Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Development of a monocular vision system for robotic drilling*

Wei-dong ZHU†[1,2], Biao MEI[1,2], Guo-rui YAN[1,2], Ying-lin KE[1,2]

([1]*Department of Mechanical Engineering, Zhejiang University, Hangzhou 310027, China*)

([2]*The State Key Lab of Fluid Power Transmission and Control, Zhejiang University, Hangzhou 310027, China*)

†E-mail: wdzhu@zju.edu.cn

Received Dec. 29, 2013; Revision accepted Mar. 3, 2014; Crosschecked July 16, 2014

Abstract: Robotic drilling for aerospace structures demands a high positioning accuracy of the robot, which is usually achieved through error measurement and compensation. In this paper, we report the development of a practical monocular vision system for measurement of the relative error between the drill tool center point (TCP) and the reference hole. First, the principle of relative error measurement with the vision system is explained, followed by a detailed discussion on the hardware components, software components, and system integration. The elliptical contour extraction algorithm is presented for accurate and robust reference hole detection. System calibration is of key importance to the measurement accuracy of a vision system. A new method is proposed for the simultaneous calibration of camera internal parameters and hand-eye relationship with a dedicated calibration board. Extensive measurement experiments have been performed on a robotic drilling system. Experimental results show that the measurement accuracy of the developed vision system is higher than 0.15 mm, which meets the requirement of robotic drilling for aircraft structures.

Key words: Vision system, Robotic drilling, Error measurement, Elliptical contour extraction, Hand-eye calibration
doi:10.1631/jzus.C1300379 **Document code:** A **CLC number:** TP391.4

1 Introduction

In aircraft manufacturing, the assembly process requires a large amount of work and is an influential factor to product quality and productivity. Aircraft parts are typically made of aluminum alloys, titanium alloys, or carbon fiber reinforced plastics, which are assembled into larger structures using riveting or bolting methods. For example, in a typical wing panel or fuselage panel, thousands of fasteners (rivets or bolts) are used to join the stringers, frames, and skin to form the panel structure. In such an assembly process, the drilling of fastener holes is a difficult task due to the large number of holes and the hard-to-machine material property of aircraft parts. To meet the high quality, high productivity, and low cost requirements of the competitive aerospace industry, robotic drilling technology has been extensively studied by researchers from industry and academia. Olsson *et al.* (2010) developed an effective robotic drilling prototype system, demonstrating the great potential of high precision industrial robot-based drilling systems in the field of aircraft assembly. Webb *et al.* (2005) developed a flexible robotic drilling unit, which can be used to drill, countersink, insert, and upset rivets on aircraft structures. It was tested and successfully demonstrated on the Bombardier CRJ700 regional jet. Electroimpact Inc. developed the ONCE robotic drilling system and used it to drill, countersink, and inspect fastener holes of the wing trailing edge flaps of Boeing's F/A-18E/F fighter aircraft (DeVlieg *et al.*, 2002).

In robotic drilling, computer-aided design (CAD) models of the production environment are used as the basis for the generation of robot programs. However, the CAD model is not ideally coincident with the real

* Project supported by the National Natural Science Foundation of China (Nos. 51205352 and 51221004)

situation regarding the shapes, positions, and orientations of the workpieces and jigs, leading to position errors of the fastener holes created by the robotic drilling system. To improve the position accuracy of drilled holes, robot programs should be created in accordance with the actual production environment. Furthermore, positioning errors of industrial robots should be compensated for since they have a marginal positioning accuracy compared with the high standard of accuracy for fastener holes in the aerospace industry (Strobl and Hirzinger, 2008; Olsson *et al.*, 2010).

Given the features of non-contact measurement, relatively low cost, and satisfactory precision, vision systems have been widely used in manufacturing and other sectors, such as robotic welding (Chen *et al.*, 2003; Pachidis and Lygouras, 2007), robotic assembly (Bone and Capson, 2003), robot calibration (Motta *et al.*, 2001), inspection and quality control (Armingol *et al.*, 2003; Neto and Nehmzow, 2007), surface roughness evaluation (Lee and Tarng, 2001), measurement and positioning (Malassiotis and Strintzis, 2003; Eberli *et al.*, 2010; Zou *et al.*, 2012; Ibarguren *et al.*, 2014), and automated micromanipulation (Zhou *et al.*, 2000; Bilen *et al.*, 2012). Particularly, several vision systems have been developed for robotic drilling applications (DeVlieg *et al.*, 2002; Zhan and Wang, 2012). The ONCE system used in the assembly of the F/A-18E aircraft can determine the position of the workpiece; however, its cost is relatively high due to the use of a laser vision system and accessory equipment (DeVlieg *et al.*, 2002). Zhan and Wang (2012) developed a vision-based positioning system for a robot drilling system, in which the calibration of hand-eye relationship is achieved using the relationship between the scene frame and the camera imaging frame. The flex track drilling system developed by Electroimpact, Inc. also integrated a vision system (Thompson *et al.*, 2005). The relationship between the coordinate system of the camera and the drill was obtained by drilling a hole in a coupon and then locating it with the vision system.

This research aims to develop a low-cost, high-accuracy monocular vision system for the measurement of reference holes, from which relative errors between the drill tool center point (TCP) and workpiece position can be acquired to enhance the position accuracy of drilled holes through error compensation (Zhu *et al.*, 2013). Although several vision systems have been tested or used in robotic drilling (DeVlieg *et al.*, 2002; Thompson *et al.*, 2005; Zhan and Wang, 2012), an accurate explanation of the working principle of vision systems was omitted in the literature. A deeper understanding of the working principle of vision systems is helpful for achieving higher measurement accuracy. Therefore, in this paper, strict mathematical analysis of the principle of relative error measurement with the monocular vision system is presented. The camera TCP of the vision system is determined so that the Abbe errors are avoided in the measurement process. The data capturing process and data processing algorithms also directly affect the final measurement accuracy of the vision system. Due to erroneous viewing direction and perspective projection, reference holes, which should theoretically be circular, are elliptic in the images captured by a 2D camera. To ensure robust and highly accurate reference hole detection, a saliency-snake elliptical contour extraction algorithm is presented to the measurement software.

System calibration is another important issue in the development of a monocular vision system. Even though camera calibration and hand-eye calibration of vision systems have been extensively studied in the computer vision community (Tsai, 1987; Dornaika and Horaud, 1998; Zhang, 2000; Strobl and Hirzinger, 2008; Strobl *et al.*, 2009; Malti, 2012; Zhao and Weng, 2013), an accurate and efficient calibration method is still lacking for vision systems used in robotic drilling. As noted in Strobl *et al.* (2009), to further increase the already high accuracy (compared with conventional applications of vision systems) of current industrial robots, cameras with a narrow field of view should be used. Since it is difficult for these cameras to gather enough evidence on perspective distortion, calibration accuracy cannot be guaranteed with the standard methods used in the computer vision community (Strobl *et al.*, 2009). For hand-eye calibration, standard methods in the computer vision community rely on constraint equations derived from the erroneous robot motions (from the viewpoint of robotic drilling) (Strobl and Hirzinger, 2008; Malti, 2012; Zhan and Wang, 2012), which inevitably affects the calibration accuracy. As a result, special calibration methods have been used in vision systems for robotic drilling (Thompson *et al.*, 2005; Zhan and Wang, 2012);

however, these methods are tedious and the calibration accuracy may be affected during the calibration process, e.g., excessive setup of calibration targets and drilling holes on a test coupon. In this paper, we propose a method for simultaneous calibration of camera internal parameters and hand-eye relationship with a dedicated calibration board. Accurate calibration of the monocular vision system is achieved by simplifying the calibration process and avoiding the requirements of perspective distortion evidence and constraint equations from robot motion.

2 Introduction to the monocular vision system

2.1 Working principle

Due to the positioning errors of the robot and installation errors of the workpiece, the actual position of the drill TCP is different from the desired hole position when the robot is driven by a program generated in an offline programming environment. To improve the positioning accuracy in robotic drilling, reference holes are pre-drilled on aircraft structures. Before the drilling process, the robot is driven to the nominal positions of the reference holes. Then the relative deviations between the drill TCP and the reference holes are measured to improve the positioning accuracy of robotic drilling through error compensation (Zhu *et al.*, 2013).

In this research, a monocular vision system (Fig. 1) has been developed to measure the deviation between the drill TCP and the reference hole. The workpiece to be measured is put perpendicular to the optical axis of the camera during measurement, the deviation of the center of the reference hole from the camera TCP of the vision system is reflected in the image captured by the camera, and the relative error can be calculated by image processing. Of course, the monocular vision system measures only deviations in the *XY* plane of the drill TCP. It suffices to measure the deviations with a monocular vision system because aircraft structures (wing panel, fuselage panel, etc.) are large in curvature radius and can be approximated by a local flat surface, and deviation along the *Z*-axis of the drill TCP does not affect the positioning accuracy of drilled holes with respect to the aircraft structure (Zhan and Wang, 2012).

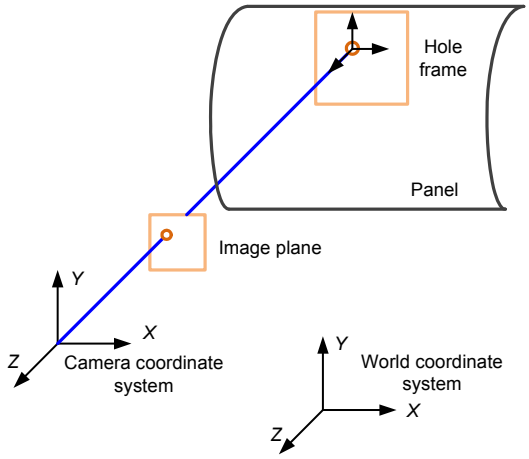

Fig. 1 Reference hole measurement with the monocular vision system

Due to the orientation errors of the robot, the camera TCP should be chosen carefully. The natural coordinate system of a camera is the camera coordinate system (camera CS); however, if it is used as the camera TCP, Abbe errors will occur in the measurement results. With homogeneous matrix representation and small angle approximation, the actual hole frame relative to the actual camera CS can be represented as

$$\begin{bmatrix} 1 & \varepsilon_z(\boldsymbol{\theta})-\varepsilon_z(i) & \varepsilon_y(i)-\varepsilon_y(\boldsymbol{\theta}) & \delta_x(i)-\delta_x(\boldsymbol{\theta})-z\varepsilon_y(\boldsymbol{\theta}) \\ \varepsilon_z(i)-\varepsilon_z(\boldsymbol{\theta}) & 1 & \varepsilon_x(\boldsymbol{\theta})-\varepsilon_x(i) & \delta_y(i)-\delta_y(\boldsymbol{\theta})+z\varepsilon_x(\boldsymbol{\theta}) \\ \varepsilon_y(\boldsymbol{\theta})-\varepsilon_y(i) & \varepsilon_x(i)-\varepsilon_x(\boldsymbol{\theta}) & 1 & z-\delta_z(\boldsymbol{\theta})+\delta_z(i) \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

where $\delta_x(\boldsymbol{\theta})$, $\delta_y(\boldsymbol{\theta})$, $\delta_z(\boldsymbol{\theta})$, $\varepsilon_x(\boldsymbol{\theta})$, $\varepsilon_y(\boldsymbol{\theta})$, $\varepsilon_z(\boldsymbol{\theta})$ are pose errors of the actual drill TCP with respect to its theoretical counterpart, $\boldsymbol{\theta}$ is the vector of joint angles, $\delta_x(i)$, $\delta_y(i)$, $\delta_z(i)$, $\varepsilon_x(i)$, $\varepsilon_y(i)$, $\varepsilon_z(i)$ are pose errors of the actual hole frame with respect to its theoretical counterpart, i is the hole index, and z is the object distance (Fig. 2).

Hence, the 2D relative errors detected by the monocular vision system are ($\delta_x(i)-\delta_x(\boldsymbol{\theta})-z\varepsilon_y(\boldsymbol{\theta})$, $\delta_y(i)-\delta_y(\boldsymbol{\theta})+z\varepsilon_x(\boldsymbol{\theta})$), where $z\varepsilon_y(\boldsymbol{\theta})$ and $z\varepsilon_x(\boldsymbol{\theta})$ are the Abbe errors.

In this study, we define the camera TCP as an offset of the camera CS along its *Z*-axis (the optical axis of the camera) by the object distance (Figs. 1 and 3). The pose of the actual hole frame relative to the camera CS is

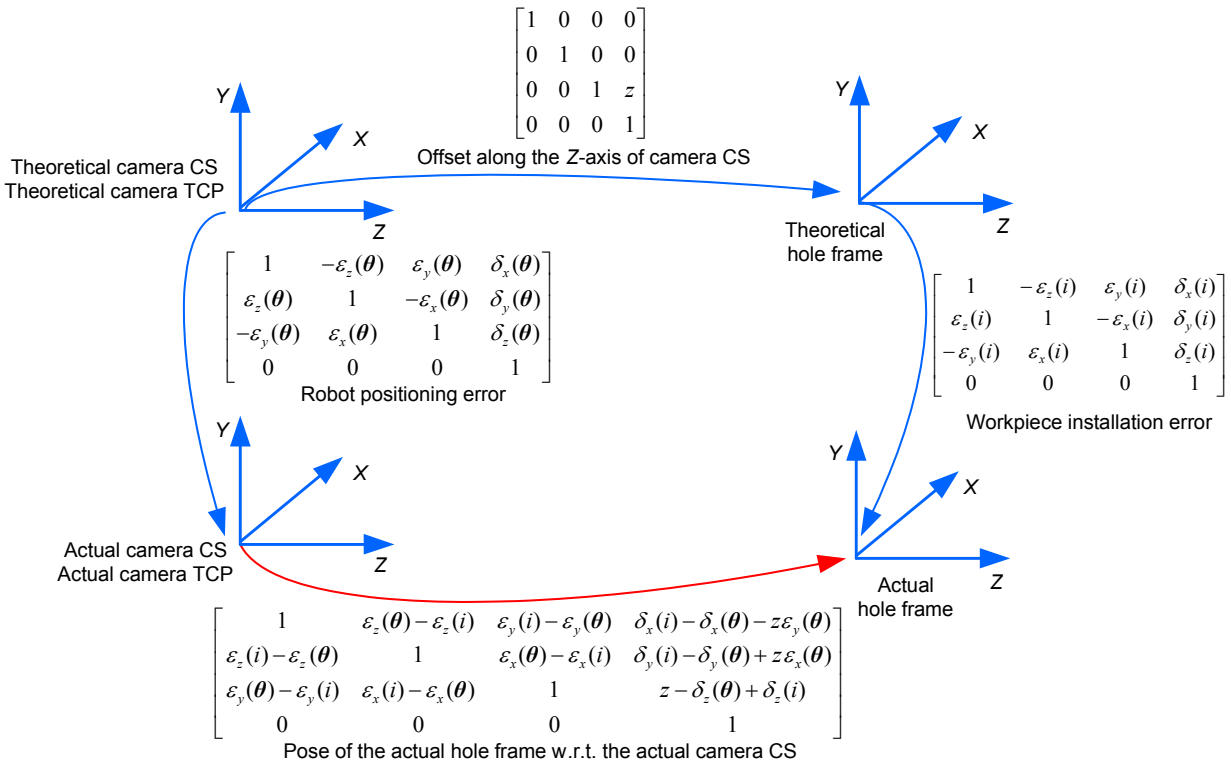

Fig. 2 Relative error measurement when the camera CS is used as the camera TCP

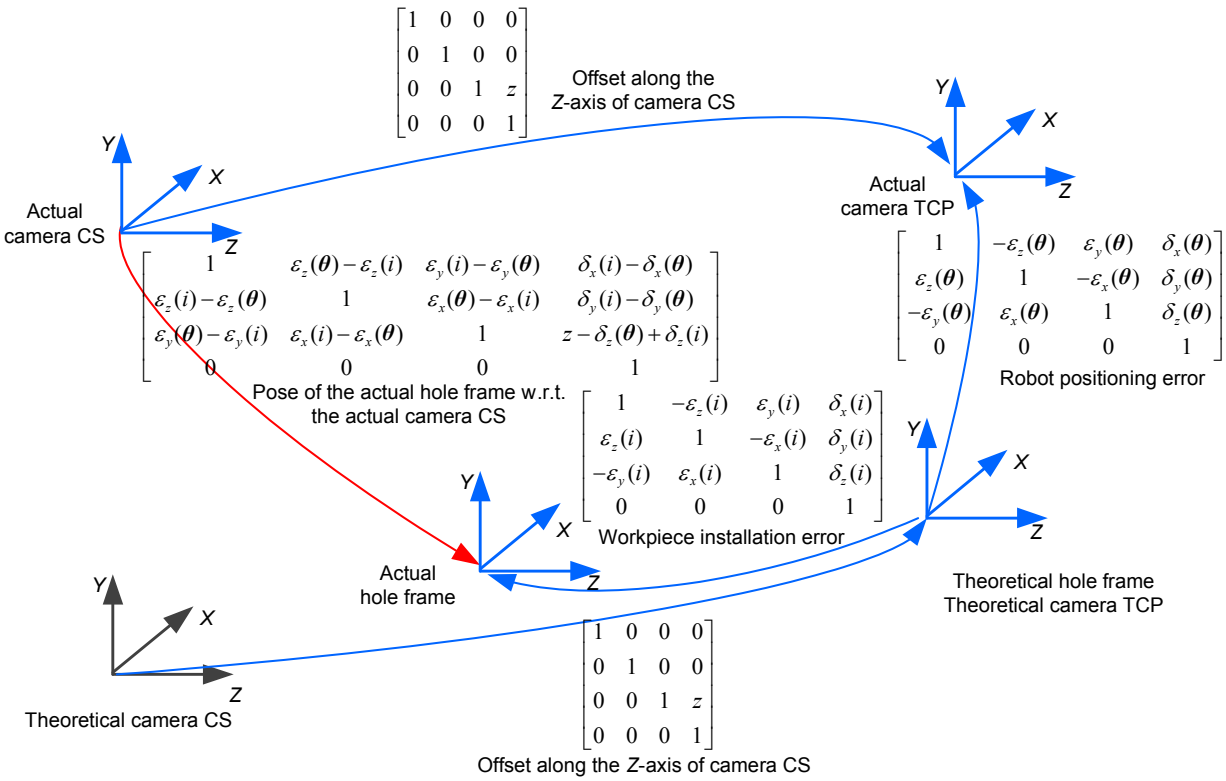

Fig. 3 Relative error measurement when the camera TCP is defined as an offset of the camera CS

$$\begin{bmatrix} 1 & \varepsilon_z(\boldsymbol{\theta})-\varepsilon_z(i) & \varepsilon_y(i)-\varepsilon_y(\boldsymbol{\theta}) & \delta_x(i)-\delta_x(\boldsymbol{\theta}) \\ \varepsilon_z(i)-\varepsilon_z(\boldsymbol{\theta}) & 1 & \varepsilon_x(\boldsymbol{\theta})-\varepsilon_x(i) & \delta_y(i)-\delta_y(\boldsymbol{\theta}) \\ \varepsilon_y(\boldsymbol{\theta})-\varepsilon_y(i) & \varepsilon_x(i)-\varepsilon_x(\boldsymbol{\theta}) & 1 & z-\delta_z(\boldsymbol{\theta})+\delta_z(i) \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

And the relative errors measured by the monocular vision system are $(\delta_x(i)-\delta_x(\boldsymbol{\theta}), \delta_y(i)-\delta_y(\boldsymbol{\theta}))$, in which no Abbe error is introduced.

2.2 Hardware, software, and system integration

The 2D vision system is mounted on a robotic drilling end-effector (Fig. 4). The end-effector consists of a base structure, linear guideways, linear scales, a spindle, a pressure foot, a 2D vision system, etc. The base structure, where other parts of the end-effector are installed, also provides the interface for mounting the end-effector onto the robot. Linear guideways and linear scales provide the precise motion of the spindle and pressure foot in the direction of the drill axis. The pressure foot is used to clamp the workpiece in the drilling process in order to stabilize the robotic drilling system and eliminate the gaps between the stacked materials of the workpiece. The vision system is used to measure reference holes pre-drilled on the aircraft structures.

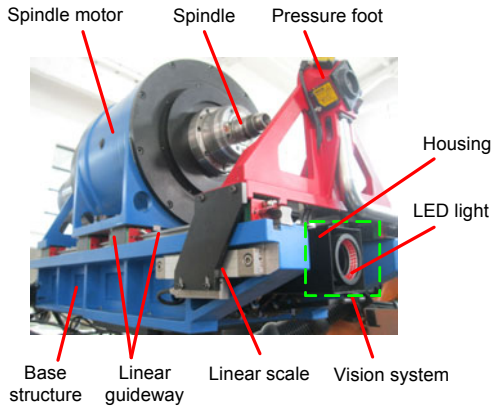

Fig. 4 Robotic drilling end-effector with the 2D vision system

The hardware components of the vision system include a Baumer industrial camera (1296×966 pixels) equipped with a 35 mm Fujinon lens, an annular LED light, a computer, a dedicated calibration board, a housing of the vision system, etc. To achieve high measurement accuracy, the camera is configured to have a small field of view of 18° (visible area is 28 mm×21 mm). The camera and the annular light are coaxially mounted inside the housing of the vision system, which is assembled on the base structure of the end-effector. The optical axis of the camera is installed to be parallel to the spindle axis. Fig. 5 shows the overall architecture of the vision system, including its hardware components, software components, and the integration of these components. The vision software includes two modules, the calibration module and the measurement module.

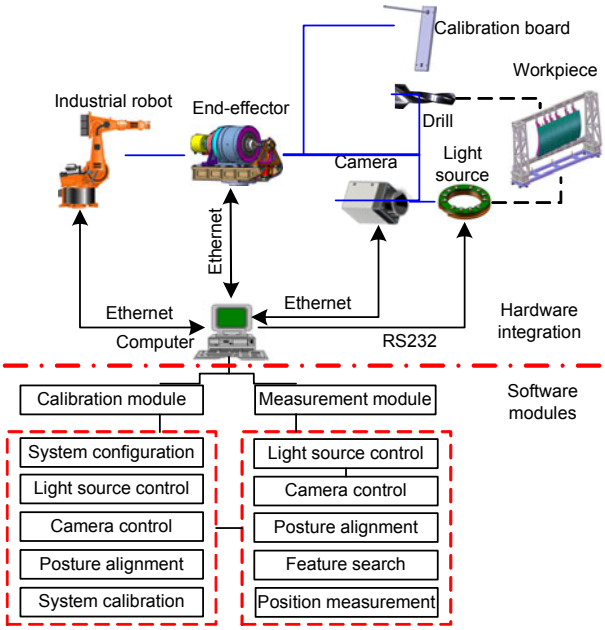

Fig. 5 Overall architecture of the vision system

In the calibration module, functions such as system configuration, light source control, camera control, and system calibration are implemented. Fig. 6 shows the user interface of the calibration module. System configuration manages various coordinate frames of the robotic drilling system (world frame, robot base frame, drill TCP, etc.) as well as the setting of names and paths of data files. The light source control function provides a convenient user interface for controlling the illumination of the calibration board in the calibration process. Major photographing parameters such as exposure and gain can be adjusted using the camera control function to ensure that high quality images are captured. Based on the captured image of the calibration board, camera internal parameters and hand-eye relationship can be obtained with the system calibration function.

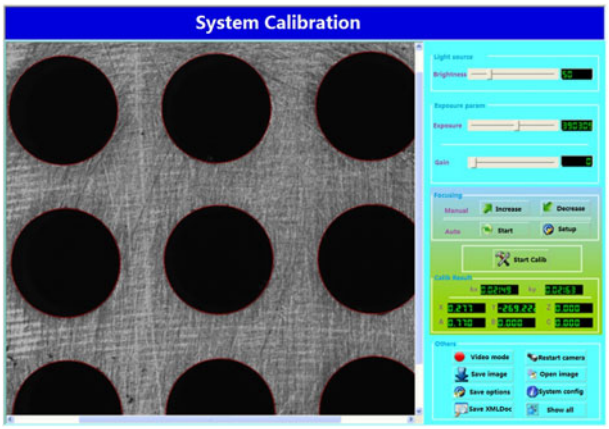

Fig. 6 User interface of the system calibration module

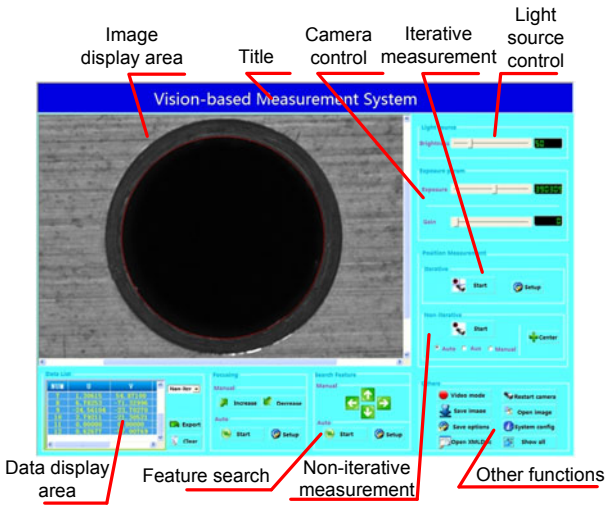

Fig. 7 User interface of the measurement module

In the measurement module (Fig. 7), in addition to light source control and camera control, major functions for reference hole measurement are provided. Feature search can be used conveniently to find the feature to be measured when it is outside the field of view of the vision system. Non-iterative and iterative position measurement functions have been developed to measure the position of reference holes. In non-iterative measurement, hole position is calculated based on a single image captured of the workpiece and is not further refined, while in the iterative measurement process the measured deviation of the reference hole from the camera TCP is used to drive the robot to reduce the deviation iteratively until the center of the reference hole coincides with the camera TCP. This process normally finishes within 3–4 iterations, and the deviation between the initial position and final position (equivalent to the position of the reference hole) of the vision system can be acquired from the robot controller and is taken as the measurement result. Iterative measurement typically achieves higher accuracy than non-iterative measurement at the expense of longer measurement time. Deviation of the reference hole from the camera TCP can be further transformed into the world coordinate system. Together with the information of the identity, name, and nominal position of the hole, this deviation information is recorded in an eXtensible Markup Language (XML) file and used to modify the nominal robotic drilling positions by error compensation (Zhu *et al.*, 2013).

3 Elliptical contour extraction

Reference hole detection involves elliptical contour extraction which directly affects the measurement accuracy of the vision system. Owing to noises and environmental disturbances in robotic drilling, robust and accurate extraction of the reference hole contour is critical to vision-based measurement. Therefore, a novel elliptical contour extraction method called the saliency-snake method is proposed, in which the robustness and accuracy of reference hole detection are improved greatly through saliency evaluation and snake-based optimization. This method consists of three main stages, i.e., salient region detection, voting, and optimization with the snake model (Fig. 8). The three stages of the saliency-snake method are as follows:

Stage 1: salient region detection. The purposes of this stage are enhancement of contrast between the background and the region of the reference hole and robust segmentation of the region of the reference hole from the background. First, estimate the saliency map S of the original image and normalize the pixel values of the map to integers between 0 and 255. Then threshold the saliency map with a fixed threshold T_s to obtain a binary image.

Stage 2: voting. The aim of this stage is to provide a robust initial contour for snake creation in implementation of the snake model. First, retrieve the contours of the binary image obtained in stage 1 using the contour following procedure developed by

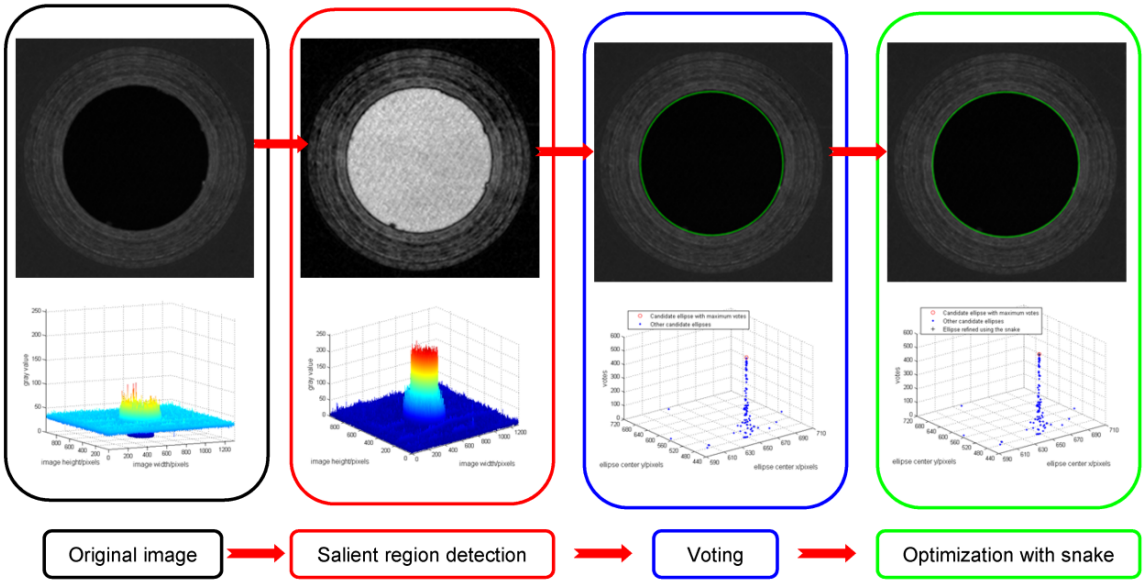

Fig. 8 Flowchart of the saliency-snake method

Suzuki (1985). Then randomly select six points (Bradski and Kaehler, 2008) from the retrieved contours and fit an ellipse E_{cand} to them. Finally, vote for the candidate ellipse E_{cand} according to the distance threshold ε_{d} between contour points and the ellipse. If the number of candidate ellipses reaches M, retrieve the votes on candidate ellipses. The ellipse with the largest number of votes is chosen as the initial contour of the snake model.

Stage 3: optimization with the snake model. To enhance the measurement accuracy for the reference hole, the snake model is used to optimize the contour of the reference hole. First, discretize the ellipse selected in stage 2 on the original image, and the pixels corresponding to the ellipse discretized are used to create the snake model. Then evolve the snake using the greedy algorithm (Williams and Shah, 1992). Finally, the snake will be driven to the real contour of the reference holes by minimizing the energy. If the termination criterion ε_{n} is satisfied, fit an ellipse to the final snake points. The ellipse thus obtained is taken as the result of reference hole detection.

To validate the performance of the proposed saliency-snake method, comparisons with the least-squares method (Fitzgibbon *et al.*, 1999) using the Otsu algorithm (Otsu, 1979) for threshold selection and the Hough transform method (Simonvsky, 2011) are made on six noisy and disturbed images captured with different exposure times (Fig. 9). The results indicate that the saliency-snake method successfully detects reference holes regardless of the variation of exposure time, burrs, and dust flocks, while the least-squares method and Hough transform method fail on some images.

4 Vision system calibration

In this section, we propose a practical and accurate calibration method for the monocular vision system. The method enables simultaneous calibration of camera internal parameters and hand-eye relationship using a high-accuracy dedicated calibration board.

Due to the large curvature radius, a small region containing a reference hole on aircraft structure can be approximated with a plane. Since the optical axis of the camera is perpendicular to the aircraft structure during measurement, the scene's relief is small relative to its average distance from the camera. Therefore, the camera can be effectively modeled with the weak perspective model (Forsyth and Ponce, 2011), in which the magnification can be taken as a constant. As shown in Fig. 10, the coordinates of a point P in the scene plane and its projected point P_0 in the image plane are related by

$$\begin{cases} x_0 = mx, \\ y_0 = my. \end{cases} \tag{3}$$

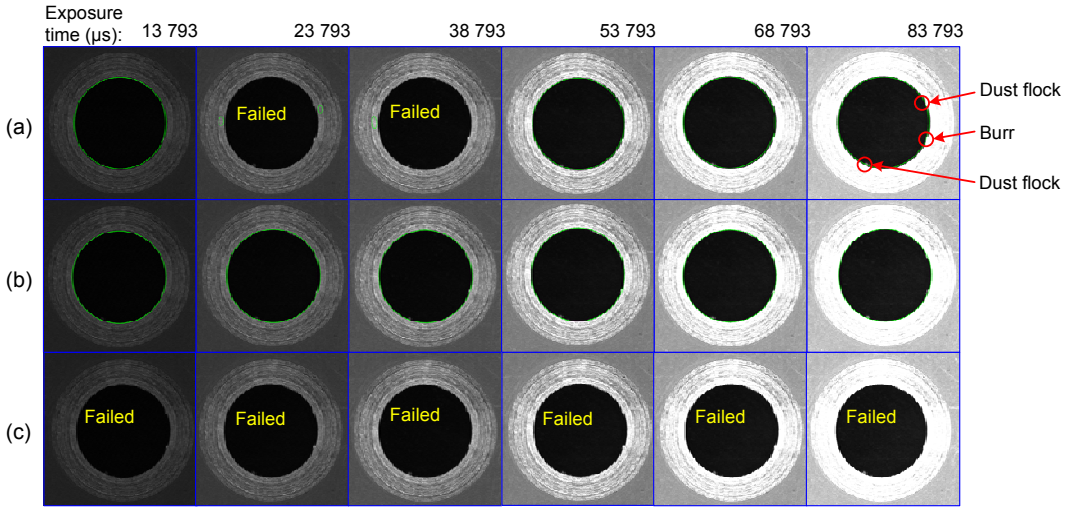

Fig. 9 Results of reference hole detection with various methods
(a) Least-squares method; (b) Saliency-snake method; (c) Hough transform method

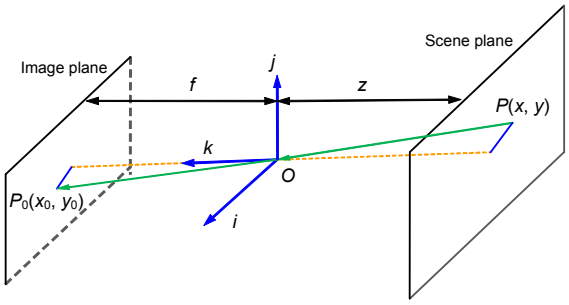

Fig. 10 Diagram of the weak perspective model

In Eq. (3), $m=f/z$, and f and z are the focal length and object distance, respectively.

Due to the improvement of lens manufacturing, skewness and lens distortion can be reasonably ignored for the camera with a small field of view (Zhan and Wang, 2012). The pixel of the camera is assumed to be rectangular instead of square, and two mapping coefficients k_x and k_y (mm/pixel) should be determined to calculate the physical distance from the captured images.

As discussed in Section 2.1, the camera TCP should be defined as an offset of the camera CS along its Z-axis by the object distance. If we first apply traditional hand-eye calibration methods to find the camera CS, and then calculate the camera TCP by offsetting the camera CS, the pose of the estimated camera TCP relative to the actual camera CS is

$$\begin{bmatrix} 1 & -\varepsilon_z & \varepsilon_y & \delta_x+z\varepsilon_y \\ \varepsilon_z & 1 & -\varepsilon_x & \delta_y-z\varepsilon_x \\ -\varepsilon_y & \varepsilon_x & 1 & z+\delta_z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

where δ_x, δ_y, δ_z, ε_x, ε_y, ε_z are pose errors of the estimated camera CS by calibration and z is the object distance of the camera (Fig. 11). Measurement errors of the vision system due to hand-eye calibration errors would be $(\delta_x+z\varepsilon_y,\ \delta_y-z\varepsilon_x)$, where $z\varepsilon_y$ and $-z\varepsilon_x$ are the Abbe errors which would be significant due to the orientation errors of the estimated camera CS with respect to the actual camera CS.

To eliminate the Abbe errors, the camera TCP is directly calibrated in this study, and the pose of the estimated camera TCP relative to the actual camera CS is

$$\begin{bmatrix} 1 & -\varepsilon_z' & \varepsilon_y' & \delta_x' \\ \varepsilon_z' & 1 & -\varepsilon_x' & \delta_y' \\ -\varepsilon_y' & \varepsilon_x' & 1 & \delta_z'+z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5}$$

where $\delta_x', \delta_y', \delta_z', \varepsilon_x', \varepsilon_y', \varepsilon_z'$ are pose errors of the estimated camera TCP determined in the direct calibration process (Fig. 12). Measurement errors of the vision system due to calibration errors are (δ_x', δ_y'), in which the Abbe errors are eliminated.

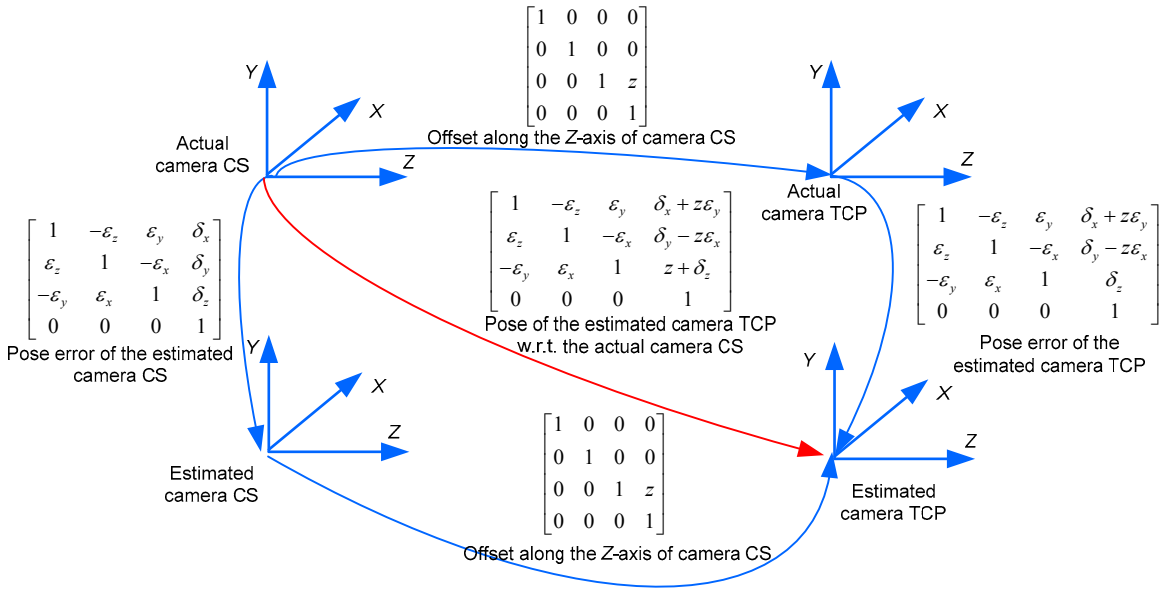

Fig. 11 Measurement errors introduced by traditional hand-eye calibration methods

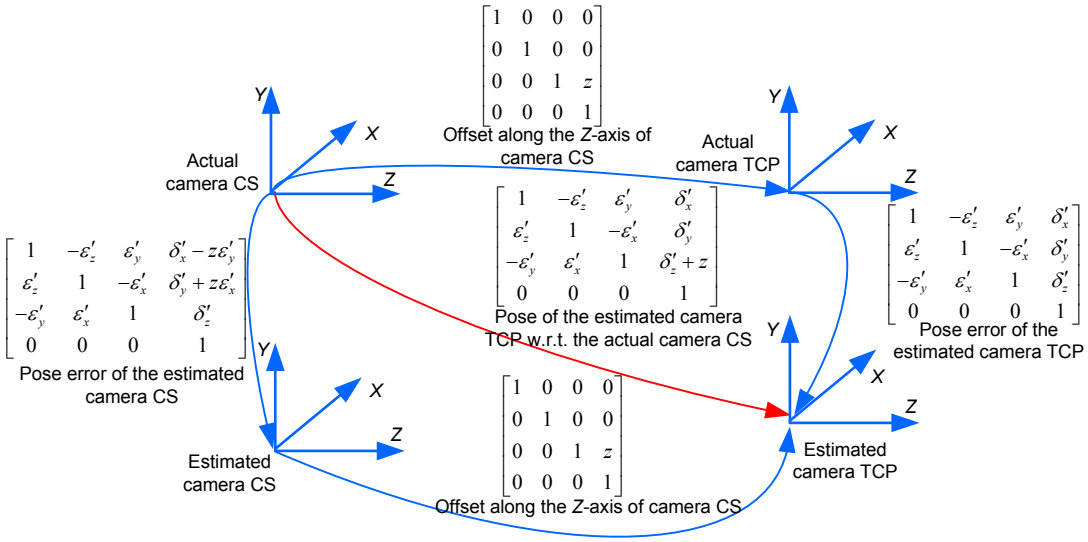

Fig. 12 Measurement errors introduced by the proposed method

For direct calibration of the camera TCP, we use a dedicated calibration board that has the necessary geometric structures for the computation of the camera internal parameters and hand-eye relationship between the drill and the camera. Fig. 13 shows the design of the calibration board. It has a mandril hole into which a cylindrical mandril is inserted. The cylindrical mandril is again connected to the spindle of the end-effector through the tool holder. To suppress the gravity-induced deformation of the calibration device, a mandril with the largest diameter (12 mm) for the tool holder is used. The array of small holes is

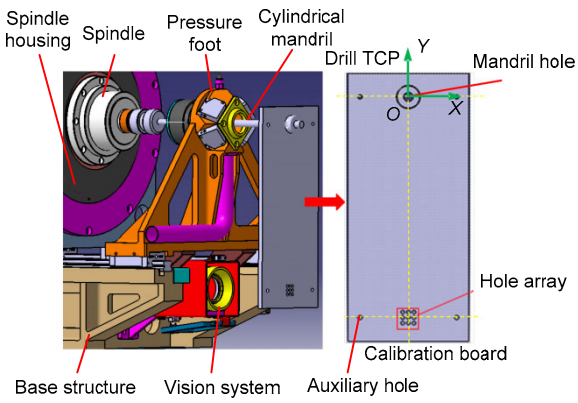

Fig. 13 Design of the calibration board

used to calibrate the internal parameters of the camera. The diameter of the holes is 5 mm and the spacing between the holes is 2 mm, which are designed according to the field of view 18° (visible area is 28 mm ×21 mm) of the camera. At the four corners of the calibration board, four auxiliary holes are designed for the mounting of an optical target when checking the positioning accuracy of the calibration board with a laser tracker during the experiments. The holes on the calibration board are manufactured to have accurate geometry, from which the hand-eye relationship between the drill and camera can be determined.

After removing the rotational degree of the spindle by fastening the spindle to its housing, the calibration board is fixed with respect to the end-effector base, and the drill TCP, intermediate coordinate system, and camera TCP are defined as shown in Fig. 14. The drill TCP is defined on the calibration board; its center is at the center of the mandril hole and its axis directions are defined by the holes on the calibration board. The intermediate coordinate system is also defined on the calibration board; its center is coincident with the center of the hole array and its axes are parallel with those of the drill TCP.

The optical axis of the camera is installed to be parallel with the drill axis. The rotational errors of the camera TCP around the X- and Y-axis of the drill TCP are not estimated in calibration but simply set to zero. This is justified because these errors are of second order to the measurement accuracy. The rotational error of the camera TCP around the Z-axis of the intermediate coordinate system is a first order error to the measurement accuracy and is calibrated in the process. The translational errors of the camera TCP along the X- and Y-axis of the intermediate coordinate system are calibrated by calculating the position of the center hole of the 3×3 hole array on the image plane. The camera TCP along the Z-axis is defined at the position where the captured image of the calibration board is the sharpest, which is ensured by controlling the position of the calibration board along the drill axis.

The XY planes of the drill TCP, the intermediate coordinate system, and the camera TCP are defined on the same plane of the calibration board (Fig. 14). The transformation from the drill TCP to the intermediate coordinate system is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -L \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6}$$

where L is the distance between the mandril hole and the center of the hole array.

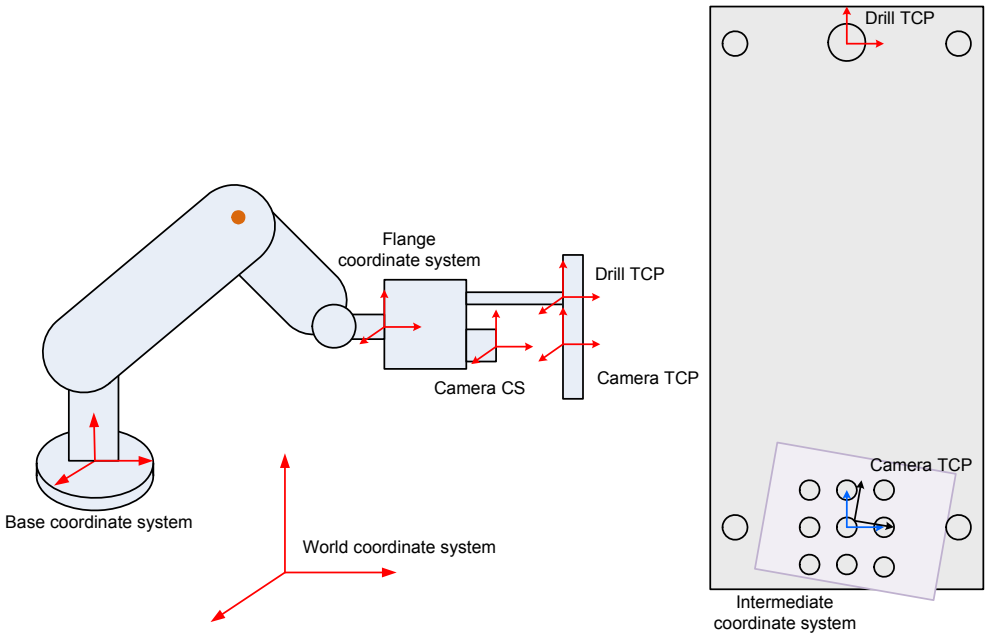

Fig. 14 Coordinate systems involved in camera calibration

The origin of the camera TCP is defined at the point whose corresponding image point is at the center of the image. The X-axis of the camera TCP is defined so that its corresponding image line is parallel to the U-axis of the image. Thus, the transformation from the camera TCP to the intermediate coordinate system is

$$\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 & x_0 \\ \sin\alpha & \cos\alpha & 0 & y_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{7}$$

where $[\cos\alpha, \sin\alpha, 0]^{\mathrm{T}}$ and $[-\sin\alpha, \cos\alpha, 0]^{\mathrm{T}}$ are direction vectors of the X- and Y-axis of the intermediate coordinate system relative to the camera TCP, respectively, and $(x_0, y_0, 0)$ is the origin of the intermediate coordinate system relative to the camera TCP.

By using matrix multiplication (Fig. 15), the hand-eye relationship between the camera TCP and the drill TCP can be calculated:

$$\begin{bmatrix} \cos\alpha & \sin\alpha & 0 & -(x_0\cos\alpha + y_0\sin\alpha) \\ -\sin\alpha & \cos\alpha & 0 & -L-(y_0\cos\alpha - x_0\sin\alpha) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{8}$$

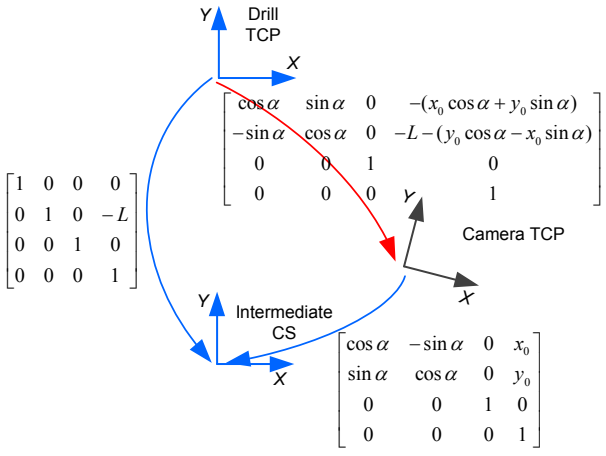

Fig. 15 Computation of the hand-eye relationship

According to the weak perspective model, these parameters can be calculated from the acquired images together with the mapping coefficients k_x and k_y. Suppose the centers of the 3×3 holes (Fig. 16) in the pixel coordinate system are (u_i, v_i), i=1, 2, …, 9, and their corresponding physical coordinates relative to the camera TCP are $(k_x u_i, k_y v_i)$, i=1, 2, …, 9. Then

$$x_0 = k_x u_5, \quad y_0 = k_y v_5, \tag{9}$$

$$\cos\alpha = \frac{k_x u_6 - k_x u_5}{\sqrt{(k_x u_6 - k_x u_5)^2 + (k_y v_6 - k_y v_5)^2}}, \tag{10}$$

$$\sin\alpha = \frac{k_y v_6 - k_y v_5}{\sqrt{(k_x u_6 - k_x u_5)^2 + (k_y v_6 - k_y v_5)^2}}. \tag{11}$$

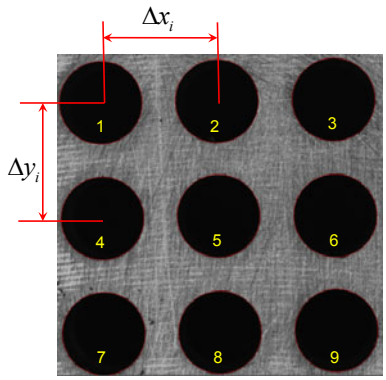

Fig. 16 Image of the 3×3 hole array of the calibration board and detected holes

The scale parameters k_x and k_y are determined from the captured image of the 3×3 hole array on the calibration board (Fig. 16). Based on the hole detection, the distance between the holes in the pixel coordinates can be obtained using the image processing algorithm. On the other hand, the physical distances between these holes are known beforehand. Since the lines passing through the hole centers are approximately parallel to the U- or V-axis on the images, the scale parameters of the camera can be calculated as

$$k_{x_i} = \frac{\Delta x_i}{\Delta u_i}, \quad k_{y_i} = \frac{\Delta y_i}{\Delta v_i}, \quad i = 1, 2, \cdots, 6, \tag{12}$$

$$k_x = \frac{1}{6}\sum_{i=1}^{6} k_{x_i}, \quad k_y = \frac{1}{6}\sum_{i=1}^{6} k_{y_i}, \tag{13}$$

where Δx_i and Δy_i are physical distances between adjacent holes in the X- and Y-direction, respectively, and Δu_i and Δv_i are pixel distances between adjacent holes in the U- and V-direction, respectively.

5 Experiments

Experiments have been conducted to verify the feasibility of the proposed hand-eye calibration method and evaluate the measurement accuracy of the developed monocular vision system. The experimental platform (Fig. 17) includes a KUKA industrial robot, a drilling end-effector developed by Zhejiang University (ZJU), a Leica laser tracker, a monocular vision system including a Baumer industrial camera and an annular LED light source (installed in a protection box), a dedicated and optimized high-accuracy calibration board (manufactured by CNC), an experimental workpiece, a ball-shaped drill, etc.

In the experiments, the camera internal parameters and hand-eye relationship between the drill TCP and the camera TCP were calibrated simultaneously. The position coordinates of the reference hole in the robot base frame were determined by the monocular vision system. Based on the measurement, the robot was driven to align the axis of the ball-shaped drill with the center line of the reference hole. The distance between the center of the optical target nested in the reference hole and the axis of the ball-shaped drill was taken as the measurement error of the monocular vision system. A concise flowchart of the experiments is shown in Fig. 18, and the detailed procedures are described as follows:

Step 1: Mount the calibration board on the drilling end-effector, move the spindle unit along the drilling direction so that the calibration board is in-focus, and shoot the calibration board.

Step 2: Calibrate the camera internal parameters and hand-eye relationship between the drill TCP and the camera TCP, and save the calibration results.

Step 3: Remove the calibration board mounted in step 1 and clamp the shank of the ball-shaped drill to the spindle of the drilling end-effector (Fig. 19).

Step 4: Move the monocular vision system to the nominal position of a reference hole by the robot for vision-based measurement. Then adjust the brightness of the annular light source, the shooting parameters of the industrial camera, and the object distance to achieve a good shooting condition.

Step 5: Execute the feature search program to ensure that the reference hole is in the shooting area, conduct a vision-based measurement program, and send the measurement results to the control software of the robotic drilling system.

Step 6: Based on the calibration results obtained in step 1 and measurement results obtained in step 5, the robot is moved to align the axis of the ball-shaped drill with the center line of the reference hole on the workpiece (Fig. 17).

Step 7: Measure the position of the optical target nested in the reference hole using the laser tracker,

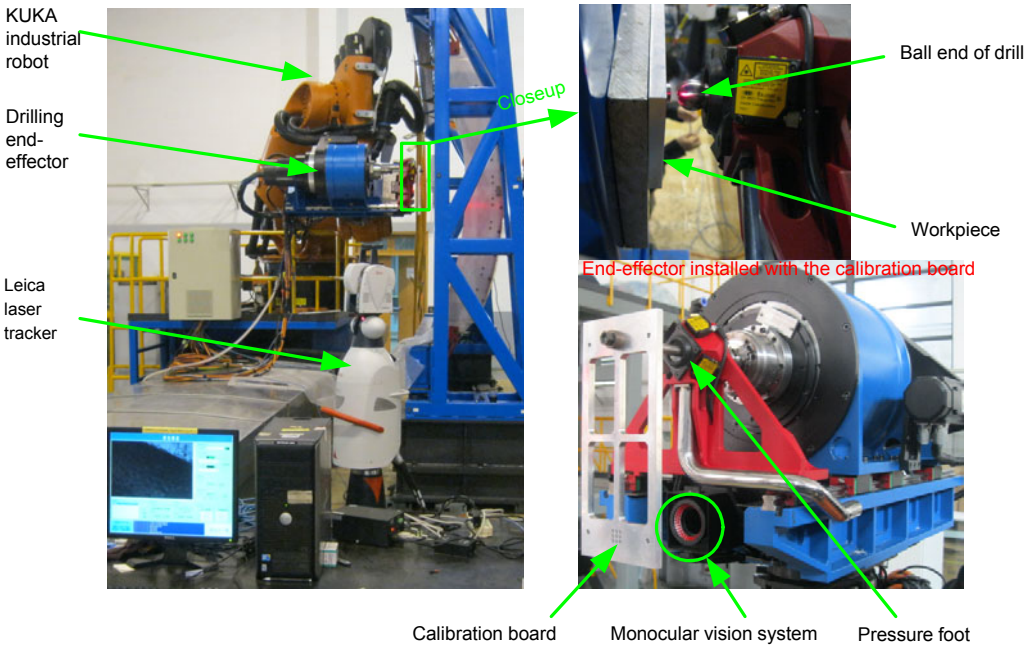

Fig. 17 Experimental platform for the monocular vision system for robotic drilling

and then fit a line to the axis of the ball-shaped drill with the data points obtained using the laser tracker.

Step 8: Compute the distance between the position of the optical target and the line of the ball-shaped drill, and take this distance as the measurement error of the vision system.

The center coordinates of the optical target, the line equations of the ball-shaped drill axis, and the resultant distances are shown in Table 1. The positioning accuracy of the vision system is higher than 0.15 mm, which suggests that the errors due to elliptical contour extraction and system calibration are reasonably small. It can be concluded that the developed monocular vision system is effective and can be used in robotic drilling for the assembly of large aircrafts, which typically requires an accuracy of ±0.5 mm for the position of fastener holes.

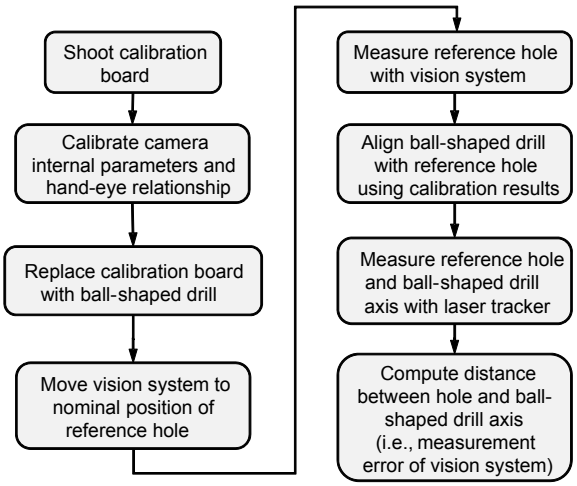

Fig. 18 Flowchart of the experiment for the monocular vision system

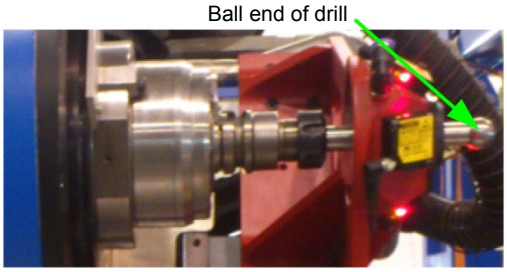

Fig. 19 Installation of the ball-shaped drill

6 Conclusions

In this research, we have developed a practical monocular vision system for measurement of the relative deviation between the drill TCP and the reference hole in robotic drilling. The principle of relative error measurement with the vision system is presented, along with a detailed discussion on the choice of the appropriate camera TCP for the vision system. The hardware components, software components, and the integration of the whole vision system are described. For accurate and robust reference hole detection, a novel saliency-snake algorithm for elliptical contour extraction, which consists of salient region detection, voting, and snake-based contour optimization, has been proposed and implemented in measurement software. For efficient and accurate calibration of the vision system, a new calibration method based on a high-accuracy dedicated calibration board has been developed for the monocular vision system. With this method, camera internal parameters and hand-eye relationship can be calibrated simultaneously, and calibration accuracy is ensured by eliminating the requirements on significant perspective distortion of captured images and constraint equations of robot motion. Experiments performed on a robotic drilling system show that the developed monocular vision system can achieve a measurement accuracy of 0.15 mm, which meets the requirement of current practice in aerospace industry.

Table 1 Experimental data of reference hole measurement

Hole No.	Center coordinates of the optical target	Parameters of the ball-shaped drill axis*	Distance (mm)
1	(−2366.422, 1413.671, 971.788)	(12.699, 23.868, 0.039)/(−2377.472, 1393.026, 971.694)	0.084
2	(−2432.467, 1446.628, 1038.065)	(14.085, 26.484, 0.034)/(−2448.287, 1417.029, 1038.061)	0.077
3	(−2299.542, 1379.751, 970.621)	(12.694, 23.855, 0.047)/(−2308.587, 1362.864, 970.648)	0.079
4	(−2365.810, 1412.536, 1036.981)	(12.691, 23.858, 0.025)/(−2376.840, 1391.779, 1037.067)	0.108
5	(−2433.221, 1447.462, 973.104)	(12.703, 23.884, 0.036)/(−2446.717, 1422.011, 972.957)	0.115
6	(−2499.462, 1480.494, 1039.459)	(12.679, 23.819, 0.040)/(−2515.722, 1449.945, 1039.548)	0.141

* Direction vector of the drill axis/one point on the axis

References

Armingol, J.M., Otamendi, J., de la Escalera, A., *et al.*, 2003. Statistical pattern modeling in vision-based quality control systems. *J. Intell. Robot. Syst.*, **37**(3):321-336. [doi:10.1023/A:1025489610281]

Bilen, H., Hocaoglu, M., Unel, M., *et al.*, 2012. Developing robust vision modules for microsystems applications. *Mach. Vis. Appl.*, **23**(1):25-42. [doi:10.1007/s00138-010-0267-y]

Bone, G.M., Capson, D., 2003. Vision-guided fixtureless assembly of automotive components. *Robot. Comput.-Integr. Manuf.*, **19**(1-2):79-87. [doi:10.1016/S0736-5845(02)00064-9]

Bradski, G., Kaehler, A., 2008. Learning OpenCV: Computer Vision with the OpenCV Library. O'reilly, Sebastopol, CA.

Chen, S., Zhang, Y., Qiu, T., *et al.*, 2003. Robotic welding systems with vision-sensing and self-learning neuron control of arc welding dynamic process. *J. Intell. Robot. Syst.*, **36**(2):191-208. [doi:10.1023/A:1022652706683]

DeVlieg, R., Sitton, K., Feikert, E., *et al.*, 2002. ONCE (ONe-sided Cell End effector) Robotic Drilling System. SAE Technical Paper 2002-01-2626. [doi:10.4271/2002-01-2626]

Dornaika, F., Horaud, R., 1998. Simultaneous robot-world and hand-eye calibration. *IEEE Trans. Robot. Autom.*, **14**(4):617-622. [doi:10.1109/70.704233]

Eberli, D., Scaramuzza, D., Weiss, S., *et al.*, 2010. Vision based position control for MAVs using one single circular landmark. *J. Intell. Robot. Syst.*, **61**(1-4):495-512. [doi:10.1007/s10846-010-9494-8]

Fitzgibbon, A., Pilu, M., Fisher, R.B., 1999. Direct least square fitting of ellipses. *IEEE Trans. Pattern Anal. Mach. Intell.*, **21**(5):476-480. [doi:10.1109/34.765658]

Forsyth, D.A., Ponce, J., 2011. Computer Vision: a Modern Approach (2nd Ed.). Prentice-Hall, Englewood Cliffs, New Jersey.

Ibarguren, A., Martínez-Otzeta, J.M., Maurtua, I., 2014. Particle filtering for industrial 6DOF visual servoing. *J. Intell. Robot. Syst.*, **74**(3-4):689-696. [doi:10.1007/s10846-013-9854-2]

Lee, B., Tarng, Y., 2001. Surface roughness inspection by computer vision in turning operations. *Int. J. Mach. Tool. Manuf.*, **41**(9):1251-1263. [doi:10.1016/S0890-6955(01)00023-2]

Malassiotis, S., Strintzis, M.G., 2003. Stereo vision system for precision dimensional inspection of 3D holes. *Mach. Vis. Appl.*, **15**(2):101-113. [doi:10.1007/s00138-003-0132-3]

Malti, A., 2012. Hand–eye calibration with epipolar constraints: application to endoscopy. *Robot. Auton. Syst.*, **61**(2):161-169. [doi:10.1016/j.robot.2012.09.029]

Motta, J.M.C.S., de Carvalho, G.C., McMaster, R., 2001. Robot calibration using a 3D vision-based measurement system with a single camera. *Robot. Comput.-Integr. Manuf.*, **17**(6):487-497. [doi:10.1016/S0736-5845(01)00024-2]

Neto, H.V., Nehmzow, U., 2007. Real-time automated visual inspection using mobile robots. *J. Intell. Robot. Syst.*, **49**(3):293-307. [doi:10.1007/s10846-007-9146-9]

Olsson, T., Haage, M., Kihlman, H., *et al.*, 2010. Cost-efficient drilling using industrial robots with high-bandwidth force feedback. *Robot. Comput.-Integr. Manuf.*, **26**(1):24-38. [doi:10.1016/j.rcim.2009.01.002]

Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.*, **9**(1):62-66. [doi:10.1109/TSMC.1979.4310076]

Pachidis, T.P., Lygouras, J.N., 2007. Vision-based path generation method for a robot-based arc welding system. *J. Intell. Robot. Syst.*, **48**(3):307-331. [doi:10.1007/s10846-006-9076-y]

Simonvsky, M., 2011. Ellipse Detection Using 1D Hough Transform. Available from http://www.mathworks.com/matlabcentral/fileexchange/33970-ellipse-detection-using-1d-hough-transform.

Strobl, K.H., Hirzinger, G., 2008. More accurate camera and hand-eye calibrations with unknown grid pattern dimensions. IEEE Int. Conf. on Robotics and Automation, p.1398-1405. [doi:10.1109/ROBOT.2008.4543398]

Strobl, K.H., Sepp, W., Hirzinger, G., 2009. On the issue of camera calibration with narrow angular field of view. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, p.309-315. [doi:10.1109/IROS.2009.5354776]

Suzuki, S., 1985. Topological structural analysis of digitized binary images by border following. *Comput. Vis. Graph. Image Process.*, **30**(1):32-46. [doi:10.1016/0734-189X(85)90016-7]

Thompson, P., Hartmann, J., Feikert, E., *et al.*, 2005. Flex Track for Use in Production. SAE Technical Paper 2005-01-3318. [doi:10.4271/2005-01-3318]

Tsai, R., 1987. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE J. Robot. Autom.*, **3**(4):323-344. [doi:10.1109/JRA.1987.1087109]

Webb, P., Eastwood, S., Jayaweera, N., *et al.*, 2005. Automated aerostructure assembly. *Ind. Robot.*, **32**(5):383-387. [doi:10.1108/01439910510614646]

Williams, D.J., Shah, M., 1992. A fast algorithm for active contours and curvature estimation. *CVGIP Image Understand.*, **55**(1):14-26. [doi:10.1016/1049-9660(92)90003-L]

Zhan, Q., Wang, X., 2012. Hand–eye calibration and positioning for a robot drilling system. *Int. J. Adv. Manuf. Technol.*, **61**(5-8):691-701. [doi:10.1007/s00170-011-3741-4]

Zhang, Z., 2000. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22**(11):1330-1334. [doi:10.1109/34.888718]

Zhao, Z., Weng, Y., 2013. A flexible method combining camera calibration and hand–eye calibration. *Robotica*, **31**(5):747-756. [doi:10.1017/S0263574713000040]

Zhou, Y., Nelson, B.J., Vikramaditya, B., 2000. Integrating optical force sensing with visual servoing for microassembly. *J. Intell. Robot. Syst.*, **28**(3):259-276. [doi:10.1023/A:1008136711577]

Zhu, W., Qu, W., Cao, L., *et al.*, 2013. An off-line programming system for robotic drilling in aerospace manufacturing. *Int. J. Adv. Manuf. Technol.*, **68**(9-12):2535-2545. [doi:10.1007/s00170-013-4873-5]

Zou, X., Zou, H., Lu, J., 2012. Virtual manipulator-based binocular stereo vision positioning system and errors modelling. *Mach. Vis. Appl.*, **23**(1):43-63. [doi:10.1007/s00138-010-0291-y]