

A Novel Material-Aware Feature Descriptor for Volumetric Image Registration in Diffusion Tensor Space

Shuai Li^{1,*}, Qiping Zhao¹, Shengfa Wang³, Tingbo Hou²,
Aimin Hao¹, and Hong Qin²

¹ State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China

² Stony Brook University, Stony Brook, New York, USA

³ Dalian University of Technology, Dalian, China

Abstract. This paper advocates a novel material-aware feature descriptor for volumetric image registration. We rigorously formulate a novel probability density function (PDF) based distance metric to devise a compact local feature descriptor supporting invariance of full 3D orientation and isometric deformation. The central idea is to employ anisotropic heat diffusion to characterize the detected local volumetric features. It is achieved by the elegant unification of diffusion tensor (DT) space construction based on local Hessian eigen-system, multi-scale feature extraction based on DT-weighted dyadic wavelet transform, and local distance definition based on PDF formulated in DT space. The diffusion, intrinsic structure-aware nature makes our volumetric feature descriptor more robust to noise. With volumetric images registration as verifiable application, various experiments on different volumetric images demonstrate the superiority of our descriptor.

1 Introduction and Motivation

With the rapid advancement of various volumetric imaging modalities, we have been witnessing the urgent need for automatic feature detection and the discriminative feature description of complex volumetric dataset in image registration, object recognition, video event detection, image retrieval, etc. To achieve this, some recent works have tried to extend 2D SIFT-like methods to 3D versions, for example, Scovanner et al. [1] created a 3D SIFT descriptor for video action recognition, Flitton et al. [2] extended the SIFT approach to 3D rigid recognition, and other applications include rigid registration of medical images [3, 4] and panoramic medical image stitching [5, 6].

Despite the limited success, certain difficulties still prevail and need to be resolved. The challenges are prompted by the facts that volumetric images typically have much more spatial flexibility, and are frequently accompanied by

* This research is supported in part by National Natural Science Foundation of China (No.61190120, No.61190121 and No.61190125) and NSF grants IIS-0949467, IIS-1047715, and IIS-1049448.

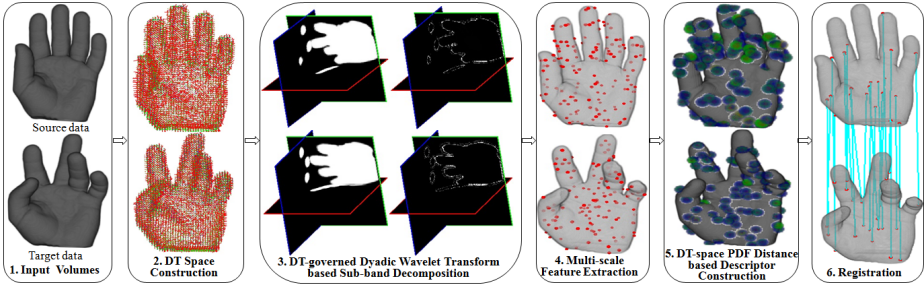


Fig. 1. The algorithmic pipeline of our method

non-rigid deformation with higher degrees of freedom. As for non-rigid registration, although some intensity/information entropy based methods can easily achieved this goal by integrating global energy optimization and deformable templates, however, a typical global approach tends not to consider local deformable feature-driven information and non-affine distortion. It remains hard for localized, feature-centric registration methods, since this requires the feature descriptor to be intrinsic, concise, informative, and discriminative. Simple statistics on local properties in intensity and gradient domain won't work, we should resort to the intrinsically physical laws underlying the embedded manifold space. Specially, the main challenges are documented as follows.

First, due to the complex topological degrees of freedom inside the volume dataset, multi-scale feature extraction based on Difference-of-Gaussian (DoG) convolution analysis frequently obtain a large number of less salient or false alarm candidates. Especially, ambiguities are unavoidable for the ones with low contrast or being poorly localized nearby an edge. More material-aware convolution kernels, which can respect the local geometry structure and its orientations, still need to be further explored for multi-scale feature extraction.

Second, most of the 3D descriptors simply imitated from 2D SIFT can only partially satisfy the rotational invariance. Although Allaire et al. [3] achieves the full rigid orientation invariance by taking 3-angle orientation (azimuth, elevation, and tilt) into account, the descriptor dimensionality is up to 16,384. From the application's viewpoint, this is less efficient and far from practical.

Third, analogous to the analysis for shape descriptor in [7], and besides the rigid transformation, the feature descriptor should take deformation into account as much as possible. However, this typically requires certain kind of mapping by parameterizing local volumetric structure with intrinsic metric over certain canonical domain, which may cause even more severe deformation effects. Thus, intrinsic metrics supporting deformation-invariant volumetric feature description are yet to be systematically explored.

To tackle the aforementioned challenges, we systematically articulate a novel material-aware feature descriptor for volumetric images. Towards the ambitious goal of isometric invariance, our observation is that, the diffusion process is intrinsically relevant to the diffusion distance metric design and the

probability of random walk, which are informative for the description of local intrinsic structure. Meanwhile, naively using the popular isotropic diffusion process (e.g., Gaussian kernel) will naturally give rise to the smooth transition between nearby regions without respecting evident clues on edges and ridges. One feasible strategy to combat this problem is to replace Gaussian kernel with the structure-aware anisotropic DT-weighted kernel during convolution. Thus, we formulate a novel descriptor by combining random walk with probability density functions in DT space. Fig. 1 illustrates the pipeline of our approach, and highlights its application in automatic registration of volumetric features (undergoing quasi-isometric deformation). The salient contributions of this paper include:

- We formulate a local diffusion tensor based on Hessian eigen-system, which can fully grasp the second order differential properties, encode the directional curvatures of local structure, intrinsically reveal the material continuity, and control the diffusion in anisotropic way.
- We devise a data-specific kernel by integrating diffusion tensor with bilateral filter, which can be employed to conduct dyadic-wavelet based direction-aware decomposition for structure-respected multi-scale feature extraction.
- We design a random walk based feature descriptor, which depicts the local material property by measuring the difference among probability density functions defined in DT space. Inherited from heat diffusion, it is robust to noise, supports isometric deformation invariance, and can better reveal the underlying material distribution statistics.

2 Related Work

2.1 Feature Descriptor Design

Existing descriptors can be roughly divided in two classes according to their level of invariance. Rigid transformation has been accommodated rather easily in different descriptors, such as phase-based descriptor, spin images, gradient location and orientation histogram [3], automated learning based descriptor [8], and combined method of logarithmic sampling with Fourier analysis [9]. As for non-rigid deformation, to our best knowledge, only in the field of surface shape analysis, some deformation-invariant descriptors have been proposed in [10–12]. Of which, most advanced approaches are the ones based upon Laplacian spectrum analysis [13], however, the required global eigen-decomposition of such methods cannot be afforded by volumetric images. Thus, analogous intrinsic feature descriptors of volumetric images still need to be systematically explored in a local and efficient way.

2.2 Image Structure Analysis

Tensor space method has great superiority in structure-respected image analysis. The structure tensor, as a measurement for edges and their orientations, has

been widely used in texture analysis. For Example, Malcolm et al. [14] generalized the tensor method to segment images by taking the Riemannian geometry of the tensor space into account. However, the structure tensor used in [14] only reflects the orientation information at a single scale and fails to discriminate textures which are varying across different scales. Most recently, the multi-scale structure tensor proposed in [15] has demonstrated successful applications in image fusion. Besides, Brox et al. [16] argued that if the local orientation is not homogeneous, the local neighborhood induced by the Gaussian filter will integrate ambiguous structure information. Thus, Bazán and Blomgren [17] proposed to perform image smoothing and edge detection by combining anisotropic diffusion and bilateral filtering. As an extension to this, Bazán et al. [18] also successfully used this technique to enhance the structure of electron tomography. Therefore, it is necessary to introduce certain tensor distance metric to govern multi-scale feature extraction.

2.3 Intrinsic Distance Metrics

Geodesic distance can measure the shortest path between two points over the curved surface, which has been widely used in graphics tasks. However, as noted in [19], the geodesic is not shape-aware, and sensitive to topological noise. Another popular metric is the diffusion distance, which has been widely employed in texture synthesis, gradient approximation, and shape matching [20]. In essence, the diffusion distance relates to diffusion time and a number of random walks in Brownian motion. The integration of diffusion distance along time [21], named commute-time distance, is also adopted on graphs. It measures the average time of the heat diffusion between two points, and relates to the Green's function of the Laplacian. As an improvement, Lipman et al. [19] proposed the bi-harmonic distance derived from the Green's function of the bi-harmonic operator. The bi-harmonic distance is locally isotropic, globally shape-aware, and isometry-invariant. However, it fails to handle local/partial shape analysis, because the Green's function is globally defined. For other distance metrics, please refer to the comprehensive survey [22]. Inspired by these, it is a robust way to devise intrinsic volumetric feature descriptor by measuring local material distribution with the help of diffusion-like distance metric.

3 Diffusion Tensor Space Construction

As already demonstrated in many previous works, the proper definition of tensor space over a scalar image will be a key to local material structure analysis and subsequent image processing. The rich differential geometry theory offers an elegant method to achieve this by treating an image as a differentiable manifold [23].

As the simplest tensor, structure tensor (Fig.2A) is derived from first-order differential analysis, which can locally characterize the predominant directions of material changes and how those directions are related to each other. However,

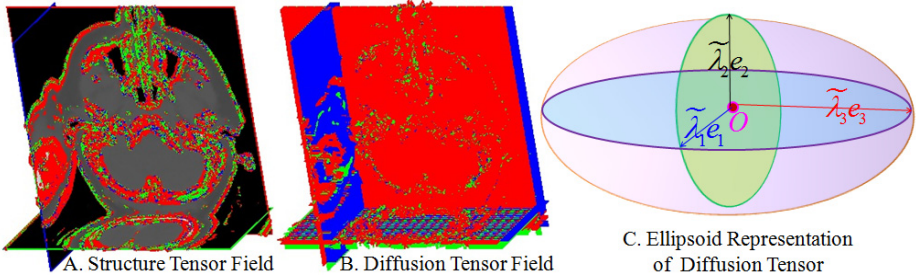


Fig. 2. Illustration of structure tensor, diffusion tensor and its physical meaning

first-order derivatives cannot fully grasp the local geometric differential property. Thus, we employ Hessian eigen-system to define the local diffusion tensor, which facilitates to the description of the second-order structure and intuitively depicts how the surface normal changes.

Hessian matrix \mathbf{H} is a symmetric matrix consisting of second-order partial derivatives, and has real-valued eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ and corresponding eigenvectors. The directions corresponding to the maximal eigenvalue of \mathbf{H} should represent the most direct change from one material to adjacent neighboring material, while the direction corresponding to the minimal eigenvalue shows the material interface and how such material flows along the interface. To suppress the diffusion when cutting across sharp material boundaries, we formulate an anisotropic diffusion tensor by a spectral representation as:

$$\mathbf{D}(p) = \tilde{\lambda}_1 \mathbf{e}_1 \mathbf{e}_1^T + \tilde{\lambda}_2 \mathbf{e}_2 \mathbf{e}_2^T + \tilde{\lambda}_3 \mathbf{e}_3 \mathbf{e}_3^T, \tag{1}$$

$$\tilde{\lambda}_i = \exp\left(-\frac{\lambda_i}{\sigma_d}\right), i = 1, 2, 3, \tag{2}$$

with diffusion parameter σ_d that controls the diffusion velocities. As shown in Fig.2C, in fact we construct an ellipsoid that encodes the direction and velocity of diffusion. According to the theory of Rayleigh quotient, the diffusion velocity from voxel p along \mathbf{e} can be viewed as the length of the vector projection onto the ellipsoid, which is expressed as

$$vel(p, \mathbf{e}) = \frac{\mathbf{e}^T \mathbf{D}(p) \mathbf{e}}{\mathbf{e}^T \mathbf{e}}. \tag{3}$$

Therefore, for a voxel inside a blob, all of its diffusion directions are principal diffusion directions. For a voxel on a boundary surface, all the directions aligning with its tangent plane are principal diffusion directions. For a voxel on a sharp edge, the direction along the edge is principal diffusion direction. For an isolated noise voxel, it will have no principal diffusion directions, as the velocities along all the directions are extremely small.

4 Multi-scale Feature Extraction

With the constructed anisotropic diffusion tensor field governing the diffusion direction and velocity, we extract multi-scale point features founded upon dyadic wavelet transform based decomposition, which comprises two steps: anisotropic wavelet kernel construction, multi-scale analysis and feature extraction.

4.1 Anisotropic Wavelet Kernel Construction

The visual perception research has indicated that the cells having directional selectivity are found in the retinas and visual cortices of the entire major vertebrate classes, thus naively using the anisotropic kernel will naturally give rise to directional information loss without having evident clues on material structure.

In order to respect the direction information embedded in the local structure during multi-scale analysis, our anisotropic wavelet kernel (AWK) is derived from the diffusion tensor and bilateral filter. AWK determines the convolution weights by considering both the directional continuity of material structures and the photometric similarity, which prefers nearby values to distant values in both spatial and material metric domain (DT space). Given two neighboring voxels located at p and q , we first define their diffusion tensor space distance as

$$d_D(p, q) = \exp(-(p - q)^T (w_{pq}(\mathbf{D}(p) + \mathbf{D}(q))^{-1}(p - q))), \quad (4)$$

w_{pq} is introduced to amend the gradient, which changes in response to the intensity change of neighboring voxels. In fact, $\mathbf{D}(p) + \mathbf{D}(q)$ describes the diffusivity and controls the diffusion directions and velocities, and w_{pq} respects the intensity variance between neighboring voxels. Therefore, we can define the AWK as

$$\Psi(p) = \frac{1}{W_p} \sum_{q \in N(p)} G_{\sigma_s}(p - q) G_{\sigma_k}(d_D(p, q)) I(q). \quad (5)$$

where W_p is a normalization factor, $G_\sigma(x) = \exp(-x^2/\sigma^2)$ is the Gaussian kernel function, and σ_k is a control parameter being set to the inverse of the maximal eigenvalues of diffusion matrices $\mathbf{D}(p)$ and $\mathbf{D}(q)$.

4.2 AWK Based Multi-scale Feature Extraction

With the built-in capability to faithfully respect material structure, and also inspired by the wavelet decomposition nature of DOG operation in the SIFT framework, we can use the proposed AWK to decompose a volumetric image into an approximation sub-band and a detail sub-band. However, only one-level decomposition is not enough to extract the feature information since images may be noisy and objects inherently comprise different details changing as a function of the observation scale. Thus, we adopt the dyadic wavelet transform to define the multi-scale form of AWK as

$$I^{n+1}(p, \sigma_s) = \frac{1}{W_p} \sum_{q \in N(q)} \omega^n(p - q, \sigma_s) G_{\sigma_k}(d_D^n(p, q)) I^n(q), \quad (6)$$

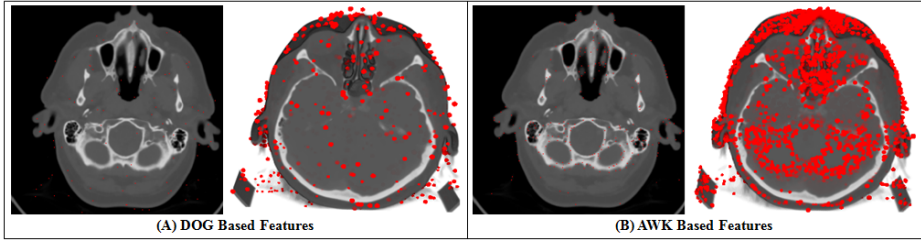


Fig. 3. Illustration of features respectively extracted by DOG and AWK operators

$$\omega^n(x, \sigma_s) = \begin{cases} G_{\sigma_s}(\|\frac{x}{2^n}\|) & \text{if } \frac{x}{2^n} \in Z^3 \text{ and } \|\frac{x}{2^n}\| < m \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

n represents the n -th level of the decomposition, W_p has the same meaning as Eq. (5), m is a threshold to control the size of neighboring region.

In implementation, it is iterated over the approximate sub-bands according to Eq. (6) and only the one-ring neighbors of each voxel are considered in each iteration. After $k + 1$ iterations, the approximate sub-band corresponding to a certain scale can be obtained, and k detail sub-bands are respectively the difference between the neighboring approximate sub-bands as

$$I(p, k\sigma) = I^{k+1}(p, \sigma) - I^k(p, \sigma). \quad (8)$$

Since point features are usually defined as local extrema of some quantities related to geometry, texture, or other information, and our multi-scale sub-band decomposition is exactly an anisotropic approximation to the Laplacian, the multi-scale point features can be obtained by extracting local minima/maxima from the detail sub-bands across scales, where a voxel will be accepted as feature if and only if all of its 80 neighbors approve that it is the brightest/darkest one respectively. Fig. 3 shows the comparison of DOG based features and AWK based features. In Fig. 3 and the other experiment figures, larger point corresponds to larger scale feature. AWK operator proves to be more informative, since the resulted features intrinsically respect sharp structures and suppress the unstable features which are poorly localized near the low contrast regions.

5 Invariant Feature Descriptor Based on PDF Distances

5.1 DT-Space PDF Distance Metrics

Inside the diffusion tensor space, the behavior of anisotropic heat diffusion can be determined by its graph Laplacian. Consider volumetric image I as an undirected graph $G = (V, E)$, the anisotropic diffusion operator \mathbf{T} can be defined as

$$\mathbf{T}(v_i, v_j) = \mathbf{S}(v_i) - \mathbf{L}(v_i, v_j), \quad (9)$$

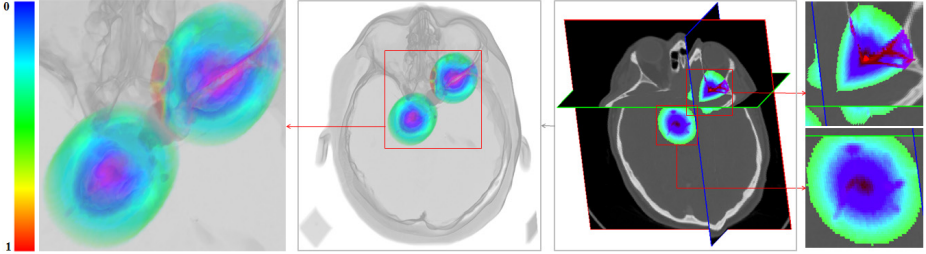


Fig. 4. Illustration of unnormalized PDFs for two feature points inside head volume

where \mathbf{L} denotes the graph Laplacian operator over G , $\mathbf{L}(v_i, v_j)$ equals to $d_D(v_i, v_j)$ (Eq. (4)), and $\mathbf{S}(v_i) = \sum_{v_j \in N_i} \mathbf{L}(v_i, v_j)$. Since \mathbf{L} is self-adjoint, the operator \mathbf{T} is self-adjoint with all non-negative entries.

From the perspective of probability in Brownian motion, \mathbf{T} is a random walk matrix with non-zero entries along the main diagonal, which allows one-step walk from a point to itself. Each entry $\mathbf{T}(v_i, v_j)$ stands for the probability of the Brownian motion moving from v_i to v_j in one step. Thus, the power \mathbf{T}^n encodes the probability of a Brownian motion from one point to another in n steps, which naturally gives rise to the random walk based probability density functions (PDF) after approximation and normalization. We formulate the PDF $\mathbf{P}_{v_i}(v_j)$ of voxel v_i as

$$\mathbf{P}_{v_i}(v_j) = \frac{\mathbf{T}^n(v_i, v_j)}{\|\mathbf{T}^n(v_i, v_k)\|_2}, \tag{10}$$

where the denominator serves for the normalization purpose, thus $\|\mathbf{P}_{v_i}(v_j)\|_2 = 1$ and $\mathbf{P}_{v_i}(v_j) > 0$. The number of random walks n is a positive integer. For fast computation, we select n from the dyadic powers 2^j . It allows to compute the matrix power \mathbf{T}^n through matrix multiplication in numerics. Since we are particularly interested in measuring the local geometry structure of volumetric image, the number of random walks n is empirically set to 2^4 . Fig. 4 illustrates the unnormalized PDFs for two feature points (the central red point). It states that PDF can efficiently reflect the material continuity, for example, the voxels belonging to the same kind of material as that of feature point have high probability, which appear red.

Consider a family of PDFs $\{\mathbf{P}_v\}_{v \in V}$ in I , if $\forall v_x, v_y \in V, v_x \neq v_y$, and there $\exists v_z \in V$, satisfies $\mathbf{P}_{v_x}(v_z) \neq \mathbf{P}_{v_y}(v_z)$, then $\{\mathbf{P}_v\}_{v \in V}$ is called *generic*, which means that no two PDFs are completely the same in a generic family of PDFs. We use the *2-norm* distance between two PDFs in $\{\mathbf{P}_v\}_{v \in V}$ as PDF metrics (PDFM):

$$d_P(v_x, v_y) = \|\mathbf{P}_{v_x}(v_z) - \mathbf{P}_{v_y}(v_z)\|_2. \tag{11}$$

Eq. (11) can also use L_p ($p > 0$) norm. Since \mathbf{P} is a vector, the range of $d_P(v_x, v_y)$ is $[0, 2^{1/p}]$. Thus it is $[0, \sqrt{2}]$ here.

In essence, PDFM describes the intrinsic material relationship, which has many attractive properties. First, inheriting from the anisotropic Laplacian

operator, it is isometry-invariant. Second, it is locally supported, since the power series \mathbf{T}^n span a scaling space in the diffusion wavelets, and the PDF $\mathbf{P}_{v_i}(v_j)$ is purely determined by a local sub-volume with v_i as a center, whose range is bounded by steps n . Third, according to the probability theory of Brownian motion and Markov chain, it is insensitive to noise, because small local changes do not have much influence to the entire set of all connected paths, hence the distribution of probabilities.

5.2 Feature Descriptor Design

PDFM is a metric naturally based on heat diffusion, and it is defined in DT space, so if the underlying material undergoes isometric deformation, the PDFM distribution in the vicinity of feature points is expected to have little or no change. We define our feature descriptor as a 2D shape context histograms comprising PDFM and the normalized image density (or gradient norm) of the volumetric image with the closest scale to that of each feature point.

We select a sub-volume centered around each feature point and compute the PDFM for all voxels in this sub-volume, and the radius is set to be the length of 8 voxels in our experiments. According to the value range of PDFM, we create 15 bins from 0 to $\sqrt{2}$ with step internal 0.1. For each bin, we take statistics of the normalized intensity (or gradient norm) of the voxels whose PDFM distance to the feature point is in current bin. Then, we create 17 intensity bins from 0 to 255 with step internal 15 or 10 gradient norm bins from 0 to 1 with step internal 0.1.

Compared with [3], which can only support rigid transformation/rotation invariance and whose descriptor dimensionality is up to 16,384, our feature descriptor is much more effective and compact (we at most need a 255-dimensional feature descriptor for each feature point).

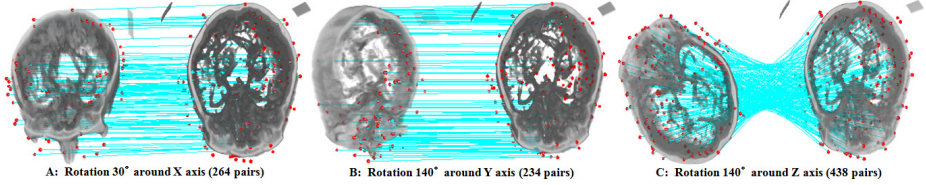
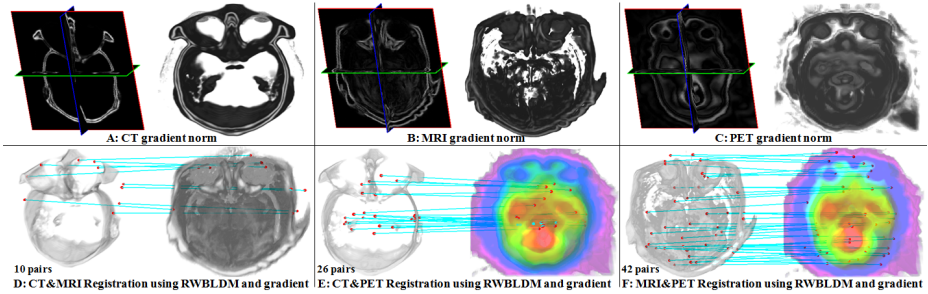
6 Applications and Experimental Results

Our prototype system is implemented using C++, and some Matlab functions are invoked to perform sparse matrix multiplication. We conduct experiments on a laptop with Intel Core (TM) i7 CPU (1.6GHz, 4 cores) and 4G RAM. Table 1 documents the time performance (in seconds) and some other experimental statistics, including Hessian matrix computation, DT construction (DTC), sub-band decomposition (SD), feature extraction (FE), number of feature points, descriptor construction (DC) and registration.

First, in order to verify the full orientation invariance with ground truth, we create various rotated volumetric images from the original one. In the interest of visual clearness, only half of the registration lines are shown in Fig. 5 (The exact number of the matching pairs is documented at the bottom of each sub-figure). During feature registration, the matching is determined by *Distance Ratio*, which is computed by comparing the distance of the closest candidate to that of second-closest candidate, we set the *Distance Ratio* to be 0.8. A match

Table 1. Time performance (in seconds) of our experiments

Dataset	Volume Size	Hessian	DTC	SD	FE	Feature #	DC	Registration
Fig.1	$128^2 \times 128$	40.9	95.3	12.7	224.6	357	220.7	14.6
Fig.5	$256^2 \times 73$	152.4	121.8	27.8	515.4	459	296.4	51.2
Fig.6	$256^2 \times 62$	150.1	105.8	24.9	567.7	559	469.3	83.2
Fig.7	$256^2 \times 64$	96.7	111.7	19.2	373.5	126	106.3	4.2
Fig.8	$128^2 \times 115$	38.9	65.5	12.1	279.4	570	480.9	88.7

**Fig. 5.** Rigid registration with full 3D orientation of head volumetric images**Fig. 6.** Multi-modality registration of Monkey head volumetric images

is deemed true when the counterpart lies within 2 voxel diagonal length of the ground truth position, the results in Fig. 5 quantitatively prove that our method can well (average 98% accuracy) support full orientation invariance.

Second, to test the capacity of our method in multi-modality volumetric image registration, we use datasets downloaded from the Laboratory of Neuro Imaging of UCLA, which have already been registered, thus offering the ground truth. Here, we compute our feature descriptor with gradient norm bin, because the gradient norm is more informative than intensity among different modality images. Fig. 6 (A-C) respectively illustrates the volumetric gradient norm of original datasets. Since the CT dataset includes less structure information than the MRI and PET datasets, the corresponding number of the matched pairs in Fig. 6 (D-E) is a bit less than that of Fig. 6 (F). In this group of experiments, we can achieve the average registration accuracy of more than 95 percents for multi-modality images.

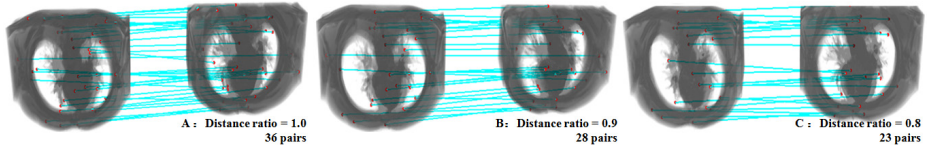


Fig. 7. Feature-based nonrigid registration of thorax MRI volumes

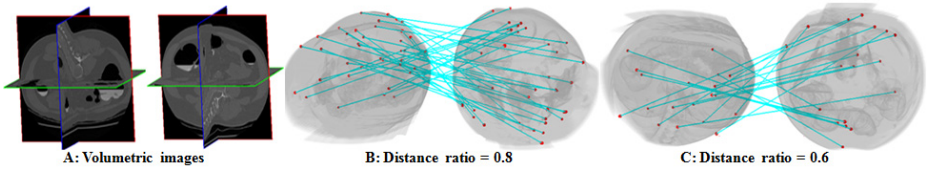


Fig. 8. Feature-based nonrigid registration of abdomen CT volumes

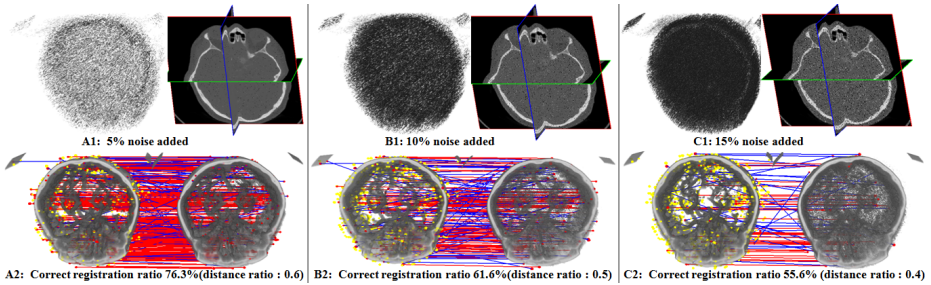


Fig. 9. Noise-perturbed datasets of head MRI volume and the registration results

Third, as for volumetric images with quasi-isometric deformation (far beyond isometric deformation), we use two experiments to qualitatively verify the isometric deformation invariance of our feature descriptor. Fig. 7 shows the non-rigid registration results of human MRI thorax volumes which are obtained from the same person before and after breathing. Although the shape of the heart and blood vessel deforms drastically, with the *Distance Ratio* reducing from 1.0 to 0.8, the mismatched pairs gradually disappear, and almost all the feature points in Fig. 7 (C) can be accepted as true. Fig. 8 shows the non-rigid registration results of two abdomen CT volumes respectively scanned in supine and prone orientations. All the matched lines in Fig. 8 (B-C) are roughly forming a cross shape, which well aligns with the orientation change (from supine to prone). When the *Distance Ratio* is set to be 0.6, most of the features located at muscle, stomach and spine can be retrieved as correct ones. It proves the superiority of our descriptor in feature-based registration with isometric deformation.

Fourth, to further examine the robustness, we respectively add 5%, 10% and 15% (of average intensity) random noise to the original volumetric images at

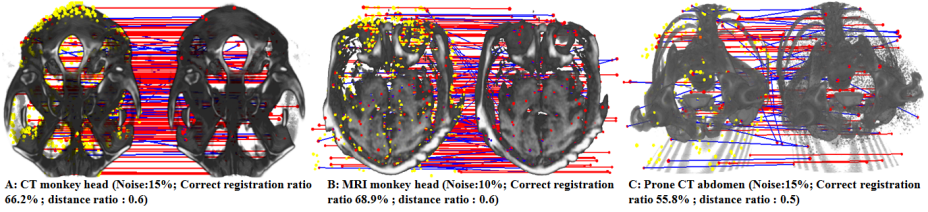


Fig. 10. More registration results for noise-perturbed datasets

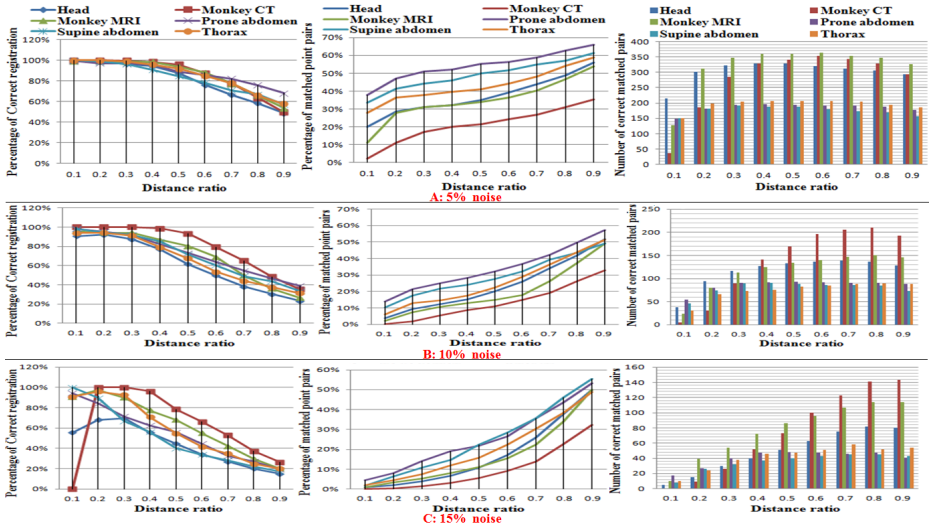


Fig. 11. The performance analysis

randomly-sampled locations. The top row of Fig. 9 shows the noise effects overlaid onto the original MRI head volume. We use the original dataset as source image and the noise-perturbed dataset as target image for feature registration. The *match ratio* is defined as the percentage of the matched pairs to the total detected feature points. As we have the ground truth, we accept the matched pair as correct ones if the distance between source point and target point is less than two voxels. For each registration result in Fig. 9, we document the *correct registration ratio* and the corresponding *Distance Ratio*. More results are also shown in Fig. 10 and our supplementary video, where red lines denote correct registration, blue lines denote incorrect registration, and feature points in yellow are the ones that cannot be matched.

Finally, we use correct registration ratios, feature matching ratios, and number of correctly-matched feature pairs to conduct quantitative evaluation. The left, middle, and right subgraph of Fig. 11 respectively reveals the relationship between the above indicators and the parameter *Distance Ratio*. For example,

the head dataset can achieve better registration performance when the *Distance Ratio* is around 0.5, since it will have both higher correct registration ratios and enough correctly-matched feature pairs even though the match ratios are not very high. As for other datasets, focusing on each type of curves, we can observe similar trends despite different noise perturbation levels and data types.

7 Conclusion

We have detailed a comprehensive feature extraction and description method for volumetric images with intrinsic properties of being material-aware. The technical originality is centered in the integration of diffusion tensor weighted dyadic wavelet transform for multi-scale analysis and the PDF distance based metric design in diffusion tensor space. At the application level, our method supports feature-based volumetric registration with full orientation invariance and isometric deformation invariance. Extensive experiments and comprehensive evaluation have demonstrated the effectiveness and robustness of our method.

For our ongoing efforts, we will continue to conduct comprehensive evaluation, and to broaden the application scope. Applications of immediate interest include local parametric representation, solid recognition, similar shapes clustering, material distance embedded meshless physical simulation, and etc.

References

1. Scovanner, P., Ali, S., Shah, M.: A 3-dimensional SIFT descriptor and its application to action recognition. In: Proceedings of International Conference on Multimedia, pp. 357–360 (2007)
2. Flitton, G., Breckon, T., Megherbi, B.N.: Object recognition using 3D SIFT in complex CT volumes. In: Proceedings of British Machine Vision Conference, pp. 11.1–11.12 (2010)
3. Allaire, S., Kim, J., Breen, S., Jaffray, D., Pekar, V.: Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis. In: Proceedings of CVPR Workshops, pp. 1–8 (2008)
4. Niemeijer, M., Garvin, M.K., Lee, K., Ginneken, B.V., Abrámoff, M.D., Sonka, M.: Registration of 3D spectral OCT volumes using 3D SIFT feature point matching. In: Proceedings of SPIE Medical Imaging, pp. 72591I.1–72591I.8 (2009)
5. Dalvi, R., Hacihaliloglu, I., Abugarbieh, R.: 3D ultrasound volume stitching using phase symmetry and harris corner detection for orthopaedic applications. In: Proceedings of SPIE Medical Imaging, pp. 762330.1–762330.8 (2010)
6. Ni, D., Chui, Y., Qu, Y., Yang, X., Qin, J., Wong, T., Ho, S., Heng, P.: Reconstruction of volumetric ultrasound panorama based on improved 3D SIFT. *Comp. Med. Imag. and Graph.* 33(7), 559–566 (2009)
7. Bronstein, M.M., Kokkinos, I.: Scale-invariant heat kernel signatures for non-rigid shape recognition. In: Proceedings of CVPR, pp. 1704–1711 (2010)
8. Winder, S.A.J., Brown, M.: Learning local image descriptors. In: Proceedings of CVPR, pp. 1–8 (2007)
9. Kokkinos, I., Yuille, A.L.: Scale invariance without scale selection. In: Proceedings of CVPR, pp. 1–8 (2008)

10. Bronstein, A.M., Bronstein, M.M., Bruckstein, A.M., Kimmel, R.: Analysis of two-dimensional non-rigid shapes. *Int. J. Comput. Vision* 78(1), 67–88 (2008)
11. Lipman, Y., Funkhouser, T.: Möbius voting for surface correspondence. *ACM Trans. Graph.* 28(3), 72.1–72.12 (2009)
12. Reuter, M., Biasotti, S., Giorgi, D., Patané, G., Spagnuolo, M.: Discrete laplace-beltrami operators for shape analysis and segmentation. *Computers and Graphics* 33(3), 381–390 (2009)
13. Aubry, M., Schlickewei, U., Cremers, D.: The wave kernel signature: A quantum mechanical approach to shape analysis. In: *Proceedings of ICCV Workshops*, pp. 1626–1633 (2011)
14. Malcolm, J., Rathi, Y., Tannenbaum, A.: A graph cut approach to image segmentation in tensor space. In: *Proceedings of CVPR*, pp. 18–25 (2007)
15. Han, S., Tao, W., Wang, D., Tai, X.C., Wu, X.: Image segmentation based on grabcut framework integrating multiscale nonlinear structure tensor. *Trans. Img. Proc.* 18(10), 2289–2302 (2009)
16. Brox, T., van den Boomgaard, R., Lauze, F., van de Weijer, J., Weickert, J., Mrázek, P., Kornprobst, P.: Adaptive structure tensors and their applications. In: Weickert, J., Hagen, H. (eds.) *Visualization and Processing of Tensor Fields*, vol. 1, pp. 17–47. Springer, Berlin (2006)
17. Bazán, C., Blomgren, P.: Image smoothing and edge detection by nonlinear diffusion and bilateral filter. Technical Report CSRCR2007-21, San Diego State University (2007)
18. Bazán, C., Miller, M., Blomgren, P.: Structure enhancement diffusion and contour extraction for electron tomography of mitochondria. *J. Struct. Biol.* 166(2), 144–155 (2009)
19. Lipman, Y., Rustamov, R.M., Funkhouser, T.A.: Biharmonic distance. *ACM Trans. Graph.* 29(3), 27.1–27.11 (2010)
20. Bronstein, A.M., Bronstein, M.M., Kimmel, R., Mahmoudi, M., Sapiro, G.: A gromov-hausdorff framework with diffusion geometry for topological-robust non-rigid shape matching. *Int. J. Comput. Vision* 89(2-3), 266–286 (2010)
21. Yen, L., Fouss, F., Decaestecker, C., Francq, P., Saerens, M.: Graph Nodes Clustering Based on the Commute-Time Kernel. In: Zhou, Z.-H., Li, H., Yang, Q. (eds.) *PAKDD 2007. LNCS (LNAI)*, vol. 4426, pp. 1037–1045. Springer, Heidelberg (2007)
22. Cha, S.H.: Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences* 1(4), 300–307 (2007)
23. Zhang, J., Zheng, J., Cai, J.: A diffusion approach to seeded image segmentation. In: *Proceedings of CVPR*, pp. 2125–2132 (2010)