20th International Conference on Information Fusion
Xi'an, China - July 10-13, 2017

Log-Euclidean Metric for Robust Multi-Modal Deformable Registration

Qiegen Liu[1,2], *Henry Leung*[1]

[1]Department of Electrical and Computer Engineering, University of Calgary, Calgary, T2N 1N4, Canada
[2]School of Electronic Information Engineering, Nanchang University, Nanchang, 330031, China

***Abstract*—Registration of images from different modalities in the presence of intra-image fluctuation and noise contamination is a challenging task. The accuracy and robustness of the deformable registration largely depend on the definition of appropriate objective function, measuring the similarity between the images. Among them the multi-dimensional modality independent neighbourhood descriptor (MIND) is a promising method, yet its ability is limited by non-uniform bias fields and image noise, etc. Motivated by the fact that Log-Euclidean metric has promising invariance properties such as inversion invariant and similarity invariant, this paper introduces an objective function that embeds Log-Euclidean similarity metric between patches to form a multi-dimensional descriptor. The Gaussian-like penalty function consisting of the log-Euclidean metric between images to be registered is incorporated to better reflect the degree of preserving feature discriminability and structure ordering. Experimental results show the advantages of the proposed method over state-of-the-art techniques both quantitatively and qualitatively.**

***Index Terms*—Multi-modal images, deformable registration, Log-Euclidean metric, Image gradients, MIND.**

I. INTRODUCTION

Compared to the advances resulted in a variety of robust and accurate methods for deformable registration techniques for scans of the same modality, the registration of images from different modalities remains a challenging task [1]. The deformation between multi-modal images describes how pixels move from a reference to a target image, which is a rich source of information for the image analysis. A good registration is beneficial for various applications such as computer-assisted surgery and disease analysis [2, 3], etc. Specifically, the registration of magnetic resonance imaging (MRI), computed tomography (CT) and positron emission tomography (PET) is very common, since combining different anatomical and functional information about the human body is indeed helpful for clinicians.

Due to the different physical phenomena that are measured by the different modalities, there is no functional relation between the intensity mapping of corresponding anatomies. The classical intensity metrics used in mono-modality such as sum-of-squared-differences (SSD) is not applicable. Since the local intensity of different modalities is exactly different, an alternative way to describe the invariance between the aligned objects is to explore their relation in global or non-local way such that the resulting quantity to be modality independent [1, 4, 5]. Specifically, Mutual information (MI) is the popularly used information theoretical algorithm that globally describe the correlation between the reference and images to be registered. MI aims to find a statistical intensity relationship across images and thereby maximizes the amount of shared information between two images. The main disadvantage is that MI is intrinsically a global measure and therefore its local estimation is difficult, which can lead to many false local optima in non-rigid registration. Hence some modality independent local quantity has been proposed to complementarily improve the alignment such as local phase, gradient orientation and multi features [6-15]. Pluim suggested integrating the gradient magnitude and orientation values into the MI measure [8]. Beijing used higher order MI [9] to include a range of other properties such as mean and median values of a pixel neighbourhood, different neighbouring pixels (left and right neighbours) and also a gradient value. Kubecka suggested using gradient-image MI, where MI is computed for both the original images (after performing illumination correction) and also for the corresponding gradient images [10]. Legg *et al.* incorporated multi-scale feature derivatives along with spatial neighbourhood knowledge into a MI framework. It consists of firstly computing the features from the fixed and moving images and then combining these together in a MI framework [11].

Structural image representation is another class of widely used approaches. Modality independent neighbourhood descriptor (MIND) assigns each pixel a structural vector that describes the central pixel in a local/nonlocal way [16-18]. Wachinger and Navab [19] proposed the entropy image and Laplacian eigenmaps based structural representation methods for image registration. In the first method, the entropy of the image patches was calculated and the entropy images based SSD was used as the similarity metric. A shortcoming is that the entropy images seem to be blurry, and it is sensitive to noise when calculating the histogram over small patches. In the second method, all the image patches were utilized to build a neighborhood graph to approximate the manifold embedded in high dimensional patch space. The low-dimensional embedding was then calculated with the graph Laplacian. Embeddings from different modalities were aligned to obtain the final representation. Yet the Laplacian images seem to be sensitive to noise in the image and their calculation involves high computational complexity. Piella [20] used the diffusion maps to obtain a unified representation that captures the geometric and spectral properties of the data for the multi-modal registration. This method can represent the complicated image features, but it is computationally intensive and sensitive to image noise.

This work was supported in part by the National NSFC under 61503176, 61362001, the international postdoctoral exchange fellowship program.

978-0-9964-5270-0©2017 ISIF

The main contribution of this paper is to propose a new descriptor termed MIND-LeM, which exploits the Log-Euclidean metric for matrices containing local feature correlations. We approximate this descriptor by applying Log-Euclidean distance between Gaussian smoothed patches in a large local neighbor of image gradients domain, thereafter forming the multi-dimensional similarity metric. The remainder of this paper is organized as follows: After reviewing some knowledge of Log-Euclidean metric In Section II, Section III introduces a new Log-Euclidean based modality independent neighborhood descriptor. Subsequently, Section IV demonstrates the performance of the proposed algorithm. Finally, conclusions and discussions are given in Section V.

II. REVIEW OF LOG-EUCLIDEAN METRICS

In this section, we briefly review the Log-Euclidean distance and its Gaussian kernels. The space of $d \times d$ Symmetric Positive Definite (SPD) matrices, Sym_d^+, is widely studied when endowed with a Riemannian metric and thus forms a Riemannian manifold [21, 23]. In such cases, the geodesic distance induced by the Riemannian metric is a more natural measure of dissimilarity between two SPD matrices than the Euclidean distance. Log-Euclidean distance [22] is the most widely used distance measure, due to the fact that it is a true geodesic distance induced by Riemannian metrics.

The log-Euclidean distance for Sym_d^+ is derived by exploiting the Lie group structure of Sym_d^+ under the group operation $Q_i \odot Q_j := \exp(\log(Q_i) + \log(Q_j))$ for $Q_i, Q_j \in Sym_d^+$ where $\exp(\cdot)$ and $\log(\cdot)$ denote the usual matrix exponential and logarithm operators, respectively. Under the log-Euclidean framework, a geodesic connecting Q_i and Q_j is defined as $\gamma(t) = \exp((1-t)\log(Q_i) + t\log(Q_j))$ for $t \in [0,1]$. The geodesic distance between Q_i and Q_j can be expressed as

$$D(Q_i, Q_j) = \left\| \log(Q_i) - \log(Q_j) \right\|_F \tag{1}$$

where $\|\cdot\|_F$ denotes the Frobenius matrix norm induced by the Frobenius matrix inner product $\langle \cdot, \cdot \rangle_F$. As can be observed, Log-Euclidean metrics correspond to Euclidean metrics in the domain of logarithms. Compared to Euclidean metrics, they satisfy a number of invariance properties, like inversion invariant, construction invariant with respect to any logarithmic multiplication and similarity-invariant (e.g., orthogonal transformation and scaling). For more details of the Log-Euclidean metrics on Sym_d^+, readers are referred to [21, 22, 24].

Furthermore, since kernel methods have been proven effective for many computer vision tasks, some researches are devoted to kernel methods for manifold-valued data. The Gaussian kernel is perhaps the most popular kernel. It would therefore seem natural to adapt this kernel to account for the geometry of Riemannian manifolds by replacing the Euclidean distance in the Gaussian kernel with the geodesic distance on the manifold. One positive definite Gaussian kernel is stated as following:

$$(Sym_d^+ \times Sym_d^+) \to R:$$
$$K(Q_i, Q_j) := \exp\left(-\frac{D^2(Q_i, Q_j)}{2\sigma^2}\right) \tag{2}$$

Numerical experiments in [23, 25] have shown that the resulting manifold kernel methods outperform the corresponding Euclidean kernel methods.

III. PROPOSED ALGORITHM

A. Review of MIND

The modality independent neighborhood descriptor (MIND) [5] was proposed by Heinrich *et al.* for multi-modal image registration. The MIND was computed based on the similarities between neighboring patches. The MIND is not rotationally invariant and it only utilizes the image intensities for similarity computation. Therefore, it will generate inaccurate registration results for corners, edges and complicated textured regions where rotations may exist between features.

Self-similarity is a concept introduced by Buades *et al.* for nonlocal means (NLM) denoising [26]. Similar to the NLM method, the MIND also explores image self-similarities by replacing the local comparison of individual pixels with the non- local comparison of image patches. In the MIND, for two pixels at x and $x+r$ in the spatial search region R of image I, the similarity $MIND(I, x, r)$ is computed as:

$$MIND(I, x, r) = \frac{1}{n}\exp\left(-\frac{D(I, x, x+r)}{V(I, x)}\right) \quad r \in R \tag{3}$$

$$D(I, x, y) = \left\| P_x - P_y \right\|^2 \tag{4}$$

where $D(I, x, x+r)$ is the Euclidean distance between two similarity windows (i.e., image patches P_x and P_{x+r}) centered at x and $x+r$, respectively. The denominator in Eq. (3) acts as the smoothing parameter and it is computed as the mean of the patch distances themselves within a six-neighborhood centered at x. From Eq. (3), we can see that the MIND computes the pixel similarity based on the differences between the intensities of image patches in the translational sense. Because this method only utilizes the intensity information, it is not robust to contrast variations. It is difficult to accurately determine self-similarity for the medical images with the complicated edge/texture features which may involve rotational self-similarity. The incorrect local image structure representation resulting from the MIND will influence the final non-rigid multi-modal registration results. To address this problem, some researchers turn to integrate some intensity-insensitive information into the MIND principle for achieving better alignment [16-18]. For instance, Li *et al.* exploited the autocorrelation of local structural information consisting of local phase and phase congruency features [17]. In this work, a non-Euclidean metric will be introduced by virtue of its rotational invariance to explore rotational self-similarity effectively by utilizing both intensities and edge/texture features of the multi-modal images.

B. Proposed MIND-LeM Descriptor

In this work, a new Log-Euclidean metric for image registration is introduced. As stated in refs. [21, 22, 24], the variable Q in log-Euclidean distance Eq. (1) should be a SPD matrix.

Usually, it can be the covariance matrix that describes the local image properties, consisting of local image features such as the intensity, gradients or high-order derivatives. In the calculation of Log-Euclidean distance, it involves the eigen decomposition. Here we sort to a simple and approximate way to approach the LeM descriptor. We view the variables in two-pixel neighbor region to form the covariance matrix as a SPD matrix with size $d=1$, similar to the strategy used in ref. [27]. Specifically, for image intensity set $\{I_{x-1}, I_x\}$ located at two neighbor points, the covariance matrix is $Q_x=(I_{x-1}-I_x)^2=(\nabla I_x)^2$. Notice that in this case the covariance matrix is truly the first-order derivatives, i.e., gradients.

After this special covariance matrix is defined, we incorporate it into the MIND descriptor, replacing the classical Euclidean metric by the Log-Euclidean metric. Accordingly, the similarity descriptor termed $MIND\text{-}LeM(I,x,r)$ is computed as:

$$MIND\text{-}LeM(I,x,r)=\frac{1}{n}\exp(-\frac{D(I,x,x+r)}{V(I,x)}) \quad r\in R \tag{5}$$

$$D(I,x,y)=\left\|\log(\nabla I_x)^2-\log(\nabla I_y)^2\right\|^2 \tag{6}$$

where $\left\|\log(\nabla I_x)^2-\log(\nabla I_y)^2\right\|^2$ is the Log-Euclidean distance between two similarity image pixels, respectively. In practical implementation, we use a Gaussian-kernel weight to convolute the whole patch around pixel ∇I_x to approximate it.

Inspired by the L1-norm and in order to attain a simpler formulation, we substitute Eq. (6) by $\left\|\log\left|\nabla I_x\right|-\log\nabla I_y\right\|$ and modify Eq. (5) to be Eq. (7), yielding:

$$MIND\text{-}LeM(I,x,r)=\frac{1}{n}\exp(-\frac{\left\|\log\left|\nabla I_x\right|-\log\left|\nabla I_y\right|\right\|}{V(I,x)}) \quad y\in R_x \tag{7}$$

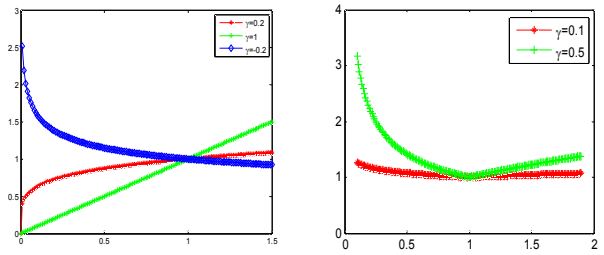
As seen from Eq. (7), MIND-LeM aims at optimizing the kernelized geodesic distance-based measure defined in Log-Euclidean space. It has intrinsic difference with the traditional measures defined in Euclidean space. In the following, we reveal their differences mathematically, and subsequently get some enlightenment why it works well for similarity metric. Particularly, by letting $\gamma=V(I,x)$ and some algebraic manipulation, Eq. (7) can be written as follows:

$$\begin{aligned}&\exp(-\gamma\left|\log\left|\nabla I_y\right|/\left|\nabla I_x\right|\right|)\\&=\left(\left|\nabla I_y\right|/\left|\nabla I_x\right|\right)^{\gamma}|_{|\nabla I_y|>|\nabla I_x|}+\left(\left|\nabla I_y\right|/\left|\nabla I_x\right|\right)^{-\gamma}|_{|\nabla I_y|<|\nabla I_x|}\end{aligned} \tag{8}$$

where $\log\left|\nabla I_y\right|/\left|\nabla I_x\right|=\log\left|\nabla I_y\right|-\log\left|\nabla I_x\right|$. As seen in Eq. (8), the objective function is divided into two parts. One for pixel pairs with $\left|\nabla I_y\right|>\left|\nabla I_x\right|$ and the other with $\left|\nabla I_y\right|<\left|\nabla I_x\right|$. The proposed metric has some promising properties. First, for the pixel pairs with $\left|\nabla I_y\right|>\left|\nabla I_x\right|$, its objective function is $\left|\nabla I_y-\nabla I_x\right|^{\gamma}$. In the meantime, for the pixel pairs with $\left|\nabla I_y\right|<\left|\nabla I_x\right|$, the associate function is $\left|\nabla I_y-\nabla I_x\right|^{-\gamma}$.

When γ is smaller than 1, i.e., $0<\gamma<1$, just like the classical power law transform function [28], for the pixel whose value $\left|\nabla I_y\right|$ is far away from the original gradient $\left|\nabla I_x\right|$, $\left(\left|\nabla I_y\right|/\left|\nabla I_x\right|\right)^{\gamma}$ will be transformed to a larger value, thus in the whole cost functional, more value will be assigned on these image pixel pairs. Specifically, the cost function in the interval of $\left|\nabla I_y\right|<\left|\nabla I_x\right|$ is amplified much higher than that in the interval of $\left|\nabla I_y\right|>\left|\nabla I_x\right|$, thus more attention will be enforced to the overlooked pixel pairs. A visual demonstration of the power law transform and the objective function with different γ are shown in Fig. 1.

Additionally, it is worth noting that when γ=0, the function is identical to 1, and the model tends to be nonsense. On the other hand, there exists $\exp(-\gamma\left|\log\left|\nabla I_y\right|/\left|\nabla I_x\right|\right|)\approx 1-\gamma\left|\log\left|\nabla I_y\right|/\left|\nabla I_x\right|\right|$, then, it should be $\gamma\le\frac{1}{\left|\log\left|\nabla I_y\right|/\left|\nabla I_x\right|\right|}$. In this work, we set it to be the mean of the patch distance within the six-neighborhood region.



(a) (b)

Fig. 1. (a) Illustration of power law transform $s=t^{\gamma}, t>0$. (b) Illustration of transform $s=t^{\gamma}|_{t>1}+t^{-\gamma}|_{0<t<1}$.

The difference between the Log-Euclidean distance Eq. (8) and Euclidean distance $(\left|\nabla I_y\right|/\left|\nabla I_x\right|-1)^2$ are shown in Fig. 2. It can be observed that our penalty function is more adaptive and meaningful.

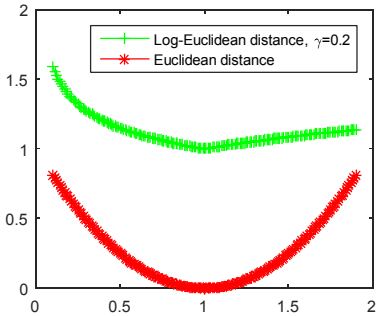


Fig. 2. Visualization of transform $s=t^{\gamma}|_{t>1}+t^{-\gamma}|_{t<1}$ and $s=(t-1)^2$.

C. Multi-modal Registration

It is worth noting that, to further improve robustness, we use the Gaussian-kernel weight to convolute the whole patch around pixel ∇I_x to approximate the gradient. Given the fixed and moving images $I_1(x), I_2(x)$, the objective of image registration is to determine the deformation field u by minimizing the following function:

$$\min_u \sum_x S(I_1(x), I_2(x+u))^2 + \alpha tr(\nabla u(x)^T * \nabla u(x)) \quad (9)$$

where $S(I_1(x), I_2(x)) = \frac{1}{|R|}\sum_{r\in R} |MIND\text{-}LeM(I_1, x, r) - MIND\text{-}LeM(I_2, x, r)|$ measures the non-local intensity difference similarity between reference and deformable image, depending on the deformation u. The second term is the diffusion deformation regularity, which favors a smooth deformation field. α is a parameter balancing the two terms.

The nonlinear function Eq. (9) can be simplified to be

$$\min_u f^T f, \quad f = [S(I_1(x), I_2(x+u)), \sqrt{\alpha}\nabla u(x)] \quad (10)$$

Since it has $S_{I_2(x+u)} \approx S_{I_2(x)} + \nabla S_{I_2(x)} u$ and $\sqrt{\alpha}\nabla u \equiv \sqrt{\alpha}\nabla u$, Its derivative is $J = [\nabla S_{I_2(x)}, \sqrt{\alpha}\nabla]$. Accordingly, by applying the Gauss-Newton, the update rule at each iteration is $J^T J u = -J^T f$, where J is the derivative of f with respect to variable u. Hence the resulting deformation field on the basis of previous value u^i is given by the following scheme

$$(\nabla S_{I_2(x)}^T \nabla S_{I_2(x)} - \alpha\Delta)u^{i+1} = -(\nabla S_{I_2(x)}^T S_{I_2(x)} - \alpha\Delta u^i) \quad (11)$$

In order to gain additional speedup and avoid being trapped in local minima, a multi-level scheme is used to represent the coarse-to-fine details of both volumes. At each step, a symmetric deformable registration is followed which prefers to obtain diffeomorphic transformations, avoiding physically implausible folding of volume occurs [29].

IV. EXPERIMENTS AND RESULTS

In this section, we test our registration algorithm on 2D and 3D data, in terms of robustness and accuracy. We evaluate our findings based on the target registration error (TRE) of anatomical landmarks. The TRE for a given transformation $u(x)$ and an anatomical landmark pair (x, x') is defined by:

$$TRE = \|x + u(x) - x'\|_2 / |L| \quad (12)$$

where L and $|L|$ are the set of anatomical landmarks and the number of landmarks in the reference image, respectively.

We first apply the method to some simulated data that needs to be registered from dataset, then perform deformable registrations on CT and MRI scans of patient and ten CT datasets of lung cancer patients. In all the experiments, the weighting parameter α in the regularization term (9) was set to 0.1 for both MIND and MIND-LeM methods. All experiments are in Matlab, on an Intel Core i7-4700MQ CPU 2.4 GHz Windows 64-bit operating system with 8 GB RAM.

A. Deformable registration of simulated data

We use the T1 and PD weighted MRI scan from the Visible Human dataset to form a synthetic data pair, the fixed image is corrupted by applying a spatially-varying intensity distortion as described by Myronenko *et al.* [30] and the moving image is contaminated by the distortion as well as Gaussian noise. As shown in Fig. 3, the high-dimensional image descriptors of MIND and MIND-LeM are displayed. It can be observed that the MIND-LeM descriptor depends less on the intensity differences within the search region around each pixel in the same modality.

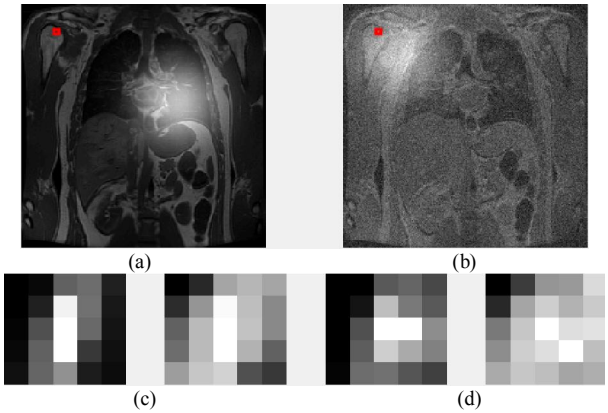


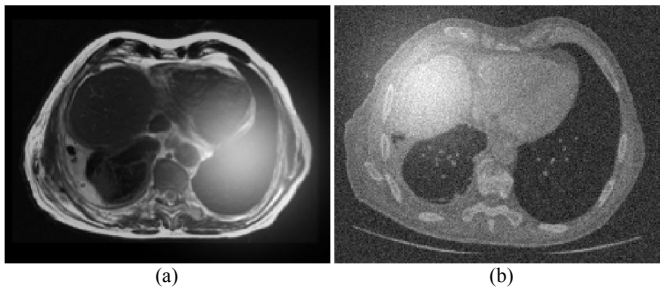
Fig. 3. Comparison of the MIND and MIND-LeM descriptor.

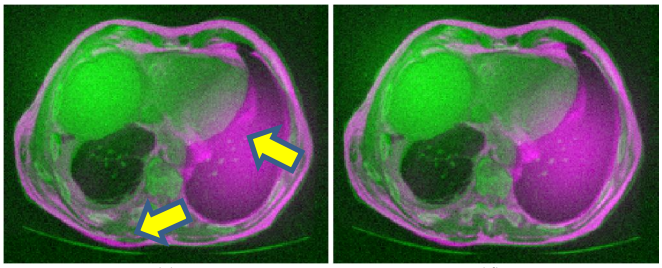
We conduct experiments in Visible Human dataset [31] and the Brain Tumor Segmentation (BRATS) challenge [32]. All the scans in both datasets are taken post-mortem, no motion is present and different modalities are aligned perfectly. To produce the artificial deformations in the synthetic data, a geometric distortion with a thin-plate spline (TPS) model is applied to the corresponding source image to get the moving image. Additionally, in the T1-PD pair, spatially-varying intensity distortion is added to the both fixed and moving images.

In the result, the reference image is displayed in green and the registered version of moving image in magenta. As expected, a well-done registration yields a gray-scale image and larger color differs imply higher misregistration. The registration comparison is shown in Fig. 4. It can be observed that MIND-LeM produces better registration results in the region with simulated intensity distortions and large spatial deformations than MIND, as indicated by the yellow arrows.

B. Deformable registration of real data

A demonstration of registering the CT-MRI is depicted in Fig. 5. We purposely add non-functional intensity distortion and Gaussian additive noise on the fixed and moving images. Compared to MIND, MIND-LeM attains better correspondence between points of similar geometry in the fixed CT and moving MR image, as denoted by yellow arrows. MIND-LeM is robust to noise and has the ability to preserve small structures.





(c) (d)

Fig. 5. Deformable registration result for CT-MRI pair. (a) simulated fixed image, (b) moving image, (c)(d) registration display between the fixed and moving image obtained by MIND and MIND-LeM.

We perform deformable registration on ten 4D thoracic CT scan pairs between inhale and exhale phase of the breathing cycle, provided by the DIR-Lab at the University of Texas [33]. Each scan pair is acquired on the thorax and upper abdomen of patients treated for esophagian cancer, between inhale and exhale phase of the breathing cycle. The slice thickness is 2.5 mm, and in-plane resolution ranges from 0.97 to 1.16 mm. For each scan 300 anatomical landmarks have been carefully annotated by thoracic imaging experts with inter-observer errors of less than 1 mm. Major challenges arise from possible contrast variations between tissue and air induced by lung compression, motion discontinuities at the lung/rib cage interface, as well as large deformations of small features such as lung vessels, airways. In the experiment, we apply registration directly between each pair of the original CT scans.

Among all the cases, the maximum average landmark error before registration is 15 mm in Case 8, the maximum displacement of a single landmark is 30 mm. The visual comparison of this shown in Fig. 6. In this visualization, the source image is shown in magenta while the reference image is shown in green. Gray scale image will emerge in the regions where the images are fully aligned. In the unregistered case, magenta and green areas can clearly be observed indicating that the morphology is not aligned. In the registered cases, these colored areas almost disappear indicating that the images have been successfully registered. Particularly, MIND-LeM largely diminishes these regions than that of MIND.

Table 1 summarizes the average TRE comparison results. The mean landmark distance and corresponding standard deviations are recorded. It can be observed that MIND-LeM achieves lower TRE value than the MIND method.

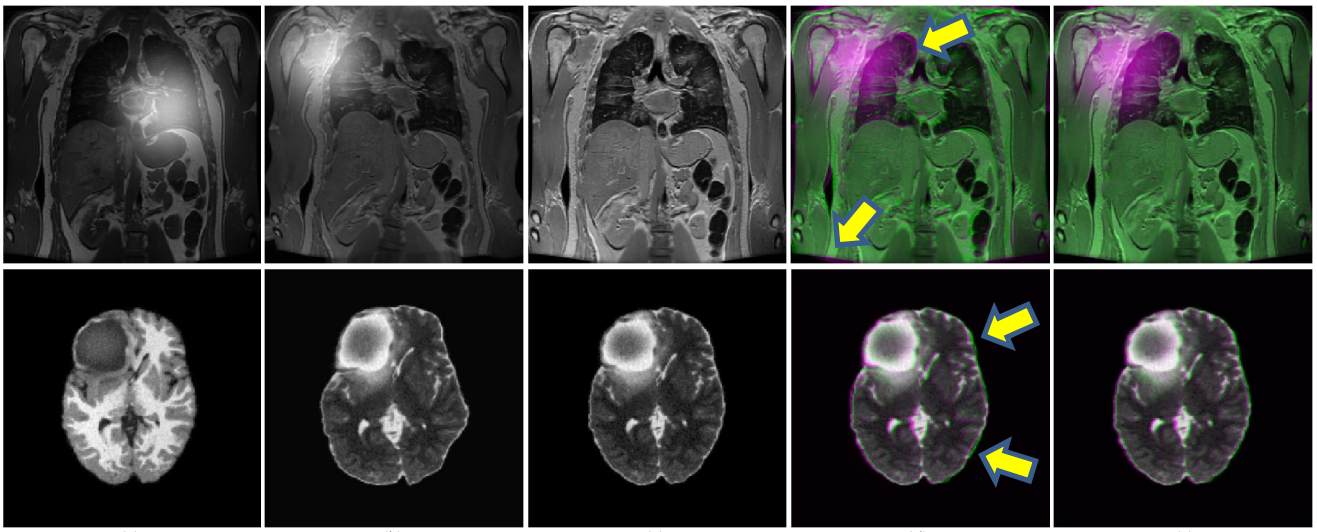
Table 1. Target registration error (in millimeters) obtained over the 10 cases of thorax CT-scans for all tested experimental conditions.

	BEFORE	MIND	MIND-LeM
TRE	8.46(6.58)	2.14(3.71)	1.58(2.58)

V. CONCLUSIONS AND FUTURE WORK

In this paper, a new descriptor termed MIND-LeM was developed for structural representation of images to be registered. MIND-LeM exploits the Log-Euclidean metric for SPD matrices consisting of patches, thereafter forming the multi-dimensional similarity metric. Its potential was illustrated in experimental results, showing the advantages of the proposed method over state-of-the-art techniques both in quantitative and qualitative perspectives.

In the current work, we present a simple class of the LeM descriptor by only considering the intensity difference in two-pixel neighbor region as image feature to form the covariance matrix and apply Gaussian weight of small patch to approximate it. In the forthcoming study, higher-order features will be investigated in the construction of covariance matrix such as to better represent the local image correlation of image cues, subsequently leading to better similarity metric for registration. In particular, it was stated in [24] that, in the procedure of computing the logarithm of the covariance matrix, there exists closed-form solution for eigen-decomposition of matrix with size d =2 and 3. Hence it provides a promising pave for exploring the high-order local features.



(a) (b) (c) (d) (e)

Fig. 4. Registration comparison of T1-PD and T1-T2 pair. (a) fixed image, (b) moving image, (c) source/reference of the moving image, (d)(e) registration display between the reference and moving image obtained by MIND and MIND-LeM.

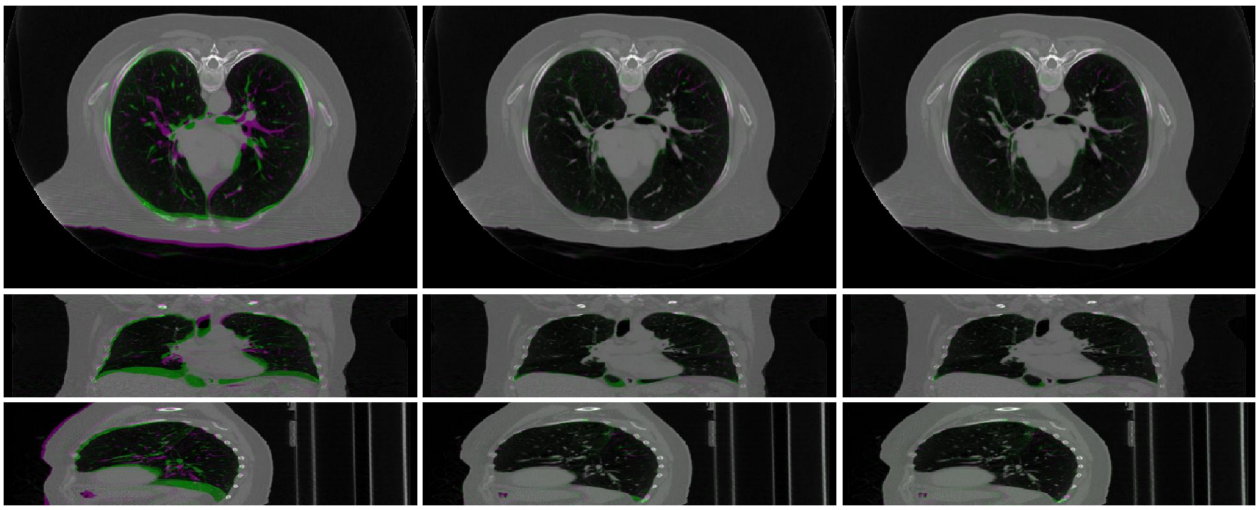


Fig. 6. Deformable registration result for Case 8 of the CT dataset. Top: axial, middle: sagittal and bottom: coronal plane. Left row: before registration, center and right row: after registration using MIND and the proposed MIND-LeM technique. The target image is displayed in magenta and the source image in green (complementary color).

VI. REFERENCES

[1] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, 16, 187–198, 1997.

[2] O. Sadowski, Z. Yaniv, L. Joskowicz, "Comparative in-vitro study of contact and image-based rigid registration for computer-aided surgery," *Computer-Aided Surgery*, 7 (4), 223-236, 2002.

[3] V. Walimbe, O. Dandekar, F. Mahmoud, R. Shekhar, "Automated 3D elastic registration for improving tumor localization in whole-body PET-CT from combined scanner," in: *Proc. IEEE EMBS*, vol. 1, pp. 2799-2802, 2006.

[4] P. Viola, W. Wells III, "Alignment by maximization of mutual information," *Int J. Comput. Vis.*, 24, 137–154, 1997.

[5] M. P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. V. Gleeson, S. M. Brady, and J. A. Schnabel, "MIND: modality independent neighbourhood descriptor for multi-modal deformable registration," *Med. Image Anal.*, 16(7), 1423-1435, 2012.

[6] M. Mellor, M. Brady, "Phase mutual information as a similarity measure for registration," *Med. Image Anal.*, 9, 330–343, 2005.

[7] E. Haber, J. Modersitzki, "Intensity gradient based registration and fusion of multi-modal images," *MICCAI* 2006, 726–733, 2006.

[8] J. P. W. Pluim, J. B. A. Maintz, M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imag.*, vol. 19, no. 8, pp. 809-814, 2000.

[9] R. Shams, P. Sadeghi, R. A. Kennedy, "Gradient intensity: A new mutual information based registration method," in: *Proc. IEEE CVPR* Workshop on Image Registration and Fusion, June 2007.

[10] C. Beijing, L. JunLi, C. Gang, "Study of medical image registration based on second-order mutual information," In: *Proc. IEEE ICME*, pp. 956–959, 2007.

[11] L. Kubecka, J. Jan, "Registration of bimodal retinal images–improving modifications," In: *IEEE Int. Conf. on Engineering in Medicine and Biology*, pp. 1695–1698, 2004.

[12] M. Staring, U. A. van der Heide, S. Klein, M. A. Viergever, J. P. W. Pluim, "Registration of cervical MRI using multifeature mutual information," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1412-1421, Sep. 2009.

[13] W. Jonghye, M. Stone, J. L. Prince, "Multimodal registration via mutual information incorporating geometric and spatial context," *IEEE Trans. Image Process.*, vol. 24, no. 2, pp. 757-769, Feb. 2015.

[14] X. Zhuang, S. Arridge, D. Hawkes, S. Ourselin, "A nonrigid registration framework using spatially encoded mutual information and free-form deformations," *IEEE Trans. Med. Imag.*, 30, 1819–1828, 2011.

[15] H. Rivaz, Z. Karimaghaloo, and D. L. Collins, "Self-similarity weighted Mutual information: A new nonrigid image registration metric," *Med. Image Anal.*, vol. 18, no. 2, pp. 343–358, Feb. 2014.

[16] S. Thiruvenkadam, "Dense multi-modal registration with structural integrity using non-local gradients," *VISAPP*, (1) 2013:258-263.

[17] Z. Li, L. J. van Vliet, and F. M. Vos, "Image registration based on autocorrelation of local structure," *IEEE Trans. Med. Imag.*, vol. 35, no. 1, pp. 63-75, Jan. 2016.

[18] B Denis de Senneville, C Zachiu, M Ries and C Moonen, "EVolution: an edge-based variational method for non-rigid multi-modal image registration," *Phys. Med. Biol.*, pp. 7377-7396, 2016.

[19] C. Wachinger, N. Navab, "Entropy and laplacian images: structural representations for multi-modal registration," *Med. Image Anal.*, 16, 1–17, 2012.

[20] G. Piella, "Diffusion maps for multimodal registration," *Sensors*, 14(6) (2014) 10562–10577.

[21] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Fast and simple computations on tensors with Log-Euclidean metrics," Research Report RR-5584, INRIA, May 2005.

[22] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-Euclidean metrics for fast and simple calculus on diffusion tensors," *Magn. Reson. Med.,* 2006.

[23] X. Pennec, P.Fillard, and N. Ayache, "A Riemannian framework for tensor computing," *Int J. Comput. Vis.*, 2006.

[24] P. Li, Q. Wang, "Local log-Euclidean covariance matrix (L2ECM) for image representation and its applications", *Proc. 12th Eur. Conf. Comput. Vis.*, pp. 469-482, 2012.

[25] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi, "Kernel methods on the Riemannian manifold of symmetric positive definite matrices," in *Proc. IEEE Conf. CVPR,* 2013.

[26] A. Buades, B. Coll, J.M. Morel, A non-local algorithm for image denoising, in: *Proc. IEEE CVPR*, June 2005, pp. 60–65.

[27] Q. Liu, J. Liu, P. Dong, D. Liang, "SGTD: Structure gradient and texture decorrelating regularization for image decomposition," *Proc. Conf. ICCV*, pp. 1081-1088, Dec. 2013.

[28] R. C. Gonzalez, R. E. Woods, Digital Image Processing. 3rd ed., Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.

[29] B. Avants, C. Epstein, M. Grossman, J. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Med. Image Anal.*, 12, 26–41, 2008.

[30] A. Myronenko, X. Song, "Intensity-based image registration by minimizing residual complexity," *IEEE Trans Med. Imag.,* 29 (11) (2010) 1882–1891.

[31] M. Ackerman, "The visible human project," *Proceedings of the IEEE*, 86, 504–511, 1998.

[32] MICCAI 2012 Challenge on Multimodal Brain Tumor Segmentation. Available online: http://www.imm.dtu.dk/projects/BRATS2012.

[33] R. Castillo *et al.*, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," *Phys. Med. Biol.*, vol. 54, no. 7, pp. 1849–1870, Apr. 2009.