A Robust Detection Method of Control Points for Calibration and Measurement With Defocused Images

Wendong Ding, Xilong Liu, De Xu, *Senior Member, IEEE*, Dapeng Zhang, and Zhengtao Zhang

***Abstract*— This paper presents a detection method of control points for the camera calibration and measurement applications, which is robust to defocus. Our method is based on a ground truth, which we call ridge invariance. That is, the positions of broad-brush lines' ridge lines are invariant to defocus blur. First, the ridge invariance is deduced in theory. Then, the methods for ridge point's detection including defocus degree estimation and salience enhancement are deduced. In calibration and measurement experiments, new marks are designed and control points are determined at intersects of the ridge curves. Experiments show that our method can obtain precise calibration and measurement results with images in a large defocus amount range. In the effective defocus amount range, the accuracy of the proposed method almost keeps unchanged to the best values. The proposed method has approximated the same performance as conventional methods at good focus values.**

***Index Terms*— Camera calibration, control point detection, defocused image, point spread function, vision measurement.**

I. Introduction

With the rising trend of vision-based measurement, the cameras are now often used as measuring instruments [33]. Accurate calibrating and measuring with defocused images are very valuable in many applications. In some visual measurement tasks, the focus process is tedious [30] or the image required for visual measurement is not well focused. In some 3-D measurement systems [28], [29], performance of traditional visual measurement methods is often very poor due to defocus blurring in images obtained at different positions.

Errors caused by defocus blurring in calibration are harmful to ensure the measurement accuracy. This is because the camera calibration is a basic step in many visual measurement tasks such as 3-D reconstruction [1], position estimation [2], [3], and robot navigation [4], [5]. In some applications, efforts to reduce the errors may bring many inconveniences when the images are required to be very well focused. For example, if we want to calibrate a camera focusing at a very short distance, the calibration pattern has to be small as well. In fact, it is hard or very expensive to ensure the precision of the pattern itself with such a small size. On the other hand, for large scale measurement tasks such as mapping of street view, the camera is set to focus at a few dozen meters. Then, the calibration pattern has to be large enough for the features to be detected in the focused region, and either the pattern or the camera has to be moved in a large range to obtain the images in various poses.

More difficulties come from the calibration of zoom camera devices which are widely used in intelligent systems such as service robots and unmanned aerial vehicle. For the zoom lens vision system calibration, there is a nonlinear relationship between the intrinsic parameters and the lens settings. The analytical relationship between lens settings and the camera intrinsic parameters is hard to model. What is more, the relationship is often different from one device to another even they have the same product model. Therefore, a large amount of calibrations with dense lens settings are necessary. As discussed in the last paragraph, many calibration blocks in different sizes corresponding to those lens settings are needed. However, this has no practical feasibility in the absence of a high defocus tolerance calibration method with acceptable precision.

The traditional measuring or calibrating methods' vulnerabilities to defocus mainly originate from the features they use for measurement, which can only be extracted and precisely localized in well focused images, such as edges or corners. Aiming at this basic problem, a new feature with the corresponding extracting method is proposed in this paper, which makes control points be precisely localized in defocus images even when the defocus amount is considerably large.

II. Related Works

A. Camera Models

In most works, a perspective camera is usually modeled [26], [32] as a pinhole imaging system if its lens distortion is corrected or negligible. A simple mathematical model is the direct linear transformation [6], [7]. The parameters are modeled as a set of linear equations. Considering the lens distortion, Zhang [8] developed a flexible planar-based

Manuscript received November 03, 2016; revised March 16, 2016; accepted: March 27, 2017. Date of publication July 14, 2017; date of current version September 13, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61503376, Grant 61227804, Grant 61303177, and Grant 61421004, and in part by the special fund of Jiangsu Province for the transformation of scientific and technological achievements under Grant BA2015144. The Associate Editor coordinating the review process was Dr. Jochen Lang. *(Corresponding author: Xilong Liu.)*

The authors are with the Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the University of Chinese Academy of Sciences, Beijing 101408, China (e-mail: xilong.liu@ia.ac.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2017.2709561

calibration technique. This technique uses a pinhole camera model to obtain the initial closed-form solution of intrinsic and extrinsic related matrices. Then two radial distortion terms are added to estimate the maximum likelihood of the refined parameters.

B. Calibration Patterns and Blocks

A traditional way of producing a calibration block is to print the pattern on a piece of paper and attach it on a planar board or display the pattern on a liquid crystal display panel [18], [26], so that the pattern maintains the plane assumption. Well-designed patterns may enhance the performance of calibration. Among various patterns, the planar chessboard is one of the most famous patterns because of the ease of feature point's detection, low cost, and simple structure. Therefore, some recent research on camera calibration has focused on developing patterns [9], [10] with distinctive feature points that can be accurately localized in the images. Placht *et al.* [9] used a circular pattern to avoid the requirement of multiview images. Datta *et al.* [11] presented a camera calibration algorithm for square, circle, and ring planar calibration patterns. An iterative refinement approach is proposed, which utilizes the parameters obtained from traditional calibration algorithms as initialization to relocalize the control points and recompute the camera parameters until convergence.

C. Calibration With Defocused Images

Some researchers have been working on calibration methods robust to defocus images. Baba *et al.* [16] presented a defocus calibration procedure based on a thin lens camera model. The position and the width of blurred features in an image are determined. However, they only added reprojection blur width error of control points into the conventional cost function. The quantitative result of calibration accuracy was not reported in their work. Ha *et al.* [17] proposed a defocused calibration method using a smart phone as a calibration block. They displayed a set of unidirectional binary patterns on the smart phone. 2-D Gaussian deconvolution is simplified to a 1-D Gaussian deconvolution problem with multiple observations and feature point locations are estimated in subpixel accuracy. This method subjects to the smart phone's size limitation and does not suit for the large object distance.

Liu *et al.* [18] estimated the optimal control points in the regions of plausibility determined by distortion bias from perspective distortion, lens distortion, and localization bias from out-of-focus blurring. The accurate detection of control points is formulated as an optimization process to globally search the optimal control points and refining camera parameters. This method needs a long optimize time when the defocus blur amount is serious.

Automatic calibration with neural network is studied by some researchers [12]–[15]. In [12], a calibrating network that is capable of solving the perspective-projection-transformation matrix between the world 3-D points and the corresponding 2-D image pixels was proposed. Compared to traditional methods, neuro-calibrations can be extended to the more difficult problem of calibrating cameras with automated active lenses. In theory, with proper training samples, neural networks can be used for measurement and calibration with defocus images. However, to the best of our knowledge there are still no such reports yet.

Different from existing methods, we start from feature extracting. Our method is based on a ground truth, which we call ridge invariance. That is, the positions of broad-brush lines' ridge lines are invariant to defocus blur. Based on this, new calibration pattern and measurement mark are designed. In both measurement and calibration procedures, the defocus degree of the input images is evaluated first. Based on the evaluated defocus degree, the size of a Gaussian convolution kernel is determined. Convolution with this kernel will make the ridge points more saliency to be detected. Second, the ridge points are obtained and then fit into polynomial curves due to a specific lens distortion model. After this, control points are determined at intersects of the curves. Finally, calibration or measurement results can be gotten. The contribution of this paper is summarized as follows. In measurement and calibration tasks with the designed marks, tolerance of defocus is significantly improved comparing to existing methods. Inconveniences due to defocus problems in calibration can be solved to a great extant. Besides, our method can apply to different kinds of defocus models without the assumption of Gaussian point spread function (PSF).

The rest of this paper is organized as follows. In Section III, the extreme conditions and ridge invariance in the defocused image are given. Then, how to get the saliency condition for the ridge and the estimation of the defocus amount are discussed. Section IV analyses the system's error. The experiments and results are demonstrated in Section V. The conclusion is drawn in Section VI.

III. Ridge in Defocused Image

A. Extremum Conditions and Ridge Invariance

Most defocus methods of camera calibration concentrate on how to recover the precise location of edges from the defocused image. This is usually done as a deconvolution procedure based on the hypothesis that the PSF is a Gaussian function [34]–[38]. As is well-known, deconvolution is not a unique solution problem and a certain solution demands for many additional assumptions. Solutions also highly rely on the supposed convolution kernel. In fact, PSF in real application optical systems is quite complex, so Gaussian approximation may not always be apropos.

In this section, it is shown that the position of broad-brush lines' skeleton, namely, ridge is invariant to defocus with different kinds of PSF. This means that using ridge for camera calibration instead of edge can avoid a deconvolution procedure and the requirement with an explicit formula assumption of PSF. In this section, we demonstrate that the only constraint of PSF is central symmetric which suits for most lens systems. Due to the central symmetric assumption, in the following discussion, the 2-D case is simplified to the 1-D case. And similarly, a central symmetric PSF is simplified to an even function.

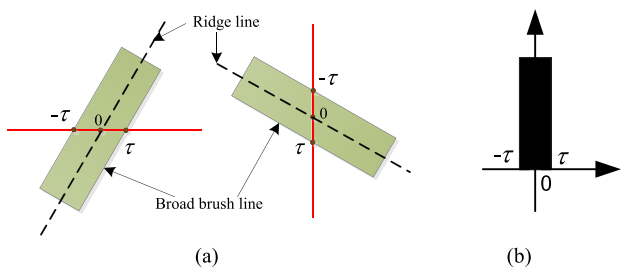


Fig. 1. Brush line and its ridge. (a) Broad-brush line (green strip) and its detection line path marked with red. (b) Gray level of a detection line.

We approximate an arbitrary PSF $f_{\mathrm{ar}}(x)$ with radical bases as

$$f_{\mathrm{ar}}(x) = \sum_{i=1}^{N} \omega_i \phi_{\sigma_i}\left(\left\|x - x_i^p\right\|\right) \tag{1}$$

where N is the number of bases. An even PSF denoted as $f_{\mathrm{even}}(x)$ satisfies

$$\begin{aligned} f_{\mathrm{even}}(x) &= f_{\mathrm{ar}}(x) + f_{\mathrm{ar}}(-x) \\ &= \sum_{i=1}^{N} \omega_i \left[\phi_{\sigma_i}\left(\left\|x - x_i^p\right\|\right) + \phi_{\sigma_i}\left(\left\|x + x_i^p\right\|\right)\right]. \end{aligned} \tag{2}$$

Denote $\phi_{\sigma_i}(\|x - x_i^p\|) + \phi_{\sigma_i}(\|x + x_i^p\|)$ as $\varphi_i(x)$, then $f_{\mathrm{even}}(x)$ can be simplified as

$$f_{\mathrm{even}}(x) = \sum_{i=1}^{N} \omega_i \varphi_i(x). \tag{3}$$

For an ideal broad-brush line, we draw a vertical or horizontal detection line path through it, as shown in Fig. 1. The length of the part covered by the broad-brush line is denoted as 2τ. Locating the original point at the intersection point of the detection line and the ridge line, gray levels change along the detection line can be written as a square wave function

$$g_s(x) = \begin{cases} A_d + B_d, & |x| < \tau \\ B_d, & |x| \geq \tau \end{cases} \tag{4}$$

where $g_s(x)$ is the gray value of the pixel at position x. A_d and B_d are constants, which refers to the amplitude and offset.

The defocus can be modeled as the following convolution:

$$g_b(x) = \left(\sum_{i=1}^{N} \omega_i \varphi_i(x)\right) \otimes g_s(x) \tag{5}$$

where $\otimes$ denotes the convolution.

Due to distributivity of convolution, (5) can be written as

$$g_b(x) = \sum_{i=1}^{N} \omega_i \varphi_i(x) \otimes g_s(x). \tag{6}$$

The integral of $\varphi_i(x)$ is described as $\Phi_i(x)$

$$\int_{-\infty}^{x} \varphi_i(t)dt = \Phi_i(x). \tag{7}$$

Substitute (7) into (6), we have

$$g_b(x) = A_d \sum_{i=1}^{N} \omega_i [\Phi_i(x+\tau) - \Phi_i(x-\tau)] + C_{\mathrm{gb}} \tag{8}$$

Where C_{gb} is a positive constant.

The derivative of $g_b(x)$ is

$$\dot{g}_b(x) = A_d \sum_{i=1}^{N} \omega_i \left[\varphi_i(x+\tau) - \varphi_i(x-\tau)\right]. \tag{9}$$

$\varphi_i(x) = \varphi_i(-x)$ since $\varphi_i(x)$ is an even function. If $x = 0$, we have $\dot{g}_b(0) = 0$ from (9). The second derivative of $g_b(x)$ is

$$\ddot{g}_b(x) = A_d \sum_{i=1}^{N} \omega_i [\dot{\varphi}_i(x+\tau) - \dot{\varphi}_i(x-\tau)]. \tag{10}$$

Considering that

$$\dot{\varphi}_i(x) = -\frac{x - x_i^p}{\sigma_i^2}\phi_{\sigma_i}\left(\left\|x - x_i^p\right\|\right) - \frac{x + x_i^p}{\sigma_i^2}\phi_{\sigma_i}\left(\left\|x + x_i^p\right\|\right). \tag{11}$$

$\dot{\varphi}_i(x)$ is odd since $\varphi_i(x)$ is an even function. It means $\dot{\varphi}_i(-x) = -\dot{\varphi}_i(x)$, so

$$\begin{aligned} \ddot{g}_b(0) &= A_d \sum_{i=1}^{N} \omega_i \left[\dot{\varphi}_i(\tau) - \dot{\varphi}_i(-\tau)\right] = 2A_d \sum_{i=1}^{N} \omega_i \dot{\varphi}_i(\tau) \\ &= 2A_d \dot{f}_{\mathrm{even}}(\tau). \end{aligned} \tag{12}$$

When $\dot{f}_{\mathrm{even}}(\tau) < 0$, we have $\dot{g}_b(0) = 0$ and $\ddot{g}_b(0) < 0$. In this case, $x = 0$ is a maximum extreme point. Now we obtain the extremum conditions and ridge invariance conclusion.

1) *Extremum Conditions:* For an arbitrary PSF, if it is an even function and satisfies $\dot{f}_{\mathrm{even}}(\tau) < 0$, then $x = 0$ is a maximum extreme point of $g_b(x)$.
2) *Ridge Invariance:* The position of the ridge point is invariant to different PSFs and defocus amount.

The above content is aiming at physical phenomenon of image formation rather than the image acquired by the camera. The physical phenomenon of image formation is modeled as a pinhole imaging system sequenced with a convolution of some specific PSFs. The theory is suitable for all cameras that can be modeled in this way regardless of the image acquired is digital or on a film. In the digital condition, the proposed "Extremum Conditions" will be present in the discretization format, namely, difference equations are used instead of differential equations.

B. Saliency Condition

When the image is focused, $g_b(x)$ will equal to $g_b(0)$ for a large range of x. In this case, there is no ridge point. With increasing of the defocus amount, $g_b(0)$ changes to be an absolute extremum.

In order to analyze the relationship between the ridge saliency and blur amount quantitatively, the gradient of $g_b(x)$ at a point near the peak is used to describe the saliency of the peak. The point is denoted as x in the following. In (6),

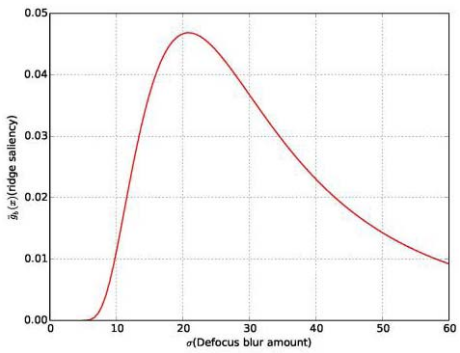


Fig. 2. One curve of saliency function.

$\sum_{i=1}^{N} \omega_i \varphi_i(x)$ is denoted as $P(x)$ for short, and $g_s(x)$ can be written as $A_d[u(x+\tau) - u(x-\tau)] + B_d$, so we have

$$\begin{aligned}\dot{g}_b(x) &= P(x) \otimes \dot{g}_s(x) \\ &= P(x) \otimes A_d[\delta(x+\tau) - \delta(x-\tau)] \\ &= A_d[P(x+\tau) - P(x-\tau)]. \end{aligned} \tag{13}$$

In applications, $P(x)$ is usually a hat-like function such as part of Gaussian function. The defocus amount is fully dependent on the width of hat-like function. They are usually measured by the radius d_a where the value of PSF decreases to smaller than 0.05 times of the center points. For Gaussian PSF, the defocus amounts is quantified as $d_a = 2\sigma$. Taken Gaussian as an example, $P(x)$ is parameterized with σ denoted as $P(x, \sigma)$. Then (13) can be written as

$$\dot{g}_b(x, \sigma) = A_d(P(x+\tau, \sigma) - P(x-\tau, \sigma)) \tag{14}$$

$\dot{g}_b$ is a function of σ, when the variables x and τ are assigned to specific values. One curve of the functions is shown in Fig. 2. In fact, the experiments indicate that the curves generated with different hat-like PSF have a similar shape.

It can be found from Fig. 2 that the peak's saliency at first increases and then decreases. In Section V, we will further show that the trends of the saliency curves are similar for different kinds of PSFs. There exists a best defocus amount where the peak is most salient, so when the defocus amount is smaller than the best value, we should reblur it to make the saliency reach its maximum case. The reblur amount depends precisely on the curve's shape and defocus amount.

As an example, Gaussian function is chosen as PSF and the reblur function. The reblurred signal is denoted as $g_{\mathrm{rb}}(x, \sigma)$, and the saliency function is denoted as $\dot{g}_{\mathrm{rb}}(x, \sigma)$. $g_{\mathrm{rb}}(x)$ can be expanded as

$$\begin{aligned} g_{\mathrm{rb}}(x, \sigma) &= g_b(x) \otimes R_G(x, \sigma_0) \\ &= P_G(x, \sigma) \otimes g_s(x) \otimes R_G(x, \sigma_0) \end{aligned} \tag{15}$$

where $P_G(x, \sigma)$ is a Gaussian function, $R_G(x, \sigma_0)$ is a Gaussian reblur function.

By taking derivative of $g_{\mathrm{rb}}(x, \sigma)$ respective to x, we have

$$\begin{aligned} \dot{g}_{\mathrm{rb}}(x, \sigma) &= P_G(x, \sigma) \otimes \dot{g}_s(x) \otimes R_G(x, \sigma_0) \\ &= P_G(x, \sigma) \otimes A_d[\delta(x+\tau) - \delta(x-\tau)] \otimes R_G(x, \sigma_0) \\ &= P_G\left(x, \sqrt{\sigma^2 + \sigma_0^2}\right) \otimes A_d[\delta(x+\tau) - \delta(x-\tau)] \\ &= P_G(x+\tau, \sigma_1) - P_G(x-\tau, \sigma_1) \end{aligned} \tag{16}$$

where $\sigma_1 = (\sigma^2 + \sigma_0^2)^{1/2}$, $P_G(x, \sigma) = (1/(2\pi)^{1/2}\sigma) \exp\ (-(x^2/2\sigma^2))$.

By taking partial derivative of $P_G(x, \sigma_1)$ to σ_0, then

$$\begin{aligned} &\frac{\partial P_G(x, \sigma_1)}{\partial \sigma_0} \\ &= \frac{\partial}{\partial \sigma_0} \frac{A_d}{\sqrt{2\pi(\sigma^2 + \sigma_0^2)}} \exp\left(-\frac{x^2}{2(\sigma^2 + \sigma_0^2)}\right) \\ &= -\frac{A_d \sigma_0}{\sqrt{2\pi}} \left(\sigma^2 + \sigma_0^2\right)^{-\frac{3}{2}} \exp\left(-\frac{x^2}{2(\sigma^2 + \sigma_0^2)}\right) \\ &\quad + \frac{A_d x^2 \sigma_0}{\sqrt{2\pi(\sigma^2 + \sigma_0^2)}} \left(\sigma^2 + \sigma_0^2\right)^{-\frac{5}{2}} \exp\left(-\frac{x^2}{2(\sigma^2 + \sigma_0^2)}\right) \\ &= -\frac{A_d \sigma_0}{\sqrt{2\pi}} \left(\sigma^2 + \sigma_0^2\right)^{-\frac{3}{2}} \exp\left(-\frac{x^2}{2(\sigma^2 + \sigma_0^2)}\right) \\ &\quad \times \left(x^2\left(\sigma^2 + \sigma_0^2\right)^{-1} - 1\right). \end{aligned} \tag{17}$$

Combining (16) and (17), we can get

$$\begin{aligned} \frac{\partial \dot{g}_{\mathrm{rb}}(x, \sigma)}{\partial \sigma_0} &= \frac{\partial P_G(x+\tau, \sigma_1)}{\partial \sigma_0} - \frac{\partial P_G(x-\tau, \sigma_1)}{\partial \sigma_0} \\ &= -\frac{A_d \sigma_0}{\sqrt{2\pi}} \left(\sigma^2 + \sigma_0^2\right)^{-\frac{3}{2}} (\Theta_{\mathrm{left}} - \Theta_{\mathrm{right}}) \end{aligned} \tag{18}$$

where

$$\Theta_{\mathrm{left}} = \exp\left(-\frac{(x+\tau)^2}{2(\sigma^2 + \sigma_0^2)}\right)\left(\frac{(x+\tau)^2}{\sigma^2 + \sigma_0^2} - 1\right)$$

$$\Theta_{\mathrm{right}} = \exp\left(-\frac{(x-\tau)^2}{2(\sigma^2 + \sigma_0^2)}\right)\left(\frac{(x-\tau)^2}{\sigma^2 + \sigma_0^2} - 1\right).$$

Here we use the Taylor series to approximate the exponential function to a polynomial of degree 1. Let $\Theta_{\mathrm{left}} = \Theta_{\mathrm{right}}$, we can get

$$\frac{2}{3}\tau\delta = \sigma^2 + \sigma_0^2. \tag{19}$$

In practice, we need a space interval to determine whether current point is the extreme point (ridge in 2-D conditions). This interval is taken as δ and its value is suggested to be about half the width of the broad-brush line in the image. For example, δ is set to seven pixels in our experiments. After the defocus amount value σ is determined, the optimal value σ_0 can be gotten according to (19).

C. Defocus Amount Estimation

In order to solve the exact value of σ_0 in (19), we must get the value of σ, namely, the defocus amount of the original image. In this paper, Zhou's model for defocus analyzing is used [21]. Here, a briefly review of this method is given

as follows. Zhou's method is based on the Gaussian gradient ratio which is robust to noise, inaccurate edge location, and interference from neighboring edges. The defocus blur is evaluated at the edge location where step edge is assumed

$$F_s(x) = k_{s1}h_s(x) + k_{s2} \tag{20}$$

$$h_s(x) = \frac{d}{dx}\max(0, x) \tag{21}$$

where F_s is a step function, k_s1 and k_s2 are the amplitude and offset of the edge.

In defocused image, an ideal point scatters into a small area called circle-of-confusion. Its diameter is denoted as d_{cs}, which is an indicator of the blur amount estimation.

The defocus blur is modeled as a convolution of an image in focus with the PSF. Taking Gaussian function $G_{\mathrm{psf}} = (x, \sigma)$ as an example, the standard deviation $\sigma = \lambda_k d_{\mathrm{cs}}$ measures the defocus amount. A blurred edge $I_b(x)$ is then given by

$$I_b(x) = F_s(x) \otimes G_{\mathrm{psf}}(x, \sigma). \tag{22}$$

In order to solve (22), the convolution of $F_s(x)$ is first calculated with a known Gaussian kernel then the ratio between the gradient magnitude of the step edge and the convolution edge is obtained. The ratio reaches the maximum value at the edge location. Using the maximum value, the amount of the defocus blur at the edge location can be computed. For convenience, [21] describes the blur estimation method for 1-D case first and then extends it to 2-D image. The gradient of the blurred edge is denoted as $\nabla I_b(x)$ and the gradient of the convolution edge as $\nabla I_{\mathrm{rb}}(x)$, their expressions are given as [21]

$$\begin{aligned}\nabla I_b(x) &= \nabla[(A_d u(x) + B_d) \otimes G_{\mathrm{psf}}(x, \sigma)] \\ &= \frac{A_d}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)\end{aligned} \tag{23}$$

$$\begin{aligned}\nabla I_{\mathrm{rb}}(x) &= \nabla[I_b(x) \otimes G(x, \sigma_{re})] \\ &= \frac{A_d}{\sqrt{2\pi(\sigma^2 + \sigma_{re}^2)}} \exp\left(-\frac{x^2}{2(\sigma^2 + \sigma_{re}^2)}\right)\end{aligned} \tag{24}$$

where σ_{re} is the standard deviation of the Gaussian convolution kernel.

The ratio R_v of the two gradients is

$$R_v = \left|\frac{\nabla I_b(x)}{\nabla I_{\mathrm{rb}}(x)}\right|_{x=0} = \sqrt{\frac{\sigma^2 + \sigma_{re}^2}{\sigma^2}}. \tag{25}$$

Reference [21] proves that the ratio R_v reaches the maximum at the edge location ($x = 0$), then σ can be solved using (25).

IV. Perspective Error Analysis

In the above sections, the precise position of the ridge line is obtained from the defocused image. However, due to perspective distortion, the ridge line of a broad-brush line in the image may not correspond to the real center line in the world frame. Fortunately, this error is usually very small that can be ignored (much smaller than one pixel).

Let the view angle corresponding to the width of the broad-brush lines be denoted as $\beta_{\mathrm{aov}} = \angle\mathrm{BOC}$. In the plane

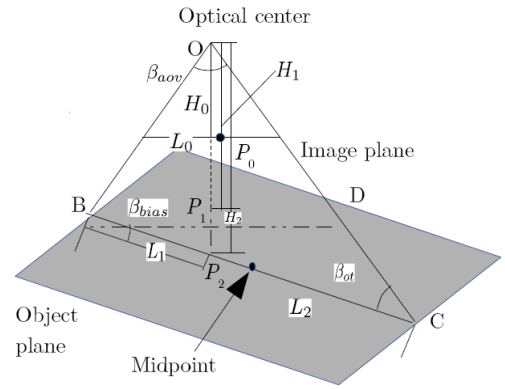


Fig. 3. Projection of the midpoint of broad-brush line.

determined by the optical center O and line segment BC in object plane, the angle between line segment BC and the image plane is denoted as β_{bias}, $\beta_{\mathrm{bias}} = \angle\mathrm{OBC} - \angle\mathrm{OBD}$ as shown in Fig. 3.

The length of line segment BC's projection on the image plane is denote as L_0. The vertical line from the optical center to image plane intersects the image plane at point P_0, the length of OP_0 is H_0, its extended line intersects with line BD at point P_1, and the length of OP_1 is denoted as H_1. OP_1's extended line intersects with line BC at point P_2, the length of OP_2 is denoted as H_2. The length of BP_2 is denoted as L_1, the length of BC is denoted as L_2. According to the geometry relationships, we have

$$L_1 = \frac{H_1 L_0}{2H_0 \cos\beta_{\mathrm{bias}}} \tag{26}$$

$$L_2 = \frac{H_1 L_0 \cos\beta_{\mathrm{bias}}}{H_0}\left(\frac{\tan\beta_{\mathrm{bias}}}{\tan\beta_{\mathrm{ot}}} + 1\right). \tag{27}$$

Assume that the camera can be described with pinhole model and its lens distortion is corrected. If a line segment is on the plane perpendicular to the optical axis of the camera, its image size is not varied with the motion of the line segment on the plane. Without the loss of generality, the line segment is moved to the image central area of the camera for the convenience of analysis. Then the triangle OBD is an isosceles triangle

$$\beta_{\mathrm{aov}} = 2\arctan\frac{L_0}{2H_0} \tag{28}$$

$$\beta_{\mathrm{bias}} + \beta_{\mathrm{ot}} = (\pi - \beta_{\mathrm{aov}})/2. \tag{29}$$

The image error $\Delta\xi_{\mathrm{mp}}$ between the ridge line and center line is defined as

$$\Delta\xi_{\mathrm{mp}} = \frac{H_0}{H_1}(L_2/2 - L_1)\cos\beta_{\mathrm{bias}} = \frac{1}{4}L_0 \sin 2\beta_{\mathrm{bias}} J \tag{30}$$

where

$$\begin{aligned}J &= \tan\ (\beta_{\mathrm{bias}} + \beta_{\mathrm{aov}}/2) - \tan\beta_{\mathrm{bias}} \\ &= \tan\ (\beta_{\mathrm{aov}}/2)\,[1 + \tan\ (\beta_{\mathrm{bias}} + \beta_{\mathrm{aov}}/2)\tan\beta_{\mathrm{bias}}] \\ &\approx \tan\ (\beta_{\mathrm{aov}}/2)(1 + \tan^2\beta_{\mathrm{bias}}) = \frac{\tan\ (\beta_{\mathrm{aov}}/2)}{\cos^2(\beta_{\mathrm{bias}})}.\end{aligned}$$

Taking J and (28) into (30), we have

$$\Delta\xi_{\text{mp}} = \frac{1}{4}L_0 \sin\ 2\beta_{\text{bias}} J \approx \frac{L_0^2}{8H_0}\frac{\sin\ 2\beta_{\text{bias}}}{\cos^2\beta_{\text{bias}}} = \frac{L_0^2}{4H_0}\tan\beta_{\text{bias}}$$

In addition, the length of BD is smaller than BC $\cdot \cos\beta_{\text{bias}}$, and BD $= H_1 L_0/H_0$, so we have

$$L_0 < \frac{H_0 L_2}{H_1}\cos\beta_{\text{bias}}. \tag{31}$$

So

$$\Delta\xi_{\text{mp}} \approx \frac{L_0^2}{4H_0}\tan\beta_{\text{bias}} < \frac{L_0 L_2}{4H_1}\sin\beta_{\text{bias}} < \frac{L_0 L_2}{4H_1}. \tag{32}$$

Typically, H_1 is approximated to the depth of the object plane, $H_1 > 500$ mm; L_2 is the width of the brush line; L_2 is about 10 mm. H_1 is at least 50 times of L_2. Usually, the largest width of the broad-brush line in the image is smaller than 20 pixels, $L_0 < 20$ pixel. In this condition, the error is smaller than 0.1 pixel. $\Delta\xi_{\text{mp}}$ can be ignored because it is always much smaller than one pixel.

Besides the pattern we used in this paper, some other patterns such as circular blob pattern are also suitable for the proposed method. Due to perspective, the detected blob center in the image will shift away. With the analysis of above content, we conclude that when the radius of the blob or the width of the broad-brush line is not very large, errors caused by perspective are small enough to be ignored.

V. Experiment

A. Validation of Ridge Points Detection Procedure

The core of the proposed method is to precisely localize the ridge points. The whole procedure can be summarized as follows. First, the defocus amount of the input image is estimated according to (25). Then, the input image is reblurred to get the best saliency according to (19). In the end, the precise position of the ridge point is obtained by calculating the extreme point according to extreme conditions and ridge invariance. In order to validate the correctness of the proposed method, simulations are designed.

Synthetic images are obtained from a broad-brush line pattern blurred by different kinds of PSFs with different defocus amount. The PSFs used in this experiment include Gaussian function, circle spot function and Ricker wavelet (i.e., Mexican hat wavelet) function. Analytic expressions of the latter two PSFs are given as follows:

1) *Ricker PSF:*

$$\psi(x, y) = \frac{1}{\pi\sigma^4}\left(1 - \frac{x^2+y^2}{2\sigma^2}\right)\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right). \tag{33}$$

2) *Circle Spot PSF:*

$$\psi\ (x, y) = \begin{cases} \dfrac{1}{\pi r^2}, & (x^2+y^2) < r^2 \\ 0, & (x^2+y^2) \geq r^2. \end{cases} \tag{34}$$

Defocus amounts with different PSFs were measured by the radius d_a where the values decrease to smaller than 0.05 times of the center points. For Gauss PSF $d_a = 2\sigma$, for Richer PSF $d_a = 0.96\sigma$, and for circle spot PSF $d_a = r$.

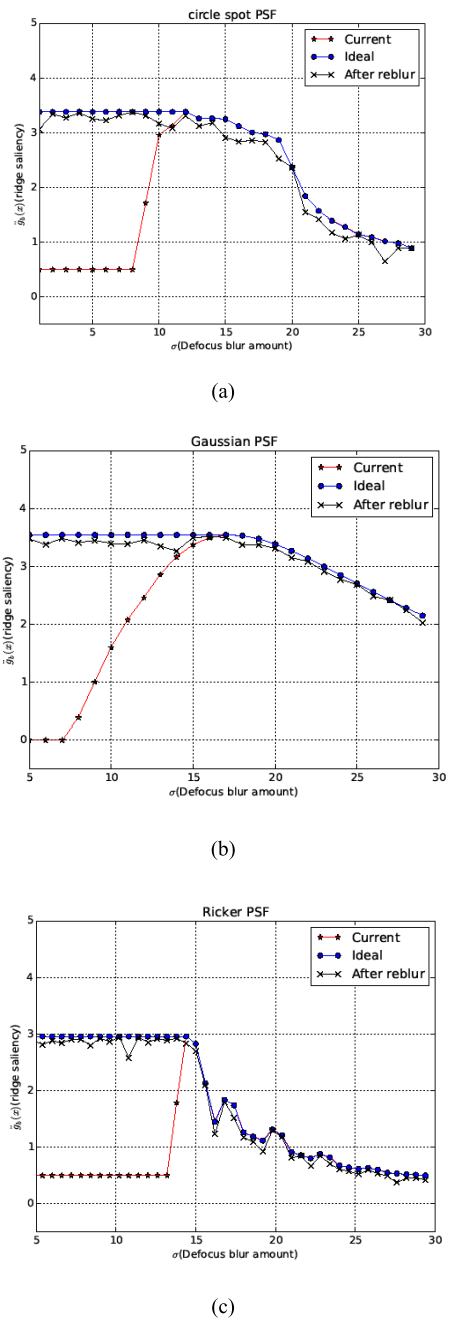


Fig. 4. Performance of saliency enhances procedure for different kinds of PSFs. (a) Circle spot PSF. (b) Gaussian PSF. (c) Ricker PSF.

When changing d_a from 0 to 30 pixels, the performance of the saliency enhance method is shown in Fig. 4. Here, the original saliency $\dot{g}_b(7, \sigma)$, the saliency after enhancement $\dot{g}_b(7, \sigma_1)$, and the theoretical maximum value of saliency $\dot{g}_b(7, \sigma_{\max})$ (see Section III-B) are shown in Fig. 4.

It can be seen from Fig. 4 that the proposed method has good performance in the experiments with the three different kinds of PSFs. When defocus amount d_a is smaller than optimal value, the saliency is significantly enhanced. The reblur procedure pulls up the red curves to the black lines, which are much nearer to the ideal maximum saliency drawn in blue lines. From the red lines we can see that the optimal defocus amount is near to $d_a = 15$ pixels. After the optimal

defocus amount, reblur procedure no longer works and the three curves almost coincide. In these experiments, when d_a is smaller than 25, errors of the ridge points' positions detected are all smaller than 0.1 pixels.

For any centrosymmetric 2-D function whose value diminishing along the radial, namely, a hat-like function has an equivalent radius to a specific Gaussian function. In our experiments, three hat-like PSFs are tested. The circle spot PSF gives a harshest testing of the proposed method. Because it is a discontinuous piecewise function that will never happen in a real system and any smoother PSF will have a better approximation to a Gaussian function.

B. Calibrations With the Proposed Method

In experiments, calibrations are done with the following camera and distortion model.

1) Camera Model: The world coordinate system $\{\mathbf{W}\}$ is established with the origin at the left bottom control point of the planar patterns. Its X_w axis is parallel to the horizontal direction of the patterns, Z_w axis vertical to the pattern plane. In the world coordinate, a point on the planar pattern is denoted as $(x_w,\ y_w,\ 0)$, its corresponding image point is denoted as $(u,\ v)$. For a pinhole camera, the relationship between a point in the world frame and its image projection is given by

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & k_s & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} n_x & o_x & p_x \\ n_y & o_y & p_y \\ n_z & o_z & p_z \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \tag{35}$$

where λ is an scale factor. k_x and k_y are the magnification factors corresponding to horizontal and vertical image axes, respectively, k_s is the skew factor. $(u_0,\ v_0)$ is the image coordinates of the principal point. $[n_x\ n_y\ n_z]^T$, $[o_x\ o_y\ o_z]^T$, and $[a_x\ a_y\ a_z]^T$ represent the unit vectors of the X_w-, Y_w-, and Z_w-axis of the world frame expressed in the camera frame. $[p_x\ p_y\ p_z]^T$ is the position vector of the origin of the world frame described in the camera frame.

When positions of control points are obtained, many methods can be used to solve intrinsic parameters k_x, k_y, k_s, and $(u_0,\ v_0)$ in (35).

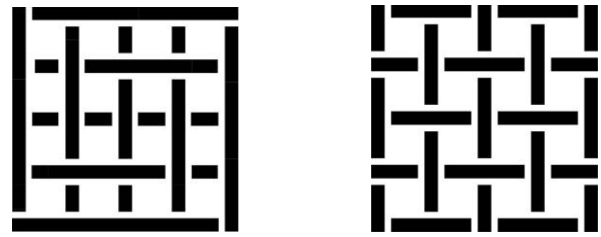
2) Distortion Model: By considering the lens distortion with radial distortion and tangential distortion models, [22] gives a relationship between the nondistorted image point $(x_u,\ y_u)$ and the corresponding distorted image point $(x_d,\ y_d)$

$$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4) \begin{bmatrix} x_d \\ y_d \end{bmatrix} + \begin{bmatrix} 2k_3 x_d y_d + k_4\left(r^2 + 2x_d^2\right) \\ k_3\left(r^2 + 2y_d^2\right) + 2k_4 x_d y_d \end{bmatrix} \tag{36}$$

where $r = (x_u^2 + y_u^2)^{1/2}$, k_1 and k_2 are the radial distortion coefficients and k_3 and k_4 are the tangential distortion coefficients.

Suppose there is a straight line on normalized image denoted as $Ax_u + By_u + C = 0$, let $Ax_u = -By_u - C = t$, then we have

$$\begin{cases} x_u = \dfrac{t}{A} \\ y_u = -\dfrac{t + C}{B} \end{cases} \tag{37}$$



(a) (b)

Fig. 5. Calibration patterns used in simulation and experiment (a) Simulation. (b) Experiment.

$$r^2 = \left(\frac{1}{A^2} + \frac{1}{B^2}\right) t^2 + \frac{2C}{B^2} t + \frac{C^2}{B^2}. \tag{38}$$

Let

$$\alpha = \frac{1}{A^2} + \frac{1}{B^2}, \quad \beta = \frac{2C}{B^2}, \quad \gamma = \frac{C^2}{B^2} \tag{39}$$

then

$$r^4 = \alpha^2 t^4 + \alpha\beta t^3 + \beta^2 t^2 + \alpha\gamma t^2 + \beta\gamma t + \gamma^2. \tag{40}$$

Combining (36)–(38), and (40) and introducing new factors $D_1 \sim D_6$, $E_1 \sim E_6$ for convenience, we can get quantic expressions of x_u and y_u

$$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = \begin{bmatrix} D_1 t^5 + D_2 t^4 + D_3 t^3 + D_4 t^2 + D_5 t + D_6 \\ E_1 t^5 + E_2 t^4 + E_3 t^3 + E_4 t^2 + E_5 t + E_6 \end{bmatrix} \tag{41}$$

where $D_1 \sim D_6$, $E_1 \sim E_6$ are the factors consisting of A, B, C, and $k_1 \sim k_4$.

According to the reports in the literature [23] and [24], most lens distortions can be dominated by the first term of the radial distortion components. It has also been found that any more elaborated modeling not only would be not helpful (negligible when compared with sensor quantization), but also would cause numerical instability [23], [24]. So here we set $k_2 = k_3 = k_4 = 0$, then $D_1 = D_2 = D_6 = 0$, and $E_1 = E_2 = E_6 = 0$. The equation (41) turns into a cubic equation.

C. Calibration Test With Simulation Images

The patterns used for calibrations are composed of broad-brush lines and feature points are detected at the intersections of their ridge lines. The lines do not intersect with each other to avoid ridge detection confusion at the cross areas. If positions of the ridge intersections are not affected, distributions of the broad-brush lines can be set randomly. Fig. 5 shows two different kinds of target patterns deigned for calibrations. In order to show the algorithm's adaptability to different distributions of broad-brush lines, the two patterns are all used in the following experiments. The pattern in Fig. 5(a) is used in simulations, as shown in Fig. 6, and the pattern in Fig. 5(b) is used in experiments with real images, as shown in Fig. 7. Because the precise detections of the ridge points are based on a convolution procedure discussed in Section III according to the proposed "Extreme Conditions" and "Saliency Condition," it is very robust to noise and blurring due to convolution nature. The proposed method may also suitable for the detections of control points in other targets such as circular blob target.

Fig. 6. Simulation image generated with different PSFs, from left to right, the Gaussian PSF, ricker PSF, circle spot, respectively.

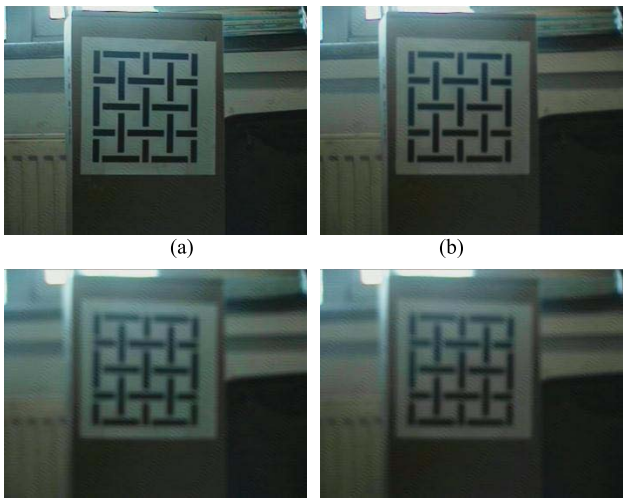


Fig. 7. Images in variant focus values. (a) Focus value set to 6155, it is the best focused. (b) Focus setting at 1000, the image is slightly blurred. (c) Focus setting at 8000, the image is quite blurred. (d) Focus setting at 9000, the image is severely blurred.

However, compared to circular blob target, broad-brush line avoids searching for the extreme points in two directions. So we suggest using broad-brush target for calibration whenever it is permitted in the specific application.

In experiments with simulation images, the accuracy of the proposed control points locating algorithm is evaluated with various blur conditions. Based on the camera model in Section II, we generate a series of images of the calibration pattern. A standard image with a specific pose is generated first. Then its pose change is simulated by performing corresponding perspective transform on the standard image. After that, different kinds of defocus blur are simulated by the convolution with different PSFs.

300 images from random viewpoints within $\pm 30°$ from the front view are generated with the pattern in Fig. 5(a). The image resolution is set to 640×480 pixels. In this simulation, the intrinsic parameters of the camera are set to $k_x = k_y = 3000$, and $u_0 = 320$, $v_0 = 240$. After enhancement with reblur, the ridge lines are detected precisely. Their intersections are taken as control points for calibration. The calibration results are presented in Table I.

It can be seen, when the defocus amount is smaller than the optimal value, calibration errors keep smaller than 0.5% of the ground truth and almost invariant to the defocus amount. After that, the errors increase slightly with defocus amount. When the defocus amount reaches 25, error of k_x and k_y increases to 0.65%. In fact, this is a quite large defocus amount, as shown in Fig. 6.

TABLE I

INTRINSIC PARAMETERS UNDER VARIANT BLUR KERNEL

d_a	3	5	7	9	11	13	15
Gaussian	2904.6	2892.9	2898.3	2899.2	2896.5	2889.3	3126.9
Circle spot	3108.0	2889.3	2901.9	2888.4	2886.6	2895.6	3120.6
Ricker	2892.9	2901.9	3116.1	3111.6	2892.9	3116.1	2874.9
d_a	17	19	21	23	25	27	29
Gaussian	2874.9	2869.5	2860.5	2848.8	2821.8	2726.4	3370.8
Circle spot	2872.2	2857.8	2850.6	2837.1	2816.4	2689.5	2549.1
Ricker	3120.6	2874.9	2851.5	2838.0	2824.5	3288.9	2609.4

TABLE II

PARAMETERS OF THE SONY FCB-ex11D COLOR CAMERA

Item	Parameters
Resolution	640×480 pixel
Lens zooms	10×
Lens focal length	4.2mm-42.0mm
Angle of view	4.6 degree to 46 degree
Serial interface	VISCA protocol
Video codec	MJPG,YUV422
Video rate	30 fps

D. Calibration With Real Images

In real data experiments, the accuracy and the validity of the proposed method are verified in various perspectives and different focal length. A Sony color camera FCB-EX11D is used in the experiment, its detail parameters are shown in Table II. The camera is set to the manual focus mode. The adjustment of the camera's focus length is done by setting a value which range from 0000 to 9999. According to the product manual, the focus length is nearly proportional to focus value with a low level of accuracy. Images with the pattern in Fig. 5(b) in variant focus values are captured, Fig. 7 presents some examples.

The lens of the camera has a 10× zoom, 4.2 mm–42 mm focal length, correspondingly, the angle of the view from 4.6° to 46°. The camera obtained image with resolution at 640×480 pixels. The analogy image signal was converted to universal serial bus format, and FFMPEG library was used to decode the MJPG video, and recovered to the checker board image. The VISCA Protocol was used to control the camera's exposure mode, iris value, exposure time, and focus mode. The frame rate of the camera is 30 fps.

Different from simulations, we cannot get the ground truth in experiments. In order to overcome this problem, the root mean square reprojection error is usually used to evaluate the accuracy. It is computed by the Euclidean distance between the extracted 2-D control points on the image and the reprojected points. The reprojected points are projected from the corresponding point in 3-D space to the image plane through the calibrated camera model with lens distortion. The root mean square reprojection error is

$$E_1 = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(x_{d,i} - \hat{x}_{d,i})^2 + (y_{d,i} - \hat{y}_{d,i})^2} \tag{42}$$

where $(x_{d,i}, y_{d,i})$ is control point on the image and $(\hat{x}_{d,i}, \hat{y}_{d,i})$ is the reprojected point, and E_1 is reprojection error.

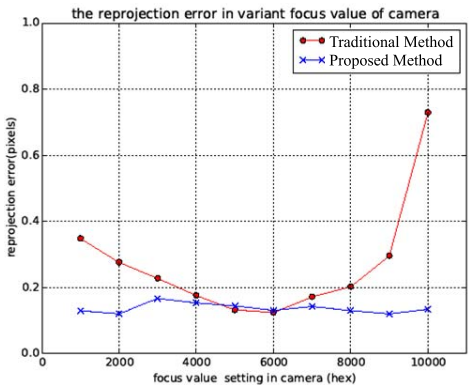


Fig. 8. Reprojection errors with the control points are extracted via the traditional method and the proposed method.

TABLE III

INTRINSIC PARAMETERS IN FIXED SETTING WITH DIFFERENT OBJECT DISTANCE

Distance (mm)		1500	2000	2500	3000	3500
Traditional Method [30]	k_x	3485.233	3472.219	**3470.911**	3488.285	3563.757
	k_y	3584.452	3575.154	**3571.069**	3588.698	3669.584
	u_0	245.184	335.928	**323.201**	341.703	372.823
	v_0	264.044	228.282	**254.465**	236.044	216.203
Proposed method	k_x	3470.059	3470.763	**3470.614**	3470.583	3470.562
	k_y	3570.671	3570.351	**3570.242**	3570.186	3570.151
	u_0	324.841	324.733	**324.697**	324.678	324.666
	v_0	354.838	354.777	**354.786**	354.812	354.845

In order to test our method adequately, two experiments are designed. The first one is to fix the object distance and change focus value. The second one is to fix focus value but change object distance. In these experiments, the proposed method is compared to the state-of-art and most widely used method [11], [26], [30], in which the control point detection is done with edge features.

In the first experiment, the distance between the camera and the target is fixed. The focus value changes from 1000 to 9999. For each focus value, the camera system is calibrated both by the propose method and conventional method. With variant focus values, the average reprojection errors is given in Fig. 8. The best focus value is about 6155 and the conventional method got the best result with the lowest reprojection error. At this value, reprojection errors of the proposed method and the conventional method almost overlap with each other. When the focus value leaves the best value, different performance of the two methods appears gradually. Fig. 8 shows that the more severe of the defocus the conventional gives the worse measure accuracy, while the proposed method has, respectively, stable accuracy for all focus values and has the approximately performance to the best condition of conventional method.

In the second experiment, we fix the focus value, but change the object distance. Here we first deposit the calibrated pattern in the best object distance which is 2500 mm where the image is in focus. Calibrated results are obtained with both the conventional and proposed methods. From Table III, we can see, the results from the two methods are very near. Then the pattern is moved to different object distances, the calibrated results are shown in Table III. The table shows that in the defocus condition the conventional gives worse measurement accuracy, while the proposed method has, respectively, stable accuracy and has the approximately performance to the best condition of traditional method.

Fig. 9. Block used for measurement, focus value set to 0000.

E. Measurement With the Proposed Method

Whenever there are broad-brush lines like patterns in the object to be measured, namely, there are detectable ridge features, the proposed method can also be used for camera pose measurement with large defocus amount which are unacceptable in traditional methods. Here, the experiments are done with a designed mark.

In P*n*P problem, the six degree of freedom can be solved with the camera's intrinsic parameters and the image coordinates and Cartesian coordinates of $n(n \geq 3)$ control points. The key is to obtain the precise image positions of these points. In following comparison experiments, the designed pattern and chessboard are printed on the same paper. As shown in Fig. 9, the left pattern is used for our method and the right one is used of traditional method. In our method, the feature points are the intersections of lines' ridges. In the traditional method, the feature points are the corners of the chessboard. The pattern's pose relative to the camera is computed with P*n*P method after the feature points are extracted.

In the experiments, the distance between the camera and the pattern keeps unchanged while the focus value of the camera changes. For each focus value, the pose of the block is measured. The pose measured by the traditional method with the best focused image was used as the benchmark value. It is (−53.3°, −7.4°, 170.9°, −220.7, −48.3, and 1280.3 mm). Table IV gives the pose values measured by the traditional and proposed methods at each focus index.

The position and orientation errors are demonstrated separately in Fig. 10(a) and (b). It is found from Table IV and Fig. 10 that the orientation errors are (−0.9, 2.1) degree and (−0.5, 1.5) degree, the position errors in z-axis are (8, 29) mm and (5.9, 14.8) mm in the results with the traditional and proposed methods. In other words, the relative position errors are less than 2.27% and 1.16% in the results with the traditional and proposed methods. It is also seen that the precision of the proposed method is almost unchanged to defocus. The proposed method performs better than the conventional method during defocus, especially when the blur

TABLE IV
MEASURED POSE WITH TRADITIONAL AND PROPOSED METHOD

Focus index	Pose (θ_x degree, θ_y degree, θ_z degree, x mm, y mm, z mm) Traditional method	Proposed method
1	-51.2, -6.3, 172.3, -226.7, -50.6,1308.3	-53.0, -7.1, 171.1, -221.9,-49.8, 1286.2
2	-52.2, -6.0, 172.0, -225.5, -50.5, 1307.2	-52.7, -7.3, 170.6, -222.1, -49.7, 1290.5
3	-51.3, -6.7, 169.9, -225.5, -48.4, 1309.3	-51.9, -7.5, 171.4, -223.3, -47.9, 1294.1
4	-53.3, -6.7, 170.3, -225.1, -48.7, 1305.0	-52.8, -7.3, 170.7, -222.9, -48.2, 1289.4
5	-53.7, -7.5, 170.8, -223.5, -48.3, 1295.9	-53.3, -7.2, 170.7, -222.3, -48.1, 1293.2
6	-52.8, -7.9, 172.2, -223.7, -47.8, 1295.5	-52.3, -7.6, 170.9, -223.3, -47.9, 1295.1
7	-52.9, -7.6, 170.8, -221.9,- 48.3, 1289.6	-53.0, -7.9, 171.1, -223.1, -48.4, 1289.4
8	-52.5, -6.9, 170.9, -222.8, -47.7, 1288.2	-52.1, -7.6, 171.1, -223.1, -47.7, 1293.3
9	-53.2, -6.7, 170.7, -222.4, -47.7, 1291.3	-53.2, -6.9, 171.3, -222.2, -47.6, 1292.4

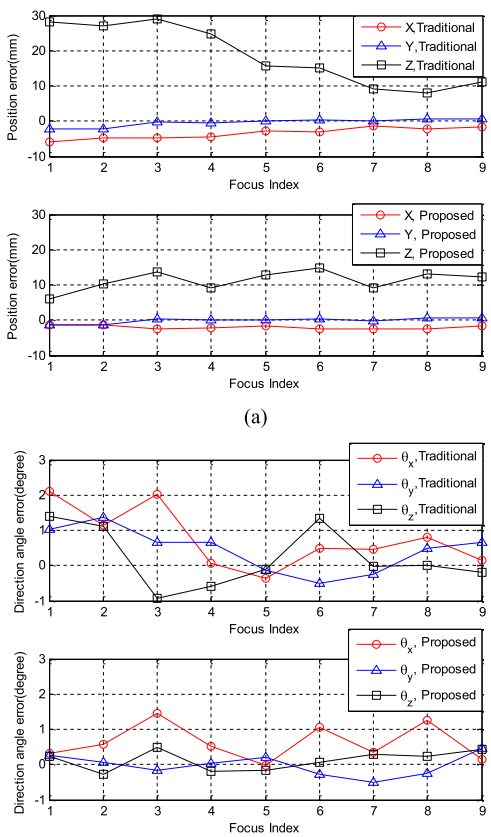


Fig. 10. Errors with the traditional method and the proposed method. (a) Position errors x, y, z. (b) Orientation errors represented in Euler angle θ_x, θ_y, θ_z.

amount is large. In the case of the image is well focused, the proposed method has the approximate performance as the conventional methods.

VI. CONCLUSION

This paper presents a detection method of control points robust to defocus and tests its applications in camera calibration and pose measurement. Experiments with synthesized images confirmed the correctness of the proposed method and real image experiments showed that our method can obtain precise calibration and measurement results with images in a large defocus amount range. In the effective defocus amount range, the accuracy of the proposed method almost keeps unchanging to the best values. Because the proposed method has approximate the same performance as conventional method at good focus values and has a much larger effective defocus amount range, it can surely be used in a wide range of applications.

REFERENCES

[1] Z. Zhang and L.-W. He, "Whiteboard scanning and image enhancement," *Digit. Signal Process.*, vol. 17, no. 2, pp. 414–432, 2007.

[2] M. Bujnak, Z. Kukelova, and T. Pajdla, "A general solution to the P4P problem for camera with unknown focal length," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[3] R. Horaud, B. Conio, O. Leboulleux, and L. B. Lacolle, "An analytic solution for the perspective 4-point problem," in *Proc. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Sep. 1989, pp. 500–507.

[4] A. I. Mourikis, N. Trawny, S. I. Roumeliotis, A. E. Johnson, A. Ansar, and L. Matthies, "Vision-aided inertial navigation for spacecraft entry, descent, and landing," *IEEE Trans. Robot.*, vol. 25, no. 2, pp. 264–280, Apr. 2009.

[5] C. Forster, M. Pizzoli, and D. Scaramuzza, "Svo: Fast semi-direct monocular visual odometry," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2014, pp. 15–22.

[6] Y. Abdel-Aziz, H. Karara, and M. Hauck, "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry," *Photogram. Eng. Remote Sens.*, vol. 81, no. 2, pp. 103–107, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0099111215303086

[7] L. Chen, C. W. Armstrong, and D. D. Raftopoulos, "An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique," *J. Biomech.*, vol. 27, no. 4, pp. 493–500, 1994. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0021929094900248

[8] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.

[9] S. Placht *et al.*, "Rochade: Robust checkerboard advanced detection for camera calibration," in *Proc. 13th Eur. Conf. Comput. Vis.*, 2014, pp. 766–779.

[10] Q. Chen, H. Wu, and T. Wada, "Camera calibration with two arbitrary coplanar circles," in *Proc. 8th Eur. Conf. Comput. Vis.*, 2004, pp. 521–532.

[11] A. Datta, J.-S. Kim, and T. Kanade, "Accurate camera calibration using iterative refinement of control points," in *Proc. IEEE 12th Int. Conf. Comput. Vis. Workshops*, Sep. 2009, pp. 1201–1208.

[12] J. Jun and C. Kim, "Robust camera calibration using neural network," in *Proc. IEEE Region 10 Conf.*, Jun. 1999, pp. 694–697.

[13] M. T. Ahmed, E. E. Hemayed, and A. A. Farag, "Neurocalibration: A neural network that can tell camera calibration parameters," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Sep. 1999, pp. 463–468.

[14] M. Ahmed and A. Farag, "A neural approach to zoom-lens camera calibration from data with outliers," *Image Vis. Comput.*, vol. 20, no. 9, pp. 619–630, 2002.

[15] M. Li and J. M. Lavest, "Some aspects of zoom lens camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 11, pp. 1105–1110, Nov. 1996.

[16] M. Baba, M. Mukunoki, and N. Asada, "A unified camera calibration using geometry and blur of feature points," in *Proc. 18th Int. Conf. Pattern Recognit.*, 2006, pp. 816–819.

[17] H. Ha, Y. Bok, K. Joo, J. Jung, and I. S. Kweon, "Accurate camera calibration robust to defocus using a smartphone," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jun. 2015, pp. 828–836.

[18] J. Liu, Y. Li, and S. Chen, "Robust camera calibration by optimal localization of spatial control points," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 12, pp. 3076–3087, Dec. 2014.

[19] M. Hu, G. Dodds, B. Yuan, and X. Tang, "Robust camera calibration with epipolar constraints," in *Proc. 7th Int. Conf. Signal Process. (ICSP)*, vol. 2. Aug. 2004, pp. 1115–1118.

[20] X. Xu, J. Xia, J. Xiong, and X. F. Xie, "Robust camera calibration with vanishing points," in *Proc. 5th Int. Congr. Image Signal Process.*, Oct. 2012, pp. 931–935.
[21] S. Zhuo and T. Sim, "Defocus map estimation from a single image," *Pattern Recognit.*, vol. 44, no. 9, pp. 1852–1858, 2011.
[22] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Aug. 1997, pp. 1106–1112.
[23] G.-Q. Wei and S. D. Ma, "Implicit and explicit camera calibration: Theory and experiments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 5, pp. 469–480, May 1994.
[24] 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. Robot. Autom.*, vol. 3, no. 4, pp. 323–344, Aug. 1987.
[25] J. A. Hesch and S. I. Roumeliotis, "A direct least-squares (DLS) method for PnP," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 383–390.
[26] S. Zhan and R. Chung, "Use of LCD panel for calibrating structured-light-based range sensing system," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 11, pp. 2623–2630, Nov. 2008.
[27] C. A. Luna, M. Mazo, J. L. Lazaro, and J. F. Vazquez, "Calibration of line-scan cameras," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 8, pp. 2185–2190, Aug. 2010.
[28] S. Zhu and G. Yang, "Noncontact 3-D coordinate measurement of cross-cutting feature points on the surface of a large-scale workpiece based on the machine vision method," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 7, pp. 1874–1887, Jul. 2010.
[29] Y. Li, Y. F. Li, Q. L. Wang, D. Xu, and M. Tan, "Measurement and defect detection of the weld bead based on online vision inspection," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 7, pp. 1841–1849, Jul. 2010.
[30] J. L. L. Galilea, J.-M. Lavest, C. A. L. Vazquez, A. G. Vicente, and B. I. Munoz, "Calibration of a high-accuracy 3-D coordinate measurement sensor based on laser beam and CMOS camera," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 9, pp. 3341–3346, Sep. 2009.
[31] X. Tao, Z. Zhang, F. Zhang, and D. Xu, "A novel and effective surface flaw inspection instrument for large-aperture optical elements," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 9, pp. 2530–2540, Sep. 2015.
[32] H. Bacakoglu and M. S. Kamel, "A three-step camera calibration method," *IEEE Trans. Instrum. Meas.*, vol. 46, no. 5, pp. 1165–1172, Oct. 1997.
[33] S. Shirmohammadi and A. Ferrero, "Camera as the instrument: The rising trend of vision based measurement," *IEEE Instrum. Meas. Mag.*, vol. 17, no. 3, pp. 41–47, Jun. 2014.
[34] Y. Oyamada and H. Saito, "Defocus blur correcting projector-camera system," in *Proc. Int. Conf. Adv. Concepts Intell. Vis. Syst.*, Berlin, Germany, 2008, pp. 453–464.
[35] Y. Oyamada and H. Saito, "Focal pre-correction of projected image for deblurring screen image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Sep. 2007, pp. 1–8.
[36] X. Zhu, S. Cohen, and S. Schiller, "Estimating spatially varying defocus blur from a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4879–4891, Dec. 2013.
[37] M. Cannon, "Blind deconvolution of spatially invariant image blurs with phase," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 24, no. 1, pp. 58–63, Jan. 1976.
[38] Z. Myles and N. da V. Lobo, "Recovering affine motion and defocus blur simultaneously," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 6, pp. 652–658, Jun. 1998.

Wendong Ding received the B.Sc. degree in electronic science and technology from the Wuhan university of Technology, Wuhan, China, in 2013. He is currently pursuing the Ph.D. degree in control science and engineering with the Institute of Automation, Chinese Academy of Sciences, Beijing, China.

His current research interests include computer vision and mobile robot localization.

Xilong Liu received the B.Sc. degree from Beijing Jiaotong University, Beijing, China, in 2009, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences (IACAS), Beijing, in 2014.

He is currently an Associate Professor with the Research Center of Precision Sensing and Control, IACAS. His current research interests include image processing, pattern recognition, visual measurement, and visual scene cognition.

De Xu (M'05–SM'09) received the B.Sc. and M.Sc. degrees in control science and engineering from the Shandong University of Technology, Jinan, China, in 1985 and 1990, respectively, and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2001.

He has been with Institute of Automation, Chinese Academy of Sciences (IACAS), Beijing, China, since 2001, where he is currently a Professor with the Research Center of Precision Sensing and Control. His current research interests include robotics and automation such as visual measurement, visual control, intelligent control, visual positioning, microscopic vision, and microassembly.

Dapeng Zhang received the B.Sc. and M.Sc. degrees from the Hebei University of Technology, Tianjin, China, in 2003 and 2006, respectively, and the Ph.D. degree from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 2011.

He is currently an Associate Professor with the Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing. His current research interests include robotics and automation, in particular, microscopic vision, and microassembly.

Zhengtao Zhang received the B.Sc. degree in control science and engineering from the China University of Petroleum, Dongying, China, in 2004, and the M.Sc. degree in control science and engineering from the Beijing Institute of Technology, Beijing, China, in 2007, and the Ph.D. degree in control science and engineering from the Institute of Automation, Chinese Academy of Sciences (IACAS), Beijing, in 2010.

He is currently a Professor with the Research Center of Precision Sensing and Control, IACAS. His current research interests include visual measurement, micro-assembly, and automation.