

# SIFT algorithm-based 3D pose estimation of femur

Xuehe Zhang, Yanhe Zhu, Changle Li, Jie Zhao and Ge Li\*

*State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin, Heilongjiang, China*

**Abstract.** To address the lack of 3D space information in the digital radiography of a patient femur, a pose estimation method based on 2D–3D rigid registration is proposed in this study. The method uses two digital radiography images to realize the preoperative 3D visualization of a fractured femur. Compared with the pure Digital Radiography or Computed Tomography imaging diagnostic methods, the proposed method has the advantages of low cost, high precision, and minimal harmful radiation. First, stable matching point pairs in the frontal and lateral images of the patient femur and the universal femur are obtained by using the Scale Invariant Feature Transform method. Then, the 3D pose estimation registration parameters of the femur are calculated by using the Iterative Closest Point (ICP) algorithm. Finally, based on the deviation between the six degrees freedom parameter calculated by the proposed method, preset posture parameters are calculated to evaluate registration accuracy. After registration, the rotation error is less than  $1.5^\circ$ , and the translation error is less than 1.2 mm, which indicate that the proposed method has high precision and robustness. The proposed method provides 3D image information for effective preoperative orthopedic diagnosis and surgery planning.

Keywords: SIFT, femur, 2D–3D registration, pose estimation

## 1. Introduction

Robotics, computer technology, and image processing technology are combined with clinical surgical operation in a new field of study called computer integrated surgical systems and technology (CIS) [1]. CIS aims at combining computed tomography (CT) or magnetic resonance imaging (MRI) information with a 3D positioning system to achieve the preoperative display, preoperative planning, and intraoperative positioning of the human anatomy. Medical robotics and computers are used to aid in surgery. Surgical operation is gradually separated from the work of the hospital surgeon and is transferred to an engineering system that involves engineering technicians and medical rehabilitation personnel [2].

A 3D pose estimation system based on 2D images significantly affects computer-aided diagnosis and virtual operation [3]. This effect can mainly be attributed to the fact that 3D imaging techniques, such as CT, have harmful radiation effects on patients, entail high medical cost, and require complex equipment operation [4]. Therefore, these approaches are unsuitable for use as routine diagnostic

---

\*Corresponding author: Ge Li, State Key Laboratory of Robotics and System, Harbin Institute of Technology, Building C1, Room 203, Science Square, 2 Yikuang Street, Harbin, Heilongjiang, China. Tel.: +86-045186413392; Fax: +86-045186414538; E-mail: lige@hit.edu.cn.

methods for orthopedic diseases. Simply using 2D images for fracture pose estimation cannot satisfy the requirement of accuracy because of the lack of 3D spatial information. For example, Caponetti [5] used B-spline interpolation to complete a 3D reconstruction of the femoral shaft on the basis of traditional radiographs. Dumas [6] used two orthogonal digital radiograph images to reconstruct a 3D pose of the femoral shaft. As a result of limited information, this method only roughly recovers the model framework and has low accuracy. Jerbi [7] presented a well-adapted method to estimate the motion of the bone using a new acquisition technique, the EOS system. The method allows the acquisition of data in weight-bearing positions. It also benefits from this technique by reducing the rate of exposure to X-ray during acquisitions.

In studies of femur surgical replacement, doctors have found that introducing a universal femur 3D model into the process of orthopedic surgery can reduce the misalignment problems of fracture break end reset attributed to positioning error. Therefore, the universal femur 3D model has an important contribution to the development of 2D image navigation surgery and minimally invasive operations. For example, Gamage [8] used a femoral 3D model that reconstructs preoperative rectification with two or more X radiographs captured during surgery, thus completing the robot visual navigation for fracture repair surgery. However, the method has unsatisfactory robustness. The proposed method is a 2D–3D registration system that uses the frontal and lateral digital radiography (DR) images of the patient femur and a universal femoral model to achieve 3D femoral model preoperative visualization. A 2D registration method that uses the scale invariant feature transform (SIFT) algorithm can solve problems associated with image rotation, translation, and other issues in medical image registration. Matching feature points are stable and can well reflect the characteristics of 2D images. The translation vector and rotation matrix are calculated by using the iterative closest point (ICP) algorithm.

## 2. An overview of the proposed algorithm

### 2.1. Main steps of femoral model 3D pose estimation

The objects of femoral 2D–3D image registration are the CT volume data of the universal femur and the patient frontal and lateral DR images. The main steps of femoral 2D–3D image registration are shown in Figure 1.

(1) *Calibration of DR imaging system.* By using the method proposed in the literature [9], calibrate the DR imaging system. The relevant parameters of the imaging system are obtained. These parameters describe the relationships between 2D images and 3D space.

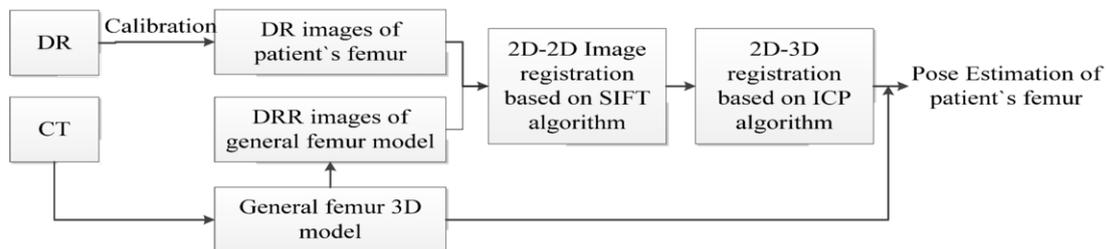


Fig. 1. Main steps of the proposed algorithm.

(2) *Obtaining femoral image data.* Frontal and lateral DR images of the patient femur are obtained. Meanwhile, a universal femur is scanned by using CT to acquire universal femoral CT volume data.

(3) *Generation of digitally reconstructed radiographs.* The calibration parameters are used to simulate a digital radiography system, project the universal femoral CT volume data to a 2D image plane, and obtain digitally reconstructed radiograph (DRR) images. The initial image registration of 2D digital radiography with 3D CT volume data is transformed into 2D DR image and 2D simulation DRR image registration.

(4) *2D–2D image registration and 2D–3D pose estimation.* The objects that need to be matched are the frontal and lateral DR images of patient femur and DRR images of the universal femur. The improved SIFT algorithm is used to detect the matching feature points of these 2D images. The corresponding registration parameters are obtained by using the ICP algorithm. The registration parameters are employed to adjust the 3D universal femoral model and then determine the 3D pose of the patient femur model that has to be reconstructed.

## 2.2. Reconstruction of the universal femur model and DRR generation

The 3D reconstruction of a universal femoral model has laid an important foundation for the realization of femoral 2D–3D image registration. The 3D universal femur model can serve as a reference model for the subsequent 3D pose estimation. In view of the reconstruction technology, the CT scan is relatively mature and can achieve high accuracy in 3D reconstruction [10]. Therefore, we can establish a 3D universal femur model with the aid of CT. The construction of the model mainly involves several steps. First, we conduct CT scan and data preprocessing. On the basis of the characteristics of patient femur, doctors select a piece of femur from the femoral sample as a universal femur. Then it is using CT equipment to accomplish the tomography of universal femur. After a preliminary image processing, universal femur tomography images are imported, and the marching cube algorithm is used to reconstruct the 3D model of universal femur. Second, we obtain the frontal and lateral projection images of the universal femoral 3D model by using the digitally reconstructed radiograph technology. The DRR images are shown in Figures 2(b) and 2(c).

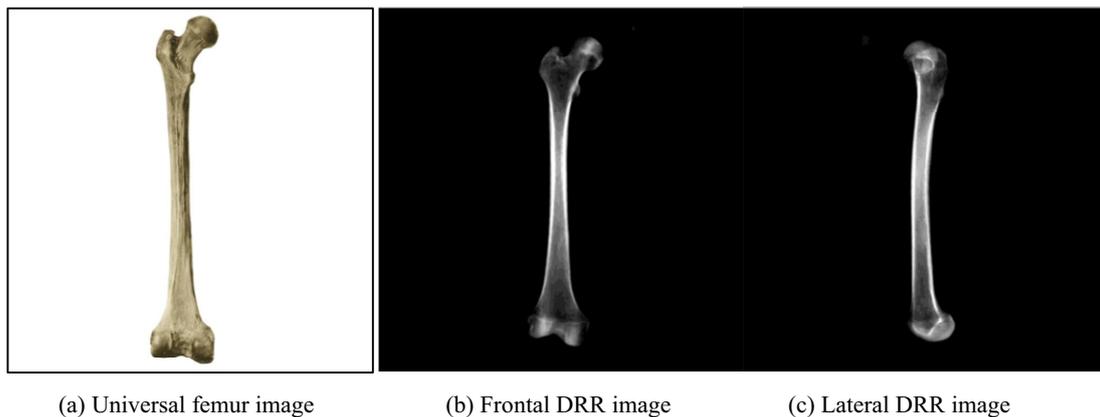


Fig. 2. Universal femur and its projection images.

### 3. 2D–2D registration of projection image based on SIFT algorithm

#### 3.1. Introduction of the SIFT algorithm

The steps of the SIFT algorithm [11] are as follows:

1) *Finding the local extrema in scale-space.* To detect the stabilized keypoints in scale-space effectively, the differential Gaussian scale-space (DOG scale-space) is proposed by using the different scales of the Gaussian kernel convolved with the image.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1)$$

To detect the local maxima and minima of  $D(x, y, \sigma)$ , each point is compared with the pixels of all its 26 neighbors. If this value is the minimum or maximum, then this point is an extremum.

2) *Locating the extreme accurately.* The SIFT algorithm determines the location and dimension of the extreme value point through the second-order Taylor expansion of the DOG function. The second-order Taylor expansion of feature points  $X_0 = (x_0, y_0, \sigma_0)$  is given by

$$D(x, y, \sigma) = D(X_0) + \frac{\partial D^T}{\partial X} X + \frac{1}{2} \frac{\partial^2 D^T}{\partial X^2} X \quad (2)$$

3) *Keypoint orientation invariance.* To ensure that the descriptor possesses orientation invariance, we specify the direction parameters by using the gradient direction distribution characteristic of keypoint neighborhood pixels.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3)$$

$$\theta(x, y) = \alpha \tan 2(((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y)))) \quad (4)$$

For each image sample  $L(x, y)$  at this scale, the gradient magnitude  $m(x, y)$  and orientation  $\theta(x, y)$  are pre-computed by using pixel differences.

4) *Generation of feature descriptor.* To achieve orientation invariance, the coordinates of the descriptor and the gradient orientations are rotated relative to the keypoint orientation. A keypoint descriptor is created by first computing the gradient magnitude and orientation at each image sample point in a region around the keypoint location (16×16 neighborhood pixels). These samples are then accumulated in orientation histograms that summarize the contents over 4×4 subregions, with the length of each arrow corresponding to the sum of the gradient magnitudes near that direction within the region. As a result, a 4×4×8 = 128 element feature vector is generated for each keypoint.

### 3.2. Feature point matching

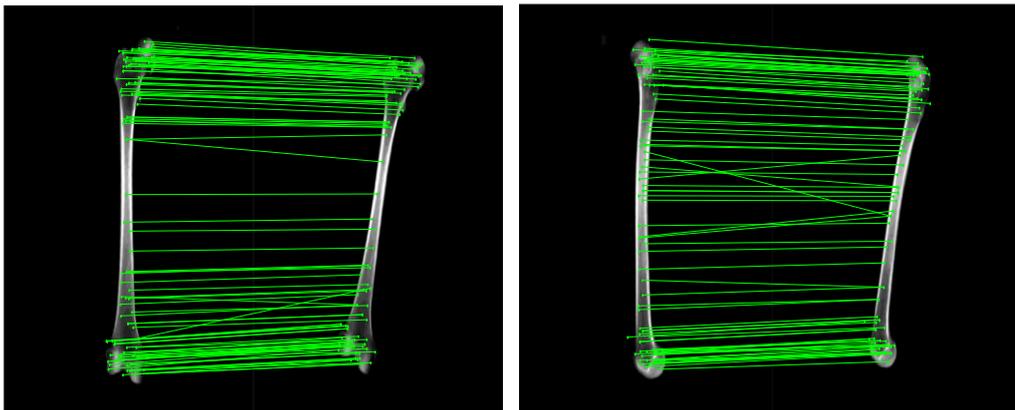
For image matching, SIFT features are first extracted from a set of reference images and then stored in a database. A new image is matched by comparing each feature individually between the new image and the previous database and finding candidate matching features on the basis of the Euclidean distance of their feature vectors.

$$dis = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (5)$$

The best candidate match for each keypoint is determined by identifying its nearest neighbor in the database of keypoints from training images. The probability that a match is correct can be determined by taking the ratio of distance from the closest neighbor to the distance of the second closest neighbor. We only need the stabilized matching feature point. Thus, we reject all matches with a distance ratio greater than 0.7, which eliminates 94% of the false matches and discards less than 4% of the correct matches.

### 3.3. Experiment on matching feature points

Two experiments were performed to meet the requirements of image registration. First, the SIFT feature detection algorithm was applied to a frontal DRR image of the universal femur 3D model and the patient frontal femur DR image. Second, feature matching was performed between a lateral DRR image of the universal femur 3D model and the patient lateral femur DR image. Experimental results are shown in Figures 3(a) and 3(b). The average correct match rate (correct matched points/total matched points) is approximately 93% and satisfactory matching results were achieved in distal femurs. Theoretically, the rotation and translation of the 2D image are required only to identify two matching points for the number of matching feature points detected by the SIFT algorithm to meet the experiment requirements. This number can also ensure the accuracy of registration.



(a) SIFT matching result of frontal image      (b) SIFT matching result of lateral image

Fig. 3. SIFT matching results of 2D femur images.

#### 4. Femur model 3D pose estimation based on ICP algorithm

Iterative Closest Point (ICP) is an algorithm employed to minimize the difference between two clouds of points. ICP is often used to reconstruct 2D or 3D surfaces from different scans [12]. Through the aforementioned preparatory work, we obtain the feature point sets  $P = \{p_i\}$  of the universal femur DRR image and feature point sets  $Q = \{q_i\}$  of the patient femur DR images.  $i = 1, 2 \dots k$  refers to the number of matching feature points. The transformation relationship of point sets  $P$  and  $Q$  can be expressed as

$$p_i = Rq_i + T \quad (6)$$

where  $R$  is the rotation matrix, and  $T$  is the translation vector. We first performed center position registration by calculating  $p_c$  and  $q_c$ , which are the centers of point sets  $P$  and  $Q$ , respectively.

$$p_c = \frac{1}{k} \sum_{i=1}^k p_i \quad (7)$$

$$q_c = \frac{1}{k} \sum_{i=1}^k q_i \quad (8)$$

We set  $\hat{p}_i = p_i - p_c$ ,  $\hat{q}_i = q_i - q_c$ . Based on the least squares criterion [13], the objective function can be expressed as

$$E(R, T) = \sum_{i=1}^n \left\| \hat{q}_i - (R\hat{p}_i + T) \right\|^2 \quad (9)$$

The core issue of 2D image registration is computing the rotation matrix and translation vector from point set  $P$  to  $Q$  and minimizing the objective function (Eq. 9). We adopted the method of the SVD matrix decomposition algorithm, which is easy to implement and produces accurate calculation results.

For the femoral 2D–3D image registration and the 3D pose estimation, the core problem is identifying the space freedom parameters for 3D pose estimation. The six degree-of-freedom registration parameters are obtained by using Euler angle representations, including rotation angle  $(\alpha, \beta, \gamma)$  and the translation vector  $(t_x, t_y, t_z)$ , along three axes. The core transformation formula is given by

$$\begin{bmatrix} X_p \\ Y_p \\ Z_p \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X'_p \\ Y'_p \\ Z'_p \\ 1 \end{bmatrix} \quad (10)$$

where  $(X'_p, Y'_p, Z'_p)$  is the 3D point coordinate of the universal femur model,  $(X_p, Y_p, Z_p)$  is the 3D point coordinate of patient femur model, and  $T_{3 \times 1} = \{t_x, t_y, t_z\}$  is the translation vector.  $R_{3 \times 3}$  is the rotation matrix and can be expressed as

$$R_{3 \times 3} = R_Z R_Y R_X = \begin{bmatrix} \cos \beta \cos \gamma & \sin \alpha \sin \beta \cos \gamma & \cos \alpha \sin \beta \cos \gamma + \sin \alpha \sin \gamma \\ \cos \beta \sin \gamma & \cos \alpha \cos \gamma + \sin \alpha \sin \beta \sin \gamma & \cos \alpha \sin \beta \sin \gamma - \sin \alpha \cos \gamma \\ -\sin \beta & \sin \alpha \cos \beta & \cos \alpha \cos \beta \end{bmatrix} \quad (11)$$

After 2D–3D image registration, we obtain the registration parameter  $(\alpha, \beta, \gamma, t_x, t_y, t_z)$ .

### 5. Experimental results and analysis

The frontal and lateral images of patient femur were captured by digital radiography. The SIFT algorithm was then used to detect the matching feature points with DRR images of the universal femur model. To calculate the registration parameters  $(\alpha, \beta, \gamma, t_x, t_y, t_z)$  on the basis of the iterative closest point algorithm, we adjusted the CT volume data of universal femur and obtained the 3D model estimation of the patient femur. Figure 4 shows the pose contrast of the femoral 3D model before and after registration. Figure 4(a) shows the original universal femur 3D model, whereas Figure 4(b) shows the 3D model of patient femur after pose estimation. After 2D–3D registration, the rotation and translation of patient femur pose estimation exhibited good recovery.

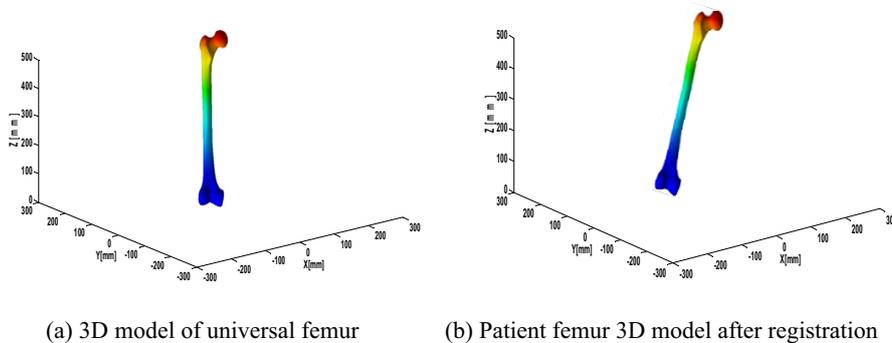


Fig. 4. 3D pose estimation result of patient femur.

Table 1  
Error estimate experiment results of Pose Estimation

	Rotation Error (°)	Translation error (mm)	Running time (s)
Experiment 1	1.54	1.21	1.25
Experiment 2	1.62	1.18	1.34
Experiment 3	1.51	1.23	1.19
Experiment 4	1.55	1.25	1.24
Experiment 5	1.58	1.15	1.35
Experiment 6	1.52	1.18	1.28
Experiment 7	1.57	1.14	1.46
Experiment 8	1.49	1.21	1.26
Experiment 9	1.51	1.24	1.38

In practical clinic applications, the parameters of the patient femur 3D pose are difficult to obtain. To verify the precision of the generated patient femur 3D pose estimation, we adopted a theoretical evaluation method by using artificial pre-set 3D pose parameters  $(\alpha, \beta, \gamma, t_x, t_y, t_z)$ . The 3D model pose of the patient femur is then adjusted on the basis of this value. Subsequently, we applied the proposed method to calculate the 2D–3D registration parameter  $(\alpha_1, \beta_1, \gamma_1, t_{x_1}, t_{y_1}, t_{z_1})$  between universal femur model and patient femur model. The error between calculated and preset parameters is used to evaluate the registration results. Table 1 shows the results of the nine experiments. The average rotational errors of the patient 3D femur pose estimation is approximately  $1.5^\circ$ , whereas the average translation error is less than 1.2 mm, and the runtime of femoral pose estimation is approximately 1.3 s.

## 6. Conclusion

This study proposed a method that uses two DR images to perform femur pose estimation on the basis of 2D–3D rigid registration, through which the preoperative visualization of the patient femur is achieved. We validated the accuracy and robustness of the proposed method through experiments. The angle rotation error of femoral model was within  $1.5^\circ$ , the translation error of femoral model was approximately 1.2 mm. The femoral pose estimation results show that the proposed method has high accuracy and good robustness. In practical clinic application, the proposed method has the advantage of lower cost, faster imaging, lower harmful radiation, and higher precision than the traditional diagnostic method of DR or CT imaging. Furthermore, the proposed method can provide effective 3D imaging information for preoperative diagnosis and surgical plans for orthopedic surgery.

## Acknowledgement

This paper is supported by the project of the State Key Laboratory of Robotics and System (SKLRS201410B) and (SKLRS201201A02).

## References

- [1] P. Kazanzides, G. Fichtinger, G.D. Hager et al., Surgical and interventional robotics-core concepts, technology, and design [Tutorial], *Robotics & Automation Magazine* **15** (2008), 122–130.
- [2] J. Enderle, S. Blanchard and J. Bronzino. Introduction to biomedical engineering, *Biomed. Optics and Lasers Contributed*, G. L. Cote, S. Racecar and L. Wang, Eds., Academic, 2000, New York, pp. 843–903.
- [3] J. Dworzak, H. Lamecker, J. von Berg et al., 3D reconstruction of the human rib cage from 2D projection images using a statistical shape model, *International Journal of Computer Assisted Radiology and Surgery* **5** (2010), 111–124.
- [4] R.C. Semelka, D.M. Armao, J. Elias et al., Imaging strategies to reduce the risk of radiation in CT studies, including selective substitution with MRI, *Journal of Magnetic Resonance Imaging* **25** (2007), 900–909.
- [5] L. Caponetti and A.M. Fanelli, Computer-aided simulation for bone surgery, *Computer Graphics and Applications* **13** (1993), 86–92.
- [6] R. Dumas, B. Blanchard, R. Carlier et al., A semi-automated method using interpolation and optimisation for the 3D reconstruction of the spine from bi-planar radiography: A precision and accuracy study, *Medical & Biological Engineering & Computing* **46** (2008), 85–92.
- [7] T. Jerbi, V. Burdin, J. Leboucher et al., 2-D-3-D frequency registration using a low-dose radiographic system for knee motion estimation, *IEEE Transactions on Biomedical Engineering* **60** (2013), 813–820.
- [8] P. Gamage, S.Q. Xie, P. Delmas et al., Intra-operative 3D pose estimation of fractured bone segments for image guided orthopedic surgery, *IEEE International Conference on Robotics and Biomimetics*, 2009, 288–293.
- [9] Z. Zhang, A flexible new technique for camera calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** (2000), 1330–1334.
- [10] G. Wang, H. Yu and B. De Man, An outlook on x-ray CT research and development, *Medical Physics* **35** (2008), 1051–1064.
- [11] P.J. Besl and N.D. McKay, Method for registration of 3-D shapes, *International Society for Optics and Photonics* **6** (1992), 586–606.
- [12] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision* **60** (2004), 91–110.
- [13] Haipeng, Liu, Jing Zhang, Dinghua Mao and Xinbo Zhao, Registration of the reconstructed CT slice model and CAD model based on the SVD-ICP algorithm, *Computer Engineering and Applications* **24** (2004), 060.