LPPCO: A Novel Multimodal Medical Image Registration Using New Feature Descriptor Based on the Local Phase and Phase Congruency of Different Orientations

LI ZHANG1, BIN LI1, (Member, IEEE), LIANFANG TIAN1, (Member, IEEE), and WENBO ZHU2

1School of Automation Science and Engineering, South China University of Technology, Guangzhou 510000, China
2Department School of Automation, Foshan University, Foshan, China
Corresponding author: Bin Li (E-mail: binlee@scut.edu.cn).

This work was supported by National Natural Science Foundation of China (61305038, 61273249), the Public Science and Technology Research Funds Projects of Ocean (201505002), the Fundamental Research Funds for the Central Universities, SCUT (No.2015ZZ028), the National Engineering Research Center for Tissue Restoration and Reconstruction and the Guangdong Key Laboratory for Biomedical Engineering (SCUT of China), Key Laboratory of Autonomous Systems and Network Control of Ministry of Education (SCUT of China), the Natural Science and Characteristic Innovation Project of the Department of Education of Guangdong Province (2017KTSCX194).

ABSTRACT At present, many feature descriptors-based registration methods have been proven to be robust in case of complex intensity distortions. However, most of these feature descriptors are only related to intensity information in a patch of neighboring pixels and ignoring the neighbor orientation information, which make the registration performance for medical images with low-resolution appear to be weak robustness and low accuracy. To improve the robustness and accuracy, a novel feature descriptor, named Local-Phase mean and Phase-Congruency values of different Orientations (LPPCO), is developed using filter-bank of Log-Gabor filters at different orientations and frequencies. Next, a similarity measure named \( nccLPPCO \) is developed using the normalized cross correlation (NCC) of the LPPCO descriptors, followed by a fast template matching techniques for detecting correspondences between the different images. Additionally, a more sensitivity of phase deviation function is presented for the calculation of phase congruency. The main steps of constructing the similarity measure \( nccLPPCO \) include: firstly, we extract the local phase mean and phase congruency values for each pixel in each orientation; secondly, local phase mean orientation histograms and phase congruency values orientation histograms over all the pixels are computed respectively, where the maximum response values are chosen to vote for the corresponding bin; thirdly, combining the two resulting histograms obtains the feature descriptor LPPCO; finally, the LPPCO descriptor is compared across images using NCC. Experimental results show that \( nccLPPCO \) is robust to complex intensity distortions between multimodal medical images and outperforms other similarity metrics such as NCC and DLSC. Furthermore, \( nccLPPCO \) based registration algorithm preformed on various types of multimodal medical image pairs shows that it outperforms the NMI-based and DLSC-based registration methods in the registration robustness and accuracy.

INDEX TERMS Local phase, Feature descriptor, Phase congruency, Image registration, Template matching

I. INTRODUCTION

Medical image processing techniques have been widely applied in clinic diagnosis and treatment. Frequently, images acquired using different modalities need to be combined for obtaining comprehensive patient information, to improve the reliability and efficiency of clinic diagnosis and treatment level. During the combination of comprehensive patient information, the robust and accurate image registration is one of the critical technologies [1]. Hence, the study of multimodal image registration has great values both in theory and clinical application. One common example of clinical application is the registration between positron emission tomography (PET) and computed tomography (CT), as it can combine the functional metabolism information provided by PET with the anatomical structural information provided by CT [2]. The primary objective of image registration is to find control points (CPs) or correspondences between different
image acquisitions in space and structure. However, the registration between images with different modalities is made difficult by the effect of significant nonlinear intensity differences, which may appear in both images due to different image acquisition protocols, organ specularity, and vignetting artifacts. In general, image registration mainly consists of the following three parts: feature space, similarity metric, and transformation model. The detail process first extracts the key features from both images to be registered, and then match them based on one similarity measure and search the optimal parameters of the transformation model.

A robust feature for multimodal medical image registration should be stable enough to be repeatedly detected and reflect the shared anatomical structure features between images. Common image features can be divided into three primarily types: point features [3, 4], line features [5, 6], and region features [7]. In recent years, local structure invariant features have been extensively adopted and rapidly developed in the medical image processing techniques. Some representative local structure invariant features such as the fast retina keypoint (FREAK) [8], the speeded up robust features (SURF) [9], the scale invariant feature transform (SIFT) [10] and the speed-up binary robust invariant scalable keypoint (SBRISK) [11] have been widely used for image registration because of their invariance to image scale and rotation changes. Nevertheless, these methods are not effective for multimodal medical image registration, because they lack robustness to complex intensity distortions and hard to obtain highly repeatable shared features between images with different modality (e.g., CT, MRI).

To overcome the problem of sensitivity to nonlinear intensity differences, many frequency-based features have been introduced in recent research. Mouats et al. proposed an edge histogram descriptor combined with phase congruency (PCEHD) incorporating frequency and shape information [12]. Phase congruency (PC), the feature operator used in PCEHD, outperformed SIFT in contrast changes [12]. Aguilera et al. [13] presented a Log-Gabor histogram descriptor (LGHD). LGHD was suitable to the task of matching features points between images with complex intensity differences because it also combined frequency and spatial information. LGHD obtained a better matching performance concerning intensity variations than PCEHD, but it was mainly suited for images rich in texture [13]. Ye et al. introduced a histogram of orientated phase congruency (HOPC) descriptor, which reflected the structural attributes of images [14-16]. The HOPC was proved to be independent of the gray distribution pattern between two images and robust to image intensity changes [15]. However, a poor matching performance of HOPC may arise if the images to be matched consist of few shape or structure information.

Therefore, to achieve an intensity invariant feature descriptor for medical images with low-resolution, refinements of the above-mentioned descriptors and additional descriptors are badly needed. Due to the invariance of frequency-based features to complex intensity distortions, local phase mean and phase congruency values of different orientations are applied to build an intensity invariant feature descriptor in our work. The local phase means at different orientations derived from the multi-scale and multi-oriented representation of its monogenic signal, which captures boundary structure feature with higher distinctiveness than other boundary features such as Sobel edge and is hardly affected by intensity distortions. Additionally, the averaged features on different scales reduce the effect of add-Gaussian noise on the estimated structure and orientation information. The phase congruency values at different orientations is another structure feature that provides an absolute measure to detect the significance of the feature. It is relatively independent of the magnitude of an image signal, making it robustness to non-linear intensity distortions which are dominant in the case of multimodal medical image pairs. To achieve a more localized response to features for allowing a more effective detection of the detail of images, a more sensitive phase deviation term is defined by giving greater weight to its sine function term for the calculation of phase congruency. To reduce the impact of low-resolution and complex intensity distortions during the image registration, the local phase mean features and the phase congruency values features are combined to develop a new feature descriptor LPPCO. The proposed descriptor uses two complementary structure representations and neighbor orientation information, which can cover the needs of the application on feature-based registration methods.

The choice of similarity metrics is vital for the accuracy and robustness of image registration. Common similarity metrics of medical image registration include the normalized mutual information (NMI) [17] and the normalized cross correlation (NCC) [18], etc. However, NCC is probably invalid for the multimodal image registration due to significant nonlinear intensity distortions. By comparison, NMI is more robust to complex intensity distortions. Unfortunately, NMI is computationally expensive as it needs to calculate the joint histogram for each window and is greatly influenced by the window size for template matching [19, 20]. To improve their robustness, these measures have been applied on image feature descriptors such as the HOPC<sub>α</sub> [15], the NCC of dense local self-similarity descriptors (DLSC) [14], the Modality independent neighborhood descriptor (MIND) [21], and the self-similarity weighted α-mutual information (SeSαMI) [22]. Inspired by these methods, to achieve a good matching performance of both robustness and accuracy for medical images with low-resolution, the NCC of the proposed feature descriptors LPPCO is used to build a new similarity measure (named LPPCO<sub>α</sub>) that represents the shape and structural similarity between images.

The contributions of this paper are as follows. Firstly, a new feature descriptor LPPCO based on local phase mean
and phase congruency values of different orientations is proposed. To enhance the contour or edge features, a more sensitivity of phase deviation function is proposed for the calculation of phase congruency. Then, the NCC of the proposed feature descriptor LPPCO is used to build a new similarity measure (named LPPCO\(_{ncc}\)). Finally, we design a fast template matching scheme based on LPPCO\(_{ncc}\) to detect CPs between images.

The remainder of the paper is divided into five sections. In Section II, the proposed similarity metric LPPCO\(_{ncc}\) is described in detail. In Section III, a robust template matching scheme is presented based on LPPCO\(_{ncc}\). In Section IV, the matching performance of LPPCO\(_{ncc}\) is evaluated. In Section V, the accuracy and robustness of the proposed registration algorithm are verified. Finally, the paper gives the conclusions in Section VI.

II. LPPCO\(_{ncc}\): FEATURE SIMILARITY METRIC

At present, many feature descriptors-based similarity metrics have been proven to be robust in the case of complex intensity distortions. However, most of these feature descriptors are only related to intensity information in a patch of neighboring pixels and ignoring the neighbor orientation information, which makes the registration performance for medical images with low-resolution appear to be weak robustness and low accuracy. To reduce the impact of low-resolution and complex intensity distortions during the image registration, the local phase mean features and the phase congruency values features are combined to develop a new feature descriptor named LPPCO. The proposed descriptor uses two complementary structure representations and neighbor orientation information, which can cover all the needs of the application on feature descriptors-based similarity metric. Since the structure and shape in the medical images with different modalities are similar despite imaging with different tissue characteristics, the NCC of the LPPCO is taken as the similarity measure (named LPPCO\(_{ncc}\)) for multimodal medical image registration.

In the section, we first present a new feature descriptor LPPCO. Then, a new feature similarity metric based on this descriptor is generated. The proposed similarity metric involves two complementary local structural features: the mean of local phase and phase congruency derived from the multi-scale image at a given direction (Section II. A). Next, combining the two features creates the descriptor named LPPCO (Section II. B). Then, the similarity metric named LPPCO\(_{ncc}\) is described (Section II. C). Finally, a fast template matching strategy based on the new similarity measure is proposed to detect CPs between multimodal medical images (Section II. D).

A. REPRESENTATION OF LOCAL STRUCTURE FEATURE

Recently, a large variety of local structure feature descriptors have been proposed in recent years, such as SIFT, Harris [3], gradient location and orientation histogram (GLOH) [18] and FREAK. Unfortunately, the above-mentioned descriptors based on gradient distributions are ineffective for multimodal medical image registration, because of its sensitivity to the significant nonlinear intensity differences. By comparison, local phase and phase congruency developed to represent local structure features are largely independent of intensity and highly robust in the presence of image non-homogeneities. The use of these features has significant advantages over the gradient-based methods. In order to robust against variations in image intensity, a novel feature descriptor named the local phase mean and phase congruency values features are combined to different orientations (LPPCO in short) is developed using filter-bank of Log-Gabor filters at different orientations and frequency scales, whose 2-D representation in the log-polar Fourier domain is shown in Eq. (1).

\[
G(\rho, \theta, s) = \exp \left( -\frac{1}{2} \left( \frac{\rho - \rho_s}{\sigma_\rho} \right)^2 \right) \exp \left( -\frac{1}{2} \left( \frac{\theta - \theta_o}{\sigma_\theta} \right)^2 \right)
\]

Here, \((\rho, \theta)\) denote the log-polar coordinates. \((\sigma_\rho, \sigma_\theta)\) denote the angular and radial bandwidths, respectively. \((\rho_s, \theta_o)\) represent the frequency center of the filters. The variable \(s\) and \(o\) are the scale and orientation selection, respectively. In addition, the real and imaginary parts of the Log-Gabor filter, acquired using the inverse Fourier transform to the Log-Gabor function, are described as the Log-Gabor odd-symmetric (sine) \(M_\rho^s\) and even-symmetric (cosine) \(M_\rho^e\) wavelets at a given scale \(l\), respectively.

If we assume \(I\) denote an image signal, a response vector can be generated by convolution of an image signal and the Sine-Cosine wavelets, as shown in Eq. (2).

\[
[c_i(x), o_i(x)] = [I(x)*M_\rho^s, I(x)*M_\rho^e]
\]

Here, \(M_\rho^s\) and \(M_\rho^e\) respectively denote the sine wavelets and cosine wavelets at a given scale acquired using the inverse Fourier transform to the Log-Gabor filter. \(c_i(x)\) and \(o_i(x)\) are the responses of \(M_\rho^s\) and \(M_\rho^e\) over a given scale \(l\) at the location \(x\), respectively. The amplitude \(A_i\) and phase angle \(\phi_i\) of the image signal at a given Log-Gabor wavelet scale \(l\) can be represented by Eq. (3) and Eq. (4), respectively.

\[
A_i(x) = \sqrt{c_i(x)^2 + o_i(x)^2}
\]

\[
\phi_i(x) = \tan^{-1}\left(\frac{c_i(x)}{o_i(x)}\right)
\]
vectors are the basis of the local representation of the signal. They can be used to calculate the phase congruency over different scales in the direction $\theta_k$, as shown in Eq. (5).

$$PC(x, \theta_k) = \frac{\sum_{j=1}^s W(x, \theta_j) [A_j(x, \theta_j) \Delta \Phi_j(x, \theta_j) - T]}{\sum_{j=1}^s A_j(x, \theta_j) + \varepsilon}$$ (5)

Here, $W$ is the weight of each pixel including the given wavelet scale, which can be used to avoid frequency spread. $A_j$ and $\phi_j$ represent the amplitude and phase of the image signal at the $j$th wavelet scale respectively. The operation symbol $\lfloor f \rfloor$ is non-negative. $\lfloor f \rfloor = f$ when $f > 0$, and $\lfloor f \rfloor = 0$ otherwise. $s$ is the total number of wavelet scales. $\varepsilon$ is a small constant used to remove division by zero. $k$ is the total number of wavelet directions. $T$ denotes a noise threshold designed to avoid the noise effect. The measure are counted in the result only for energy values above the noise threshold $T$. In Eq. (5), $\phi$ represents the local phase mean in the direction $\theta_k$, it is calculated over multiple wavelet scales using Eq. (6).

$$\phi(x, \theta_k) = a \tan \left( \frac{\sum_{j=1}^s e_j(x)}{\sqrt{(\sum_{j=1}^s e_j(x))^2 + (\sum_{j=1}^s o_j(x))^2}} \right)$$ (6)

$\Delta \Phi_j$ is a more sensitive phase deviation term than the traditional one, it is defined as the Eq. (7).

$$\Delta \Phi_j(x, \theta_j) = \cos(\phi(x, \theta_j) - \phi(x, \theta_j)) - \xi^2 \sin(\phi(x, \theta_j) - \phi(x, \theta_j))$$ (7)

Where $\xi (\xi > 1)$ denotes the newly added coefficient of the phase deviation sine function. Making use of the coefficient $\xi$ will increase the sensitivity of phase deviation function. The function $\Delta \Phi$ falls away much more rapidly than that of the traditional one, which means changing the sine coefficient is helpful to generate a rather sensitive measure of phase congruency. Fig. 1 plots the function at $\xi = \sqrt{2}$ and $\xi = 2$ along with the traditional phase deviation function with a comparison. Fig. 2 gives an example of how a more highly localized response to features is produced by the PC measure, which allows a more effective detection of the detail in an image.

The proposed method uses local phase mean (see Eq. (6)) and phase congruency values (see Eq. (5)) over multiple scales at a given orientation for constructing significant structural representation. The local phase mean provides mainly contour and shape features of the image [23]. Furthermore, it is invariant to the variations of image intensity and insensitive to the image distortion [24]. Additionally, the effect of add-Gaussian noise on the estimated phase information is reduced by averaging over different scales. Phase congruency values over multiple scales in a given orientation is another structure feature that provides an absolute measure to detect the significance of the feature. Moreover, it is relatively independent of the
magnitude of an image signal, making it robustness to nonlinear intensity differences and global trends which are dominant in the case of multimodal medical image pairs. Phase congruency is in the range between 0 and 1, which assume that features are present at points where the Fourier components are maximally in phase, and they can be applied to detect features of different strength. Fig. 3 illustrates the phase congruency values calculated from an MR-T2 image in six directions including 0 (horizontal), 30, 60, 90 (vertical), 120, and 150 degrees.

**B. LPPCO DESCRIPTOR IMPLEMENTATION**

Indeed, a mere local phase mean representation used in multimodal medical image registration can still suffer from noise and be limited because of no information about the structural significance of images. Neither phase congruency values of different orientations nor local phase mean by itself describes the coherence of anatomical structure characteristics in a neighborhood. Based on a combination of the local phase mean representation and the phase congruency of different orientations representation, a new descriptor representation is proposed named LPPCO. The LPPCO descriptor is calculated based on LGHD presented by Aguilera et al. [13]. The three major procedures to extract the LPPCO descriptor are presented as follows.

1) Select a template window in one image and then calculate the local phase mean and phase congruency values for each pixel at given orientation in the template window. Local phase mean and phase congruency values respectively are grouped into four orientations.

2) Divide the template window into 16 smaller sub-regions (4×4), where local phase mean histograms and phase congruency histograms are computed for every sub-region, respectively. Within each sub-region, two orientation histograms of four bins are computed using the maximal of the local phase mean and phase congruency values, respectively.

3) Accumulate the two histograms of each sub-region (64×2) and combine the two resulting histograms to obtain 128 bins feature vector that can be applied to template match algorithm.

The LPPCO descriptor can be more stable in describing the shared characteristics between images and be more effective to avoid being affected by nonlinear intensity differences than the LGHD descriptor. Fig. 4 gives an example of the resulting histogram calculated for each sub-region from two different medical imaging modalities. A pair of MR-T1 and MR-T2 brain images in the Fig. 4: The MR-T1 image (Fig. 4a) and MR-T2 image (Fig. 4b) are obtained simultaneously by one scan session. Therefore, the structures in these two images should be aligned perfectly although the modalities are different [25]. It can be seen from the Fig. 4c—Fig. 4f that there are more strong similarities between the two LPPCO histograms than the LGHD despite the different intensity distribution of patterns of images from different modalities. To explicitly show the amount of overlap between the two statistical histograms, Bhattacharyya coefficient is used to quantify the relative closeness, as shown in Eq. (8). The higher BC values the two statistical histograms have, the better the relative closeness is. The BC value required between the LGHD histograms is 0.9667, and the BC value required between the LPPCO histograms is 0.9946 which is higher than that of the LGHD. It proves that the LPPCO histogram is more robust to complex intensity differences than the LGHD.

\[
BC = \sum_{i=1}^{n} \sqrt{H_1 \cdot H_2}
\]

Here, \(H_1\) and \(H_2\) respectively denote the \(n\)-bin histogram calculated for each sub-region from two different medical imaging modalities.

**C. SIMILARITY MEASURE BASED ON LPPCO DESCRIPTOR**

LPPCO descriptors capture the structural and neighbor orientation features of images. Since these features are robust to complex intensity differences, the LPPCO descriptor is suitable for multimodal medical image registration because they have similar shapes. Thus, the NCC of the LPPCO descriptors can be used to build a new
similarity measure (named LPPCO\textsubscript{ncc}) for image registration, as defined in Eq. (9).

\[
\text{LPPCO}\textsubscript{ncc} = \frac{\sum_{k=1}^{n} (H_F(k) - \bar{H}_F)(H_M(k) - \bar{H}_M)}{\sqrt{\sum_{k=1}^{n} (H_F(k) - \bar{H}_F)^2} \sqrt{\sum_{k=1}^{n} (H_M(k) - \bar{H}_M)^2}}
\]

(9)

Where \(H_F\) and \(H_M\) respectively denote the LPPCO histogram descriptors of the image region \(F\) and \(M\). \(\bar{H}_F\) and \(\bar{H}_M\) are the means of \(H_F\) and \(H_M\), respectively. To try to make it clear what \(H_F\) and \(H_M\) mean, a figure is given to illustrate the main processing for calculating the similarity measure LPPCO\textsubscript{ncc}, as shown in Fig. 5.

![Fig. 5 Main processing for calculating the similarity measure LPPCO\textsubscript{ncc}](image)

To illustrate the superiority of the proposed similarity measure in multimodal medical image registration, LPPCO\textsubscript{ncc} is compared with NCC and DLSC [14] by the similarity curve, which can make a qualitative analysis of the matching performance of similarity measures [15]. A pair of MR-T2 and MR-T1 brain images obtained simultaneously by one scan session is chosen in the test, whose structures are aligned perfectly. A template window is first selected from the MR-T2, where the window size is 100×100 pixels. Then, NCC, DLSC and LPPCO\textsubscript{ncc} are calculated for the translation (-10 to 10 pixels) along x-axis direction in a predefined search region (e.g., 20×20 pixel) of the second image (i.e., MR-T1 image), sequentially. Fig. 6 gives the similarity curves of NCC, DLSC and LPPCO\textsubscript{ncc}. It can be observed that NCC is unable to detect the control point, and DLSC exhibits some location errors caused by the complex nonlinear intensity changes. By comparison, LPPCO\textsubscript{ncc} not only accurately discovers the right CP but also obtains a more smooth similarity curve and more distinct curve peak. The example shows that LPPCO\textsubscript{ncc} is the most robust against complex intensity distortions than other similarity measures. A further analysis of the matching performance of LPPCO\textsubscript{ncc} will be presented in the section V.

![FIGURE 6. Similarity curves of NCC, DLSC and LPPCO\textsubscript{ncc} are compared.](image)

**D. FAST TEMPLATE MATCHING SCHEME**

The process of fast template matching based on LPPCO\textsubscript{ncc} is divided into three steps. The first step is to determine a template window in the fixed image. The second step is to search for a correspondence window within a small search region of the moving image using the proposed similarity metric LPPCO\textsubscript{ncc}. It is obvious that the LPPCO descriptors of search window pixel-by-pixel need to be computed in the matching process between each pair of template windows. Even worse, there are many overlapping pixels between adjacent template windows which requires a large number of repetitive computations. In order to reduce the repeated calculation and accelerating the template matching algorithm, we first define a block region, which is centered by each pixel within a small search region of the moving image, and then the LPPCO descriptors for each block (named a block-LPPCO descriptor) is extracted. More specifically, each pixel can first extract a feature vector (i.e., a LPPCO descriptor), and all the extracted feature vectors are ranged in pixel index order to form a 3D descriptor for the entire image (named the block-LPPCO image). In the third step of the matching process, the block-LPPCO descriptor is collected at a certain
interval sampling (e.g., 1/2 block width) to construct the LPPCO descriptor for the template window.

Much of repetitive computation for adjacent template windows can be eliminated using our proposed template matching scheme, which will be good for accelerating the speed of the template matching. In the following, the calculating efficiency of our template matching scheme will be discussed by comparing with the traditional template matching scheme. In the template matching processing, each template window \((N \times N\) pixels) corresponds to a search window \((M \times M\) pixels). The traditional scheme spends \(O(M^2N^2)\) operations due to the fact the template window slides pixel-by-pixel along a search window. In contrast, the calculation cost taken from our template matching scheme mainly consists of two aspects: (1) calculate the block-LPPCO descriptors for all pixels within the search window of size \((M + N)^2\) pixels and (2) collect the block-LPPCO descriptor at a certain interval sampling for all the template window applied to match. The calculation cost is negligible in the second aspect, since it simply collects the block-LPPCO descriptor at a certain interval sampling. The first aspect requires \(O\left(\tau(M+N)^2\right)\) operations, where \(\tau\) represents the time taken from calculating the block-LPPCO descriptor for a pixel. In all the experiments of this paper, \(M\) is in a range from 60 to 120 pixels, \(N\) is in a range from 10 to 50 pixels. The calculating time \(\tau\) doesn’t exceed 1s in our experiment results. By the time expression comparison, our scheme is of great advantage to the traditional scheme in the large size of template or search window because the time ratio \(\frac{\tau(M+N)^2}{M^2N^2} = \tau\left(\frac{1}{M} + \frac{1}{N}\right)^2\) is far less than 1.

III. MULTIMODAL REGISTRATION ALGORITHM BASED ON LPPCO_ncc

In the section, a new robust registration algorithm based on LPPCO_ncc is presented for multimodal medical images, which comprises the following five main steps. The flow diagram of our proposed registration algorithm is presented in Fig. 7. In the flowchart, two images from different modalities (e.g., MR-T1, MR-T2, MR-PD, or CT) are taken. The image to be registered is called the moving image and the other image is regarded as the fixed image.

![Flow diagram of the proposed registration algorithm.](image)

**FIGURE 7.** Flow diagram of the proposed registration algorithm.

**Step 1:** Search the interest points inside the fixed image. This is done through the block-based SURF detector, which outperforms most of the other corner detectors with respect to robustness, repeatability, and distinctiveness, and can be computer faster. In order to generate the interest points with uniform distribution, a block-based SURF searching approach is designed. Firstly, the fixed image is partitioned into \(s \times s\) non-overlapping blocks of fixed size. Secondly, the candidate interest points are calculated using the SURF in each block. Lastly, \(n\) interest points are selected from candidate points through equal index interval sampling method, where the sampling point index interval is equal to the total number of candidate points divided by \(n - 1\). If the maximum number of candidate points is less than \(n\), all candidate points are chosen as the interest point sets. Therefore, approximately \(s \times s \times n\) interest points are obtained from the fixed image. An example of interest points obtained using the block-based SURF detector is shown in Fig. 8b. By comparison with the traditional SURF detector (Fig. 8a), the block-based strategy obtains the uniform distribution interest points over the image.

**FIGURE 8.** The interest points extracted by two different schemes: (a) traditional SURF detector using fixed threshold and (b) block-based SURF searching strategy where the CT image is separated in 8x8 blocks.

**Step 2:** Find corresponding control point pairs between the fixed image and moving image. LPPCO_ncc is utilized to search control points using the proposed fast template matching strategy within a small search window of the
moving image. In detail, for a given interest point \( p_1 \) in the fixed image, its corresponding interest point \( p_2 \) in the moving image is obtained by the maximum of \( \text{LPPCO}_{\text{ncc}} \) between the template window in the fixed image and the search window in the moving image.

Step 3: Choose an appropriate global transformation model for the consistency check. Because of existing many uncertainty elements, such as noise, shadow or occlusion, the error between above obtained CPs is unavoidable. To remove the CPs with large errors, the key for the consistency check is to choose a right transformation model according to the type of relative spatial distortions between images. Since the actual distortions between the multimodal brain images are a little complex, the projective model is selected for the consistency check to handle common global transform in this paper.

Step 4: Remove the mismatched CPs through an iterative refining procedure. Firstly, a projective model using the least squares method is set up with all the CPs. Then, the root mean square error (RMSE) and the residual error of CPs are calculated. Next, the CP pair with the maximum residual error is removed. The aforementioned steps are repeated until the RMSE is no larger than a predetermined threshold.

Step 5: Specify a spatial transform to correct the moving image. After the removal of mismatched CPs, specifying a spatial transform is needed for correcting the moving image. In this paper, the piecewise linear (PL) transform model is selected because it can effectively handle the local deformations caused by different modes.

IV. EXPERIMENTAL RESULTS: EVALUATING MATCHING PERFORMANCE OF LPPCO_{ncc}

To evaluate the matching performance, \( \text{LPPCO}_{\text{ncc}} \) is compared with NMI and DLSC within the same registration framework. To analyze the accuracy and effectiveness, quantitative comparisons of the two metrics are given using different combinations of medical image multimodal, including the correct match rate (CMR) and the calculation efficiency. In this section, the experiments primarily focus on the following two goals: 1) study on the effects of the various parameters for \( \text{LPPCO}_{\text{ncc}} \); 2) compare \( \text{LPPCO}_{\text{ncc}} \) with the advanced similarity metrics such as NMI and DLSC. In this section, two type datasets are used for the comparison in Section IV. A. Then, the details of the implementation are illustrated in Section IV. B. Subsequently, the experimental analysis is described in Section IV. C.

A. DATA SET DESCRIPTION

Two types of multimodal medical image pairs are applied to assess the matching performance of \( \text{LPPCO}_{\text{ncc}} \), including synthetic brain data with complex intensity distortions and clinical multimodal brain images.

1) Synthetic brain data sets: To generate synthetic moving images with the nonlinear intensity differences, we corrupted the fixed image according to the formula (10).

\[
I_{\text{Moving}}(x, y) = \frac{\alpha}{\sqrt{2\pi\sigma}} \exp\left(-\frac{I_{\text{fixed}}(x, y) - \xi}{2\sigma^2}\right) (10)
\]

Here, \( \alpha \) denotes a normalization constant of the formula, \( \xi \) and \( \sigma > 0 \) are parameters.

To validate the performance of the proposed methods against complex intensity distortions, five pairs of synthetic images generated by an MR-T2 image (256×256 pixels) are used to perform the registration experiments. The parameters for these experiments are: \( \sigma_1 = 0.15, \xi_1 = 0.20 \) (Fig. 9b), \( \sigma_2 = 0.15, \xi_2 = 0.44 \) (Fig. 9c), \( \sigma_3 = 0.20, \xi_3 = 0.39 \) (Fig. 9d), \( \sigma_4 = 0.25, \xi_4 = 0.8 \) (Fig. 9f). When \( \xi \) and \( \sigma > 0 \) are chosen, the synthetic moving images are generated, as shown in Fig. 9. It can clearly see that these synthetic moving images expose the significant nonlinear intensity distortions.

2) Clinical multimodal brain data sets including three groups of MR-T2 & MR-T1 image pairs, eight groups of MR-PD & MR-T2 image pairs and three groups of CT & MR-T1 image pairs. All the images shown in Fig. 10 are of size 256×256 pixels and rendered from Keith A. Johnson and J. Alex Becker of ‘The Whole Brain Atlas’ database. CT images are sensitive to higher tissue density and bone structures while MR describes parenchyma of the object of interest [26]. All of the image pairs have been captured simultaneously by one scan session. Therefore, the structures in these two images have no obvious difference in terms of their structural deformation. However, complex intensity differences are presented between multimodal medical images for the reason that they are captured from different imaging modalities.
distortions. For each image pair to be registered, 20-30 evenly distributed check points are manually labelled to determine the CM using a projective transformation model. Furthermore, RMSE of the correctly matched point pairs is calculated for accuracy evaluation, which is defined as the Eq. (11).

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{N} (x_F - \mu(x_M))^2 + (y_F - \mu(y_M))^2}{N}}
\]  

Here, \(N\) denotes the total number of the correctly matched point pairs, \((x_F, y_F)\) and \((x_M, y_M)\) represent the pixel location in the fixed and moving images, respectively. \(\mu\) is the transform model used in the matching process.

The LPPCO \(_{nc}\) measure has four primarily parameters, i.e., the total number of wavelet directions \(k\) (\(k = 2, 3, 4, 5, 6\)), the total number of wavelet scales \(s\) (\(s = 1, 2, 3, 4, 5\)), the noise threshold \(T\) (\(T = 1, 2, 3, 4, 5\)) and the coefficient of the sine of the phase deviation \(\xi\) (\(\xi = 1, \sqrt{2}, 2, \sqrt{2}, 4\)), where each parameter takes six values as shown in the parentheses following the parameter. Their effects on the performance of LPPCO \(_{nc}\) have been tested by two pairs of synthetic data sets and three pairs of clinical data sets. The CMR mean \(\overline{CMR}\) of all experimental results are shown in Table I. If one of the parameters is used as a variable, others will be taken a default value. In this paper, these default values are taken to be \(k = 5\), \(s = 4\), \(T = 1\) and \(\xi = 2\).

![CPs founded by LPPCO \(_{nc}\) using the template with size 100×100 pixels (clinical datasets). (a) MR-T2 to MR-T1, (b) MR-T2 to CT, (c) MR-T1 to CT.](image)

**FIGURE 10.** CPs founded by LPPCO \(_{nc}\) using the template with size 100×100 pixels (clinical datasets). (a) MR-T2 to MR-T1, (b) MR-T2 to CT, (c) MR-T1 to CT.

<table>
<thead>
<tr>
<th>Variable</th>
<th>v1</th>
<th>v2</th>
<th>v3</th>
<th>v4</th>
<th>v5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k)</td>
<td>0.7214</td>
<td>0.9032</td>
<td>0.9032</td>
<td><strong>0.9194</strong></td>
<td>0.887</td>
</tr>
<tr>
<td>(s)</td>
<td>0.258</td>
<td>0.548</td>
<td>0.8710</td>
<td><strong>0.9194</strong></td>
<td>0.887</td>
</tr>
<tr>
<td>(T)</td>
<td>0.9194</td>
<td>0.9194</td>
<td>0.9194</td>
<td><strong>0.9194</strong></td>
<td>0.9194</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(\overline{CMR})</th>
<th>v1</th>
<th>v2</th>
<th>v3</th>
<th>v4</th>
<th>v5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T)</td>
<td>0.9194</td>
<td>0.9194</td>
<td>0.9194</td>
<td><strong>0.9194</strong></td>
<td>0.9194</td>
</tr>
</tbody>
</table>

Testing results indicate that the LPPCO \(_{nc}\) with the parameters \(k = 5\), \(s = 4\) and \(\xi = 2\) obtained the optimal CMR value. Consequently, the above parameter settings are used for all subsequent experiments. Specifically, the change of noise threshold \(T\) has no effect on the performance of LPPCO \(_{nc}\). The noise compensation \(T\) applied in Eq. (5) is only effective to additive Gaussian noise, yet there are usually other noise in medical image. Thus, our proposed method has better robust for complex intensity distortions. However, the disadvantages require further improvement in the noise robustness, which is a major research field in our future work.

**C. EXPERIMENTAL ANALYSIS**
To verify the precision, LPPCO\textsubscript{ncc} is compared with two similarity measures NMI and DLSC. The curves of CMR values varying with template size are represented in Fig. 11.

![CMR curves](image)

**CMR values versus the template sizes of NMI, DLSC, and LPPCO\textsubscript{ncc} for the clinical multimodal image pairs. (a) MR-T2 to MR-T1. (b) MR-T2 to MR-PD. (c) MR-T2 to CT.**

Fig. 11 shows the comparative CMR values versus the template sizes of NMI, DLSC, and LPPCO\textsubscript{ncc} for the clinical multimodal image pairs in Fig. 10. It can be observed that LPPCO\textsubscript{ncc} performs the best in any template sizes, followed by DLSC. This indicates that the similarity metric capturing the local phase mean and phase congruency of different orientations (used for LPPCO\textsubscript{ncc}) is less sensitive against variations in image intensity and contrast changes compared with local self-similarity feature (used for DLSC). NMI obtains the lowest CMR values because it is very sensitive to the number of samples taken (determined by the template size). This is due to the relatively small template size which implies very sparse data in the MI bins [19]. Compared with NMI, DLSC represents a slight advantage because the anatomical structure and shape properties (used by DLSC) are more robust against the global characteristics of gray-level statistics (used by NMI). Furthermore, the CMR values of the three similarity measures increase as the template size increases. Also, the presented matching performance of these three similarity measures will be affected by different test datasets due to the differences in image intensity distribution. For MR-T2 to MR-T1 matching sets and MR-T2 to MR-PD matching sets obtained simultaneously by one scan session, LPPCO\textsubscript{ncc} achieves a large CMR values even in a small template size [Fig. 10 (a) and (b)]. This is because the structures in the two matching sets have no obvious difference in terms of their structural deformation and phase feature. Therefore, LPPCO\textsubscript{ncc} presents the most robust against the complex intensity differences than other similarity measures. For MR-T2 to CT matching sets, the CMR values of LPPCO\textsubscript{ncc} decreases obviously compared with the first two matching sets due to quite low-resolution CT image. This shows the performance of LPPCO\textsubscript{ncc} depends on anatomical characteristics presented by medical images and could decline when images contain few anatomical characteristics. Nevertheless, LPPCO\textsubscript{ncc} still has an advantage over other similarity metrics. In general, combined with the result of the upper experiment, LPPCO\textsubscript{ncc} is robust against the complex intensity distortions between multimodal medical images.

Computational efficiency is another key factor to evaluate the matching performance of similarity measures. Fig. 12 illustrates that the run time is taken from NMI, DLSC and LPPCO\textsubscript{ncc} versus the template size. The experiments have been implemented on an HP -2008 workstation with Intel Xeon(R) CPU E3-1225(3.30 GHz) and RAM 4 GB.
Table II reports the registration results using the template window size of 90×90 pixels for the synthetic brain data sets and clinical multimodal brain data sets, respectively. The average of the number of matched CPs and the RMSE values over all the synthetic image pairs and all the clinical multimodal image pairs are respectively represented in Table II.

<table>
<thead>
<tr>
<th>Category</th>
<th>Method</th>
<th>CPs(Num.)</th>
<th>RMSE(Pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD/T2</td>
<td>Proposed</td>
<td>117</td>
<td>0.5345</td>
</tr>
<tr>
<td></td>
<td>DLSC</td>
<td>104</td>
<td>0.6621</td>
</tr>
<tr>
<td></td>
<td>NMI</td>
<td>96</td>
<td>0.6823</td>
</tr>
<tr>
<td></td>
<td>Before-registration</td>
<td>25</td>
<td>1.2848</td>
</tr>
<tr>
<td>T1/T2</td>
<td>Proposed</td>
<td>98</td>
<td>0.7998</td>
</tr>
<tr>
<td></td>
<td>DLSC</td>
<td>81</td>
<td>0.9049</td>
</tr>
<tr>
<td></td>
<td>NMI</td>
<td>75</td>
<td>0.9848</td>
</tr>
<tr>
<td></td>
<td>Before-registration</td>
<td>25</td>
<td>1.4491</td>
</tr>
<tr>
<td>CT/T2</td>
<td>Proposed</td>
<td>51</td>
<td>0.7440</td>
</tr>
<tr>
<td></td>
<td>DLSC</td>
<td>39</td>
<td>0.9005</td>
</tr>
<tr>
<td></td>
<td>NMI</td>
<td>37</td>
<td>1.0377</td>
</tr>
<tr>
<td></td>
<td>Before-registration</td>
<td>25</td>
<td>3.0578</td>
</tr>
<tr>
<td>Synthetics</td>
<td>Proposed</td>
<td>75</td>
<td>0.8675</td>
</tr>
<tr>
<td></td>
<td>DLSC</td>
<td>63</td>
<td>1.4142</td>
</tr>
<tr>
<td></td>
<td>NMI</td>
<td>62</td>
<td>1.4142</td>
</tr>
<tr>
<td></td>
<td>Before-registration</td>
<td>25</td>
<td>1</td>
</tr>
</tbody>
</table>

It can be clearly observed in Table II that the registration method based on LPPCO<sub>rec</sub> achieves the lowest RMSE value and the most matched check points. One reason for this outcome is that the LPPCO<sub>rec</sub>–based registration algorithm obtains the highest accuracy than other registration algorithms, which is beneficial to the local structure and shape similarity captured by LPPCO<sub>rec</sub>. These characteristics are robust to complex intensity differences and image local distortions. The performance of DLSC mainly depends on shape properties of images, so a low accuracy is achieved by DLSC-based registration algorithm in multimodal medical images due to these images containing low-intensity contrast. NMI-based registration algorithm ignores spatial structure information and local characteristics of gray-level statistics, which may lead to a decline of registration robustness. Another reason for the result is that the LPPCO<sub>rec</sub>–based registration algorithm finds many more CPs than other registration methods, which benefits from the SURF for extracting the interest points with special properties of the structure, scale, and orientation, and the PL transform for handling the local distortions between images. On the whole, the LPPCO<sub>rec</sub>–based registration algorithm is effective for the multimodal medical image registration and outperforms DLSC and NMI in registration accuracy and robustness.

VI. CONCLUSION

In our work, a similarity measure named LPPCO<sub>rec</sub> for multimodal medical image registration is developed to deal with complex intensity distortions between images with low-

It can be clearly observed in the Fig. 12 that NMI takes the longest run time among the similarity metrics because NMI needs to calculate the joint histogram for each window pair to be matched. The run time of LPPCO<sub>rec</sub> is less than that of NMI within 100×100 pixels of template size. This is good for real application due to the fact the small template size requires short calculation time for image registration, and the CMR values of LPPCO<sub>rec</sub> increase slowly when the template size is larger than a certain range (e.g., 100×100 pixels). LPPCO<sub>rec</sub> requires a little more run time than DLSC primarily because LPPCO is a high dimension of feature vector, which can be accelerated by using FFT to improve the computational efficiency of LPPCO<sub>rec</sub>.

V. EXPERIMENTAL RESULTS: MULTIMODAL REGISTRATION

To evaluate the registration accuracy, LPPCO<sub>rec</sub>–based registration algorithm is compared with DLSC-based registration method and NMI-based registration algorithm. RMSE of the correctly matched point pairs is employed to represent the registration accuracy. Two types of multimodal medical image pairs (synthetic brain data with complex intensity distortions and clinical multimodal brain images) are applied to validate the proposed registration method. The descriptions of data sets used in our experiments are given in Section IV.A. The five synthetic images individually moved a pixel to the right are used to the moving images and the original image is used as the fixed image in the registration process. In the LPPCO<sub>rec</sub>–based and DLSC-based registration algorithm, the block-based detector is set to extract 64 uniformly distributed interest points. In the NMI-based registration, a histogram with 32 bins is used to estimate NMI, because it obtains the optimal matching performance for the data sets used. For each image pair to be registered, 25 evenly distributed check points are manually labelled to determine the CM using a projective transformation model.
resolution. Our method first generates the LPPCO descriptor on the basis of local phase mean and phase congruency values of different orientations. To enhance the contour or edge features, a more sensitivity of phase deviation function is proposed for the calculation of phase congruency. Then, the NCC of LPPCO descriptors is used to define a new similarity measure (named LPPCO$_{ncc}$) that indicates that shape and structural similarity between images. Finally, a fast template matching strategy based on LPPCO$_{ncc}$ is developed to find CPs between images. LPPCO$_{ncc}$ has been evaluated with various types of multimodal medical image pairs (T2/PD, T2/T1, and T2/CT) and compared with the two popular similarity metrics including NMI and DLSC. Experimental results show that LPPCO$_{ncc}$ is robust to complex intensity distortions between multimodal medical images and outperforms the two popular similarity metrics. Furthermore, a registration algorithm based on LPPCO$_{ncc}$ for multimodal medical images is introduced. The registration results using various types of multimodal medical image pairs demonstrate that the proposed registration algorithm outperforms the NMI-based and DLSC-based registration methods in both the number of matched CPs and the registration accuracy.

REFERENCES


LI ZHANG received the B.S. degree in applied mathematics from Henan Normal University (HNU), Xinxiang, China, in 2012. She is currently working toward the Ph.D. degree at the School of Automation Science and Engineering, South China University of Technology (SCUT), Guangzhou, China. Her current research interests include multimodal medical image registration, medical image processing, and pattern recognition.

BIN LI received the B.S. and Ph.D. degrees from the School of Automation Science and Engineering, South China University of Technology (SCUT), Guangzhou, China, in 2002 and 2007, respectively. He is currently an Associate Professor at the School of Automation Science and Engineering, SCUT. His current research interests include information visualization, medical image processing and pattern recognition.

LIANFANG TIAN received the B.S. and M.S. degrees in mechanical engineering from Shandong University of Technology, Jinan, China, and the Ph.D. degree in mechanical engineering and electrical engineering from Harbin Institute of Technology (HIT), Harbin, China, in 1991, 1994, and 1997, respectively. He is currently a Professor at the School of Automation Science and Engineering, South China University of Technology (SCUT). His current research areas include biomedical image processing, medical image processing and pattern recognition.

WENBO ZHU received the B.S. and M.S. degree in pattern recognition and intelligent system from Harbin Engineering University, and the Ph.D. degree in pattern recognition and intelligent system from South China University of Technology (SCUT), China. He is currently a lecturer at the School of Automation School, FOSHAN University, China. His current research interests include medical image processing, pattern recognition and artificial intelligence.