implies that the estimator is not well defined at places where the following equation holds:

$$\begin{aligned}K_{xx}K_{yy}K_{zz} + K_{xy}K_{yz}K_{xz} + K_{xy}K_{yz}K_{xz} \\ -K_{yy}K_{xz}^2 - K_{zz}K_{xy}^2 - K_{xx}K_{yz}^2 = 0.\end{aligned} \tag{7}$$

Mathematically, it can be proved that: (1) $\widehat{M}'_x(x_i, y_j, z_k)$, $\widehat{M}'_y(x_i, y_j, z_k)$ and $\widehat{M}'_z(x_i, y_j, z_k)$ would converge to $M'_x(x, y, z)$, $M'_y(x, y, z)$ and $M'_z(x, y, z)$, respectively, in regions where M has continuous first order derivatives, when n gets larger and larger, and (2) if $\widehat{M}'_x(x_i, y_j, z_k)$, $\widehat{M}'_y(x_i, y_j, z_k)$ and $\widehat{M}'_z(x_i, y_j, z_k)$ are replaced by $M'_x(x, y, z)$, $M'_y(x, y, z)$ and $M'_z(x, y, z)$ in (7), then M satisfies equation (7) in the neighborhood $O(x, y, z; h_n)$ if and only if there is a continuously differentiable univariate function ψ and a constant ρ such that

$$\begin{aligned}M(x', y', z') = \psi(\rho x' + y'), \\ \text{for any } (x', y', z') \in O(x, y, z; h_n).\end{aligned} \tag{8}$$

Intuitively, if M satisfies (8) in $O(x, y, z; h_n)$, then its intensity levels are the same on the line segment $\rho x' + y' = \rho_0$ in $O(x, y, z; h_n)$, for any appropriate constant ρ_0 such that the line segment is contained in $O(x, y, z; h_n)$. In such cases, the trivariate function M is degenerate locally in $O(x, y, z; h_n)$, and it is impossible to uniquely determine $\mathbf{T}(x, y, z)$ because any small move along the line direction would not change the value of $M(x', y', z')$ for any $(x', y', z') \in O(x, y, z; h_n)$. In this paper, a voxel $(x, y, z) \in \Omega$ is called a *local degenerate voxel* of M if M has partial derivatives at (x, y, z) and there exists a neighborhood $O(x, y, z; h_n)$ such that equation (8) holds. Other voxels at which M has partial derivatives are called *local non-degenerate voxels*. Similarly, we can define local degenerate voxels and local non-degenerate voxels for the reference image R. Therefore, around local degenerate voxels, the image registration problem is actually not well defined in the sense that the geometrical transformation $\mathbf{T}(x, y, z)$ cannot be uniquely determined.

1.3. Proposed Image Registration Procedure

From the description in Section 1.2, it can be seen that the geometrical transformation $\mathbf{T}(x, y, z)$ can be properly estimated only around local non-degenerate voxels or places where the image intensity function of the reference image R is not smooth. Based on that result, we propose an IBIR procedure consisting of four major steps, as described below.

Step 1 Detect edge voxels for the observed reference image Z_R using an edge detector. See Chapter 6 in [32] for a discussion about existing edge detectors.

Step 2 At a given voxel (x, y, z) in R, consider its circular neighborhood with radius r_1, denoted as $O(x, y, z; r_1)$. If the number of detected edge voxels in $O(x, y, z; r_1)$ is smaller than $[nr_1]$, where $[s]$ denotes the integer part of s, then (x, y, z) is regarded as a continuity voxel of R. In such a case, if the denominator on the right-hand-side of equation (6) (after M is replaced by R) is larger than or equal to a pre-specified threshold value μ_n, then (x, y, z) is regarded as a local non-degenerate voxel of R; otherwise, it is regarded as a local degenerate voxel.

$$\begin{pmatrix}\widehat{b}(x,y,z)\\ \widehat{c}(x,y,z)\\ d(x,y,z)\end{pmatrix} = \frac{\begin{bmatrix} K_{yy}K_{zz}-K_{yz}^2 & -(K_{xy}K_{zz}-K_{yz}K_{xz}) & K_{xy}K_{yz}-K_{yy}K_{xy}\\ -(K_{xy}K_{yy}-K_{xy}K_{yz}) & K_{xx}K_{zz}-K_{xz}^2 & -(K_{xx}K_{yz}-K_{xy}K_{xz})\\ K_{xy}K_{yz}-K_{xy}K_{yy} & -(K_{xx}K_{yz}-K_{xz}K_{xy}) & K_{xx}K_{yy}-K_{xy}^2\end{bmatrix}}{(K_{xx}K_{yy}K_{zz}+K_{xy}K_{yz}K_{xz}+K_{xy}K_{yz}K_{xz}-K_{yy}K_{xz}^2-K_{zz}K_{xy}^2-K_{xx}K_{yz}^2)}\begin{pmatrix}K_1^*\\ K_2^*\\ K_3^*\end{pmatrix} \tag{6}$$

Step 3 Let D be the set of all local non-degenerate voxels of R or voxels whose circular neighborhoods with radius r_1 contain at least $[nr_1]$ detected edge voxels. Then, for any $(x,y,z) \in D$, $\mathbf{T}(x,y,z)$ is computed by the following algorithm. For any voxel $x',y',z' \in O(x,y,z;r_1)$ of the moved image, consider its circular neighborhood $O(x',y',z';r_2)$, where r_2 is a radius that could be different from r_1. Compute the mean squared difference (MSD)

$$\mathrm{MSD}((x',y',z');(x,y,z)) = \frac{1}{\tilde{N}} \sum_{(x'+s,y'+t,z'+u)\in O(x',y',z';r_2)} [Z_M(x'+s,y'+t,z'+u) - Z_R(x+s,y+t,z+u)]^2,$$

where $\tilde{N}$ is the number of voxels in $O(x',y',z';r_2)$. Then, $\widehat{T}(x,y,z)$ is defined to be the minimizer of

$$\min_{(x',y',z')\in O(x,y,z;r_1)} \mathrm{MSD}((x',y',z');(x,y,z)).$$

See Figure 1 for a demonstration.

Step 4 If (x,y,z) is a local degenerate point of R, then $\widehat{\mathbf{T}}(x,y,z)$ is defined as follows. First, find a voxel in D that is closest to (x,y,z), which is denoted as $(x^{(1)},y^{(1)},z^{(1)})$. Then, let

$$\widehat{\mathbf{T}}^*(x,y,z) = (x,y,z) + (\widehat{b}(x^{(1)},y^{(1)},z^{(1)}), \widehat{c}(x^{(1)},y^{(1)},z^{(1)}), d(x^{(1)},y^{(1)},z^{(1)}))$$

Define $\widehat{\mathbf{T}}(x,y,z) = \widehat{T}^*(x,y,z)$ if

$$\mathrm{MSD}(\widehat{\mathbf{T}}^*(x,y,z);(x,y,z)) \le \mathrm{MSD}((x,y,z);(x,y,z)).$$

Otherwise, define $\widehat{\mathbf{T}}(x,y,z) = (x,y,z)$.

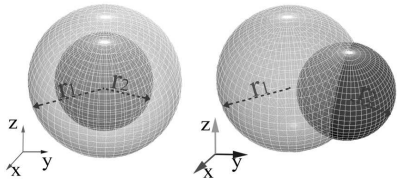

Fig. 1. A demonstration of Step 3 of the proposed 3D image registration algorithm.

Note that the above IBIR procedure can only properly handle interior voxels in Ω whose Euclidean distances from the border of Ω are at least $r_1 + r_2$. For a boundary voxel of Ω whose Euclidean distance from the border of Ω is smaller than $r_1 + r_2$, we define the geometrical transformation at that voxel to be the same as that at the interior voxel who is closest to the boundary voxel.

At the end of this section, we make several remarks about the above image registration procedure. First, at local non-degenerate voxels of R, instead of using formulas (3) and (6), we use the searching algorithm described in Step 3 for computing $\widehat{\mathbf{T}}(x,y,z)$. That is because formula (6) is valid only when $\|\mathbf{T}(x,y,z) - (x,y,z)\|$ is small. From our numerical studies, the estimator by the searching algorithm would perform better than the one by (3) and (6) for most realistic $\mathbf{T}(x,y,z)$ functions. Second, from the discussion in Section 1.2, it seems that, to define $\mathbf{T}(x,y,z)$ properly, (x,y,z) should be a non-degenerate voxel of M, instead of R. Again, this discussion is based on the assumption that $\|\mathbf{T}(x,y,z) - (x,y,z)\|$ is small so that the Taylor's expansion is valid. In practice, it is more reasonable to require (x,y,z) to be a non-degenerate voxel of R. Third, in Step 3, MSD is used as the matching criterion. Actually, other criteria, including the cross correlation and entropy of image difference (cf., their definitions in Section 3), can also be considered here. Fourth, computation involved in the searching algorithm described in Step 3 is actually not very extensive because voxels in D represent only about 10% of all voxels for a typical image. Fifth, performance of the proposed image registration procedure depends on the three parameters h_n, r_1 and r_2. Based on many numerical experiments, we found that the results are reasonably good if we choose $h_n/n \in [0.1, 0.2]$, $r_1 = 2r_2$, and $r_2/n \in [0.1, 0.2]$.

2. Numerical Study

We have performed numerical studies with more than 10 different pairs of 3D test images. The evaluation is made in comparison with certain methods included in the well-known software packages 3D Slicer and ANTS/SyN. The existing methods from 3D Slicer considered here include the free-form deformation method based on B-splines, denoted as B-Spline, the IR method using the affine invariant geometric transformation, denoted as Affine, and the method using the rigid-body geometric transformation, denoted as Rigid. The existing methods from ANTS/SyN include the rigid+affine+deformable syn, denoted as RAD-SyN, and the rigid+affine+deformable B-spline syn, denoted as RADB-SyN. These methods represent different state-of-the-art 3D IBIR methods in the literature. The version of the 3D Slicer software used here is its latest release 4.5.0, that is available at http://www.slicer.org/. The version of ANTS/SyN package used here is its latest release 2.1.0, that is available at http://www.picsl.upenn.edu/ANTS/. Our proposed method is denoted as NEW. For comparison purposes, we also include results when no image registration is performed. This case is denoted as No-Registration.

To evaluate the numerical performance of all related methods, we use three popular measures, including the root residual mean squares (RRMS), the correlation coefficient (CC), and the entropy of image difference (EID). RRMS is the conventional and most widely used measure. When evaluating an estimator

$\widehat{\mathbf{T}}(x,y,z)$ of the geometrical transformation $\mathbf{T}(x,y,z)$, it is defined to be

$$\text{RRMS} = \left\{\frac{1}{n^3}\sum_{i,j,k=1}^{n}\left[Z_R(x_i,y_j,z_k) - Z_M(\widehat{\mathbf{T}}(x_i,y_j,z_k))\right]^2\right\}^{\frac{1}{2}}.$$

Basically, RRMS is the Euclidean distance between $\{Z_R(x_i,y_j,z_k)\}$ and $\{Z_M(\widehat{\mathbf{T}}(x_i,y_j,z_k))\}$. Therefore, if its value is smaller, then the registration is regarded better. The CC measure is defined to be the Pearson's sample correlation coefficient of the bivariate data $\{Z_R(x_i,y_j,z_k), Z_M(\widehat{\mathbf{T}}(x_i,y_j,z_k))\}$. Intuitively, if the estimator $\widehat{\mathbf{T}}(x_i,y_j,z_k)$ is good, then $Z_M(\widehat{\mathbf{T}}(x_i,y_j,z_k))$ would be close to $Z_R(x_i,y_j,z_k)$. Consequently, the CC measure would be close to its maximum value 1. The EID measure became popular recently. It is defined by

$$\text{EID} = -\sum_{d\in D} p(d)\log p(d),$$

where $D = \{Z_R(x_i,y_j,z_k) - Z_M(\widehat{\mathbf{T}}(x_i,y_j,z_k)),\ i,j,k = 1,2,\ldots,n\}$. So, EID is basically the negative entropy of D. Intuitively, if $\widehat{\mathbf{T}}(x_i,y_j,z_k)$ is a good estimator of $\mathbf{T}(x_i,y_j,z_k)$, then the randomness in the elements of D should be large. Consequently, EID should be small because the entropy $\sum_{d\in D} p(d)\log p(d)$ is a good measure of the randomness of D. Therefore, by this measure, the registration is better if the EID value is smaller.

For the methods B-Spline, Affine, and Rigid in the software package 3D Slicer and the methods RAD-SyN and RADB-SyN in the software package ANTS/SyN, we use their default settings in the software. The proposed method NEW has three parameters h_n, r_1 and r_2. Based on the practical guidelines given in Section 1.3, we choose $h_n = 10$, $r_1 = 20$, and $r_2 = 10$.

In the first example, the 3D reference image is downloaded from the web page http://www.slicer.org/slicerWiki/images/3/31/CT-chest.nrrd, which is a human chest image with $128\times128\times69$ voxels. Its 10 slices at $z = 10, 15, 20, 25, 30, 35, 40, 45, 50, 55$ are shown in the first row of Figure 2. First, we apply the following transformation to all slices along the $z-$axis to obtain the moved image: for $y = 32, 33, \ldots, 75$, move the voxel at (x,y,z) to $(x + 5\sin(4\pi y/180), y, z)$. By this transformation, voxels located between 32nd and 75th rows along the $y-$axis move along the $x-$axis by the amount of $5\sin(4\pi y/180)$, for all slices along the $z-$axis. Then, we apply rotate, translation, and scale transformations on the above image by randomly generating values of the transformation parameters as follows: each of (α,β,γ) was generated from the Uniform distribution on $[3,3]$, each of $(\Delta x,\Delta y,\Delta z)$ was generated from the Uniform distribution on $\{3,2,1,0,1,2,3\}$, and the scale factor s was generated from the Uniform distribution on $[0.9,1]$. The corresponding parameter values are -2.54, 1.27, -1.60, -1, -3, -1, 0.98 for $(\alpha,\beta,\gamma,\Delta x,\Delta y,\Delta z,s)$. The 10 slices at $z = 10, 15, 20, 25, 30, 35, 40, 45, 50, 55$ of the moved image are shown in the second row of Figure 2. From the images, we can see that the chest shape has been changed in the moved image and it becomes more asymmetric.

We then apply the proposed IBIR method NEW to the two images, along with the methods B-Spline, Affine, Rigid, No-Registration, RAD-SyN, and RADB-SyN. The corresponding slices of the restored reference image by NEW, defined as $M(\widehat{\mathbf{T}}(x,y,z))$, are shown in the last row of Figure 2. It can be seen that the restored reference image is very close to the original reference image.

To compare different methods, the values of the performance evaluation criteria RRMS, CC and EID of the methods No-Registration, B-Spline, Affine, Rigid, RADB-SyN, RAD-SyN and NEW are shown in Table 1. From the table, it can be seen that (i) NEW is uniformly better than No-Registration, B-Spline, Affine, and RAD-SyN in all three performance criteria, and (ii) NEW is better than RADB-SyN in terms of RRMS and CC in quite large margins and slightly worse in terms of EID. The 20th and 40th slices along the $z-$axis of the residual images, defined as $R(x,y,z) - M(\widehat{\mathbf{T}}(x,y,z))$, of the related methods are shown in Figure 3. The conclusions from these results are consistent with those from Table 1.

Table 1. Performance measures of the seven IR methods in the chest image local distortion example.

	No-Registration	B-Spline	Affine	Rigid	RADB-SyN	RAD-SyN	NEW
RRMS	40.504	28.534	37.517	31.347	28.243	27.695	**8.739**
CC	0.692	0.852	0.742	0.807	0.839	0.843	**0.982**
EID	4.855	3.854	4.167	3.852	3.205	3.459	**3.445**

Next, we consider a real IR problem in which an AIDS patient took 3D MRI brain images before and after a medical treatment. The image before the treatment is used as a reference image, and the one after the treatment is used as the moved image. Each image has $128\times128\times88$ voxels. The 1st, 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and 88th slices along the $z-$axis of the reference image are shown in the first row of Figure 4, and the corresponding slices of the moved image are shown in the second row of Figure 4. It can be seen that the slices of the two images look quite different, implying that the relative position between the imaging device and the patient's head is different when taking the two images. The corresponding slices of the restored image by the proposed method NEW are shown in the last row of Figure 4. It can be seen that they look much more similar to the ones in the first row, compared to the ones in the second row. This example shows that the method NEW is effective in registering the two real 3D MRI images in this example. The performance measures of the 7 competing methods are presented in Table 2, from which it can be seen that the proposed method NEW is the best in terms of all performance measures in this example. The 30th and 60th slices of the residual images of the 7 methods are shown in Figure 5. Again, the ones of the

Table 2. Performance measures of the seven IBIR methods in the AIDS patient example.

	No-Registration	B-Spline	Affine	Rigid	RADB-SyN	RAD-SyN	NEW
RRMS	25.347	16.499	17.598	17.739	16.263	16.221	**13.812**
CC	0.745	0.912	0.899	0.896	0.914	0.915	**0.920**
EID	6.133	5.606	5.646	5.653	5.582	5.585	**5.377**

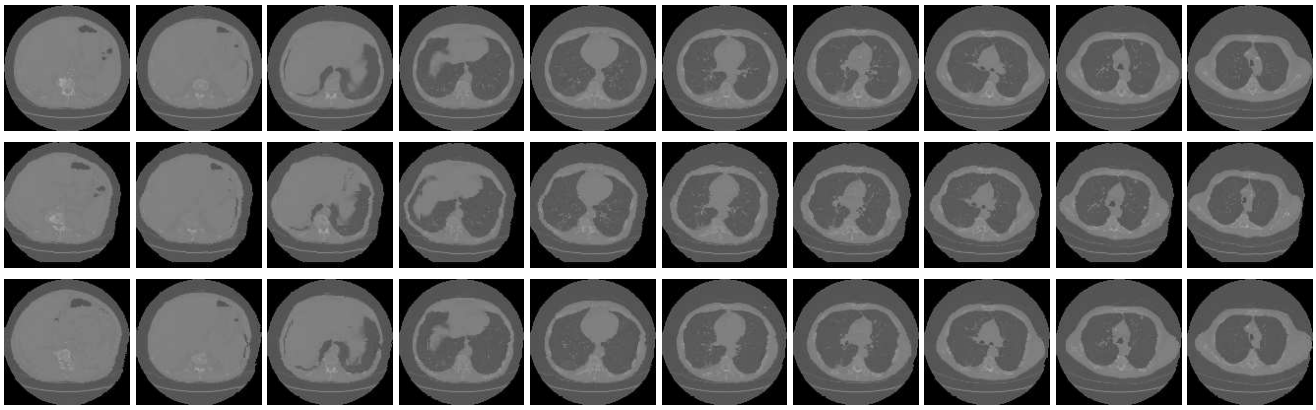

Fig. 2. Chest image local distortion example. Images in the first row are the 10 slices at $z = 10, 15, 20, 25, 30, 35, 40, 45, 50, 55$ of the reference image. The ones in the second row are the corresponding slices of the moved image, and the ones in the last row are the corresponding slices of the restored reference image by NEW, defined as $M(\widehat{\mathbf{T}}(x, y, z))$.

Fig. 3. Chest image local distortion example. (a)-(g) The 20th (top) and 40th (bottom) slices of the residual images of the methods No-Registration, B-Spline, Affine, Rigid, RADB-SyN, RAD-SyN, and NEW.

method NEW seem to contain less pattern compared the ones of the other 6 methods.

3. Conclusion

In the previous sections, we have described our proposed method for intensity-based 3D image registration. The proposed method is flexible in the sense that it does not impose any parametric form on the related geometric transformation. Numerical examples show that it works well in different applications. There are still some issues about the proposed method to be properly addressed in our future research. For instance, our proposed method NEW has three parameters to choose before the method can be actually used. Some data-driven parameter selection procedures might be helpful. The current version of NEW does not require any restrictive assumptions on the geometric transformation $\mathbf{T}(x, y, z)$, and it even allows $\mathbf{T}(x, y, z)$ to be discontinuous. However, in most applications, it is reasonable to assume that $\mathbf{T}(x, y, z)$ is a continuous transformation. It needs much future research effort to suggest proper modifications of NEW so that such reasonable regularity assumptions can be accommodated while the flexibility of the method can still be kept.

Acknowledgment

The authors acknowledge grant support from the National Natural Science Foundation of China (No.61305012) and grant support from US National Science Foundation (DMS-1405698).

References

[1] B. Zitova, J. Flusser, Image registration methods: a survey, Image and Vision Computing 21 (2003) 977–1000.
[2] J. Modersitzki, Fair: Flexible Algorithms for Image Registration, SIAM: Philadelphia, 2008.
[3] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M. C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration, Neuroimage 46 (2009) 786–802.
[4] H. Li, B. S. Manjunath, S. K. Mitra, A contour-based approach to multisensor image registration, IEEE Transactions on Image Processing 4 (1995) 320–334.
[5] L. Liu, T. Jiang, J. Yang, C. Zhu, Fingerprint registration by maximization of mutual information, Image Processing IEEE Transactions on 15 (2006) 1100–1110.
[6] F. Dufaux, J. Konrad, Efficient, robust, and fast global motion estimation for video coding, IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society 9 (2000) 497–501.
[7] M. Irani, S. Peleg, Motion analysis for image enhancement: Resolution, occlusion, and transparency, Journal of Visual Communication and Image Representation 4 (1993) 324–335.
[8] R. J. Althof, M. G. Wind, D. J. Rd, A rapid and automatic image registration algorithm with subpixel accuracy, IEEE Transactions on Medical Imaging 16 (1997) 308–16.
[9] M. H. Davis, A. Khotanzad, D. P. Flamig, S. E. Harms, A physics-based coordinate transformation for 3-d image matching, IEEE Transactions on Medical Imaging 16 (1997) 317–328.
[10] P. Qiu, C. Xing, Feature based image registration using non-degenerate pixels, Signal Processing 93 (2013) 706–720.

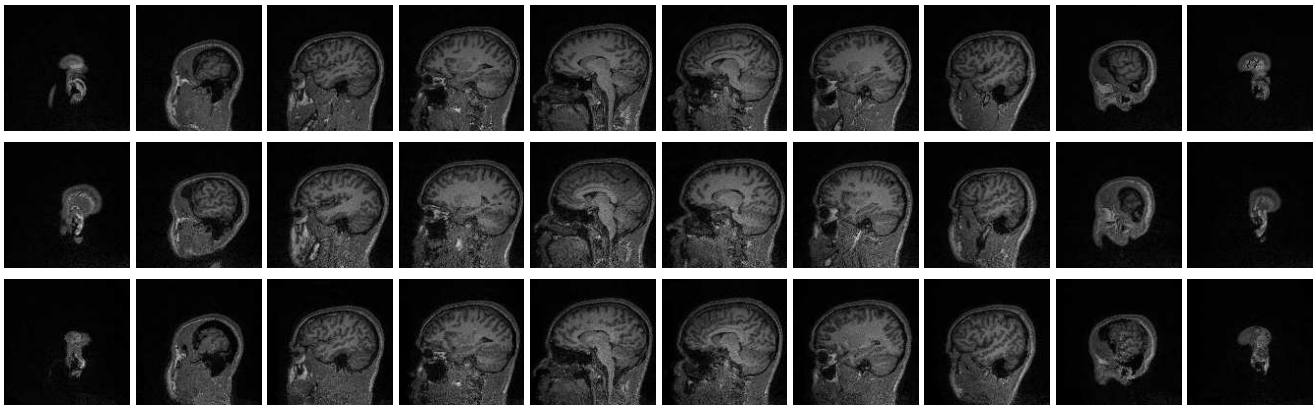

Fig. 4. 3D MRI images of an AIDS patient example. Images in the first row are the 10 slices at $z = 1, 10, 20, 30, 40, 50, 60, 70, 80, 88$ of the reference image. The ones in the second row are the corresponding slices of the moved image, and the ones in the last row are the corresponding slices of the restored reference image by NEW, defined as $M(\widehat{\mathbf{T}}(x, y, z))$.

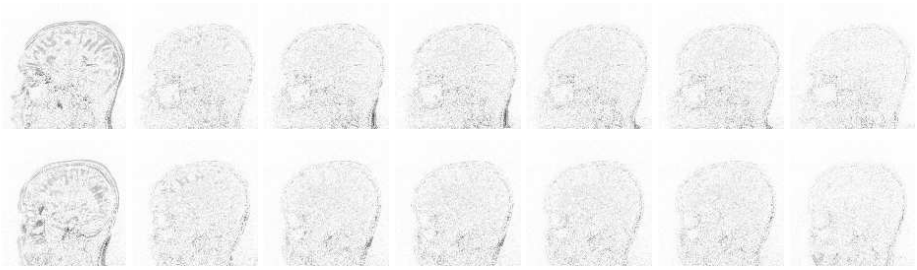

Fig. 5. 3D MRI images of an AIDS patient example. (a)-(g) The 30th (top) and 60th (bottom) slices of the residual images of the methods No-Registration, B-Spline, Affine, Rigid, RADB-SyN, RAD-SyN, and NEW.

[11] G. Wu, F. Qi, D. Shen, Learning-based deformable registration of mr brain images, IEEE Transactions on Medical Imaging 25 (2006) 1145–1157.

[12] J. W. Hsieh, H. Y. M. Liao, K. C. Fan, M. T. Ko, Y. P. Hung, Image registration using a new edge-based approach, Computer Vision and Image Understanding 67 (1997) 112–130.

[13] N. Saeed, Magnetic resonance image segmentation using pattern recognition, and applied to image registration and quantitation, NMR in Biomedicine 11 (1998) 157–167.

[14] E. R. Denton, L. I. Sonoda, D. Rueckert, S. C. Rankin, C. Hayes, M. O. Leach, D. L. Hill, D. J. Hawkes, Comparison and evaluation of rigid, affine, and nonrigid registration of breast mr images, Journal of Computer Assisted Tomography 23 (1999) 800–805.

[15] B. Avants, C. Epstein, M. Grossman, J. Gee, Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain, Medical Image Analysis 12 (2008) 26–41.

[16] W. Pan, K. Qin, Y. Chen, An adaptable-multilayer fractional fourier transform approach for image registration., Pattern Analysis and Machine Intelligence IEEE Transactions on 31 (2009) 400–14.

[17] P. Qiu, C. Xing, On nonparametric image registration, Technometrics 55 (2013) 174–188.

[18] A. Rajwade, A. Banerjee, A. Rangarajan, Probability density estimation using isocontours and isosurfaces: applications to information-theoretic image registration, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (2008) 475–91.

[19] N. Tustison, B. Avants, J. Gee, Directly manipulated free-form deformation image registration, IEEE Transactions on Image Processing 18 (2009) 624–635.

[20] C. Xing, P. Qiu, Intensity based image registration by nonparametric local smoothing, IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (2011) 2081–2092.

[21] P. J. Besl, N. D. Mckay, A method for registration of 3-D shapes, 1992.

[22] J. Yang, H. Li, Y. Jia, Go-icp: Solving 3d registration efficiently and globally optimally, in: IEEE International Conference on Computer Vision, 2013, pp. 1457–1464.

[23] J. Ashburner, A fast diffeomorphic image registration algorithm, NeuroImage 38 (2007) 95–113.

[24] P. Markelja, D. Tomazevica, B. Likara, F. Pernusa, A review of 3d/2d registration methods for image-guided interventions, Medical Image Analysis 16 (2012) 642–661.

[25] H. Song, P. Qiu, Parametric intensity-based 3d image registration method for magnetic resonance imaging, Signal, Image and Video Processing (2016) in press.

[26] T. Vercauteren, X. Pennec, A. Perchant, N. Ayache, Non-parametric diffeomorphic image registration with the demons algorithm, International Conference on Medical Image Computing and Computer-Assisted Intervention (2007) 319–326.

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

[28] W. R. Crum, T. Hartkens, D. L. Hill, Non-rigid image registration: theory and practice, British Journal of Radiology 77 Spec No 2 (2004) S140.

[29] Hall, Peter, Qiu, Peihua, Blind deconvolution and deblurring in image analysis, Statistica Sinica 17 (2007) 1483–1509.

[30] P. Qiu, A nonparametric procedure for blind image deblurring, Computational Statistics and Data Analysis 52 (2008) 4828–4841.

[31] P. Mukherjee, P. Qiu, Efficient bias correction for mri image denoising, Statistics in Medicine 32 (2013) 2079–2096.

[32] P. Qiu, Image Processing and Jump Regression Analysis, New York: John Wiley and Sons, 2005.