Automatic Measurement Based on Stereo Vision System Using a Single PTZ Camera

Rui Wang(Member, IEEE), Suizhou Feng
Key Laboratory of Precision Opto-mechatronics Technology, Ministry of Education, School of Instrumentation Science and Opto-electronics Engineering, Beihang University
Beijing 100191, China
E-mail: wangr@buaa.edu.cn, 11171150@buaa.edu.cn

***Abstract*—A parallel binocular stereo vision system and its application to dimensional measurement is introduced in this paper. The proposed system is established by controlling a single PTZ (pan-tilt-zoom) camera linear motion on a one dimensional precision displacement platform. To achieve accurate measurement results, an improved matching algorithm called as DC-SURF (Distortion Compensated-Speeded Up Robust Feature) by using adaptative filtering in terms of the image distortion during generating scale space is presented. Furthermore, the classic eight-point algorithm and the designed zoom strategy for PTZ camera are adopted for the measurement. The experimental results show that the proposed matching algorithm has better performance than other two classic matching algorithms over distorted images. Furthermore, the experiments demonstrate that without any reference materials the stereo vision dimensional measurement system proposed in this paper can be successfully applied to human height measurement with high precision. The results indicate a potential possibility of our approach to be used in other computer vision applications.**

Keywords—stereo vision; PTZ camera; adaptative filtering; zoom strategy; dimensional measurement

I. INTRODUCTION

Recently the high-accuracy dimensional measurement is playing an important role in a variety of computer vision applications, for instance, human-computer interface, biometric authentication, robot vision, etc. Current dimensional measurement techniques can be divided into two main types: active and passive. Generally, active measurement adopts laser scanning or structure illumination, which will not be convenient in many applications. However, passive measurement techniques overcome this deficiency by requiring simple equipment [1]. In addition, passive measurement could be achieved based on monocular vision, binocular stereo vision or multi-view stereo vision, a monocular vision based object dimensional measurement method is presented in [2], which requires a known size object to calculate the size of the measured target. Compared to monocular vision, binocular vision based method can obtain accurate results without any prior knowledge, e.g., a 3D measurement method proposed in [3] can be used to measure the pins' positions in production lines without prior information. To utilize the advantages of binocular stereo vision and simplify the calibration of the camera for binocular vision , a structure model of single camera binocular vision sensor using mirrors is proposed in [4], which is equivalent to the traditional binocular vision sensor and could obtain two virtual images reflected by two mirrors of the object in a single shot for further measurement, yet the disadvantage of this system is that the installation of the mirrors and the calibration of the angel between the mirrors is complicated. From above analysis, in this paper, we established a parallel binocular stereo vision system based on a single PTZ (pan-tilt-zoom) camera, which is constructed by controlling a single PTZ camera's linear motion on a one-dimensional precision displacement platform during the process of obtaining 3D coordinates for further measurement.

On the other hand, in the process of the passive measurement based on stereo vision, image matching is crucial which partly determines the accuracy of the overall procedure. Generally, matching techniques are classified into three major types: pixel-based, area-based and feature-based [5,6]. A pixel-based method [7] implements the matching at each pixel with only the intensity of a single pixel. Area-based algorithms [8,9] find correspondences based on similarities between areas in the right and left images, which assumes that disparities within a rectangular window centered at a pixel are constant. A feature-based method known as SIFT (Scale-invariant feature transform) [10] and its improved version SURF (Speeded Up Robust Features) [11] which uses integral images and has less dimensions of the descriptor can both provide invariance to common image transformations such as scale, rotation, illumination, and minimal viewpoint changes. Unfortunately, the wildly used SIFT and SURF approaches are not invariant to the radial distortion present in images acquired by cameras. Addressing this problem, we proposed an improved version for SURF that could compensate the distortion in the images and achieve superior matching results. Hence, based on our superior matching results, we can acquire a better measurement results by means of the proposed parallel binocular stereo vision system and the classic eight-point algorithm [12,13].

The remainder of the paper is organized as follows: Section II summarizes the system establishment and the measurement process. Section III describes our developed matching algorithm. In section IV, an overview of eight-point algorithm is given and the zoom strategy for measurement is presented. Section V shows experiment results and the analysis. Finally, a conclusion is given in Section VI.

II. MEASUREMENT SYSTEM SETUP

To accomplish the automatic dimensional measurement, we establish a parallel binocular stereo vision system based on a single PTZ camera. The measurement process consists of four steps and is shown in Fig. 1.

A. PTZ Camera Self-calibration

Firstly, we calibrate the PTZ camera through the method presented in [14], which doesn't require 2D or 3D calibration target and only needs to capture two images of the same scene in orientation 1 and orientation 2 after a relative rotation (pan and tilt). Two principle equations are listed as follows

$$(D_1 + \eta D_2 + \eta^2 D_3)h = 0 \tag{1}$$

Here h is a column vector collecting the nine elements of the homography H between the two images captured in two orientations, η is distortion coefficient based on division distortion model [15], D_1, D_2, D_3 are 2×9 matrixes related to the distorted image coordinate and the detailed expression is presented in [14]. Hence, the relation between intrinsic matrix K and homography H can be represented as

$$Q_3Q_2Q_1K^{-1}H \sim K^{-1}, \quad K = \begin{bmatrix} \alpha f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

In (2), Q_3, Q_2, Q_1 are Givens rotations [16] which can simplify the calculation, f is focal length of the camera, α is the aspect ratio and (u_0,v_0) is the coordinate of principle point. Hence, the intrinsic parameters (α,f,u_0,v_0,η) of the PTZ camera can be obtained by solving (1) and (2).

B. Stereo Vision System Setup

In our system, HIWIN KK50 ball screw displacement platform, which can be driven by permanent-magnet AC servo motor control system for linear motion, with absolute positioning accuracy ±0.01mm and repeated positioning accuracy 0.02mm is used to hold up the single PTZ camera. The value of the linear displacement of the mounted single PTZ camera is the length of the baseline in the established binocular parallel stereo vision system. In our system, the precise linear displacement of the camera is proportional to the angular displacement of Mitsubishi HF-KP13 servo motor controlled by Mitsubishi PLC. Thus using PLC programming to control the displacement-distance of displacement platform, it improves the degree of automation and displacement accuracy, i.e., the accuracy of the baseline in our stereo vision system. We can capture two images on two endpoints of the baseline for subsequent matching and measurement.

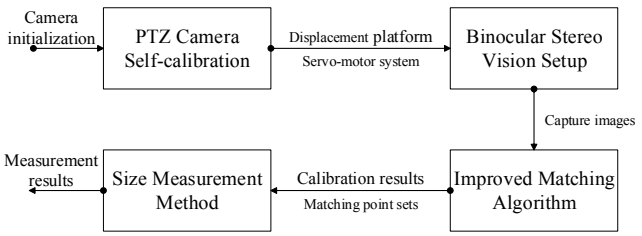


Fig. 1. Flow diagram of measurement process.

C. Improved Matching Algorithm

The lens manufacture error and the components assembly error of PTZ camera can lead to distortion in captured images. However, the state-of-the-art methods SIFT and SURF suffer a considerable deterioration for a certain amounts of distortion, which has a negative impact in the matching performance. Considering our demand of high accuracy real-time measurement, we proposed an improved stereo matching algorithm which is based on SURF and compensates the distortion in images by implementing adaptative filtering during generating the scale space to acquire accurate matching results.

D. Measurement Algorithm

Based on the above steps, we adopted the classic eight-point algorithm [12,13] to estimate the 3D coordinates of the measured object and further to calculate the dimension of the object, in addition, we design the zoom strategy of the PTZ camera for high-accuracy measurement.

III. IMPROVED SURF UNDER RADIAL DISTORTION

Despite of the fact that the SURF algorithm is not invariant to lens distortion, its operation efficiency is of great significance in real-time measurement. In order to obtain more accurate measurement results, we improved original SURF by using adaptative box filters during building the scale space over distorted images, dubbed as DC-SURF (Distortion Compensated-Speeded Up Robust Feature), which will be described in detail.

A. Division Model for Radial Distortion

In our method, we adopt a first-order division model [15] to represent the effect of lens distortion in images. The degree of distortion can be quantified by a distortion parameter η (generally, $\eta < 0$), and the distortion center is assumed to be closely approximated to the image center. Let image center be the origin of the image coordinate system in images, $p_u = (u,v)^T$ is the coordinates in undistorted image, $p_d = (x,y)^T$ is the coordinates of corresponding point in distorted image, hence

$$p_u = p_d / (1 + \eta r_d^2) \tag{3}$$

Here η is the distortion parameter, r_d is the distortion radius which denotes the distance from the distorted point to the image center and $r_d = || p_d ||_2$, the corresponding undistorted radius $r_u = r_d/(1+\eta r_d^2)$.

B. Adaptative Filtering and Scale Space Generating

We introduce a new method for image adaptative filtering which accounts for the distortion based on the above division model. The purpose is to generate a scale space which is equal to the one that would be obtained by filtering the image in the absence of distortion, followed by applying the distortion over all the layers of the SURF pyramid. To achieve our objective, we implement an implicit distortion correction method by adapting the convolution kernel that is used over the distorted image during the scale space generation. Due to the use of box filters and integral images [11], we apply box filters of any size at exactly the same speed directly on the original image

and the scale space is generated by up-scaling the filter size without changing the image size. The output of the 9×9 box filter is considered as the initial scale layer, to which we will refer as scale σ = 1.2 (approximating Gaussian derivatives with σ = 1.2). The following layers in Fig. 2 are acquired by filtering the original image with box filter whose size is gradually increasing, the number of octaves is set to be 4 and the box filter size in each layer is rounded to the nearest integer. Let $(h,k)^T$ be the pixel coordinate in each layer to be convolved, the scale space considering distortion in image is given by

$$L(h,k;\sigma)=\sum_x\sum_y I(x,y)G(\frac{h-x}{1+\eta r^2},\frac{k-y}{1+\eta r^2};\sigma) \qquad (4)$$

The expression can be rewritten by transforming the Gaussian convolution kernel G by G'

$$G'=G(x,y;(1+\eta r^2)\sigma),\quad D'=D(x,y;(1+\eta r^2)s) \qquad (5)$$

Here r is the corresponding distortion radius in each layer, D' is the box filter kernel approximate to the adaptative Gaussian kernel G', s is the size of box filter. While in the original SURF detection in the image is filtered using standard isotropic box filter with standard size s, in our method the standard size can be represented as $(1+\eta r^2)s$ that decreases as a function of the image radius. The convolution kernel follows the deformation caused by distortion, and emphasizes the contribution of pixels increasingly closer to the convolution point while the filter moves far from the center of distortion.

C. Distortion Compensated Speeded Up Robust Feature

As described earlier, the novel DC-SURF algorithm could be accomplished by the following steps

1) Fast interest point detection: Select the location and scale of interest points according to the determinant of the Hessian matrix in each pixel.

2) Build the novel scale space: Implement the adaptative box filtering to buid the scale space accounting for distortion.

3) Interest point descriptor: In a first step, find interest points in scale space by applying non-maximum suppression in a 3×3×3 neighborhood; in a second step, construct a circular region around the detected interest points and compute the principle orientation by using Haar wavelet theory; in a third step, construct the DC-SURF descriptor vector which has a length of 64 by extracting square regions around the detected interest points.

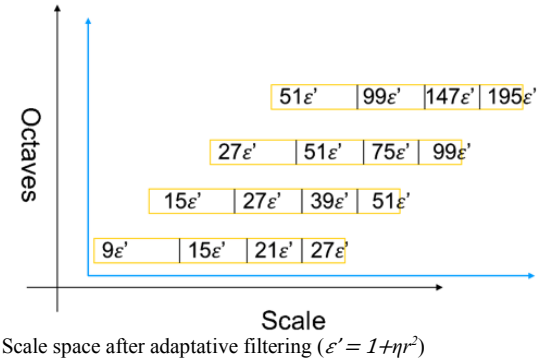


Fig. 2. Scale space after adaptative filtering ($\varepsilon'=1+\eta r^2$)

4) Feature descriptors matching: In two corresponding images, calculate the Euclidian distance to all feature descriptors in the second image for every interest point in the first image, if the ratio of the nearest neighbor to the second-nearest neighbor is smaller than a predetermined threshold, a match is considered to be correct.

We compare the proposed DC-SURF algorithm with SURF and SIFT in INRIA datasetsd to validate its performance in section V.

IV. MEASUREMENT

Based on the procedures mentioned in previous two sections, we subsequently execute dimensional measurement through estimating the fundamental matrix optimized by RANSAC [17] and deducing the formula of dimension calculation, during which we adopt classic eight-point algorithm that reconstructs a scene from two projections firstly proposed by Longuet-Higgins in [12] and enhanced by Hartley in [13].

A. Eight-point Algorithm

The model of our parallel binocular vision system is illustrated in Fig. 3. Let O-XYZ be the world coordinate system, which is coincide with the camera coordinate system of the left camera, the visible point $P(X,Y,Z)^T$ in the scene is projected onto the image planes, resulting in points $p_l(x_l,y_l)^T$ and $p_r(x_r,y_r)^T$. The steps of eight-point algorithm are explained as follows

1) Select points and estimate the fundamental matrix: Given a set of point correspondences in homogeneous coordinates $\{p_l'\leftrightarrow p_r'\}_i$, $i=1,2,\ldots,8$, between two uncalibrated perspective views of a rigid scene, the objective is to estimate the 3×3 fundamental matrix F satisfying the epipolar constraints

$$p_r'^{T}Fp_l'=0 \qquad (6)$$

Where p_l' and p_r' are the normalized coordinates of p_l and p_r, the F should be of rank-2 because of the epipolar geometry constraint that all epipolar lines must intersect at a point. In order to obtain optimal measurement results, we implement RANSAC algorithm to estimate the accurate fundamental matrix during this step.

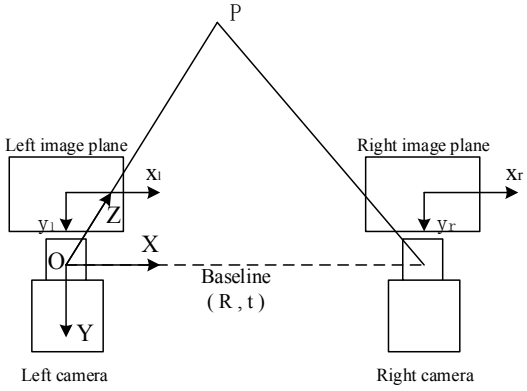


Fig. 3. Binocular parallel stereo vision system.

2) Estimate the rotation matrix and translation vector between two cameras: Fundamental matrix and essential matrix describe the position, orientation and properties of the two cameras in stereo vision, through transforming the fundamental matrix to essential matrix by (7), it is convenient to obtain the rotation matirx R and translation vector t between two cameras after solving (8) by using a SVD method proposed in [18]

$$E = K_r^T F K_l \tag{7}$$

K_r and K_l are the intrinsic parameters matrix of two cameras in (2), in our system $K_r = K_l$.

$$E = RS \tag{8}$$

Here S is a 3×3 skew-symmetric matrix of t.

3) Calculate the 3D coordinates and dimension of object: The 3D world coordinates of the point on object is given by

$$\begin{cases} X = \dfrac{x_l\left(f_r R_{1j} - x_r R_{3j}\right)\cdot t}{f_l\left(f_r R_{1j} - x_r R_{3j}\right)\cdot p_l^{'}}, \quad Y = \dfrac{y_l\left(f_r R_{1j} - x_r R_{3j}\right)\cdot t}{f_l\left(f_r R_{1j} - x_r R_{3j}\right)\cdot p_l^{'}} \\ Z = \dfrac{\left(f_r R_{1j} - x_r R_{3j}\right)\cdot t}{\left(f_r R_{1j} - x_r R_{3j}\right)\cdot p_l^{'}}, \quad R = \begin{pmatrix} R_{1j} & R_{2j} & R_{3j} \end{pmatrix}^T \end{cases} \tag{9}$$

Where f_l and f_r are the respective focal length of two cameras and $f_l = f_r$. Finally, through selecting two points on the object and calculate the distance between them, we can accomplish the dimension measurement

$$dimension = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2} \tag{10}$$

In (10), $(X_1,Y_1,Z_1)^T$ and $(X_2,Y_2,Z_2)^T$ are the 3D world coordinates of the chosen points on object.

B. PTZ camera zoom strategy during the measurement

The image size of the measured object partly determines the measurement accuracy. Therefore, it is crucial to formulate a zoom strategy under which the PTZ camera can capture the measured object with a proper size in image plane after zoom operation. We define a variable to represent the ratio of the object's image size to the overall image size as follows

$$\delta(\%) = 100 \times S_c / S_i \tag{11}$$

Here, δ is the ratio, S_c is the measured object's image size decomposed in the linear direction decided by the two select points, S_i is the overall image size decomposed in the corresponding linear direction. By setting two thresholds t_{max} and t_{min}, we implement the zoom operation to satisfy the condition $t_{min} < \delta < t_{max}$ so that we can capture images with apposite measured object's image size for dimensional measurement. These two thresholds are determined in terms of the measurement data in the next section.

V. EXPERIMENT AND ANALYSIS

In this section, we firstly present results comparing DC-SURF to standard SIFT and SURF in INRIA datasets. Secondly, we summarize the dimensional measurement results to validate our algorithm.

A. DC-SURF Performance Evaluation

Recall–Precision is one of the most prevalent metrics for evaluating detection tasks [19], we will use *Recall–Precision* to evaluate the matching performance of the DC-SURF algorithm developed.

The *recall* is the proportion of the number of the correctly matched interest regions to the number of the corresponding interest regions between two images of the same scene

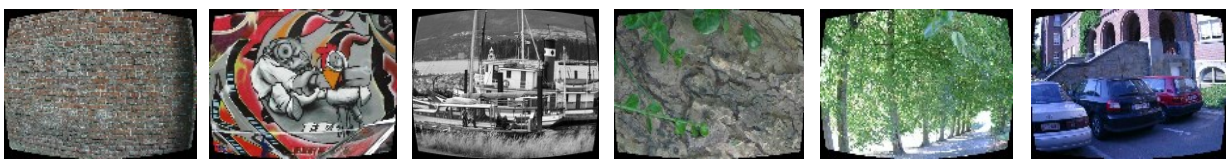
$$recall = (\#correct \quad matches) / (\#correspondences) \tag{12}$$

1-precison is the proportion of the number of the false matches to the total number of matches

$$1 - precision = (\#false \quad matches) / (\#total \quad matches) \tag{13}$$

We generate the *recall* vs. *1-precision* graphs for our experiments by varying the threshold for each descriptor.

The dataset for evaluation downloaded from INRIA datasets contains image sequences with different geometric and photometric transformations, including viewpoint change, zoom and rotation, blur, light change. According to the problems our method focuses on, two images from each sequence are chosen and distorted by using (3) (see as Fig. 4 (a)-(f)). Specifically, the DC-SURF is compared with two existing algorithms in the literature: 1) the standard SIFT; 2) the standard SURF. Fig. 5(a)-(f) exhibit the results of the matching experiments under various transformations. In Fig. 5 (a)-(b) the performance is measured for distorted images with viewpoint change, we can observe that DC-SURF is dramatically better at handling distorted images for almost all values of 1-precision. Fig. 5 (c)-(d) demonstrate the influence of zoom and rotation in *boat* and *bark* scene under distortion, results clearly show that DC-SURF dominates the other two algorithms. Finally, we can see from Fig. 5 (e)-(f) that while all of the representations are well-suited to capturing these two variations, the performance of DC-SURF is slightly better in the case of blur and light change. As a result, we can draw a conclusion that compared to classic SURF and SIFT, DC-SURF algorithm has superior robustness to distortion.



(a) (b) (c) (d) (e) (f)

Fig. 4. Distorted version of images from the INRIA datasets for evaluation: (a)wall+(b)graffi: viewpoint change; (c)boat+(d)bark: zoom+rotation; (e): blur; (f)leuven: light change.

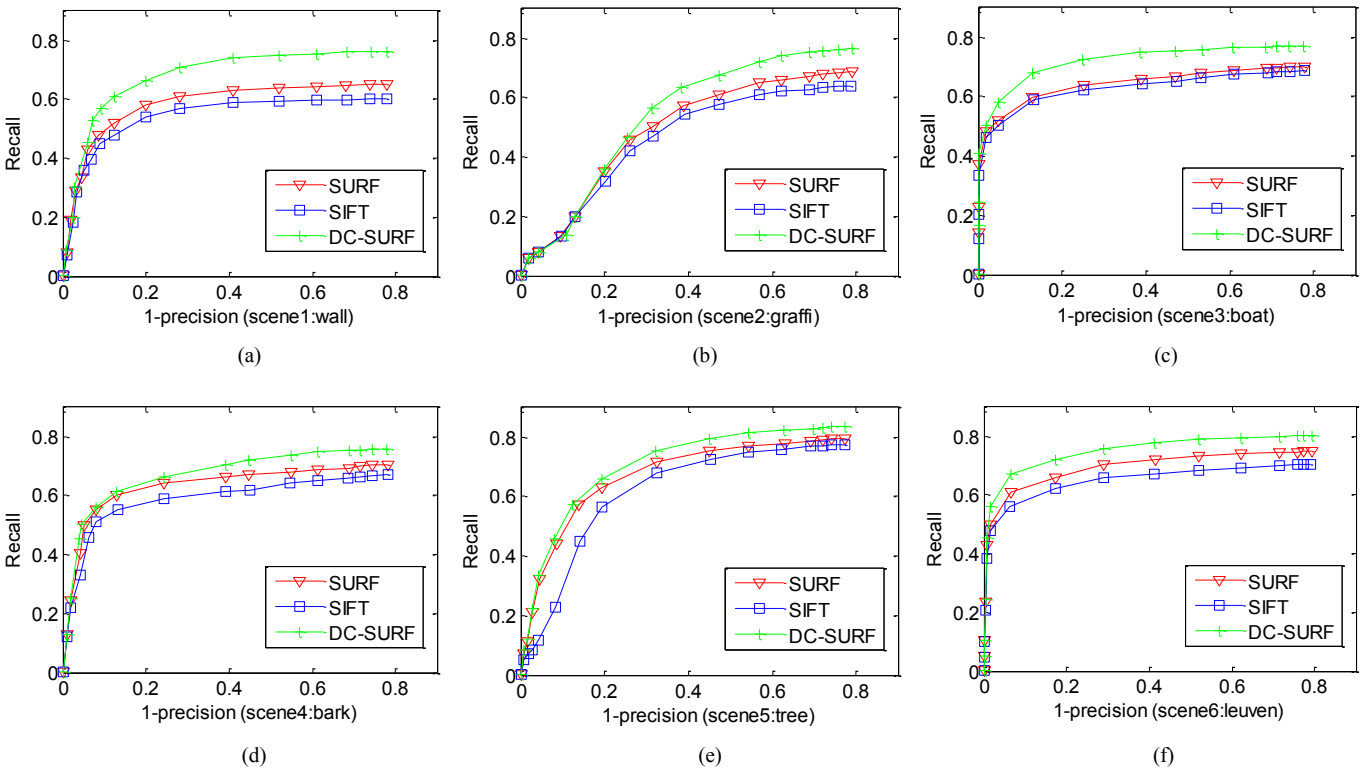


Fig. 5. Illustration of the performance of *Recall vs. 1-precision* curves: (a) and (b), viewpoint change; (c) and (d), zoom + rotation; (e) image blur; (f) light change.

(a) (b)

Fig. 6. Two camera sample images of different target: (a) the image for dimensional measurement; (b) the image for height measurement.

B. Dimensional Measurement Experiment

The parameters of the parallel binocular stereo vision system are given as follows: PTZ camera used in system is SONY EVI-D70P, length of the system baseline is 200mm (theoretically, there is a negative correlation between baseline length and measurement error. In our experiments, the selected baseline length can guarantee relatively accurate measurement results), resolution of image obtained from PTZ camera is 320×240 pixels, the size of each pixel is 6.5μm×6.5μm, the focal length of PTZ camera can be adjusted according to the zoom strategy. Thus we can capture two images by the PTZ camera on two endpoints of baseline and extract relevant points on object manually to complete dimensional measurement.

To form an appropriate zoom strategy during measurement for PTZ camera, we firstly choose a planar target with standard grid size as the measured object. We select appropriate number of squares in Fig. 6 (a) for dimensional measurement with every square size is 30mm×30mm, and the distance between the object and the camera is increased from initial 1110mm to different integral values shown in Fig. 7. Secondly, we apply our zoom strategy summarized from dimensional measurement to the human height measurement to validate the proposed algorithm. Fig. 6 (b) displays an images collected by PTZ camera for height measurement and the man in the picture has a standard height of 165cm.

The object dimensional measurement results at several distances are illustrated in Fig. 7, in which we increase the focal length of camera to obtain larger ratio δ and estimate the measurement error at every distance. We can see that the measurement error has obvious negative correlation with the ratio δ in any positions of the object, which indicates we can reduce error by increasing the focal length of PTZ camera to obtain a larger ratio for measurement.

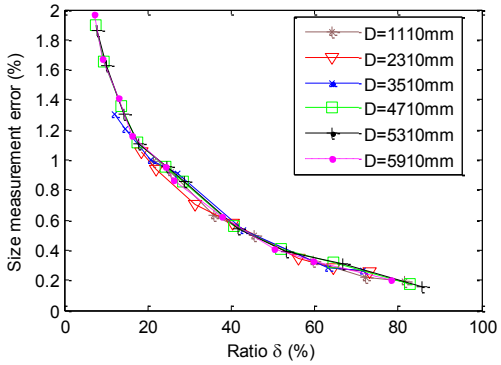


Fig. 7. Dimensional measurement results after PTZ camera zoom operation (in figure legend the letter D is the distance between object and left camera).

TABLE I. HUMAN HEIGHT MEASUREMENT RESULTS (REAL HEIGHT H_0=1650.00mm)

Distance(mm)	4110	4710	5310	5910	5310
Focal Length(mm)	2.53	3.01	3.45	3.45	3.79
Measured Height(mm)	1644.37	1655.12	1655.17	1655.75	1655.17
Measurement error(%)	0.34	0.31	0.31	0.35	0.31
Ratio δ(%)	60.83	64.58	64.17	57.50	65.00

It can be observed from Fig. 7 that when the ratio δ is higher than 20%, the dimension error is below 1%. Furthermore, we conclude from our experiment that it will be difficult to match the public areas in two different view from endpoints of the baseline for the same measured target if the ratio δ is too large. In order to balance the matching difficulty with measurement accuracy, we set the threshold t_{min} is equal to 40% and the threshold t_{max} is equal to 65%, which is testified in Fig. 7 that dimension error is approximately between 0.3% and 0.5% as long as the ratio δ satisfies the condition $t_{min} < \delta < t_{max}$. For the objective to attain results with higher accuracy, we stipulate the zoom strategy making the ratio in (11) near 65% after zoom operation in our system.

Based on the above zoom strategy, we implement the human height measurement by using our developed parallel binocular stereo vision system. The results are listed in Table I, from which we can summarize that when the ratio is near 65% (from 57.5% to 65% in our experiment), we can accomplish accurate height measurement with the errors staying below 0.35% by using our system and the proposed algorithm.

VI. CONCLUSION

In this paper we established a parallel binocular stereo vision system for automatic dimensional measurement. During the measurement procedure, an enhanced stereo matching algorithm accounting for image distortion is presented and the zoom strategy used for high-accuracy measurement is formulated. The performance of the improved matching algorithm in distorted images had been verified to be superior to standard SIFT and SURF algorithms in the INRIA dataset. Furthermore, the high precision results of the dimensional measurement confirmed that our system could be surely applied to field of computer vision application.

ACKNOWLEDGMENT

The authors thank the anonymous reviewers for helping. This work was supported by a grant from National Natural Science Foundation of China (61673039).

REFERENCES

[1] Petrov M, Talapov A, Robertson T, et al. Optical 3D digitizers: Bringing life to the virtual world[J]. IEEE Computer Graphics & Applications, 1998, 18(3):28-37.

[2] Kang M S, Lee C H, You B M, Chung Y S. A 3D object measurement method using a single view camera[C]// International Conference on Information and Communication Technology Convergence. IEEE, 2015:790-792.

[3] Stroppa L, Cristalli C. Stereo Vision System for Accurate 3D Measurements of Connector Pins' Positions in Production Lines[J]. Experimental Techniques, 2016:1-10.

[4] Cheng Y, Fuqiang Z. Structure Design of Binocular Vision Sensor Using Mono-camera with Mirrors[J]. Journal of Mechanical Engineering, 2011, 47(22):7.

[5] Dhond U. Aggarwal J: Structure from stereo-a review[J].

[6] Wei G Q, Brauer W, Hirzinger G. Intensity- and gradient-based stereo matching using hierarchical Gaussian basis functions[J]. Pattern Analysis & Machine Intelligence IEEE Transactions on, 1998, 20(11):1143-1160.

[7] Cox I J, Hingorani S L, Rao S B, Maggs B M. A maximum likelihood stereo algorithm ☆, ☆☆[J]. Computer Vision & Image Understanding, 1996, 63(3):542-567.

[8] Kanade T, Okutomi M. A stereo matching algorithm with an adaptive window: Theory and experiment[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1994, 16(9):920-932.

[9] Okutomi M, Katayama Y, Oka S. A simple stereo algorithm to recover precise object boundaries and smooth surfaces[C]// IEEE Workshop on Stereo and Multi-Baseline Vision. IEEE Computer Society, 2001:158.

[10] Lowe D G. Distinctive image features from scale-invariant keypoints[M]. Kluwer Academic Publishers, 2004.

[11] Bay H, Tuytelaars T, Gool L V. SURF: Speeded up robust features[J]. Computer Vision & Image Understanding, 2006, 110(3):404-417.

[12] Longuet-Higgins H C. A computer algorithm for reconstructing a scene from two projections[C]// Morgan Kaufmann Publishers Inc. 1987:61-62.

[13] Hartley R I. In defense of the eight-point algorithm[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1997, 19(6):580-593.

[14] Jiaoru Yang, Rui Wang. High-accuracy homography estimation for robust pan-tilt-zoom camera calibration[C]// 2017 International Conference Communication Computer Networks and Intelligent Computing, in press

[15] Fitzgibbon A W. Simultaneous linear estimation of multiple view geometry and lens distortion[C]// Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE, 2003:I-125-I-132 vol.1.

[16] Junejo I N, Foroosh H. Optimizing PTZ camera calibration from two images[M]. Springer-Verlag New York, Inc. 2012.

[17] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Readings in Computer Vision, 1987:726-740.

[18] Wang R, Yuan Y. Image quality assessment using full-parameter singular value decomposition[J]. Optical Engineering, 2011, 50(5):-.

[19] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2005, 27(10):1615.