3D Shape Reconstruction of Lumbar Vertebra From Two X-ray Images and a CT Model

Longwei Fang, Zuowei Wang, Zhiqiang Chen, Fengzeng Jian, Shuo Li, and Huiguang He, Senior Member, IEEE

Abstract-Structure reconstruction of 3D anatomy from biplanar X-ray images is a challenging topic. Traditionally, the elastic-model-based method was used to reconstruct 3D shapes by deforming the control points on the elastic mesh. However, the reconstructed shape is not smooth because the limited control points are only distributed on the edge of the elastic mesh. Alternatively, statistical-model-based methods, which include shape-model-based and intensity-model-based methods, are introduced due to their smooth reconstruction. However, both suffer from limitations. With the shape-model-based method, only the boundary profile is considered, leading to the loss of valid intensity information. For the intensity-based-method, the computation speed is slow because it needs to calculate the intensity distribution in each iteration. To address these issues, we propose a new reconstruction method using X-ray images and a specimen's CT data. Specifically, the CT data provides both the shape mesh and the intensity model of the vertebra. Intensity model is used to generate the deformation field from X-ray images, while the shape model is used to generate the patient specific model by applying the calculated deformation field. Experiments on the public synthetic dataset and clinical dataset

Manuscript received January 23, 2019; revised March 26, 2019; accepted April 2, 2019. This work was supported in part by The National Key Research and Development Program of China (2018YFC2001302), the National Natural Science Foundation of China (61976209), CAS International Collaboration Key Project (173211KYSB20190024), and Strategic Priority Research Program of CAS (XDB32040000). Recommended by Associate Editor Chenglin Liu. (Corresponding author: Huiguang He.)

Citation: L. W. Fang, Z. W. Wang, Z. Q. Chen, F. Z. Jian, S. Li, and H. G. He, "3D shape reconstruction of lumbar vertebra from two X-ray images and a CT model," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 4, pp. 1124–1133, Jul. 2020.

- L. W. Fang and Z. Q. Chen are with the Research Center for Brain-Inspired Intelligence & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: fanglongwei2014@ia.ac.cn; chenzhiqiang14@mails.ucas.ac.cn).
- Z. W. Wang is with the Department of Neurosurgery, Xuanwu Hospital, Capital Medical University, Division of Spine, China International Neurological Institute, Beijing 100053, and also with Department of Neurosurgery, Beijing Hospital, National Center of Gerontology, Beijing 100730, China (e-mail: wzw6855@163.com).
- F. Z. Jian is with the Department of Neurosurgery, Xuanwu Hospital, Capital Medical University, Division of Spine, China International Neurological Institute, Beijing 100053, China (e-mail: fengzengjian@hotmail.com).
- S. Li is with the Department of Medical Biophysics, Schulich School of Medical and Dentistry, and also with the Department of Medical Imaging, Schulich School of Medical and Dentistry, University of Western Ontario, ON N6A 4V2, Canada (e-mail: slishuo@gmail.com).
- H. G. He is with the Research Center for Brain-Inspired Intelligence & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, and with the University of Chinese Academy of Sciences, Beijing 100049, and also with the Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Beijing 100190, China (e-mail: huiguang.he@ia.ac.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2019.1911528

show that the average reconstruction errors are 1.1 mm and 1.2 mm, separately. The average reconstruction time is 3 minutes.

Index Terms—2D/2D registration, 2D/3D registration, 3D reconstruction, vertebra model, X-ray image.

I. INTRODUCTION

MAGE guided radiotherapy is widely applied in the hospital-setting because it provides valuable anatomical model information to help doctors understand pathology [1]–[4]. Images shown in 3D form provide more clear anatomical structures and spatial information than 2D images [4], [5]. However, it is inconvenient to get 3D form data during surgery since 3D data collecting devices used in the surgery are so expensive that ordinary hospitals cannot afford it. For ordinary hospitals, a 2D X-ray image collecting device is widely used in the surgery. Recently, 3D anatomy shape reconstruction using 2D X-ray images has been developed [6]–[9]. Several corresponding methods have been proposed, such as the elastic-mesh-based method and statistical based model.

For elastic-mesh-based methods, an priori elastic mesh model, which is also called the base model, is introduced to reconstruct the 3D shape. Control points are distributed along the edges of the base model. For reconstruction, the based model is deformed through an optimal procedure so that the projection of control points match the corresponding points in X-ray images. The most commonly used strategy for this kind of method is using the stereo or non-stereo correspondence of control points [10]–[12]. Mitton et al. [10] and Mitulescu et al. [11] deformed an elastic mesh by using stereocorresponding points (SCPs) and non-stereo-corresponding points (NSCPs) that available in different projections. The correspondence was built by rigid registration of control points with the identified points in X-ray images. Then 3D SCPs were reconstructed by applying the direct linear transformation (DLT) algorithm. Subsequently, NSCPs and their 3D coordinates were reconstructed by defining lines connecting the projections of anatomical landmarks and the ray source in the 2D image. The optimization procedure was achieved by finding the best position on these lines by calculating shape similarity with the base elastic mesh. Laporte et al. [13] improved the NSCP by offering a nonstereo corresponding contour (NSCC) in their method. Instead of using non-stereo points, they applied non-stereo contours to increase the number of control points. The main idea of their method was to build the correspondence of 2D contours which were generated from radiographs to 3D outlines that were projected from the surface of the elastic mesh. A coarse to fine strategy was also used to improve reconstruction accuracy.

Mitton *et al.* [14] improved NSCC by progressively reconstructing a coarse, intermediate, and fine personalized model. The elastic-model-based methods just use a single priori model to reconstruct the patient-specific model. The advantage of elastic-mesh-based methods is that they only use a single priori model for the reconstruction. However, the limitation is that the control points are only distributed along the model edge, which may lead to the un-smooth reconstruction.

For statistical-model-based methods, a deformable mean model and characteristic factors are extracted from a series of prior models. The mean model is deformed through a characteristic factors selection optimal procedure until the projection of the deformed mean model can match the X-ray images. According to the status of the mean model, statisticalmodel-based methods can be classified into two categories: statistical shape model (SSM) methods which use shape model as the mean model and statistical intensity model (SIM) methods which use an intensity model as the mean model. SSM based methods project the outlines of the mean model to match the profiles of the X-ray images. A distance metric is widely used in the SSM method for building the correspondence between profiles of the X-ray images and projection outlines of the training geometry [15], [16]. Fleute et al. [17] built the correspondence by using an iterative closest point (ICP) algorithm. Zhu et al. [18] restricted the SSM to the contours of X-ray images by using a constraint equation. Unlike ICP, this equation did not deform the projected outlines. Baka et al. [19] did not build the correspondence using a distance metric which was used in [17], [18]. Instead, the correspondences were weighted by the orientation difference of the X-ray images edges and the projected silhouettes. The limitation of SSM based methods is that they only use the boundary information from X-ray images. Thus, the abundant intensity information inside X-ray images is wasted. Moreover, the restriction is only distributed along the projection outlines, which may not provide adequate restrictions for a good reconstruction. SIM-based methods can avoid this limitation. The projection of the intensity mean model is known as digital reconstructed radiography (DRR). The optimal procedure of SIM-based methods is achieved until the intensity distribution of the projected DRR can match the intensity distribution of the X-ray image. SIM-based methods make full use of the intensity information from 2D X-ray images, and they can achieve a higher reconstruction accuracy. Whitmarsh et al. [20] used an intensity-based 2D/3D registration strategy to register the DRRs onto two bone densitometry (DXA) images. The best model was found by adjusting the parameters of the mean intensity model which resulted in the DRRs having the smallest difference compared to the DXA images. Yao et al. [21] and Hurvitz et al. [22] shared a similar reconstruction strategy with Whitmarsh, except Yao built the shape and intensity model into one equation, and Hurvitz built an active appearance model (AAM), which consists of three related components: the template intensity image, the template bone surface, and the statistical shape model. SIM-based methods make full use of the intensity information from 2D X-ray images, which can achieve higher reconstruction accuracy. However, the

reconstruction costs a large amount of time because the probability distribution of the intensity model needs to be calculated in each iteration in 3D space.

There are also other reconstruction methods to reconstruct 3D model from X-ray images, such as Purisha *et al.* [23], who reconstructed a 3D model from sparse data. Vukicevic *et al.* [24] reconstructed directly from X-ray images without using any 3D prior model. However, none of these methods are suitable for reconstruction of a complex structure like vertebra.

In this paper, a novel reconstruction method is proposed by using two X-ray images and the CT data of a vertebra specimen. The introduced CT data could provide both the elastic shape mesh and the intensity model of the vertebra. DRRs are generated from the intensity model, which are used to calculate the difference from patient X-ray images. While the elastic shape model is deformed by the calculated difference to generate the patient specific model. The novelty of our method is described in three parts:

- 1) By taking the advantage of both the intensity-model-based and elastic-model-based methods, the proposed algorithm only uses one prior model and makes good use of all X-ray images.
- 2) Compared with the elastic-model-based method, the control points densely distribute on the whole elastic mesh. Thus, the reconstruction result is smooth.
- 3) Compared with the statistical-intensity-based method, the proposed method largely reduces the computational cost.

II. MATERIALS AND METHODS

A. A Brief Introduction to Reconstruction Procedure

The novelty of the proposed method is that we introduce the CT data of a vertebra specimen in reconstruction. The CT data can provide both the elastic mesh and the intensity model. Specifically, we use 2D/3D registration to estimate the projection parameters and 2D/2D registration to estimate the deformation field between the intensity model and two perpendicular X-ray images. Then we use the deformation field and projection parameters to deform the elastic mesh to generate the patient-specific 3D model. The whole reconstruction process is completed in four steps: Firstly, projection parameters are calculated by 2D/3D registration between the intensity model and X-ray images. Secondly, correspondences are built between the intensity model and DRRs. Thirdly, 2D/2D registration between X-ray images and DRRs is applied to calculate the deformation fields. Lastly, the elastic model is deformed using the calculated deformation field, projection parameters and correspondences. The flowchart of the vertebra reconstruction is shown in Fig. 1.

B. Specimen Description

The vertebra specimen is provided by Beijing Hospital. It is the third lumbar vertebra of the human body. Thirteen landmarks are labeled on the vertebra, six of them are posted on the vertebral body, two of them on the transverse process, two of them on the superior articular process, and the last one on the spinous process. Those landmarks are used to register the CT data and X-ray images. The resolution of CT is

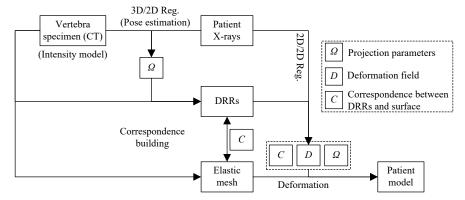


Fig. 1. The flowchart of vertebra reconstruction.

 $0.24~\text{mm} \times 0.24~\text{mm} \times 0.7~\text{mm}$, which is collected from the GE Discovery HD720. Fig. 2(a) shows the real vertebra specimen model and its landmarks location. Fig. 2(b) is the 3D model reconstructed from the CT data and the landmarks' location.

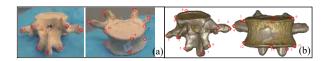


Fig. 2. The location of landmarks on vertebra specimen: (a) on the physical model; (b) on the reconstructed 3D model. Red circles are the location of landmarks, the numbers near the red circles are the serial number of the landmarks, which are fixed.

C. Pose Estimation

Projection parameters include the projection direction and the position of ray sources when putting the 3D CT and X-ray images into the same space. Projection parameters are calculated in two steps: labeling landmarks on segmented vertebrae and 2D/3D landmark based rigid registration between the CT data and the segmented vertebra. The schematic of pose estimation is shown in Fig. 3.

1) Vertebra Segmentation and Landmarks Labeling: Precisely segmenting the vertebra from X-ray images is a challenging job because of the complex background and the overlap between adjacent vertebras. In this paper, Intelligent Scissorsis [25] is used to segment the reconstruction part. Intelligent Scissoris is an interactive algorithm which uses mouse gesture motions to extract objects from the image. A live-wire boundary wraps around the segmented object when the mouse location coming proximately to the object edge. Fig. 4(a) is the presentation of Intelligent Scissors on original image and Fig. 4(b) is the segmentation result.

The next step is labeling the landmarks on the segmented images. The position of the landmarks is based on the markers' position on the physical model, which is shown in Fig. 2(a). The principle of labeling landmarks on two perpendicular segmented images is that 1) seen from the ray source, only visible landmarks are labeled; 2) landmark numbers which are labeled in two images can be different. Fig. 4(c) shows the labeled landmarks on the two perpendicular segmented images. The number in two images is different.

2) 2D/3D Registration Based on Landmarks: A 2D/3D landmark-based registration strategy is applied to compute the projection parameter between X-ray images and the intensity model. In this paper, two X-ray images are from perpendicular positions, which called the sagittal image and coronal image. As shown in Fig. 4(a). Q represents projection parameters, P represents the projection of landmarks from the intensity model, and Q represents the landmarks labeled on the X-ray image, then projection parameters can be solved by

$$\hat{\Omega} = \arg\min_{\Omega} \sum_{\substack{i,l\\k \neq (s,c)}} \left(\theta_{l(k)}^{i}(\Omega) + \alpha \left\| P_{k}^{8}(\Omega) - Q_{k}^{8}(\Omega) \right\|^{2} \right)$$
(1)

where $\theta_{l(k)}^{i}$ represents the angle between the projection line and landmark line on the k plane. Here, the projection line means that the line connects two projection points, and the landmark line means that the line connects two landmarks. The unit of the angle is radians. k = s and k = c mean the sagittal and coronal plane, respectively. i and l represents landmark sets (see in Fig. 2). When l = 8, $i \in \{1, 2, 3, 4, 9, 10\}$, and l = 13, $i \in \{5, 6, 7, 11, 12\}$. We should notice that the value of i is not a constant value and it can be a partial or entire value in its set. The value depends on the landmarks' position which are marked on X-ray images. For example, as shown in Fig. 5, we only use seven landmarks to estimate the pose of the CT data, and the landmarks in two projection planes are different. $\| \bullet \|$ means the Euclidean distance, α is the ratio to balance the distance and angle in the optimization process, $\alpha = 0.017$ is empirically found to be effective in registration. $\Omega = (\Omega_s, \Omega_c)$, Ω_s and Ω_c are the projection parameters of two ray source, which includes three translation parameters and three rotations along three axes. P_k^8 represents the projection of the 8th landmark on the kth plane, $P_k^8 = [x, y]^T$, Q_k^8 represents the 8th landmark on the kth X-ray image plane, $Q_k^8 = [x, y]^T$, and $[x, y]^T$ represents the pixel location. The center of the CT data is the original coordinate, and the coordinate of ray source is $S = [0,0,t_z]^T$. d represents the distance value from the projection plane to the ray source. Three points are chosen from the projection plane, which are $p_0 = [x_0, y_0, t_z - d]^T$, $p_1 = [x_1, y_1, t_z - d]^T$, and $p_2 = [x_2, y_2, t_z - d]^T$. The coordinate of the *i*th landmark is $l_i = [x_i, y_i, z_i]^T$. For the projection of landmark i on sagittal plane, the transformed ray source S_s is calculated in

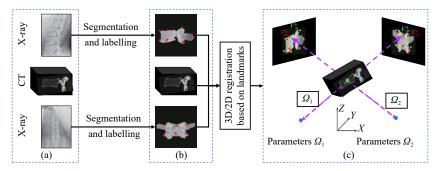


Fig. 3. Flowchart of pose estimation: (a) input images; (b) segmentation results (Red dots are the labeled landmarks); (c) the schematic of 2D/3D registration (Red dots are manually added landmarks in X-ray images, the green dots are the projection of landmarks from CT image).

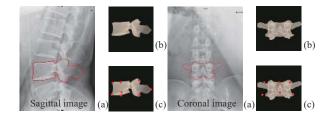


Fig. 4. Vertebra segmentation and landmarks labeling: (a) The presentation of Intelligent Scissors on original image (b) segmentation result by Intelligent Scissors (c) the result of adding landmarks to segmented images.

$$S_s = T(\Omega_s)S = [x_s, y_s, z_s]^T.$$
 (2)

The transformed points in the projection plane are calculated by

$$\begin{cases} p'_0 = T(\Omega_s)p_0 = [x'_0, y'_0, z'_0]^T \\ p'_1 = T(\Omega_s)p_1 = [x'_1, y'_1, z'_1]^T \\ p'_2 = T(\Omega_s)p_2 = [x'_2, y'_2, z'_2]^T. \end{cases}$$
(3)

So, the projection of the landmark on the transformed projection plane is the intersection point of the line and plane, which is $P' = [x', y', z']^T$, calculated by

$$P' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x_s - x_i & x'_1 - x'_0 & x'_2 - x'_0 \\ y_s - y_i & y'_1 - y'_0 & y'_2 - y'_0 \\ z_s - z_i & z'_1 - z'_0 & z'_2 - z'_0 \end{bmatrix}^{-1} \begin{bmatrix} x_s - x'_0 \\ y_s - y'_0 \\ z_s - z'_0 \end{bmatrix}. \tag{4}$$

Then after transforming back to the transformed projection plane, the projection of point i is $\bar{P} = [\bar{x}, \bar{y}, \bar{z}]^T$, calculated by

$$\bar{P} = [\bar{x}, \bar{y}, \bar{z}]^T = T^{-1}(\Omega_S)P'. \tag{5}$$

We assume the pixel space in the X-ray image is $\mu = [\mu_x, \mu_y]^T$, and the projection of landmark $P_s^i = [x, y]^T$ is represented as

$$P_S^i = [x, y]^T = [\frac{\bar{x}}{\mu_x}, \frac{\bar{y}}{\mu_y}]^T.$$
 (6)

The projection in the coronal plane can be calculated in the same way. Equation (1) is a non-convex function, in this paper, two optimization stages are used to solve the equation. We first optimize the distance component and then optimize the angle component. Specifically, the optimization of the distance component makes the projection of landmark 8 overlap with the manually labeled landmark 8 on the X-ray image. We then fix the translation parameters and optimize

the rotation parameters. Gradient descent is used to iteratively optimize the distance component. For rotation parameters, greedy search is used to seek the optimal parameters. Particularly, we first search in a search domain every 5 degrees and select 2 rotation parameters which make the top 2 minimal values for (1) as the new search center. Then we search in these sub-regions by every 1 degree, and select the rotation parameters that minimize the (1) as the search domain and search in this subregion in every 0.2 degrees to get the optimal result. These two stages are optimized iteratively. In this method, the number of iterations is set to 3. Fig. 5 is the schematic of 2D/3D registration based on landmarks. Red dots are the landmarks labeled by hand in X-ray images, and green dots are the projection of landmarks.

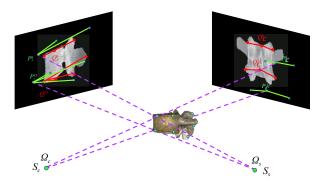


Fig. 5. The schematic of 2D/3D registration based on landmarks; red dots are the landmarks labeled by hand in X-ray images while green dots represent the projection of landmarks.

D. Correspondence Building

As we have mentioned before, CT data can be used to generate the surface mesh of the specimen and the intensity model. Because the surface mesh and projection images are generated from the same source, we can build a connection between pixels of the projection image and vertexes of the surface mesh. This process can be achieved in three steps: extract the elastic mesh from the CT data, generate DRRs from the intensity model, and build the correspondence between the projection of control points and DRRs. The flowchart of the whole procedure is shown in Fig. 6, red points are examples where control points are projected in two planes and blue dots represent the ray sources.

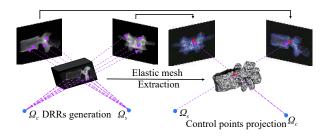


Fig. 6. Flowchart of correspondence building between the elastic mesh and DRRs. Red points are examples that a control point projected in two planes while blue dots represent the ray sources.

- algorithm to segment lumbar vertebra model from the CT data of the vertebra specimen and then use open source software Medical Image Interaction Toolkit (MITK)¹ to extract its surface. OTSU is a threshold-based segmentation method that can automatically choose the best threshold to segment the data. This algorithm is integrated into MITK. MITK is a medical image processing software that was written based on the open source code Visualization Toolkit (VTK) and Insight Segmentation and Registration Toolkit (ITK). The extracted surface, also called an elastic mesh, is stored in a triangular format. The mesh vertexes are the control points to control the deformation of the elastic model.
- 2) DRRs Generation: DRRs, also called projection images, are the projections of the intensity model. A ray casting [27] is used to generate the DRRs. Ray-casting algorithm simulates the generation process of the X-ray image. It equidistant samples the volume and calculates the attenuation value of the sampling point when the ray gets through the CT volume, and then synthesizes the value of all the sampling points along the ray path to generate the pixel intensity of this ray. Projection parameters calculated from (1) are used to generate the DRRs. Two perpendicular DRRs are shown in Fig. 7(a).
- 3) Control Points Projection: After we generate DRRs and the elastic mesh, we can build the correspondence between control points and two DRRs. Extending the lines, which connects the ray source and control points, to the projection plane, the intersection points are the projected control points. Basically, the calculated value is not an integer, thus we use linear interpolation to find the correspondence location in the DRR image. In this way, the projection of control points moves with the moving of correspondence pixels in DRRs. Projections which are generated from the same control point in two DRRs are a control point pair, and are used to generate the 3D point by back projection. The projection of the control points is shown in Fig. 7(b). Blue points are the projection of control points, and the red ones in two planes are generated from the same control point.

E. Registration

Deformable registration between DRRs and X-ray images involves finding deformation fields that can deform the DRRs into X-ray images. A lot of algorithms are proposed to improve the deformable registration. Du *et al.* [28] built the

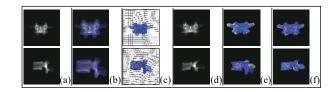


Fig. 7. The intermediate results for computing deformation fields of projection images: (a) DRRs that generated in two perpendicular directions; (b) project control points into DRRs; (c) the deformation field from DRRs to X-ray images; (d) the overlap of the X-ray images and the DRRs before registration; (e) the overlap of the X-ray images and the DRRs after registration. The red dots shown in the figure represent the projection points generated from the same control point.

dynamic population graph for accurate correspondence detection between X-rays images. In their following works [29], [30], they proposed two new algorithms to handle the noise and outliers of the image. Balakrishnan *et al.* [31] used a learning strategy to register the medical image. Here, B-Spline and mutual information [32] techniques are used to register these two modalities images. Mutual information is commonly used to register images from different modality, and achieve a good result.

The deformation field from DRRs to X-ray images is shown in Fig. 7(c). The overlap of the X-ray images and the DRRs before the registration are shown in Fig. 7(d), and the overlap of the X-ray images and the DRRs after the registration are shown in Fig. 7(e). Red dots in Fig. 7 represent the projection points generated from the same control point.

F. Mesh Deformation

After we determine the deformation field, the elastic mesh of the vertebra which is generated in Section II-D-1) can be deformed to build the patient-specified vertebra. The projections of control point i in the two planes are represented by $V_c^i(\Omega_c) = [x_c^i, y_c^i]^T$ and $V_s^i(\Omega_s) = [x_s^i, y_s^i]^T$, respectively. D_c^i, D_s^i represents the deformation vector in the coronal projection image and sagittal projection image respectively, where $D_c^i = [d_{cx}^i, d_{cy}^i]^T$, $D_s^i = [d_{sx}^i, d_{sy}^i]^T$, and $[d_{*x}, d_{*y}]^T$ is the deformation in position $[x, y]^T$, and N_c^i, N_s^i is the new position after deformable registration in two planes. Thus, we can calculate the new position by

$$\begin{cases} N_C^i(\Omega_C) = V_C^i(\Omega_C) + D_C^i = [x_c^i + d_{cx}^i, y_C^i + d_{cy}^i]^T \\ N_S^i(\Omega_C) = V_S^i(\Omega_S) + D_S^i = [x_s^i + d_{sx}^i, y_S^i + d_{sy}^i]^T \end{cases}$$
(7)

Assuming the ray sources of the coronal direction and sagittal direction are S_c and S_s , the intersection point M^i of the line $\overline{S_CN_C^i(\Omega_C)}$ and line $\overline{S_SN_S^i(\Omega_S)}$ is the control point i's new position. The whole reconstruction is finished after moving all the control points. Fig. 8 is the schematic diagram of the reconstruction of a specific patient vertebra. Red dots are the example of projections of one control point before deformation, while yellow dots are the example of the same control point's projections after deformation.

III. EXPERIMENTS

In the proposed method, we do not use the learning strategy

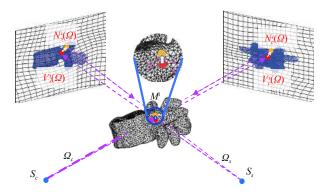


Fig. 8. The schematic diagram of mesh deformation; red dots are the original location and yellow dots are the reconstruct control point.

for reconstruction, the proposed method does not involve any training phase. We just use the cadaver data collected to design our reconstruction algorithm, and then validate the performance on the clinical dataset and public dataset.

A. Validation on the Public Synthetic Dataset

The proposed method is tested on the second dataset of spine public dataset² [33]. This dataset includes ten spine CT scans and their manual segmentation, while the spine CT includes the lumbar vertebrae and the thoracic vertebrae. The resolution of the CT data is $0.31 \text{ mm} \times 0.45 \text{ mm} \times 1 \text{ mm}$. We need to generate the simulated X-ray image since this dataset only provides CT data. The coronal and sagittal position DRR images are used as X-ray images in two perpendicular positions. The resolution of the DRR image is 0.58 mm × 0.58 mm. In this paper, we reconstruct L2, L3, and L4 lumbar vertebrae. The amount of control points is set to 3000. The manual segmentation provided by the dataset is used to calculate the reconstruction accuracy, the average distance and its standard deviation of the control points to ground truth surface is regarded as reconstruction error. They are calculated by the open source software CloudCompare³.

B. Validation on the Clinical Dataset

X-ray images in the two perpendicular positions of six patients acquired from the Beijing Hospital are used to validate our reconstruction algorithm. One of them is collected from a cadaver, and the other five data are real patient data. The X-ray image collection device is Siemens Axiom Multix M, and the resolution of the image is 0.16 mm × 0.16 mm. To validate the reconstruction results, the corresponding CT data is also collected. The resolution of the CT data is 0.61 mm × 1 mm. L2, L3, and L4 vertebra are reconstructed. This study was approved by the ethical committee of Institute of Automation, Chinese Academy of Sciences. All of the subjects signed consent forms in advance. If not specified in the following paragraphs, the influence of parameters are evaluated on the clinical dataset.

C. The Influence of Control Points' Number

In previous experiments, the amount of control points is set

²http://opendatacommons.org/licenses/pddl/1.0/ ³http://www.danielgm.net/cc to 3000. We choose a different amount to validate the influence of the number of control points on reconstruction accuracy. The numbers we choose are 500, 1000, 2000, 3000, 5000, 10 000, 20 000, and 30 000.

D. The Influence of Projection Parameters

The projection parameters are calculated by 2D/3D rigid registration. However, the calculated projection parameters may possibly have a deviation from the real projection parameters since the landmarks are labeled manually. Therefore, the robustness of the proposed method is tested by using different projection parameters near real projection parameters. Specifically, we test this on the public synthetic dataset mentioned in Section III-A. We manually generate two perpendicular images as input, and the projection parameters are regarded as real projection parameters. Then we shift the projection parameters in a small range. In this experiment, we neglect the influence of translation and examine the influence of rotation because of the 2D/2D deformable registration robustness to translation. We shift the rotation angle from real projection parameters for $\pm 1^{\circ}$, $\pm 2^{\circ}$, $\pm 3^{\circ}$, respectively. Then we have 117 649 results. In order to display the results, the shifts of the rotation angle along three axes are simultaneous. For example, the initial projection parameter in the sagittal plane is $r_{sx} = 0$, $r_{sy} = 0$, $r_{sz} = 0$, where the result of a shift of 1° is $r_{sx} = 1$, $r_{sy} = 1$, $r_{sz} = 1$, and the number of results is reduced to 49.

E. The Influence of Image Sizes on Reconstruction Time

The image size has a big influence on registration time since the time cost of 2D/2D deformable registration between DRRs and X-ray images relies on image size. Here, we evaluate the influence of different image sizes on reconstruction time. Image sizes we choose are 512×512 , 450×450 , 350×350 , and 220×220 . The default image size is 512×512 .

We run the experiments on the CentOS6.3 operating system, and the configuration of the machine is: Intel(R) Xeon(R) CPU e5-2687w, the frequency of CPU is 3.10 GHz, the memory of the machine is 64G RAM. We use C++ and MATLAB to write the code.

IV. RESULT AND DISCUSSION

The average reconstruction error on the public synthetic dataset is 1.1 ± 0.9 mm, and the average reconstruction error on clinic dataset is 1.2 ± 1.0 mm. For the reconstruction time, each vertebra requires 3 minutes. This time cost includes the time used for manual segmentation and labeling. The visualization of reconstruction results on the public dataset and clinical dataset are listed in Fig. 9, and L2, L3, and L4 represent the second, third and the fourth vertebra in the human body. The final reconstructions are listed in the left column, and ground truths with the colored error are listed in the right column.

A. The Influence of Control Points' Number

The reconstruction time and accuracy vary with the number of control points. Fig. 10(a) shows the relationship between reconstruction time and the number of control points. Seen from the figure, the reconstruction accuracy increased with the

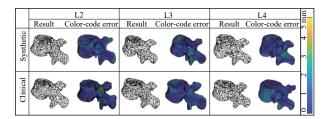


Fig. 9. The reconstruction results in the synthetic dataset and clinical dataset. L2, L3, and L4 represent the second, third, and fourth vertebra in the human body. Reconstruction results are listed in left and ground truths with the colored error in right.

increase of the number of control point. But the reconstruction accuracy does not increase significantly when the number is more than 3000. This can be easily explained by examining our method. When we use fewer control points, which means the elastic mesh has less vertex, it is hard to present anatomic details clearly using limited vertices, so the reconstruction accuracy is lower. When the number of the control points increasing, anatomy detail can be represented clearly, thus increasing the accuracy. However, after the number of control points is sufficient to describe the detail of the elastic model, the accuracy will hardly increase by increasing the vertex number. On the other hand, the reconstruction time has a positive relationship with the control points' number, as shown in Fig. 10(a). The reconstruction of the patient-specific model uses the back projection strategy, it needs to back project each point pair to generate the vertex of the model in 3D space; thus calculation time has a positive relationship to control points' number. From Fig. 10(a), we also find that the reconstruction error decreases sharply from 1.2 mm to 1 mm when the control points number increases from 500 to 3000, but remains stable at 1 mm between 3000 and 10 000. When the control points number is more than 10 000, it needs more points to reduce the error by 0.1 mm. On the other hand, reconstruction time increases with control points' number. Balancing the reconstruction error and time cost, in our experiments, 3000 is the best choice to fulfill the accuracy and time demand.

B. The Influence of Projection Parameters

Table I shows the reconstruction error by different projection parameters. The average reconstruction error does not increase much (the biggest error divergence is 0.2 mm) even when the shift in three directions is three degrees. We also randomly test the projection parameters that do not shift the same angle simultaneously, and find that the error divergence is also within 0.2 mm. For elastic-model-based methods, which restrict the mean shape model only by using boundaries of X-ray images, the contour that is generated by projecting the mean shape model is prone to be affected by projection parameters. For example, a slight variation in projection parameters will lead to a large shape variation in the projection contour and thus will cause the wrong correspondence being built between projection contours and boundaries of the X-ray image. In our method, the projection of the control points is densely distributed on the whole

vertebra. 2D/2D registration uses the full information of the X-ray images, which is more stable than the boundary information. Therefore, our method is robust to the bias of projection parameters.

C. The Influence of Image Sizes on Reconstruction Time

Fig. 10(b) shows the result of different image sizes on reconstruction time. The reconstruction time has a positive relationship with the image size. For the proposed method, the reconstruction time can be divided into five parts. Manual segmentation, 2D/3D registration, DRRs generation and elastic mesh extraction, 2D/2D registration, and elastic mesh deformation. The first three parts use a fixed amount of time, which is not influenced by different variables in the proposed method. The fifth part, elastic mesh deformation, is influenced by the control points' number, as discussed in Section IV-A. The fourth part, 2D/2D registration between X-ray images and DRRs, is the most time-consuming step aside from the manual segmentations. For a large image size, it needs to calculate more interpolation points and needs more time to calculate the probability distribution. Thus, the time cost increases with the increase of image size.

V. CONCLUSION

In this paper, we propose a new method to reconstruct the lumbar vertebra by using the CT data of a vertebra specimen and X-ray images. The introduced CT data can provide both the elastic shape mesh and the intensity model of the vertebra. DRRs are generated from the intensity model, and they are used to calculate the difference from X-ray images. Finally, the elastic mesh model is deformed by the calculated difference to generate the patient specific model. Our method is evaluated on both the public synthetic dataset and clinical dataset. The average reconstruction errors on these two datasets are 1.1 mm and 1.2 mm, respectively, and the average construction time for both datasets is 3 minutes. This reconstruction accuracy is a commonly achieved level [34]–[36] but we use much less time. For application in the surgery, the reconstruction error needs to be less than 1.5 mm, and our method can fulfill this requirement. We also discuss the influence of some parameters in the proposed method on the reconstruction accuracy and reconstruction time, like the number of control points, the deviation of projection parameters, and the image size.

The main contribution of the proposed technique is that we use the CT data of a vertebra specimen as prior knowledge. The introduced CT data not only provides an intensity model of the vertebra, which is used to generate the DRRs for 2D/2D intensity-based deformable registration but also can provide the elastic mesh. The proposed method is a combination of the elastic-mesh-based method and statistical-intensity-model-based method. Compared with the elastic-mesh-based method, the control points are evenly distributed on the surface of the whole mesh, not just on the edge of the mesh. This will improve the robustness of the method with regards to the projection parameters and can provide fine control when deforming the vertebral mesh. Moreover, the proposed method makes full use of all information on X-ray images, not

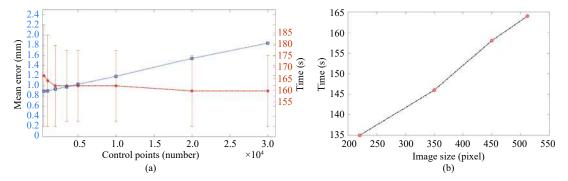


Fig. 10. Reconstruction accuracy and time by different parameters: (a) Reconstruction time and accuracy by different vertex number; (b) Reconstruction time by different image size.

TABLE I
THE RECONSTRUCTION ERROR BY DIFFERENT PROJECTION PARAMETERS

Sagittal	Coronal						
	-3	-2	-1	0	1	2	3
-3	1.0 ± 0.7	1.0 ± 0.8	1.0 ± 0.7	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8
-2	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8	1.1 ± 0.8
-1	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8	1.1 ± 0.8	1.1 ± 0.8
0	1.0 ± 0.7	1.0 ± 0.8	1.0 ± 0.7	1.0 ± 0.7	1.0 ± 0.7	1.0 ± 0.8	1.1 ± 0.8
1	1.0 ± 0.7	1.0 ± 0.7	1.0 ± 0.7	1.0 ± 0.7	1.0 ± 0.7	1.0 ± 0.7	1.0 ± 0.8
2	1.0 ± 0.7	1.0 ± 0.7	1.0 ± 0.7	0.9 ± 0.7	1.0 ± 0.7	1.0 ± 0.7	1.0 ± 0.7
3	0.9 ± 0.7	1.0 ± 0.7	1.0 ± 0.7				

Note: the unit is mm.

just the segmented boundary. On the other hand, compared with the SIM-based method, the proposed method does not need to calculate the intensity distribution in each iteration, which needs less time for reconstruction.

Despite its appealing aspects, the proposed method is limited by manual operation when calculating projection parameters. The proposed method needs to segment the vertebra from X-ray images manually to reduce the segment error caused by the overlap between two adjacent vertebrae. In the future, we will mainly focus on how to segment vertebrae automatically, e.g., using a neural network for automatic segmentation.

ACKNOWLEDGMENT

The authors would like to thank Emily Shaljian and Xunyun Wen for writing assistance.

REFERENCES

- [1] Y. Allam, J. Silbermann, F. Riese, and R. Greiner-Perth, "Computer tomography assessment of pedicle screw placement in thoracic spine: Comparison between free hand and a generic 3D-based navigation techniques," *European Spine J.*, vol. 22, no. 3, pp. 648–653, 2013.
- [2] G. Cui, Y. Wang, T.-H. Kao, Y. Zhang, Z. Liu, B. Liu, J. Li, X. Zhang, S. Zhu, N. Lu, et al., "Application of intraoperative computed tomography with or without navigation system in surgical correction of spinal deformity: A preliminary result of 59 consecutive human cases," Spine, vol. 37, no. 10, pp. 891–900, 2012.
- [3] E. Van de Kelft, F. Costa, D. Van der Planken, and F. Schils, "A prospective multicenter registry on the accuracy of pedicle screw placement in the thoracic, lumbar, and sacral levels with the use of the

- o-arm imaging system and stealthstation navigation," *Spine*, vol. 37, no. 25, pp. E1580–E1587, 2012.
- [4] S. W. Tohtz, P. Rogalla, M. Taupitz, C. Perka, T. Winkler, and M. Putzier, "Inter-and intraobserver variability in the postoperative evaluation of transpedicular stabilization: Computed tomography versus magnetic resonance imaging," *The Spine Journal*, vol. 10, no. 4, pp. 285–290, 2010.
- [5] V. Amato, L. Giannachi, C. Irace, and C. Corona, "Accuracy of pedicle screw placement in the lumbosacral spine using conventional technique: Computed tomography postoperative assessment in 102 consecutive patients," J. Neurosurgery: Spine, vol. 12, no. 3, pp. 306–313, 2010.
- [6] V. Pomero, D. Mitton, S. Laporte, J. A. de Guise, and W. Skalli, "Fast accurate stereoradiographic 3D-reconstruction of the spine using a combined geometric and statistic model," *Clinical Biomechanics*, vol. 19, no. 3, pp. 240–247, 2004.
- [7] S. Benameur, M. Mignotte, S. Parent, H. Labelle, W. Skalli, and J. de Guise, "3D/2D registration and segmentation of scoliotic vertebrae using statistical models," *Computerized Medical Imaging and Graphics*, vol. 27, no. 5, pp. 321–337, 2003.
- [8] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, N. Ayache, et al., "Articulated spine models for 3-D reconstruction from partial radiographic data," *IEEE Trans. Biomedical Engineering*, vol. 55, no. 11, pp. 2565–2574, 2008.
- [9] S. Kadoury, F. Cheriet, and H. Labelle, "Personalized X-ray 3-D reconstruction of the scoliotic spine from hybrid statistical and imagebased models," *IEEE Trans. Medical Imaging*, vol. 28, no. 9, pp. 1422–1435, 2009.
- [10] S. Kolta, A. Le Bras, D. Mitton, V. Bousson, J. A. de Guise, J. Fechtenbaum, J. Laredo, C. Roux, and W. Skalli, "Three-dimensional X-ray absorptiometry (3D-XA): A method for reconstruction of human bones using a dual X-ray absorptiometry device," *Osteoporosis*

- International, vol. 16, no. 8, pp. 969–976, 2005.
- [11] A. Mitulescu, I. Semaan, J. A. De Guise, P. Leborgne, C. Adamsbaum, and W. Skalli, "Validation of the non-stereo corresponding points stereoradiographic 3D reconstruction technique," *Medical and Biological Engineering and Computing*, vol. 39, no. 2, pp. 152–158, 2001.
- [12] D. Mitton, C. Landry, S. Veron, W. Skalli, F. Lavaste, and J. A. De Guise, "3D reconstruction method from biplanar radiography using nonstereocorresponding points and elastic deformable meshes," *Medical* and *Biological Engineering and Computing*, vol. 38, no. 2, pp. 133–139, 2000
- [13] S. Laporte, W. Skalli, J. De Guise, F. Lavaste, and D. Mitton, "A biplanar reconstruction method based on 2D and 3D contours: Application to the distal femur," *Computer Methods in Biomechanics & Biomedical Engineering*, vol. 6, no. 1, pp. 1–6, 2003.
- [14] D. Mitton, S. Deschenes, S. Laporte, B. Godbout, S. Bertrand, J. A. de Guise, and W. Skalli, "3D reconstruction of the pelvis from biplanar radiography," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 9, no. 1, pp. 1–5, 2006.
- [15] S. Benameur, M. Mignotte, H. Labelle, and J. A. De Guise, "A hierarchical statistical modeling approach for the unsupervised 3-D biplanar reconstruction of the scoliotic spine," *IEEE Trans. Biomedical Engineering*, vol. 52, no. 12, pp. 2041–2057, 2005.
- [16] G. Zheng, X. Dong, K. T. Rajamani, X. Zhang, M. Styner, R. U. Thoranaghatte, L.-P. Nolte, and M. A. G. Ballester, "Accurate and robust reconstruction of a surface model of the proximal femur from sparsepoint data and a dense-point distribution model for surgical navigation," *IEEE Trans. Biomedical Engineering*, vol. 54, no. 12, pp. 2109–2122, 2007.
- [17] M. Fleute, S. Lavallee, and R. Julliard, "Incorporating a statistically based shape model into a system for computer-assisted anterior cruciate ligament surgery," *Medical Image Analysis*, vol. 3, no. 3, pp. 209–222, 1999.
- [18] Z. Zhu and G. Li, "Construction of 3D human distal femoral surface models using a 3D statistical deformable model," *J. Biomechanics*, vol. 44, no. 13, pp. 2362–2368, 2011.
- [19] N. Baka, B. L. Kaptein, M. de Bruijne, T. van Walsum, J. Giphart, W. J. Niessen, and B. P. Lelieveldt, "2D-3D shape reconstruction of the distal femur from stereo X-ray imaging using statistical shape models," *Medical Image Analysis*, vol. 15, no. 6, pp. 840–850, 2011.
- [20] T. Whitmarsh, L. Humbert, L. M. D. R. Barquero, S. Di Gregorio, and A. F. Frangi, "3D reconstruction of the lumbar vertebrae from anteroposterior and lateral dual-energy X-ray absorptiometry," *Medical Image Analysis*, vol. 17, no. 4, pp. 475–487, 2013.
- [21] J. Yao and R. Taylor, "Deformable 2D-3D medical image registration using a statistical model: Accuracy factor assessment," *American J. Science and Engineering*, vol. 1, no. 2, pp. 1–13, 2012.
- [22] A. Hurvitz and L. Joskowicz, "Registration of a CT-like atlas to fluoroscopic X-ray images using intensity correspondences," *Int. J. Computer Assisted Radiology and Surgery*, vol. 3, no. 6, pp. 493–504, 2008.
- [23] Z. Purisha, S. S. Karhula, J. H. Ketola, J. Rimpeläinen, M. T. Nieminen, S. Saarakkala, H. Kröger, and S. Siltanen, "An automatic regularization method: An application for 3-D X-ray micro-CT reconstruction using sparse data," *IEEE Trans. Medical Imaging*, vol. 38, no. 2, pp. 417–425, 2019.
- [24] A. M. Vukicevic, S. Çimen, N. Jagic, G. Jovicic, A. F. Frangi, and N. Filipovic, "Three-dimensional reconstruction and nurbs-based structured meshing of coronary arteries from the conventional X-ray angiography projection images," *Scientific Reports*, vol. 8, no. 1, pp. 1711, 2018.

- [25] W. A. Barrett and E. N. Mortensen, "Interactive live-wire boundary extraction," *Medical Image Analysis*, vol. 1, no. 4, pp. 331–341, 1997.
- [26] N. Ohtsu, "A threshold selection method from gray-level histograms," IEEE Trans. Systems Man and Cybernetics, vol. 9, no. 1, pp. 62–66, 1979
- [27] J. Meyer-Spradow, T. Ropinski, J. Mensmann, and K. Hinrichs, "Voreen: A rapid-prototyping environment for ray-casting-based volume visualizations," *IEEE Computer Graphics and Applications*, vol. 29, no. 6, pp. 6–13, 2009.
- [28] S. Du, Y. Guo, G. Sanroma, D. Ni, G. Wu, and D. Shen, "Building dynamic population graph for accurate correspondence detection," *Medical Image Analysis*, vol. 26, no. 1, pp. 256–267, 2015.
- [29] S. Du, J. Liu, B. Bi, J. Zhu, and J. Xue, "New iterative closest point algorithm for isotropic scaling registration of point sets with noise," *J. Visual Communication and Image Representation*, vol. 38, pp. 207–216, 2016.
- [30] D. Shaoyi, X. Guanglin, Z. Sirui, Z. Xuetao, G. Yue, and C. Badong, "Robust rigid registration algorithm based on pointwise correspondence and correntropy," *Pattern Recognition Letters*, 2018.
- [31] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "An unsupervised learning model for deformable medical image registration," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 9252–9260, 2018.
- [32] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "Pet-CT image registration in the chest using free-form deformations," *IEEE Trans. Medical Imaging*, vol. 22, no. 1, pp. 120–128, 2003.
- [33] J. Yao, J. E. Burns, H. Munoz, and R. M. Summers, "Detection of vertebral body fractures based on cortical shell unwrapping," in *Proc.* Int. Conf. Medical Image Computing and Computer Assisted Intervention, pp. 509–516, 2012.
- [34] L. Humbert, J. A. De Guise, B. Aubert, B. Godbout, and W. Skalli, "3D reconstruction of the spine from biplanar X-rays using parametric models based on transversal and longitudinal inferences," *Medical Engineering & Physics*, vol. 31, no. 6, pp. 681–687, 2009.
- [35] D. C. Moura, J. Boisvert, J. G. Barbosa, H. Labelle, and J. M. R. Tavares, "Fast 3D reconstruction of the spine from biplanar radiographs using a deformable articulated model," *Medical Engineering & Physics*, vol. 33, no. 8, pp. 924–933, 2011.
- [36] S. Prakoonwit, "Towards multiple 3D bone surface identification and reconstruction using few 2D X-ray images for intraoperative applications," Int. J. Art, Culture and Design Technologies (IJACDT), vol. 4, no. 1, pp. 13–31, 2014.



Longwei Fang is a Ph.D. candidate at the Research Center for Brain-Inspired Intelligence & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA). He received the master degree from Beihang University in 2013. He received the bachelor degree from China University of Mining & Technology in 2010. His research interests include medical image analysis and image reconstruction.



Zuowei Wang received the bachelor degree from Hunan Medical Science University in 1998, and then received the master and Ph.D. degrees in Huazhong University of Science and Technology and Capital Medical University, respectively. Currently, he is an Associate Professor in Peking University, and he is also an Associate Chief Physician of neurosurgery in Beijing Hospital. His main research interest is minimally invasive spine surgery.



Zhiqiang Chen is a Ph.D. candidate at the Research Center for Brain-Inspired Intelligence & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA). He received the bachelor degree from University of Science and Technology Beijing in 2014. He received the master degree from University of Chinese Academy of Sciences in 2017. His research interests include medical image analysis and brain-like intelligence.



Fengzeng Jian received the bachelor degree from Shandong Medical University in 1990, and then work in Beijing Hospital. During this time, he received the master and Ph.D. degrees. Currently, he is a Professor in Capital Medical University. He is also a Chief Physician of neurosurgery in Xuanwu Hospital. His main research interests include cervical and lumbar spondylosis.



Shuo Li is the Scientific Director of the Digital Imaging Group of London, U.K. He is also an Adjunct Research Professor in Department of Medical Biophysics and Medical Imaging of the University of Western Ontario, Canada. He is also a Research Scientist and Project Manager in the leading healthcare provider: GE Healthcare. His current research interest is focused on the computer automated medical image analysis and visualization.



Huiguang He (M'04–SM'10) received the B.S. and M.S. degrees from Dalian Maritime University (DMU) in 1994 and 1997, respectively, and the Ph.D. degree (Hons.) in pattern recognition and intelligent systems from the Institute of Automation, Chinese Academy of Sciences (CASIA). From 1997 to 1999, he was an Associate Lecturer with DMU. From 2003 to 2004, he was a Post-Doctoral Researcher with the University of Rochester, USA. From 2014 to 2015, he was a Visiting Professor with

the University of North Carolina at Chapel Hill, USA. He is currently a Full Professor with the Research Center for Brain-Inspired Intelligence, National Laboratory of Pattern Recognition, CASIA. He is also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, and the Center for Excellence in Brain Science and Intelligence Technology, CAS. His research has been supported by several research grants from the National Science Foundation of China. He has authored or co-authored over 150 peer-reviewed papers. His current research interests include pattern recognition, medical image processing, and brain—computer interfaces. He was a Recipient of the Excellent Ph.D. Dissertation of CAS in 2004, the National Science and Technology Award in 2003 and 2004, the Beijing Science and Technology Award in 2002 and 2003, the K.C. Wong Education Prizes in 2007 and 2009, and the Jia-Xi Lu Young Talent Prize in 2009. He is an Excellent Member of Youth Innovation Promotion Association, CAS in 2016.