Metrol. Meas. Syst., Vol. 27 (2020) No. 3, pp. 495–512
DOI: 10.24425/mms.2020.134593

METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.wat.edu.pl

DESIGN AND EVALUATION OF A NEEDLE TIP MEASUREMENT SYSTEM BASED ON BINOCULAR VISION

Yuyang Lin[1], Yunlai Shi[1], Jun Zhang[2], Fugang Wang[1], Haichao Sun[1,3]

1) *Nanjing University of Aeronautics and Astronautics, State Key Laboratory of Mechanics and Control of Mechanical Structures, Nanjing 210016, China* (*lamyuyeung@163.com*, ✉ *shiyunlai950438@nuaa.edu.cn*, +86 25 8489 1812, *wolfgang-d@foxmail.com*, *shc@nuaa.edu.cn*)
2) *Northwest A&F University, College of Mechanical and Electronic Engineering, Yangling 712100, China,* (*zhangjun871113@163.com*)
3) *Pingdingshan University, Institute of Electrical and Mechanical Engineering, Pingdingshan 467000, China*

Abstract

The precise location of the needle tip is critical in robot-assisted needle-based percutaneous interventions. An automatic needle tip measuring system based on binocular vision technology with the advantages of non-contact, excellent accuracy and high stability is designed and evaluated. First the measurement requirements of the prostate intervention robot are introduced. A laser interferometer is used as the reference for measuring the position of the needle tip whose relative position variation is described as the needle tip distance in the time domain. The parameters of the binocular cameras are obtained by Zhang's calibration method. Then a robust needle tip extraction algorithm is specially designed to detect the pixel coordinates of the needle tip without installing the marked points. Once the binocular cameras have completed the stereo matching, the 3D coordinates of the needle tip are estimated. The measurement capability analysis (MCA) is used to evaluate the performance of the proposed system. The accuracy of the system can be controlled within 0.3621 mm. The agreement analysis is conducted by the Bland–Altman analysis, and the Pearson correlation coefficient is 0.999847. The P/T ratio value is 16.42% in the repeatability analysis. The results indicate that the accuracy and stability of the binocular vision needle tip measuring system are adequate to meet the requirement for the needle tip measurement in percutaneous interventions.

Keywords: needle tip measurement, binocular vision system, measurement capability analysis, non-contact optical positioning.

© 2020 Polish Academy of Sciences. All rights reserved

1. Introduction

Prostate cancer is one of the most frequently diagnosed cancer among men in the world and ranks second most frequent cancer with 1.3 million new cases and the fifth leading cause of cancer death in men according to the status report on the global cancer statistics in 2018 [1]. Robot-assisted percutaneous prostate intervention under *Magnetic Resonance Image* (MRI)-guidance is

Copyright © 2020. The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (CC BY-NC-ND 4.0 https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits use, distribution, and reproduction in any medium, provided that the article is properly cited, the use is non-commercial, and no modifications or adaptations are made.

becoming a promising therapy method because of its advantages such as safety, accuracy, and efficiency [2, 3]. The MRI-guided robot-assisted prostate intervention can be divided into two phases *i.e.* needle placement and needle penetration. The accuracy in the placement phase is critical for the efficiency of the penetration phase. However, the needle tip cannot be detected by the MRI before the needle is inserted into the prostate. Therefore, a measurement system has to be established to inspect the position of the needle tip during the placement phase.

Stereo vision, as a typical measurement technology, is widely used in many areas, such as geometry measuring, robot path correction, and quality inspection [4–13]. Xiang *et al.* [14] used the binocular vision technology and the relative measurement principle to determine the bayonet sizes of large automobile brake pads. The tests of repeatability, bias, and real-time performance were conducted to validate that the proposed system met the requirements of online measurement. Li *et al.* [15] proposed an automatic in-line ball height and coplanarity inspection method based on stereo vision technology to detect the potential connectivity in semi-conductor units. The measurement results were compared with those from the laser-scanning and confocal inspection tools. The *measurement capability analysis* (MCA) procedure was applied and showed high accuracy, repeatability, and reproducibility of the inspection method.

The stereo vision can be used in the inspection of medical devices. Lin *et al.* [16] proposed a near-infrared optical tracking system to measure the position of surgical tools. The position of a surgical tool was marked by near-infrared light-emitting diode chips installed on it. The accuracy of the position of the surgical tool regarding light-emitting points was evaluated and the mean square errors were less than 0.31 mm. Jiang *et al.* [17] used the binocular vision measurement principle to investigate medical punctures. The targets were marked with circular points which were transformed into elliptic points in the actual image process and a target extraction algorithm was proposed. The experiment results showed that the deviation of the 3 axes was about 2 mm, 2 mm, and 3 mm. Zhou *et al.* [18] designed a stereo vision tracking system to navigate the surgical instruments. Optical infrared LEDs mounted on an surgical instrument were used to mark its position. An algorithm was designed to extract the marker's pixel coordinates. An accuracy simulation was conducted and the results showed that the system satisfies the requirements of surgical navigation.

The common method used in the navigation of a surgical instrument is installing marked points which can be easily identified by cameras. Spatial coordinates of the marked points are extracted with a specifically designed algorithm. However, most of the marked points are difficult to mount on the tip of a standard insertion needle and some electrical marked points such as LEDs or light-emitting chips are strictly forbidden in the MRI environment for the MR safeuse. In addition, the spatial coordinates extraction from some marked points would induce superfluous detection errors. Thus, a needle tip position detection method specifically for the insertion needle without any marked points is required. Most of the non-contact needle tip position detection methods in the robot-assisted prostate intervention are medical imaging technologies such as *ultrasound* (US), *computed tomography* (CT) and *magnetic resonance image* (MRI). These technologies are expensive and lack universality since they are only used in hospitals with special requirements regarding the environment and needle insertion robot. For example, some needle insertion robots with ferromagnetic material are forbidden in the MRI environment. The stereo vision method is an alternative non-contact method of precise detection of the position of the needle tip. Compared with the medical imaging technologies, the stereo vision method with its low cost, real-time detection and high accuracy features can be easily used in laboratories in the development of the percutaneous needle insertion robot as well as and satisfy the measuring accuracy requirements of the needle tip positioning. Therefore, based on the stereo vision technology, a needle tip measurement system used for the prostate intervention robot is proposed. Accuracy is the most

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 495–512
DOI: 10.24425/mms.2020.134593

common indicator for evaluating the performance of the majority of the stereo vision system. There are few researches on the stability and agreement of the stereo vision system. Inspired by [14] and [15], an MCA having the accuracy, agreement, and repeatability is proposed as a means to evaluate the position of the needle tip.

In this paper, an accurate binocular vision system used for inspecting the needle tip position of a prostate intervention robot is designed and evaluated. The rest of the paper is organized as follows. The prostate intervention robot and the system establishment are introduced briefly in Section 2. The design details of the binocular vision system are presented in Section 3. The evaluation of the proposed system by using the MCA is described in Section 4. The conclusions are given in Section 5.

2. System overview

2.1. The prostate intervention robot

The proposed binocular vision system is used for needle tip detection in a prostate intervention robot. The expected accuracy of the robot determined by the size of the prostate cancer foci is essential for the design of the binocular vision system. As a classical progressive disease, the malignant potential of prostate cancer tends to increase with tumor volume. Some researchers reported that prostate cancer tumor volume and Gleason Grades are two vital factors used to predict cancer progression. It is not clinically significant when tumor volume is not more than 0.5 cc and Gleason Grades is under 4 [19, 20]. The prostate cancer tumors are usually assumed to be spherical and the 0.5 cc tumor volume corresponds to a sphere with a radius of approximately 4.92 mm. According to [21–24], the placement accuracy of the most precise percutaneous needle insertion robot is about 2–3 mm.

The photography of the prostate intervention robot which is supposed to be evaluated with the proposed binocular vision system is shown in Fig. 1. The needle mounted on the robot is a standard 18-gauge titanium needle, 150.0 mm in length and 1.67 mm in diameter, and specifically used for prostate intervention. The robot is fully driven by ultrasonic motors with a theoretical linear displacement resolution of 0.488 μm in 3 orthogonal directions. The designed strokes of the robot for the 3 directions are 28 mm, 22 mm, and 15 mm respectively. The proposed binocular vision system should be adequate to measure the needle tip position of the robot.

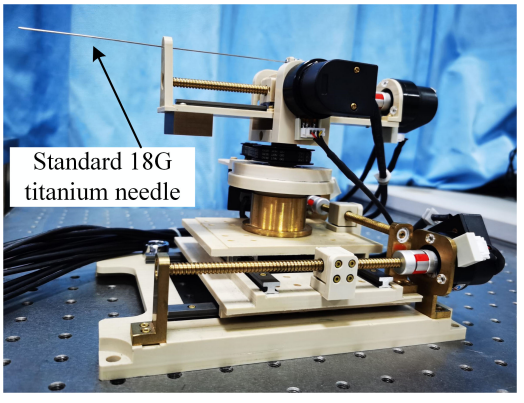


Fig. 1. The photography of the prostate intervention robot.

2.2. System establishment

The binocular vision system includes a standalone computer, two CMOS cameras (U-300C, MindVision) with a resolution of 2048 × 1436 pixels and a pixel size of 3.2 μm. Each camera is equipped with a 6 mm f/1.4 lens. The cameras are mounted on aluminum support, and the distance between the two cameras is about 72 mm. The background is white and the fluorescent lamp, common in many venues including laboratories and MRI scanner rooms, is used as the lighting condition. The titanium needle may reflect the light ray which can affects the gray value of the pixel in the needle image. The influence can be eliminated by adjusting the relative positions of the needle and the fluorescent lamp and modifying the binarization threshold during the needle tip extraction process.

To evaluate the performance of the binocular vision system, an experimental system was constructed as shown in Fig. 2a. The experimental system includes the binocular vision system, the needle insertion robot, and a laser interferometer. The needle tip is in the field of view of the two cameras in the design stroke of the robot. The needle images acquired by the two cameras are illustrated in Fig. 2c and Fig. 2d. The measurement capability of the binocular vision system was evaluated by a laser interferometer system (Renishaw XL-80) with a linear measurement accuracy of ±0.5 μm/m and a resolution of 1 nm. In the laser interferometer system, the fixed-length reference arm is formed by the beam-splitter and the first retro-reflector while the second retro-reflector is mounted on the robot parallel to the insertion direction of the needle.

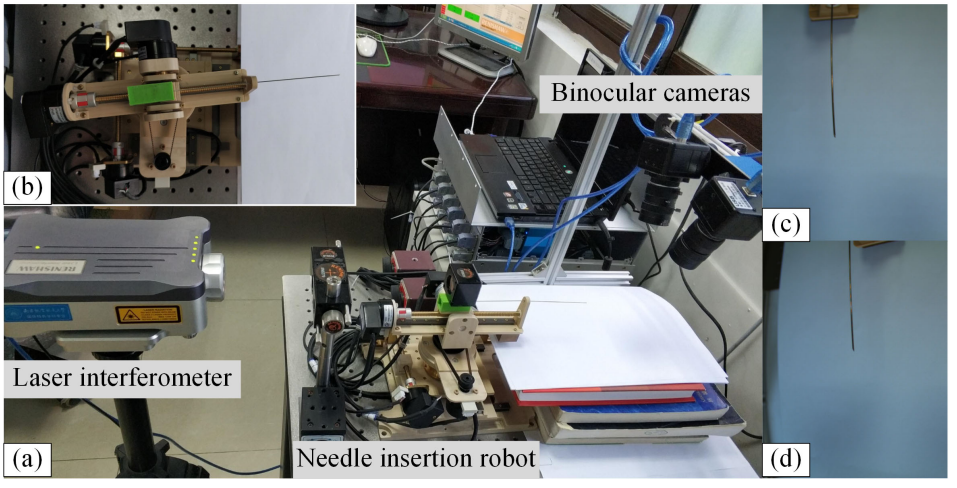


Fig. 2. The measurement capability evaluation system. a) Overall experiment setup. b) Top view of the needle insertion robot. c) Image of the needle in the left camera. d) Image of the needle in the right camera.

The spatial position of the needle tip sampled by the binocular vision system is denoted as P_{i-1} $(x_{i-1}, y_{i-1}, z_{i-1})$, while the position obtained by the laser interferometer is l_{i-1}. After the needle was driven for a certain distance in the insertion direction, the position of the needle tip was sampled by the binocular vision system and the laser interferometer again and represented as P_i (x_i, y_i, z_i) and l_i respectively. The Euclidean distance of the needle tip (retro-reflector) obtained by the binocular vision system and the laser interferometer is represented as d_{vision} and

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 495–512
DOI: 10.24425/mms.2020.134593

d_{laser} given by:

$$d_{vision} = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2}, \tag{1}$$

$$d_{laser} = |l_i - l_{i-1}| . \tag{2}$$

The distances describe the relative position variation of the needle tip in the time domain. The distance obtained with the laser interferometer is used as a reference for evaluating the accuracy of the binocular vision system, the error of the binocular vision system is calculated by

$$\Delta = d_{vision} - d_{laser} . \tag{3}$$

3. System design

3.1. Pinhole camera model and calibration

The pinhole camera model which describes the coordinate relationship of a physical point and the corresponding projection onto the image plane is depicted in Fig. 3. The pinhole camera model includes the *pixel coordinate system* [PCS, (u, v)], the *image plane coordinate system* [IPCS, (x, y)], the *camera coordinate system* [CCS, (x_c, y_c, z_c)] and the *world coordinate system* [WCS, (x_w, y_w, z_w)].

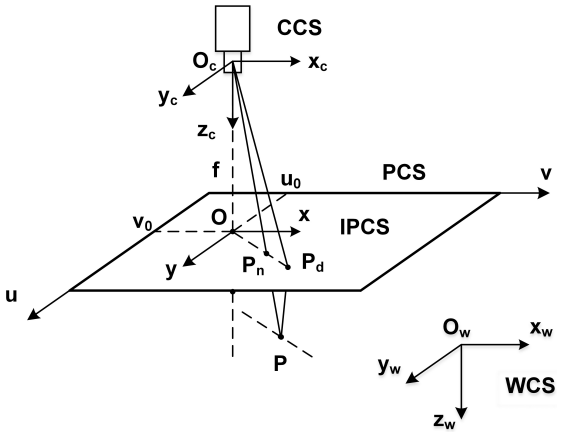


Fig. 3. Pinhole camera model.

The camera coordinate system is located in the optical center O_c, in which z_c and the image plane are assumed to be mutually perpendicular and x_c is parallel to the image horizontal axis. The 3D rigid body transformation is adopted to change the world coordinates of a physical point into camera coordinates as follows:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T \tag{4}$$

in which the orientation of the world coordinate system is turned into that of the camera coordinate system by the 3×3 rotation matrix R, while the offset of the world coordinate origin and the optical center O_c is denoted by the 3×1 translation vector T.

When the lens distortions are not considered, the physical point P with camera coordinates is mapped to the image plane linearly along a line starting at the optical center O_c to the point P, and the intersection of the line with the image plane is a 2D point P_n on the image plane. The geometrical relationship of the similar triangles is utilizing to represent the transformational relation of the camera coordinate and the corresponding image plane coordinate as follows:

$$x = f\frac{x_c}{z_c}, \qquad y = f\frac{y_c}{z_c}, \qquad z = f, \tag{5}$$

the relationship written in the matrix form is given by:

$$z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}, \tag{6}$$

where f denotes the focal length of the camera which indicates the distance from the optical center O_c to the image plane.

The transformational relation of the pixel coordinates and the image plane coordinates is given by:

$$\begin{aligned} u &= \frac{x}{dx} + u_0 \\ v &= \frac{y}{dy} + v_0 \end{aligned} \tag{7}$$

and the matrix form is expressed as:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 00 & 1 & \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{8}$$

where dx, dy represents the pixel size on the image plane in the orientation of x and y, while the coordinate of the principal point which represents the origin of the image plane coordinate system is denoted by (u_0, v_0) with respect to the left bottom corner of the image plane.

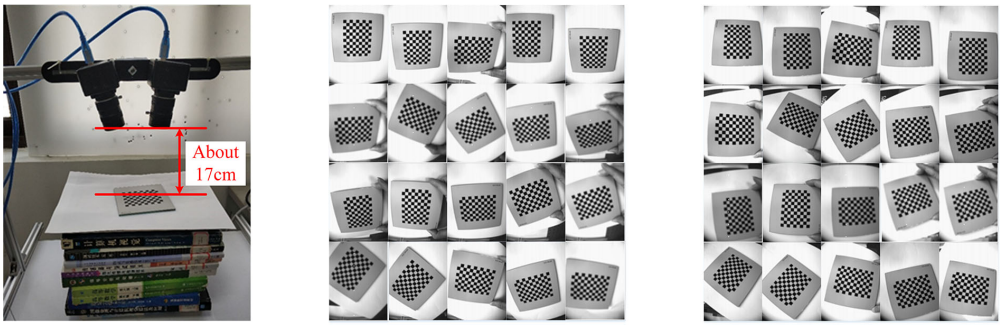
The overall projection equation describing the conversion relationship between the world coordinates and the corresponding pixel coordinates is given by:

$$\begin{aligned} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= \frac{1}{z_c} \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} \\ &= I_M \begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix}, \end{aligned} \tag{9}$$

where I_M denotes the camera intrinsic parameters, and the structural parameters of the camera are represented as R, T.

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 495–512
DOI: 10.24425/mms.2020.134593

Zhang's calibration method [25] was adopted to conduct the stereo calibration of the binocular vision system. As shown in Fig. 4, a standard chessboard placed on a white background plane has 9×12 divisions, each of which is a 6 mm $\times$ 6 mm square. The distance between the lens and the chessboard is about 17 cm which is close to the distance between the lens and the needle mounted on the robot. Twenty sets of the chessboard images with different angles and positions were sampled by the two cameras. The binocular vision system is calibrated and the results are listed in Table 1, where I_l, I_r are intrinsic matrices of the two cameras, R, and T are regarded as the structure parameters describing the relative position and posture of the two cameras.



(a) (b) (c)

Fig. 4. Calibration. (a) calibration setup. (b) Left camera calibration images. (c) Right camera calibration images.

Table 1. Camera calibration results.

Calibration parameter	Value
Left camera intrinsic matrix I_l	$\begin{bmatrix} 1909.84 & 0 & 932.97 \\ 0 & 1910.22 & 657.90 \\ 0 & 0 & 1 \end{bmatrix}$
Right camera intrinsic matrix I_r	$\begin{bmatrix} 1923.38 & 0 & 935.88 \\ 0 & 1914.32 & 729.04 \\ 0 & 0 & 1 \end{bmatrix}$
Rotation structure matrix R	$\begin{bmatrix} 0.862 & -0.020 & -0.507 \\ 0.036 & 0.999 & 0.022 \\ 0.506 & -0.037 & 0.861 \end{bmatrix}$
Translation structure matrix T	$\begin{bmatrix} 72.047 & 2.138 & 14.041 \end{bmatrix}^T$

There are 88 corner points in a single chessboard image. The pixel coordinates of the corner points on the chessboard can be calculated using the calibration parameters. The comparison of the calculated coordinates and the pixel coordinates of the corner points extracted from the images captured by the cameras are accomplished. The reprojection errors of the coordinates described in Fig. 5 indicates that the reprojection errors of the two cameras are mostly within the range of ±0.5 pixels and that the mean reprojection errors of both cameras are 0.052 pixels and 0.055 pixels which proves adequate accuracy of the calibration.

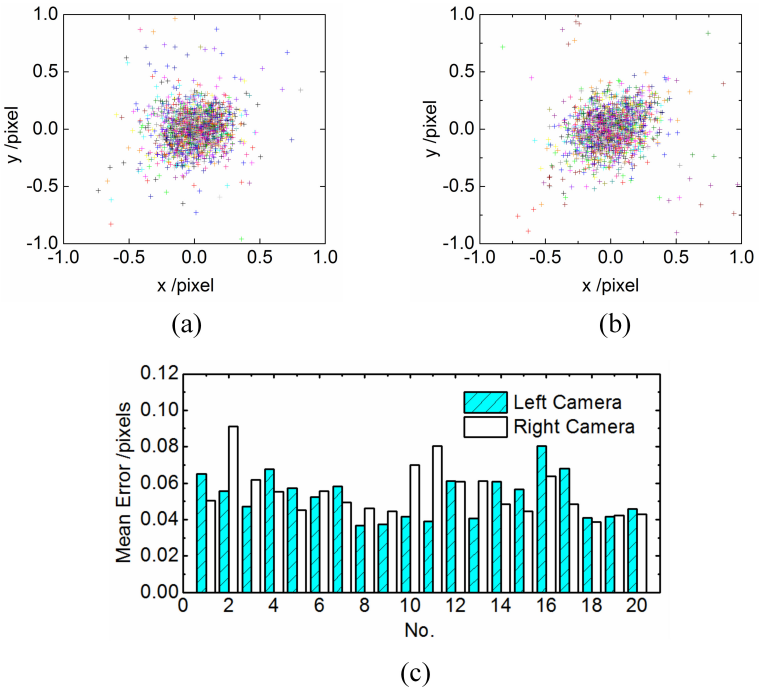


Fig. 5. Reprojection errors. a) Left camera. b) Right camera. c) Mean errors in 20 group chessboard sets.

3.2. Image processing and needle tip extraction

It is necessary to process the images to extract the pixel coordinates of the needle tip automatically and accurately. Figure 6 shows a scheme for image processing and pixel coordinates extraction for the needle tip.

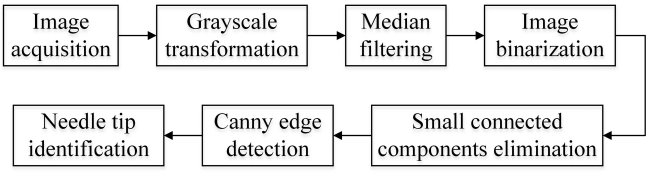


Fig. 6. Needle tip pixel coordinates extraction.

At first the needle was placed between the white background plane and the lens. The color needle image on the white plane with random background noises was collected by the cameras as shown in Fig. 7a. By means of the weighted average method which simulates the sensitivity of the human eye to different colors, the color needle image was converted into the grayscale image as described in Fig. 7b to improve the contrast of the image and speed up the calculation.

The original image contains a mass of random noise generated during the image acquisition. The undesirable noise can easily induce the geometric distortion in the image and directly affect the accuracy of pixel coordinate extraction of the needle tip. Median filtering is applied to blur the image and suppress the noise. According to the Median filtering, the gray values of pixels

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 495–512
DOI: 10.24425/mms.2020.134593

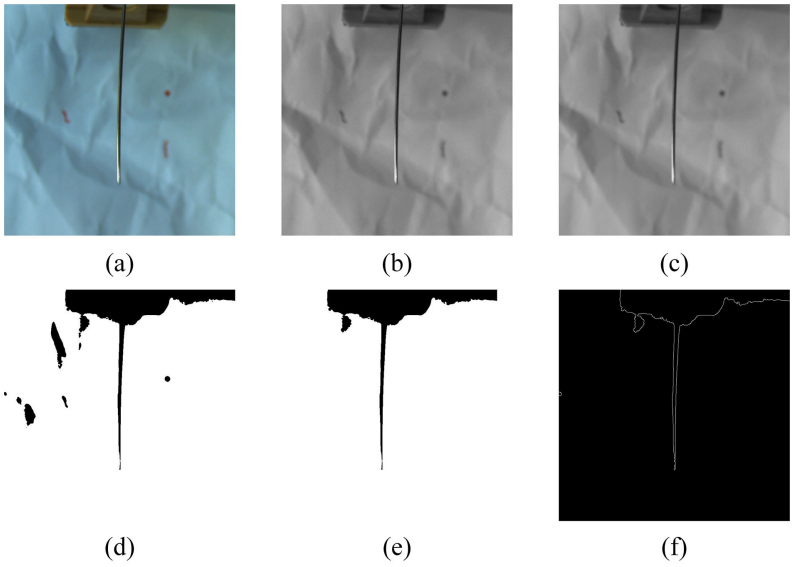


Fig. 7. Image processing results. a) Original image. b) Grayscale image. c) Median filtered image. d) Binary image. e) Small connected components elimination image. f) Needle feature edge image.

inside a filter window were replaced with the median of gray values of the pixels to eliminate isolated noise points. The median filtering is given by:

$$g(x, y) = median\{f(x - i,\ y - j)\}, \qquad (i, j) \in W, \tag{10}$$

where W is the filter window, $g(x, y)$, $f(x, y)$ denote the gray value of the pixels in the filter window. Figure 7c shows the image after the median filtering. Binarization of the image was performed to distinguish the extracted needle and its background according to the differences in grayscale characteristics. Base on a threshold T, the grayscale image is transformed into a binary image of only two colours (gray values 0 and 255). The binarization process converts the pixel grayscale in the image into 255 when it is greater than or equal to the threshold T, while transforming the pixel grayscale under the threshold T into 0. The binarization operation is given by:

$$f(x, y) = \begin{cases} 255 & f(x, y) \geq T \\ 0 & f(x, y) < T \end{cases} \tag{11}$$

where $f(x, y)$ denotes the gray value of the pixel. The grayscale image converted into the binary image is shown in Fig. 7d where the threshold value T is determined as 105 according to the light condition in the experiment.

As Fig. 7d shows, there is more than one connected component in the binary image. Most of the small connected components are background noises and of no interest and should be eliminated. The statistics of the area of each connected component were obtained and the components in which the area is less than a certain threshold were deleted. The threshold area value was set to 4500 pixels and the effect of the elimination is depicted in Fig. 7e.

The selection of the edge detection algorithm is of importance in image processing. Some operators such as Prewitt operator, Sobel operator, and Robert operator are sensitive to noise and

failed to form a closed edge area [26, 27]. Canny operator [28] can be used in many different scenarios and has satisfied the requirements for edge detection effects with the merits of low error detection rate, low missing detection rate, and high accuracy of edge positioning. The Canny edge operator extracts edges through searching for a local gradient maximum of the image and creating a binary mask that includes the edges of the needle by the first threshold of 40 and the second threshold of 80. The edge features of the needle in the binary image were extracted by the Canny edge operator as shown in Fig. 7f. All the pixel coordinates of the needle feature edges were obtained and the down corner position pixel coordinate (x_0, y_0) of the needle feature edge pixel coordinates is regarded as the pixel coordinate of the needle tip.

3.3. Stereo matching

In the binocular vision system, the images of the needle at different positions and orientations were acquired by the two cameras. The needle tips in the two images were matching and the camera coordinates of the needle tip were estimated.

The epipolar geometry constraint is extensively used in stereo matching. The geometry relationship of the two cameras is illustrated in Fig. 8, a physical point P_w in space is projected to the 2D points P_l in the left image and P_r in the right image, the optical centers of the two cameras are visualized by the points O_{cl} and O_{cr}. All the points mentioned above satisfy the epipolar geometry constraint [29–31].

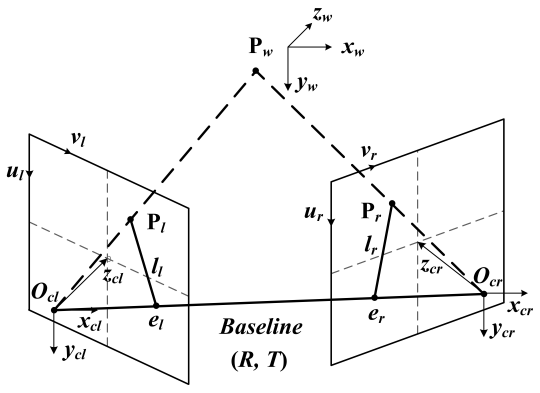


Fig. 8. Epipolar geometry diagram.

The epipolar constraint depends on the cameras' intrinsic and structure parameters. Once the parameters are given, the epipolar line in the right image with respect to point P_l in the left image can be estimated and the searching space is compressed to the 1D line. According to the epipolar geometric relationship, there is a mapping between the 2D point in the left image and the epipolar line in the right image. The fundamental matrix F with a rank of 2 is an algebraic representation defined to represent the mapping relationship [25–27]. The relationship of the fundamental matrix F and the coordinates of the 2D projection points $P_l[u_l, v_l, 1]$, $P_r[u_r, v_r, 1]$ is as follows:

$$P_l^T F P_r = \begin{bmatrix} u_l & v_l & 1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = 0. \tag{12}$$

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 495–512
DOI: 10.24425/mms.2020.134593

The matrix equation can be expanded as:

$$u_l u_r f_{11} + u_l v_r f_{12} + u_l f_{13} + v_l u_r f_{21} + v_l v_r f_{22} + v_l f_{23} + u_r f_{31} + v_r f_{32} + f_{33} = 0. \tag{13}$$

The fundamental matrix F is transferred into the column vector form as follows:

$$\begin{bmatrix} u_l u_r & u_l v_r & u_l & v_l u_r & v_l v_r & v_l & u_r & v_r & 1 \end{bmatrix} \cdot F^* = 0, \tag{14}$$

where $F^* = (f_{11}, f_{12}, f_{13}, f_{21}, f_{22} f_{23}, f_{31}, f_{32}, f_{33})$. Given the set of n sets of points, the following equation can be obtained:

$$A \cdot F^* = \begin{bmatrix} u_l^1 u_r^1 & u_l^1 v_r^1 & u_l^1 & v_l^1 u_r^1 & v_l^1 v_r^1 & v_l^1 & u_r^1 & v_r^1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_l^n u_r^n & u_l^n v_r^n & u_l^n & v_l^n u_r^n & v_l^n v_r^n & v_l^n & u_r^n & v_r^n & 1 \end{bmatrix} \cdot F^* = 0. \tag{15}$$

Theoretically, it only takes eight sets of feature point coordinates to solve the fundamental matrix F [32]. In the actual stereo matching process, more than eight sets of feature point coordinates are used. By solving the equation (15) with the least-squares method, the fundamental matrix is:

$$F = \begin{bmatrix} -5.3121 \times 10^{-8} & 1.2930 \times 10^{-6} & -0.0011 \\ 2.2319 \times 10^{-6} & -2.1695 \times 10^{-7} & 0.0103 \\ -0.0017 & -0.0137 & 1 \end{bmatrix}. \tag{16}$$

Due to the error induced from the camera distortion, the camera calibration, the feature points extraction, and the fundamental matrix estimation, the actual feature point falls on the region with a distance from the epipolar line as illustrated in Fig. 9. The distance $d(P_r, l_r)$ from the actual feature point P_r to the corresponding epipolar line l_r is calculated by:

$$d(P_r, l_r) = \frac{\left|(P_r)^T \cdot F \cdot P_l\right|}{\sqrt{(F \cdot P_l)_1^2 + (F \cdot P_l)_2^2}}. \tag{17}$$

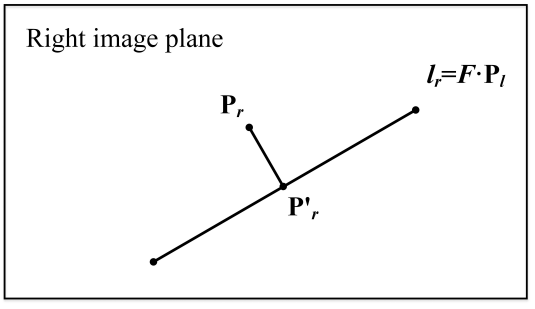


Fig. 9. The error of the stereo matching.

There are 88 corner points in the chessboard calibration target with 9×12 divisions. Twenty sets of chessboard images were sampled by the binocular vision system, the 1760 sets of corner point pixel coordinates in the left chessboard images, as well as the corresponding epipolar line in the right chessboard image, were extracted and generated, the distance from the corner point

to the corresponding epipolar line was estimated. The distance from the 1760 corner points is shown in Fig. 10, the mean distance of the epipolar line and the corresponding corner point is 0.3807 pixels, the maximum distance is 0.5750 pixels. The results indicate that the right matching points are close to the epipolar lines generated by the left matching points. By means of epipolar constraints, the right matching points are searched along the corresponding epipolar lines with the region of 2 pixels and the position of the needle tip is validated.

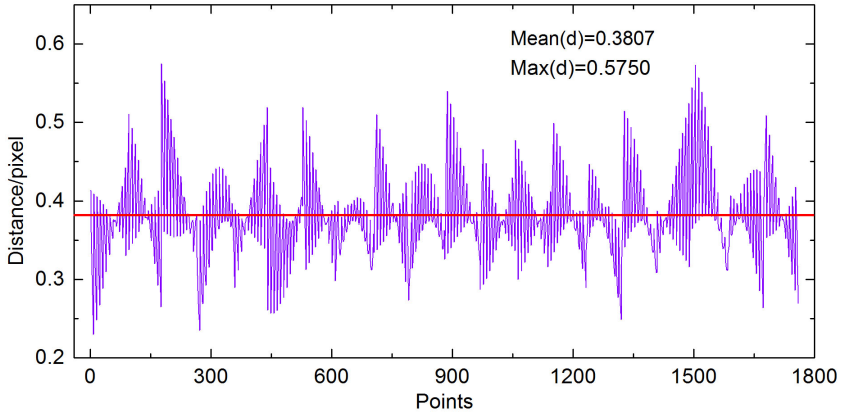


Fig. 10. The distance from the corner points to the corresponding epipolar lines.

3.4. Camera coordinate estimation for the needle tip

The camera coordinates of the needle tip are acquired and calculated with the parameters obtained from the binocular cameras. After the image processing, the pixel coordinates of the needle tip in the two cameras are extracted accurately and satisfy the epipolar constraint. The corresponding image plane coordinates denoted by $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$ arere calculated by equation (7). The camera coordinates of both cameras are $P_{c1}(x_{c1}, y_{c1}, z_{c1})$ and $P_{c2}(x_{c2}, y_{c2}, z_{c2})$ respectively from which the information on the depth z_{c1}, z_{c2} needs to be calculated. The conversion relation of the camera coordinates and the image plane coordinates of the two cameras can be derived from equation (5) as:

$$\begin{aligned} x_{c1} &= \frac{x_1 z_{c1}}{f_1}, \qquad y_{c1} = \frac{y_1 z_{c1}}{f_1}, \\ x_{c2} &= \frac{x_2 z_{c2}}{f_2}, \qquad y_{c2} = \frac{y_2 z_{c2}}{f_2}, \end{aligned} \tag{18}$$

where f_1, f_2 is the focal length of the two cameras.

The position and orientation relationship between the two cameras as well as the camera coordinates of the feature points can be depicted as:

$$P_{c1} = RP_{c2} + T, \tag{19}$$

where the rotation matrix R and the translation vector T can be denoted as:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \qquad T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}. \tag{20}$$

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 495–512
DOI: 10.24425/mms.2020.134593

Equations (18), (19) and (20) were combined and simplified as:

$$\begin{bmatrix} \dfrac{x_1 z_{c1}}{f_1} - t_x \\ \dfrac{y_1 z_{c1}}{f_1} - t_y \\ z_{c1} - t_z \end{bmatrix} = \frac{z_{c2}}{f_2} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ f_2 \end{bmatrix}. \tag{21}$$

Then the equation (21) can be transformed into:

$$\begin{bmatrix} \dfrac{x_1}{f_1} & t_x \\ \dfrac{y_1}{f_1} & t_y \\ 1 & t_z \end{bmatrix} \begin{bmatrix} \dfrac{z_{c1}}{z_{c2}} \\ -\dfrac{1}{z_{c2}} \end{bmatrix} = \begin{bmatrix} \dfrac{r_{11}x_2}{f_2} + \dfrac{r_{12}x_2}{f_2} + r_{13} \\ \dfrac{r_{21}x_2}{f_2} + \dfrac{r_{22}x_2}{f_2} + r_{23} \\ \dfrac{r_{31}x_2}{f_2} + \dfrac{r_{32}x_2}{f_2} + r_{33} \end{bmatrix}. \tag{22}$$

Equation (22) was solved through the *singular value decomposition* (SVD) method and the depth information z_{c1}, z_{c2} was estimated. The camera coordinates of the needle tip were reconstructed automatically from the images acquired by the binocular vision system.

4. System evaluation

The needle tip measurement system was established as described in Section 2.2. The intrinsic and structure parameters of the binocular cameras were obtained after Zhang's calibration illustrated in Section 3.1. Then the needle was placed in the field of view of the two cameras and the images were acquired by the cameras. Based on the needle images, the algorithm mentioned in Sections 3.2–3.4 was executed to estimate the needle tip coordinates. The average cost time of the algorithm is under 0.46s based on the C++ OpenCV code on a personal computer with a Core i7-7700 CPU 3.60 GHz, 8 GB RAM. The results of cost time calculation indicate that the proposed algorithm can meet the requirements of real-time needle tip detection.

The performance of the needle tip measurement system was evaluated using the *measurement capability analysis* (MCA) procedure [33, 34]. In the MCA procedure, the needle tip displacement results measured by the binocular vision system are compared with those collected by the laser interferometer. The MCA procedure used in the binocular vision system evaluation consists of accuracy analysis, agreement analysis, and repeatability analysis.

4.1. Accuracy analysis

In the accuracy analysis, the needle was commanded to move in 2 mm increments 14 times in the x-direction of the robot. The errors are calculated by equation (3) and shown in Fig. 11. The maximum error and the mean error of the binocular vision system is 0.3621 mm and 0.2824 mm compared to the laser interferometer. The correlation between the two comparison results is evaluated by the R-squared value method. The two results are correlated when the R-squared value is larger than 0.75. The R-squared value of the needle displacement between the two methods is 0.8794 by statistical analysis, indicating that the binocular vision system exceeds the accuracy criterion significantly and is adequately accurate for needle tip automated inspection.

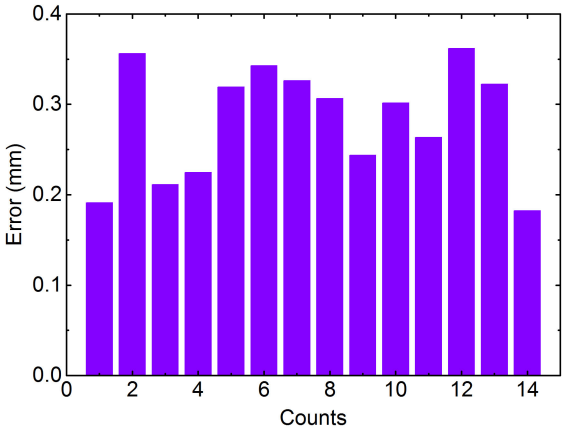


Fig. 11. Accuracy analysis of displacement results of the binocular vision system and the laser interferometer.

4.2. Agreement analysis

In the accuracy analysis, the needle was moved by a fixed distance of 2 mm. The relative position of two points at a random distance in the designed stroke was not considered. Thus, the agreement analysis was proposed to represent the degree of consistency of the two measurement results at random distances. The Bland–Altman analysis [35] and the Pearson correlation coefficient were used to evaluate the agreement of measured displacement between the binocular vision system and the laser interferometer. The needle tip was driven 5 times with a step length of 2 mm within a displacement scope of 2 mm to 28 mm each time. The differences of the needle displacement between the binocular vision system and the laser interferometer were calculated and compared. The Bland–Altman plot for the results is shown in Fig. 12. The mean value of the differences between the two methods is 0.29 mm, the 96% confidence interval of the mean value of the displacement difference is [0.15, 0.43]. The Pearson correlation coefficient is 0.999847 which means that the two methods are in high agreement and linear correlation.

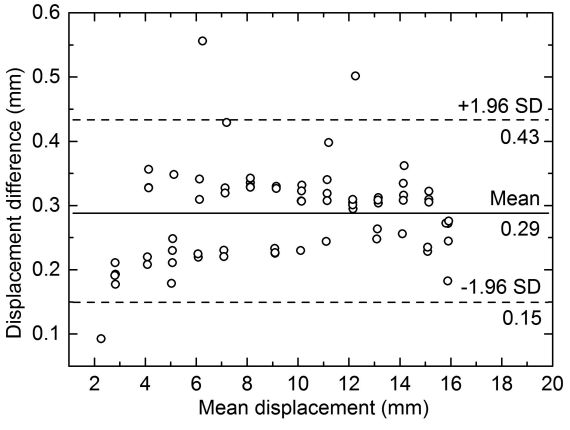


Fig. 12. Bland–Altman plot showing the differences between the binocular vision system and laser interferometer.

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 495–512
DOI: 10.24425/mms.2020.134593

4.3. Repeatability analysis

The acceptability and stability of the inherent error of the proposed binocular vision system in the short term are evaluated by a repeatability analysis. The measurement number for the repeatability analysis is selected between 4 to 10 according to [33]. Five measurement processes were tested and the repeatability analysis was conducted. The results of the repeatability test are shown in Fig. 13. The repeatability analysis is evaluated using the P/T ratio:

$$P/T = \frac{6\sigma_{ms}}{USL - LSL} \times 100\%, \tag{23}$$

where σ_{ms} represents the mean variance of 70 displacements of the needle tip measured, USL–LSL is the tolerance range of the displacement in the pre-defined specification limits. The repeatability of the measurement system is proved when the P/T ratio is less than 20%. The mean variance of the displacement σ_{ms} is 0.017 mm, while the tolerance range of the displacement is 0.62 mm. The P/T ratio value of the binocular vision system on the needle tip displacement is 16.42% which validates the adequate repeatability of the binocular vision system.

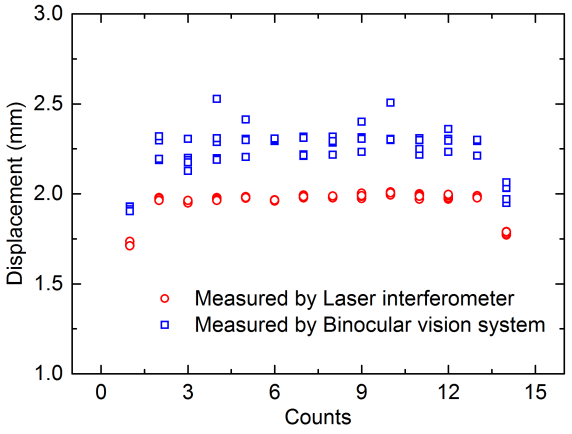


Fig. 13. Repeatability test results of the displacement measuring by the binocular vision system and the laser interferometer.

5. Conclusions

In order to inspect the needle tip position of the robot used in percutaneous prostate intervention accurately, an automatic robust measuring system based on the binocular cameras is designed and evaluated. The proposed binocular vision system is supposed to improve the accuracy of the needle tip position in the preliminary robot research. Furthermore, the system can work with the MRI technology to detect the needle tip position by coordinate transformation during the real prostate intervention. The needle tip measuring system is composed of a standalone PC, two CMOS cameras and a white background plane. The PC controls the cameras and implements the image processing algorithms.

The proposed system is calibrated by Zhang's calibration method firstly and the parameters of the binocular cameras are obtained. The needle images on the white background are acquired and a needle tip extraction algorithm is designed to obtain the pixel coordinates of the needle tip

without any marked point, the stereo matching is then completed and the spatial coordinates of the needle tip are estimated with the epipolar constraint.

A laser interferometer with high accuracy and resolution is used to evaluate the position accuracy of the needle tip. The position inspection is converted into distance measurement and the distances acquired by the laser interferometer are defined as reference values for eliminating the coordinate transformation error. The MCA procedure characterized by accuracy, agreement, and repeatability is used to evaluate the system comprehensively. The maximum error and the mean error of the proposed system are 0.3621 mm and 0.2824 mm respectively and the R-squared value of the needle displacement is 0.8794 which proves good accuracy of the system. The agreement analysis with the laser interferometer is conducted by using the Bland–Altman analysis and Pearson correlation coefficient. The 96% confidence interval of the mean value of the displacement difference is [0.15, 0.43] and the Pearson correlation coefficient is 0.999847 which means that the two methods are in high agreement and linear correlation. The P/T ratio value of the binocular vision system on the needle tip displacement is 16.42% proving the adequate repeatability of the system in the repeatability analysis. According to the evaluation results, the binocular vision system is appropriate for the needle tip measuring application in the needle insertion robot development.

Acknowledgments

This paper is supported by the National Sciences Foundation of China (Grant No. 51975282).

References

[1] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68, 394–424.

[2] Adam, A. (2002). Interventional radiology in the treatment of hepatic metastases. *Cancer Treatment Reviews*, 28(2), 93–99.

[3] Aschoff, A.J., Merkle, E.M., Emancipator, S.N., Petersilge, C.A., Duerk, J.L., Lewin, J.S. (2002). Femur: MR imaging-guided radio-frequency ablation in a porcine model-feasibility study. *Radiology*, 225(2), 471–478.

[4] Wang, Q., Wang, Z., Yao, Z., Forrest, J., Zhou, W. (2016). An improved measurement model of binocular vision using geometrical approximation. *Measurement Science and Technology*, 27(12), 125013.

[5] Mcguire, K., Croon, G.D., Wagter, C.D., Tuyls, K., Kappen, B. (2017). Efficient Optical Flow and Stereo Vision for Velocity Estimation and Obstacle Avoidance on an Autonomous Pocket Drone. *IEEE Robotics & Automation Letters*, 2(2), 1070–1076.

[6] Sanchez-Rodriguez, J.P., Aceves-Lopez, A. (2018). A survey on stereo vision-based autonomous navigation for multi-rotor MUAVs. *Robotica*, 36(8), 1225–1243.

[7] Li, X., Liu, W., Pan, Y., Ma, J., Wang, F. (2019). Binocular vision-based 3D method for detecting high dynamic and wide-range contouring errors of CNC machine tools. *Measurement Science And Technology*, 30(12), 125019.

[8] Bellandi, P., Docchio, F., Sansoni, G. (2013). Roboscan: a combined 2D and 3D vision system for improved speed and flexibility in pick-and-place operation. *The International Journal of Advanced Manufacturing Technology*, 69(5), 1873–1886.

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 495–512
DOI: 10.24425/mms.2020.134593

[9] Luo, Z., Zhang, K., Wang, Z., Zheng, J., Chen, Y. (2017). 3D pose estimation of large and complicated workpieces based on binocular stereo vision. *Applied Optics*, 56(24), 6822–6836.

[10] Xia, R., *et al*. (2020). An accurate and robust method for the measurement of circular holes based on binocular vision. *Measurement Science and Technology*, 31(2), 025006.

[11] Pełczyński, P., Ostrowski, B., Rzeszotarski, D. (2012). Motion Vector Estimation of a Stereovision Camera with Inertial Sensors. *Metrology and Measurement Systems*, 19(1), 141–150.

[12] Lin, G., Tang, Y., Zou, X., Xiong, J., Fang, Y. (2019). Color-, depth-, and shape-based 3D fruit detection. *Precision Agriculture*, 21, 1–17.

[13] Tang, Y., Chen, M., Wang, C., Luo, L., Zou, X. (2020). Recognition and Localization Methods for Vision-Based Fruit Picking Robots: A Review. *Frontiers in Plant Science*, 11, 1–17.

[14] Xiang, R., He, W., Zhang, X., Wang, D., Shan, Y. (2018). Size measurement based on a two-camera machine vision system for the bayonets of automobile brake pads. *Measurement*, 122, 106–116.

[15] Li, J., Bennett, B.L., Karam, L.J., Pettinato, J.S. (2016). Stereo Vision Based Automated Solder Ball Height and Substrate Coplanarity Inspection. *IEEE Transactions on Automation Science and Engineering*, 13(2), 757–771.

[16] Lin, Q., Cai, K., Yang, R., Chen, H., Wang, Z., Zhou, J. (2016). Development and Validation of a Near-Infrared Optical System for Tracking Surgical Instruments. *Journal of Medical Systems*, 40(107), 1–14.

[17] Jiang, G., Luo, M., Bai, K. (2019). Optical positioning technology of an assisted puncture robot based on binocular vision. *International Journal of Imaging Systems and Technology*, 29(2), 180–190.

[18] Zhou, Z., Wu, B., Duan, J., Zhang, X., Zhang, N., Liang, Z. (2017). Optical surgical instrument tracking system based on the principle of stereo vision. *Journal of Biomedical Optics*, 22(6), 065005.

[19] Stamey, T.A., Freiha, F.S., Mcneal, J.E., Redwine, E.A., Whittemore, A.S., Schmid, H.P. (2015). Localized prostate cancer. Relationship of tumor volume to clinical significance for the treatment of prostate cancer. *Cancer*, 71(S3), 933–938.

[20] Stamey, T.A., McNeal, J.E., Freiha, F.S., Redwine, E. (1988). Morphometric and clinical studies on 68 consecutive radical prostatectomies. *Journal of Urology*, 139(6), 1235–1240.

[21] Krieger, A., Song, S., Cho, N.B., Iordachita, I., Guion, P., Fichtinger, G. (2013). Development and Evaluation of an Actuated MRI-Compatible Robotic System for MRI-Guided Prostate Intervention. *IEEE/ASME Transactions on Mechatronics*, 18(1), 273–284.

[22] Elhawary, H., Tse, Z. T. H., Rea, M. Zivanovic, A., Davies, B.L., Besant, C., Souza, N., McRobbie, D., Young, I., Lamperth, M.U. (2010). Robotic System for Transrectal Biopsy of the Prostate: Real-Time Guidance Under MRI. *IEEE Engineering in Medicine and Biology Magazine*, 29(2), 78–86.

[23] Seifabadi, R., Song, S., Krieger, A., Cho, N.B., Tokuda, J., Fichtinger, G., Iordachita, I. (2012). Robotic system for MRI-guided prostate biopsy: feasibility of teleoperated needle insertion and ex vivo phantom study. *International Journal of Computer Assisted Radiology and Surgery*, 7(2), 181–190.

[24] Chen, Y., Squires, A., Seifabadi, R., Xu, S., Agrawal, H., Bernardo, M., Pinto, P., Choyke, P., Wood, B., Tse, Z.T.H. (2017). Robotic System for MRI-Guided Focal Laser Ablation in the Prostate. *IEEE/ASME Transactions on Mechatronics*, 22(1), 107–114.

[25] Zhengyou, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), 1330–1334.

[26] Raman, M., Aggarwal, H. (2009). Study and Comparison of Various Image Edge Detection Techniques. *International Journal of Image Processing*, 3(1), 1–11.

[27] Zhang, X. H., Li, G., Li, C. L., Zhang, H., Zhao, J., Hou, Z. X. (2015). Stereo Matching Algorithm Based on 2D Delaunay Triangulation. *Mathematical Problems in Engineering*, 2015(PT.17), 137191–137193.

[28] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698.

[29] Lu, J., Cai, H., Lou, J., Li, J. (2007). An Epipolar Geometry-Based Fast Disparity Estimation Algorithm for Multiview Image and Video Coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(6), 737–750.

[30] Han, J. H., Park, J. S. (2000). Contour matching using epipolar geometry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(4), 358–370.

[31] Zhang, Z. (1998). Determining the Epipolar Geometry and its Uncertainty: A Review. *International Journal of Computer Vision*, 27(2), 161–195.

[32] Schmid, C., Zisserman, A. (2000). The Geometry and Matching of Lines and Curves Over Multiple Views. *International Journal of Computer Vision*, 40(3), 199–233.

[33] Dong, S.C. (2006). *Measurement System Analysis: Theory, Method and Applications*. China Metrology Press, 143–148.

[34] Wu, Z.G. (2004). *Measurement System Analysis*. China Standards Press, 86–95.

[35] Bland, J.M., Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *International Journal of Nursing Studies*, 47(8), 931–936.