Fully automatic algorithm for region of interest location in camera calibration

Xinyu Kou
Zhong Wang
Mingzhou Chen
Shenghua Ye, MEMBER SPIE
Tianjin University
State Key Laboratory of Precision Measuring Technology and Instruments
Tianjin, China, 300072

Abstract. We present an automatic method for region of interest (ROI) location in camera calibration used in computer vision inspection. An intelligent ROI location algorithm based on the Radon transform is developed to automate the calibration process. The algorithm remains robust even if the anchor target has a notable rotation angle in the target plane. This method functions well although the anchor target is not carefully positioned. Several improvement methods are studied to avoid the algorithm's huge time/space consumption problem. The algorithm runs about 100 times faster if these improvement methods are applied. Using this method fully automatic camera calibration is achieved without human interactive ROI specification. Experiments show that this algorithm can help to calibrate the intrinsic parameters of the zoom lens and the camera parameters quickly and automatically. [DOI: 10.1117/1.1476327]

Subject terms: fully automatic camera calibration; zoom lenses; Hough transforms; Radon transforms; intelligent calibration; computer vision.

Paper 010131 received Apr. 10, 2001; revised manuscript received June 20, 2001; accepted for publication Dec. 10, 2001.

1 Introduction

Camera calibration is an important task in computer vision fields. The camera or lens must be carefully calibrated before application so that the internal parameters such as the focal length of the lens, the optical principle point, the sampling scale factor, and lens distortions can be available and precise. Different calibration algorithms have been studied. Typical algorithms include direct linear transform (DLT) algorithm[1] and the radial alignment constraint (RAD) method.[2] In these algorithms, most methods focus on the calibration precision, complexity, and robustness. All these algorithms assume that the region of interest (ROI) of the anchor target can be acquired, mostly by human interactive operations. These assumptions match well to cases when a fixed focus length lens is used. Recently, zoom lenses are used more widely so that the vision system can be more adaptable and flexible.[3–6] In these cases, traditional interactive ROI location methods seem to be time consuming and fatiguing. A fully automatic ROI location method should be developed to locate the ROI intelligently and automatically. This paper presents a novel ROI location method based on the Radon transform in the calibration of the zoom lens. This algorithm can also be applied in a vision system composed of fixed-focus lenses.

2 Location Methods for Regions of Interest

2.1 *Traditional Interactive ROI Location Method*

In calibration of fixed-focus lenses, the anchor targets are usually located at several discrete positions (see Fig. 1). At each position, the calibrator uses an interactive method to draw small rectangular regions to specify the ROI. For each circle feature in Fig. 1(b), at least two mouse clicks are required. For a coplanar anchor target array, the calibrator must repeat this process until all ROIs are specified. This repetitive operation seems to be tiresome and inefficient, especially when calibrating zoom lenses where the anchor targets should be positioned at numerous different positions. Even in fixed-focal-lens calibration, to achieve better precision, the ROI specification must also be carried out at several discrete positions to cover the depth of field [see Fig. 1(a)]. Obviously, such tiresome operations can be done automatically through computation. The key problem is to plot out the regions of the features automatically. Then different image processing algorithm in each ROI can be applied to provide detailed calibration information.

2.2 *Automatic ROI Location Algorithm in Ideal Calibration Positions*

As discussed, the key problem is to identify each circle position and plan out each ROI. Because the frequently used anchor target is the circle array on a coplanar plane, a natural idea is to recognize each circle first. The generalized Hough transform is a favorable method for detecting circular or elliptical features; however, its primary problem lies in its great time and memory consumption. Some scholars have studied some algorithms to eliminate the memory and time cost,[7–10] yet we are still a long way from satisfaction. Reference 10 introduces an improved method for detecting three elliptical features on a Pentinum 133-MHz personal computer, but it still takes 15 s. As is known, the Hough transform has an advantage in determining whether circle features exist, yet it seems too complex for acquiring a detailed geometrical parameter (such as a radius or center positions) if we already know that circle features do exist. A dual (horizontal and vertical) projection method

 0091-3286/2002/$15.00

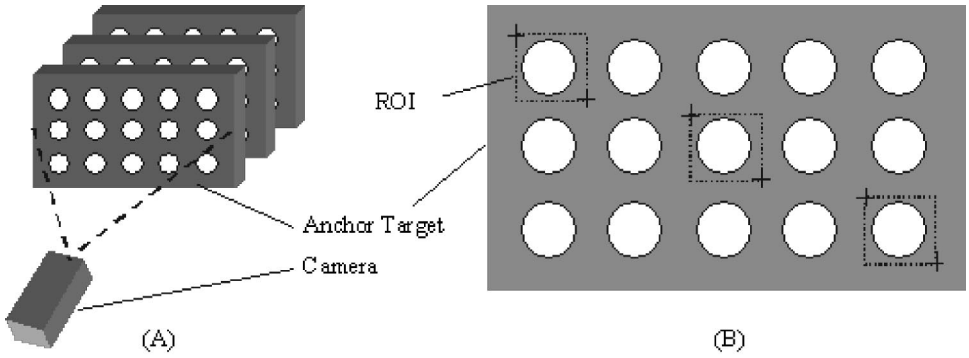


Fig. 1 Traditional ROI location method.

(DPM) followed by the fast comb-shaped signal processing algorithm (FCSPA) is developed here to simplify the ROI location problem.

Suppose that the anchor target can be positioned parallel to the CCD pixel array directions first (see Fig. 2). Then we can project the image along the vertical direction and compute the integration along each pixel line. The integration result becomes zero whenever the projected image crosses position (3) or (3′) in Fig. 3, and reaches its peak at position (1); and at position (2), the value is between the peak and minimum values. Figure 2 shows an original image captured in ideal calibration status (or after careful adjust); Fig. 4 illustrates the result of applying the Sobel algorithm; and Figs. 5 and 6 are the projection integrations along the vertical and horizontal directions. As we can see, the projection results are comb-shaped signals. There is clear division between each ROI and by applying the FCSP algorithm, the ROI location can be dynamically plotted out. As we are not aiming at introducing details of the FCSPA, we just refer to its conclusion and will not provide a detailed explanation. Figure 7 shows the results after applying the DPM and an FCSPA. We can see that each ROI has been easily identified and located. Also the fragmentary circle features around the boundary of the image are intelligently eliminated. The total time consumption for this automatic location of ROIs (30 circle features) is 154 μs on a Pentium 266-MHz PC.

3 Fully Automatic ROI Location Algorithm in Random Calibration Positions

As discussed in Sec. 2.2, the precondition of the DPM and an FCSPA is that the circle array direction must be parallel to the CCD pixel array direction. This requires that the

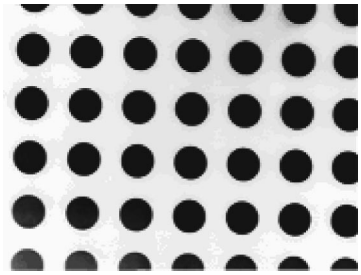


Fig. 2 Original anchor target (inversed).

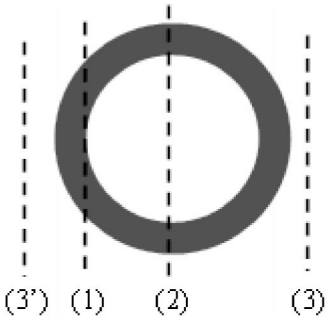


Fig. 3 Projection along a direction.

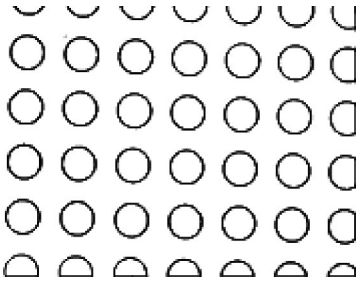


Fig. 4 Results of applying the Sobel algorithm (inversed).

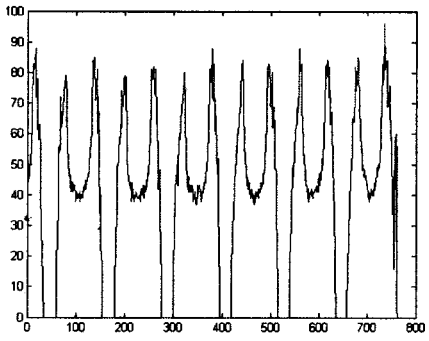


Fig. 5 Projection results along the vertical direction.

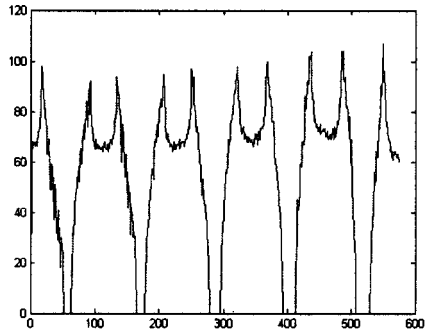


Fig. 6 Projection results along the horizontal direction.

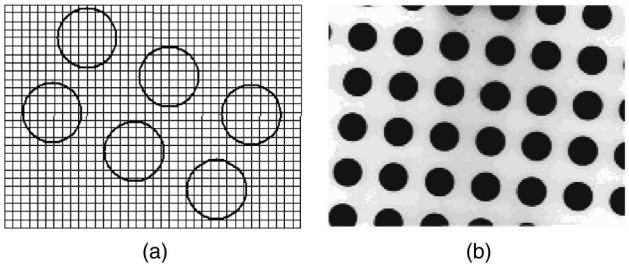


Fig. 8 (a) Random target position and (b) image of general target position.

anchor target be adjusted carefully before calibration; otherwise, this method will fail to locate the ROI. Figure 8 illustrates the random position of the anchor target. Obviously, the DPM and FCSPA will not function well at this task; the projection in both of the dual directions has overlapped and there will be no clear division along the vertical or horizontal direction.

Directed by the idea of the DPM and FCSPA as discussed, if we can find the correct oblique angle between the circle array direction and the camera pixel array direction and rotate the image to adjust it to vertical and horizontal status, then the DPM and FCSPA will still function well. The key is to acquire the oblique angle. Here we employ the Radon transform to get this angle. The Radon transform (RT) can be described as

$$R_\theta(x)=\int_{-\infty}^{+\infty} f(x'\cos\theta-y'\sin\theta, x'\sin\theta+y'\cos\theta)\mathrm{d}y', \tag{1}$$

where

$$\begin{bmatrix} x' \\ y' \end{bmatrix}=\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}. \tag{2}$$

Figure 9 is the RT result of image Fig. 8(b). The abscissa is angle θ ranging from 0 to 180 deg in 1-deg steps. The y coordinate is the image pixel coordinate on axis x'; and the gray-scale level shows the integration results. As we can see, there are two strong peaks in the RT matrix in Fig. 9. The locations of these strongest peaks, at around 85 and 175 deg, correspond to the orientation angle of the original image. Note that there are also two other strong peaks at around 40 and 130 deg, which also stand out to visual perception. For a captured picture, however, only two of the strongest peaks should be considered, and each should correspond to integration orientation along the horizontal and vertical directions.

Careful readers would probably ask what the criterion is for the strongest peak. This is indeed a problem of concern. Although we can easily find the strong peaks in the visual results of the RT, qualitative criteria must be brought forward to measure or describe the strongest peak position in a detailed computation. We here employ a simple target function, the standard deviation value, as the guideline:

$$\sigma_s=\left\{\frac{1}{n-1}\sum_{k=0}^{n-1}[R_\theta^{(i)}(x)-\overline{R_0(x)}]^2\right\}^{1/2}, \tag{3}$$

$$\overline{R_\theta(x)}=\frac{1}{n}\sum_{i=0}^{n-1}R_\theta^{(i)}(x), \tag{4}$$

where $\overline{R_\theta(x)}$ is the mathematical expectation of the vector, $R_\theta^{(i)}(x)$ is the i'th element of the RT vector, and n is the length of RT vector.

Figure 10 shows the RT vector derived at 175, 5, and 10 deg according to Eq. (1), in Figs. 10(a), 10(b), and 10(c), respectively. Obviously the standard deviation value of vector [Fig. 10(a)] will be the maximum of all RT vectors, which corresponds to the strongest peak, namely, corresponding to the rotation angle between the circle arrays and the camera pixel arrays. Then we can infer that the original image has a 5-deg rotation along the x axis. If we rotate the

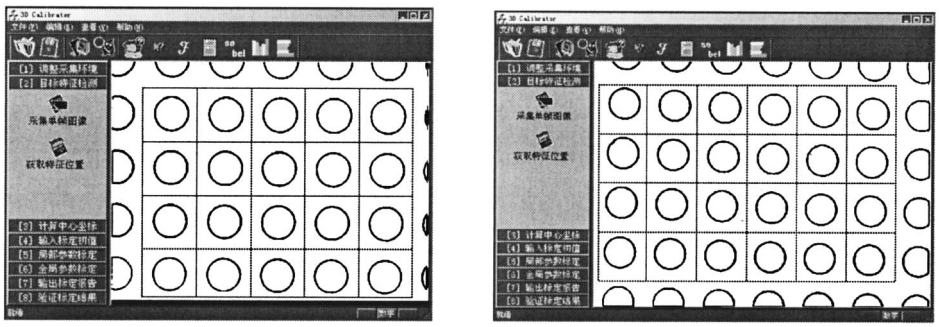


Fig. 7 Automatic location of ROI examples using the DPM and an FCSPA (inversed).

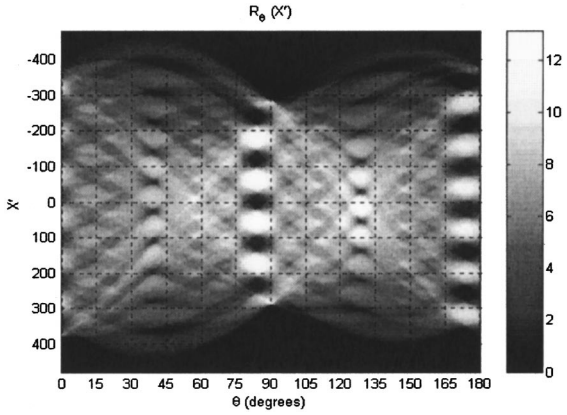


Fig. 9 Radon transform of the calibration target.

original image $-\theta$, we will achieve the ideal horizontal image (as in Fig. 3) on which the DPM and FCSPA can be applied.

The fully automatic ROI location method in camera calibration procession can be described as the two flow charts shown in Figs. 11 and 12. It is effective and robust in any random positions and without requiring strict adjustment.

4 Problems and Improvements

The preceding method can automatically locate the ROI in camera calibration, yet several problems remain. The main question concerns the huge time and memory consumption of this algorithm. We applied this algorithm in the Matlab environment and the Visual C++ environment, respectively, using a Pentium II 266-MHz (128-M RAM) personal computer. The original image is of 768 pixels in width and 576 pixels in height, and the processing time is 121.22 s in Matlab and 115.86 s in Visual C++.

The reason why this method takes so much time and space is closely related to the complexity of RT. For a 768- $\times$576-pixel image, corresponding to each searching angle θ, there must be at least $768\times576\times2$ floating multiplicative operations to compute new x and y coordinates in the x' and y' coordinate system. This is a great consumption source. For all the searching angles ranging from 0 to 180 deg with 1-deg steps, the operation requires at least $768\times576\times2\times180$ multiplicative operations. This also adds to the complexity of the operation.

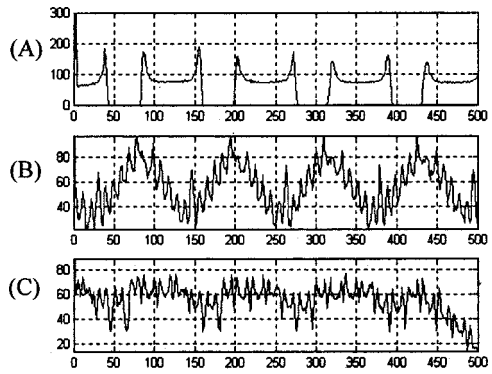


Fig. 10 RT result vectors at different angles.

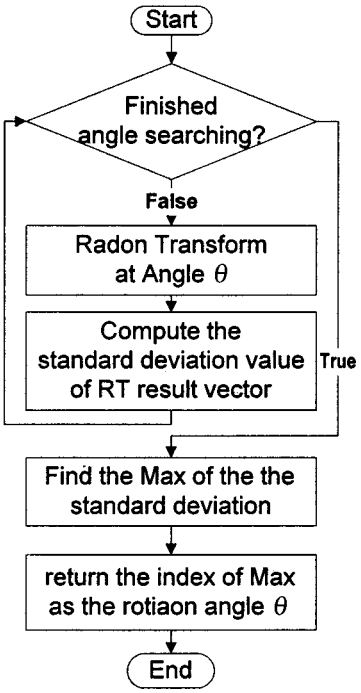


Fig. 11 Searching rotation angle θ.

Consider these two consumption sources. The original image size is 768×576 pixels, but in searching the rotation angle using the RT, the pixel information appears redundant (or strongly correlative) to some extent; that is, even if only part of the image is used to search the angle, the method also succeeds. Thus, a natural idea is to use only the center part of image to compute the RT results. This may reduce the consumption to 1/9 or 1/10 of the original computation.

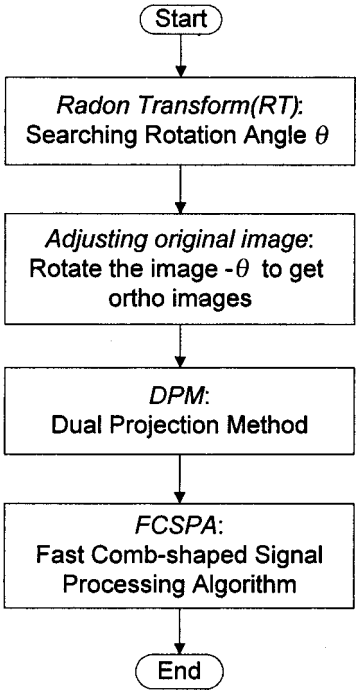


Fig. 12 Fully automatic location of ROI.

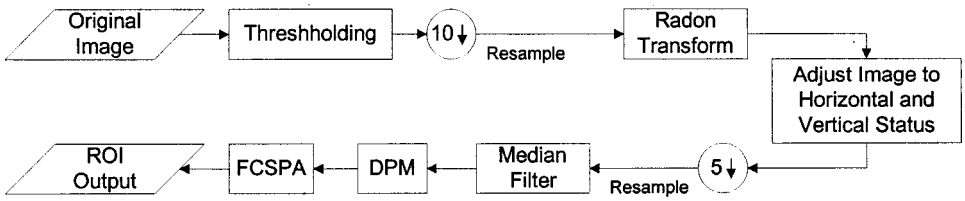


Fig. 13 Using a resampling technique to reduce the huge time and memory consumption.

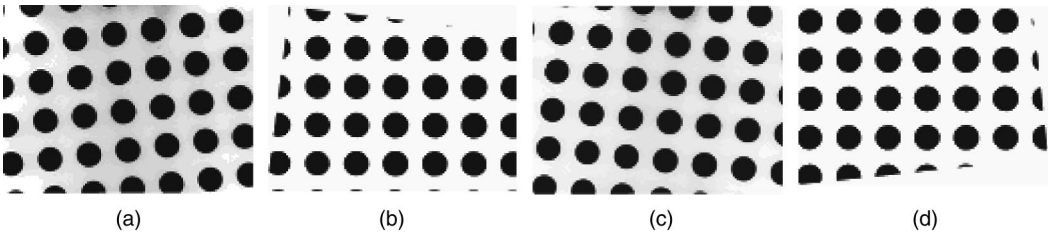


Fig. 14 Original image and adjusted image.

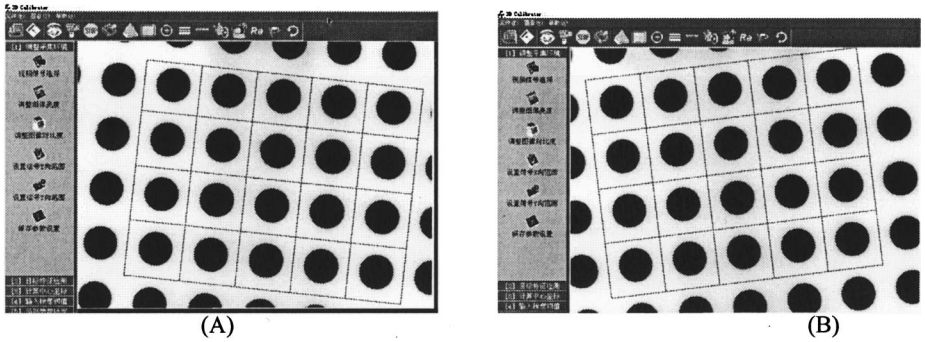


Fig. 15 Original image and adjusted image.

Table 1

		StartX	StartY	EndX	EndY
Vertical line border point of each ROI	1	99	504	148	40
	2	218	516	267	53
	3	337	529	386	65
	4	456	541	504	77
	5	575	554	624	90
	6	695	567	744	103
Horizontal line border point of each ROI	1	99	504	695	567
	2	112	385	707	448
	3	124	268	720	330
	4	137	149	732	211
	5	148	40	744	103

Table 2

		StartX	StartY	EndX	EndY
Vertical line border point of each ROI	1	91	543	33	70
	2	210	529	152	56
	3	328	514	270	41
	4	447	500	389	27
	5	565	485	507	12
	6	684	471	626	0
Horizontal line border point of each ROI	1	91	543	684	471
	2	77	425	670	352
	3	62	307	655	234
	4	48	189	641	116
	5	33	70	626	0

Table 3 Multiplicative times and actual time costs before and after improvement.

	Before Improvement				After Improvement			
Multiplicative operation times	1.59×10^{08}				1.55×10^{05}			
Actual time cost (s)	1	2	3	4	1	2	3	4
	115.86	116.98	114.12	114.66	0.376	0.367	0.368	0.365

This, however, is still far away from a practical application, which counts for 10 to 15 s. If the image in searching the angle (i.e., in RT computation) can be even smaller, the consumption will continue to decrease. However, the image cannot be further reduced because we require at least three or four circle features in the central subimage or the RT will yield the same result along any direction. A good way to settle this is to resample the original image uniformly instead of picking up only the centeal image. Then the resampled image can still represent the whole feature of the original image and it is less redundant after this resampling procession.

A similar resampling can be carried out before the DPM is employed because the original image pixels are redundant in the DPM. Figure 13 shows the improved process. Here the original image was subsampled by 10 and 5 before the RT and DPM were used.

The subsampling rate should not be too high or the DPM and FCSPA may fail to function. Too large a sampling rate will worsen the algorithm's robustness, and the automatic location accuracy will be damaged, which leads to incorrect ROI location. As we know, the accuracy of the feature point in each ROI is crucial to the final calibration performance. Thus, the automatic ROI location algorithm must match the calibration precision. An appropriate subsampling rate must be carefully chosen to compromise between accuracy and efficiency.

Another step to reduce the computation is to limit the searching angle range. Here we limit this to ±10 deg, which can be easily met in actual calibration without careful adjustment.

To achieve robust results, a dynamic threshold is first applied to isolate the background and the circle feature; a median filter is used to eliminate the random noise.

5 Experiments and Conclusions

Figure 14 shows the experimental result of this fully automatic method for ROI location. Figures 14(a) and 14(c) are two original images, and Figs. 14(b) and 14(d) are their adjusted images. Figures 14(a) and 14(c) are processed according the flow chart of Fig. 13. By applying the RT, the rotation angle is derived, then we rotate the original image to orthogonal status [see Figs. 14(b) and 14(d)]. We can use the same method as described in Sec. 2.2 to dynamically locate all the ROIs. The visual location results are shown in Figs. 15(a) and 15(b). Edge points (in pixels) for each ROI are shown in Tables 1 and 2. The multiplicative times and actual time costs before and after the improvement are shown in Table 3.

Experiments show that this improved method can be used in fully automatic location of ROI in camera calibration. This is especially useful with calibration zoom lens cameras.

Acknowledgments

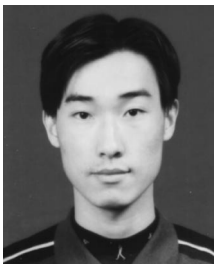
The work reported in this paper was supported by the National Natural Fund Committee under Grant No. 50075063 and by the Key Project Fund Committee of Ministry of Education under Grant No. 99140, China. The encouragement and guidance of Dr. Duan Fajie is appreciated.

References

1. Y. I. Abdel-Aziz and H. M. Karara, ''Direct linear transformation into object space coordinates in close-range photogrammetry,'' in *Proc. Symp. Close-Range Photogrammetry*, pp. 1–18 (1971).
2. R. Y. Tsai, ''A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses,'' *IEEE J. Rob. Autom.* **RA-3**(4), 323–344 (1987).
3. T.-H. Tsai and K.-C. Fan, ''A mathematical model and procedure for calibrating a 3D measurement System with a zoom-lens camera,'' *J. Chin. Soc. Mech. Eng.* **21**(3), 223–227 (2000).
4. H. Kondo and L. Zhang, ''Image edge sharpening with phase correction,'' *IEICE Trans. Inf. Syst.* **E82-D**(8), 1200–1209 (1999).
5. T. Numao, Y. Nakatani, and M. Okutomi, ''Calibration of a pan/tilt/zoom camera by a simple camera model,'' *J. Inst. Image Inf. TV Eng.* **52**(9), (1998).
6. K. Tarabanis, R. Y. Tsai, and D. S. Goodman, ''Calibration of a computer controlled robotic vision sensor with a zoom lens,'' *Comput. Vis. Image Underst.* **59**(2), 226–241 (1994).
7. Y. Chen and F. Qi, ''New detection method of ellipse by using randomized Hough transform,'' *Int. J. Infrared Millim. Waves* **19**(1), 43–47 (2000).
8. ''Memory-efficient circle and ellipse detection in digital images,'' in *IEEE Int. Conf. Systems, Man and Cybernetics* **4711**, 4262–4267 (1995).
9. B. Araabi, N. Babak, and N. Kehtarnavaz, ''Hough array processing via fast multi-scale clustering,'' *Real-Time Imaging* **6**(2), 129–141 (2000).
10. P. Yang, H. Fuqiao, and L. Jiegu, ''Fast extraction of ellipse,'' *J. Shanghai Jiaotong Univ.* **32**(9), 61–63 (1998).

Xinyu Kou received his MS degree in precision measuring technology and instruments from Tianjin University in 2000. He has been involved in several research projects in 3-D computer vision and the opto-electronic inspection field. The paper described here is part of his PhD dissertation at Tianjin University. His current research interests include computer vision, automatic on-spot camera calibration for 3-D vision, image processing and application, and 3-D object reconstruction.

Zhong Wang received his MS degree in precision measuring technology and instruments from Tianjin University in 1991 and he is currently an associate professor at the College of Precision Instrument. His primary research interests are 3-D vision, precision metrology, CAD/CAM, and reverse engineering.



Mingzhou Chen received his BS degree in precision measuring technology and instruments from Tianjin University in 1999 and he is currently an MS candidate in the Department of Precision Instruments. His research interests include image processing, reverse engineering, and 3-D vision. He is now working on the development of optical 3-D measurement systems.

Shenghua Ye directs the State Key Laboratory of Precision Measuring Technology and Instruments, China. He has published over 110 papers in journals at home and abroad. He has conducted the vision inspection research for tens of years and his major research interests include computer vision inspection, 3-D vision technology, and optoelectronic inspection.