Article:
Song, S, Du, C, Liu, X et al. (7 more authors) (2019) Deep motion tracking from multiview angiographic image sequences for synchronization of cardiac phases. Physics in Medicine and Biology, 64 (2). ARTN 025018. ISSN 1361-6560

https://doi.org/10.1088/1361-6560/aafa06

Physics in Medicine and Biology

Deep motion tracking from multiview angiographic image sequences for synchronization of cardiac phases

To cite this article before publication: Shuang Song *et al* 2018 *Phys. Med. Biol.* in press https://doi.org/10.1088/1361-6560/aafa06

Deep Motion Tracking from Multiview Angiographic Image Sequences for Synchronization of Cardiac Phases

Shuang Song[1], Chenbing Du[1], Xinxin Liu[1], Yong Huang[1], Hong Song[2], Yurong Jiang[1], Danni Ai[1,*], Alejandro F. Frangi[3], IEEE Fellow, Yongtian Wang[1], Jian Yang[1,*]

[1]Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

[2]School of Computer Science & Technology, Beijing Institute of Technology, Beijing 100081, China

[3]CISTIB Center for Computational Imaging & Simulation Technologies in Biomedicine, School of Computing and School of Medicine, University of Leeds, Leeds LS2 9JT, UK.

*Corresponding author: Jian Yang, Email: jyang@bit.edu.cn; Danni Ai, Email: danni@bit.edu.cn

Abstract: In the diagnosis and interventional treatment of coronary artery disease, the 3D+time reconstruction of the coronary artery on the basis of X-ray angiographic image sequences can provide dynamic structural information. The synchronization of cardiac phases in the sequences is essential for minimizing the influence of cardiorespiratory motion and realizing precise 3D+time reconstruction. Key points are initially extracted from the first image of a sequence. Matching grid points between consecutive images in the sequence are extracted by a multi-layer matching strategy. Then deep motion tracking of key points is achieved by local deformation based on the neighboring grid points of key points. The local deformation is optimized by the Random Sample Consensus algorithm. Then, a simple harmonic motion model is utilized to distinguish cardiac motion from other motion sources (e.g. respiratory, patient movement, etc.). Next, the signal which is composed of cardiac motions is filtered by a band-pass filter to reconstruct the cardiac phases. Finally, the synchronization of cardiac phases from different imaging angles is realized by piece-wise linear transformation. The proposed method was evaluated using clinical X-ray angiographic image sequences from 13 patients. 85% matching points can be accurately computed by the deep motion tracking method. The mean peak temporal distance between the reconstructed cardiac phases and the electrocardiograph signal is 0.027s. The correlation between the cardiac phases of the same patient is over 89%. Compared with three other state-of-the-art methods, the proposed method accurately reconstructs and synchronizes the cardiac phases from different sequences of the same patient. The proposed deep motion tracking method is robust and highly effective in synchronizing cardiac phases of angiographic image sequences captured from different imaging angles.

Key words: coronary arteries, X-ray angiographic image sequence, cardiac phase, synchronization, deep motion tracking

1. Introduction

With its fast imaging speed and high-resolution capability, X-ray angiography has been regarded as the gold standard for diagnosis and interventional treatment of coronary artery disease in clinical practice (Kurra et al., 2010; Chen et al., 2014). However, owing to the perspective projection principle, 2D X-ray angiographic images lose 3D information of the coronary artery. 3D+time reconstruction of coronary artery in sequences can provides a dynamic 3D structure for the clinicians to realize the preoperative surgical planning. In 3D reconstruction of the vasculature in 3D space,

the angiographic images should be captured at the same time and at different imaging angles (Cong et al., 2016; Yang et al., 2009; Yang et al., 2014). However, in clinical practice, the widely utilized mono-plane imaging device can only obtain a single sequence from a specific imaging angle in a single acquisition procedure. To realize the 3D+time reconstruction of vasculature in sequences, cardiac phases of images in different sequences should be synchronized to minimize the influence of cardiorespiratory motion to reconstruction (Cimen et al., 2016). In clinical practice, the synchronization can be handled by cardiac electrocardiogram (ECG) gating (Lauritsch et al., 2006). But the ECG-gating device is usually an optional extra for an X-ray imaging system and cannot normally be obtained in a regular hospital. Meanwhile, the use of an ECG-gating device in an operation may raise concerns because of its complexity and cost. Hence, image-based cardiac phase reconstruction and synchronization are highly necessary.

In the past two decades, to realize the cardiac phase synchronization, numerous methods have been proposed to measure cardiac motion. Lehmann et al. (Lehmann et al., 2006) applied histogram equalization to enhance vessels in an angiographic sequence and measured the superior–inferior component of a weighed centroid to track cardiac motion. Sundar et al. (Sundar et al., 2009) estimated cardiac and respiratory motion between successive images by a phase correlation-based method. By assuming that the motion of the structures only exhibits translation, cardiac phases can be computed with the sum of the cross-power spectrum of successive images. However, this assumption is not complex enough for the coronary artery. Considering the motions of both the coronary sinus catheter and coronary artery in angiographic image sequences are all highly related to cardiac motion, Toth et al. (Toth et al., 2017) reconstructed cardiac phases by a mask-PCA method (Panayiotou et al., 2014) by setting the threshold segmented and dilated coronary artery region as the mask. In addition, the mask-PCA method is proposed by Panayiotou et al. (Panayiotou et al., 2014) in the estimation of cardiac motion from angiographic image sequences that only contain a coronary sinus (CS) catheter. In the cardiac motion estimation of CS catheter sequences, hierarchical manifold learning (Panayiotou et al., 2013) is also proposed which is similar to the mask-PCA method. In the two methods, the catheter is initially enhanced by vesselness filter (Frangi et al., 1998) and then dilated. The intensity of the dilated regions is employed to form a matrix. After reducing the dimensionality of the matrix by hierarchical manifold learning or PCA method, the first or second principal components are used for describing cardiac and respiratory motion, respectively. Panayiotou et al. (Panayiotou et al., 2013) also proposed the Track-PCA to track the positions of CS catheter in the angiographic sequence and then exploited PCA to acquire cardiac and respiratory motion. Due to the possible absence of the catheter, Panayiotou et al. proved that Mask-PCA is more accurate than Track-PCA. However, when the influence of contrast agent washing in and out within the coronary arteries was considered, the three methods proposed by Panayiotou et al. (Panayiotou et al., 2014; Panayiotou et al., 2013; Panayiotou et al., 2013) were challenged by image intensity fluctuations and introduce errors to the cardiac phases reconstructed from the angiographic image sequences. Brost et al. (Brost et al., 2011) utilized a boosted classifier to segment the catheter and track the catheter by rigid registration in successive images. The catheter trajectory can compensate cardiac and respiratory motions. However, the extraction of coronary artery from angiographic images remains extremely challenging due to the coexistence of multi-organs (Chen et al., 2016). In the angiographic image sequences of coronary artery, multi-organ interference, non-uniform contrast agent infusion and complex motion of coronary artery all exist in the images. The above mentioned methods cannot accurately estimate the cardiac motion in the sequences.

In this paper, we propose a novel deep motion matching (DMT) method for synchronization of the cardiac phase by estimating the cardiac motion. Initially, in a sequence, the image with whole coronary artery is regarded as the first image and key points are extracted from the image. Meanwhile, matching grid points between consecutive images are computed by a multi-layer matching strategy. Next, an octree model is utilized to search for the neighboring grid points around each key point. Then, for the neighboring grid points of each key point, a local projective transformation is computed by the neighboring grid points. In computing the transformation, Random Sample Consensus (RANSAC) algorithm is utilized to discard the matching gird points that cause errors to the transformation. The local transformation is applied to the corresponding key point and obtain the matching key point in the next image. By repeating computing the transformation of different key points throughout the sequence, all the key points can be tracked in the sequence. After this, a simple harmonic motion (SHM) model is utilized to estimate the cardiac motions which constitute the final cardiac phases. Finally, for the cardiac phases reconstructed from different sequences, a piece-wise linear transformation is computed to synchronize the cardiac phases.

The proposed algorithm has three main contributions. First, cardiac phase synchronization is achieved by point motion tracking. The tracking is very effective to alleviate the influence of non-uniform contrast agent infusion. Second, the motion tracking is realized through the projective deformation of the dense correspondences of local region. It is highly robust to solve the challenging non-rigid motion within weak texture regions (without anatomical structures). Third, a simple harmonic motion model is utilized to compute the motion velocity between consecutive images in a sequence. It can effectively distinguish the cardiac motion from other motion sources (e.g. respiratory, patient movement, etc.).

2. Methods

In this study, sequences are comprised of images with full coronary artery. Let $\{I_e|e = 1, \cdots, E\}$ denote the coronary artery images, where I_e is the e^{th} image, and E is the total number of the images belong to a sequence. In a sequence, cardiac phases are reconstructed by tracking the motion of the key points that are extracted from I_1. In addition, cardiac phases reconstructed from different sequences are synchronized by matching the peaks and valleys of the cardiac phases.

2.1 Key Point Detection

The key points are extracted from I_1, and then the motion of the key points in the whole sequence are tracked. In this paper, we utilize a constraint to extract the key points that are mainly distributed on the coronary artery. The constraint can effectively identify the tubular structures in the images and distinguish the points belong to the coronary artery.

First, I_1 is enhanced by the vesselness based enhancement filter (Frangi et al., 1998), and the enhanced image is denoted as EI_1. The gradient matrix $G(x, y)$ of I_1 and Hessian matrix $H(x, y)$ of EI_1 can be computed as:

$$\begin{cases} G(x,y) = \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} \\ H(x,y) = \begin{pmatrix} EI_{xx} & EI_{xy} \\ EI_{yx} & EI_{yy} \end{pmatrix} \end{cases} \quad (1)$$

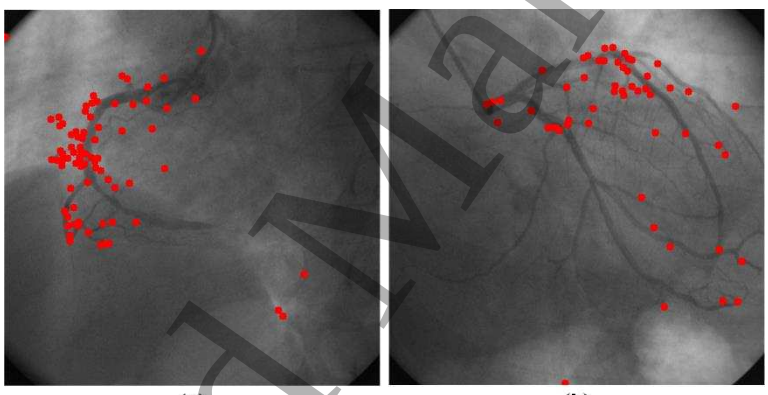
where I_x and I_y represent the first-order partial derivatives of the intensity in image I_1, and $EI_{xx}, EI_{xy}, EI_{yx}$ and EI_{yy} represent the second-order partial derivatives of the intensity in image E_k. The eigenvalues $\lambda_{1,2}^G$ and $\lambda_{1,2}^H$ of the gradient and Hessian matrices are computed:

$$\begin{cases} \lambda_{1,2}^G(x,y) = \left(I_x^2 + I_y^2 \pm \sqrt{\left(I_x^2 - I_y^2\right)^2 + 4\left(I_x I_y\right)^2}\right) \\ \lambda_{1,2}^H(x,y) = \left(EI_{xx} + EI_{yy} \pm \sqrt{\left(EI_{xx} - EI_{yy}\right)^2 + 4EI_{xy}^2}\right) \end{cases} \quad (2)$$

Then, we extract the key points $\{p_u^1 | u = 1,2,\cdots,U\}$ in image I_1, where p_u^1 refers to u^{th} the key point in I_1, U is the number of detected key points. In addition, the key points are computed according to the following constraints:

$$\begin{cases} R = \lambda_1^G \lambda_2^G - \alpha * (\lambda_1^G + \lambda_2^G)^2 \\ |\lambda_1^H| \approx 0 \\ |\lambda_1^H| \ll |\lambda_2^H| \end{cases} \quad (3)$$

where R is decided by the gray variation of local region (Harris et al., 1988) and is computed for every point in image I_1. α is a weighted value. When R is beyond a threshold value, the corresponding point is regarded as the initial key points. Then, the other two constraints are utilized to remove the points that are distributed in the tubular structures. The final remaining points are the extracted key points. $|\cdot|$ is the absolute value. Fig. 1 shows an example of the extracted key points in the angiograms.



(a) (b)

Fig. 1. Examples of the extracted key points in the angiograms when α=0.09, $R > 5$. Number of key points in (a): 54 and in (b): 77.

2.2 Deep Motion Tracking

Considering angiograms in the sequence are achieved by a perspective projection procedure, we compute the motion of each key point in $\{p_u^1 | u = 1,2,\cdots,U\}$ by the local projective transformation in two consecutive angiograms iteratively throughout the whole sequence. In addition, the local projective transformation can be computed by the matching point pairs of each key point in two consecutive angiograms.

The matching point pairs between two consecutive angiograms are computed by the multi-layer matching strategy (Revaud et al., 2016) and refer to the matching grid points. In the strategy, points per 4 pixels in the first image constitute the grid points. For images of size 512×512 pixels, the grid points are denoted as $\{2,6,10,\ldots,510\} \times \{2,6,10,\ldots,510\}$. Matching is realized by the bottom-up correlation pyramid computation and top-down correspondence extraction. In bottom-up procedure, a series of grid-point-centered non-overlapped patches ($4 * 4$ pixels) are extracted from the first image and convolved with the patches with the same size at all points in the second image

to generate the bottom correlation maps, respectively. $3 * 3$ max-pooling and averaging are applied to the 4 neighboring patches to generate the correlation maps in a higher layer. In the bottom-up correlation pyramid computation, 6 repetitions are used in the convolution and max-pooling procedures when the image size is of $512 * 512$ pixels. In top-down procedure, by extracting the correspondences with largest correlation values in the top layer of the pyramid, the correspondences at the lower layer are computed by searching for the maximum correlation values in the local region of the corresponding 4 sub-patches. By propagating searching procedure for the maximum correlation values from top layer to the bottom layer, the matching grid points between two consecutive images can be obtained. Meanwhile, in the searching procedure, incorrect correspondences are successively discarded.

For a specific sequence, grid points are extracted in images $\{I_e|e = 1, \cdots, E-1\}$, respectively. The matching grid points are then computed by the multi-layer strategy, and are denoted as $\left\{\{(q_k^e, q_k^{e+1})|k = 1, \cdots, K^e\}|e = 1, \cdots, E-1\right\}$, where (q_k^e, q_k^{e+1}) is k^{th} matching grid points in I_e and I_{e+1}, respectively, K^e is the number of matching grid points in I_e. By assuming that the key points and grid points are distributed in a plane that is perpendicular to the z-axis in 3D space, the octree-based algorithm (Meagher et al., 1982) is utilized to search for the neighboring grid points of each key point. For a key point p_u^1, $\{(q_{a,u}^1, q_{a,u}^2)|a = 1,2, \cdots, A\}$ is a subset of $\{(q_k^1, q_k^2)|k = 1, \cdots, K^1\}$ and denotes the neighboring grid points of p_u^1, where $q_{a,u}^1$ is a^{th} neighboring grid points of u^{th} key point in image I_1, $q_{a,u}^2$ is the matching grid point of $q_{a,u}^1$ in I_2, A is the number of the neighboring grid points. For each key point, we compute the same number of neighboring grid points. Fig. 2 shows an example of key point tracking based on the neighboring grid points when $A = 8$.

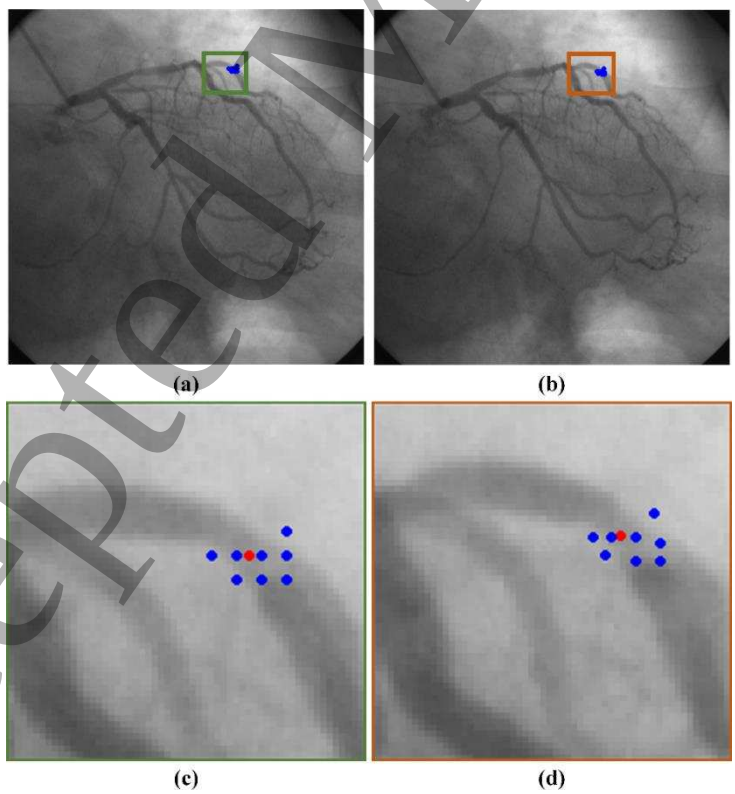


Fig. 2. An example of the key point tracking based on 8 neighboring grid points in the consecutive images. (a and b) first image and second image; (c) enlarged view in the green rectangular box from Fig. (a); (d) enlarged view in the

orange rectangular box from Fig. (b); (c and d) points in red color: matching key points; points in blue color: 8 neighboring grid points of the matching key points, respectively.

For the matching key points p_u^1, $B_u^{1,2}$ denotes the projective transformation. Hence, the matching key point p_u^2 in I_2 can be calculated as follows:

$$\begin{pmatrix} x_u^2 \\ y_u^2 \\ 1 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} * \begin{pmatrix} x_u^1 \\ y_u^1 \\ 1 \end{pmatrix} \quad (4)$$

where $\{b_{z_1 z_2,} | z_1, z_2 = 1,2,3\}$ are the free variables of matrix $B_u^{1,2}$, while (x_u^1, y_u^1) and (x_u^2, y_u^2) are the coordinates of key points p_u^1 and p_u^2, respectively. $B_u^{1,2}$ is computed by the neighboring grid points. However, as the A neighboring grid points may not correspond to the same projective transformation, the RANSAC optimization (Fischler et al., 1981) method is then used to search for the optimal projective transformation.

To find the optimal projective transformation, we randomly select four points from the A grid points to compute $B_u^{1,2}$. Then, all the A grid points in I_1 are transformed by current $B_u^{1,2}$. If the difference values between the computed point and the matching grid point in I_2 are less than the pre-defined threshold, the point is then considered as an inlier point, otherwise, it is an outlier point. The threshold function can be defined:

$$la(q_{a,u}^1) = \begin{cases} 1 & \left\| B_u^{1,2} * q_{a,u}^1 - q_{a,u}^2 \right\| \le \varepsilon \\ 0 & else \end{cases}, a = 1, \cdots, A \quad (5)$$

where $\|\cdot\|$ is the Euclidean distance between points $B_u^{1,2} * q_{a,u}^1$ and $q_{a,u}^2$. $la(q_{a,u}^1)$ is the label to define whether $q_{a,u}^1$ is an inlier or outlier, ε is the threshold, a is the order number of the neighboring grid points. After iteratively computing $B_u^{1,2}$ by randomly selecting four points from the A grid points, the matrix $B_u^{1,2}$ that has the largest number of inliers is the optimal:

$$\underset{B_u^{1,2}}{\operatorname{argmin}} \sum_u la(q_{a,u}^1) \quad (6)$$

After computing the optimal $B_u^{1,2}$, p_u^2 can be obtained by equation (4). All the computed matching key points in I_2 are regarded as the key points when computing the matching key points in I_3. By repeating computing the matching key points in the consecutive images in the sequence, the motion of the key points extracted in I_1 can be tracked in the whole sequence.

2.3 Cardiac Phase Reconstruction

Integrated cardiac motion comprises four major parts, namely, expansion, contraction, rotation, and twisting. Potel et al. (Potel et al., 1984) found that the expansion and contraction of the ventricle wall are far more significant than rotation or twisting. He concluded that expansion and contraction accounts for 90% of cardiac motion and assumed that the heart has a spherical shape (Potel et al., 1984; Chen et al., 1994). Given that the coronary artery distributes on the cardiac surface, the points on the coronary artery move towards and away from the center along with the cardiac systolic and diastolic movement. Hence, by regarding the point on the static cardiac surface as the motion center, the motion of points on the coronary artery can be modeled as simple harmonic motion (SHM) (Marion et al., 2013) that moves towards and away from the cardiac center.

In the SHM model, the intersections of extension lines or the reverse extension lines of the velocities define the cardiac center O_{e+1} in I_{e+1}. As shown in Fig. 3(a) and (b), the velocity $\overrightarrow{v_u^{e+1}}$ of p_u^{e+1} forms an angle θ_u^{e+1} with the vector $\overrightarrow{p_u^{e+1} O_{e+1}}$. Hence, the center O_{e+1} of cardiac motion can be obtained by minimizing the sum of θ_u^{e+1} of all key points:

$$\underset{O_{e+1}}{\operatorname{argmin}} \sum_{u=1}^{U} \theta_u^{e+1} \tag{7}$$

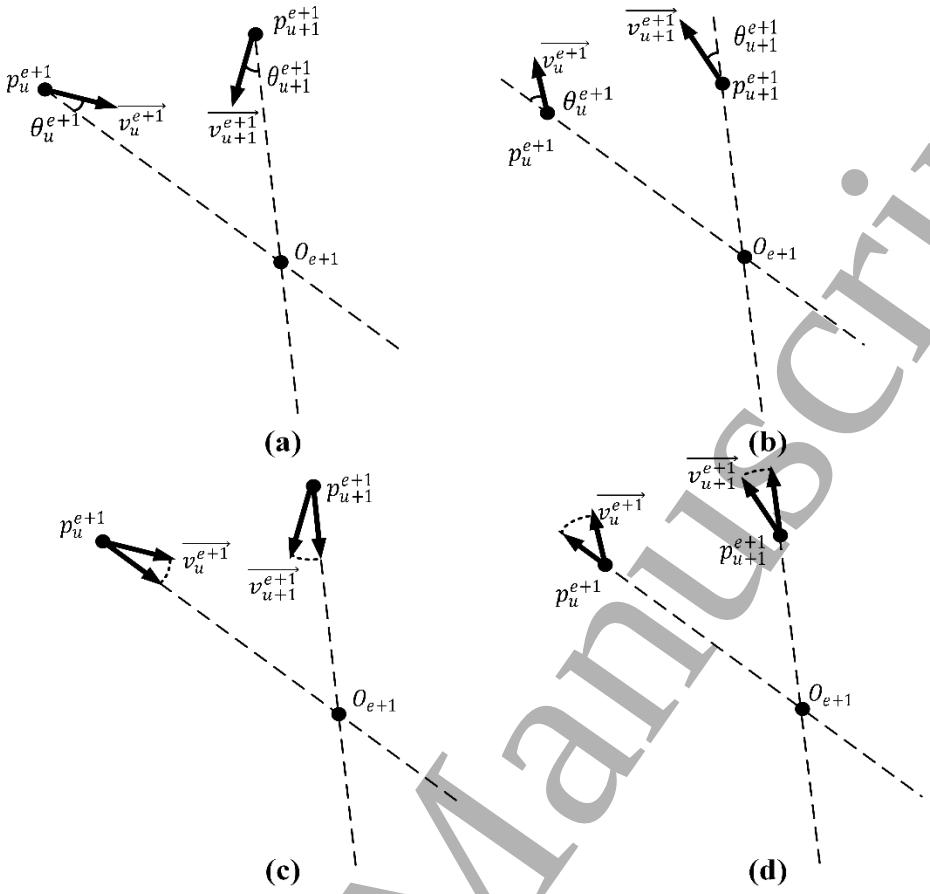


Fig. 3. Estimation of cardiac motion center and cardiac velocity: (a) and (b) cardiac motion center in systolic and diastole stage, respectively; (c) and (d) cardiac velocity in systolic and diastole stage, respectively.

As shown in Figs. 3(c) and (d), the velocity for each the key point can be projected towards the center O_{e+1}, which can be defined as:

$$v_u^{e+1} = \frac{\overrightarrow{v_u^{e+1}} * \overrightarrow{p_u^{e+1} O_{e+1}}}{\left|\overrightarrow{p_u^{e+1} O_{e+1}}\right|} \tag{8}$$

where $\overrightarrow{p_u^{e+1} O_{e+1}}$ is the vector from O_{e+1} to p_u^{e+1}, and v_u^{e+1} is the scalar quantity. If $v_u^{e+1} > 0$, the heart is in the systolic stage; otherwise, the heart is in the diastolic stage. The sum of all the v_u^{e+1} refers to the overall motion tendency. Thus, the cardiac motion velocity V_{e+1} from image I_e to I_{e+1} can be calculated as:

$$V_{e+1} = \frac{\sum_{u=1}^{U} v_u^{e+1}}{U} \tag{9}$$

The cardiac phases can be reconstructed by connecting the cardiac motion velocity values. To further completely remove the motion (respiratory, patient movement) from the cardiac phases, a 2^{nd} order Butterworth band-pass filter (Hernandez-Sabate et al., 2011) is utilized. After normalizing the filtered cardiac phases to $[-1,1]$, the final cardiac phases can be obtained. The cardiac phases from m^{th} sequence is denoted as C_{m}. The peaks $\left\{t_{j,pks}^{C_{\mathrm{m}}} | j = 1,2,\cdots,J_m\right\}$ of C_{m} represents the end-systole (Sundar et al., 2009), where J_m is the number of peaks in m^{th} sequence, pks

refers to the peaks.

2.4 Synchronization

The synchronization includes two parts: one is the synchronization between the reconstructed cardiac phases and the ECG signal. The corresponding purpose is to validate the accuracy of cardiac phase reconstruction. The other is the synchronization between different reconstructed cardiac phases from the sequences of the same patient. The purpose is to synchronize the images from different imaging angles.

Given that a time delay may exist between the ECG signal and coronary artery sequence, a time offset should be added to the ECG signal. Hence, the synchronization between the ECG signal and reconstructed cardiac phases can be realized as:

$$\underset{dt}{\operatorname{argmin}} \sum_{j=1}^{J} \left| t_j^{ECG_{\mathrm{m}}} + dt - t_{j,pks}^{C_{\mathrm{m}}} \right| \tag{10}$$

where dt is the time offset between the ECG signal and the reconstructed cardiac phases. $t_j^{ECG_{\mathrm{m}}}$ refers to the closest peak in the ECG signal to the j^{th} peak in the reconstructed cardiac phases.

In the second kind of synchronization, after the cardiac phase of all the sequences from a same patient are extracted, the peaks and valleys of the cardiac phases are all extracted. The sequence of the cardiac phases that have the largest number of peaks is regarded as the reference, and we assume the reference sequence is m^{th} one. For the reference sequence, the time of peaks in the cardiac phases are denoted as $\left\{ t_{j,pks}^{C_{\mathrm{m}}} | j = 1,2,\cdots,J \right\}$, and the time of valleys in the cardiac phases are denoted as $\left\{ t_{o,vys}^{C_{\mathrm{m}}} | o = 1,2,\cdots,O \right\}$, where vys refers to the valleys. For another m'^{th} sequence from the same patient, the peaks are denoted as $\left\{ t_{j',pks}^{C_{\mathrm{m}'}} | j' = 1,2,\cdots,J' \right\}$, and the valleys are denoted as $\left\{ t_{o',vys}^{C_{\mathrm{m}'}} | o' = 1,2,\cdots,O' \right\}$. The synchronization between the cardiac phases are realized by a piece-wise linear transformation. Considering that the first local extreme value in the cardiac phases may be a peak or a valley, the four endpoints for computing first linear transformation is decided as follows:

$$\begin{cases} \left(\left[t_{1,pks}^{C_{\mathrm{m}'}}, t_{1,vys}^{C_{\mathrm{m}'}} \right], \left[t_{1,pks}^{C_{\mathrm{m}}}, t_{1,vys}^{C_{\mathrm{m}}} \right] \right) & if\ t_{1,pks}^{C_{\mathrm{m}}} < t_{1,vys}^{C_{\mathrm{m}}}, t_{1,pks}^{C_{\mathrm{m}'}} < t_{1,vys}^{C_{\mathrm{m}'}} \\ \left(\left[t_{1,vys}^{C_{\mathrm{m}'}}, t_{1,pks}^{C_{\mathrm{m}'}} \right], \left[t_{1vys}^{C_{\mathrm{m}}}, t_{2,pks}^{C_{\mathrm{m}}} \right] \right) & if\ t_{1,pks}^{C_{\mathrm{m}}} < t_{1,vys}^{C_{\mathrm{m}}}, t_{1,pks}^{C_{\mathrm{m}'}} > t_{1,vys}^{C_{\mathrm{m}'}} \\ \left(\left[t_{1,pks}^{C_{\mathrm{m}'}}, t_{1,vys}^{C_{\mathrm{m}'}} \right], \left[t_{1,pks}^{C_{\mathrm{m}}}, t_{2,vys}^{C_{\mathrm{m}}} \right] \right) & if\ t_{1,pks}^{C_{\mathrm{m}}} > t_{1,vys}^{C_{\mathrm{m}}}, t_{1,pks}^{C_{\mathrm{m}'}} < t_{1,vys}^{C_{\mathrm{m}'}} \\ \left(\left[t_{1,vys}^{C_{\mathrm{m}'}}, t_{1,pks}^{C_{\mathrm{m}'}} \right], \left[t_{1,vys}^{C_{\mathrm{m}}}, t_{1,pks}^{C_{\mathrm{m}}} \right] \right) & if\ t_{1,pks}^{C_{\mathrm{m}}} > t_{1,vys}^{C_{\mathrm{m}}}, t_{1,pks}^{C_{\mathrm{m}'}} > t_{1,vys}^{C_{\mathrm{m}'}} \end{cases} \tag{11}$$

where $([\cdot,\cdot],[\cdot,\cdot])$ refers to the two intervals to compute the first linear transformation, $[\cdot,\cdot]$ is an interval comprised by two endpoints. The transformation transforms the values in the first interval to the values in the second interval. After obtaining the endpoints of the first linear transformation, the endpoints of the remaining linear transformation can be obtained in chronological order. After the piece-wise linear transformation is computed, the cardiac phases of m'^{th} sequence can be synchronized to the cardiac phases of m^{th} sequence.

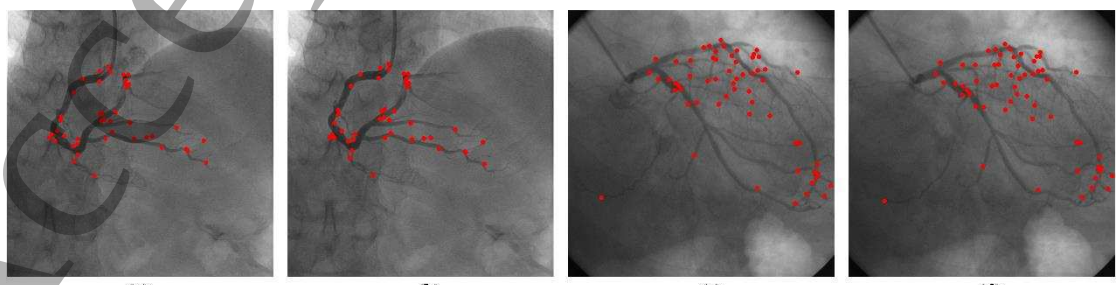
2.5 Dataset and Evaluation Criteria

A) Clinical Datasets

The clinical coronary artery angiographic image sequences from 13 patients were used in our experiments. All the sequences were captured by using a monoplane cardiac X-ray angiographic device (Philips Medical System, The Netherlands) in the Peking Union Hospital, Beijing, China. In the angiograms, multiple types of motion, including cardiac, respiratory and patient motion coexist. In addition, structures such as heart present large weak texture regions. Furthermore, multiple physical structures, including heart, bones, ribs, diaphragm and vessels coexist. All the sequences are initiated by injecting the contrast agent to the catheter. The frequency of image acquisition is 15fps. The size of images is $512 * 512$ pixels. The pixel sizes of the images are $0.3 * 0.3$ mm. The range of magnification is $[1.3, 1.4]$ when acquiring image sequences. The imaging angles of a sequence are described by two angles, which were left/right anterior oblique angles (LAO/RAO) and caudal/cranial angles (CAU/CRA), respectively. The sequences are divided into two datasets, viz. Dataset with ground truth (DWGT) and Dataset without ground truth (DWOGT). DWGT contains 7 sequences from the same patient who underwent ECG-gated examination. The embedded ECG signal is continuously recorded, and images in the corresponding sequence are captured at several specific time points of the ECG signal. The number of images in DWGT varies from 49 to 63. The imaging angles of sequences in DWGT vary from LAO22.2° to RAO41.4° and CAU30.1° to CRA42.2°. DWOGT contains 75 sequences from 12 another patients without ECG signal. For the 12 patients, the number of sequences varies between 12 and 2. The number of images in the sequences varies from 42 to 90. The imaging angles of sequences in DWOGT vary from LAO37.6° to RAO46.5° and CAU36.9° to CRA40.7°. The 13 patients are denoted as P1 to P13, respectively. For each patient, we denote the sequences as $\{Data1, \cdots, DataN\}$, and N is the number of the sequences belong to the same patient.

B) Evaluation Criteria

The proposed deep motion tracking (DMT) method will be compared with the techniques, viz. Multi-resolution image registration (MIR) (Nejati et al., 2014), DeepFlow (Weinzaepfel et al., 2013) and EpicFlow (Revaud et al., 2015), using a set of matching points. The points are manually labelled in two groups of consecutive images from left coronary artery (LCA) and right coronary artery (RCA) sequences, respectively. The labelled points are distributed on the vascular structures, including the intersection, branching, large scale and small scale segments. As for each kind of segment, the points are labelled randomly. We achieve DeepFlow (https://thoth.inrialpes.fr/src/deepflow/) and EpicFlow (https://thoth.inrialpes.fr/src/epicflow/) using available public implementations. We re-implemented MIR in strict accordance with the method in the original paper. The Euclidean distance between the computed point and the labelled point is utilized to evaluate the accuracy of motion tracking. Fig. 4 shows the manually labelled points in two groups of images.



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

Fig. 4. Manually labelled points in two groups of images. (a) and (b): two consecutive images from a RCA sequence. (c) and (d): two consecutive images from a LCA sequence.

The DMT-based cardiac phase reconstruction will be compared with the approaches, viz. Track-PCA (T-PCA) (Panayiotou et al., 2013), Mask-PCA (M-PCA) (Panayiotou et al., 2014) and Phase-Correlation (PCR) (Sundar et al., 2009), using the sequences from DWGT. Meanwhile, to evaluate the effectiveness of the DMT-based cardiac phase reconstruction on branching points (B-DMT), we also manually select the branching points from I_1 and compare the reconstructed cardiac phases with the proposed method. We re-implement the three approaches in strict accordance with the original paper and utilize the same band-filter to obtain the final cardiac phases. Before the evaluation, the ECG signal is synchronized to the reconstructed cardiac phases, as described in section 2.4. The mean peak temporal distance (MPTD) is utilized to evaluate the accuracy of cardiac phase reconstruction. The MPTD refers to the distance between the peak time of the reconstructed cardiac phases and of the R-waves of the ECG signal. MPTD is computed using the following equation:

$$\mathrm{MPTD}(m, ECG_m) = \frac{\sum_{e=1}^{E} \left| t_{e,pks}^{C_\mathrm{m}} - t_e^{ECG_\mathrm{m}} \right|}{E} \tag{12}$$

where E is the number of peaks in the reconstructed cardiac phases. $t_{e,pks}^{C_\mathrm{m}}$ is the time of m^{th} peak in the reconstructed cardiac phases; and $t_e^{ECG_\mathrm{m}}$ is the time of the m^{th} peak in the ECG signal. $|\cdot|$ is the absolute value. If the cardiac phases and ECG signal are accurately aligned, the MPTD equals to 0. However, for the reason that the ECG signal is continuously recorded while the images in the sequences are captured in specific time points, MPTD cannot strictly equal to 0. Hence, the smaller MPTD is, more accurately the cardiac phases is reconstructed. The standard deviation of the peak temporal distance (SDPTD) is computed to evaluate the variability for each sequence.

The DMT-based cardiac phase synchronization will also be compared with methods, viz. T-PCA, M-PCA and PCR, using the sequences both from DWGT and DWOGT. Before the comparison, we utilize the spline interpolation method (Smith et al., 2012) to interpolate the synchronized cardiac phases and generate a group of new cardiac phases with the same length. Pearson correlation coefficient (PCC) (Benesty et al., 2009) is then utilized to evaluate the correlation between two synchronized cardiac phases from the same patient. PCC is computed using the following equation:

$$\mathrm{PCC}(m, m') = \frac{cov(C_m, C_{m'})}{\sigma_{C_m} \sigma_{C_{m'}}} \tag{13}$$

where $cov(\cdot)$ is the covariance between interpolated cardiac phases C_m and $C_{m'}$; σ_{C_m} and $\sigma_{C_{m'}}$ are the standard deviation of C_m and $C_{m'}$, respectively. If cardiac phases are reconstructed accurately, the correlation is close to 100%. The mean and standard deviation of the PCC (MPCC$\pm$SDPCC) are computed to evaluate the correlations between the cardiac phases from the sequences of each patient.

3. Results

All the algorithms were implemented in C++ under the Ubuntu environment, and all the experiments were conducted on a relatively low-cost PC with 16 GB RAM and 3.2 GHz Intel CPU. For the proposed method, parameters $\alpha = 0.09$, $\epsilon = 3.0$. The parameters are the same for the sequences of all the patients.

A. Evaluation of Motion Tracking

Fig. 5 shows the Euclidean distances between the labelled points and computed points by methods, viz. MIR, DeepFlow, EpicFlow and DMT, respectively. In Fig. 5(a), the percentage of the point distances less than 3 pixels is 90.7%. In Fig. 5(b), the percentage of the point distances less than 3 pixels is 85%. The points computed by DMT are very close to the labelled points which indicates DMT can accurately obtain the matching points in two consecutive images. While most of the points computed by MIR, DeepFlow and EpicFlow are far away from the labelled points. Fig. 6 shows the points whose distances is beyond 3 pixels by DMT. Especially in the small scale vascular segments, when the number of the vascular segments is very large, or the vascular segments between images disappear, the matching points cannot accurately obtained by DMT.

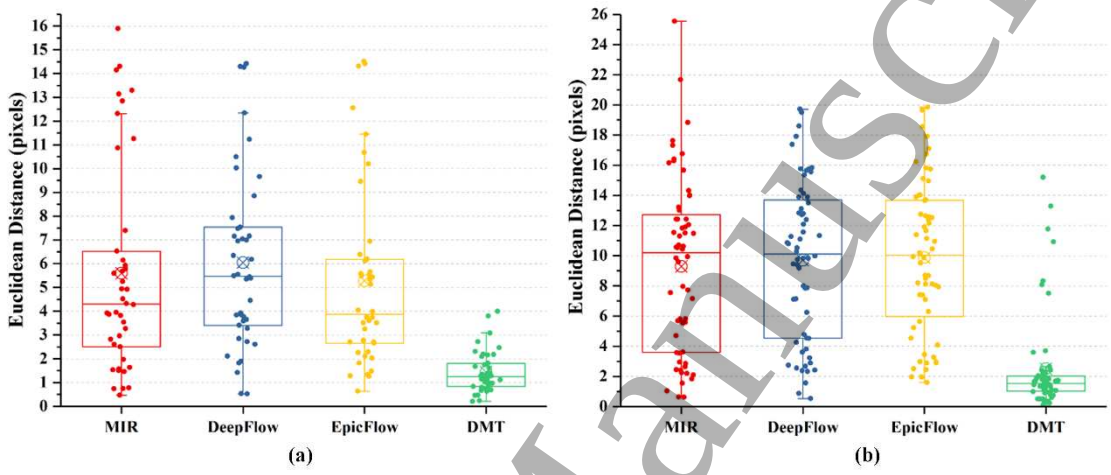


Fig. 5. Euclidean distances between the manually labelled points and the computed points by methods MIR, DeepFlow, EpicFlow and DMT. (a): Distances of points in Fig. 3(b); (b): Distances of points in Fig. 3(d).

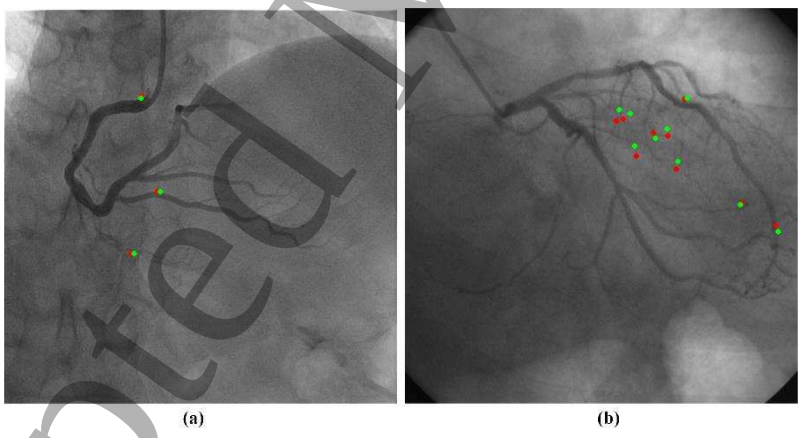


Fig. 6. Points whose distances with the labelled points are beyond 3 pixels in Fig. 3(b) and Fig. 3(d), respectively. Red color: labelled points; Green color: computed points by DMT.

Fig. 7 shows the Euclidean distances between the manually labelled points and the computed points by DMT method when ε varies from 0.0 to 8.0. As can be seen from the figure, when ε changes, the error of point matching does not change much. This indicates that the motion tracking is not sensitive to value ε.

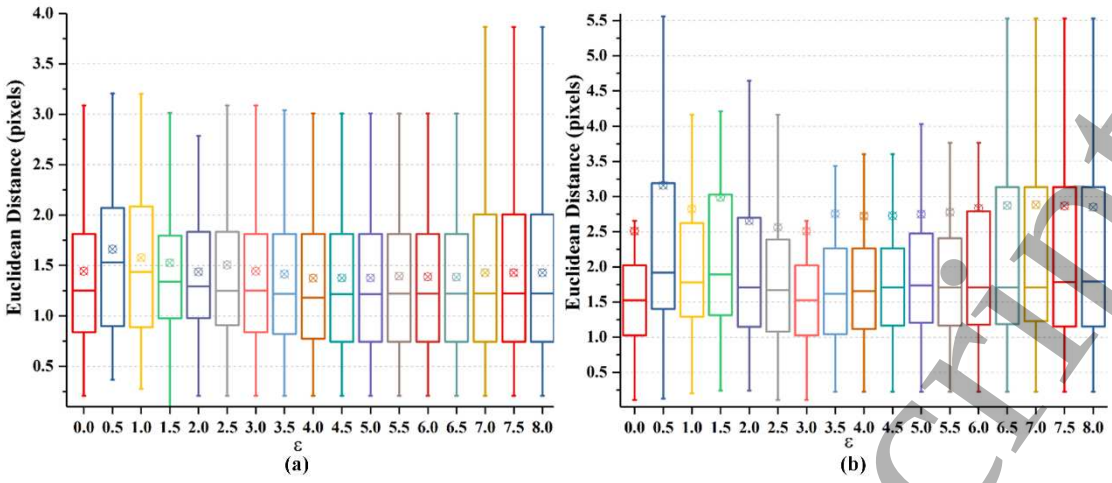


Fig. 7. Euclidean distances between the manually labelled points and the computed points by DMT methods when ε varies from 0.0 to 8.0. (a): Distances of points in Fig. 3(b); (b): Distances of points in Fig. 3(d).

B. Evaluation of Cardiac Phase Reconstruction

Fig. 8 shows the synchronization between the reconstructed cardiac phases and ECG signal for sequence Data2 of patient P1. The magnitudes of the reconstructed cardiac phases refer to the normalized cardiac motion velocity. In the figure, the cardiac phases and ECG signal are both linearly normalized to [0,1]. In computing cardiac velocity, the changing of patient motion in each frame results in the magnitude difference even in the same cardiac phase, as can be seen in Figs. 8 and 10. From full infusion of coronary artery to the dissipation beginning of contrast agent, the images cover 4 R-wave of the ECG signal. The time of the peaks in the ECG signal are 1.467s, 2.219s, 2.967s and 3.720s, respectively. The MPTD±SDPTD between the reconstructed cardiac phases and the ECG signal are 0.023 ± 0.021s. The four peaks of the reconstructed cardiac phases are effectively aligned with the peaks of the ECG signal, as shown in Fig. 8.

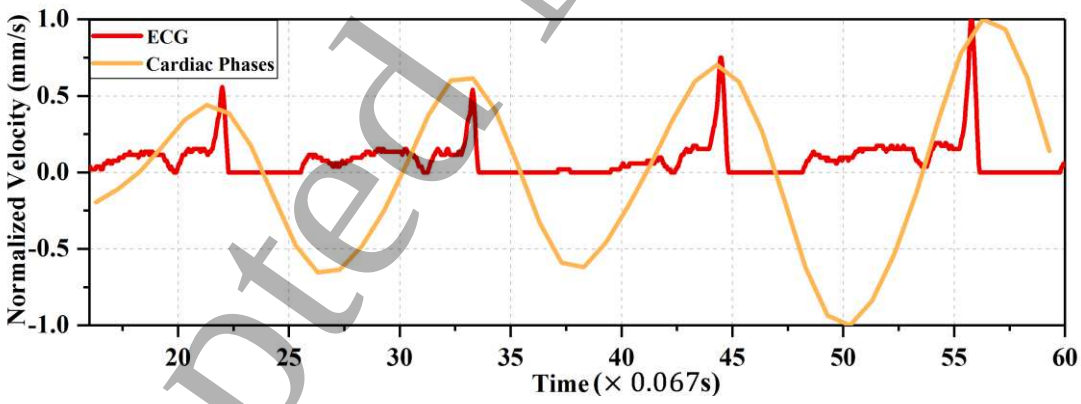


Fig. 8. An example of the reconstructed cardiac phases synchronized with the ECG signal for sequence Data2 of patient P1. Red curve: the ECG signal; Orange curve: reconstructed cardiac phases; Data2: LCA sequence.

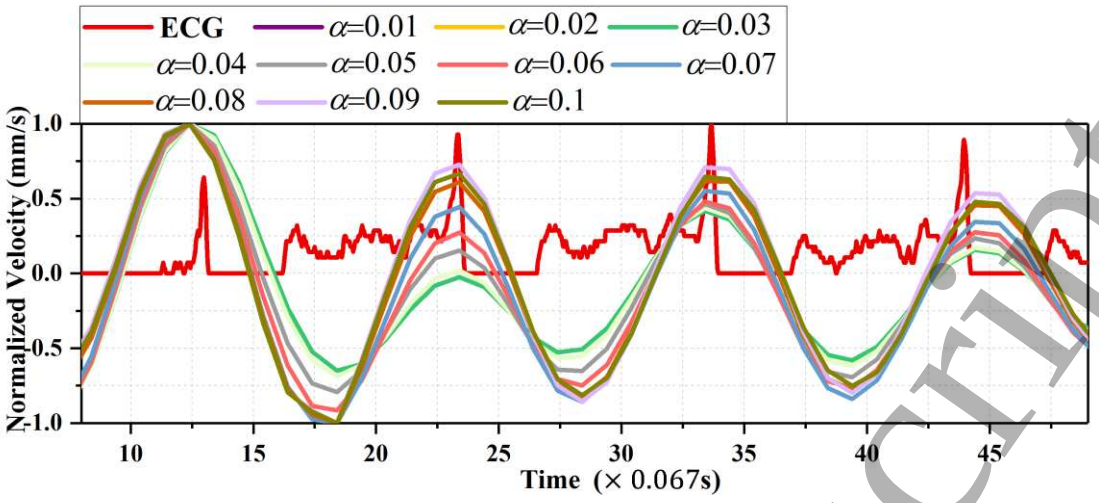


Fig. 9. Reconstructed cardiac phases along with the increase of α, synchronized with the ECG signal for sequence Data4 of patient P1. Data4: LCA sequence.

Fig. 9 shows the synchronization between the ECG signal and the reconstructed cardiac phases along with the increase of α from 0.01 to 0.1. The time of the peaks in the ECG signal are 0.870s, 1.564s, 2.256s and 2.944s, respectively. For each reconstructed cardiac phases at different α, the MPTD$\pm$SDPTD between the cardiac phases and the ECG signal are all 0.023 ± 0.015s. In addition, for the other sequences in DWGT, the reconstructed cardiac phases at different α also have the same accuracy which indicates that the accuracy of cardiac phase reconstruction is not influenced by α.

Fig. 10 shows the synchronization between the ECG signal and the cardiac phases by T-PCA, M-PCA, PCR, and DMT, respectively. Cardiac cycles in the ECG signal are computed by the time difference between the peaks in two successive R-waves. The mean and standard deviation of the cardiac cycles are 0.744 ± 0.094s, indicating that the cardiac cycles change much throughout the sequence, as shown in Fig. 10. According to the order of T-PCA, M-PCA, PCR, B-DMT and DMT, the MPTD$\pm$SDPTD between the cardiac phases and ECG signal are 0.076 ± 0.093s, 0.059 ± 0.083s, 0.037 ± 0.045s, 0.015 ± 0.013s and 0.015 ± 0.013s, respectively. The peaks of the cardiac phases achieved by DMT is the closest to the peaks of the R-waves in the ECG signal. In addition, the cardiac phases by B-DMT is coincident with the phases by DMT.

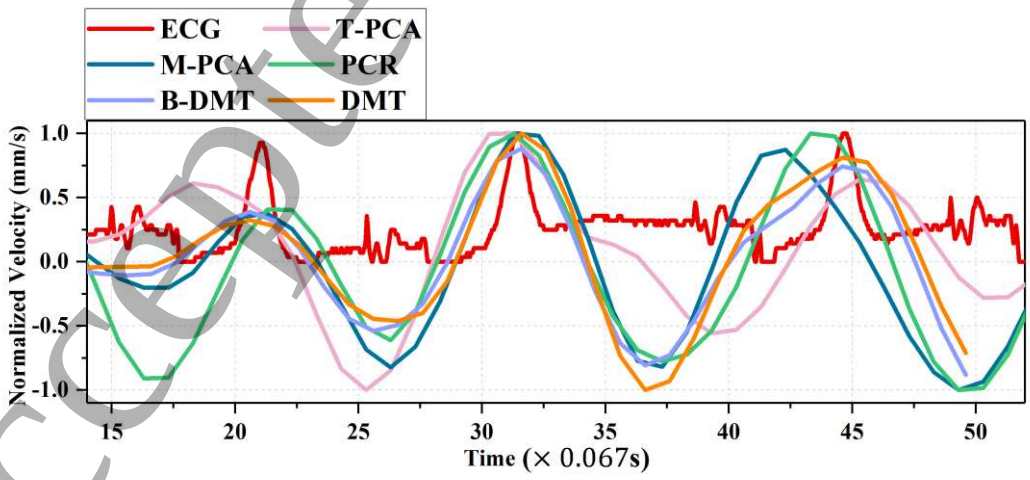


Fig. 10. An example of reconstructed cardiac phases synchronized with the ECG signal of sequence Data5 of patient P1 by four methods, T-PCA, M-PCA, PCR, B-DMT and DMT, respectively. Cardiac phases obtained by each method

are synchronized with ECG signal, respectively. Data5: RCA sequence.

Table 1. MPTD±SDPTD (s) between the cardiac phases and corresponding ECG signals from 7 sequences in DWGT by four methods, viz. T-PCA, M-PCA, PCR, and DMT, respectively. Num.: Number of covered R-waves. Data1-Data3: LCA sequences; Data4-Data7: RCA sequences. Avg.: Average values.

Data	ECG Signal		T-PCA	M-PCA	PCR	DMT
	Cycles (s)	Num.	MPTD±SDPTD (s)			
Data1	0.723 ± 0.004	3	0.098 ± 0.083	0.129 ± 0.185	0.187 ± 0.253	0.031 ± 0.045
Data2	0.746 ± 0.010	4	0.033 ± 0.024	0.167 ± 0.254	0.117 ± 0.101	0.023 ± 0.021
Data3	0.642 ± 0.005	2	0.018 ± 0.019	0.049 ± 0.067	0.382 ± 0.227	0.018 ± 0.019
Data4	0.686 ± 0.005	4	0.029 ± 0.023	0.312 ± 0.357	0.123 ± 0.171	0.023 ± 0.015
Data5	0.744 ± 0.094	3	0.076 ± 0.093	0.059 ± 0.083	0.037 ± 0.045	0.015 ± 0.014
Data6	0.695 ± 0.005	4	0.039 ± 0.045	0.167 ± 0.177	0.155 ± 0.175	0.034 ± 0.037
Data7	0.700 ± 0.012	4	0.055 ± 0.043	0.114 ± 0.125	0.039 ± 0.017	0.039 ± 0.017
Avg.	0.710 ± 0.045	-	0.053 ± 0.067	0.161 ± 0.204	0.141 ± 0.160	0.027 ± 0.024

For each sequence from dataset DWGT, Table 1 shows the MPTD±SDPTD between the reconstructed cardiac phases and the corresponding ECG signals by T-PCA, M-PCA, PCR, and DMT, respectively. The Cycles in the second column of Table 1 show the mean and standard deviation of cardiac cycles in each ECG signal, and Num. refers to the number of covered R-waves from the full contrast filling of coronary artery to the start of contrast agent washing out of the coronary artery. In Table 1, a large heart rate variation is observed when Data2, Data5, and Data7 are acquired. In particular, the heart rate varies much in Data5. Despite the variations in heart rate, the cardiac phases reconstructed by DMT can always effectively synchronized with the corresponding ECG signals. The cardiac phases reconstructed by T-PCA have a large deviation when the heart rate varies much. In M-PCA and PCR, cardiac phases have a larger deviation with the ECG signals in most sequences. In addition, MPTD±SDPTD by B-DMT are the same with DMT.

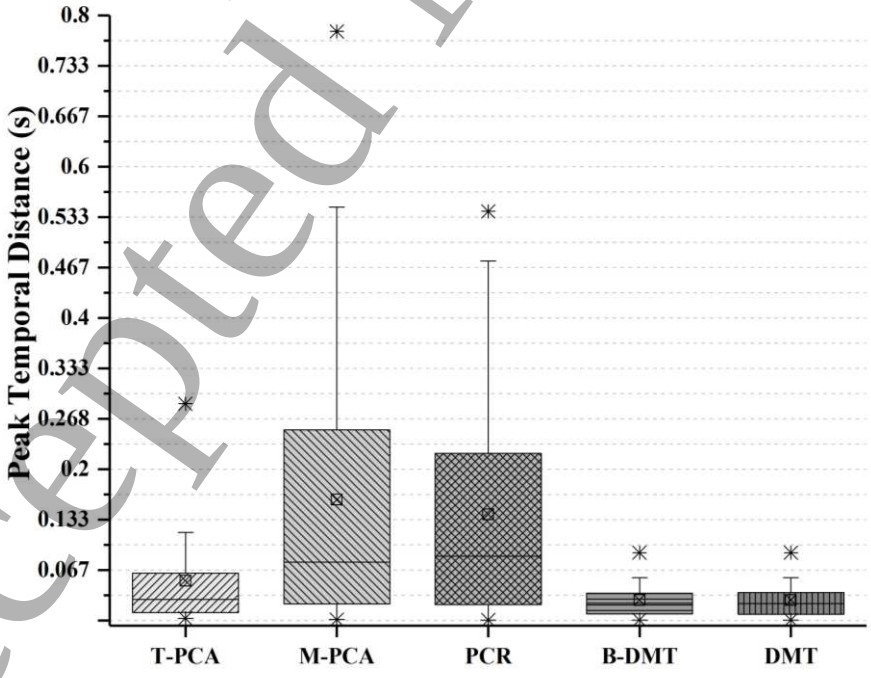


Fig. 11. Distribution of the peak temporal distances between the reconstructed cardiac phases and ECG signal by four methods, T-PCA, M-PCA, PCR, B-DMT and DMT, respectively.

Fig. 11 shows the performance of cardiac phase reconstruction by computing the peak temporal

distance between the cardiac phases and ECG signals. After synchronizing the peaks of all the data in DWGT with the corresponding peaks in the ECG signal, the peak temporal distances are displayed in Fig. 11. According to the order of T-PCA, M-PCA, PCR, B-DMT and DMT, the MPTD$\pm$SDPTD are 0.053 ± 0.066s, 0.160 ± 0.203s, 0.140 ± 0.159s, 0.027 ± 0.024s and 0.027 ± 0.024s, respectively. This indicates that the cardiac phases by DMT are in maximum agreement with the ECG signals. B-DMT can also achieve the same accuracy with DMT. Several peaks that are computed by T-PCA deviate with the ECG signal. As to the cardiac phases computed by M-PCA and PCR, large errors are caused due to the methods highly dependent on the intensity of images. We also performed paired t-test on the peak temporal distances to evaluate the performance of the cardiac phase reconstruction. The p-values between DMT and T-PCA, M-PCA, PCA are 0.026, 0.002 and 0.001, respectively. It can be seen that all the p-values are smaller than the commonly accepted significant value of 0.05, which demonstrates significant differences between DMT and other three methods.

C. Evaluation of Cardiac Phase Synchronization

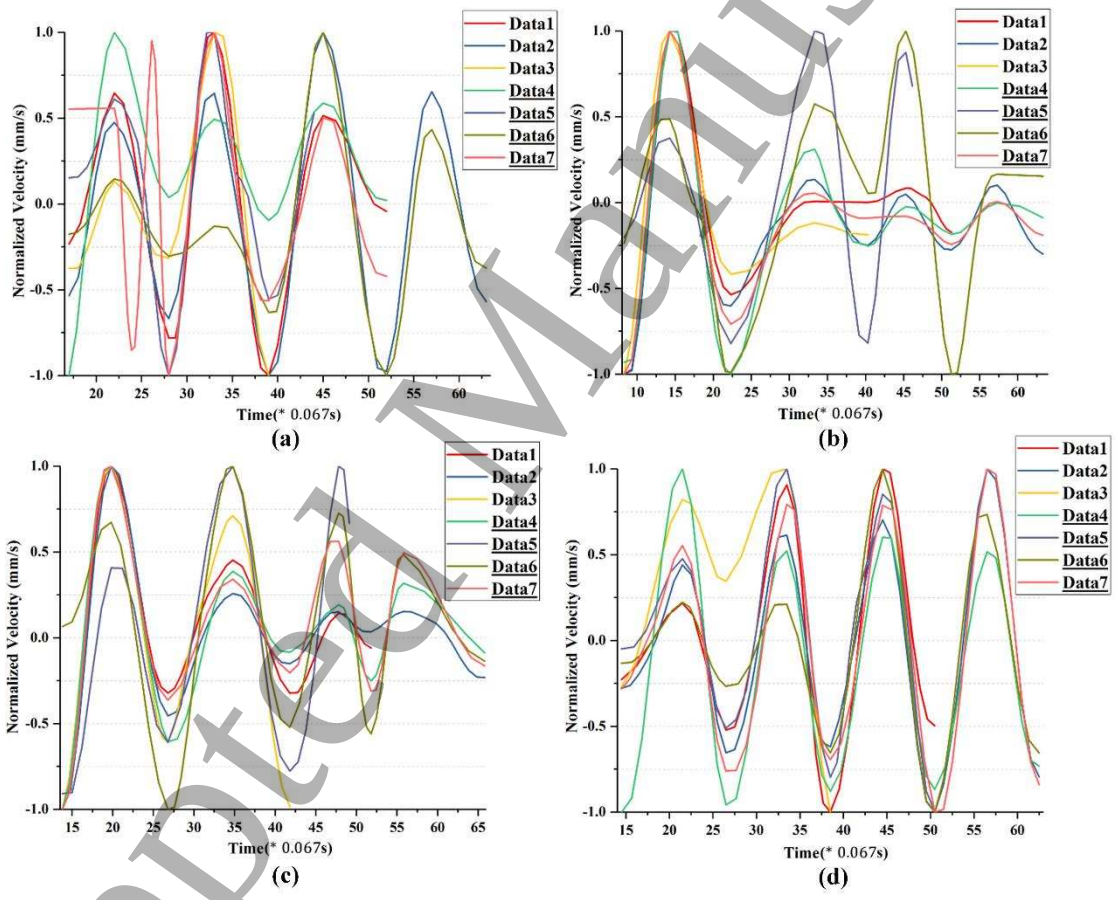


Fig. 12. Synchronization of the cardiac phases reconstructed from all the sequences in DWGT. Cardiac phases are reconstructed by four methods, (a) T-PCA, (b) M-PCA, (c) PCR, and (d) DMT, respectively. Data2: reference cardiac phases. Data1-Data3: LCA sequences. Data4-Data7: RCA sequences.

Fig. 12 shows the synchronization of the cardiac phases from the sequences in dataset DWGT by T-PCA, M-PCA, PCR, and DMT, respectively. In Figs. 12(a)-(d), cardiac phases extracted from Data2 are regarded as the reference, and the other cardiac phases are synchronized with it by the method in Section 2.4. As can be seen from Fig. 12(d), cardiac phases computed by DMT are highly

correlated with each other. In Fig. 12(a), the cardiac phases from Data7 differ much with other cardiac phases. In Fig. 12(b) and 12(c), cardiac phases from Data5 and Data6 deviate much with respect to cardiac phases of Data2. In two different sequences, if the cardiac phases are reconstructed accurately, the spline interpolation will not change the increasing or decreasing rule of the cardiac phases. Hence, the cardiac phases present high correlation after the synchronization, as can also be seen from Fig. 12(d).

According to the order of T-PCA, M-PCA, PCR and DMT, the MPCC±SDPCC between different cardiac phases are $76.42\% \pm 16.24\%$, $81.03\% \pm 22.66\%$, $81.33\% \pm 17.87\%$ and $89.45\% \pm 9.22\%$, respectively. As shown in Fig. 13, with either stable or varying heart rate, DMT could obtain an effective synchronization between different cardiac phases. The synchronization between RCA and LCA sequences is also very robust. Paired t-test is calculated on the correlations of DMT and other three methods T-PCA, M-PCA, and PCR. The p-values are all much smaller than 0.05, which demonstrates the significance of the differences in performance between DMT and other three methods.

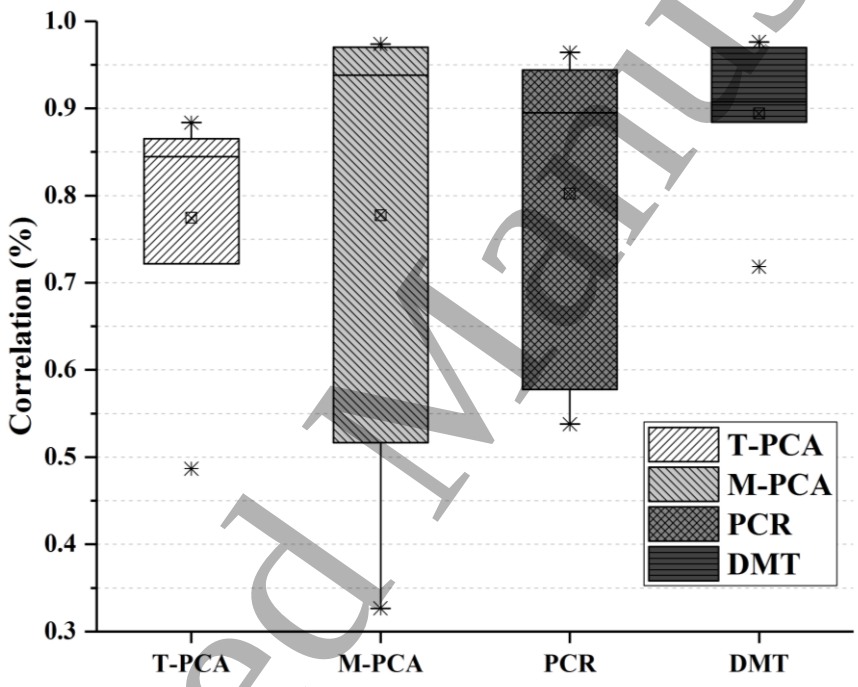


Fig. 13. PCC between the reconstructed cardiac phases from sequences in DWGT by methods, T-PCA, M-PCA, PCR and DMT, respectively.

To validate the wide applicability of DMT, Fig. 14 compares the accuracy of the cardiac phase synchronization from the sequences of patient P2 in DWOGT. Fig. 14(a) shows the cardiac phase synchronization by DMT, which indicates that cardiac phases from different sequences are highly correlated with each other. Fig. 14(b) shows the correlations by T-PCA, M-PCA, PCR, and DMT, respectively. The MPCC±SDPCC by the four methods are $65.33\% \pm 12.71\%$, $77.03\% \pm 7.75\%$, $79.32\% \pm 7.04\%$ and $90.31\% \pm 3.83\%$, respectively. Fig. 15 also compares the accuracy of cardiac phase synchronization from the sequences of patient P9 in DWOGT. The MPCC±SDPCC by the four methods are $67.54\% \pm 12.71\%$, $88.42\% \pm 7.03\%$, $72.08\% \pm 15.31\%$ and $98.17\% \pm 1.48\%$, respectively. Table 2 displays the correlations of cardiac phases from the sequences of the remaining 10 patients in DWOGT. DMT achieves the largest correlation for all the

sequences from the 10 patients. It is demonstrated that DMT could effectively synchronize the cardiac phases from different RCA and LCA sequences. Paired t-test is again applied on the PCC of all the sequences from DWOGT between DMT and T-PCA, M-PCA and PCR, respectively. The p-values are all much smaller than 0.05, which also demonstrates statistical significance of the differences between DMT and other three methods.

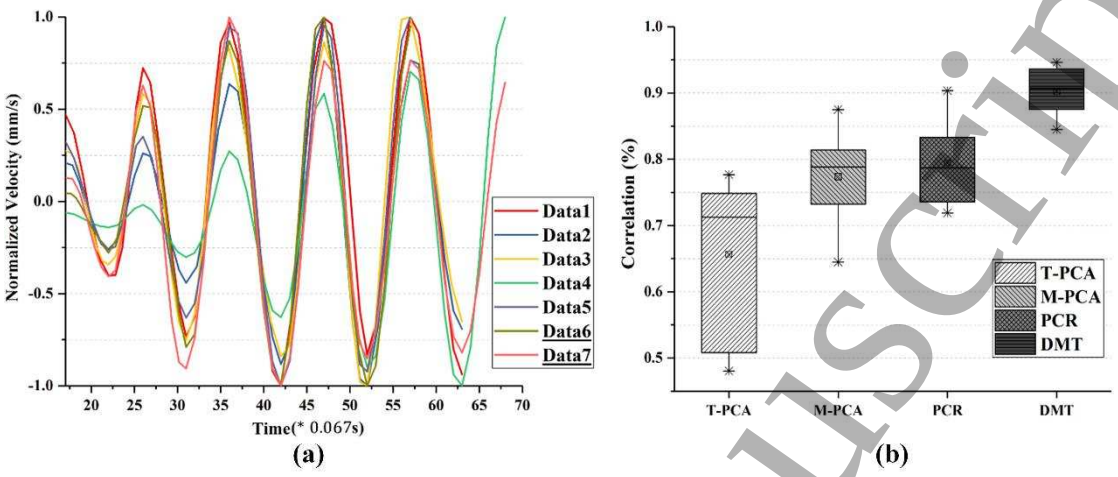


Fig. 14. Synchronization of the cardiac phases reconstructed from the sequences of patient P2 in DWOGT. (a) Synchronized cardiac phases. (b) Correlations of the cardiac phases by four methods, T-PCA, M-PCA, PCR, and DMT, respectively. Data1-Data5: LCA sequences; Data6-Data7: RCA sequences.

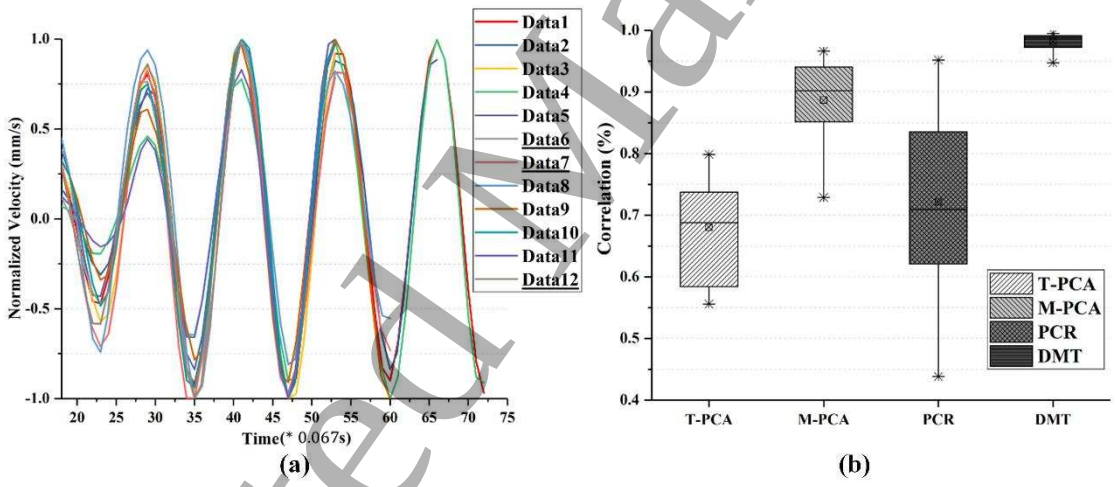


Fig. 15. Synchronization of the cardiac phases reconstructed from the sequences of patient P9 in DWOGT. (a) Synchronized cardiac phases. (b) Correlations of the cardiac phases by methods T-PCA, M-PCA, PCR, and DMT, respectively. Data1-Data5, Data8-Data11: LCA sequences; Data6-Data7, Data12: RCA sequences.

Table 2. MPCC±SDPCC (%) of synchronized cardiac phases from the sequences in DWOGT by four methods, T-PCA, M-PCA, PCR, and DMT, respectively. Avg.: average value.

Data	T-PCA	M-PCA	PCR	DMT
P3	82.45 ± 16.37	80.67 ± 14.50	65.71 ± 2.66	$\mathbf{96.90 \pm 0.39}$
P4	84.27 ± 5.73	62.81 ± 5.94	79.30 ± 25.59	$\mathbf{93.58 \pm 0.79}$
P5	87.51 ± 10.59	68.03 ± 18.25	58.64 ± 42.49	$\mathbf{93.48 \pm 3.34}$
P6	84.26 ± 5.91	75.56 ± 10.50	81.54 ± 9.41	$\mathbf{94.35 \pm 2.17}$
P7	92.43 ± 7.19	71.21 ± 18.10	90.20 ± 5.26	$\mathbf{94.30 \pm 5.13}$
P8	71.47 ± 12.60	80.84 ± 7.78	72.18 ± 17.07	$\mathbf{96.35 \pm 2.16}$

P10	77.15 ± 12.60	83.28 ± 12.20	79.78 ± 6.52	$\mathbf{97.39 \pm 1.42}$
P11	89.69 ± 11.06	89.62 ± 3.43	62.37 ± 12.22	$\mathbf{95.47 \pm 3.98}$
P12	86.92 ± 10.99	77.24 ± 10.40	84.96 ± 3.72	$\mathbf{97.88 \pm 1.04}$
P13	83.06 ± 8.58	57.17 ± 4.79	91.62 ± 3.75	$\mathbf{92.87 \pm 4.02}$
Avg.	79.29 ± 13.25	76.89 ± 14.14	73.78 ± 22.55	$\mathbf{95.05} \pm 3.61$

Table 3 shows the time of motion tracking, cardiac phase reconstruction and cardiac phase synchronization, respectively, when $\alpha = 0.01$ and $\alpha = 0.1$. For different sequences, the time efficiency differs by the number of angiograms in the sequences. For the same sequence, the time efficiency differs by the number of detected key points and the number is decided by α.

Table 3. Computational time (s) of motion tracking, cardiac phase reconstruction and cardiac phase synchronization, respectively, when $\alpha = 0.01$ and $\alpha = 0.1$.

Data		$\alpha = 0.01$	$\alpha = 0.1$
DMT	Multi-Layer Matching	539.79 ± 52.16	539.79 ± 52.16
	Motion Tracking	40.23 ± 3.79	2.54 ± 0.99
Cardiac Phase Reconstruction		0.40 ± 0.05	0.02 ± 0.01
Cardiac Phase Synchronization		0.19	0.19

4. Discussion

In this manuscript, we proposed a novel and robust deep motion tracking technique for synchronization of cardiac phases from multiview angiographic images and demonstrated its application in X-ray angiographic image sequences from different imaging angles. Our technique was validated by using DWGT which contains 7 clinical sequences of the same patient who underwent ECG-gated examination. DWOGT was also used for the evaluation, which contains 75 clinical sequences from another 12 patients. For the motion tracking, 85% and 90.7% matched points are accurately computed. For the reconstruction of cardiac phases, we established peak temporal distance of 0.027 ± 0.024s and correlations of $89.45\% \pm 9.22\%$ for 7 sequences in DWGT. We also obtained correlations with lowest values of $90.31\% \pm 3.83\%$ and highest values of $98.17\% \pm 1.48\%$ for the 75 X-ray sequences in DWOGT.

For motion tracking, we performed a comparative quantitative evaluation on two groups of consecutive images randomly selected from clinical datasets to validate whether our tracking method is superior to previously published MIR, DeepFlow and EpicFlow methods. Ma et al. (Ma et al., 2015) proposed that the maximum coronary diameter is 5-7 mm. With the minimum magnification of 1.3, the diameter of coronary artery in images is 6.5-9.1mm (22-30 pixels in diameter). Hence, the Euclidean distance less than 3 pixels between the computed point and labelled point can be regarded as accurate matching. However, since DMT is based on the local transformation, tracking errors can be caused when the structures does not appear both in consecutive angiograms. Besides, the tracking errors can also appear when matching point pairs appear in regions with repetitive textures. When compared with other methods, the results showed that our proposed tracking method is much better than MIR, DeepFlow and EpicFlow methods. For MIR method, intensity based similarity metric and affine transformation are utilized. The metric is not suitable for the coronary artery sequences in which patient motion and contrast agent non-uniform injection both cause intensity variation in images. Meanwhile, the motion in coronary artery sequences is non-rigid and cannot accurately computed by an affine transformation. For both DeepFlow and EpicFlow methods, gray value assumed constancy throughout the sequence. However, this assumption is not true in most

coronary artery sequences (Meijering et al., 1999; Meijering et al., 1999). These two methods utilize variational optimization techniques, which cannot accurately deal with regions without anatomical structures (Baker et al., 2011). These may be the reasons that DeepFlow and EpicFlow cannot accurately track points in the images of coronary artery sequences.

For the evaluation of cardiac phases, a comparison was performed on all clinical X-ray angiographic image sequences. The purpose was to determine whether the proposed technique is superior to previously developed T-PCA, M-PCA, or PCR methods. The results indicated that our proposed technique outperforms the T-PCA method, which relies on tracking key points throughout the sequence. In T-PCA, the coordinates of the key points in all the images comprise the matrix. Given that some key points deviates from the centerline of coronary artery, the motion largely differs from homologous points owing to cardiac motion. This may cause the large errors on T-PCA. In the M-PCA and PCR, image intensity may be changed during imaging procedure by other disturbance including clinician operation catheters and non-uniform contrast agent infusion. This may be a reason why M-PCA and PCR methods reconstruct cardiac phases incorrectly.

For the time efficiency of the proposed method, the mean and standard deviation of deep motion tracking and cardiac phase reconstruction are 542.33 ± 53.15s and 0.02 ± 0.01s, respectively. The time of cardiac phase synchronization is 0.19s for 7 sequences. Since the mutli-layer matching based motion tracking has low time efficiency, the current method cannot be applied in the intra-operative image-guided surgical navigation. However, considering cardiac phase reconstruction and synchronization in the proposed method require considerably low computational time, an extension of motion tracking based on deep learning will be considered. By combining the high accuracy and efficiency of deep learning, the proposed method can then be suitable for intra-operation.

The proposed technique is clinical-workflow-friendly and requires no fiducial markers. In clinical practice, the technique has the potential to synchronize X-ray angiographic images from sequences at different imaging angles. The process is very important for the 3D+t reconstruction of the coronary artery because it generates clear and dynamic structural information for physicians.

5. Conclusions

We presented a novel and potentially clinically useful cardiac phase synchronization technique based on deep motion tracking and applied it to the automatic synchronization in X-ray image sequences. Unlike previously developed synchronization methods, our technique is robust to motion of complicated weak textures and multiple motion coexistence. Thus, it is suitable to X-ray angiographic images that contain different types of motions, large weak texture regions, and multiple physical structures. One major limitation of the proposed method is based on the motion tracking. When the motion tracking are not accurately computed, the precision of the cardiac phase reconstruction will be greatly affected. This situation often occurs when the vascular structures present missing segments.

Acknowledgement

This work was supported by the National Key R&D Program of China (2017YFC0107900), and the National Science Foundation Program of China (61672099, 81627803, 61501030, 61527827). AFF is partially funded by the Engineering and Physical Sciences Research Council (EP/N026993/1), the Royal Society (International Exchanges Program IE141258), and the European Commission (H2020-SC1-PM-16-2017-777119).

Authors' contributions

SS, YJ and DCB conceived and conducted the experiments; SS, DCB, LXX and ADN analyzed the results and wrote the paper; YJ, AFF, HY, SH, JYR, ADN and WYT reviewed the manuscript and provided many thoughtful suggestions to improve the manuscript.

Conflict of Interest

The authors declare that they have no competing interests.

Ethical statement

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

Kurra, V., Kapadia, S.R., Tuzcu, E.M., Halliburton, S.S., Svensson, L., Roselli, E.E., and Schoenhagen, P 2010 Pre-procedural imaging of aortic root orientation and dimensions: comparison between X-ray angiographic planar imaging and 3-dimensional multidetector row computed tomograp JACC: Cardiovascular Interventions **3** 105-113

Chen, Y., Shi, L., Feng, Q., Yang, J., Shu, H., Luo, L., Coatrieux, J.-L., and Chen, W 2014 Artifact suppressed dictionary learning for low-dose CT image processing IEEE Transactions on Medical Imaging **33** 2271-2292

Cong, W., Yang, J., Ai, D., Chen, Y., Liu, Y., and Wang, Y 2016 Quantitative Analysis of Deformable Model-Based 3-D Reconstruction of Coronary Artery From Multiple Angiograms. IEEE Transactions on Biomedical Engineering **63** 1389-1400

Yang, J., Wang, Y., Liu, Y., Tang, S., and Chen, W 2009 Novel approach for 3-D reconstruction of coronary arteries from two uncalibrated angiographic images IEEE Transactions on Image Processing **18** 1563-1572

Yang, J., Cong, W., Chen, Y., Fan, J., Liu, Y., and Wang, Y 2014 External force back-projective composition and globally deformable optimization for 3-D coronary artery reconstruction Physics in Medicine and Biology **59** 975-1003

Çimen, S., Gooya, A., Grass, M., and Frangi, A.F 2016 Reconstruction of coronary arteries from X-ray angiography: A review Medical Image Analysis **32** 46-68

Lauritsch, G., Boese, J., Wigström, L., Kemeth, H., and Fahrig, R 2006 Towards cardiac C-arm computed tomography IEEE Transactions on Medical Imaging **25** 922-934

Lehmann, G.C., Holdsworth, D.W., and Drangova, M 2006 Angle-independent measure of motion for image-based gating in 3D coronary angiography Medical Physics **33** 1311-1320

Sundar, H., Khamene, A., Yatziv, L., and Xu, C 2009 Automatic image-based cardiac and respiratory cycle synchronization and gating of image sequences International Conference on Medical Image Computing and Computer-Assisted Intervention **2009** 381-388

Toth, D., Panayiotou, M., Brost, A., Behar, J.M., Rinaldi, C.A., Rhode, K.S., and Mountney, P 2017 3D/2D Registration with superabundant vessel reconstruction for cardiac resynchronization therapy Medical Image Analysis **42** 160-172

Panayiotou, M., King, A.P., Housden, R.J., Ma, Y., Cooklin, M., O'Neill, M., Gill, J., Rinaldi, C.A., and Rhode, K.S 2014 A statistical method for retrospective cardiac and respiratory motion gating of

interventional cardiac X-ray images Medical Physics **41** 071901

Panayiotou, M., King, A.P., Bhatia, K.K., Housden, R.J., Ma, Y., Rinaldi, C.A., Gill, J., Cooklin, M., O'Neill, M., and Rhode, K.S 2013 Extraction of cardiac and respiratory motion information from cardiac x-ray fluoroscopy images using hierarchical manifold learning International Workshop on Statistical Atlases and Computational Models of the Heart **2013** 126-134

Frangi, A.F., Niessen, W.J., Vincken, K.L., and Viergever, M.A 1998 Multiscale vessel enhancement filtering International Conference on Medical Image Computing and Computer-Assisted Intervention **1998** 130-137

Panayiotou, M., King, A., Ma, Y., Housden, R., Rinaldi, C., Gill, J., Cooklin, M., O'Neill, M., and Rhode, K 2013 A statistical model of catheter motion from interventional x-ray images: application to image-based gating Physics in Medicine and Biology **58** 7543-62

Brost, A., Wimmer, A., Liao, R., Hornegger, J., and Strobel, N 2011 Motion compensation by registration-based catheter tracking SPIE Medical Imaging **7964** 79641O

Chen, Y., Zhang, Y., Yang, J., Cao, Q., Yang, G., Chen, J., Shu, H., Luo, L., Coatrieux, J.-L., and Feng, Q 2016 Curve-like structure extraction using minimal path propagation with backtracking IEEE Transactions on Image Processing **25** 988-1003

Harris, C., and Stephens, M 1988 A combined corner and edge detector Alvey Vision Conference **1988** 10-5244

Revaud, J., Weinzaepfel, P., Harchaoui, Z., and Schmid, C 2016 Deepmatching: Hierarchical deformable dense matching International Journal of Computer Vision **120** 300-323

Meagher, D. Geometric modeling using octree encoding 1982 Computer Graphics and Image Processing **19** 129-147

Fischler, M.A., and Bolles, R.C 1981 Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography Communications of the ACM **24** 381-395

Potel, M.J., MacKAY, S.A., Rubin, J.M., Aisen, A.M., and Sayre, R.E 1984 Three-dimensional left ventricular wall motion in man. Coordinate systems for representing wall movement direction. Investigative Radiology **19** 499-509

Chen, C.W., Huang, T.S., and Arrott, M 1994 Modeling, analysis, and visualization of left ventricle shape and motion by hierarchical decomposition IEEE Transactions on Pattern Analysis and Machine Intelligence **16** 342-356

Marion, J.B 2013 Classical dynamics of particles and systems Academic Press **2013**

Hernandez-Sabate, A., Gil, D., Garcia-Barnes, J., and Marti, E 2011 Image-based cardiac phase retrieval in intravascular ultrasound sequences IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **58** 60-72

Nejati, M., and Pourghassem, H. 2014 Multiresolution image registration in digital X-Ray angiography with intensity variation modeling Journal of Medical Systems **38** 10

Weinzaepfel, P., Revaud, J., Harchaoui, Z., and Schmid, C 2013 DeepFlow: Large displacement optical flow with deep matching IEEE Conference on Computer Vision **2013** 1385-1392

Revaud, J., Weinzaepfel, P., Harchaoui, Z., and Schmid, C 2015 Epicflow: Edge-preserving interpolation of correspondences for optical flow IEEE Conference on Computer Vision and Pattern Recognition **2015** 1164-1172

Smith, P.W 2012 A Practical Guide to Splines (Carl de Boor) Siam Review **22** 520-521

Benesty, J., Chen, J., Huang, Y., and Cohen, I 2009 Pearson correlation coefficient Noise Reduction in Speech Processing **2009** 1-4

Ma, H., Dibildox, G., Schultz, C., Regar, E., and van Walsum, T 2015 PCA-derived respiratory motion surrogates from X-ray angiograms for percutaneous coronary interventions International Journal of Computer Assisted Radiology and Surgery **10** 695-705

Meijering, E.H., Niessen, W.J., and Viegever, M 1999 Retrospective motion correction in digital subtraction angiography: a review IEEE Transactions on Medical Imaging **18** 2-21

Meijering, E.H., Zuiderveld, K.J., and Viergever, M.A 1999 Image registration for digital subtraction angiography International Journal of Computer Vision **31** 227-246

Baker, S., Scharstein, D., Lewis, J.P., Roth, S., Black, M.J., and Szeliski, R 2011 A Database and Evaluation Methodology for Optical Flow International Journal of Computer Vision **92** 1-31