PERGAMON

Pattern Recognition 36 (2003) 1155–1164

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

A new easy camera calibration technique based on circular points☆

Xiaoqiao Meng[a,*], Zhanyi Hu[b]

[a] *Department of Computer Science, University of California, Los Angeles, CA 90095, USA*
[b] *National Laboratory of Pattern Recognition, Institute of Automation, Beijing 100080, China*

Received 16 June 2000; received in revised form 28 March 2001; accepted 18 March 2002

Abstract

Inspired by Zhang's work on flexible calibration technique, a new easy technique for calibrating a camera based on circular points is proposed. The proposed technique only requires the camera to observe a newly designed planar calibration pattern (referred to as the model plane hereinafter) which includes a circle and a pencil of lines passing through the circle's center, at a few (at least three) different unknown orientations, then all the five intrinsic parameters can be determined linearly. The main advantage of our new technique is that it needs to know neither any metric measurement on the model plane, nor the correspondences between points on the model plane and image ones, hence the whole calibration process becomes extremely simple. The proposed technique is particularly useful for those people who are not familiar with computer vision. Experiments with simulated data as well as with real images show that our new technique is robust and accurate.

Keywords: Computer vision; Projective geometry; Circular points; Camera calibration; 3D reconstruction

1. Introduction

The computation of camera's intrinsic parameters, which is usually called camera calibration, is one of the most important issues in computer vision [1,2]. A lot of literature on camera calibration has appeared during the last decade [3,4]. With the increasing popularity of cameras used in offices and families, many people who are not experts in computer vision look more and more concerned with an easy and cheap calibration technique to help them to perform vision tasks.

To this end, Zhang recently proposed a flexible camera calibration technique by replacing an expensive classical calibration grid with a planar pattern [3]. Zhang's technique needs to print a dotted sheet (Fig. 1) as the model plane, and the Euclidean coordinates of every dot on the model plane should be measured accurately. After taking a few images of the model plane at different orientations by moving either the model plane or the camera, the homographies between the model plane and its projections can be determined, then the camera's intrinsic parameters can be derived linearly from these homographies. Zhang's technique is easy to apply, and its accuracy is generally higher than self-calibration. Hence it is particularly desirable for desktop vision systems (DVS) which are applied in offices and families. However a major drawback in Zhang's technique is that it requires the user to manually establish correspondences of the projected corners between different images. This process is time-consuming and inconvenient to users, in particular to those who are not familiar with computer vision. To overcome this drawback, a novel model plane (Fig. 1) is designed and used in this paper, it is composed of a circle and a pencil of lines passing through the circle's center (referred to as the model circle and the model lines later on). Employing this new model

☆ An earlier version of the paper appeared in BMVC'00.

* Corresponding author. Tel.: +1-310-206-3091; fax: +1-310-794-5057.
E-mail address: xqmeng@cs.ucla.edu (X. Meng).

0031-3203/03/$30.00
PII: S0031-3203(02)00225-X

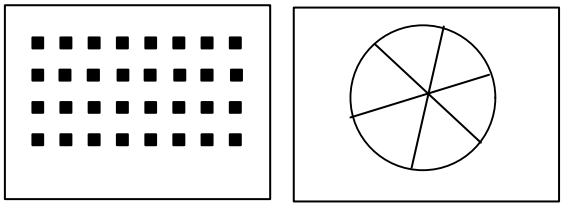

Fig. 1. Zhang's calibration pattern (left) and our proposed calibration pattern (right).

plane, our calibration technique needs neither physical measurements on the model plane, nor any correspondence information, and the whole calibrating process can be done totally automatically.

2. Theory

2.1. Camera model and concept of circular points

The camera is modeled as a pinhole one. The following notation is used in this paper: an image point is denoted by $\mathbf{m} = [u, v]^{\mathrm{T}}$, a 3D point is denoted by $\mathbf{M} = [X, Y, Z]^{\mathrm{T}}$, and their homogeneous coordinates (or projective coordinates) are denoted by $\tilde{\mathbf{m}} = [u, v, t]^{\mathrm{T}}$ and $\tilde{\mathbf{M}} = [X, Y, Z, t]^{\mathrm{T}}$, respectively. Then the imaging process from a 3D point $\mathbf{M}$ to its image $\mathbf{m}$ can be expressed as

$$s\tilde{\mathbf{m}} = \mathbf{K}[\mathbf{R} \quad \mathbf{t}]\tilde{\mathbf{M}}, \tag{1}$$

where s is a non-zero scale factor, $\mathbf{R}$, $\mathbf{t}$ are the rotation matrix and translation vector from the world system to the camera system, respectively. $\mathbf{K}$ is the camera matrix with the following explicit form [4]

$$\mathbf{K} = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We will first consider that the world is embedded in a 3D projective space. In this projective space, points satisfying the equation $t = 0$ are called *points at infinity*. They form the *plane at infinity* which is a 2D projective subspace embedded in the 3D projective space. In this plane at infinity, points satisfying $\tilde{\mathbf{M}}^{\mathrm{T}}\tilde{\mathbf{M}} = 0$ constitute the absolute conic $\boldsymbol{\omega}$. By using Eq. (1), we can easily verify that the image of absolute conic $\boldsymbol{\omega}$ (IAC for short) is the conic $\mathbf{K}^{-\mathrm{T}}\mathbf{K}^{-1}$. This indicates that IAC encapsulates all the information on the camera's intrinsic parameters. Hence if we have determined IAC, we can easily derive all the intrinsic parameters, e.g., via Cholesky factorization.

Without loss of generality, we assume that the model plane lies on the X–Y plane in the world coordinate system, so the equation of the model plane is $Z = 0$. Let us denote

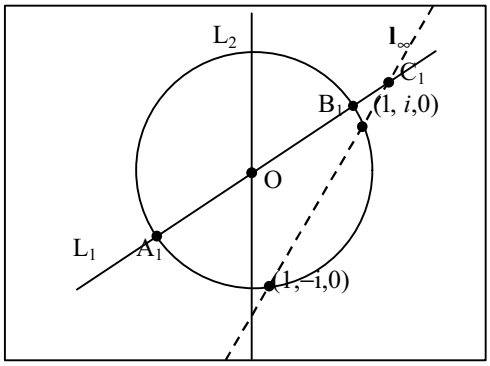

Fig. 2. The model plane.

the ith column of the rotation matrix $\mathbf{R}$ by $\mathbf{r}_i$, we have

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3 \quad \mathbf{t}]\begin{bmatrix} x \\ y \\ 0 \\ t \end{bmatrix}$$

$$= \mathbf{K}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}]\begin{bmatrix} x \\ y \\ t \end{bmatrix}. \tag{2}$$

From Eq. (2), we can see that point $[x, y, z, t]^{\mathrm{T}}$ on the model plane also can be defined by 2D homogenous coordinates $[x, y, t]^{\mathrm{T}}$. Based on projective geometry theory, points on the model plane satisfying the equation $t = 0$ form the *line at infinity* of the model plane. We denote it by $\mathbf{l}_\infty$. Now we consider two specific points $\mathbf{I}(1, i, 0, 0)$, $\mathbf{J}(1, -i, 0, 0)$ on $\mathbf{l}_\infty$ (which are generally called *circular points*) [5]. Obviously both $\mathbf{I}$ and $\mathbf{J}$ satisfy the equation $\tilde{\mathbf{M}}^{\mathrm{T}}\tilde{\mathbf{M}} = 0$, so $\mathbf{I}$ and $\mathbf{J}$ are points on the absolute conic $\boldsymbol{\omega}$. If the images of $\mathbf{I}$ and $\mathbf{J}$ are denoted by $\mathbf{I}_m$ and $\mathbf{J}_m$, then $\mathbf{I}_m$ and $\mathbf{J}_m$ should lie on IAC, which yields

$$\mathbf{I}_m^{\mathrm{T}}\mathbf{K}^{-\mathrm{T}}\mathbf{K}^{-1}\mathbf{I}_m = 0,$$

$$\mathbf{J}_m^{\mathrm{T}}\mathbf{K}^{-\mathrm{T}}\mathbf{K}^{-1}\mathbf{J}_m = 0. \tag{3}$$

Since $\mathbf{I}$ and $\mathbf{J}$ are conjugated points, under the perspective transformation, $\mathbf{I}_m$ and $\mathbf{J}_m$ are also conjugated ones. Hence the two equations in Eq. (3) are actually identical. We note that any equation in Eq. (3) can actually produces two linear constraints on IAC by enforcing both its real and imaginary parts to be zero. i.e.,

$$\mathrm{Re}(\mathbf{I}_m^{\mathrm{T}}\mathbf{K}^{-\mathrm{T}}\mathbf{K}^{-1}\mathbf{I}_m) = 0,$$

$$\mathrm{Im}(\mathbf{I}_m^{\mathrm{T}}\mathbf{K}^{-\mathrm{T}}\mathbf{K}^{-1}\mathbf{I}_m) = 0.$$

Now let us look at the proposed new model pattern (Fig. 1) which can be characterized by Fig. 2, i.e., a pencil of lines passing through the circle's center $\mathbf{O}$. If we assume

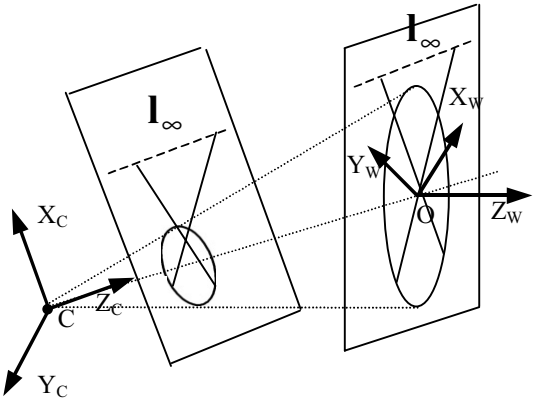

Fig. 3. From model plane to image plane.

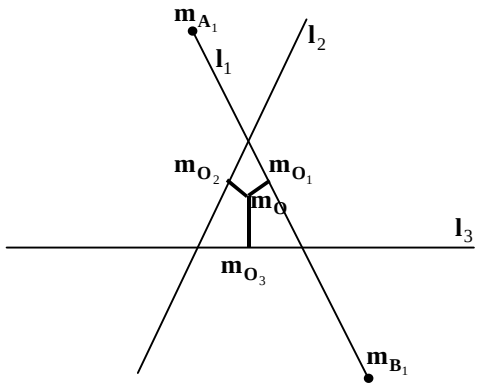

Fig. 4. $\{\mathbf{l}_i,\ i=1,2,\ldots,n\}$ do not intersect at the same point.

that the coordinates of $\mathbf{O}$ is $(O_x, O_y, 0)$ and the circle's radius is r which is unknown, then the circle equation expressed in homogenous coordinates is

$$(x - O_x t)^2 + (y - O_y t)^2 = r^2 t^2. \tag{4}$$

To compute the intersecting points of the line at infinity $\mathbf{l}_\infty$ with circle $\mathbf{O}$, we combine the equation of $\mathbf{l}_\infty(t=0)$ and (4) to give

$$x^2 + y^2 = 0.$$

The solution is $y = \pm \mathrm{i}x$. That is to say, the two intersecting points can be expressed as $(1, \pm\mathrm{i}, 0)$ in homogeneous coordinates, which are independent of O_x, O_y and r. It also means that *any circle in the model plane should intersect the line at infinity of the model plane at two circular points* [6] (see Fig. 3). Accordingly, in the image plane, the image of the line at infinity should intersect the image of the model circle at the image of the two circular points.

In most cases, the image of the model circle is an ellipse, which can be directly extracted from the image. Hence if we can get the image of the line at infinity (generally called the vanishing line), then we can obtain the images of the two circular points.

2.2. Computing the vanishing line

As sketched in Fig. 2, line $\mathbf{L}_1$ which is passing through the model circle's center $\mathbf{O}$ intersects the circle at point $\mathbf{A}_1, \mathbf{B}_1$, and it also intersects the vanishing line $\mathbf{l}_\infty$ at $\mathbf{C}_1$. Based on the theory of projective geometry, the cross-ratio of the four collinear points $\mathbf{A}_1, \mathbf{B}_1, \mathbf{O}, \mathbf{C}_1$ will be -1, i.e.,

$$(\mathbf{A}_1\mathbf{B}_1, \mathbf{O}\mathbf{C}_1) = \frac{\mathbf{A}_1\mathbf{O}}{\mathbf{B}_1\mathbf{O}} \Big/ \frac{\mathbf{A}_1\mathbf{C}_1}{\mathbf{B}_1\mathbf{C}_1} = -1.$$

Since $\mathbf{O}$ is the midpoint of $\mathbf{A}_1\mathbf{B}_1$ and $\mathbf{C}_1$ is a point at infinity, we can also say that points $\mathbf{A}_1, \mathbf{B}_1$ harmonically conjugate with respect to points $\mathbf{O}, \mathbf{C}_1$. Suppose the corresponding image points of $\mathbf{A}_1, \mathbf{B}_1, \mathbf{O}, \mathbf{C}_1$ are $\mathbf{m}_{\mathbf{A}_1}, \mathbf{m}_{\mathbf{B}_1}, \mathbf{m}_{\mathbf{O}}, \mathbf{m}_{\mathbf{C}_1}$, respectively, and since both collinearity and cross-ratio are projective invariants, we may easily conclude that $\mathbf{m}_{\mathbf{A}_1}, \mathbf{m}_{\mathbf{B}_1}, \mathbf{m}_{\mathbf{O}}, \mathbf{m}_{\mathbf{C}_1}$ are also collinear and their cross-ratio is kept as -1. Based on these two properties, we obtain the following two equations:

$$\left|\frac{\mathbf{m}_{\mathbf{A}_1}\mathbf{m}_{\mathbf{O}}}{\mathbf{m}_{\mathbf{B}_1}\mathbf{m}_{\mathbf{O}}}\right| \Big/ \left|\frac{\mathbf{m}_{\mathbf{A}_1}\mathbf{m}_{\mathbf{C}_1}}{\mathbf{m}_{\mathbf{B}_1}\mathbf{m}_{\mathbf{C}_1}}\right| = -1,$$

$$(\mathbf{m}_{\mathbf{A}_1} \times \mathbf{m}_{\mathbf{B}_1}) \cdot \mathbf{m}_{\mathbf{C}_1} = 0. \tag{5}$$

$\mathbf{m}_{\mathbf{C}_1}$ can be easily solved through the above two equations, and obviously $\mathbf{m}_{\mathbf{C}_1}$ should be on the vanishing line. If there exist other lines on the model plane similar to $\mathbf{L}_1$, more points on the vanishing line can be similarly obtained. Finally, we use all these points to compute the vanishing line by a least squares fitting.

As described above, during computing $\mathbf{m}_{\mathbf{C}_1}$, we have assumed that the point $\mathbf{m}_{\mathbf{A}_1}, \mathbf{m}_{\mathbf{B}_1}$ and $\mathbf{m}_{\mathbf{O}}$ are collinear. In practice, such collinearity usually does not hold due to noise and errors introduced in the line detecting process, and consequently extracted image lines $\{\mathbf{l}_i,\ i=1,2,\ldots,n\}$ do not in practice always intersect with each other precisely at the image of the model circle's center (see Fig. 4). In this case, we minimize the following cost function:

$$E = \sum_i d^2(\mathbf{m}_{\mathbf{O}}, \mathbf{l}_i),$$

where $d(\mathbf{m}_{\mathbf{O}}, \mathbf{l}_i)$ is the distance from point $\mathbf{m}_{\mathbf{O}}$ to line $\mathbf{l}_i$. Minimizing this cost function becomes a standard nonlinear optimization problem, and it can be solved by standard optimization algorithms such as Levenberg–Marquardt. After obtaining $\mathbf{m}_{\mathbf{O}}$, we back-project $\mathbf{m}_{\mathbf{O}}$ onto every $\mathbf{l}_i$ to compute $\mathbf{m}_{\mathbf{O}_i}$. During our using Eq. (5) to compute $\mathbf{m}_{\mathbf{C}_i}$, we replace $\mathbf{m}_{\mathbf{O}}$ by $\mathbf{m}_{\mathbf{O}_i}$ in order to ensure $\mathbf{m}_{\mathbf{A}_1}, \mathbf{m}_{\mathbf{B}_1}$ and $\mathbf{m}_{\mathbf{O}}$ are indeed collinear.

2.3. Extracting ellipse and lines

Accurately extracting the ellipse and lines from images is an essential step for our new technique. Here we use random sample consensus paradigm (RANSAC) [7] to extract lines

and it proves to be performing well. In contrast to extracting lines, extracting the ellipse is a bit more involved. In our experiments reported in this paper, we use a least-squares fitting technique based on algebraic distance to extract the ellipse after we have removed all the lines from the images. Assume there are n image points $\{\mathbf{x}_i\} = \{(x_i, y_i),\ i = 1, 2, \ldots, n\}$ on the ellipse, and the equation of the ellipse is

$$Q(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F,$$

we then minimize the cost function $F = \sum_{i=1}^{n} Q^2(x_i, y_i)$ subject to the constraint of $A^2+B^2+C^2+D^2+E^2+F^2=1$ to extract the ellipse. More detailed discussions on extracting ellipse are in Refs. [8–10].

It is worth noting that using Euclidean distance as the minimizing criterion will generally outperform using algebraic distance in feature extraction. However, since our model plane is a white sheet containing a black circle and several black lines, its projected image is quite ideal for applying traditional feature extracting methods. We emphasize that we detect the ellipse only after all the lines have been extracted and all the points of the lines have been removed. Thus in this case, extracting ellipse by minimizing the algebraic distance proves to be accurate enough. However, we do believe using more involved ellipse detecting techniques will help improve the accuracy of final calibration results.

2.4. Determining the intrinsic parameters

As discussed in Sections 2.1 and 2.2, once the image of the model circle and the vanishing line are obtained, the images of the two circular points $\mathbf{I}_m, \mathbf{J}_m$ can be easily derived which are the intersecting points between the ellipse and the vanishing line.

Let $\mathbf{I}_m = [I_{m1}, I_{m2}, I_{m3}]^{\mathrm{T}}$, and

$$\mathbf{C} = \mathbf{K}^{-\mathrm{T}}\mathbf{K}^{-1} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{12} & C_{22} & C_{23} \\ C_{13} & C_{23} & C_{33} \end{bmatrix}.$$

From Eq. (3), we have

$$[I_{m1}, I_{m2}, I_{m3}]\mathbf{C}[I_{m1}, I_{m2}, I_{m3}]^{\mathrm{T}} = 0. \tag{6}$$

Because matrix $\mathbf{C}$ is symmetric, we define a 6×1 vector

$$\mathbf{c} = [C_{11}, C_{12}, C_{22}, C_{13}, C_{23}, C_{33}]^{\mathrm{T}}$$

and rewrite Eq. (6) as

$$\mathbf{Ac} = 0 \tag{7}$$

with $\mathbf{A} = [I_{m1}I_{m1},\ I_{m1}I_{m2} + I_{m2}I_{m1},\ I_{m2}I_{m2},\ I_{m3}I_{m1} + I_{m1}I_{m3},\ I_{m3}I_{m2} + I_{m2}I_{m3},\ I_{m3}I_{m3}]$.

Since $\mathbf{A}$ is a complex vector, $\mathbf{c}$ is a real vector, Eq. (7) is equivalent to the following two homogeneous equations:

$$\begin{bmatrix} \mathrm{Re}(\mathbf{A}) \\ \mathrm{Im}(\mathbf{A}) \end{bmatrix} \mathbf{c} = 0. \tag{8}$$

If n images of the model plane are taken, by stacking n such equations, we have

$$\mathbf{Vc} = 0, \tag{9}$$

where $\mathbf{V}$ is a $2n \times 6$ matrix. If $n \geqslant 3$ and $rank(\mathbf{V}) \geqslant 5$, $\mathbf{c}$ can be determined uniquely up to a scale factor in the least-squares sense (if the skew factor s can be assumed to be zero in advance, then two images are sufficient for the calibration). The result can also be interpreted this way as shown in [11]. In general, 5 points are required to fit the image of the absolute conic, since each image can only provide two such points, the minimum number of images required is: $\frac{5}{2} + 1 = 3$. The solution to Eq. (9) is well known as the eigenvector of $\mathbf{V}^{\mathrm{T}}\mathbf{V}$ associated with the smallest eigenvalue [12].

Once vector $\mathbf{c}$ is obtained, $\mathbf{K}^{-1}$ can be computed using Cholesky factorization [12], and $\mathbf{K}$ can then be obtained by inverting $\mathbf{K}^{-1}$. This $\mathbf{K}$ is equal to the matrix of the real camera intrinsic parameters up to a scale. We can obtain the actual intrinsic parameters matrix by normalizing $\mathbf{K}$ such that $k_{33} = 1$.

2.5. Recovering partial extrinsic parameters

As indicated in Section 2.1, any point $(x, y, 0)$ in the model plane satisfying $y = \pm \mathrm{i}x$ (x, y are complex number) should correspond to $\mathbf{I}_m$ (or $\mathbf{J}_m$). Relating this to Eq. (2) we have

$$\lambda_1 \mathbf{I}_m = \mathbf{K}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}] \begin{bmatrix} x \\ \pm \mathrm{i}x \\ 0 \end{bmatrix} = x(\mathbf{Kr}_1 \pm \mathbf{Kr}_2 \cdot \mathrm{i}), \tag{10}$$

where λ_1 is a constant, let $\lambda_1/x = a + b\mathrm{i}$, we then have

$$\mathbf{r}_1 = \lambda_2 \mathbf{K}^{-1}(a\,\mathrm{Re}(\mathbf{I}_m) + b\,\mathrm{Im}(\mathbf{I}_m)),$$

$$\mathbf{r}_2 = \lambda_3 \mathbf{K}^{-1}(b\,\mathrm{Re}(\mathbf{I}_m) + a\,\mathrm{Im}(\mathbf{I}_m)) \quad (\lambda_2, \lambda_3 \text{ are constants}),$$

where a, b are unknown constants, so $\mathbf{r}_1$, $\mathbf{r}_2$ can not be uniquely obtained. It is worth noting that $\mathbf{r}_1^{\mathrm{T}}\mathbf{r}_2 = 0$ is always true regardless of the takings of a and b, hence it cannot add any new more constraints to a and b. In the contrast, $\mathbf{r}_3$ can be derived from $\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$,

$$\mathbf{r}_3 = \lambda_4[\mathbf{K}^{-1}\,\mathrm{Re}(\mathbf{I}_m)] \times [\mathbf{K}^{-1}\,\mathrm{Im}(\mathbf{I}_m)],$$

where λ_4 can be obtained by $\|r_3\| = 1$.

We can also recover $\mathbf{t}$ up to an unknown scale factor λ_t by $\mathbf{m_O}$ which is the image point of the model circle's center

$$\mathbf{t} = \lambda_t \mathbf{K}^{-1}\mathbf{m_O}.$$

Now we can see that the extrinsic parameters $\mathbf{r}_1, \mathbf{r}_2$ are lost when using our model plane. It is mainly due to the central-symmetry of the model plane. As described in Fig. 3, we assume $\mathbf{C}$ is the camera's optical center, and the origin of the camera coordinate system $X_C Y_C Z_C$ is located at $\mathbf{C}$ also. From Section 2.1, we know that the coordinates of the circular points on the model plane are independent of the

model circle's center **O**, thus without loss of generality, let axis Z_C pass through the circle's center **O** and let **O** be the origin of the world coordinate system $X_W Y_W Z_W$, and axis Z_W be perpendicular to the model plane. Under such a setup, vectors $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ of the rotation matrix **R** correspond to the unit vector of X_W-, Y_W-, Z_W-axis in the camera coordinate system, and the image of the model plane is only affected by Z_W-axis, while independent of X_W-, Y_W-axis because of the central-symmetry of the model plane. As a result, once the camera's intrinsic parameters and the image of the model plane are obtained, we can only compute $\mathbf{r}_3$ from them. Fortunately, in most cases, the camera's extrinsic parameters are not required by DVS, so the loss of $\mathbf{r}_1, \mathbf{r}_2$ is not a major deficiency for our new technique.

2.6. Degenerate configurations

A complete analysis of all possible degenerate configurations in self-calibration has been carried out by Sturm [13] and Ma [14]. In our method, degenerate configurations will only occur when the projected images of circular points in different images are identical. From Eq. (10) we can see that $\mathbf{I}_m$ depends only on rotation matrix **R** and is independent of translation vector **t**. Hence in order to avoid degenerate configurations, **R** should be altered at each image taking. In other words, the relative orientation between the camera and the model plane must be different for different image taking.

Another degenerate configuration occurs when the image plane is parallel to the model plane. In this case, any point at infinity of the model plane $(x, y, 0, 0)$ will correspond to

$$\mathbf{K}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}] \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \mathbf{K} \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha x + \gamma y \\ \beta y \\ 0 \end{bmatrix},$$

where $[\alpha x + \gamma y, \beta y, 0]^{\mathrm{T}}$ is a point at infinity of the image plane. It means that the vanishing line coincides with the line at infinity of the image plane. In this case, the method of computing the vanishing line in Section 2.1 becomes invalid. Such a degenerate configuration can be easily detected and avoided as follows. As shown in Fig. 5, $\mathbf{m_O}$ is the image of the model circle's center, its corresponding polar line associated with the ellipse (the projection of the model circle) is the vanishing line $\mathbf{l}_\infty$. **E** is the center of the ellipse whose corresponding polar line is the line at infinity of the image plane (denoted as $\mathbf{L}_\infty$). According to the principle of polarity transformation in projective geometry, a polar line is uniquely determined by its corresponding polar associated with any proper quadric conic and vice versa. Hence, $\mathbf{l}_\infty$ will coincide with $\mathbf{L}_\infty$ as well as $\mathbf{m_O}$ coincides with **E** while the image plane is parallel to the model plane. Based on this property, by checking the coincidence of **E** and $\mathbf{m_O}$, the degenerate configuration can be detected and avoided easily.

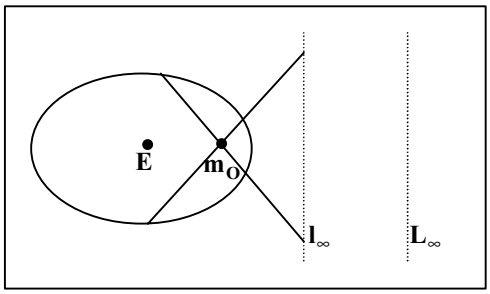

Fig. 5. Line at infinity and vanishing line in image plane.

2.7. Algorithm outline

The proposed calibration algorithm is outlined as following.

1. Print a circle and a pencil of lines passing through the circle's center on a white sheet, and attach the sheet to a planar surface;
2. Take 3 or more images of the sheet at different orientations.
3. For each image:
   - 3.1 Extract the ellipse and the pencil of lines, then check whether the configuration is a degenerate one as shown in Section 2.6.
   - 3.2 Compute the vanishing line.
   - 3.3 Compute the image points of the two circular points by intersecting the vanishing line with the ellipse (Fig. 5).
4. Determine vector **c**, then compute matrix **K** as shown in Section 2.4.

3. Experiments

In this section, a number of experiments are reported. Both synthetic and real images are used to evaluate the accuracy and robustness of the proposed technique.

3.1. Simulation results

In all the subsequent computer simulations, the camera's setup is: $f_u = 1200$, $f_v = 1000$, $s = 0.2$, $u_0 = v_0 = 0$. The image resolution is 1000×1000. The orientation of the model plane is characterized by a rotating axis r and a rotating angle θ about this axis (unit: degree). The position of the model plane is denoted by a 3D vector t (unit: centimeter). The model plane includes a circle with radius 50 (unit: centimeter) and 10 lines passing through the circle's center. The including angle between any two adjacent model lines is equal to 36°.

Table 1
Calibrating results at different noise level

Noise level	f_u	f_v	s	u_0	v_0
0.4	1201.320	1001.351	0.201	−0.544	0.000
0.8	1201.609	998.817	0.263	−1.991	0.448
1.2	1198.048	1003.224	0.429	−2.396	0.852
1.6	1203.699	1006.150	0.605	−3.245	2.087
2.0	1207.070	985.183	0.630	−6.949	2.664
2.4	1182.346	1017.202	0.714	−7.567	4.369
2.8	1219.099	1018.161	0.838	−8.536	4.433
3.2	1221.168	1024.087	0.947	−15.345	9.673

3.1.1. Noise influence

In this experiment, three images of the model plane are taken by the camera at the following three different orientations: $\mathbf{r}_1 = [170, 50, 10]^T, \theta = 15°, \mathbf{t}_1 = [20, 20, 0]^T; \mathbf{r}_2 = [40, 50, 160]^T, \theta = 15°, \mathbf{t}_2 = [30, 20, 5]^T; \mathbf{r}_3 = [20, 70, 70]^T, \theta = 15°, \mathbf{t}_3 = [10, 30, 30]^T$. A Gaussian noise with 0 mean and σ standard deviation (noise level) is added to each projected image point. The noise level varies from 0.4 to 3.2 pixels. The results are shown in Table 1. All the results in Table 1 are the average value of 100 independent trials.

In order to investigate the accuracy of the proposed technique, the standard deviations of the five intrinsic parameters at each different noise level are computed and shown in Fig. 6. From this figure, we can see that though the noise level increases to 6.0 pixels, the standard deviations of the intrinsic parameters are still low. This indicates that the proposed technique is accurate enough even with the presence of a high degree of noise.

3.1.2. Influence of the number of used images

This experiment investigates the performance regarding to the number of the images used for calibration. The number of the images used for calibration varies from 3 to 18. For the first 3 images, the camera's intrinsic parameters and the extrinsic parameters are the same as in the experiments in Section 3.1.1. For other images, rotation axis, rotation angle and position of the model plane for each image are altered and chosen at random, while the camera's intrinsic parameters keep unchanged. For each number, 100 independent trials are done. A Gaussian noise with 1.0 standard deviation is added in each trial. The results shown in Fig. 7 are average relative errors. From this figure, we can see that in general using more images will help increase the accuracy of calibration results. However, once image number is more than 6, such an improvement becomes insignificant.

3.1.3. Sensitivity to the deformation of the model circle

We also investigate the performance of the proposed technique with respect to the deformation of the model circle. For convenience, we only model a deformed circle as an ellipse. We use the eccentricity e of an ellipse to describe the ellipse's deviation from a circle (e is defined by $e = \sqrt{1 - (b/a)^2}$ where a, b are the long axe and short axe of an ellipse respectively. When $e = 0$, the ellipse reduces to a circle and there is no deformation). We increase e from 0 to 0.3. When e is equal to 0.3, the corresponding model circle is greatly deviated from a true one, and such a case is nearly impossible to happen in practice since the model pattern is usually designed with high precision. For each e, 100 independent trials are done. A Gaussian noise with 1.0 standard deviation is added in each trial, and the average relative errors are computed and shown in Fig. 8. From this figure, we can see that the performance of the proposed technique is

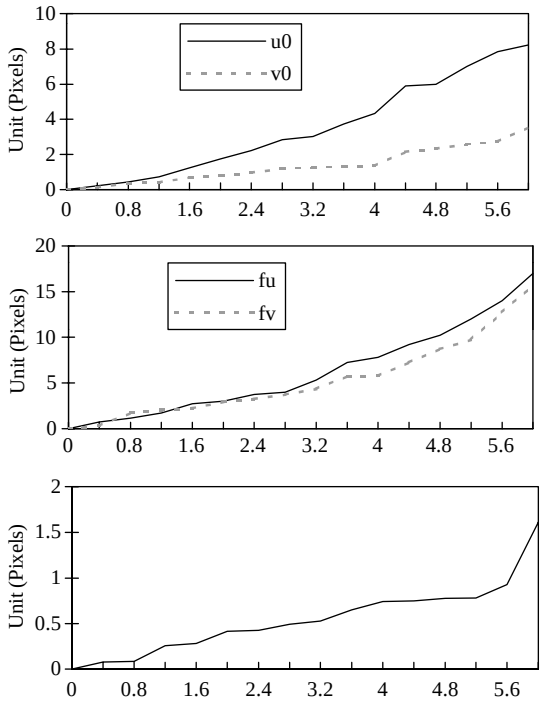

Fig. 6. Calibrating results (u_0, v_0, f_u, f_v, s) at different noise levels.

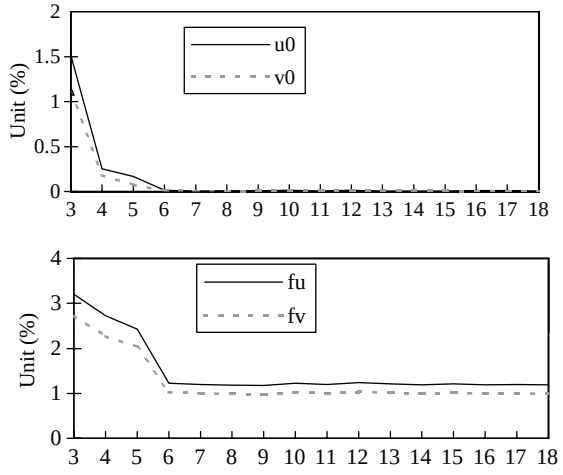

Fig. 7. Calibrating results (u_0, v_0, f_u, f_v) at different image numbers.

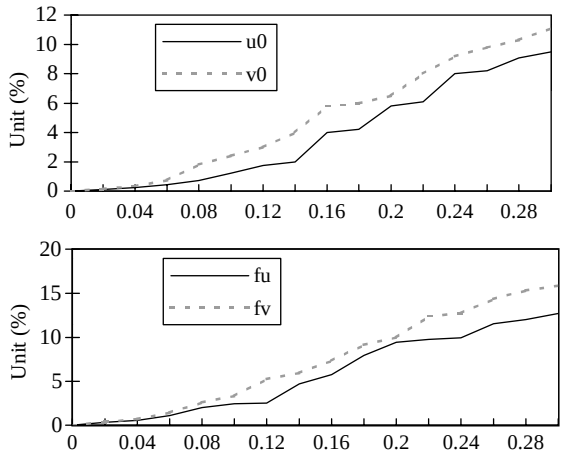

Fig. 8. Calibrating results (u_0, v_0, f_u, f_v) at different eccentricity of circle.

still satisfactory when the deformation of the model circle is not too serious.

3.1.4. A comparison with Zhang's technique

In order to make the comparison meaningful, all the parameters, e.g. the camera's intrinsic and extrinsic parameters, noise level, image number are set identical during the calibrating processes of applying the two techniques. In addition, the corner points used by Zhang's technique are directly selected from the model plane employed in our technique (totally 13 corners). Note that in both Zhang's and our techniques, we do not apply any nonlinear optimization methods to refine the results as mentioned in Ref. [3]. We show the relative errors of u_0 and f_u in Fig. 9. From this figure, we clearly see that the two techniques are comparable when noise level is low. However, once the noise level

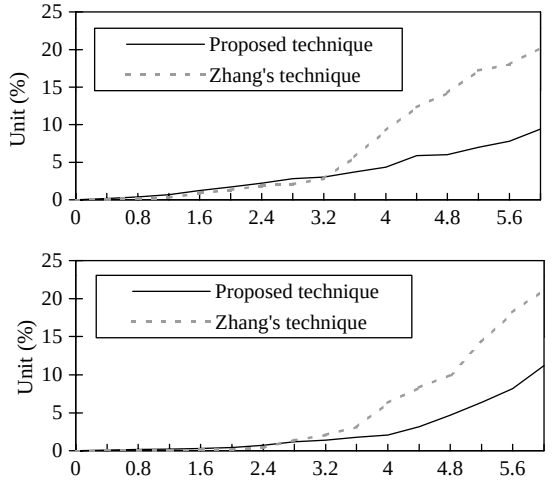

Fig. 9. Calibrating results (relative error of u_0 and f_u, respectively) compared to Zhang's technique.

Table 2
Calibrating results from the 8 images (4 of them are shown in Fig. 10)

f_u	f_v	s	u_0	v_0
1396.220	1400.731	−0.827	568.401	453.079

increases to as high as 3.0 pixels, the relative errors of Zhang's technique are distinctly higher than those of our technique.

3.2. Experiments with real images

The images were taken by a KODAK-DC120 digital camera. The image resolution is 1280×960. We printed a circle with 6 lines passing through the circle's center on a white paper with a laser printer, and attached the paper to the wall. 8 images of the model plane were taken at different orientations (4 of them are shown in Fig. 10). Then we applied the proposed technique to these 8 images. The calibration results are shown in Table 2.

In order to evaluate the calibration results in Table 2, a well structured calibration object was reconstructed. Fig. 11 shows two images of the calibration object taken by the previous calibrated camera. We manually picked 9 corresponding points from each of the two visible sides (marked by cross). Applying the structure-from-motion algorithm (SFM) as described in Ref. [15], we reconstructed the two visible sides. Two views of the reconstructed object from different viewpoints are shown in Fig. 12. We can clearly see that the reconstructed points on the same side of the calibration object are indeed coplanar. In addition, the computed including angle between the two reconstructed sides is 91.02°, which accords well with the ground truth 90.0°.

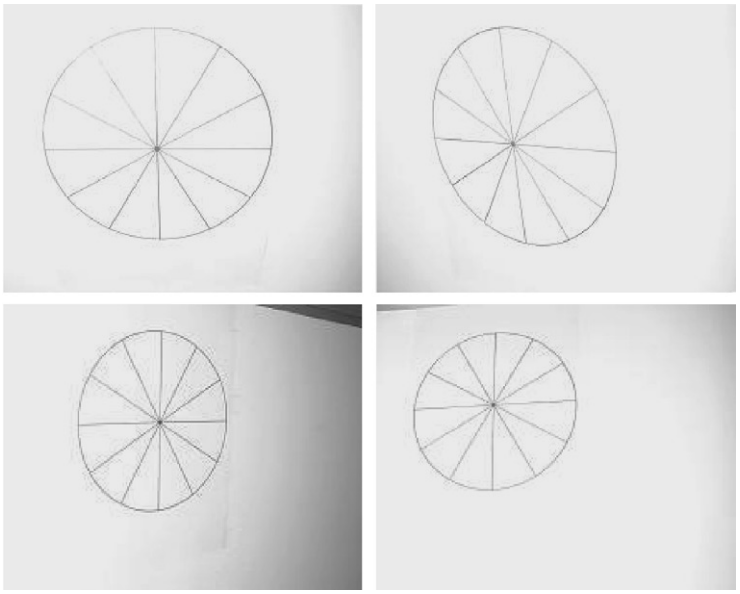

Fig. 10. Four real images taken by a digital camera.

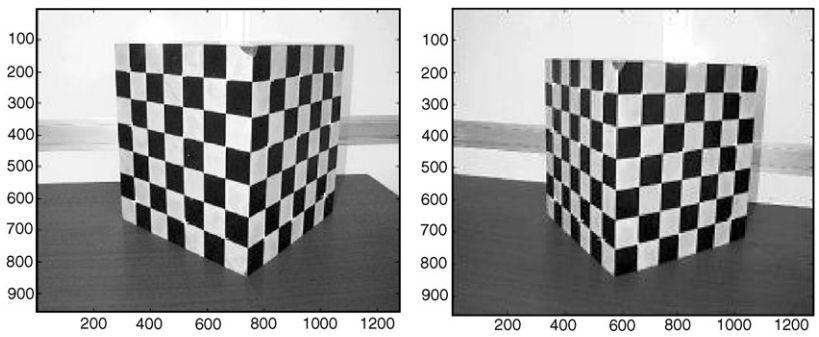

Fig. 11. Two images of a calibration object taken by a digital camera.

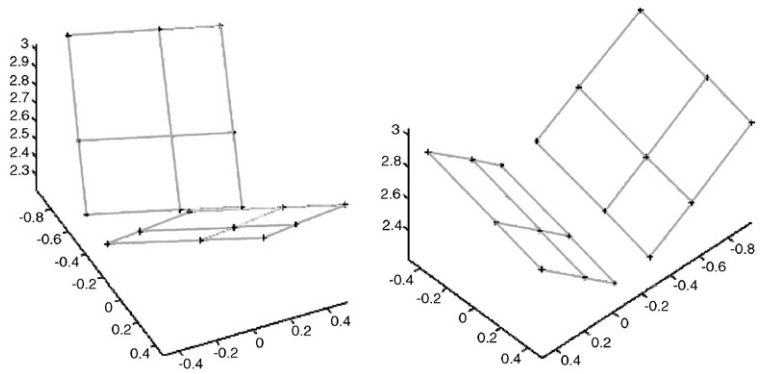

Fig. 12. Two views of the reconstructed calibration object.

Fig. 13. Two images of a box and a puppet taken by a digital camera.

Fig. 14. Two views of the reconstructed box and puppet.

This proves indirectly that the calibrating results in Table 2 are accurate.

Another experiment on real images is also carried. We used the same digital camera to take two images of a box and a puppet as shown in Fig. 13. Then we reconstructed the box and the puppet using the calibration results in Table 2. Fig. 14 shows two views of the reconstructed results under two different orientations. The reconstructed results seem realistic.

4. Conclusion

Inspired by Zhang's technique, we propose a new flexible calibrating technique based on the concept of circular points. Our new technique is based on a novel model plane consisting of a circle and a pencil of lines passing through the circle center. The main points of our technique are: (1) There is no need to establish point correspondences between the model points and resulting image ones, hence the calibration process becomes extremely simple; (2) there is no need to measure the circle center's coordinates and radius, hence it is more convenient for the design and making of a model plane. The proposed technique is particularly useful for those people who are not familiar with computer vision. Experiments with simulated data as well as with real images show that our new technique is robust and accurate.

5. Summary

In this paper, we propose a novel camera calibration technique. This technique is inspired by Dr. Zhengyou Zhang's work on flexible calibration technique. This technique requires a newly designed 2D calibration pattern which includes a circle and a pencil of lines passing through the circle's center. The camera only needs to observe the planar calibration pattern at more than two different unknown orientations, then all the five intrinsic parameters are determined linearly. The main advantage of the proposed technique is that it needs to know neither any metric measurement on the model plane, nor any corresponding information between the model plane and images. The only burden of the technique is involving some ellipse and line extraction from the images. After this feature extraction, the whole calibration process becomes a straightforward computation. Due to the simplicity of this technique, it is particularly useful for those people who are not familiar with computer vision. Experiments with simulated data as well as with real images show that our new technique is robust and accurate.

References

[1] D.C. Brown, Close-range camera calibration, Photogrammetric Eng. 37 (8) (1971) 855–866.
[2] W. Faig, Calibration of close-range photogrammetry system: mathematical formulation, Photogrammetric Eng. Remote Sensing 41 (12) (1975) 1479–1486.
[3] Z. Zhang, Flexible camera calibration by viewing a plane from unknown orientations, Proceedings of the Fifth International Conference on Computer Vision, 1999, pp. 666–673.
[4] S.D. Ma, Z.Y. Zhang, Computer Vision—Computational Theory and Algorithm, Science Press, Beijing, 1998 (in Chinese).
[5] J.G. Semple, G.T. Kneebone, Algebraic Projective Geometry, Clarendon Press, Oxford, 1952 (reprinted 1979).
[6] D. Liebowitz A. Zisserman, Metric rectification for perspective images of planes. Proceedings of the Computer Vision and Pattern Recognition, 1998, pp. 482–488.
[7] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with application to image analysis

and automated cartography, Commun. Assoc. Comput. Mach. 24 (1981) 381–395.
[8] F.L. Bookstein, Fitting conic sections to scattered data, Comput. Graphics Image Process. 9 (1979) 56–71.
[9] Z. Zhang, Parameter estimation techniques: a tutorial with application to conic fitting, Technical Report, INRIA, October, 1995.
[10] K. Kanatani, Geometric Computation for Machine Vision, Oxford Science Publications, Oxford, 1993.
[11] O. Faugeras, L. Quan, P. Sturm, Self-calibration of a 1D projective camera and its application to the self-calibration of a 2-D projective camera, Proceedings of the Fifth European Conference on Computer Vision, 1998, pp. 36–54.
[12] G. Golub, C. Van Loan, Matrix Computations, 3rd Edition, The John Hopkins University Press, Baltimore, MD, 1996.
[13] P. Sturm, Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction, Proceedings of the Conference on Computer Vision and Pattern Recognition, 1997, pp. 1100–1105.
[14] Y. Ma, Stefano Soatto, Jana Kosecka, Shankar Sastry, Euclidean reconstruction and reprojection up to subgroups, Proceedings of the International Conference of Computer Vision, 1999, pp. 773–780.
[15] Z. Zhang, Motion and structure from two perspective views: from essential parameters to Euclidean motion via fundamental matrix, J. Opt. Soc. Am. A 14 (11) (1997) 2938–2950.

About the Author—XIAOQIAO MENG is currently a Ph.D. student in the University of California, Los Angeles (UCLA). He received his B.S. degree in Automatic Control from the University of Science and Technology of China in 1998, and his M.S. degree in Pattern Recognition and Intelligent Control from the Institute of Automation, Chinese Academy of Sciences, P. R. China. His interests include pattern recognition, wireless networks and communications.

About the Author—ZHANYI HU is a professor in the Institute of Automation, Chinese Academy of Sciences, P. R. China. He holds a B.S. degree in Automatic Control from the North China University of Technology in 1985, P. R. China, and a Ph.D. degree (Docteur d'Etat) in Computer Science from the University of Liege, Belgium, in 1993. His research interests include Hough transform, image processing, vision guided robot navigation and etc.