

Research Article

Effective Volumetric Feature Modeling and Coarse Correspondence via Improved 3DSIFT and Spectral Matching

Peizhi Chen¹ and Xin Li²

¹ Xiamen University of Technology, Xiamen 361024, China

² Louisiana State University, Baton Rouge, LA 70803, USA

Correspondence should be addressed to Xin Li; xinli@cct.lsu.edu

Received 12 July 2014; Revised 1 October 2014; Accepted 1 October 2014; Published 20 October 2014

Academic Editor: Erik Cuevas

Copyright © 2014 P. Chen and X. Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a nonrigid coarse correspondence computation algorithm for volumetric images. Our matching algorithm first extracts then correlates image features based on a revised and improved 3DSIFT (I3DSIFT) algorithm. With a scale-related keypoint reorientation and descriptor construction, this feature correlation is less sensitive to image rotation and scaling. Then, we present an improved spectral matching (ISM) algorithm on correlated features to obtain a one-to-one mapping between corresponded features. One can effectively extend this feature correspondence to dense correspondence between volume images. Our algorithm can benefit nonrigid volumetric image registration in many tasks such as motion modeling in medical image analysis and processing.

1. Introduction

Establishing spatial correspondence between volume images is an important and fundamental computational problem in 3D image and video processing. It can greatly benefit many tasks such as nonrigid image registration [1–4], registration evaluation [5–8], volumetric ultrasound panorama [9], object recognition [10], and three-dimensional classification [11], to name a few. In biomedical computation, many medical scans (CT and MRI) are volumetric images, whose spatial and temporal correspondence are critical in clinical monitoring, diagnosis, and treatment planning [12–14]. However, robust volumetric images registration is often challenging due to the nonrigid characteristics of the deformations of the organs and tissues.

A nonrigid registration is often formulated as a nonlinear optimization which is usually nonconvex and very expensive to solve. Feature correspondence suggests desirable initial guess to this optimization problem which can improve the computation efficiency. Furthermore, it can be adopted as an extra guidance in the dense registration computation

to overcome many undesirable local minima and obtain a reliable global solution.

In 2D image registration, the SIFT descriptor and its variants [15, 16] are widely adopted feature modeling methods due to their great robustness and discrimination power. 3DSIFT [17–19] generalizes the 2D SIFT into volumetric images and has been demonstrated to be effective in volume image matching problems. However, the existing 3DSIFT algorithms are still inadequate in handling dynamically deforming volumetric images such as aforementioned medical images, due to their sensitivity against image rotation and scaling. Also, the extracted feature correspondences using local descriptors are usually many-to-many and include mismatches. We need a reliable correspondence selection/refinement algorithm to remove mismatches and obtain an objective mapping between feature point sets extracted from the source and target images. In the recent computer vision literature, such a correspondence refinement procedure is often formulated as a graph matching problem; Leordeanu and Hebert [20] proposed an efficient and effective spectral algorithm to solve this graph matching problem.

In this work, we propose a volumetric image matching algorithm in a three-step pipeline: (1) feature extraction and initial correlation, (2) feature matching and correspondence refinement, and (3) extending the coarse correspondence (feature matching) to dense intervolumetric correspondence. Specifically, consider the following:

- (1) we develop an improved 3DSIFT (I3DSIFT) descriptor that is less sensitive to rotation and scaling, to reliably extract feature correspondences between given volumes;
- (2) we propose an improved spectral matching (ISM) algorithm that can more effectively refine the initial feature correlation and obtain correct matching between corresponding features;
- (3) we implement a spline fitting algorithm to extend these coarse correspondences to dense intervolumetric correspondences.

The remainder of this paper is organized as follows. We briefly review related work in Section 2. After giving an overview of our algorithm pipeline in Section 3, we elaborate the three steps of this pipeline in Sections 4, 5, and 6. Some volumetric matching results are shown in Section 7 and one application on the motion modeling of medical images is shown in Section 8.

2. Related Work

Our algorithm of computing dense correspondences between volumetric data is guided by coarse correspondences between features. The major challenges and our contributions in this work lie in the robust extraction and matching of features. Therefore, we mainly discuss related work in these fields.

2.1. Extraction of Correlated Features. Volumetric feature extraction is the process to build coarse correspondences between volume images. It has been widely studied.

Scovanner et al. [17] first proposed the 3DSIFT descriptor and showed its application in action recognition. Their work is only on 3DSIFT descriptor construction and does not suggest feature correspondences construction. Furthermore, their descriptor is sensitive to scaling and rotation change. Ni et al. [9] implemented a 3DSIFT algorithm for feature extraction and applied it in volumetric ultrasound panorama. They extended scale space extreme detection and dominant direction reorientation step of SIFT to 3DSIFT. However, they did not utilize scale information so that their method is sensitive to scaling change. Cheung and Hamarneh extended SIFT to the so-called N-dimension SIFT (N-SIFT) [18]. They demonstrated its effectiveness using synthetic experiments on volumetric MRI and 4DCT data matching. However, N-SIFT algorithm is still sensitive to both scaling and rotation. Allaire et al. [19] proposed a *full 3DSIFT* algorithm. They filtered out potential keypoints that either have low contrast or locate on edges or ridges. But many correct corresponding keypoints are filtered out unnecessarily. As a result, only 0.1%~8% of potential keypoints were kept, which is sometimes too few for reliable volumetric matching.

Also, this 3DSIFT is sensitive to scale change. Cheung and Hamarneh [18] also proposed the global histogram feature and reoriented global histogram feature. Both of them use single histogram on a keypoint region, while the latter has an additional reorientation step. The reoriented global histogram feature normalizes the direction of the gradients with respect to the highest peak in the histogram to increase robustness to orientation change. Unfortunately, both are less discriminative than N-SIFT for the lack of space information. Han [21] proposed a method based on 3DSURF. It is based on the partial derivatives of local regions and their absolute values. Therefore, 3DSURF is sensitive to noise.

2.2. Feature Correspondence Refinement. Feature extraction and matching algorithms depending on local descriptors usually compute a set of correspondences between feature pairs. Some correspondences are many-to-many and some are mismatches. In order to provide correct and one-to-one matching for the subsequent dense correspondence computation, a refinement is needed. The RANdom SAMple Consensus (RANSAC) algorithm [22, 23] is a widely adopted randomized algorithm that iteratively estimates correspondences and eliminates outliers. It starts with a small feasible dataset and enlarges this set with consistent data correspondences when possible. The RANSAC algorithm is nondeterministic and is only stable in a probabilistic sense. In Xu et al. [24], we used 3DSIFT to extract features and applied RANSAC to obtain the feature matching.

Spectral matching [20, 25], on the other hand, is a deterministic algorithm that has been reported with better efficiency [20]. In spectral matching, a correspondence set is treated as a bipartite graph. It is based on the observation that correct matches correlate with each other strongly, while incorrect matches correlate with each other accidentally. Despite its successful application when the bipartite graph is connected, how well the algorithm works when the graph is not connected (which is the case in our pipeline) is not clear.

3. Algorithm Overview

In the paper, we propose a pipeline to find dense correspondences between volumetric images. The pipeline has three steps.

- (1) Given two volume images I_F and I_M , we extract an initial coarse correspondence set $L(I_F, I_M)$ using a revised 3DSIFT (I3DSIFT) algorithm. This initial coarse correspondence set L contains all the potential correspondences between features on I_F and features on I_M . In the following L is called the *potential correspondence set* or *initial correspondence set* (Section 4).
- (2) Some correspondences in L are mismatches; therefore, we perform a revised spectral matching (ISM) algorithm to filter them out and get a refined coarse correspondence set L' (Section 5).

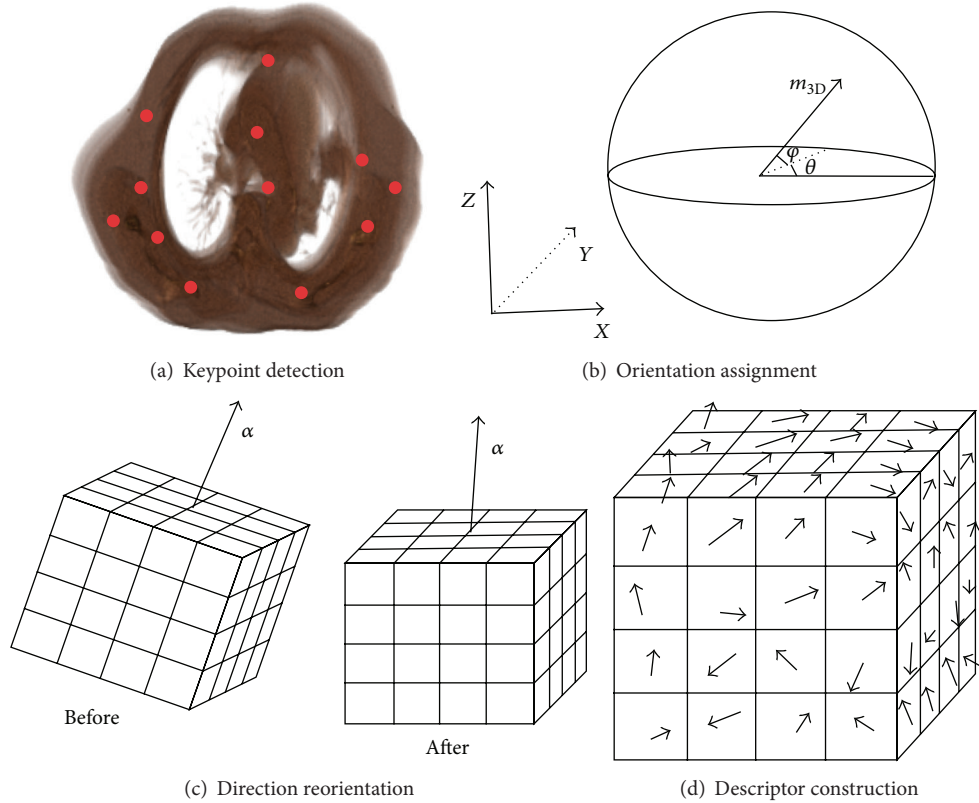


FIGURE 1: The four-step pipeline of I3DSIFT. (a) A small amount of the detected keypoints is illustrated. (b) The orientation of a pixel is computed. (c) A keypoint's surrounding region is reoriented according to its dominant direction α . (d) 3DSIFT descriptor construction. Note that we draw fixed size regions in (c) and (d) for convenient visualization while both steps actually use scale-related region in I3DSIFT.

- (3) Finally, we perform a B-spline fitting algorithm, guided by L' , to compute the dense correspondences between I_F and I_M (Section 6).

4. Revised 3DSIFT

As discussed in Section 2, a reliable feature extraction algorithm should be insensitive to scaling and rotation change. We implement a revised 3DSIFT (I3DSIFT) algorithm consisting of four main steps: keypoint detection, orientation assignment, direction reorientation, and descriptor construction. Figure 1 illustrates the feature extraction through these steps. The main changes on I3DSIFT are in the direction reorientation and descriptor construction. We propose a scale-related region definition method in these two steps. Moreover, 5-linear interpolation and feature normalization are introduced in the descriptor construction step to improve the robustness. Compared with the original 3DSIFT with fixed-size region definition, I3DSIFT is insensitive to rotation and scaling transformations. We will elaborate these four steps in the following.

Keypoint detection is an essential procedure in volumetric image matching. In 3DSIFT, scale space extrema are identified as keypoints. The detection is based on the scale space theory [26, 27]. In real implementation, DoG (difference of Gaussian) [15, 28] is usually adopted to detect

keypoints. Given a volume image, its DOG images can be defined as a stack of images that are smoothed using series of DoG functions with the variable σ . For efficiency, an image pyramid with N octaves and S levels per octave is implemented to represent DoG images. The extremes of $3 \times 3 \times 3 \times 3$ across scale in each octave are defined as keypoints. The σ of the Gaussian image in which the keypoint shows up is selected as the keypoint's scale. For example, a keypoint that shows up on the s th level image, its scale c is assigned as $c = 2^{s/S} \sigma$. The scale of a keypoint can represent its surrounding feature region size (e.g., keypoint with big scale has a big feature region and vice versa). Therefore, we implement a scale-related region definition method in the following direction reorientation and descriptor construction steps.

Orientation assignment provides pixel attributes for descriptor construction. It has been shown that gradient-based descriptors, such as SIFT [15], are more discriminative than intensity-based descriptors [16]. 3DSIFT is a gradient-based descriptor. Voxel gradient can be depicted according to its azimuth (θ , $\theta \in [0, 360)$), elevation (φ , $\varphi \in [0, 180)$), and magnitude (m_{3D}).

Direction reorientation is an essential procedure to deal with image rotation. It includes two steps: dominant direction definition and keypoint region reorientation. In the previous 3DSIFT algorithms [17], the size of the keypoint regions is

defined as a constant (e.g., sized $16 \times 16 \times 16$). However, this definition does not distinguish keypoints with different feature regions. Keypoints with small feature regions have the same size of keypoint regions with the large ones. Accordingly, the definition of the dominant direction is not accurate. In the paper, inspired by SIFT [15], we implement a direction reorientation step that is related to scale. Our size of the keypoint region for orientation definition is defined as $K * c$, where K is a user-defined parameter and c is the scale of the keypoint defined above. The definition of K is related to the content of volume images. In our lung CT images application, we set $K = 6$. Consequently, this definition is more accurate.

Descriptor construction is to build description for each keypoint. In order to design a 3DSIFT descriptor that is insensitive to scale change, similarly, we add a scale-related descriptor region. We set the region size $12c \times 12c \times 12c$. After subdivision of the descriptor region into $4 \times 4 \times 4$ subregions, we add magnitudes of the pixels in each subregion (its size is $3c$) according to its θ and φ . In order to avoid all boundary effects that the descriptor abruptly changes as a sample shifts from being within one subregion to another or one orientation to another during this step, we implement a 5-linear interpolation method to interpolate within three subregion dimensions (X , Y , and Z) and two gradient orientations ($\theta(x, y, z)$ and $\varphi(x, y, z)$). Each entry into a bin is multiplied by a weight of $1-d$ for each dimension, where d is the distance of the sample from central value of the bin as measured in units of the histogram bin spacing. Then, we join the feature of these subregions and get the descriptor. In order to discard information due to noise or illumination, we normalize the descriptor to $[0, 1]$ and discard features below a threshold (0.2 in the paper) by setting it as 0. Finally we normalize the descriptor into $[0, 1]$ and get the final descriptor.

5. Revised Spectral Matching

After initial feature extraction and matching (Section 4), we obtain an initial correspondence set L , indicating potential matches between features on I_F and I_M . To filter out incorrect matches, we can use the spectral matching algorithm. Compared with the well-known RANSAC [22] algorithm, spectral matching is in general more efficient [20, 25], especially when a good initial correspondence set L is given.

5.1. Basic Idea of Spectral Matching. Spectral matching is an efficient spectral method to find consistent correspondences in an initial correspondence set L . An assignment graph G is assigned to represent L . Each node represents a correspondence, while the weights on the edges represent the pairwise agreements between the correspondences. G can be represented by an affinity matrix M . The definition of M is shown as follows:

$$M(a, b) = \begin{cases} 4.5 - \frac{(d_{ij} - d_{i'j'})^2}{2\sigma_d^2} & \text{if } |d_{ij} - d_{i'j'}| < 3\sigma_d \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where $a = (i, i')$ and $b = (j, j')$ are two correspondences from the potential correspondence set. d_{ij} is the distance between points i and j , the same as $d_{i'j'}$. M is a nonnegative symmetric matrix and each m_{ab} encodes the pairwise agreement of right matches between a and b correspondences.

The basic idea of spectral matching is from an observation: correct assignments are likely to establish link with each other and thus form a strongly connected cluster, while incorrect correspondences establish links with others only accidentally so that they are unlikely to belong to strongly connected clusters. Accordingly, correct matches form a strong cluster in M , while incorrect matches do not. This phenomenon can be illustrated by the *principal eigenvector* of M (denoted as v^p). The elements of v^p represent the confidence of each match. Correct matches have big values in v^p , while incorrect matches have small ones. Consequently, a greedy selection method based on v^p is presented to pick out the correct matches. Intuitively, the matches are picked based on the elements from big value to small value, while the one-one matching constraint is considered.

5.2. A Revision on Spectral Matching. The basic idea of spectral matching is strongly based on the analysis of the affinity matrix M and the assignment graph G . When most of the elements in M are larger than 0, G usually corresponds to a connected graph; when most of them are 0, G usually corresponds to a unconnected graph.

The basic spectral matching algorithm works well when G is highly connected. However, when a discriminative descriptor was used to establish the potential correspondence set, the mismatch pairs have been significantly reduced, and G is often sparsely connected or even not fully unconnected (see Figures 2(a)–2(c)). For such graphs, the basic spectral matching algorithms have the following problems.

The first problem is the theoretical error on unconnected graphs for spectral matching. As was mentioned above, spectral matching can be illustrated by the principal eigenvector of M . However, this is not true when M is not a primitive matrix (i.e., M corresponds to a unconnected graph) according to the Perron-Frobenius theorem [29]. Directly applying spectral matching on unconnected graphs is theoretically unsupported.

The second problem is the failure on filtering out unobvious wrong matches. Spectral matching considers all the positive values in v^p . It can filter out those matches whose values equal zero or which do not meet the matching constraint. However, when wrong matches have some connection with others they have positive value in v^p . Spectral matching can not filter them out correctly.

The existing spectral matching algorithms [20, 25] can not work well when G is unconnected, specifically, when initial correspondences in L are almost one-one matching. In this work we present a revised spectral matching (ISM) to overcome these problems (see Figures 2(d)–2(f)).

First, M is decomposed into several submatrices. Suppose that M has r , $r \geq 1$ connected components; it can be decomposed into r primitive submatrices. Hence, it is

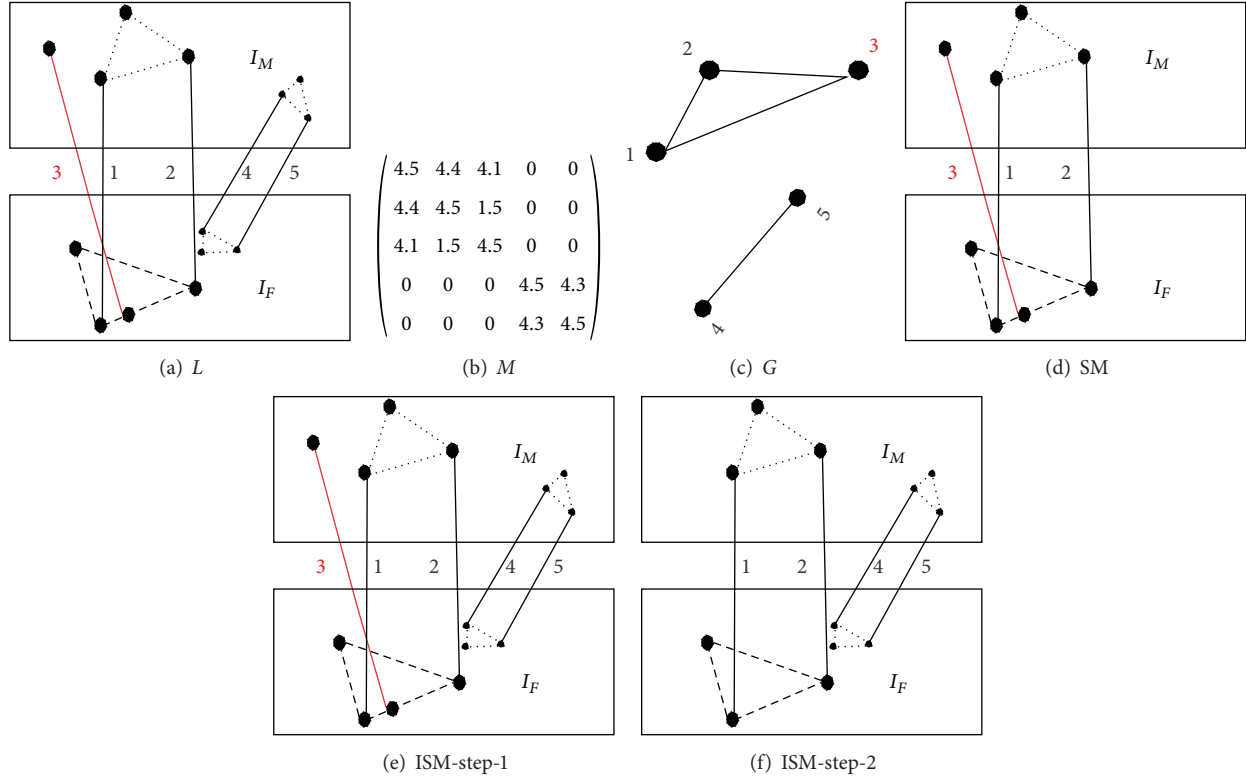


FIGURE 2: An simple example demonstrating the effectiveness of ISM. (a) shows an initial correspondence set L . (b) is the affinity matrix computed from L . (c) shows the assignment graph G which is unconnected. (d) shows the result of SM. (e) and (f) show the result of two revised steps of ISM, respectively.

theoretically supported to apply spectral matching on each submatrix.

Second, a threshold is given to filter out unobvious wrong matches. In particular, we add a reject ratio γ , $\gamma \in [0, 1]$ to discard mismatches. γ represents the rate of matches that we want to discard. A small γ could not filter out mismatches, while a larger γ leads to a higher accuracy but could reject many correct matches at the same time. We find that a suitable γ should be correlated with the accuracy rate α of L . We always choose $\gamma = 1 - \alpha$ or a slightly larger value. The accuracy rate α can always be estimated by randomly choosing a small amount of correspondences for verification. For example, in our biomedical application (see Section 8), we model the deformation of lung volume images during the respiratory cycles, where organs will not deform too drastically. And medical therapists suggest that correspondences should not have the offset more than 3 voxel in z -direction and 20 voxels in x - or y -directions. These type of rules can be used as guidance to estimate α and therefore γ .

6. Feature-Guided Image Registration

This algorithm is based on our previous paper [24]. However, unlike [24], which focuses on solving a continuous parameterization and smooth trajectory of an image sequence, we solve this registration to evaluate the effectiveness of our coarse correspondence results. Hence we used a simplified

model with only the intensity term and regularization term. The computation of this model is hence simpler.

Assume that we have two volume images. One image, the moving image $I_M(\mathbf{x})$, is deformed to fit the other image, the fixed image $I_F(\mathbf{x})$. $I_F(\mathbf{x})$ and $I_M(\mathbf{x})$ are defined on their own spatial domain: $\Omega_F \subset \mathbb{R}^3$ and $\Omega_M \subset \mathbb{R}^3$, respectively. We are going to find a transformation $\mathbf{T}(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x})$ that makes $I_M(\mathbf{T}(\mathbf{x}))$ spatially aligned to $I_F(\mathbf{x})$. The transformation is defined as a mapping from the fixed image to the moving image; that is, $T: \Omega_F \subset \mathbb{R}^3 \rightarrow \Omega_M \subset \mathbb{R}^3$. The quality of alignment is defined by a cost function C with a regularization term P . In here, we will adopt the sum of squared differences with a feature regularization term:

$$\hat{T} = \arg \min_T C(T; I_F, I_M) + \lambda P(T), \quad (2)$$

where λ weighs similarity against feature constraints. To solve the above minimization problem, there are basically two approaches: parametric and nonparametric. In here we will consider the parametric methods by introducing a parametrization model of the transformation. The original optimization problem thus becomes

$$\hat{T}_\mu = \arg \min_{T_\mu} C(T_\mu; I_F, I_M) + \lambda P(T_\mu), \quad (3)$$

where the subscript μ indicates that the transformation has been parameterized. The vector μ contains the values of the

“transformation parameters.” In here, we use the nonrigid B-spline transformation. Consider

$$T_\mu = x + \sum_{x_k \in N_x} p_k \beta^3 \left(\frac{x - x_k}{\sigma} \right) \quad (4)$$

with x_k the knot points, $\beta^3(x)$ the cubic multidimensional B-spline polynomial, p_k the B-spline coefficient vectors (the control point displacements), σ the B-spline control point spacing, and N_x the set of all knot points within the compact support of the B-spline at x . The parameters μ are formed by the B-spline coefficients p_k .

The cost function is defined as

$$C(T_\mu; I_F, I_M) = \frac{1}{|\Omega_F|} \sum_{x_i \in \Omega_F} (I_F(x_i) - I_M(T_\mu(x_i)))^2, \quad (5)$$

with Ω_F the domain of the fixed image I_F and $|\Omega_F|$ the number of voxels. Given a transformation T , this measure can easily be implemented by looping over the voxels in the fixed image $I_F(x_i)$, calculating $I_M(T_\mu(x_i))$ by interpolation, and adding the squared difference to the sum.

Basing on our proposed feature matching algorithm, we can get the feature correspondences between two volume images. They will guide the spline-fitting. The regularization term therefore minimizes the distance of two corresponding feature points:

$$P(T_\mu) = \frac{1}{|\{x_i^F\}| |\{x_i^M\}|} \sum |x_i^M - T_\mu(x_i^F)|^2, \quad (6)$$

where $|\{x_i^F\}|$ indicates the size of the correspondence set and x_i^F, x_i^M are corresponding features from the fixed and moving images, respectively.

7. Experimental Results

We denote the first two steps of our algorithm (that use I3DSIFT and ISM) as I3DSIFT-ISM. This section demonstrates the effectiveness of our feature matching (the first two steps of our algorithm). In our experiments, we generate an image pyramid with $N = 3$ octaves and $S = 6$ levels, where we set $\sigma = 1.6$ and $k = 2^{1/3}$. We set $\gamma = 0.2$ for I3DSIFT-ISM. Our experiment materials are lung volumetric images. We perform experiments on two sets of data: S1 and S2. S1 includes 3 real volume images (SE_1, SE_5, and SE_8) from different respiratory periods of the same person. S2 is from the public POPI data [30]. The volume images in S1 have a resolution of $465 \times 465 \times 20$, while images S2 are $482 \times 360 \times 141$. In preprocessing, we extract region of interests (ROI) through the graph-cut image segmentation and rescale the gray level of each volume image to $[0, 255]$.

In our experiments, we detect scale space extremes as keypoints. We match descriptors based on NN (nearest neighborhood) search and only those matches whose ratio of NN distance and the second NN distance is smaller than a threshold T ($T = 0.8$ in our paper). We perform the NN search in two directions and a matched pair is kept

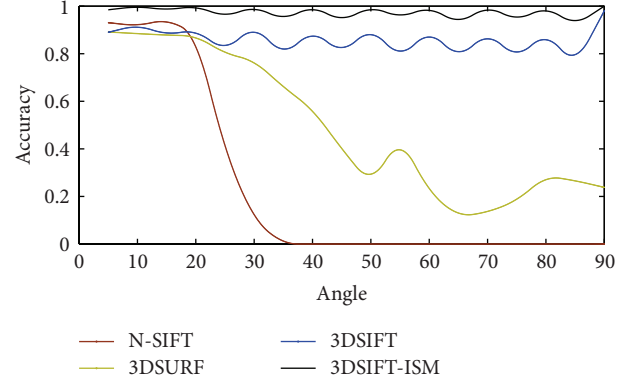


FIGURE 3: Accuracy on rotation change. Accuracy comparison is shown on a volume image and its synthetic images with varying rotation angles around z -axis.

only if it satisfies the NN optimality in both directions. We conduct NN search using a k-d tree algorithm for the sake of efficiency. We adopt a 1.5 voxel matching accuracy in the experiments on synthetic images. In the paper, most existing volumetric image matching algorithms are included to make a full range of comparison on volumetric image matching. The algorithms include global histogram (Global) [18], reoriented global histogram (Reor.) [18], N-SIFT [18], and 3DSURF [21].

We implement a 3DSURF descriptor based on the paper [21] and use a source code provided in [18] on global histogram, reoriented global histogram, and N-SIFT. I3DSIFT and I3DSIFT-ISM are implemented in C++ and ITK. We show the performance of I3DSIFT and I3DSIFT-ISM on synthetic images to demonstrate the performance with scaling and rotation change. Moreover, we show results on real data.

7.1. On Synthetic Images. Rotation insensitive is an important property to evaluate volumetric matching methods. In order to demonstrate the performance of I3DSIFT-ISM with rotation change, we do matching on SE_1 (from the first dataset S1) and its synthetic images with transforms of varying angles about z -axis. Figure 3 shows the result. In Figure 3, I3DSIFT gets an accuracy about 80% while N-SIFT [18] and 3DSURF [21] will fail when rotation change is large (larger than 30°). I3DSIFT outperforms N-SIFT [18] and 3DSURF [21]. I3DSIFT performs much better when the rotation angle is times of 10° , because we define the dominant direction every 10 degrees. As a result, we can change the way we define the dominant direction, for example, every 5° definition instead. However, we should make a compromise between accuracy and efficiency. I3DSIFT-ISM gets an accuracy of almost 100% in every rotation angle. Above all, I3DSIFT is less insensitive to rotation change to the state of the art, while, furthermore, I3DSIFT-ISM is nearly rotation-invariant.

Scale insensitive is another important property to evaluate volumetric matching methods. In order to demonstrate the performance of I3DSIFT-ISM with scale change, we show matching results on SE_1 and its down-sampling synthetic images by a varying factor along each of the 3 axes.

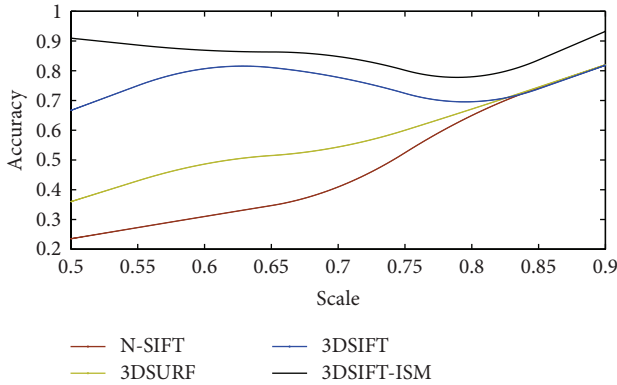


FIGURE 4: Accuracy on scale change. Accuracy comparison is shown on a volume image and its synthetic images with varying scale change.

Figure 4 shows the result. I3DSIFT outperforms N-SIFT [18] and 3DSURF [21]. There are some further improvement when using I3DSIFT-ISM. I3DSIFT is relatively robust to scale change and I3DSIFT-ISM further improves the robustness.

In conclusion, I3DSIFT-ISM is very robust against rotation and insensitive to scaling.

7.2. Matching Real Data. Our dataset images S1 are volumetric CT images scanned during respiratory cycles. SE_i indicates the i th frame. We pick three frames 1, 5, and 8, and use I3DSIFT-ISM to match their corresponding pairs: SE_15, SE_58, and SE_18. The gray values of corresponding points vary little (i.e., the lighting condition is stable) in the volume images. Therefore, we compute mean of Sum of Square Difference D to evaluate the matching results. Effective matching should produce small D value. Consider

$$D = \frac{1}{N} \sum_{1 \leq i \leq N} (I_F(x_{Fi}) - I_M(x_{Mi}))^2, \quad (7)$$

where I_F and I_M are the fixed image and the moving image, (x_{Fi}, x_{Mi}) is a matching pair between the two images, and N is the match number. Table 1 shows the results.

I3DSIFT finds matches about 2 times of N-SIFT and 1.5 times of 3DSURF. Moreover, I3DSIFT is more accurate. The I3DSIFT-ISM filters out mismatches of the matching result of I3DSIFT. The I3DSIFT-ISM is the most accurate. Figure 5 shows some matching results on volumetric images using I3DSIFT.

8. Application on Feature-Guided Image Registration

Integrating the step-3, we now apply our whole volumetric matching algorithm on the publicly available POPI dataset. In the 3D volume at time frame t , the coherent landmarks (a set of 3D points, denoted as $P_t = \{p_{t,1}, p_{t,2}, \dots, p_{t,|P_t|}\}$) are available and can be used to evaluate the registration.

In our experiments, we adopt the 3D consecutive image registration with the alignment of corresponded features enforced. The registration results are evaluated by the *mean*

TABLE 1: Mean of Square Difference D on 3 real volume image pairs. SE_{ij} represents a volume image pair, that is, SE_i is the fixed image and SE_j the moving image. The volume images are from different respiratory periods of the same person. Note that this error is small; for example, out of the gray value range of 255, $D = 32.04$ indicates a mean gray difference of only 5.66.

| Method | Dataset | Match | D |
|-------------|---------|-------|--------|
| 3DSURF | SE_15 | 239 | 79.45 |
| | SE_58 | 592 | 58.66 |
| | SE_18 | 210 | 86.88 |
| N-SIFT | SE_15 | 201 | 98.06 |
| | SE_58 | 396 | 36.64 |
| | SE_18 | 185 | 110.71 |
| I3DSIFT | SE_15 | 399 | 81.79 |
| | SE_58 | 701 | 36.18 |
| | SE_18 | 365 | 83.81 |
| I3DSIFT-ISM | SE_15 | 288 | 57.88 |
| | SE_58 | 480 | 32.04 |
| | SE_18 | 264 | 79.58 |

target registration error (MTRE) between the set of landmark points $\{P_0, \dots, P_9\}$. MTRE is defined by

$$E_{r,t} = \frac{1}{|P_t|} \sum_{p_{t,i} \in P_t} \|T^{r,t}(p_{r,i}) - p_{t,i}\|, \quad (8)$$

where $p_{t,i}$ is a landmark i in time t and $T^{r,t}$ is the transformation from time r to time t . In our experiments, we set the control weight in (3) as $\lambda = 0.1$. Table 2 shows the comparison results between feature-guided matching and the registration algorithm without any feature constraints $\lambda = 0$. By comparing the MTRE errors, we can see that the method with both I3DSIFT and I3DSIFT-ISM outperforms (by 8%) the method that does not consider the feature correspondences.

Also, the method with I3DSIFT-ISM outperforms the method with I3DSIFT. This demonstrates that

- (i) correct feature correspondences (computed by our first two-steps) are very important in improving the volumetric matching accuracy;
- (ii) better (more matches, higher correctness rate) feature correspondences lead to better matching.

Based on our feature-guided image registration, we can do the dynamic tumor motion modeling. With our volumetric registration between consecutive images, we can build up (by linear interpolation) a temporal deformation model. The motion and dynamics of the tumor can be analyzed using this model. The coherent tracking of the deforming tumor is visualized in Figure 6.

9. Conclusion

We propose a volumetric feature modeling and coarse correspondence algorithm for volumetric images. The pipeline

TABLE 2: The registration error in mm, on 40 landmarks among POPI-data. $E_{i,j}$ is the matching error from i th to j th frame; \bar{E} is the mean error for the whole sequence. We show the result using B-spline fitting without feature as well as the result using B-spline fitting constrained with the correspondence result of I3DSIFT and I3DSIFT-ISM.

| | $E_{0,1}$ | $E_{1,2}$ | $E_{2,3}$ | $E_{3,4}$ | $E_{4,5}$ | $E_{5,6}$ | $E_{6,7}$ | $E_{7,8}$ | $E_{8,9}$ | \bar{E} |
|------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Without feature | 0.71 | 0.80 | 1.42 | 1.41 | 1.71 | 1.21 | 1.51 | 1.41 | 1.13 | 1.25 |
| With I3DSIFT | 0.69 | 0.69 | 1.40 | 1.40 | 1.16 | 1.18 | 1.50 | 1.30 | 1.08 | 1.15 |
| With I3DSIFT-ISM | 0.67 | 0.65 | 1.35 | 1.32 | 1.12 | 1.17 | 1.34 | 1.30 | 1.07 | 1.11 |

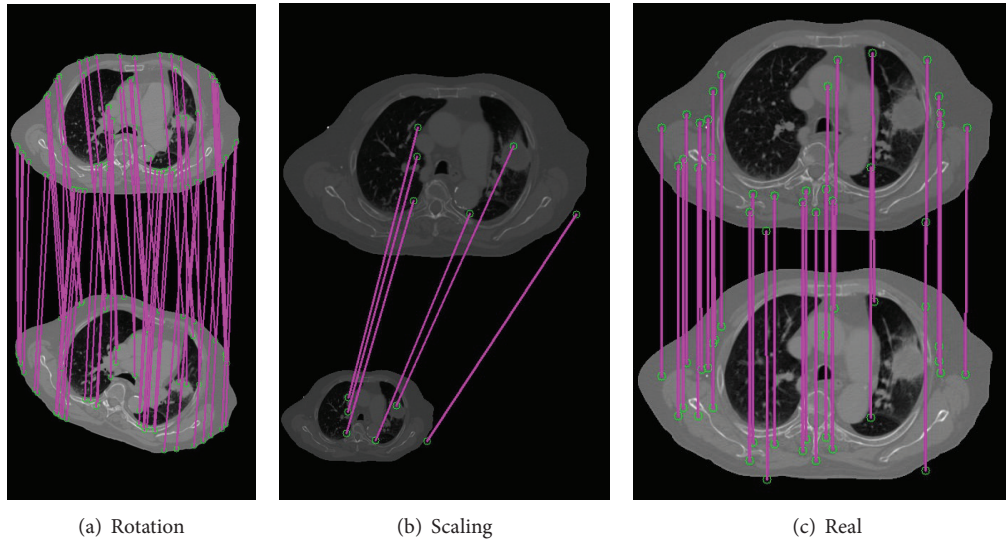


FIGURE 5: Matching examples on volumetric images using I3DSIFT. Note that we show the results in planar images for convenience, although I3DSIFT is applied to volumetric images.

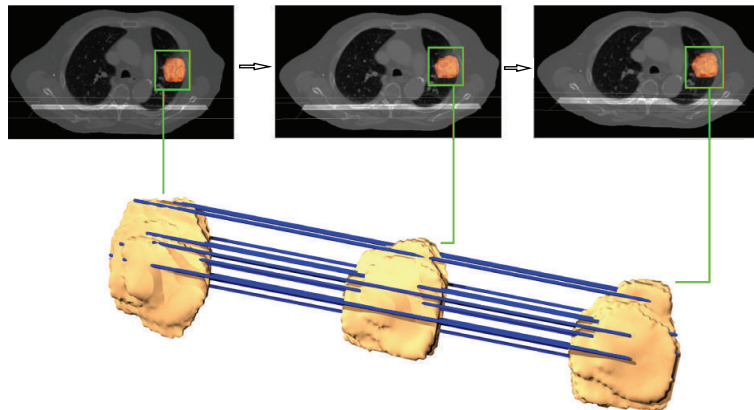


FIGURE 6: Temporal image registration guided by our feature matching results in 0th, 5th, and 9th time frame (experiments conducted on CT scans of lung during respiratory cycles).

aims to facilitate general nonrigid volumetric image registration by providing a reliable coarse feature matching. Our algorithm includes three main steps. (1) We develop a revised 3DSIFT (I3DSIFT) that is less sensitive to scaling and rotation. (2) We propose a revised spectral matching (ISM) that more reliably refines the initially extracted feature correspondences. (3) We use a B-spline function to solve the dense correspondences from the coarse (feature) correspondences.

We demonstrate the effectiveness of our coarse matching computation on both synthetic and real data by showing that our algorithm outperforms the existing volumetric matching algorithm, such as N-SIFT [18], 3DSURF [21], global histogram [18], and the reoriented global histogram [18]. Our method can be used for reliable volumetric matching. As its application, we show effective feature-guided image registration using our volumetric matching pipeline. We

show that it outperforms the image registration method without feature guidance.

A limitation of the current pipeline is the expensive computational cost in 3DSIFT. Feature extraction based on this needs a convolution in 3-dimensional space so it is quite slow. However, in this multiscale space, the computation can be easily conducted in a parallel manner. In the near future, we will explore its GPU/multithread implementation to achieve more efficient computation.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is partially supported by the US National Science Foundation IIS-1320959 and IIS-1251095; US Louisiana Board of Regents LEQSF(2009-12)-RD-A-06, LEQSF-EPS(2009)-PFUND-133, and LEQSF-EPS(2013)-PFUND-312; National Natural Science Foundation of China no. 61170323 and no. 61373147; Science and Technology Research Projects of XMUT YKJ11012R and YKT10037R; and the 12th Five-year Plan Project of Xiamen Education Science no. 1250.

References

- [1] K. Rohr, H. S. Stiehl, R. Sprengel, T. M. Buzug, J. Weese, and M. H. Kuhn, "Landmark-based elastic registration using approximating thin-plate splines," *IEEE Transactions on Medical Imaging*, vol. 20, no. 6, pp. 526–534, 2001.
- [2] M. Urschler, J. Bauer, H. Ditt, and H. Bischof, "Sift and shape context for feature-based nonlinear registration of thoracic CT images," in *Computer Vision Approaches to Medical Image Analysis*, vol. 4241 of *Lecture Notes in Computer Science*, pp. 73–84, 2006.
- [3] S. Chambon, A. Moreno, A. P. Santhanam, J. P. Rolland, E. Angelini, and I. Bloch, "CT-PET landmark-based lung registration using a dynamic breathing model," in *Proceedings of the 14th Edition of the International Conference on Image Analysis and Processing (ICIAP '07)*, pp. 691–696, Modena, Italy, September 2007.
- [4] M. Urschler, C. Zach, H. Ditt, and H. Bischof, "Automatic point landmark matching for regularizing nonlinear intensity registration: application to thoracic ct images," *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '06)*, vol. 9, no. 2, pp. 710–717, 2006.
- [5] K. Murphy, B. van Ginneken, J. Pluim, S. Klein, and M. Staring, "Semi-automatic reference standard construction for quantitative evaluation of lung CT registration," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2008*, pp. 1006–1013, 2008.
- [6] R. Castillo, E. Castillo, R. Guerra et al., "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," *Physics in Medicine and Biology*, vol. 54, no. 7, pp. 1849–1870, 2009.
- [7] Y. Yin, E. A. Hoffman, and C.-L. Lin, "Mass preserving nonrigid registration of CT lung images using cubic B-spline," *Medical Physics*, vol. 36, no. 9, pp. 4213–4222, 2009.
- [8] X. Huang, N. Hill, J. Ren, G. Guiraudon, D. Boughner, and T. Peters, "Dynamic 3d ultrasound and mr image registration of the beating heart," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2005*, vol. 3750 of *Lecture Notes in Computer Science*, pp. 171–175, Springer, New York, NY, USA, 2005.
- [9] D. Ni, Y. Qu, X. Yang et al., "Volumetric ultrasound panorama based on 3d sift," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2008*, vol. 5242 of *Lecture Notes in Computer Science*, pp. 52–60, Springer, New York, NY, USA, 2008.
- [10] G. Flitton, T. P. Breckon, and N. Megherbi, "Object recognition using 3D SIFT in complex CT volumes," in *Proceedings of the 21st British Machine Vision Conference (BMVC '10)*, September 2010.
- [11] J. Knopp, M. Prasad, G. Willems, R. Timofte, and L. van Gool, "Hough transform and 3D SURF for robust three dimensional classification," in *Computer Vision—ECCV 2010*, vol. 6316 of *Lecture Notes in Computer Science*, pp. 589–602, Springer, Berlin, Germany, 2010.
- [12] D. Sarrut, "Deformable registration for image-guided radiation therapy," *Zeitschrift für Medizinische Physik*, vol. 16, no. 4, pp. 285–297, 2006.
- [13] S. Iyengar, X. Li, H. Xu et al., "Toward more precise radiotherapy treatment of lung tumors," *Computer*, vol. 45, no. 1, pp. 59–65, 2012.
- [14] H. Xu and X. Li, "A symmetric 4d registration algorithm for respiratory motion modeling," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*, vol. 8150 of *Lecture Notes in Computer Science*, pp. 149–156, Springer, Berlin, Germany, 2013.
- [15] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [16] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [17] P. Scovanner, S. Ali, and M. Shah, "A 3-dimensional sift descriptor and its application to action recognition," in *Proceedings of the 15th ACM International Conference on Multimedia (MM '07)*, pp. 357–360, September 2007.
- [18] W. Cheung and G. Hamarneh, "N-sift: N-dimensional scale invariant feature transform for matching medical images," in *Proceedings of the 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 720–723, Arlington, Va, USA, 2007.
- [19] S. Allaire, J. J. Kim, S. L. Breen, D. A. Jaffray, and V. Pekar, "Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '08)*, pp. 1–8, Anchorage, Alaska, USA, June 2008.
- [20] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, vol. 2, pp. 1482–1489, Beijing, China, October 2005.
- [21] X. Han, "Feature-constrained nonlinear registration of lung CT images," in *Medical Image Analysis for the Clinic: A Grand Challenge*, pp. 63–72, 2010.

- [22] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the Association for Computing Machinery*, vol. 24, no. 6, pp. 381–395, 1981.
- [23] S. Choi, T. Kim, and W. Yu, "Performance evaluation of RANSAC family," in *Proceedings of the British Machine Vision Conference*, 2009.
- [24] H. Xu, P. Chen, W. Yu, A. Sawant, S. S. Iyengar, and X. Li, "Feature-aligned 4D spatiotemporal image registration," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR '12)*, pp. 2639–2642, November 2012.
- [25] M. Leordeanu, *Spectral graph matching, learning, and inference for computer vision [Ph.D. thesis]*, Carnegie Mellon University, 2010.
- [26] J. J. Koenderink, "The structure of images," *Biological Cybernetics*, vol. 50, no. 5, pp. 363–370, 1984.
- [27] T. Lindeberg, "Scale-space theory: a basic tool for analyzing structures at different scales," *Journal of Applied Statistics*, vol. 21, no. 1-2, pp. 225–270, 1994.
- [28] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.
- [29] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, Mass, USA, 2012.
- [30] J. Vandemeulebroucke, D. Sarrut, and P. Clarysse, "The POPI-model, a point-validated pixel-based breathing thorax model," in *Proceedings of the 15th International Conference on the Use of Computers in Radiation Therapy (ICCR '07)*, pp. 195–199, Toronto, Canada, 2007.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

