Simple and efficient method of calibrating a motorized zoom lens

Yong-Sheng Chen[a,c,1], Sheng-Wen Shih[b,2], Yi-Ping Hung[a,c,*], Chiou-Shann Fuh[a,3]

[a]*Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan*
[b]*Department of Computer Science and Information Engineering, National Chi Nan University, Nantou, Taiwan*
[c]*Institute of Information Science 20, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan*

Abstract

In this work, three servo motors are used to independently control the aperture, zoom, and focus of our zoom lens. Our goal is to calibrate, efficiently, the camera parameters for all the possible configurations of lens settings. We use a calibration object suitable for zoom lens calibration to deal with the defocusing problem. Instead of calibrating the zoom lens with respect to the three lens settings simultaneously, we perform the monofocal camera calibration, adaptively, over the ranges of the zoom and focus settings while fixing the aperture setting at a preset value. Bilinear interpolation is used to provide the values of the camera parameters for those lens settings where no observations are taken. The adaptive strategy requires the monofocal camera calibration only for the lens settings where the interpolated camera parameters are not accurate enough, and is hence referred to as the calibration-on-demand method. Our experiments show that the proposed calibration-on-demand method can provide accurate camera parameters for all the lens settings of a motorized zoom lens, even though the camera calibration is performed only for a few sampled lens settings. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Camera calibration; Lens calibration; Motorized zoom lens; Camera parameters; Active vision

1. Introduction

Motorized zoom lenses have great potential in the applications of active vision [1–3], three-dimensional reconstruction [4–6], and visual tracking [7,8]. In such applications, the aperture, zoom, and focus of the lens can be adjusted to different lighting conditions or to the desired field of view, depth of field, spatial resolution, or focused distance. Although a motorized zoom lens is more flexible and useful than a monofocal lens, it is not an easy job, in general, to calibrate a motorized zoom lens.

The goal of the motorized zoom lens calibration is to determine the relationship between the lens settings (control parameters for the driving motors) and the camera parameters (CPs). Unfortunately, a motorized zoom lens usually consists of some compound lens groups, such as the focusing lens group, the variator lens group, the compensator lens group, and the relay lens group [9–11]. There is no accurate model that describes the paths of the rays passing through the lens. Moreover, various mechanical assembly and driving motors can be used to adjust the aperture, zoom and focus. Consequently, the relationship between the lens settings and the CPs becomes quite complicated [12].

One way to determine the relationship between the lens settings and the CPs is to perform the monofocal camera calibration for each lens setting. That is, each configuration of a zoom lens can be treated as a monofocal lens and its CPs can be calibrated by using well-known camera calibration techniques, such as [13] and [14]. However, this method is extremely inefficient and hardly ever feasible because a motorized zoom lens can usually have a large number of configurations. For example, there will be one million configurations when there are one hundred settings for each of the three motors controlling the aperture, zoom, and focus.

In the past, Tarabanis et al. [15] used a special optical bench to calibrate a zoom lens. They constructed a sparse table storing the CPs calibrated for the sampled lens settings. CPs for other lens settings can be obtained via interpolation. The zoom lens they used was mounted on a robot gripper. In Ref. [15], they also presented a hand-eye calibration method.

* Corresponding author. Address: Institute of Information Science 20, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan. Tel.: +886-2-27883799 ext. 1718; fax: +886-227824814.

E-mail addresses: yschen@iis.sinica.edu.tw (Y.-S. Chen), stone@csie.ncnu.edu.tw (S.-W. Shih), hung@iis.sinica.edu.tw (Y.-P. Hung), fuh@csie.ntu.edu.tw (C.-S. Fuh).

[1] Tel.: +886-2-27883799 ext. 1518; fax: +886-2-27824814.
[2] Tel.: +886-49-2910960 ext. 4832; fax: +886-49-2915226.
[3] Tel.: +886-2-23625336 ext. 327; fax: +886-2-23628167.

Willson and Shafer [12,16] also developed zoom lens calibration techniques. They calibrated two zoom lenses driven by microstepping motors. First, they used an auto-collimated laser for locating the image center. Then, Tsai's method [13] was applied to calibrate the eleven CPs for 25 and 121 regularly sampled lens settings of the two zoom lenses, respectively. For each CP, they approximately modeled the relationship between the CP and the lens settings (zoom and focus) with a bivariate polynomial function. The calibrated CP values of the sampled lens settings were then used to determine the coefficients of the polynomial functions. These polynomial functions can provide CPs for continuous ranges of the lens settings. The order of the polynomial function depends on the lens used and it has been determined empirically in Willson and Shafer's work (from order zero to order five) [12]. One of the major advantages of this method is that the memory consumption is very efficient. Only the coefficients of the polynomial functions need to be stored. Moreover, these functions could give very accurate CPs according to their experiments, in which average prediction error of less than 0.14 pixels was achieved.

In this work, we designed and built a motorized zoom lens with computer-controlled aperture, zoom, and focus. In order to have faster response for the visual surveillance application, we used servo motors, instead of microstepping motors as in Refs. [3,12], for driving the aperture, zoom, and focus of the zoom lens. Also, we designed a calibration object to accommodate various spatial resolutions and field of views, which are unique to zoom lens calibration. The image of the calibration object is often blurred, especially when the depth of field is small. Centroid estimation for the image coordinates of the calibration points was used for dealing with this defocusing problem. In this paper, we fixed the aperture setting and considered only the zoom and the focus settings when determining the relationship between the lens settings and the CPs. For representing this relationship, a two-dimensional table was adaptively constructed by using a simple and efficient method. This table is sparse and each entry stores the calibrated values of the variable CPs as well as the corresponding residual error of calibration. For those lens settings with respect to which the table has not recorded the corresponding CPs, the desired values of the CPs can be obtained via interpolation. To reduce the calibration time, the lens settings for which the camera calibration needs to be performed were adaptively determined according to the accuracy of the CP values interpolated by using the table. Our experimental results have demonstrated that the table constructed by using our calibration-on-demand method can well represent the relationship between the lens settings and the CPs. This table provides accurate CPs of any lens setting for many applications.

This paper is organized as follows. We introduce the camera models of the monofocal and zoom lenses in Sections 2.1 and 2.2, respectively. In Section 3, we propose a simple and efficient adaptive method, the calibration-on-
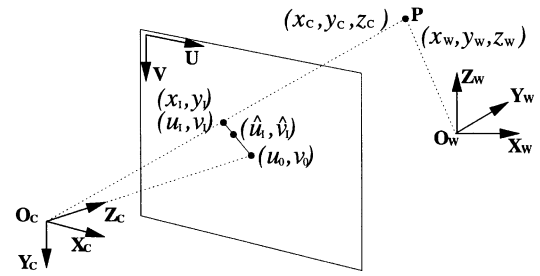

Fig. 1. The perspective projection geometry with the pinhole camera model.

demand method, for calibrating a motorized zoom lens. Section 4 shows the experimental results, and Section 5 gives the conclusions.

2. Camera model

2.1. Monofocal lens

Given a 3-D point in the world coordinate system, the camera model and the associated CPs can predict the projected 2-D image coordinates of the 3-D point in the image coordinate system. For a monofocal lens, the pinhole camera model with radial distortion is considered here. As shown in Fig. 1, the 3-D world coordinates (x_w, y_w, z_w) of a point P can be transformed into the 3-D camera coordinates (x_c, y_c, z_c) as follows:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \mathbf{t}, \tag{1}$$

where $\mathbf{R}$ is a 3×3 rotation matrix that is a function of the three CPs: θ_x, θ_y, and θ_z (X–Y–Z Euler angles), and $\mathbf{t} = [t_x, t_y, t_z]^T$ is a translation vector from the origin of the world coordinate system to the origin of the camera coordinate system. In the camera coordinate system, the origin, O_c, is located at the perspective center and the Z axis is aligned with the optical axis of the lens. By using perspective projection, the coordinates (x_I, y_I) in the image plane is given by the equations

$$x_I = f \frac{x_c}{z_c}, \tag{2}$$

$$y_I = f \frac{y_c}{z_c}, \tag{3}$$

where f is the effective focal length. Then, without considering lens distortion, the coordinates (x_I, y_I) can be transformed into (u_I, v_I) in the image coordinate system by using the equations

$$(u_I - u_0)s_u = x_I, \tag{4}$$

$$(v_I - v_0)s_v = y_I, \tag{5}$$

where s_u and s_v are the horizontal and vertical pixel width

and (u_0, v_0) is the image coordinates of the piercing point where the optical axis pierces the image plane. If the lens distortion is considered, Eqs. (4) and (5) become nonlinear. In this work, only the first coefficient of radial distortion, κ, is considered [13,14], and Eqs. (4) and (5) should be replaced by:

$$(1 - \kappa \rho^2)(\hat{u}_I - u_0)s_u = x_I, \tag{6}$$

$$(1 - \kappa \rho^2)(\hat{v}_I - v_0)s_v = y_I, \tag{7}$$

where $(\hat{u}_I, \hat{v}_I)$ is the actual image coordinates measured in the image and $\rho^2 = [(\hat{u}_I - u_0)s_u]^2 + [(\hat{v}_I - v_0)s_v]^2$.

Totally, there are twelve CPs in this camera model: u_0, v_0, s_u, s_v, f, κ, θ_x, θ_y, θ_z, t_x, t_y, and t_z. Among these CPs, the vertical pixel width, s_v, can be obtained from the specification of the CCD camera. The remaining eleven CPs can be estimated by using Weng et al.'s method [14], provided that a set of known calibration points are observed in the image. In the following, the 3-D world coordinates of the calibration points, (x_w, y_w, z_w), and their corresponding image measurements, $(\hat{u}_I, \hat{v}_I)$, will be referred to as the *calibration data*.

2.2. Zoom lens

2.2.1. Aperture setting

Ideally, we would like to have a zoom lens whose aperture setting does not affect the physical values of the CPs. However, practical lenses suffer a variety of aberrations. Some of the aberrations, such as spherical aberration, coma, and astigmatism, become more serious when the size of the aperture stop is larger [9–11,17]. Furthermore, trying to correct the spherical aberration will usually result in the zonal aberration, which may lead to the focus shift phenomenon where the focal plane moves along the optical axis when adjusting the aperture setting [10,11]. Therefore, changing the aperture setting alone may alter the value of the focal length, even though the zoom and the focal settings remain unchanged. Also, the aberrations may affect the quality of the acquired image and hence the accuracy of the estimated image coordinates of the calibration points. As a result, the aperture setting may indeed influence the estimated values of the CPs.

However, some researchers have found that the influence of the aperture setting is negligible for many applications. For example, Li and Lavest [18] reported that the aperture setting does not change the focal length significantly. Also, our experiment described in Ref. [19] showed that the aperture setting did not much affect the estimates of the CPs. Hence, in this paper, we have chosen to fix the aperture setting while calibrating our motorized zoom lens. With this simplification, the obtained CPs can be stored in a two-dimensional table indexed by the zoom and focus settings, instead of a three-dimensional table indexed by the zoom, focus, and aperture settings. Even if the influence of the aperture setting is not negligible, the variation of the

CPs, when adjusting the aperture setting, is usually quite smooth, as described in Ref. [12]. In that case, we can simply choose a couple of aperture settings and repeat our adaptive calibration method described in Section 3. Another possibility is to extend the proposed method to the additional axis of the aperture setting.

2.2.2. Zoom and focus settings

Once an aperture setting is chosen and fixed, our goal is to determine the relationship between the CPs and the zoom-focus setting. For each combination of the zoom and focus settings, the zoom lens can be treated as a monofocal lens and its CPs can be calibrated individually. When we adjust the zoom or the focus setting, some of the CPs remain unchanged during the adjustment. For example, the pixel width, s_u and s_v, and the extrinsic CPs, θ_x, θ_y, θ_z, t_x, t_y, and t_z, are not supposed to vary. Hence, once these CPs are determined with the initial zoom lens setting, they can be fixed during the remaining procedure of zoom lens calibration.

When adjusting the zoom or the focus setting, the focal length will change accordingly. Thus, f and κ should be estimated for each combination of the zoom and focus settings. Furthermore, the lens groups of the zoom lens are moved forward and backward along the optical axis when we adjust the zoom or the focus setting. This longitudinal movement will make the perspective center, i.e. the origin of the camera coordinate system move along the optical axis relative to the fixed image plane. Consequently, we have to estimate the displacement, Δt_z, of the perspective center on the Z axis for different zoom and focus settings. Besides, the lens groups will be moved back and forth by gradual rotation during the adjustment. If the optical axis deviates from the rotation axis, the image coordinates of the piercing point, u_0 and v_0, may change and they should also be estimated for each combination of the zoom and focus settings.

To achieve a smaller 2-D residual error, all the twelve CPs can be freely optimized during the calibration procedure. However, this will increase the calibration effort and the memory storage of the table. Furthermore, smaller 2-D residual errors do not necessarily result in more accurate estimates of the physical camera parameters. Therefore, our method estimates only five camera parameters, u_0, v_0, f, κ, and Δt_z, in the zoom lens calibration procedure, which will be referred to as the variable camera parameters (VCPs), and fixes all the remaining CPs to be those determined with the initial zoom lens setting.

3. Zoom lens calibration

3.1. Calibration object

There are two major issues when designing the calibration object: the structure of the calibration object and the

calibration patterns to be measured. To accurately estimate the CPs, the calibration data has to be obtained by having the calibration points at different distances from the lens, rather than just at the precisely focused distance. Therefore, the calibration object is usually made of a cube containing calibration patterns on several of its surfaces [5,20] or a calibration plate mounted on a translation stage [13,14,21]. The calibration patterns on the calibration object can be grid lines [5], squares [13,14], or circles [20,21]. Image coordinates of the calibration points can be obtained by measuring the intersections, corners, or centroids of the calibration patterns in the image.

3.1.1. Defocusing problem caused by aperture stop

Without considering aberrations, the points sitting on the focused plane can be projected onto sharp image points in the image plane. Other points deviated from the focused plane will be imaged as filled circles (or, to be more precise, as the shape of the aperture stop) and the acquired image is blurred. In this case, it is not easy to accurately estimate the image coordinates, $(\hat{u}_I, \hat{v}_I)$. This defocusing problem is more serious when we use a larger focal length or a larger size of the aperture stop, due to the resulted smaller depth of field. For zoom lens calibration, this defocusing problem becomes even more serious due to the combination of the following three requirements: (1) accurate CP estimation requires calibration data obtained at different distances, rather than just at the focused distance; (2) camera calibration has to be performed for different focus settings, which will make the focused distances deviate from the calibration object and result in blurred images; and (3) camera calibration has to be performed for different zoom settings, while larger focal length will result in shorter depth of field, thus blur the image.

When the blurring is symmetric, we can still obtain accurate image coordinates, $(\hat{u}_I, \hat{v}_I)$, of the calibration points from the blurred image. Let $(\hat{u}_I, \hat{v}_I)$ be where in the image plane the chief ray projects. The chief ray is defined as the ray passing through the center of the aperture stop. Starting from an out-of-focus point object, the chief ray always project on the center of the circle in the image plane when the aperture stop is at its ideal location. That is, the aperture stop is perpendicular to the optical axis and its center is located at the optical center. For different aperture settings, the projected circles are of different sizes but are all concentric. If the center of the circle is measured as the image coordinates, $(\hat{u}_I, \hat{v}_I)$, these coordinates will not vary with the size of the circle. Consequently, the defocusing problem due to the aperture stop will not influence the calibration result, when the aperture stop is ideally located and none of the aberrations are considered.

For a practical lens, the center of the aperture stop may not locate at the optical center. The estimates of the CPs can be slightly changed when adjusting the aperture stop. More discussions is given in Appendix A.
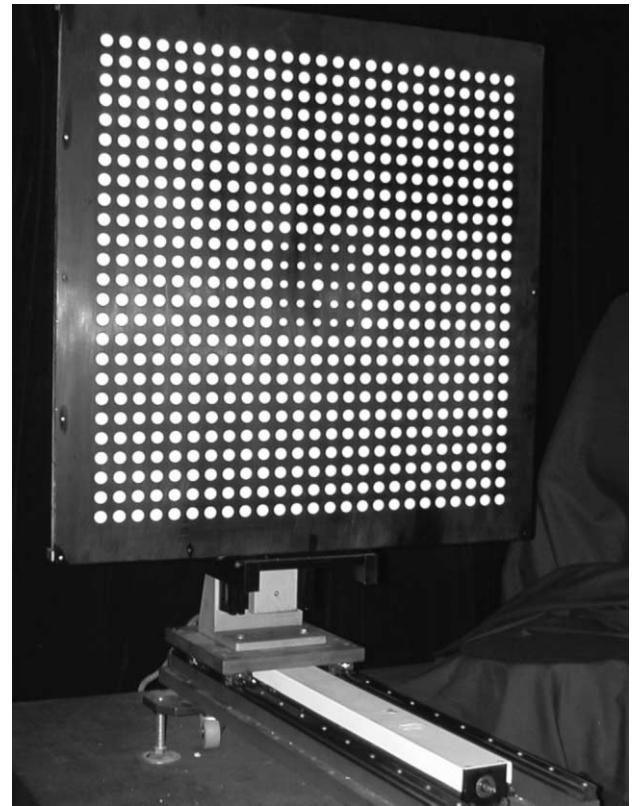

Fig. 2. This figure shows the calibration object we used in this work, which is a calibration plate mounted on a translation stage. The translation direction of the stage is perpendicular to the calibration plate. The calibration plate is a dark board containing 25×25 white circles. The diameters of the circles are 15 mm except for the 24 smaller circles in the middle, whose diameters are 10.5 mm. The distances between the centers of the neighboring circles are 25 mm.

3.1.2. Design of calibration object

In this work, we constructed a black calibration plate mounted on a computer-controlled translation stage, as shown in Fig. 2. The normal direction of the calibration plate, the translation direction of the stage, and the direction of the optical axis of the lens are aligned to be parallel. The calibration plate can be moved along the direction of the stage such that the calibration data measured at different distances can be obtained. On the calibration plate, there are many circles used as the calibration patterns. Each circle projected in the image plane will be symmetrically blurred if the circle is out of focus. Edge detection followed by circle fitting is not suitable to estimate the circle center because there is no sharp edge around the blurred circle. Fortunately, accurate image coordinates of the circle center can be obtained by measuring the centroid of the circle in the image, as mentioned in Section 3.1.1.

For better accuracy of camera calibration, the calibration patterns had better be distributed all over the image and the number of the calibration patterns appearing in the image should be as many as possible [18]. For a zoom lens, it is hard to meet these requirements because of the varying field of view and spatial resolution during the adjustment of the
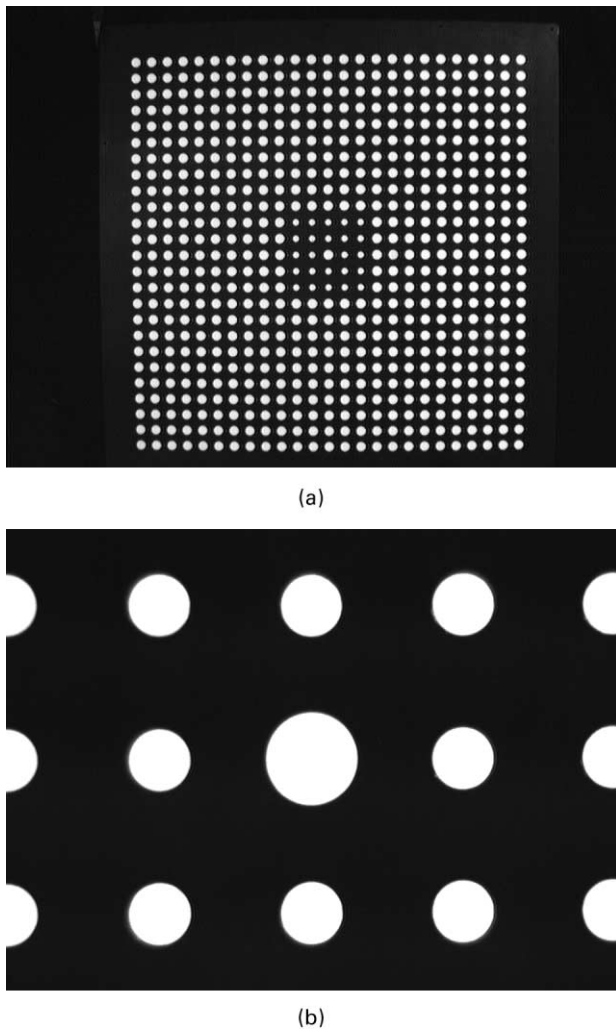
(a)

(b)

Fig. 3. Images of the calibration plate acquired with a zoom lens when its zoom is set to be (a) wide-angle and (b) telephoto.

zoom setting. To remedy this problem, circles of two different sizes were used as shown in Fig. 2. When the zoom of the lens is set to be wide-angle, many larger circles as well as the smaller circles in the middle of the calibration plate appear in the image, as shown in Fig. 3(a). These larger circles can provide more accurate centroid positions estimated in the image. When the zoom of the lens, on the other hand, is set to be telephoto, only a few smaller circles in the middle of the calibration plate appear in the image, as shown in Fig. 3(b). These smaller circles still can provide accurate centroid positions due to larger magnification. Instead of crowding more small circles in the middle region of the calibration plate, we put the same number of circles, per unit area, as what was put in the outside region, to leave larger gap between the circles. This will prevent the neighboring circles from overlapping each other when the out-of-focus circles are seriously blurred in the image.

3.1.3. Image coordinates estimation for calibration patterns

Image coordinates of the calibration point, i.e. the center of the circle, can be easily estimated by using the following image processing techniques: image binarization, blob analysis, and centroid estimation. As the field of view varies, part of the calibration plate may become invisible in the image. However, we still have to identify the circles appearing in the image in order to obtain their corresponding 3-D world coordinates. For easy identification, the larger circle in the middle, which is surrounded by twenty-four smaller circles, is chosen to be the fiducial circle.

The origin of the world coordinate system, O_w, is located at the center of the fiducial circle. The 3-D coordinates of the other circles are then described with respect to this origin. If the camera is properly aligned such that the fiducial circle always appear in the image for all the zoom settings, the task of locating the fiducial circle in the image can be easily accomplished by utilizing the sizes of the circles. Each circle appearing in the image can then be identified and the calibration data, including 3-D world coordinates and 2-D image coordinates, can be obtained.

Notice that if the lens distortion is large enough to severely bias the estimate of the centroid position, one should rectify the image by using the initially estimated lens distortion coefficients before estimating the centroid position. The initial estimates of the lens distortion coefficient, together with those of the other CPs, are obtained by using the estimates of the centroid positions without rectifying the images. The rectification–estimation procedure may have to be repeated a few times until convergence. However, for the motorized zoom lens used in our experiments, we simply use the initial estimates of the CPs because the lens distortion is not severe enough to adopt the above iterative procedure.

3.2. Zoom lens calibration procedure

This section describes the proposed procedure of motorized zoom lens calibration. For the zoom lens used in our experiments, there are three thousand steps for both the zoom and the focus motors, and totally nine million lens settings in combination. To reduce the number of the monofocal camera calibration performed and the storage required to record the VCPs, we have adopted a sampling and interpolation technique over the entire working range of the lens settings. Our goal is then to create a two-dimensional table, indexed by the values of the lens settings, to store the VCPs estimated at the sampled lens settings. The VCPs for unsampled lens settings, that is, where the monofocal camera calibration has not been performed, can be approximated via bilinear interpolation.

Because the values of the VCPs usually vary nonlinearly with respect to the lens settings, a uniformly sampled table might not be able to efficiently represent the relationship between the VCPs and the lens settings. A more efficient sampling strategy is to perform the monofocal camera calibration only when the bilinear interpolation cannot give an accurate enough result. This strategy leads to an adaptive

focus setting

F_{END}	2	10 13		5
			15	8
$\frac{F_{START}+F_{END}}{2}$	3	11	14	6
				7
F_{REF}		0		
F_{START}	1	9	12	4

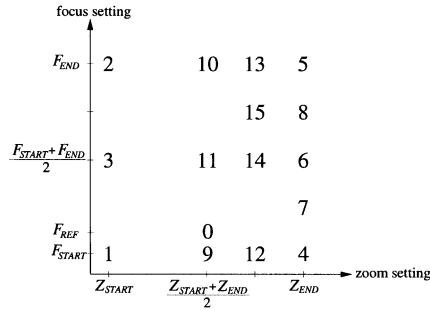Z_{START} $\frac{Z_{START}+Z_{END}}{2}$ Z_{END} → zoom setting

Fig. 4. This figure shows an example of the sample positions. The number in the table means the order at which the camera calibration is performed.

method, referred to as the calibration-on-demand method, which allows adaptive sample positions and locally varying sampling rates. That is, each entry of the VCP table will be adaptively created, by trial and error, to store the VCPs estimated at the sampled lens setting and the corresponding calibration residual error. Our criterion for entry creation is based on the residual error. In the following, we denote the *residual error*, e, as the function $g(\cdot)$, which calculates the average difference between the predicted and measured image coordinates of the calibration points. That is,

$$e = g(I, \beta; z, f),$$

where z denotes the zoom setting, f the focus setting, I the image set used for measuring the image coordinates of the calibration points, and β denotes the VCPs used for predicting the image coordinates of the calibration points.

There are three kinds of residual error used in our zoom lens calibration procedure. The first one is the residual error $\hat{e}$, where $\hat{e} = g(I, \hat{\beta}; z, f)$. In this case, the VCPs $\hat{\beta}$ are calculated by performing the monofocal camera calibration in which the required image coordinates of the calibration points are measured in the image set I. The second one is the residual error $\tilde{e} = g(I, \tilde{\beta}; z, f)$. Here, the VCPs $\tilde{\beta}$ are interpolated values computed by using the values of the VCPs that have already been stored in the table. The last one is the residual error $\bar{e}$, which is the interpolated values by directly using the values of the residual error $\hat{e}$ stored in the table.

Fig. 4 gives an example of how we determine the sample positions. At first, we adjust the aperture stop to match the lighting condition and then fix it in the subsequent calibration. Then, we set the zoom setting to be the middle value within its range, "[Z_{START}, Z_{END}]," that is, $Z_{START} + Z_{END}/2$. Next, the calibration plate is placed at an appropriate distance and the focus setting is adjusted within its entire working range, [F_{START}, F_{END}], until the image of the calibration plate becomes the sharpest. This focus setting is referred to as F_{REF}. A monofocal camera calibration is then performed (as illustrated by sample position number 0 in Fig. 4) to obtain all the CPs. During the following zoom lens calibration procedure, eight CPs, that is, s_u, s_v, θ_x, θ_y, θ_z, t_x, t_y, and t_z, are fixed and the remaining five VCPs, u_0, v_0, f, κ, Δt_z, are to be calibrated.

For each zoom setting to be calibrated, the monofocal camera calibration is first performed at the start and end of the focus setting. For example, when the zoom setting is set at Z_{START} at the next stage, the start and end of the focus setting are F_{START} and F_{END}, as illustrated by sample positions number 1 and 2 in Fig. 4. Both the obtained VCPs $\hat{\beta}$ and the residual error $\hat{e}$ will be stored in the table. Next, the focus setting is set to the middle value of its range. The VCPs $\tilde{\beta}$ at this focus setting can be computed by interpolation with the VCPs obtained previously at the start and end of the focus setting, but we have to evaluate whether this interpolated VCPs, $\tilde{\beta}$, are accurate enough. To evaluate the accuracy of the interpolation, we first acquire the images of the calibration plate and estimate the 2-D image coordinates of the calibration points observed in those images. Then, we compute the residual error $\tilde{e}$ between the observed image coordinates and the predicted positions calculated by using the interpolated VCPs, $\tilde{\beta}$, that is, $\tilde{e} = g(I, \tilde{\beta}; z, f)$, where z and f are set at sample position number 3 shown in Fig. 4. Large $\tilde{e}$ implies that the interpolated VCPs, $\tilde{\beta}$, may not be accurate enough, which means a monofocal camera calibration had better be performed at sample position number 3. In our calibration-on-demand method, the monofocal camera calibration is performed only when the residual error $\tilde{e}$ is larger than a threshold, $k_f \bar{e}$, where k_f is a scaling factor and $\bar{e}$ is the residual error computed via interpolation with the values of the residual error $\hat{e}$ obtained at the start and end focus settings. The above procedure is recursively repeated for each middle focus setting between two calibrated focus settings (for example, see sample positions number 4–8 in Fig. 4, where the zoom setting is set to be Z_{END}) until the interpolated VCPs, $\tilde{\beta}$, are accurate enough, that is, $\tilde{e}$ is small enough. Notice that the decision of whether $\tilde{e}$ is small enough depends on $\bar{e}$, which varies for different focus settings.

For the dimension of the VCP table along the zoom setting, we adopt an adaptive sampling strategy similar to the one used for the focus setting. Initially, the VCPs corresponding to all the focus settings at $z = Z_{START}$ and $z = Z_{END}$ are calibrated by using the procedure described in the previous paragraph, which is illustrated by the sample positions in the first and the last columns shown in Fig. 4. Next, the zoom setting is set to be "$z_{START} + z_{END}/2$" and the focus setting is initially set to be F_{REF}. Then, the images of the calibration plate are acquired and the observed 2-D image coordinates of the calibration points are estimated. With the observed 2-D image coordinates, the residual error $\tilde{e}$ can be computed by using the interpolated VCPs, $\tilde{\beta}$, computed by using the VCPs recorded in the first and last columns in the table. The interpolated VCPs, $\tilde{\beta}$, will be considered as accurate enough if $\tilde{e}$ is small, that is, $\tilde{e} < k_z \bar{e}$, where k_z is a scaling factor. Otherwise, a monofocal camera calibration will be performed to obtain more accurate VCPs. Again, the

above procedure is repeated recursively for each middle zoom setting between two calibrated zoom settings until the interpolated VCPs are accurate enough, that is, $\tilde{e}$ is small enough.

The proposed procedure for motorized zoom lens calibration is summarized below.

Zoom lens camera calibration

procedure intrinsic-camera-calibration (z, f)
  begin
    set the lens settings (z, f) and perform the monofocal camera calibration
    record the VCPs $\hat{\beta}$ $(u_0, v_0, f, \kappa,$ and $\Delta t_z)$ and residual error $\hat{e}$ in the table with indices (z, f)
  end
procedure focus-calibration (f_{start}, f_{end})
  begin
    focus setting $f := (f_{start} + f_{end})/2$
    interpolate the VCPs, $\tilde{\beta}$, from those of f_{start} and of f_{end}
    evaluate the residual error $\tilde{e}$
    if $\tilde{e}$ is not small enough
      begin
        intrinsic-camera-calibration (z, f)
        focus-calibration (f_{start}, f)
        focus-calibration (f, f_{end})
      end
  end
procedure zoom-calibration (z_{start}, z_{end})
  begin
    focus setting $f := F_{REF}$
    zoom setting $z := (z_{start} + z_{end})/2$
    interpolate the VCPs, $\tilde{\beta}$, from those of z_{start} and of z_{end}
    evaluate the residual error $\tilde{e}$
    if $\tilde{e}$ is not small enough
      begin
        intrinsic-camera-calibration (z, F_{START})
        intrinsic-camera-calibration (z, F_{END})
        focus-calibration (F_{START}, F_{END})
        zoom-calibration (z_{start}, z)
        zoom-calibration (z, z_{end})
      end
  end
  begin
    perform the initial camera calibration to obtain the CPs to be fixed: $s_u, s_v, \theta_x, \theta_y, \theta_z, t_x, t_y,$ and t_z
    intrinsic-camera-calibration (Z_{START}, F_{START})
    intrinsic-camera-calibration (Z_{START}, F_{END})
    focus-calibration (F_{START}, F_{END})
    intrinsic-camera-calibration (Z_{END}, F_{START})
    intrinsic-camera-calibration (Z_{END}, F_{END})
    focus-calibration (F_{START}, F_{END})
    zoom-calibration (Z_{START}, Z_{END})
  end

Fig. 5. This figure shows the motorized zoom lens used in our experiments, which is a Fujinon TV zoom lens H6 × 12.5R with its aperture, zoom, and focus driven by three servo motors through pushrod links.

4. Experimental results

4.1. Motorized zoom lens

In this work, we constructed a motorized zoom lens by using a Fujinon TV zoom lens H6 × 12.5R, as shown in Fig. 5. Three servo motors were used for driving the aperture, zoom, and focus of the lens through pushrod links. We adopted servo motor because of its faster response. Each servo motor contains three thousand steps, indicated by the optical encoder. We controlled these servo motors with a PC through the Advantech PCL-832 control cards.

To avoid the hysteresis problem [12] due to the backlash, we always set the motor to the desired setting in the same direction. That is, we first drive the motor to the setting a little smaller than the desired one and then drive it to the desired setting.

4.2. Constructing and assessing the table

In our experiments, the proposed calibration-on-demand method was used to calibrate the motorized zoom lens described in Section 4.1. In the first experiment, the parameters k_f and k_z, which determine the sampling density in the focus and zoom axes, were set to be 1.5 and 2, respectively. A table of 120 entries was created and the sample positions are shown in Fig. 6. The VCPs stored in the table are shown in Fig. 7, where t_{z0} is the value of the fixed CP, t_z, obtained in the initial camera calibration. The mean residual error $\hat{e}$ for the acquired table is 0.19 pixels.

To evaluate the repeatability of the motorized zoom lens, we repeatedly set the zoom and focus settings to each sampled settings previously recorded in the table and then took the images of the calibration palate to check how large the residual error $\hat{e}$ became. In this repeatability experiment, the mean residual error $\tilde{e}$ did increase to 0.44 pixels. However, if we randomly set the focus and zoom settings, which are unlikely to be the sample positions recorded in the table, the residual error $\tilde{e}$ does not increase much further. For example, Fig. 8 shows the residual error $\tilde{e}$ of four
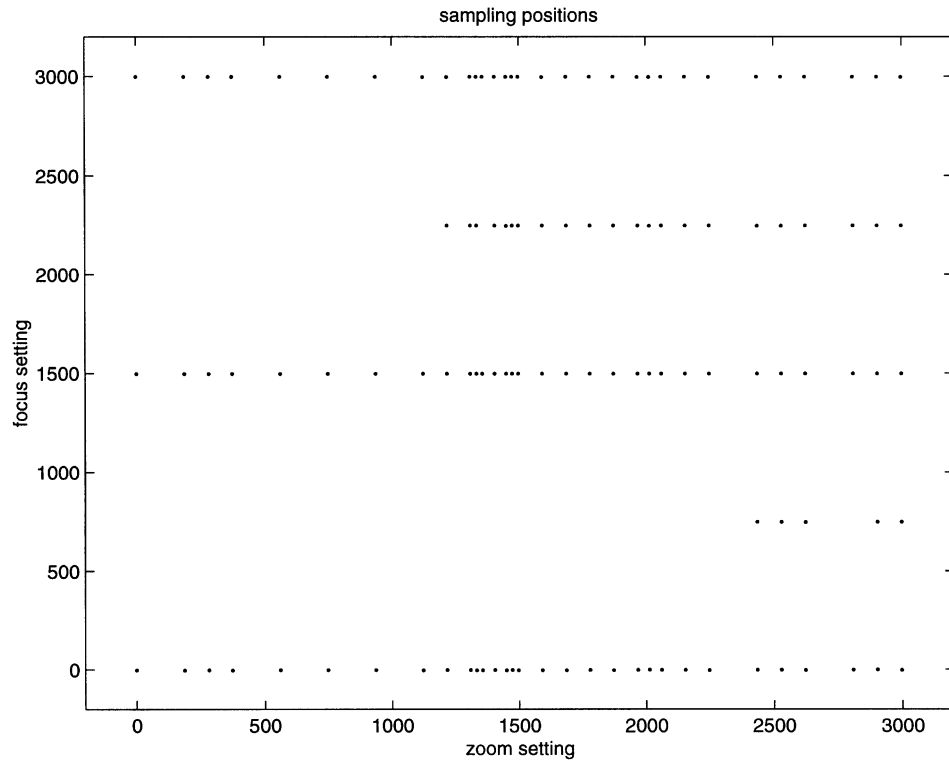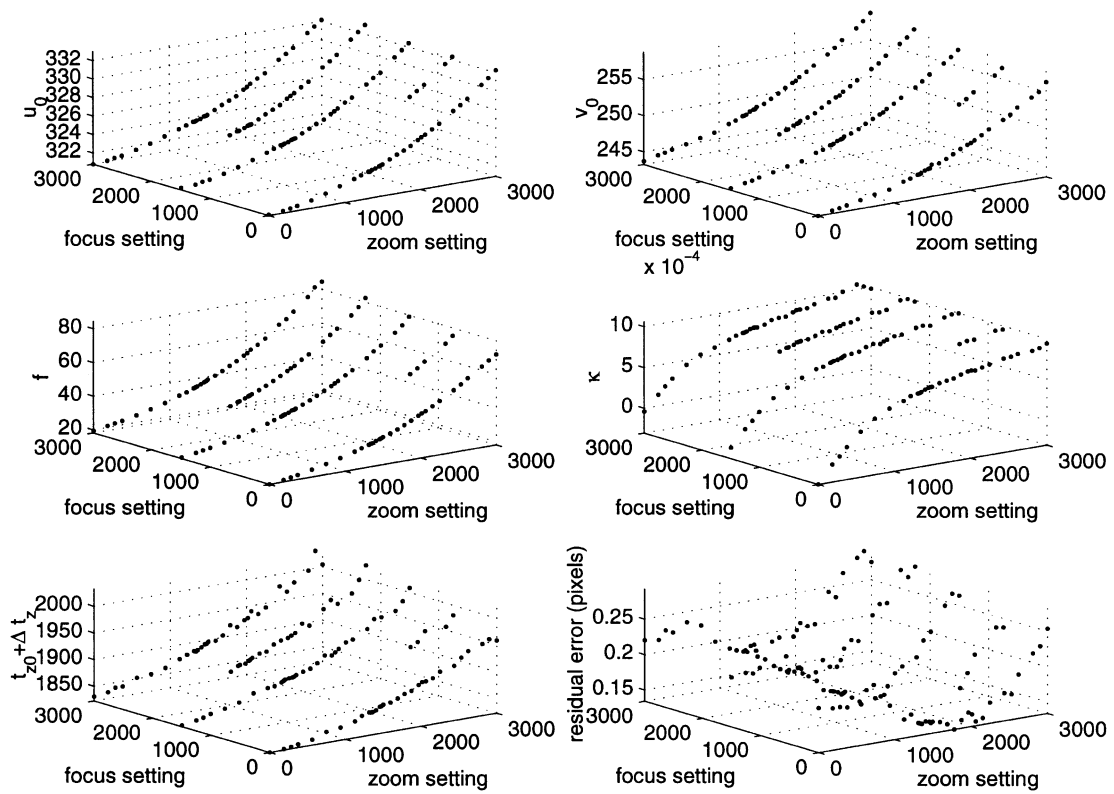
Fig. 6. The sample positions with respect to the zoom and focus settings.

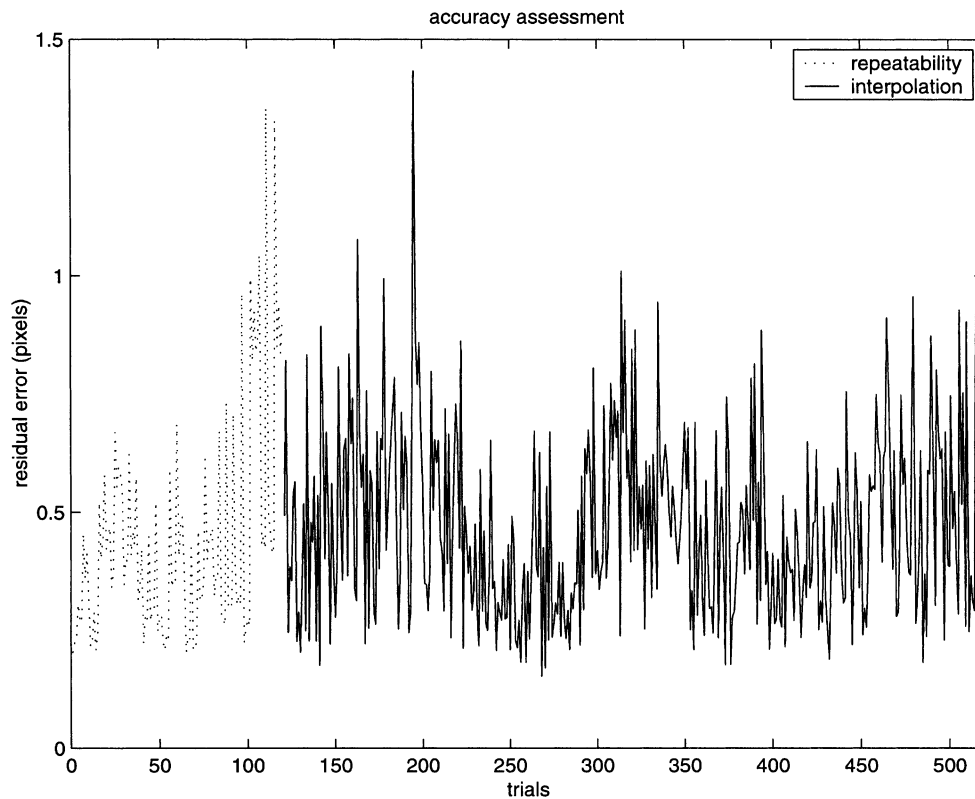Fig. 7. The VCPs $\hat{\beta}$ and the residual error $\hat{e}$ stored in the table.

Fig. 8. The residual error $\tilde{e}$ of repeatability and interpolation are shown in dotted and solid lines, respectively.

hundred random trials, and the mean residual error $\hat{e}$ is 0.475 pixels, which is only 8% increase compared to the residual error $\tilde{e}$ caused by repeatability error.

Another experiment was conducted by using different set of k_f and k_z, which were set to be 2 and 3, respectively. In this case, a table of smaller size (52 entries) was created at the expense of larger error, where the residual error $\tilde{e}$ was 0.54 pixels (that is, a 22.7% increase).

5. Conclusions

In this paper, we have proposed a calibration-on-demand procedure for calibrating a motorized zoom lens, which can adaptively create a table for representing the relationship between the lens settings and the camera parameters. The major advantage of the proposed method is in its ability to reduce, in an adaptive manner, the amount of image acquisition and monofocal camera calibration, which is a time-consuming task, while maintaining the required calibration accuracy. Our experiments have shown that the proposed method can provide accurate intrinsic camera parameters for each lens setting of a motorized zoom lens by performing a much smaller number of monofocal camera calibrations. Our experiment has shown that the average residual error of the camera parameters given by the acquired table (having only 120 entries) is less than half pixels.

Another contribution of this work is that we have

proposed a calibration object which consists of circles distributing on a calibration plate. These circles are of different sizes and have been designed to accommodate various spatial resolutions and field of views, which are unique problems to zoom lens calibration. Another advantage of this calibration object is that accurate image coordinates of the calibration points can be obtained by estimating the centroids of the circles, even when the circles are blurred due to the defocusing problem.

Appendix A

Here, we analyze the influence of the aperture stop on the resulted images. For an out-of-focus point object, the position of the aperture stop and the aperture setting influence the size and position of the resulted circle in at least three ways:

1. *Off optical center.* A zoom lens consists of a few lens groups, and the aperture stop is located somewhere between lens groups. As shown in Fig. 9, if the center of the aperture stop is away from the optical center but still on the optical axis, the aperture stop may asymmetrically shield the pencil of rays, which may then shift the center of the circle.
2. *Off optical axis.* As shown in Fig. 10, if the optical axis does not pass through the center of the aperture stop, the
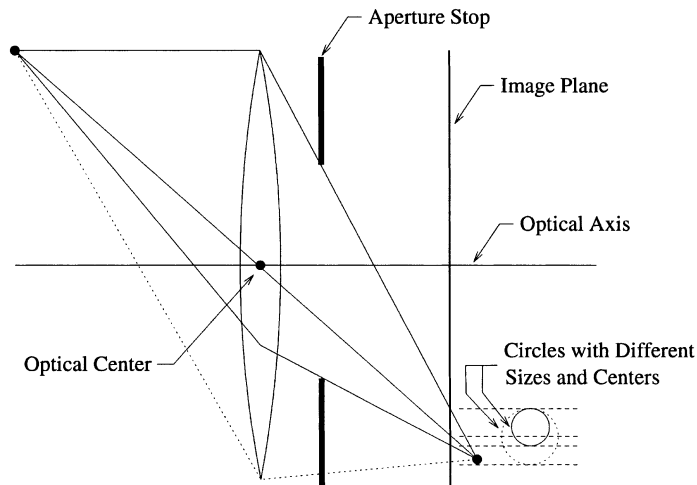
Fig. 9. When the center of the aperture stop is away from the optical center but still on the optical axis, the aperture stop asymmetrically shields the pencil of rays and the center of the resulted circle is shifted.
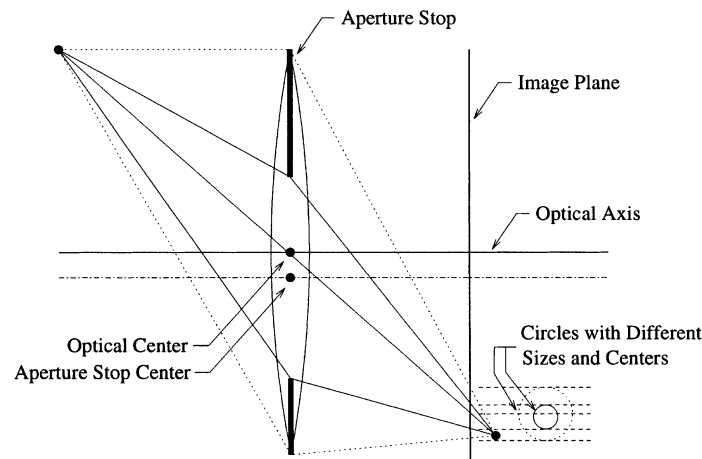

Fig. 10. When the center of the aperture stop does not pass the optical axis, the aperture stop asymmetrically shields the pencil of rays and the center of the resulted circle is shifted.

aperture stop may asymmetrically shield the pencil of rays, which may then shift the center of the circle.

3. *Vignetting*. As shown in Fig. 11, if the size of the aperture stop is too large, the lens mount or other lens groups may asymmetrically shield the pencil of rays. This phenomenon is called vignetting [9,11]. Not only the center but also the shape of the resulted circle are changed.

To sum up, each out-of-focus point is projected onto a filled circle in the image plane. If the center of the aperture stop is not ideally located at the optical center, or if the vignetting phenomenon occurs, the center of the circle will not be the image position where the chief ray projects. Moreover, the shape of the circle will not be symmetric when the vignetting phenomenon occurs.

In the following, we will quantitatively analyze the influence of the aperture stop on the projected image position, without considering vignetting. As shown in Fig. 12, pencil of rays from a point light source pass through the lens with
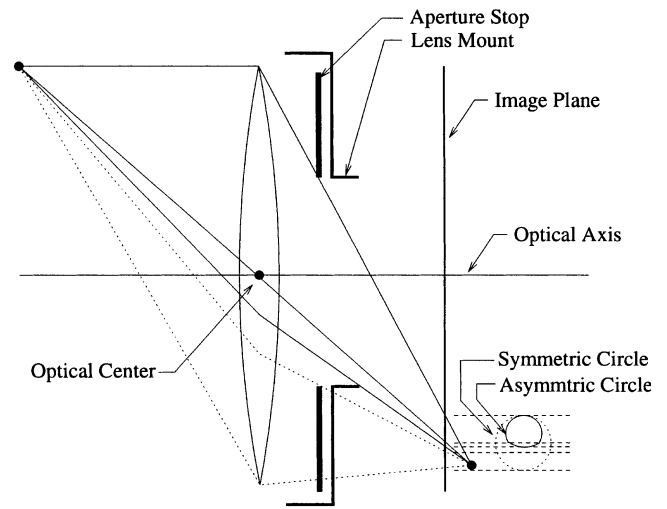

Fig. 11. When the size of the aperture stop is too large, the lens mount may asymmetrically shields the pencil of rays. Not only the center but also the shape of the resulted circle are changed.
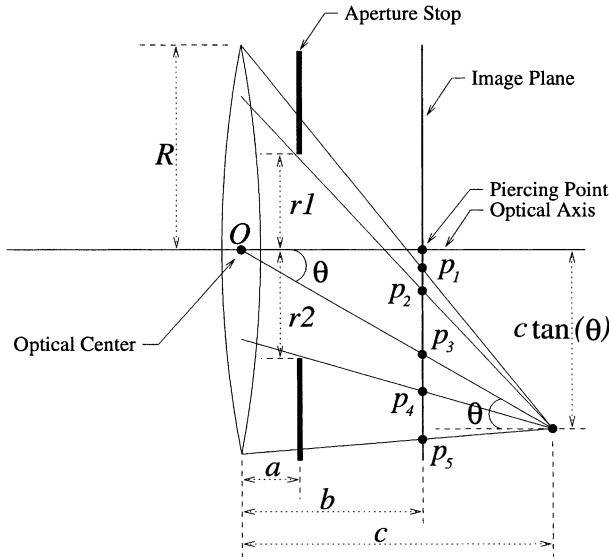
Fig. 12. We quantitatively analyze the movement of the projected circles with or without shielding by the aperture stop.

optical center O. When the aperture stop is open, this point light source projects in the image plane as a circle between p_1 and p_5 centered at p_3. After shielding by the aperture stop, the resulted circle is between p_2 and p_4. Let the origin be the piercing point in the image plane. The vertical coordinates, $y_{p_1}, \ldots, y_{p_5}$, of the points $p_1, \ldots, p_5$ are:

$$y_{p_1} = c\tan(\theta) - \frac{c\tan(\theta) + R}{c}(c - b),$$

$$y_{p_2} = c\tan(\theta) - \frac{c\tan(\theta) + r1}{c - a}(c - b),$$

$$y_{p_3} = c\tan(\theta) - \tan(\theta)(c - b) = b\tan(\theta),$$

$$y_{p_4} = c\tan(\theta) - \frac{c\tan(\theta) - r2}{c - a}(c - b),$$

$$y_{p_5} = c\tan(\theta) - \frac{c\tan(\theta) - R}{c}(c - b),$$

where R is the radius of the lens, $r1$ and $r2$ are the distances from the top and the bottom of the aperture stop to the optical axis, θ is the angle between the optical axis and the ray emerging from O, a, b, and c are the distances from the lens to the aperture stop, the image plane, and the converging point of the pencil of rays from the point light source, respectively.

Assume that the center of the aperture stop is on the optical axis, i.e. $r1 = r2$. The center p_3' of the circle between p_2 and p_4 can be estimated in the image as the middle point between p_2 and p_4:

$$y_{p_3}' = \frac{c(b - a)}{c - a}\tan(\theta) = k \times y_{p_3},$$

where $k = c(b - a)/b(c - a)$. Notice that k is irrelevant to θ. That is, the aperture stop will cause the estimated image

coordinates of the calibration data be enlarged or shrunk with a constant ratio k. This is equivalent to elongating or shortening the effective focal length, f, of the CPs. However, this ratio k depends on c. For the points at different distances, the positions of their projected circles are shifted in different ratios. This makes zoom lens calibration harder when the calibration data is obtained at different distances.

References

[1] E.P. Krotkov, Exploratory Visual Sensing for Determining Spatial Layout with an Agile Stereo Camera System, PhD thesis, University of Pennsylvania, Pennsylvania, April 1987, Technical Report MS-CIS-87-29.

[2] K. Pahlavan, Active Robot Vision and Primary Ocular Processes, PhD thesis, Royal Institute of Technology, Stockholm, Sweden, May 1993.

[3] S.-W. Shih, Y.-P. Hung, W.-S. Lin, Calibration of an active binocular head, IEEE Transactions on Systems, Man, and Cybernetics 28 (4) (1998) 426–442.

[4] A.L. Abbott, Dynamic Integration of Depth Cues for Surface Reconstruction from Stereo Images, PhD thesis, University of Illinois at Urbana-Champaign, Illinois, January 1991.

[5] J.-M. Lavest, G. Rives, M. Dhome, Three-dimensional reconstruction by zooming, IEEE Transactions on Robotics and Automation 9 (2) (1993) 196–207.

[6] G. Surya, M. Subbarao, Depth from defocus by changing camera aperture: A spatial domain approach, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, June 1993, pp. 61–67.

[7] K. Hosoda, H. Moriyama, M. Asada, Visual servoing utilizing zoom mechanism, Proceedings of the International Conference on Robotics and Automation, Nagoya, Aichi, Japan, May 1995, pp. 178–183.

[8] J.A. Fayman, O. Sudarsky, E. Rivlin, Zoom tracking, Proceedings of the International Conference on Robotics and Automation, Leuven, Belgium, May 1998, pp. 2783–2788.

[9] R. Kingslake, Lens Design Fundamentals, Academic Press, San Diego, 1978.

[10] R. Kingslake, A History of the Photographic Lens, Academic Press, San Diego, 1989.

[11] W.J. Smith, Modern Lens Design, McGraw-Hill, New York, 1992.

[12] R.G. Willson, Modeling and Calibration of Automated Zoom Lenses, PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania, January 1994, Technical Report CMU-RI-TR-94-03.

[13] R.Y. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, IEEE Journal of Robotics and Automation RA-3 (4) (1987) 323–344.

[14] J. Weng, P. Cohen, M. Herniou, Camera calibration with distortion models and accuracy evaluation, IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (10) (1992) 965–980.

[15] K. Tarabanis, R.Y. Tsai, D.S. Goodman, Calibration of a computer controlled robotic vision sensor with a zoom lens, CVGIP: Image Understanding 59 (2) (1994) 226–241.

[16] R.G. Willson, S.A. Shafer, Precision imaging and control for machine vision research at Carnegie Mellon University, Proceedings of SPIE Conference on High-Resolution Sensors and Hybrid Systems, San Jose, CA, February 1992, vol. 1656, pp. 297–314.

[17] C.C. Slama (Ed.), Manual of Photogrammetry 4th ed, American Society of Photogrammetry, 1980. Location: Falls Church, Virginia.

[18] M. Li, J.-M. Lavest, Some aspects of zoom lens camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (11) (1996) 1105–1110.

[19] Y.-S. Chen, S.-W. Shih, Y.-P. Hung, C.-S. Fuh, Camera calibration

with a motorized zoom lens, Proceedings of the International Conference on Pattern Recognition, Barcelona, Spain, September 2000, vol. 4, pp. 495–498.

[20] J. Heikkilä, O. Silvén, A four-step camera calibration procedure with implicit image correction, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Puerto Rico, June 1997, pp. 1106–1112.

[21] S.-W. Shih, Kinematic and Camera Calibration of Reconfigurable Binocular Vision Systems, PhD thesis, National Taiwan University, Taipei, Taiwan, June 1996.