

# Real-time monitoring system with accelerator controlling: An improvement of radiotherapy monitoring based on binocular location and classification

Lei Chai<sup>a</sup>, Da Chen<sup>a</sup>, Xiao-bing Tang<sup>a,\*</sup>, Yun Ge<sup>b</sup>, Ying Chen<sup>b</sup> and Jun Li<sup>a</sup>

<sup>a</sup>*Department of Nuclear Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China*

<sup>b</sup>*School of Electronic Science and Engineering, Nanjing University, Nanjing, China*

Received 21 June 2016

Revised 3 August 2016

Accepted 17 October 2016

**Abstract.** Real-time monitoring and amendment of patient position is important for the radiotherapy. However, using electronic portal imaging device (EPID) and cone beam computer tomography (CBCT) in the clinical practice generate different degrees of delay, so that they cannot achieve the purpose of real-time application. Meanwhile, a few products come with the function of the real-time monitoring and amendment, such as CyberKnife, which is too expensive for the common people. The objective of this study is to develop and test a novel independent system to monitor treatment center and amend the position of patient, which is applicable to most accelerators, based on binocular location. The system monitors the treatment center by tracking the markers attached to the patient. Once the treatment center shifts, the system uses the magic finger, which is developed to control the treatment bed automatically to adjust the treatment bed position. To improve the monitoring accuracy, we trained the data collected from the clinic based on SVM (Support Vector Machine). Thus, the training results assist users to adjust the feasible degree of the monitoring. The experiment results showed that using this new monitoring system, the monitoring resolution reached 0.5 mm, and the error ratio of the judgment was less than 1.5%.

**Keyword:** Real-time monitoring, binocular system, magic finger, Support Vector Machine

## 1. Introduction

Positioning is very important to the radiotherapy and its accuracy significantly influences the treatment efficiency. With the frequent two and three-dimensional imaging during a course of radiation treatment, such as the electronic portal imaging device (EPID) and the cone beam computed tomography (CBCT), patient can receive a high-precision localization under guidance of the imaging coordinates of the actual radiation treatment plan. Such a process is called image-guided radiation therapy (IGRT) [1–3]. With two orthogonal pictures taken by EPID, the radiation therapist compare such EPID images with digitally reconstructed radiograph (DRR) to determine the geometric shift

---

\*Corresponding author: Xiao-bing Tang, PhD, Department of Nuclear Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China. Tel.: +86 025 84895988; Fax: +86 025 52112626; E-mail: tangxiaobin@nuaa.edu.cn.

between the planned and the actual lesion and adjust the position of patient according to the shift [4]. By using a kilo-voltage CBCT to acquire the interested anatomy of each patient which takes less time than CT [5, 6] and then matching it with DRRs from the planning CT to calculate the positioning error, the physician adjust the position of patient until he or she is properly aligned in the treatment room.

However, during a treatment, the change of patient posture is ignored commonly. And because the x-ray imaging always be time consuming, the monitoring of patient cannot be real-time. In addition to the feature of imaging delay, such imaging will expose the patients to an additional dose, which may bring an irreversible damage to them. Considering of the above problem, a few products with the function of the real-time monitoring and amendment have been available, such as CyberKnife [7, 8]. While due to the high cost of CyberKnife, most people, especially those people in the developing countries, can't afford to it and consequently can't receive an accurate patient positioning which is provided such a system.

This paper presents a novel independent system to monitor treatment center and amend the position of patient. Such system is based on a binocular system, by using two cameras to track the infrared sensitive markers which are fixed to the patient and indirectly to monitor the treatment center. Once the treatment center has a certain degree of shift, i.e. it is not aligned with the isocenter of a linear accelerator, the presented system will use an in-house controller, called as the magic finger, to move the treatment bed automatically based on the shift to correct position until the shift is within the clinical tolerance.

## 2. System composition

The system consists of three components as follows: a binocular system, a software which monitors the duration of the patient positioning in real time and a magic finger as shown in Fig. 1. The binocular system includes two cameras to take pictures of the interested region and then transfers the images to

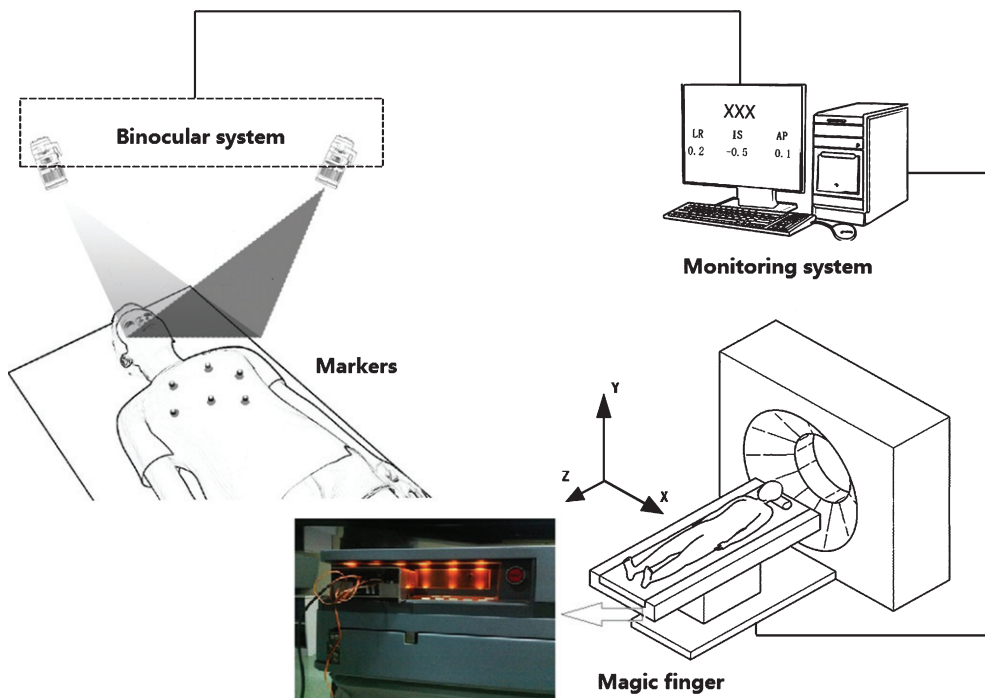


Fig. 1. System Composition. Three components are included in the system: a binocular system, a software to monitor the process of the patient positioning and a magic finger to move the treatment bed according to instruction from the software.

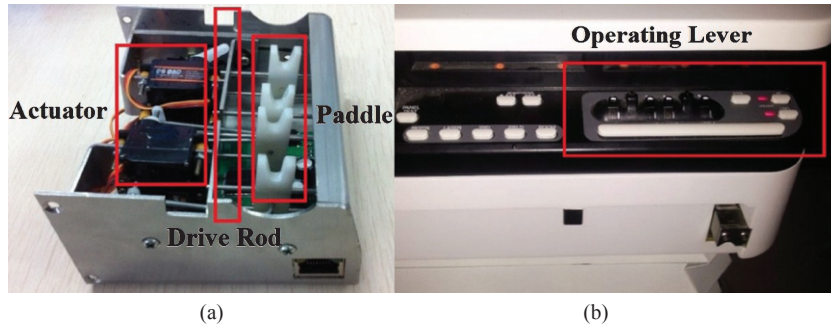


Fig. 2. Magic Finger Components. (a) An exhibition of the Magic Finger. The magic finger consists of three control units and each unit contains an actuator, a paddle and a drive rod. Each unit controls one of the three levers on the control board of the treatment bed. (b) An exhibition of the control board of the treatment bed. The three levers on the board are designed to control the movement of the treatment bed along anterior-posterior, inferior-superior and left-right directions in a three-dimensional space.

the software. The software processes the images including extracting the markers and reconstruct the markers in tri-dimensional space, and subsequently computes the geometric shift between the positions of the planned and the actual lesion by matching the acquired data in real time with the planning CT data. Note that before moving the treatment bed, the software would judge whether the calculated shift is credible. If it is credible, the treatment bed will be adjusted to the correct position, vice versa.

The space resolution and the monitoring scope were taken into account when the binocular system was designed. We chose the 1/2.8 in. gray color Complementary Metal Oxide Semiconductor (CMOS) camera (PointGrey Flea3). With the specific type of cameras, the acquisition of images was synchronized at the hardware level with the infrared light-emitting diodes array. The image size was 2048 pixels  $\times$  1536 pixels, with the pixel size of 2.5  $\mu\text{m}$   $\times$  2.5  $\mu\text{m}$ . The two cameras were mounted rigidly with a distance of 660 mm and an angle of 13°.

The Magic Finger was designed to control the movement of a treatment bed based on an advanced RISC (reduced instruction set computing) machine (ARM) processor. It contains three control units and each unit, consisting of an actuator, a paddle and a drive rod, was designed to control one of the three levers on the control board of treatment bed as shown in Fig. 2. And, the Magic Finger communicated with the software by cable.

### 3. Camera calibration and three-dimensional reconstruction

As shown in Fig. 3, the reconstruction work can be divided into two parts: the camera calibration and the real-time three-dimensional (3D) reconstruction. Before the system works, we should complete the calibration of the two cameras to estimate their intrinsic and extrinsic parameters. In the real-time

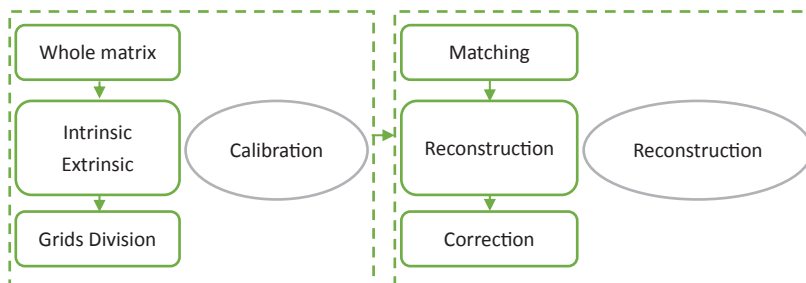


Fig. 3. Work flow of a three-dimensional reconstruction. The reconstruction contains two steps: the camera calibration and the real-time reconstruction.

3D reconstruction part, firstly, we capture the synchronous images from the two cameras. Then after preprocessing the images, we extract the centers of the markers and subsequently match the centers between left and right images. In the end, we complete the 3D reconstruction and the correction for the result.

### 3.1. Camera calibration

#### 3.1.1. Camera model

The camera is assumed to be as a standard perspective camera, so that the transformation between the world coordinates and the image coordinates can be seen in Equation (1) [9, 10]:

$$s \cdot m = A \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} M \quad (1)$$

where  $s$  is the scale factor.  $m = [u \ v \ 1]^T$  represents a two-dimensional (2D) point position in the image coordinates, while  $M = [x \ y \ z \ 1]^T$  is a 3D point position in the world coordinates. The  $3 \times 4$  matrix  $A$  contains five intrinsic parameters. And the extrinsic parameters are described by a rotation matrix  $R$  and a translation vector  $t$ , which denote the transformations from the 3D world coordinates to the 3D image coordinates. The intrinsic matrix  $A$  can be written as shown in Equation (2) [9, 10]:

$$A = \begin{bmatrix} \alpha_x & \gamma & \mu_0 & 0 \\ 0 & \alpha_y & \nu_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2)$$

in which  $\alpha_x = f/d_x$  and  $\alpha_y = f/d_y$ , where  $d_x$  and  $d_y$  are the scale factor relating pixels to distance and  $f$  is the focal length in terms of length.  $\mu_0$  and  $\nu_0$  represent the principal point which is the intersection point of the optical axis and the imaging plane.  $\gamma$  is the obliquity between the  $u$  and  $v$  image axes and it is zero ideally.

Although cameras are calibrated based on the camera model as shown in Equation (1), we cannot complete the reconstruction correctly. In real cameras, there are radial distortion and tangential distortion. In this paper, we choose a non-analytical solution to solve these distortions.

#### 3.1.2. Calibration method

Based on the proposed camera model, the intrinsic and extrinsic parameters of the cameras can be obtained easily. The calibration method is similar to the method proposed by Tsai [9], especially in estimating the intrinsic parameters. The sequence of steps can be summarized as followed:

1. We take pictures of  $17 \times 17 \times 17$  points of  $1 \text{ m} \times 1 \text{ m} \times 1.5 \text{ m}$  space, divided by 17 planes.
2. With extracting points from pictures and getting the image coordinates of the points, we can get the world coordinates of the corresponding points based on the 3D Coordinate Measuring Machine (3D CMM).
3. Compute the whole matrix encompassing intrinsic and extrinsic parameters, i.e.  $H = A \begin{bmatrix} R & t \end{bmatrix}$ .
4. Compute the extrinsic and intrinsic parameters using the method proposed by Tsai [9].
5. Divide the space into  $4 \times 4 \times 4$  grids as shown in Fig. 4, reconstruct the 3D coordinates of the eight corner points of each divided grid using the whole matrix  $H$  given by step 3, and then store them.
6. Finally divide the four grids on the corner into  $2 \times 2 \times 2$  grids further as shown in Fig. 4, reconstruct them and store them as well.

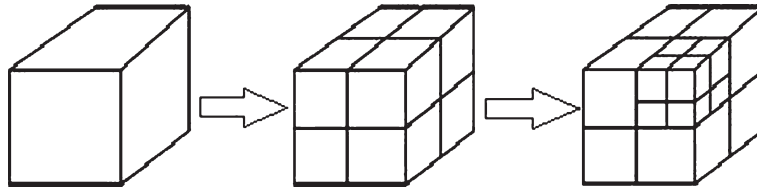


Fig. 4. Sketch map of the grids division. The three-dimensional space is divided into several grids of uneven size. Firstly, the space was divided into  $4 \times 4 \times 4$  grids. Then, four corners are divided into smaller grids for correction the distortion of cameras.

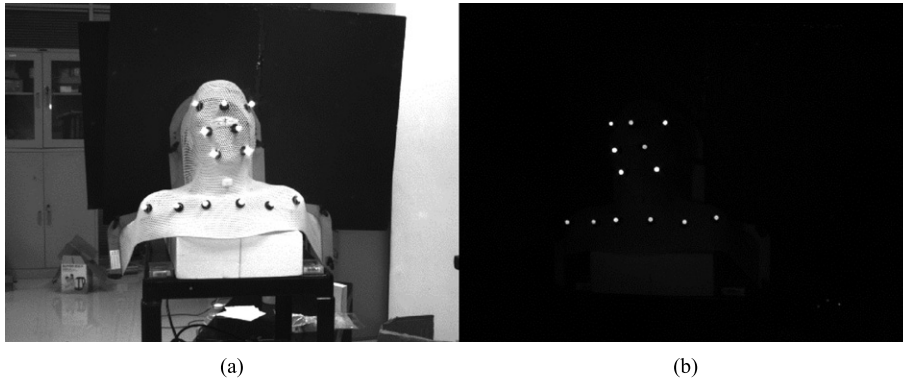


Fig. 5. Camera Images. (a) Camera image without an infrared cut-off filter. (b) Camera image after an infrared cut-off filter works on the light which is received by the cameras.

### 3.2. 3D reconstruction

By using the infrared light-emitting diodes array and the infrared filter in the front of camera, we can get the images where only the infrared wavelength is passed while the visible light is cutoff. Because the markers are infrared sensitive, after the infrared filter works on the light which is received by cameras, as shown in Fig. 5(b), only the markers are relatively visible in the images. By processing on such images with the open source computer vision (OpenCV), a library of programming functions for real-time computer vision, we can easily calculate the positions of the marker centers in the images.

#### 3.2.1. Matching method

Assuming that there are many points in the real scene, denoted as  $n$  balls, then we will get  $n$  points in a pair of images from the left and right eyes of the cameras. Without matching the points from the left and the right images, we have to compute  $n \times n$  times to reconstruct the 3D scene. Therefore, it is very important to match the points from the left and right images. However there is a constraint in binocular imaging called epipolar geometry [11–13]. As shown in Fig. 6, there is a base line, defined as a line connecting the two camera centers, in the binocular system. The epipolar plane is defined as the plane containing the base line and an arbitrary point in the space. An epipolar point is defined as the intersection point of the imaging plane and the line which connects the camera centers. And an epipolar line is defined as the intersection line of the epipolar plane and the imaging plane, while the intersection point of the all epipolar lines is the epipolar point. Ideally, the imaging point of an arbitrary point in the space should be always in the epipolar line which is the intersection line of the epipolar plane and the imaging plane as shown in Fig. 6 [11, 14]. This relationship is evolved to Equation (3) [11]:

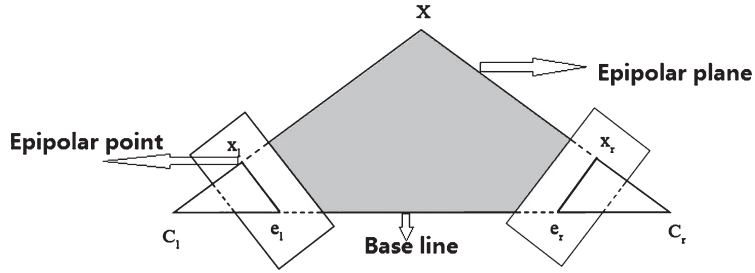


Fig. 6. Epipolar geometry. Any point in the view space of the binocular system will always encompass a epipolar plane with the camera centers \$C\_1\$ and \$C\_2\$. The line connecting \$C\_1\$ and \$C\_2\$ is called a base line. And the intersection points \$x\_1\$ and \$x\_2\$ both are the epipolar points.

$$x^t T F x = 0 \tag{3}$$

where \$x^t T\$ and \$x\$ are the image coordinates of the left and right cameras respectively. And \$F\$ is the base matrix which is defined as:

$$F = k^{t-T} R \left[ R^T t \right] \times K^{-1} \tag{4}$$

where \$K^{t-T}\$ and \$K\$ are the intrinsic parameters matrices of the left and right cameras respectively, \$R\$ is the rotation matrix of the right camera, \$t\$ is the translation matrix of the right camera, \$[\ ]\_{\times}\$ is the operation of a antisymmetric matrix. \$F\$ is defined as above on the premise that the world coordinate system is defined by the left camera coordinate system. In Equation (3), \$x^t T F\$ is the epipolar line in the right camera, so that \$x^t T F = 0\$ means the imaging point is in its epipolar line. However, in the practical situation, Equation (3) will never be true. So instead, we use the distance between the imaging point and its epipolar line to judge whether the two points are matched.

### 3.2.2. Three-dimensional reconstruction and correction

As the matching is done, we can use the whole matrix \$H\$ which is computed in the process of the calibration and use the Least Squares (LS) method to complete the preliminary reconstruction. As we get the 3D coordinates from the preliminary reconstruction, we can locate the points in the grids that we have divided. In order to eliminate the effects of the camera distortion, we develop a near-linear correction method. As shown in Fig. 7, the transformation from the grid \$A\$ to \$A'\$ is a transformation with distortion rather than a rigorous affine transformation \$H\$, so an arbitrary point in \$A\$ cannot be transformed by \$H\$. But, if the grid is infinitely small, it can be assumed that any point in the grid can be transformed by \$H\$. Based on the above analysis, if the grid is divided to be enough small, the transformation of any point inside the grid is linearly varying with the location within a narrow range. As a result, we can compute the transformation by the tri-linear interpolation. The computation formula can be written as Equation (5):

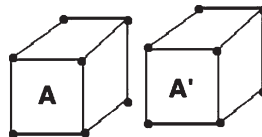


Fig. 7. Correction by transformation. \$A\$ is the actual position in the space, and the \$A'\$ is the calculation position effected by the camera distortion.

$$H = \sum_{i=1}^6 \left( \frac{1}{d_i} \right) / \left( \sum_{i=1}^6 \frac{1}{d_i} \right) \times H_i \quad (5)$$

where,  $d_i$  is the distance between the  $i$ th point and the reference plane,  $H_i$  represents the transformation matrix from  $i$ th plane of  $A$  to the corresponding plane of  $A$ . If  $d_i = 0$ , the transformation of the point is the transformation of the plane where the point is inside it.

## 4. Tracking and accelerator control

### 4.1. Tracking

After the reconstruction, the next step is to realize the real-time monitoring. Note that prior to CT scan, several markers are attached to the thermoplastic mask. In order to monitor the course of the patient positioning, we matched the markers in CT data to those acquired from the monitoring system. And then we perform some matrix operations to calculate the real-time rotation and translation matrices which represent the translation from the CT coordinates to the optical coordinates, which is shown in Equation (6) and Equation (7):

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = P_{cp} \begin{bmatrix} x_{ct} \\ y_{ct} \\ z_{ct} \\ 1 \end{bmatrix} \quad (6)$$

$$P_{cp} = \begin{bmatrix} \cos \varphi_{cp} \cos \theta_{cp} & \sin \varphi_{cp} \cos \theta_{cp} & -\sin \theta_{cp} & T_x \\ \cos \varphi_{cp} \sin \theta_{cp} \sin \gamma_{cp} - \sin \varphi_{cp} \cos \gamma_{cp} & \sin \varphi_{cp} \sin \theta_{cp} \sin \gamma_{cp} + \cos \varphi_{cp} \cos \gamma_{cp} & \cos \theta_{cp} \sin \gamma_{cp} & T_y \\ \cos \varphi_{cp} \sin \theta_{cp} \cos \gamma_{cp} + \sin \varphi_{cp} \sin \gamma_{cp} & \sin \varphi_{cp} \sin \theta_{cp} \cos \gamma_{cp} - \cos \varphi_{cp} \sin \gamma_{cp} & \cos \theta_{cp} \cos \gamma_{cp} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7)$$

where  $x_{ct}$ ,  $y_{ct}$  and  $z_{ct}$  are the 3-D coordinates of any point in the CT coordinate system while  $x$ ,  $y$  and  $z$  represent the 3-D coordinates of any point in the optical coordinate system. The transformation is based on the  $4 \times 4$  matrix  $P_{cp}$ . Then,  $\gamma_{cp}$ ,  $\varphi_{cp}$  and  $\theta_{cp}$  are the rotation angles of the roll, pitch, and yaw respectively.  $T_x$ ,  $T_y$  and  $T_z$  are the translations in left-right (LR), inferior-superior (IS) and anterior-posterior (AP) directions respectively. After taking in account the treatment center coordinates in Equation 6, we can obtain the real-time treatment center in monitoring coordinate system, and the deviation to the isocenter in LR, IS, and AP directions consequently which are the geometric shift.

### 4.2. Accelerator control

Once the transformation from the CT coordinates to the optical coordinates is calculated, the treatment center bias, defined as the deviation between the linear accelerator isocenter and the treatment center, can be figured out approximately. In this paper, we judge whether the translation is credible by using support vector machine (SVM). If the translation is credible, we will move the treatment bed according to the above treatment center bias.

#### 4.2.1. SVM classification

In order to judge whether the treatment center bias is credible, we choose the SVM, a supervise learning model to classify the credible value and the incredible one. For the two training classes that

are linearly separable in the selected feature space, the rationale of SVM is to find a hyperplane to divide the two classes. The hyperplane is defined as [15]:

$$f(x) = w^T * x + b = 0 \quad (8)$$

There may be an unlimited number of such planes that can divide the feature space into two classes, and SVM aims to find the best plane which maximizes the distance from it to the nearest data point on each side. The object is evolved to [15, 16]:

$$\begin{aligned} & \min_{w,b} \|w\|^2/2 \\ \text{s.t.} \quad & \begin{cases} - (w^T x_i^{(1)} + b) \geq 1, & i = 1, 2, \dots, m \\ + (w^T x_i^{(1)} + b) \geq 1, & i = 1, 2, \dots, m \end{cases} \end{aligned} \quad (9)$$

However, in the actual situation, there is no hyperplane can divide the two classes completely. To deal with this problem, the SVM introduced a mapping function and a penalty function. And then the hyperplane and the object are evolved to [17]:

$$\begin{aligned} & f(x) = w^T * \varnothing(x) + b = 0 \\ & \min_{w,b} \|w\|^2/2 + C \sum_{i=1}^m \xi_i \\ \text{s.t.} \quad & \begin{cases} - (w^T x_i^{(1)} + b) \geq 1 - \xi_i, & i = 1, 2, \dots, m \\ + (w^T x_i^{(1)} + b) \geq 1 - \xi_i, & i = 1, 2, \dots, m \\ \xi_i \geq 0, & i = 1, 2, \dots, m \end{cases} \end{aligned} \quad (10)$$

where  $w$  is the normal vector to the hyperplane,  $\varnothing()$  is the mapping function called as kernel function, and the  $\xi_i$  is the tolerance error allowed for the sample point being on the wrong side of the hyperplane. There are five common types of kernel functions, and in this paper, we only choose three types of kernel functions, i.e. the linear kernel, polynomial kernel and radial basis function (RBF) kernel, to enroll into the investigation. The three function can be expressed as [17–20]:

$$\begin{aligned} \text{Linear: } & K(x, x') = x^T * x' \\ \text{Polynomial: } & K(x, x') = (\gamma x^T * x' + c)^q \\ \text{RBF: } & K(x, x') = \exp(-\|x - x'\|^2/2\sigma^2) \end{aligned} \quad (11)$$

in which  $c$  is a constant value.  $q$  is the order of the polynomial and it is set to 3 in this paper.  $\gamma$  is the gamma coefficient. And  $\sigma$  is the width of RBF.

As the prototype function of the hyperplane and the kernel function is fixed [21], the supervised learning is to determine the optimal parameters of such a model. Once the parameters are all finalized, the judgement depends on the value of  $w^T * \varnothing(x) + b$ .

#### 4.2.2. Magic finger

Since we have judged the computation credibility, it is known that how much we should shift the treatment bed. However, all the commercially available accelerators don't provide the signal interface to control the movement of the treatment bed to the users. As a result, we can't control the bed movement directly by transmitting the control signal to the accelerator. The only solution to deal with this problem is to manipulate the control panel which is offered by the accelerator to control the treatment bed. Therefore, we design a novel device called Magic Finger to achieve the goal of



controlling the treatment bed movement automatically. As shown in Fig. 2, Magic Finger has three paddles to control the three levers on the panel respectively, and each paddle was driven by an actuator which is connected by a drive rod.

## 5. Experimental results and discussion

### 5.1. Reconstruction results

The accuracy of the reconstruction is analyzed by performing a set of experiments which consists of tracking a single marker in different positions with three dimensional measuring machine which has the accuracy of 0.05 mm. The different positions are all evenly distributed in the monitoring space with a height of 1.0 m, a width of 1.0 m and a depth of 1.5 m. As shown in Fig. 8, before the correction which will correct the reconstruction coordinates, the average error of the reconstruction using the whole matrix directly is 0.79 mm, and the maximal error is 2.3 mm. However, after the correction, the average error is corrected down to 0.24 mm and the maximal error is controlled within 1.0 mm.

We sorted the experiment points according to the z value and the study shows that the reconstruction accuracy is affected at various degrees by the distortions of cameras at different displacements. As shown in Fig. 9, before correction, the reconstruction error increases with the increase of z value, and the bigger the distance between the marker and the center axis is, the greater the reconstruction error is. And, after correction, the influence of camera distortion was eliminated as shown in Fig. 9.

### 5.2. Classification results

In this study, 64 sets of clinical data of 6 patients from Najing Drum Tower Hospital are collected, including: the CT images, the synchronized coordinates of the six markers which are attached to the thermoplastic mask of patients from the proposed system and the CT respectively, the treatment bias detected by a cone beam computed tomography (CBCT) and the corresponding calculated bias from the proposed system at the same time. With the study, we find that the credibility of the calculated

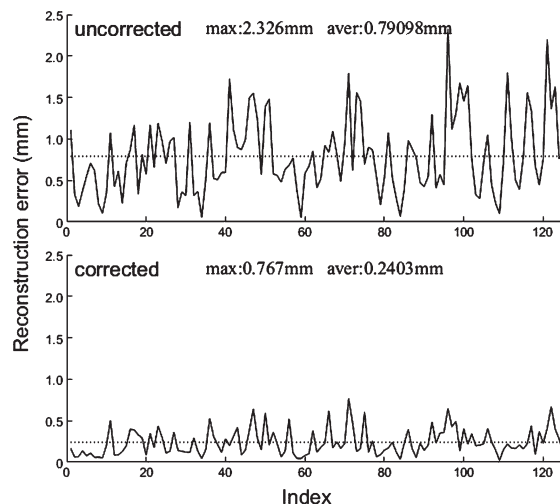


Fig. 8. The reconstruction errors before correction and after correction. Before the correction, the average error is 0.79 mm, and the max error reached up to 2.33 mm. After the correction, the average error decreased to 0.24 mm, and the max error was controlled within 1.0 mm.

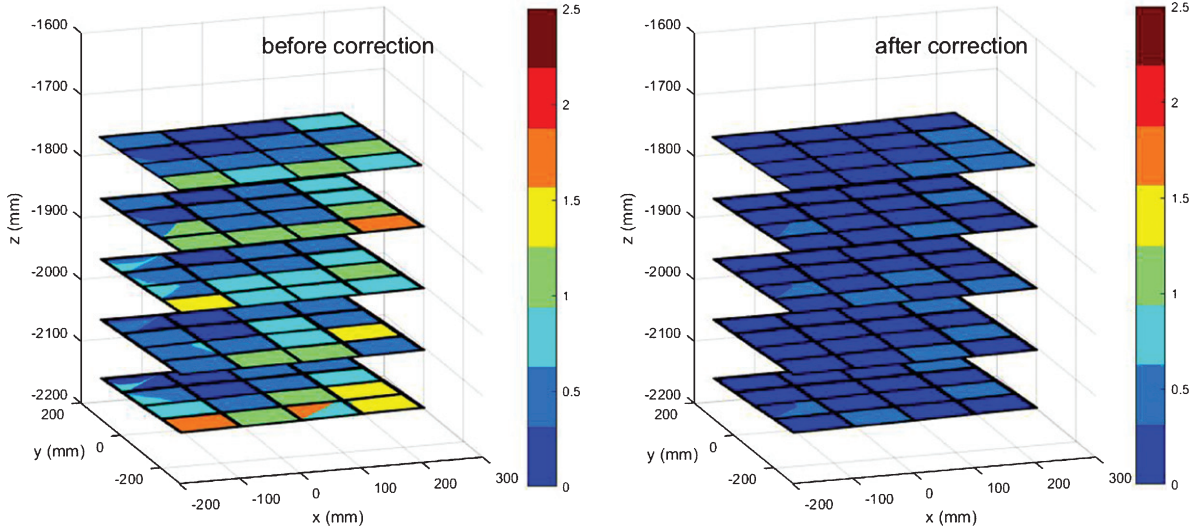


Fig. 9. The distribution of the reconstruction errors before correction and after correction. Before the correction, the influence of the distortion focused on the margin of the monitoring space. And with the distance between the monitoring point and the binocular system increasing, the influence becomes more and more serious. After the correction, the influence of distortion was almost eliminated. The colorbar in the figure represents the deviation of the 3D spatial positions between the reconstructed point and its actual value. The cold color shows the small deviation, while the warm color shows the large deviation.

results is concerned with the deformation of the mask which can be reflected by the markers attached to the mask. The classification input vector is defined as follows:

$$X = (\text{MinD}, \text{MaxD}, \text{MidD}, \text{AverD}, \text{VarD}) \quad (12)$$

| No | mark  | remark                |
|----|-------|-----------------------|
| 1  | MinD  | Minimum distance      |
| 2  | MaxD  | Maximum distance      |
| 3  | MidD  | Median of distances   |
| 4  | AverD | Average of distances  |
| 5  | VarD  | Variance of distances |

where the distance of each marker is defined as:

$$\text{Distance} = \|H^{-1} * P_w - P_{ct}\|^2 \quad (13)$$

in which the  $H^{-1}$  is the reverse transformation matrix from the optical coordinate system (CS) to the CT coordinate system.  $P_w$  is the 3D coordinates in the optical CS of an arbitrary marker and  $P_{ct}$  is the 3D coordinates in CT CS of the corresponding marker.

By binary processing on the Euclidian distance between the translation bias detected by CBCT and those detected by the proposed system with a threshold value of 1.0 mm, we classified the detected translation bias from the proposed system into two labels: a credible one (i.e., Euclidian distance  $\leq 1.0$  mm) and an incredible one (i.e., Euclidian distance  $> 1.0$  mm).

To validate the classification method, we selected a quarter of the clinical data randomly as the validation set and the remainder as the training set. For the three different kernel functions, we used

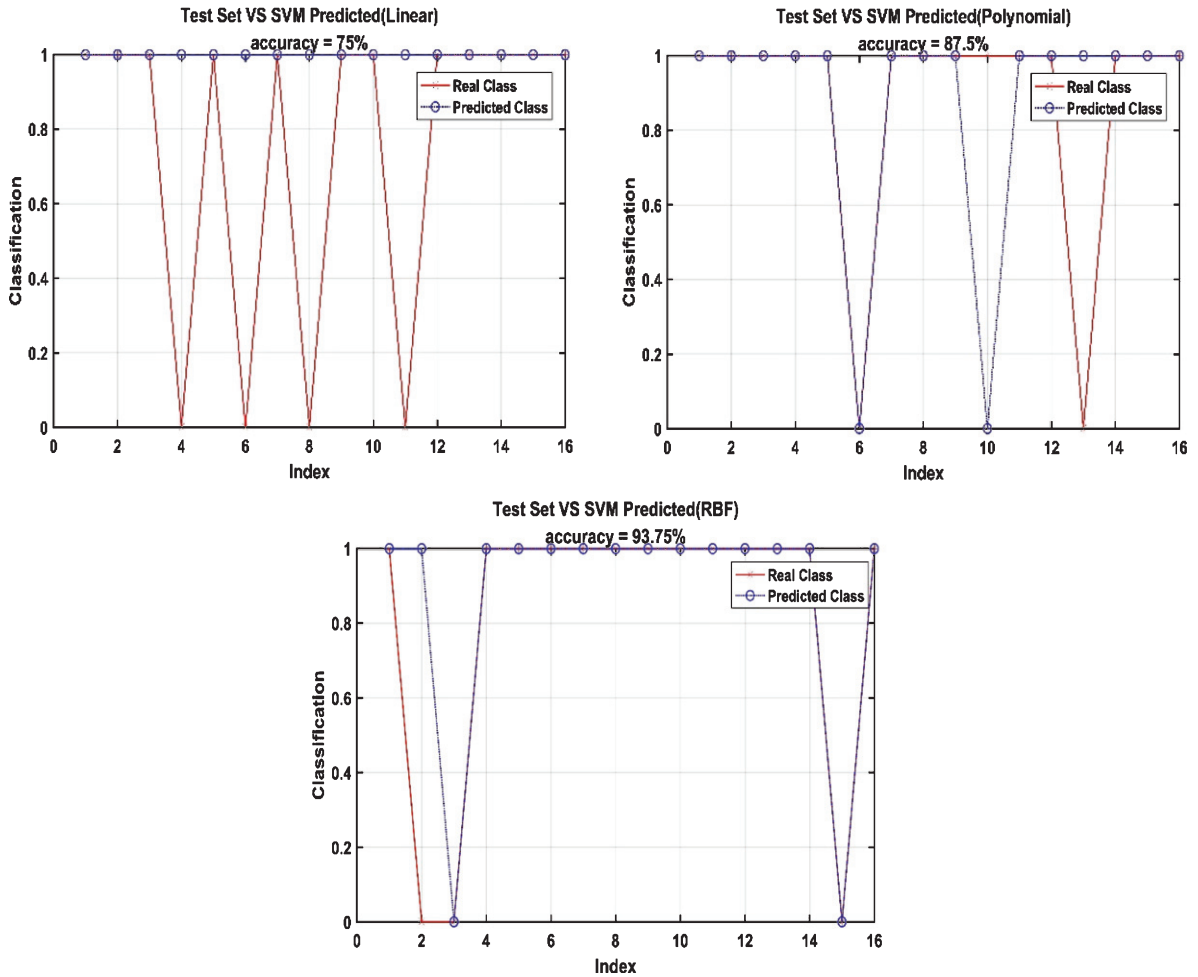


Fig. 10. The accuracy of the SVM prediction with different kernel functions. The three kernel functions perform different accuracy. The two nonlinear kernel functions result in higher accuracy: 87.5% for polynomial kernel and 93.75% for RBF kernel.

a cross-validation method to choose the best parameter  $c$ ,  $g$  and  $r$  for each function, where  $c$  is the factor of the penalty function in Equation (10),  $g$  is the gamma coefficient in polynomial kernel, and is the width of RBF in RBF kernel,  $r$  is the constant coefficient in polynomial kernel. Such an ergodic process takes several cycles. The results can be seen in Fig. 10, the classification accuracy of the linear, polynomial and RBF kernel is 75%, 87.5% and 93.75% respectively. Based on it, we can infer that the RBF kernel function performs better than other two types of kernel functions in clinics.

## 6. Summary

This paper presents a novel system to monitor the patient positioning and to correct the patient posture in real-time. Monitoring patient positioning is based on an in-house binocular system and we propose a novel convenient calibration method to eliminate the camera distortion of such binocular system easily. To reduce the computational load of the data processing in real time, we introduced the epipolar geometry to match the characteristic point between the left and the right images from the two cameras of the binocular system. And to reduce the probability of wrong operation of the in-house

magic finger, we classify the monitoring results into two categories: the credible and the incredible. To achieve the monitoring in a higher accuracy, in our ongoing research, we will use the higher-resolution camera lens and CMOS. And we will reduce the response time of the magic finger in the further research for the synchronous correction of the patient posture.

## References

- [1] L.B. Liu, J.X. Wu, Y. Qu, et al., Reserch and compare the application of the electronic portal imaging device and cone beam CT in IGRT for the nasopharyngeal carcinoma patients, *Chinese Journal of Medical Physics* **6** (2012) 005.
- [2] B.M. Barney, R.J. Lee, D. Handrahan, et al., Image-guided radiotherapy (IGRT) for prostate cancer comparing kV imaging of fiducial markers with cone beam computed tomography (CBCT), *Int J Radiat Oncol Biol Phys* **80**(1) (2011), 301–305.
- [3] H. Zhou, J. Shen, B. Li, et al., Analysis of precision in tumor tracking based on optical positioning system during radiotherapy, *J Xray Sci Technol* **24**(3) (2016), 443–455.
- [4] B.J.M. Heijmen, K.L. Pasma, M. Kroonwijk, et al., Portal dose measurement in radiotherapy using an electronic portal imaging device (EPID), *Phys Med Biol* **40**(11) (1995), 1943–1955.
- [5] R. Topolnjak, J.J. Sonke, J. Nijkamp, et al., Breast patient setup error assessment: Comparison of electronic portal image devices and cone-beam computed tomography matching results, *Int J Radiat Oncol Biol Phys* **78**(4) (2010), 1235–1243.
- [6] H. Guan and H. Dong, Dose calculation accuracy using cone-beam CT (CBCT) for pelvic adaptive radiotherapy, *Phys Med Biol* **54**(20) (2009), 6239–6250.
- [7] J.R. Adler, S.D. Chang, M. Murphy, et al., The Cyberknife: A frameless robotic system for radiosurgery, *Stereotactic and Functional Neurosurgery* **69**(1) (1998), 124–128.
- [8] S.D. Chang, W. Main, D.P. Martin, et al., An analysis of the accuracy of the CyberKnife: A robotic frameless stereotactic radiosurgical system, *Neurosurgery* **52**(1) (2003), 140–147.
- [9] R. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, *IEEE Journal on Robotics and Automation* **3**(4) (1987), 323–344.
- [10] Z. Zhang, A flexible new technique for camera calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(11) (2000), 1330–1334.
- [11] Z. Guangjun, Machine vision. 2005, Beijing: Science Press.
- [12] M. Bjorkman and J.O. Eklundh, A real-time system for epipolar geometry and ego-motion estimation, in Computer Vision and Pattern Recognition, 2000. *Proceedings. IEEE Conference on. 2000. IEEE*.
- [13] S. Negahdaripour, Epipolar geometry of opti-acoustic stereo imaging, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**(10) (2007), 1776–1788.
- [14] M. Hansard and R. Horaud, Cyclopean geometry of binocular vision. *JOSA A* **25**(9) (2008), 2357–2369.
- [15] X. Zhao, E.K. Wong, Y. Wang, et al., A support vector machine (SVM) for predicting preferred treatment position in radiotherapy of patients with breast cancer, *Med Phys* **37**(10) (2010), 5341–5350.
- [16] A. Pella, R. Cambria, M. Riboldi, et al., Use of machine learning methods for prediction of acute toxicity in organs at risk following prostate radiotherapy, *Med Phys* **38**(6) (2011), 2859–2867.
- [17] M. Hussain, S.K. Wajid, A. Elzaart and M. Berbar, A comparison of SVM kernel functions for breast cancer detection. in Computer Graphics, Imaging and Visualization (CGIV), *2011 Eighth International Conference on. 2011. IEEE*.
- [18] X. Zhao, D.Kong, G. Jozsef, et al., Automated beam placement for breast radiotherapy using a support vector machine based algorithm, *Med Phys* **39**(5) (2012), 2536–2543.
- [19] T. Lin, R. Li, X. Tang, et al., Markerless gating for lung cancer radiotherapy based on machine learning techniques. *Phys Med Biol* **54**(6) (2009), 1555–1563.
- [20] U.K. Bhowmik and R.P. Adhami, Correlation and support vector machine based motion artifacts mitigation in 3D computer tomography, *J Xray Sci Technol* **20**(2) (2012), 141–160.
- [21] Y. Cui, J.G. Dy, B. Alexander and S.B. Jiang, Fluoroscopic gating without implanted fiducial markers for lung cancer radiotherapy based on support vector machines, *Phys Med Biol* **53**(16) (2008), N315–N327.