SYSTEMS-LEVEL QUALITY IMPROVEMENT

An Accurate Recognition of Infrared Retro-Reflective Markers in Surgical Navigation

Han Wu [1] · Qinyong Lin [1] · Rongqian Yang [1,2,3] · Yuan Zhou [1] · Lingxiang Zheng [1] · Yueshan Huang [1] · Zhigang Wang [4] · Yonghua Lao [1] · Jinhua Huang [5]

Abstract

Marker-based optical tracking systems (OTS) are widely used in clinical image-guided therapy. However, the emergence of ghost markers, which is caused by the mistaken recognition of markers and the incorrect correspondences between marker projections, may lead to tracking failures for these systems. Therefore, this paper proposes a strategy to prevent the emergence of ghost markers by identifying markers based on the features of their projections, finding the correspondences between marker projections based on the geometric information provided by markers, and fast-tracking markers in a 2D image between frames based on the sizes of their projections. Apart from validating its high robustness, the experimental results show that the proposed strategy can accurately recognize markers, correctly identify their correspondences, and meet the requirements of real-time tracking.

Keywords Optical tracking system · Marker recognition · Accurate stereo-matching · Ghost-markers elimination · Fast-tracking

Introduction

Optical tracking system (OTS) has been widely used in many surgical operations [1–8], such as Parkinson's disease therapy and neurosurgery, owing to its capability of tracking patient and instruments with high location accuracy by affixing retro-reflective markers on them [9–15]. The retro-reflective markers can be captured and

reconstructed their three-dimensional (3D) coordinates by the OTS using the triangulation method when the coordinates and correspondences of projections of the retro-reflective markers in two camera image planes and the calibration parameters of the OTS are knowing. However, there are the problems of recognizing marker projections and finding the correspondence of marker projections on the two cameras of the OTS in some cases. These problems will lead to the appearance of ghost markers.

Despite the prevalent usage of professional marker-based OTSs, such as Polaris (Northern Digital Inc., Canada) and ExacTrac (BrainLAB, Feldkirchen, Germany) [16], ghost markers are appeared in some cases when these systems are used to tracked the retro-reflective markers [16–20]. Yan et al. [16] noticed that the ability of a commercial OTS detecting markers varied with the number of markers and marker configuration. Dieterich et al. [18] found that ghost markers are frequently appeared when markers are close to each other. According to a study by Soete et al. [17], when evaluating the correlations between external markers and internal tumor by the same group, only 3–5 markers can be used due to the presence of ghost markers, even though it is advantageous to use more markers. Soete et al. [19] and Linthout et al. [20] reveal that ghost markers have a great impact on the operation procedures in their studies. The emergence of ghost markers may be attributed to high-intensity reflections caused by

---

Han Wu and Qinyong Lin Authors contributed equally to this work.

This article is part of the Topical Collection on *Systems-Level Quality Improvement*

✉ Rongqian Yang
   bmeyrq@gmail.com

1 Department of Biomedical Engineering, South China University of Technology, Guangzhou, Guangdong, China

2 School of Medicine, Yale University, New Haven, CT 06520, USA

3 Guangdong Engineering Technology Research Center for Translational Medicine of Mental Disorders, Guangzhou, Guangdong, China

4 Guangzhou Aimooe Technology Co., Ltd., Guangzhou, Guangdong, China

5 Sun Yat-Sen University Cancer Center, Guangzhou, Guangdong, China

surrounding objects with a glossy surface or some markers which are coplanar with two optical centers of OTS. Owing to the near infrared (NIR) filter and the reflective coating in the surfaces of the markers, the visible light is prevented by the NIR filter and the near infrared light reflected by the markers can pass through the NIR filter. Thus, the images captured by the cameras are presented as tiny bright blobs with dark backgrounds. Nevertheless, the professional marker-based OTSs are unable to identify whether tiny bright blobs are the projections of markers or high-intensity reflections caused by surrounding objects, such as surfaces of surgical instruments and electronic devices, thereby leading to misrecognition and the presence of ghost markers. Besides, incorrect correspondences of marker projections in two images may emerge when two or more markers are coplanar with two optical centers of OTS, which is called false stereo-match. Some methods have been developed to address the stereo-match problem [21–24]. An epipolar constraint (EC) is generally used to search for stereo correspondence [21, 22]. However, it is unable to determine the correct correspondences when two or more markers are coplanar with two optical centers, because marker projections have similar white areas with dark backgrounds. Other constraints, such as the geometrical distance constraint between points [24] and epipolar combined with the ordering constraint (EOC) [23], are used to find the correct correspondences. However, geometrical distance constraints are not always consistent because stray markers are independent and can move in a 3D space, while EOC is unable to confirm whether the correspondences are correct because the reflective markers are stray distributed in a 3D space. A stereo matching method has been proposed in our previous work [25], but it is too complex and time consuming to be applied for the OTS.

To make the OTS work well in tracking patient and instruments, the aforementioned problems should be solved. The misrecognition and incorrect correspondences are primarily attributed to the low robustness of the marker recognition and stereo-matching methods. Therefore, this paper proposes a novel strategy to eliminating ghost markers. The first part of the proposed strategy is marker recognition, which is designed based on the features of marker projections. Meanwhile, the correspondences are found based on the geometric information of markers in the second part of the proposed strategy. The third part of the proposed strategy is able to quickly recognize the markers and reduce the processing workload by introducing region of interest (ROI) for the next frame according to the positions of the markers in the current frame that the markers in the next frame can be recognized in the corresponding region.

This paper is structured as follows. Section 2 presents the system setup and the proposed strategy in detail. Section 3 describes the implementation process and evaluation experiment. Section 4 presents the discussion and conclusion.

Method

A marker-based OTS is constructed in this study, as shown in Fig. 1. This system comprises two cameras with near infrared (NIR) filters and two NIR light sources. Spherical retro-reflective markers covered by reflective materials are fixed on a customized pedestal with some markers adhered to a human phantom.

Marker recognition

As shown in Fig. 2(a), the application environment of the OTS is complex that the marker recognition process is more prone to error. Most of the existing marker-based OTSs use two cameras with NIR filters to prevent the interruption of environment light, as shown in Fig. 2(b). However, high-intensity reflections surrounding the marker projections are observed in some images. Moreover, the surface of surgical instruments may reflect IR light during the operation, thereby greatly influencing the marker recognition in the process of tracking patient and surgical instruments. Therefore, high-intensity reflections caused by surrounding objects must be eliminated in the stage of marker recognition.

The features of marker projections show that the edges of these projections are similar to a circle and the region inside this edge is high intensity. Hough transform is considered a classical method for circle detection [26, 27], but it is time consuming and sometimes inaccurately identifying the markers, especially when dealing with images with complex background containing high-intensity objects [28]. To address this issue, an algorithm is proposed to automatically detect the markers in the original images.
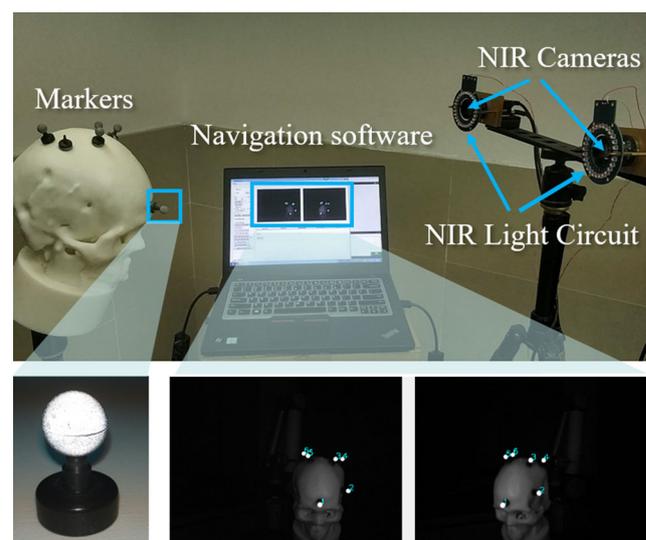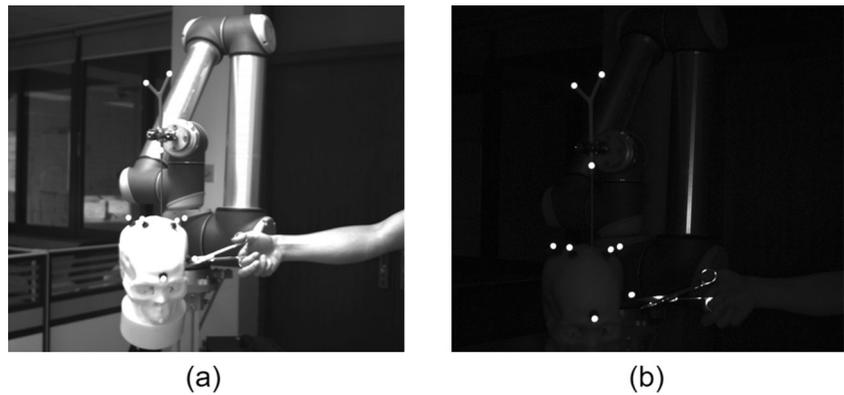

Fig. 1 Setup of the optical tracking system and the head phantom

Fig. 2 Comparison of images from cameras with and without an IR filter. (a) An image without an IR filter; (b) an image with an IR filter

(a)   (b)

Rectification of distortion

Every camera suffers a number of non-linear properties introduced by the camera lens, which result in distortion in the original images captured by the camera. The distortion is consisted of radial and tangential distortions that should be rectified. Using pinhole camera model, the imaging process of a camera can be described by the internal and external parameters, including focal length f_c, principal point c_c, skew coefficient α_c, and distortion coefficient k_c [29]. Four coordinate systems involve in the pinhole camera model, namely, the world coordinate system S_w, the camera coordinate system S_c, the image coordinate system S_i, and the pixel coordinate system S_p. Set $v_n(x_n, y_n)$ as the normalized projection in S_i, and let $u_n^2 = x_n^2 + y_n^2$. By including lens distortion, the new normalized coordinate v_d of points can be defined as

$$v_d = \left[1 + k_{c1}u_n^2 + k_{c2}u_n^4 + k_{c5}u_n^6\right]v_n + v_t, \tag{1}$$

where $v_n(x_n, y_n)$ is the normalized projection in S_i and $u_n^2 = x_n^2 + y_n^2$. Meanwhile, the vector of tangential distortion, v_t, is calculated as

$$v_t = \begin{bmatrix} 2k_{c3}x_ny_n + k_{c4}\left(u_n^2 + 2x_n^2\right) \\ k_{c3}\left(u_n^2 + 2y_n^2\right) + 2k_{c4}x_ny_n \end{bmatrix}. \tag{2}$$

After applying distortion, the pixel coordinate $v_p = [x_p, y_p]^{\mathrm{T}}$ of the projection of point p on S_i can be expressed as

$$v_p = Kv_d, \tag{3}$$

where K is known as the camera matrix [29] defined by

$$K = \begin{bmatrix} f_{c1} & \alpha_c f_{c1} & c_{c1} \\ 0 & f_{c2} & c_{c2} \\ 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

v_n is the target value and v_d is known, but the implicit formula Eq. (1) cannot be solved directly. Therefore, the iterative method is used to solve v_n. The parameter v_i ($i =$ 1,2,…) represents the ith iteration result of v_n, and v_{i+1} is given by

$$v_{i+1} = \frac{v_i - v_t}{1 + k_{c1}u_n^2 + k_{c2}u_n^4 + k_{c5}u_n^6}, \tag{5}$$

where v_d is the initial input ($v_0 = v_d$). The iterative terminating conditions are $\|v_{i+1} - v_i\| < t_{ite}$, where t_{ite} is the terminating threshold of iteration.

Marker detection

The marker edges have to extract using the method of image segmentation before performing marker recognition. Although threshold segmentation can be implemented swiftly, the image segmentation method using single- or multi-threshold makes the results instability when the OTS is carried out in an environment with a complex background. The Canny edge detector [30] is the common method to extract the marker edges. I^c denotes the result of an original image applied by the Canny filter. The distorted contour set $G^d = \{g^d_i\}$ in I^c is then obtained by the border following the method in [31], where $g^d_i = \{p^d_j\}$ denotes the ith distorted contours and p^d_j denotes the single pixel in g^d_i. It is time consuming to rectify the distortion of the entire original image that only the distortion of G^d needs to be rectified. Therefore, p^d_j is transformed into the image coordinate v^d_i using Eq. (3) and the distortion is rectified by Eq. (5). The undistorted coordinate v_i in S_i, the undistorted contour $g_i = \{v_j\}$, and the undistorted contour set $G = \{g_i\}$ are eventually acquired.

The first step of marker recognition is to determine whether the edge is approximate to a circle. If the edge g_i is a perfect circle, then the distance d_k from the pixel q_k on the circle edge to the circle center c_i is equivalent with any other distance d_j ($j \neq k$). By contrast, the distance d_j is not equivalent to any other distance d_j ($j \neq k$) when g_i does not perfectly represent a circle. The average distance of the set $\{d_j\}$ is denoted as e_j, and the root mean square error (RMSE) of $|d_j - e_j|$ can be used to determine the similarity

between g_i and the circle with radius e_j. Let n_i be the number of pixels in $\{g_i\}$ and c_i be the mass center of g_i:

$$c_i = \sum_{j=1}^{n_i} \frac{q_j}{n_i}. \tag{6}$$

Suppose $d_j = \| q_j - c_i \|$. Given that those contours with the same shape yet different sizes can be classified by an RMS, let r_j be the ratio of d_j deviating from e_j:

$$r_j = \frac{(d_j - e_j)}{e_j}. \tag{7}$$

Then, the RMS of r_j can be represented the extent of similarity between the contour and a circle, which can be denoted as $RMS(r)_i$:

$$RMS(r)_i = \sqrt{\sum_{j=1}^{n_i} r_j^2 / n_i}, \tag{8}$$

g_i can be regarded as a circle if $RMS(r)_i$ is less than a specific threshold t^c.

The second step of marker recognition is to determine whether the intensity value of the region inside the edge is higher than a specific threshold value. If the average intensity value is higher than the specific threshold t^p and $RMS(r)_i < t^c$, then g_i is considered as the edge of a marker projection. Therefore, the undistorted candidate marker contour set $G^m = \{g^m_k\}$ can be obtained.

Calculation of the projection center

The rays, which are emitted from the optical center and tangential with a sphere, form a cone and intersect with the image plane. The intersecting region is an ellipse which is the projection of the sphere. As shown in Fig. 3, the ellipse center c_1,
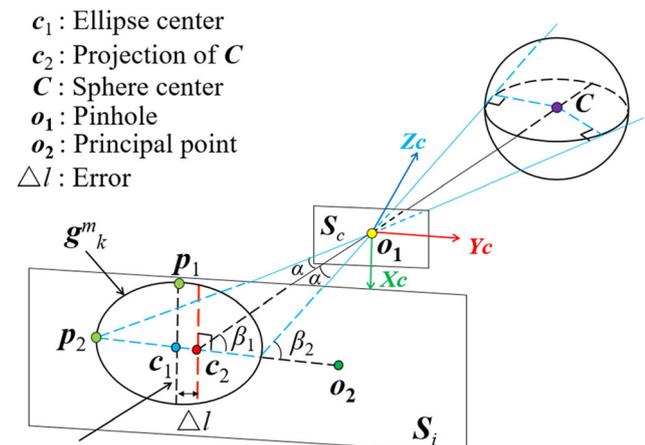

c_1 : Ellipse center
c_2 : Projection of C
C : Sphere center
o_1 : Pinhole
o_2 : Principal point
$\triangle l$: Error

Fig. 3 The projection of a marker in the image plane S_i is an ellipse, where c_1 denotes the ellipse center and c_2 denotes the projection of the marker center

the length of the major axis l, and the length of the minor axis s can be obtained by fitting the edge point set g^m_k by using the method described in [32]. One of the end point p_1 in the minor axis and one of the end point p_2 in the major axis can then be obtained.

The center of ellipse c_1 and the projection of sphere center c_2 are generally not located in the same position due to the line o_1c_2 that divides the major axis of the ellipse into two segments with different lengths. Assume that $\Delta l = \| c_1 - c_2 \|$ is the error between $\overrightarrow{o_2c_1}$ and $\overrightarrow{o_2c_2}$, $C = [x_c, y_c, z_c]^T$ is the 3D coordinate of the marker, and d is the distance between C and o_1. According to the geometric relationship presented in Fig. 3, we obtain $\alpha = \arcsin(r^m/d)$, $\beta_1 = \arcsin[z_c/d]$, and $\beta_2 = \beta + \alpha$. The error of the computed projection center Δl can be obtained as follows based on the sine theorem:

$$\Delta l = \left| \frac{1}{2} - \frac{\sin\alpha\sin(\pi - \beta_2)}{\sin 2\alpha\sin(\pi - \beta_1)} \right| l. \tag{9}$$

c_2 can then be computed as

$$\overrightarrow{o_2c_2} = \overrightarrow{o_2c_1} - \overrightarrow{c_2c_1}. \tag{10}$$

Accurate stereo-matching

The EC method is commonly used to find the correspondences between two images. Without the misrecognition of markers, any marker that satisfies the epipolar constraint should lie on a common epipolar line. The epipolar constraint provides a reduced searching space from a 2D image to an epipolar line. In enforcing this constraint, all those cases where n ($n \geq 2$) markers lie on the same epipolar line can be sorted out, while those cases with only one marker on the corresponding epipolar line can be directly applied to reconstruct the 3D coordinate, thereby reducing the implementation time. However, the case where n markers are lying on the same epipolar line remains an issue to be dealt with in the situation of the lack of information to determine the corresponding relationship between the marker points. The EC method, which is applied in Polaris, reserves $n \times n$ correspondences that contain n real markers and $n \times (n-1)$ ghost markers. The EOC method is used to find n correspondences in the case of n ($n \geq 2$) markers lying on the same epipolar line, but m ($m \geq 2$, $m \leq n$) ghost markers may be presented. Meanwhile, the misrecognition of markers gives rise to more ghost markers. The incorrect correspondences can be explained by the fact that both the EC and EOC methods only consider the 2D coordinates of marker centers without sufficient information to find the correct correspondences.

To address this issue, this paper proposes a stereo-matching method based on EC. The marker projections in a 2D image are presented as white blobs with circular shapes and different

sizes. The distance d can be estimated based on the geometric information related to the size of the blobs in the 2D image. The 3D locations of markers, which includes real and ghost markers that belong to the same marker blob in the 2D image, are restricted to the projection direction l_p. The coordinate of real marker can be estimated according to l_p and d. After obtaining the estimated coordinates of markers $\boldsymbol{P}^e = \{p^e_l\}$ in the 3D space, the coordinates $\boldsymbol{P} = \{p_{lr}\}$, which are reconstructed by the triangulation method and contain real and ghost markers, are computed and compared with $\boldsymbol{P}^e$. The ghost markers in $\boldsymbol{P}$ can be eliminated because the error between the estimated coordinate and the coordinate of real marker along the same projection direction is less than the permissible error. Therefore, all correct coordinates are obtained by using the proposed stereo-matching method.

Estimation of 3D coordinates by using a single camera

In Fig. 4, $\boldsymbol{D}$ denotes the line that is parallel to the minor axis, which passes through c_2 and intersects with the ellipse. According to the results of elliptical fitting, the length of $\boldsymbol{D}$ can be computed as

$$\|D\| = s\sqrt{1 - \left(\frac{2\Delta l}{l}\right)^2}. \qquad (11)$$

Due to $\Delta o_1 p_3 c_2$ and $\Delta o_1 p_3 \boldsymbol{C}$ are similar triangles, the relationship between $\|\boldsymbol{D}\|$ and d can be expressed as

$$\frac{\|D\|/2}{1/\cos\theta} = \frac{r^m}{d\cos\alpha}, \qquad (12)$$

where d is the distance between the marker and optical center, r^m is the radius of the marker, and θ is the angle between the projection path $\overrightarrow{o_1 C}$ and optical axis. By combining Eq. (11)

c_1 : Ellipse center
c_2 : Projection of C
C : Sphere center
o_1 : Pinhole
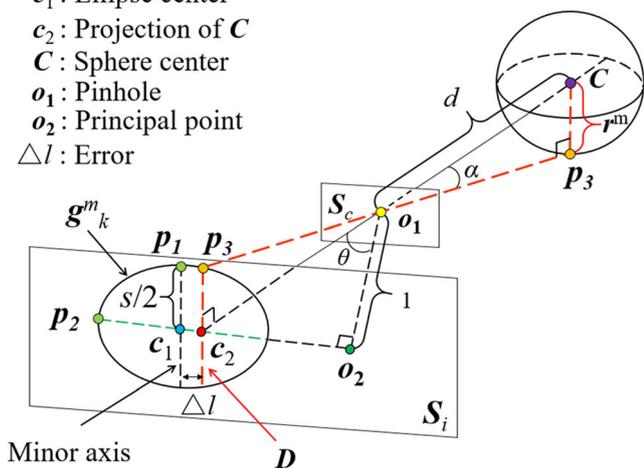o_2 : Principal point
$\triangle l$: Error

Fig. 4 The relationship between s and d is inversely proportional

and Eq. (12), the relationship between s and d can be expressed as

$$d = \frac{2r^m}{s\cos\alpha\cos\theta\sqrt{1 - (2\Delta l)^2/(l)^2}}. \qquad (13)$$

The left center and minor axis length s of the marker projections in $\boldsymbol{S}_i$ are denoted by $\boldsymbol{C}^L = \{c_l\}$ and $\boldsymbol{S}^L = \{s_l\}$, respectively, where $c_l = [x_l, y_l, 1]^T$. The estimated distance between the marker and optical center of the left camera is obtained from Eq. (13), while the estimated coordinates in the left camera frame can be computed as

$$p^e_l = \frac{d^e_l c_l}{\sqrt{x^2_l + y^2_l + 1}}. \qquad (14)$$

Generally, the estimated coordinates are not equal to the real coordinates due to various types of noise, such as model error, image noise, and lens distortion. Nevertheless, the error between the estimated and real coordinate is less than the permissible error. Thus, $\boldsymbol{P}^e_l$ can be used to eliminate the ghost markers in $\boldsymbol{P}$.

The proposed method does not consider the situation of marker occlusions, the occlusion part of markers introduces errors in the calculation of center coordinates. Supposing that markers A and B are coplanar with the two optical centers and given the coordinate p^A of marker A, marker B may be located in the four regions ($\boldsymbol{R}_1$, $\boldsymbol{R}_2$, $\boldsymbol{R}_3$, and $\boldsymbol{R}_4$) of plane $\boldsymbol{O}_{AB}$ that comprise markers A and B and optical centers p_1 and p_2 as shown in Fig. 5(a).

The situation of marker B in $\boldsymbol{R}_1$ is initially examined. Among all markers located in $\boldsymbol{R}_1$, B_1 is located closest to marker A. In Fig. 5(b), o_l and o_r denote two optical centers, p^A denotes the center of marker A, p^B_1 denotes the center of marker B_1, l_1 to l_6 are straight lines on plane $\boldsymbol{O}_{AB}$, l_2 and l_5 are tangential to markers A and B_1, and l_1, l_3, l_4, and l_6 pass through p^A and p^B_1. According to these geometric constraints, l_1 to l_6 and p^B_1 are confirmed. Therefore, the coordinates of p^c_1 and p^c_2 and the line $\boldsymbol{E}^p_1 : p^c_1 p^c_2$ can be obtained, and their relationship can be described as $e^p_1 = \| p^c_1 - p^A \| = \| p^c_2 - p^A \| = \| \boldsymbol{E}^p_1 \|/2$ as shown in Fig. 5(c). To distinguish marker A and ghost marker g on projection direction l_4, the estimated coordinate p^e of marker A should lie on $\boldsymbol{E}^p_1$.

If the estimated coordinate p^e lies on $\boldsymbol{E}^p_1$, that is, $\| p^e - p^A \| < e^p_1$, then p^A can be regarded as the coordinate of a real marker. The other three situations of marker p^B can be deduced in a similar way and three permissible errors e^p_2, e^p_3, and e^p_4 are eventually obtained. Summarily, the range of the four permissible errors is acquired through each 3D coordinate, while the ghost markers can be removed by using the permissible error and estimated coordinate.
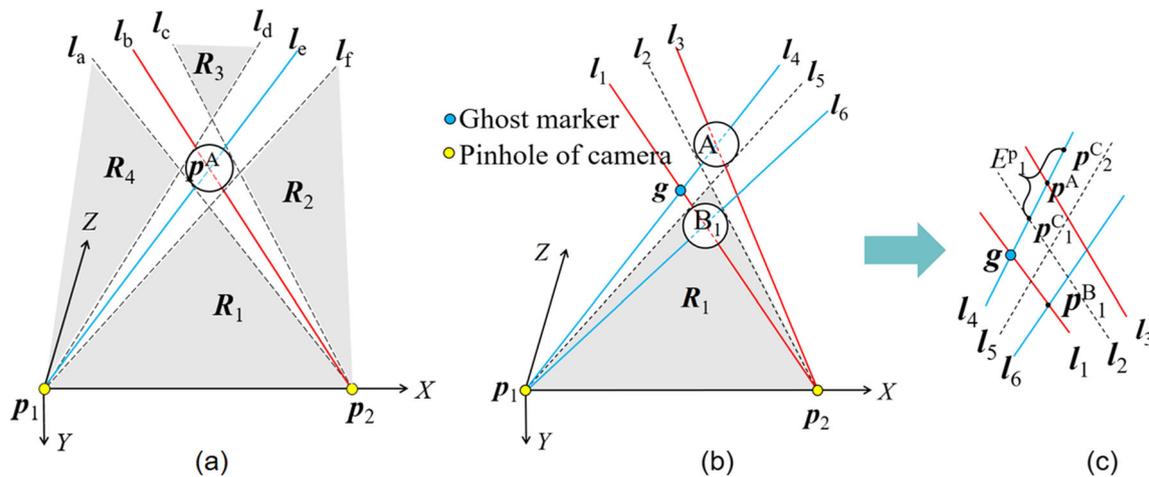
Fig. 5 Calculating the permissible error between the estimated and correct coordinates. (a) Four regions of marker B are located on the lower, upper, left, and right sides of marker A, respectively. Lines l_a to l_f are tangential to the markers; (b) the situation of marker B in R_1, and g is a ghost marker lying on l_4; (c) p^A and p^B_1 denote the marker centers, the line segment E^p_1 can be acquired from the geometric constraint relation

Obtaining correct 3D coordinates

In [29], the triangulation method is applied to reconstruct 3D coordinates through the left and right 2D coordinates. For the case that n points in the left image and m points in the right image are in the same epipolar line, $n \times m$ pairs of correspondence are established using the EC method. Then $n \times m$ 3D points $\boldsymbol{P}^{3D} = \{p_{lr} \mid l = 1, \ldots, n; r = 1, \ldots, m\}$, which contain many ghost markers, are reconstructed via the triangulation method. The coordinates $P^e = \{p_l^e\}$ in the left camera frame are transformed into coordinates in a 3D space $Q^e = \{q_l^e\}$ by $q_l^e = Rp_l^e + T$, where $\boldsymbol{R}$ and $\boldsymbol{T}$ denote the rotation matrix and translation vector that transform the point coordinates in the left camera frame to a 3D space. Ghost markers can be removed because the estimated error between coordinate p_l^e and the corresponding coordinate of the real marker is less than the permissible error e^p. For p_{lr}, identifying which of e^p_1, e^p_2, e^p_3, and e^p_4 can be treated as the permissible error formula is difficult. Therefore, $e^p_{lr} = \min\{e^p_{1lr}, e^p_{2lr}, e^p_{3lr}, e^p_{4lr}\}$ is assumed to be the permissible error between p_{lr} and p_l^e on the lth projection direction in the left camera frame.

Only one real marker can be found at most in a single projection direction. Compute the error $e_{lr} = \| \boldsymbol{p}^e_l - \boldsymbol{p}_{lr} \|$ between the estimated coordinate and the m coordinates $\boldsymbol{p}_{lr}$ ($r = 1, \ldots, m$) in the lth projection direction in the left camera frame and then compare e_{lr} with the corresponding permissible error e^p_{lr} ($r = 1, \ldots, m$). If $r \in \{1, \ldots, n\}$ makes $e_{lr} < e^p_{lr}$, then $\boldsymbol{p}_{lr}$ is the coordinate of the only real marker in the lth projection direction. Otherwise, $\boldsymbol{p}_{lr}$ is inferred to be a ghost marker. Ultimately, a correct coordinate set $\boldsymbol{Q}$ without ghost markers is obtained.

The reflections on smooth surfaces in surgical environment, such as the screw cap of a medical device or a circular dental tool, can introduce some marker-like interference projections that may possibly lead to the reconstruction of ghost markers. These projections cannot be eliminated directly during the marker recognition process. In this section, the estimated coordinates are evaluated based on projection sizes and utilized to eliminate the ghost markers reconstructed by marker-like interference projections which sizes differ from those of markers.

Fast-tracking of markers

To meet the real-time requirement, marker-based OTSs must process the image as fast as possible. Therefore, the fast-tracking method, which combines image

Fig. 6 Fast-tracking of markers. (a) Red boxes denote the ROIs in the original image; and the (b) blue boxes denote the ROIs in the next compressed frame
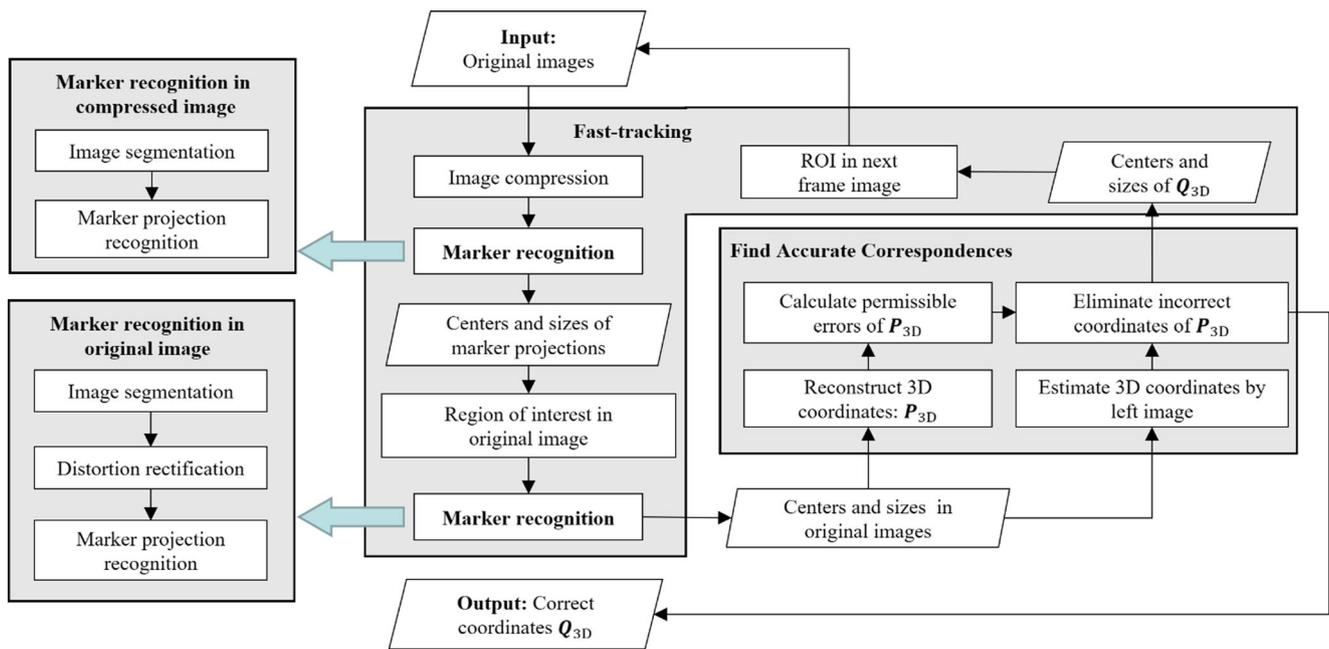

(a)   (b)

Fig. 7 Flowchart of the proposed strategy

Fig. 8 Experiments using different methods to identify markers. The yellow crosses are denoted as the centers of misrecognitions, while the red crosses are denoted as the centers of correct recognitions. (a) The experiment scene in laboratory, including a head phantom, a surgical robot, and a pair of scissors; (b) recognizing markers and speculate misrecognitions based on the output of 3D coordinates obtained from NDI Polaris, and marking the misrecognitions in a 2D image; (c) the outcome of processing the edge image by the method of Hough transform, the recognized red circles with yellow cross inside it are the mistaken recognition which is showed inside the blue box; (d) the outcome of processing the edge image used the proposed method for markers recognition, where $t_c = 0.05$ and $t_p = 200$; (e) surgery scene of liver cancer ablation; (f) results of the proposed method for markers recognition
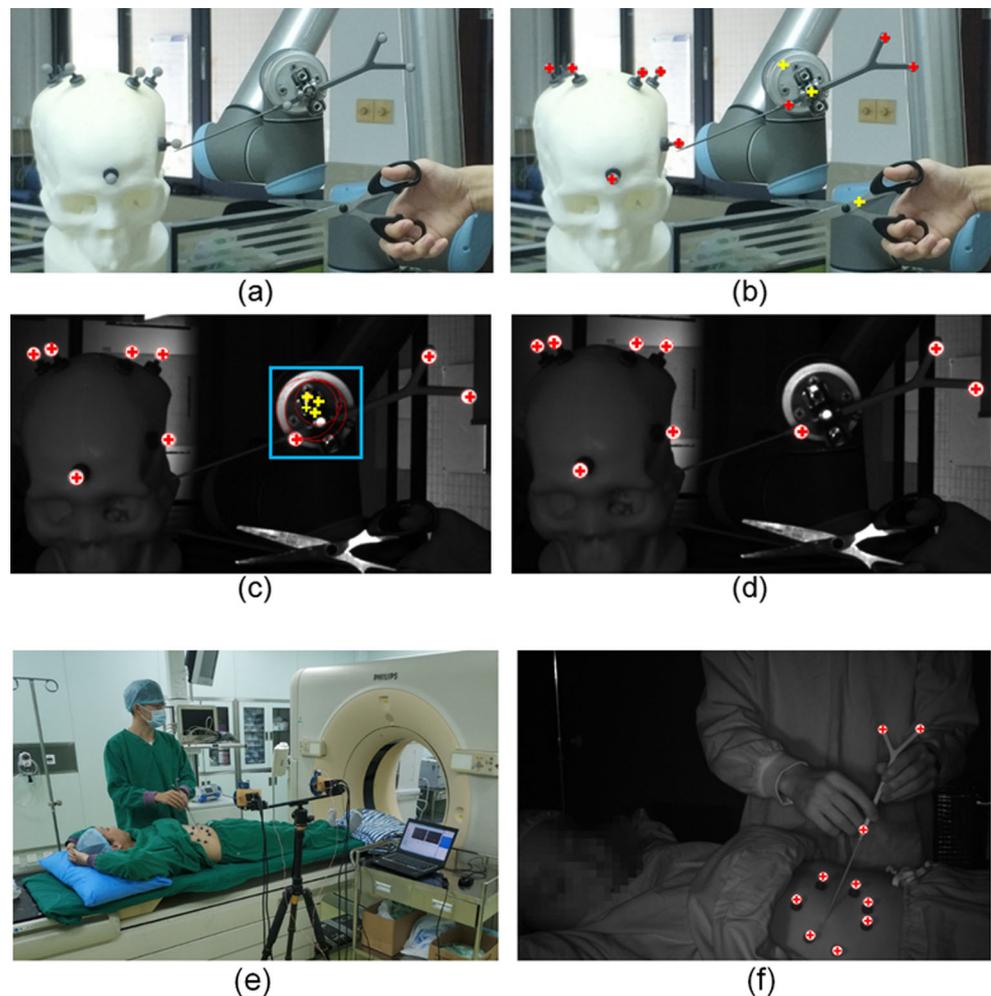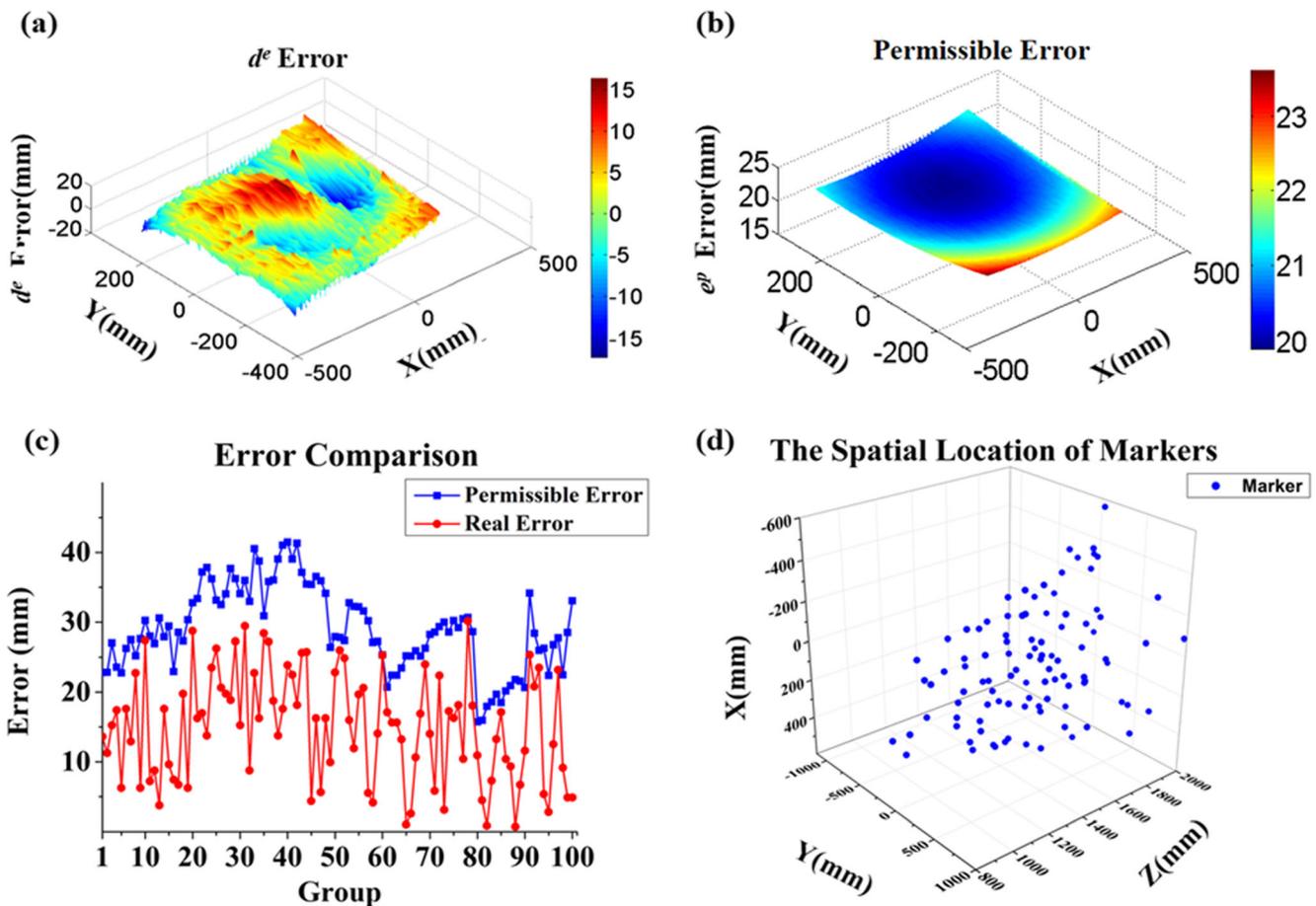
Fig. 9 Validation of Eqs. (13) and (14). (a) and (b) present the results of the first part of the experiment where the markers are on the plane of Z = 1100 mm. (c) and (d) present the results of the second part of the experiment where the markers are in random positions. (a) Errors between the estimated and correct coordinates. (b) Permissible error $\{e^p_k\}$. (c) Permissible $\{e^p_w\}$ and real errors $\{e_w\}$. (d) The spatial locations of 100 markers

compression with the local searching of ROI, is proposed to increase the speed of OTSs up to 50 to 60 frames per second.

Processing the whole original image for every frame is a time-consuming process. Therefore, the fast-tracking method is applied to track the markers in compression images and to search for the accurate radii and centers of markers in ROIs, which are created based on the sizes and centers obtained in

compression images. The original images are initially compressed to $1/n^c$ to obtain compression images. Due to the radii and centers calculated from compressed images can lead to increasing the coordinate extraction error, the accurate centers and projection radii are obtained via local searching of ROIs (as shown in Fig. 6(a), ROIs are blue squares centered the marker centers, and they are slightly larger than marker projections) in the original image.

Fig. 10 The red lines in (b) and (c) are epipolar lines. (a) The head phantom; (b) and (c) projections of markers in the left and right images, respectively
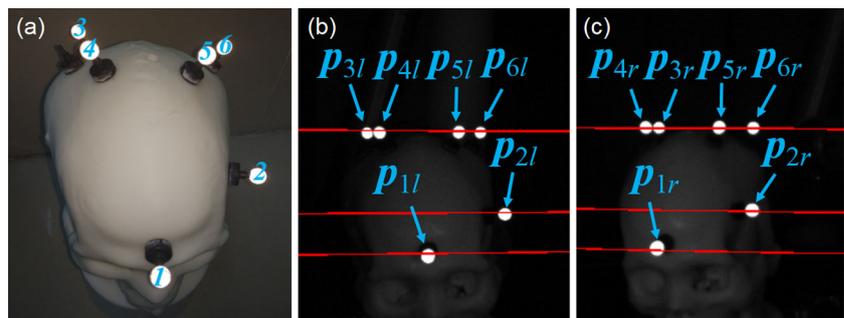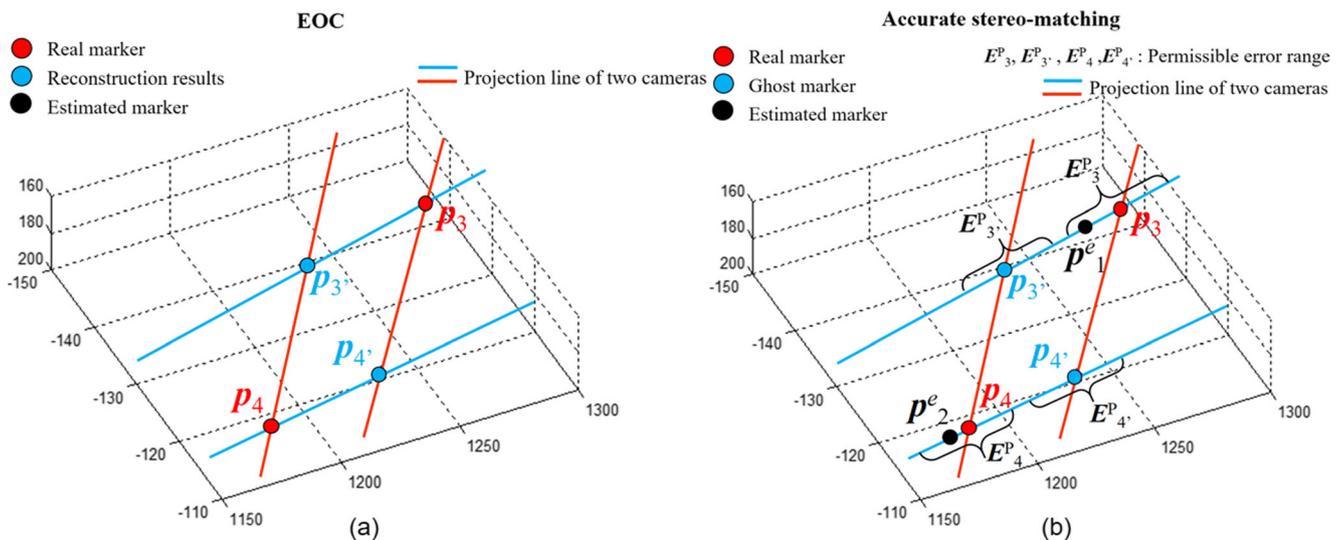
Fig. 11 Comparison of the correspondence results obtained by the EOC and proposed methods. (a) The results obtained by the EOC method; (b) the results obtained by the proposed accurate stereo-matching method,

two coordinates p_3 and p_4 are reserved while $p_{3'}$ and $p_{4'}$ are eliminated because the estimated coordinates p^e_1 and p^e_2 are within the permissible error ranges of E^P_3 and E^P_4

Applying marker recognition in each frame of the whole image wastes much time in processing the image background. Therefore, the whole image must be processed only once per m_f frame to ensure that the new markers appearing within the field of view of the system will be recognized. Only the ROIs of the other $m_f - 1$ frames will be processed to save time. The centers and projection radii are obtained via local searching of ROI (as shown in Fig. 6(b), ROIs are blue squares centered the marker centers, and they are larger than the maximum displacements of the marker projections between two frames) in the next compressed image. Meanwhile, the repetitive rough centers are removed. As shown in Fig. 6, the closer the marker is, the larger the projection radius and ROI will be.

Conclusion

The flowchart of the proposed strategy is shown in Fig. 7. The first frame of the compressed image is processed to obtain the rough center and size of the marker projections, which in turn will be used to create ROIs in the original image. Accurate centers and sizes will be searched in the ROIs of the original image. The correspondences between the left and right marker projections are identified by the EC method, and the point set P^{3D} is reconstructed using the triangulation method. The estimated coordinates and permissible error are computed to remove the ghost markers in P^{3D}. Accurate 3D coordinates Q^{3D} are eventually obtained. In the next m^f-1 frames of compressed images, the ROIs in the current frame are created according to the centers and sizes of marker projections in the last frame.

Experiment

To test the availability and robustness of the proposed method, two opening experimental environments are set up. The proposed method is initially applied in a complex environment with a head phantom to verify its robustness. Afterward, this method is applied in the surgery scene of liver cancer ablation to check its effectiveness.

Table 1 Data of coordinates p_3, $p_{3'}$, p_4, and $p_{4'}$

p_u:(x_u, y_u, z_u)	s/mm	d/mm	d^e/mm	$\lvert d^e - d \rvert$/mm	e^p/mm	Reserve or Not
p_3:(195.39, −142.12, 1294.35)	0.00446	1316.71	1295.12	21.60	24.54	R
p_4:(177.04, −112.36, 1174.58)	0.00491	1193.15	1181.53	11.64	22.30	R
$p_{3'}$:(186.11, −135.65, 1233.07)	0.00446	1254.39	1295.08	72.89	23.46	N
$p_{4'}$:(185.45, −117.45, 1230.21)	0.00491	1249.64	1181.49	68.14	23.28	N

Fig. 12 The red lines in (b) and (c) are epipolar lines. (a) Surgery scene of liver cancer ablation; and (b) and (c) projections of markers in the left and right images, respectively. p_{1l}, p_{2l}, p_{1r}, and p_{2r} are projections of p_1 and p_2 in two images
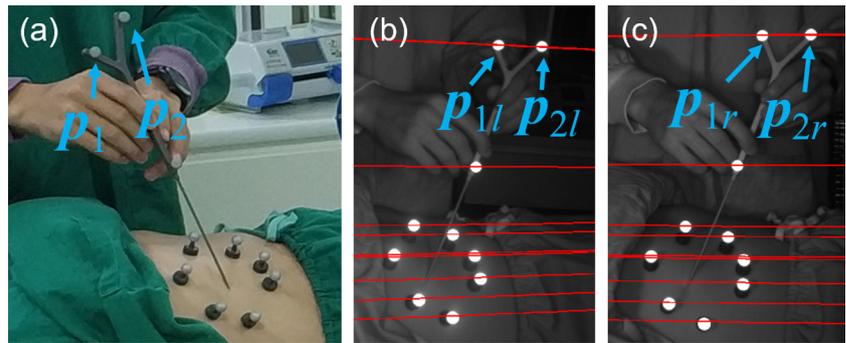
Validation of marker recognition

To validate the robustness of marker recognition, the recognition results of the proposed algorithm are compared with those used the method of Hough transform and NDI Polaris, of which the results obtained from NDI Polaris find the presence of ghost markers due to misrecognition as shown in Fig. 8(b). Meanwhile, as can be seen in Fig. 8(c), some high-intensity regions are recognized as markers by the method of Hough transform in the step of extracting marker. From Figs. 8(b) to 8(d), the results show that the proposed method is more accurate than the method of Hough transform and the method used by NDI Polaris. Moreover, for the same image in the same laptop, the processing time to recognize the markers is 16 ms and 2 ms for the method of Hough transform and the proposed method. Therefore, the proposed method can recognize markers accurately and rapidly.

Validation of estimation distances and calculation of errors

An experiment is designed to verify whether Eqs. (13) and (14) can be applied to estimate the marker coordinates at any location within the field of view. The Universal Robot (Denmark) is used to control a separate marker moving at a plane of $z = 1100$ in the left camera frame (the z coordinates of the markers are floating within 1100 ± 1 mm as a result of shaking the robot base), and the marker coordinates $\{p_k\}$ and minor axis length $\{s_k\}$ of the projection are recorded in real time. Meanwhile, the estimated distances $\{d_k^p\}$ are calculated by Eq. (13), the errors $\{e_k\}$ are computed by $e_k = \|d_k^e - \|p_k - o_1\|\|$ as shown in Fig. 9(a) where o_1 is the left optical center, and the permissible errors $\{e_k^p\}$ are calculated according to $\{p_k\}$ as shown in Fig. 9(b). Figs. 9(a) and 9(b) show that $\{e_k^p\}$ exceeds 20 mm while the corresponding $\{e_k\}$ is below 18 mm.

The Universal Robot (Denmark) is then used to control the separate marker to move randomly to 100 different positions as shown in Fig. 9(d). The coordinates $\{p_w\}$ and minor axis length $\{s_w\}$ are recorded, and the errors of estimated coordinates $\{e_w\}$ and permissible error $\{e_w^p\}$ are calculated. Fig. 9(c) shows that $e_w < e_w^p$ where $w = 1,2\ldots,100$. Summarily, Eqs. (13) and (14) can be used to estimate the coordinates of markers at any location within the field of view.

Verification of accurate stereo-matching

Eqs. (13) and (14) are applied to address the circular interference and the situation where multiple markers are coplanar with two optical centers. As shown in Fig. 10, six markers are located on the surface of the head phantom and marker 3, 4, 5, and 6 are coplanar with two optical centers. In this opening experimental environment, the EC, EOC, and the proposed accurate stereo-matching methods are separately applied to find the correspondences. When using the EC method (applied by NDI
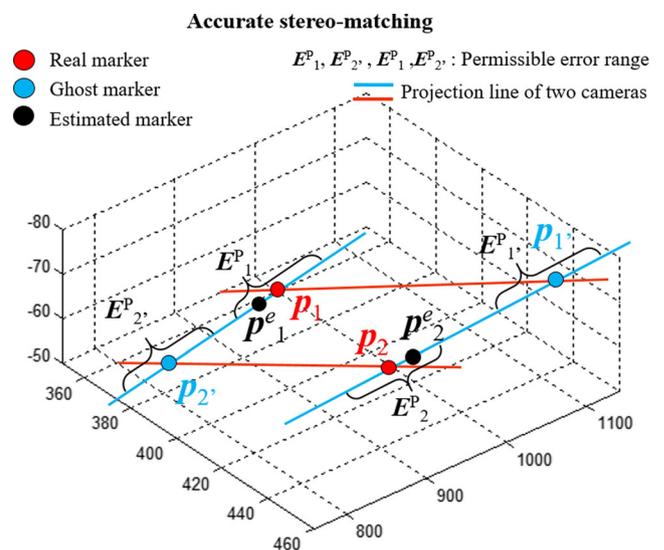

Fig. 13 Results of the proposed accurate stereo-matching method. p_1 and p_2 are real markers while $p_1\cdot$ and $p_2\cdot$ are ghost markers. Two coordinates p_1 and p_2 are reserved, while $p_1\cdot$ and $p_2\cdot$ are eliminated because the estimated coordinates p_1^e and p_2^e are within the permissible error range of E_1^P and E_2^P

Table 2 Data of the four coordinates reconstructed by p_{1l}, p_{2l}, p_{1r}, and p_{2r} in Fig. 13

p_u:(x_u, y_u, z_u)	s/mm	d/mm	d^e/mm	$\lvert d^e - d\rvert$/mm	e^p/ mm	Reserve or Not
p_1:$(-66.10, 388.08, 918.68)$	0.00631	999.47	990.64	8.83	18.55	R
p_2:$(-66.38, 438.15, 912.83)$	0.00624	1014.71	1033.61	18.90	18.95	R
$p_{1'}$:$(-57.72, 383.08, 798.13)$	0.00720	887.18	990.64	103.46	16.48	N
$p_{2'}$:$(-77.70, 453.71, 1077.01)$	0.00534	1171.26	1033.61	137.65	21.84	N

Polaris) to search correspondences, 4×4 3D coordinates in which 12 markers are ghost markers are reconstructed as the final output. This outcome will cause erroneous tracking of medical instruments and patient during treatment.

Only four 3D coordinates (including two ghost markers) reconstructed by the projections of markers 3 and 4 are selected for the analysis. The process adopted for the other 3D coordinates are also used in the reconstruction of these four coordinates. The EOC method finds two false correspondences $[p_{3l}, p_{4l}] \leftrightarrow [p_{4r}, p_{3r}]$, and reconstructs two ghost markers as shown in Fig. 10(a). This method fails to recognize the coordinates of two real markers. When using the proposed accurate stereo-matching algorithm, the EC method finds 4×4 correspondences, which are then used to eliminate the false correspondences. Fig. 11(b) and Table 1 present the process of eliminating the ghost markers, where p_3 and p_4 are denoted as the real markers while p_3' and p_4' are denoted as the ghost markers. The estimated markers p^e_1 and p^e_2 are within the permissible error ranges E^p_3 and E^p_4, respectively. Therefore, p_3 and p_4 are reserved while p_3' and p_4' are eliminated.

p_3 and p_4 are reserved because the error between the estimated coordinates and p_3 and p_4 are less than the permissible errors.

Fig. 11(a) shows that 7 markers are affixed to the patient's skin, and 3 markers are affixed on the medical instrument in the surgery scene of liver cancer ablation. When the medical instrument or the markers in patient's skin is moving in the operation, some markers may be coplanar with two optical centers as shown in Figs. 12(b) and 12(c). Four 3D coordinates, which are reconstructed by p_{1l}, p_{2l}, p_{1r}, and p_{2r} in Figs. 12(b) and 12(c), respectively, are analyzed and described in Fig. 13 and Table 2.

p_1 and p_2 are reserved because the error between the estimated coordinates and p_1 and p_2 are less than the permissible errors.

The marker recognition procedure is unable to eliminate the projections of marker-like interferential objects. However, the mistaken recognition can be removed based on the permissible error. Three coins with different sizes and a reflective surface, shown in Fig. 14, are used as interferential objects and placed in different locations. The 3D coordinates are then reconstructed and the estimated coordinates are calculated according to the left image. Table 3 presents the results.

p^c_1, p^c_2, and p^c_3 are eliminated because the error between the estimated coordinates and p^c_1, p^c_2, and p^c_3 are larger than the permissible errors.

The above finding indicates that the proposed accurate stereo-matching method can accurately identify the correspondences and coordinates of markers.

Verification of image processing acceleration

The process of compressing the original image with 1200×1596 pixels to an image with 800×1064 pixels will save half of processing time. In the experiments, markers are distributed at a distance of about 1.3 m away from the cameras. The average time to process the original left and right images takes 50 ms, and the average time to process the corresponding compressed images takes 28 ms. However, the processing time taken by ROIs method in the next original frame is less than that in the next compressed frame. The results are shown in Fig. 15.

Processing the compressed images and setting the ROIs in the next frame can significantly accelerate the tracking of OTS. When the number of markers is less than 25 in the visual field, the acceleration algorithm can provide a speed of more than 50 frames per second. When the number of markers is 26 to 40 in the visual field, the tracking

Fig. 14 Three different-sized coins that are used as interferential objects.

Table 3 Three coordinates of coins p^c_1, p^c_2, and p^c_3

p^c_u:(x_u, y_u, z_u)	r/mm	d/mm	d^e/mm	$\lvert d^e - d\rvert$/mm	e^p/mm	Reserve or Not
p^c_1:(158.04, 82.50, 1104.18)	0.00793	1118.48	734.39	384.09	20.15	N
p^c_2:(120.34, 35.64, 1117.55)	0.00637	1124.58	908.75	215.83	20.42	N
p^c_3:(133.68, 40.97, 1132.35)	0.00553	1132.35	1046.8	85.55	20.69	N

speed still keeps in 50 frames per second based on the time curve trend. In surgeries such as liver cancer ablation and neurosurgery, the number of markers used for tracking is usually less than 20. Therefore, the OTS can track the patient and medical instruments using the proposed method in real-time during operations. Moreover, in actual operations, the movement speed of markers affixed on the medical instruments and patient is less than 1500 mm/s. Therefore, the proposed strategy allows the marker-based OTS to track separate markers with a movement speed of less than 1500 mm/s.

Discussion and Conclusion

Although the optical navigation system is widely used, the mainstream commercial products, such as NDI Polaris, are unable to deal with the phenomenon of mistaken recognition and ghost marker that have an effect on operations. Some methods have been introduced to address the situation where the markers are coplanar with the optical centers. However, these methods are not as effective as expected. In this case, marker-based OTSs still face challenges of mistaken recognition and ghost marker.

To solve these problems, this paper proposes a method with high robustness for maker recognition and ghost marker. By taking advantage of the geometric information of markers as well as the relationship between the spatial position and projection size of markers, this method not only recognizes markers and estimates their 3D coordinates but also effectively eliminates ghost markers and mistaken recognitions. This method also analyzes the maximum possible movement speed of markers during the operation and adopts the radii of marker projections to create the ROIs to accelerate the tracking speed, thereby reducing the average processing time for each frame. Experiments are performed to verify the feasibility of the proposed method.

In conclusion, the proposed method can effectively eliminate ghost markers when two or more markers are coplanar with optical centers. This method can satisfy the requirement of obtaining marker coordinates in real time during operations with high speed and accuracy. In the future, additional clinical experiments, such as in neurosurgery and lung cancer ablation, will be conducted to verify the benefits of OTSs in operations.
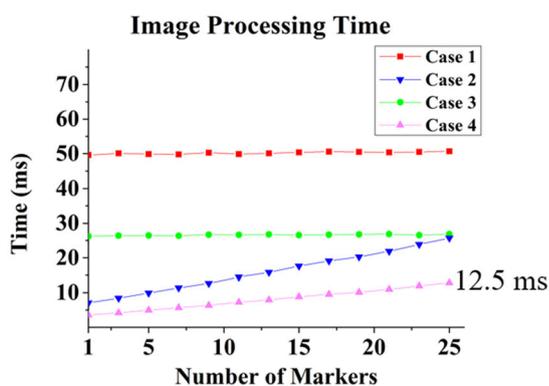
(a)   (b)

Fig. 15 Processing time of left and right images corresponding to the number of markers within the field of view. The processing outputs are 3D coordinates. (a) Surgery scene of liver cancer ablation, 6 to 10 markers are affixed to the patient's skin; (b) the relationship between the number of markers and the processing time of two images. Case 1 represents the processing time for the original image, Case 2 represents the processing time using the ROIs method in the next original frame, Case 3 represents the processing time for the compressed image, and Case 4 represents the time consumed by the proposed fast-tracking method in creating the ROI in the next compressed frame

Compliance with ethical standards

Conflicts of interest The authors declare no conflicts of interest related to this article.

References

1. Wagner, T. H. et al., Optical tracking technology in stereotactic radiation therapy. Med. Dosim. 32(2):111–120, 2007.

2. Meeks, S. L. et al., Optically guided patient positioning techniques. Semin. Radiat. Oncol. 15(3):192–201, 2005.

3. Peters, T. M., Image-guidance for surgical procedures. Phys. Med. Biol. 51(14):505–540, 2006.

4. Shamir, R. R., and Joskowicz, L., Geometrical analysis of registration errors in point-based rigid-body registration using invariants. Med. Image Anal. 15(1):85–95, 2011.

5. Shamir, R. R. et al., Localization and registration accuracy in image guided neurosurgery: a clinical study. Int. J. Comput. Assist. Radiol. Surg. 4(1):45–52, 2009.

6. Huang, Q. H., Wu, B. W. et al., Fully automatic three-dimensional ultrasound imaging based on conventional B-scan. IEEE T Biomed Cirs S. 12(2):426–436, 2018.

7. Huang, Q. H., Lan, J. L. et al., Robotic arm based automatic ultrasound scanning for three-dimensional imaging. *IEEE Trans on Industrial Informatic*s 2018:1-1, 2018.

8. Lin, Q., Cai, K., Yang, R. et al., Geometric calibration of markerless optical surgical navigation system. Int J Med Robotics Comput Assist Surg 15(2):e1978, 2019.

9. Hong, J., and Hashizume, M., An effective point-based registration tool for surgical navigation. Surg. Endosc. 24(4):944–948, 2010.

10. Tauscher, S. et al., OpenIGTLink interface for state control and visualisation of a robot for image-guided therapy systems. Int. J. Comput. Assist. Radiol. Surg. 10(3):285–292, 2015.

11. Pflugi, S. et al., A Cost-effective surgical navigation solution for periacetabular osteotomy (PAO) Surgery. Int. J. Comput. Assist. Radiol. Surg. 11(2):271–280, 2016.

12. Chen, X. et al., Development of a surgical navigation system based on augmented reality using an optical see-through head-mounted display. J. Biomed. Inf. 55(C):124–131, 2015.

13. Pflugi, S. et al., Augmented marker tracking for peri-acetabular osteotomy surgery. Int. J. Comput. Assist. Radiol. Surg. 13(2):291–304, 2018.

14. Hong, J. et al., Medical navigation system for otologic surgery based on hybrid registration and virtual intraoperative computed tomography. IEEE Trans. Biomed. Eng. 56(2):426–432, 2009.

15. Lin, Q. et al., Strategy for accurate liver intervention by an optical tracking system. Biomed. Opt. Express. 6(9):3287–3302, 2015.

16. Soete, G. et al., Initial clinical experience with infrared-reflecting skin markers in the positioning of patients treated by conformal radiotherapy for prostate cancer. Int. J. Radiat. Oncol. Biol. Phys. 52(3):694–698, 2002.

17. Yan, H. et al., A phantom study on the positioning accuracy of the Novalis body system. Med. Phys. 30:3052–3060, 2003.

18. Yan, H., Zhu, G. et al., The investigation on the location effect of external markers in respiratory-gated radiotherapy. J. Appl. Clin. Med. Phys. 9(2):57–68, 2008.

19. Dieterich, S. et al., Respiratory skin motion tracking in stereotactic radiosurgery with the CyberKnife (WIP). Med.Phys. 1256(1256):130–136, 2003.

20. Linthout, N. et al., Six dimensional analysis with daily stereoscopic x-ray imaging of intrafraction patient motion in head and neck treatments using five points fixation masks. Med. Phys. 33(2):504–513, 2006.

21. Liere, R. V., and Mulder, J. D., Optical Tracking Using Projective Invariant Marker Pattern Properties. In: *IEEE Virtual Reality*, pp. 191–198, Los Angeles, 2003.

22. Steinicke, F., et al "Generating optimized marker-based rigid bodies for optical tracking systems", in *Second International Conference on Computer Vision Theory & Applications*, pp. 387–395, Barcelona, Spain, 2007

23. Ribo, M., Pinz, A., Fuhrmann, A. L. (2001) "A new optical tracking system for virtual and augmented reality applications," *the 18th IEEE conference on Instrumentation and Measurement Technology*, pp. 1932–1936, IEEE, Budapest, Hungary.

24. Yan, G., Li, J., Huan, Y. et al (2014) "Ghost marker detection and elimination in marker-based optical tracking systems for real-time tracking in stereotactic body radiotherapy," *Med. Phys.* **41**(10):101713–1–101713-10.

25. Lin, Q. et al., Robust stereo-match algorithm for infrared markers in image-guided optical tracking system. IEEE Access 6:52421–52433, 2018.

26. Ioannou, D. et al., Circle recognition through a 2D hough transform and radius histogramming. Image Vis. Comput. 17(1):15–26, 1999.

27. Lestriandoko, N. H., and Sadikin, R., (2017) "Circle detection based on hough transform and Mexican Hat filter", *International Conference on Computer*, pp. 153–157, IEEE, Tangerang, Indonesia.

28. Zhang, M., and Cao, H., "A new method of circle's center and radius detection in image processing", *IEEE International Conference on Automation & Logistics*, pp. 2239–2242, IEEE, Qingdao, Peoples Republic of China 2008

29. Cai, K. et al., Near-Infrared camera calibration for optical surgical navigation. J. Med. Syst. 40(3):1–12, 2016.

30. Canny, J., A computational approach to edge detection. IEEE T Pattern Anal. 8(6):679–698, 1986.

31. Suzuki, S., and Be, K., Topological structural analysis of digitized binary images by border following. Computer Vision, Graphics, and Image Processing. 30(1):32–46, 1985.

32. Fitzgibbon, A. W., et al "Direct least squares fitting of ellipses", *International Conference on Pattern Recognition*. pp. 253–257, IEEE Comput. Soc. Press, Vienna, Austria, 1996

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.