[3] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Advances in Neural Information Processing Systems 14*. Cambridge, MA: MIT Press, 2001, pp. 585–591.
[4] F. R. K. Chung, *Spectral Graph Theory,* 1997, AMS, Regional Conf. Ser. Math.
[5] J. Duchene and S. Leclercq, "An optimal transformation for discriminant and principal component analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 6, pp. 978–983, Jun. 1988.
[6] R. O. Duda, P. E. Hart, and D. G. Stork*, Pattern Classification*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2000.
[7] G. H. Golub and C. F. Van Loan*, Matrix computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.
[8] R. Gross, J. Shi, and J. Cohn, "Where to go with face recognition," in *Third Workshop on Empirical Evaluation Methods in Computer Vision*, Kauai, HI, Dec. 2001.
[9] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.
[10] X. He and P. Niyogi, "Locality preserving projections," in *Advances in Neural Information Processing Systems 16.*. Cambridge, MA: MIT Press, 2003.
[11] Q. Liu, R. Huang, H. Lu, and S. Ma, "Face recognition using Kernel based fisher discriminant analysis," presented at the 5th Int. Conf. Automatic Face and Gesture Recognition, Washington, DC, May 2002.
[12] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 228–233, Feb. 2001.
[13] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 696–710, Jul. 1997.
[14] H. Murase and S. K. Nayar, "Visual learning and recognition of 3-d objects from appearance," *Int. J. Comput. Vis.*, vol. 14, 1995.
[15] P. J. Phillips, "Support vector machines applied to face recognition," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 1998, pp. 803–809.
[16] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
[17] T. Shakunaga and K. Shigenari, "Decomposed eigenface for face recognition under various lighting conditions," presented at the IEEE Conf. Computer Vision and Pattern Recognition, Dec. 2001.
[18] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illuminlation, and expression database," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1615–1618, Dec. 2003.
[19] J. Tenenbaum, V. d. Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
[20] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
[21] M. Turk and A. P. Pentland, "Face recognition using eigenfaces," presented at the IEEE Conf. Computer Vision and Pattern Recognition, Maui, HI, 1991.

Camera Calibration Using Symmetric Objects

Xiaochun Cao and Hassan Foroosh*, Senior Member, IEEE*

***Abstract*—This paper proposes a novel method for camera calibration using images of a mirror symmetric object. Assuming unit aspect ratio and zero skew, we show that interimage homographies can be expressed as a function of only the principal point. By minimizing symmetric transfer errors, we thus obtain an accurate solution for the camera parameters. We also extend our approach to a calibration technique using images of a 1-D object with a fixed pivoting point. Unlike existing methods that rely on orthogonality or pole-polar relationship, our approach utilizes new interimage constraints and does not require knowledge of the 3-D coordinates of feature points. To demonstrate the effectiveness of the approach, we present results for both synthetic and real images.**

***Index Terms*—Camera calibration, homography, stereo reconstruction.**

I. Introduction

Traditional camera calibration methods require a calibration object with a fixed 3-D geometry [1]. Recently, more flexible plane-based calibration methods [2]–[4] are proposed, which use a planar point pattern shown at a few different orientations. Some recent methods in this category use also nonplanar shapes [5]–[7]. Zhang [8] has also presented a method for calibration using 1-D objects. Another branch of study, referred to as self-calibration [9], does not use a calibration object. Their aim is to provide further flexibility by not requiring a prior knowledge of the 3-D to 2-D correspondences. Various methods [10]–[13] have been proposed in this category that rely on scene or motion constraints, most of which require good initialization, and multiple views.

The approach described herein needs a calibration object, but it does not require the knowledge of 3-D coordinates of the object. Different from existing methods, our approach extends the current state-of-the-art calibration methods to situations where only one vanishing point and no vanishing line are known per view. The proposed technique requires the camera to observe a symmetric object only at a few (at least two) different orientations. We will show that such configuration provides sufficient information to solve the problem using only interimage homographies. Given a configuration described shortly, our method can also be applied to 1-D objects [14], which is useful in calibration of a large number of cameras [8].

II. Preliminaries

A homogeneous 3-D point $\tilde{\mathbf{M}} \sim [\,X \quad Y \quad Z \quad 1\,]^T$ and its corresponding homogeneous image projection $\tilde{\mathbf{m}}$ in a pinhole camera are related via a 3×4 projection matrix $\mathbf{P}$. More comprehensive imaging models including radial distortion [15] are out of the scope of this paper. When the world points are coplanar (e.g., without loss of generality in the plane $Z = 0$), we get

$$\tilde{\mathbf{m}} \sim \underbrace{\mathbf{K}[\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3 \quad \mathbf{t}\,]}_{\mathbf{P}}\tilde{\mathbf{M}} = \underbrace{\begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}[\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}\,]}_{\mathbf{H}_w}\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{1}$$

Manuscript received October 7, 2005; revised May 28, 2006. This work was supported by the National Science Foundation (NSF) under Grant IIS-0644280. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Luca Lucchese.

The authors are with the Computational Imaging Lab, School of EECS, University of Central Florida, Orlando, FL 32816-2362 USA (e-mail: xccao@cs.ucf.edu; foroosh@cs.ucf.edu).

Color versions of Figs. 2–7 are available online at http://ieeexplore.ieee.org.
Digital Object Identifier 10.1109/TIP.2006.881940

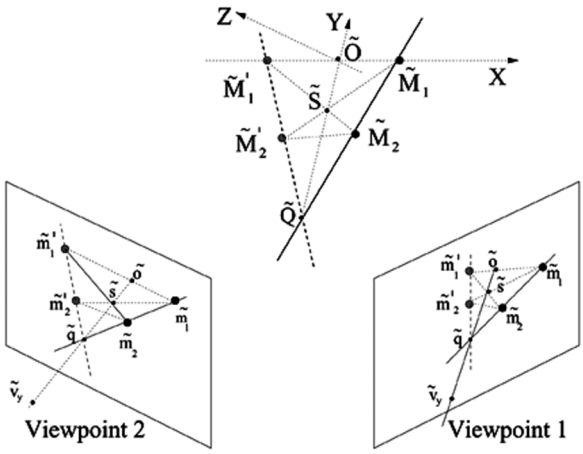


Fig. 1. Points $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$ are two arbitrary feature points on the calibration object, and $\tilde{\mathbf{M}}'_1$ and $\tilde{\mathbf{M}}'_2$ are their symmetric counterparts. Alternatively, $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$ are two feature points on a line through the point $\tilde{\mathbf{Q}}$, and $\tilde{\mathbf{M}}'_1$ and $\tilde{\mathbf{M}}'_2$ are their position after pivoting the line around the point $\tilde{\mathbf{Q}}$. In order to simplify the illustration, we choose the plane of symmetry as $X = 0$ and world coordinate frame as follows: X axis along the line $\tilde{\mathbf{M}}'_1\tilde{\mathbf{M}}_1$, Y axis along the line $\tilde{\mathbf{S}}\tilde{\mathbf{O}}$, where $\tilde{\mathbf{S}}$ is the intersection of $\tilde{\mathbf{\Pi}}$ and the line $\tilde{\mathbf{M}}'_2\tilde{\mathbf{M}}_1$, and the Z axis given by the right-hand rule.

where $\sim$ indicates equality up to a nonzero scale factor, $\mathbf{r}_i$ are the columns of the rotation matrix $\mathbf{R}$, $\mathbf{t}$ is the translation vector, $\mathbf{K}$ is the camera intrinsic matrix including the focal length f and the principal point $\tilde{\mathbf{c}} \sim [\mathbf{c}^T, 1]^T = [\,u_0 \quad v_0 \quad 1\,]^T$ (assuming a unit aspect ratio and zero skew), and $\mathbf{H}_w$ is the homography mapping the world plane $Z = 0$ to the image plane.

We assume that our cameras are observing an object that has a plane of symmetry $\tilde{\mathbf{\Pi}}$, so that for each object point $\tilde{\mathbf{M}}$, there exists a point $\tilde{\mathbf{M}}'$ on the object that is symmetrically situated with respect to $\tilde{\mathbf{\Pi}}$. Assume that two such pairs, $\tilde{\mathbf{M}}_1\tilde{\mathbf{M}}'_1$ and $\tilde{\mathbf{M}}_2\tilde{\mathbf{M}}'_2$, are viewed by a camera (see Fig. 1). Clearly, the two parallel and coplanar lines $\tilde{\mathbf{M}}'_1\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}'_2\tilde{\mathbf{M}}_2$ are perpendicular to $\tilde{\mathbf{\Pi}}$. The configuration of four such points $\tilde{\mathbf{M}}'_1\tilde{\mathbf{M}}'_2\tilde{\mathbf{M}}_2\tilde{\mathbf{M}}_1$ on a symmetric object as described above yields in general an isosceles trapezoid [16].

Note that this configuration of coplanar points may also be viewed as two separate configurations of colinear points in the 3-D space. In other words, if we take the lines $\tilde{\mathbf{M}}_1\tilde{\mathbf{M}}_2$ and $\tilde{\mathbf{M}}'_1\tilde{\mathbf{M}}'_2$ as two different positions of a 1-D object pivoting around the point $\tilde{\mathbf{Q}}$, then we can also apply the technique proposed herein to the images of this 1-D object. The only difference then is that the set of four coplanar points $\tilde{\mathbf{M}}_1$, $\tilde{\mathbf{M}}_2$, $\tilde{\mathbf{M}}'_1$, and $\tilde{\mathbf{M}}'_2$, are captured in two separate images. Therefore, the number of images required in the 1-D case is twice as many as the number of images required in the 2-D case, i.e., a minimum of four: one pair from the same viewpoint but with the 1-D object at two distinct positions, and another pair from a different viewpoint but the same two positions of the object that were used in the first pair.

III. Our Method

Our approach is based on expressing the interimage homography as a function of only the principal point. In the coordinate frame described above, $\mathbf{P}$ in (1) reduces to $\mathbf{H}_w$, which has the form [17]

$$\mathbf{H}_w = [\,r_{31}\tilde{\mathbf{v}}_x \quad r_{32}\tilde{\mathbf{v}}_y \quad t_z\tilde{\mathbf{o}}\,] \tag{2}$$

where $\tilde{\mathbf{v}}_x \sim [\,v_{xx} \quad v_{xy} \quad 1\,]^T$ and $\tilde{\mathbf{v}}_y \sim [\,v_{yx} \quad v_{yy} \quad 1\,]^T$ are the homogeneous x and y vanishing points in the image coordinate system, and $\tilde{\mathbf{o}} \sim [\,o_x \quad o_y \quad 1\,]^T$ is the homogeneous image point corresponding to the projection of the world origin $\tilde{\mathbf{O}}$. We can now verify that [5], [16]

$$\tilde{\mathbf{v}}_x \sim \left(\tilde{\mathbf{m}}_1 \times \tilde{\mathbf{m}}'_1\right) \times \left(\tilde{\mathbf{m}}_2 \times \tilde{\mathbf{m}}'_2\right) \tag{3}$$

$$\tilde{\mathbf{v}}_y^T(\tilde{\mathbf{s}} \times \tilde{\mathbf{q}}) = 0 \tag{4}$$

$$\tilde{\mathbf{o}} \sim \left(\tilde{\mathbf{m}}_1 \times \tilde{\mathbf{m}}'_1\right) \times (\tilde{\mathbf{s}} \times \tilde{\mathbf{q}}). \tag{5}$$

In our case, we only have the vanishing point $\tilde{\mathbf{v}}_x$. From the orthogonality constraint $\tilde{\mathbf{v}}_y^T\boldsymbol{\omega}\tilde{\mathbf{v}}_x = 0$ with respect to the image of the absolute conic $\boldsymbol{\omega}$, and (4), it immediately follows that

$$\tilde{\mathbf{v}}_y \sim [\tilde{\mathbf{l}}]_\times\boldsymbol{\omega}\tilde{\mathbf{v}}_x, \text{ where } \boldsymbol{\omega} \sim \begin{bmatrix} \mathbf{I} & -\mathbf{c} \\ -\mathbf{c}^T & f^2 + \|\mathbf{c}\|^2 \end{bmatrix} \tag{6}$$

where $\tilde{\mathbf{l}} \sim \tilde{\mathbf{s}} \times \tilde{\mathbf{q}}$, and $[\cdot]_\times$ is the usual notation for the skew symmetric matrix characterizing the cross product.

As shown in Fig. 1, the four points $\tilde{\mathbf{v}}_y$, $\tilde{\mathbf{q}}$, $\tilde{\mathbf{s}}$, and $\tilde{\mathbf{o}}$, are collinear, and, hence, their cross ratio is preserved under perspective projection. Therefore, we have the following equality between two given images:

$$\frac{|\tilde{\mathbf{v}}_y\tilde{\mathbf{q}}||\tilde{\mathbf{s}}\tilde{\mathbf{o}}|}{|\tilde{\mathbf{v}}_y\tilde{\mathbf{s}}||\tilde{\mathbf{q}}\tilde{\mathbf{o}}|} = \frac{|\tilde{\mathbf{v}}'_y\tilde{\mathbf{q}}'||\tilde{\mathbf{s}}'\tilde{\mathbf{o}}'|}{|\tilde{\mathbf{v}}'_y\tilde{\mathbf{s}}'||\tilde{\mathbf{q}}'\tilde{\mathbf{o}}'|} \tag{7}$$

where $|\cdot|$ denotes the distance between two points, and the primes indicate corresponding points in the second image. Equation (7) provides two solutions for f^2 in terms of $\mathbf{c}$, of which only one is correct. The ambiguity is resolved by minimizing the symmetric transfer errors as discussed later.

The camera external parameters can also be expressed in terms of $\mathbf{c}$ as follows. First, due to the fact that $\tilde{\mathbf{v}}_x \sim \mathbf{K}[\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3 \quad \mathbf{t}\,][\,1 \quad 0 \quad 0 \quad 0\,]^T = \mathbf{K}\mathbf{r}_1$ (similarly $\tilde{\mathbf{v}}_y \sim \mathbf{K}\mathbf{r}_2$), $\mathbf{r}_1$ and $\mathbf{r}_2$ in (1) can be computed as

$$\mathbf{r}_1 = \pm\frac{\mathbf{K}^{-1}\tilde{\mathbf{v}}_x}{\|\mathbf{K}^{-1}\tilde{\mathbf{v}}_x\|}, \text{ and } \mathbf{r}_2 = \pm\frac{\mathbf{K}^{-1}\tilde{\mathbf{v}}_y}{\|\mathbf{K}^{-1}\tilde{\mathbf{v}}_y\|}. \tag{8}$$

Since the vanishing point $\tilde{\mathbf{v}}_y$ has already been defined as a function of f and $\mathbf{c}$ in (6), and f depends only on the principal point $\mathbf{c}$ as shown in (7), $\mathbf{r}_1$ and $\mathbf{r}_2$ depend only on $\mathbf{c}$, with two-fold sign ambiguities eliminated again by minimizing the symmetric transfer errors as discussed shortly.

As for the scale factor t_z of the third column of $\mathbf{H}_w$ in (2), we use the fact that in our case the world origin is mapped to a finite point in the image plane, and, hence, the t_z of the viewpoints can not be close to zero. Therefore, since the scale factor of the third column, i.e., t_z, does not introduce ambiguity in the rotation angles [18], we can assume without loss of generality a unit t_z for one of the viewpoints, in which case the following result holds.

Proposition 1: Let the true world-to-image homography be $\mathbf{H}_w = [\,\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3\,]$, and the estimated world-to-image homography be $\hat{\mathbf{H}}_w = [\,\mathbf{h}_1 \quad \mathbf{h}_2 \quad (1/t_z)\mathbf{h}_3\,]$. Then:

1) $\hat{\mathbf{H}}_w^{-1} = \text{diag}(1, 1, t_z)\mathbf{H}_w^{-1}$.
2) Given an inhomogeneous image point $\mathbf{m}$, its corresponding true inhomogeneous world point $\mathbf{M}$ and its estimated inhomogeneous world point $\hat{\mathbf{M}}$, are related via $\hat{\mathbf{M}} = 1/t_z\mathbf{M}$.

From proposition 1, we can compute the t_z of the other viewpoints by forcing correspondences to be projected to the same 3-D world point. Therefore, using (3), (5), (6)–(8), and proposition 1, we have $\mathbf{H}_w$ of each image as a function of the principal point only. In order to find the principal point, and, hence, other camera parameters, we formulate our problem in terms of the interimage homography that minimizes the symmetric transfer errors

$$\min_{\mathbf{c}} \sum_i d(\tilde{\mathbf{m}}_i, \mathbf{H}^{-1}\tilde{\mathbf{m}}'_i))^2 + d(\tilde{\mathbf{m}}'_i, \mathbf{H}\tilde{\mathbf{m}}_i)^2$$

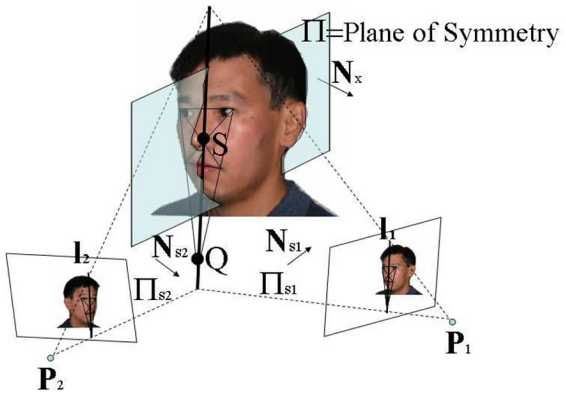


Fig. 2. Our configuration: the plane $\mathbf{\Pi}_s$ that defines the image line $\tilde{\mathbf{l}}$, is different from the plane of symmetry $\mathbf{\Pi}$, whose normal $\mathbf{N}_x$ defines the vanishing point $\tilde{\mathbf{v}}_x$. As a result, in our configuration, $\tilde{\mathbf{l}} \neq \boldsymbol{\omega}\tilde{\mathbf{v}}_x$.

where $d(\cdot,\cdot)$ is the Euclidean distance between the image points and $\mathbf{H} = \hat{\mathbf{H}}'_w \hat{\mathbf{H}}_w^{-1}$ is the interimage homography, which is only a function of $\mathbf{c}$.

The solution is, therefore, found by resorting to standard nonlinear minimization technique [19] with the initial solution obtained by assuming the principal point at the camera center. Once the correct principal point is found, all other camera parameters can be easily computed since they are expressed in terms of $\mathbf{c}$, and the last column of the world-to-image homography is computed up to a scale factor [16].

IV. Relation to Existing Methods and Degenerate Configuration

One approach that seems to resemble our method is the work in [5] using surfaces of revolution. The resemblance, however, is superficial as described below. In [5], Wong *et al.* use the line $\tilde{\mathbf{l}}_s$ corresponding to the projection of the axis of revolution in the image plane for calibration. In their case, this line is the vanishing line of the plane containing the camera center and the axis of revolution. By choosing their $\tilde{\mathbf{v}}_x$ as the vanishing point along the normal to this plane, they obtain a pole-polar relationship between $\tilde{\mathbf{v}}_x$ and $\tilde{\mathbf{l}}_s$, i.e., $\tilde{\mathbf{l}}_s = \boldsymbol{\omega}\tilde{\mathbf{v}}_x$ (see [5] for derivation). This pole-polar relationship provides two linear constraints per image. Therefore, Wong *et al.* can solve for the intrinsic parameters using two images. In our case, however, as depicted in Fig. 2, $\tilde{\mathbf{v}}_x$, is not the vanishing point along the normal of the plain $\mathbf{\Pi}_s$ containing the camera center and the line $\tilde{\mathbf{S}}\tilde{\mathbf{Q}}$. Therefore, it does not have a pole-polar relationship with the vanishing line $\tilde{\mathbf{l}}$ of that plane.

Formally, in our configuration

$$\tilde{\mathbf{v}}_x \sim \mathbf{P}\mathbf{N}_x \text{ and } \mathbf{N}_x \sim \boldsymbol{\Omega}\mathbf{\Pi} \tag{9}$$

where $\boldsymbol{\Omega}$ denotes the absolute dual quadric [10], but, clearly (as seen in Fig. 2), in general, back-projecting the line $\tilde{\mathbf{l}}$ would not yield the plane of symmetry $\mathbf{\Pi}$, i.e.,

$$\mathbf{\Pi} \neq \mathbf{P}^T\tilde{\mathbf{l}}. \tag{10}$$

From (9) and (10), we deduce that, in our case

$$\tilde{\mathbf{l}} \neq \boldsymbol{\omega}\tilde{\mathbf{v}}_x. \tag{11}$$

Therefore, we do not have a pole-polar relationship between the image line $\tilde{\mathbf{l}}$ and the vanishing point $\tilde{\mathbf{v}}_x$. As a result, in our configuration, instead of the two constraints provided by the pole-polar relationship per image, we get one weak constraint in (6). In order to be able to still solve the problem using two images, we resort to the invariance of a cross ratio (7), which provides a nonlinear constraint from two images while introducing a two-fold ambiguity, and solve the problem by minimizing the symmetric transfer errors.

Note that, if the two lines $\tilde{\mathbf{l}}$ and $\boldsymbol{\omega}\tilde{\mathbf{v}}_x$ become parallel or identical, then we loose the constraint in (6), and cannot find the inhomogeneous $\tilde{\mathbf{v}}_y$ as the intersection of these two lines. Line $\tilde{\mathbf{l}}$ is the imaged projection of the 3-D Y axis and, therefore, passes through the image of the world origin $\mathbf{K}[\,\mathbf{r}_1 \;\; \mathbf{r}_2 \;\; \mathbf{r}_3 \;\; -\mathbf{R}\mathbf{C}][\,0 \;\; 0 \;\; 0 \;\; 1\,]^T = -\mathbf{K}\mathbf{R}\mathbf{C}$, where $\mathbf{C}$ is the camera center in the world coordinate system (not to be confused with the principal point, for which we used a lower case character). It also passes through the vanishing point $\tilde{\mathbf{v}}_y \sim \mathbf{K}\mathbf{r}_2$. along the 3-D Y-axis direction. Similarly, line $\boldsymbol{\omega}\tilde{\mathbf{v}}_x$ is the image projection of the 3-D plane of symmetry (Y-Z plane) and, therefore, passes through two vanishing points $\tilde{\mathbf{v}}_y$ and $\tilde{\mathbf{v}}_z (\sim \mathbf{K}\mathbf{r}_3)$. When two lines are parallel to each other, it indicates that the two lines intersect at infinity. By using the property that the cofactor matrix is related to the way matrices distribute with respect to the cross product [17], we have

$$\begin{aligned}
\tilde{\mathbf{l}} \times (\boldsymbol{\omega}\tilde{\mathbf{v}}_x) &\sim ((\mathbf{K}\mathbf{r}_2) \times (\mathbf{K}\mathbf{R}\mathbf{C})) \\
&\quad \times ((\mathbf{K}\mathbf{r}_2) \times (\mathbf{K}\mathbf{r}_3)) && (12)\\
&\sim (\mathbf{K}^*([\mathbf{r}_2]_\times \mathbf{R}\mathbf{C})) \times (\mathbf{K}^*(\mathbf{r}_2 \times \mathbf{r}_3)) && (13)\\
&\sim (\mathbf{K}^*([\,-\mathbf{r}_3 \;\; 0 \;\; \mathbf{r}_1\,]\mathbf{C})) \times (\mathbf{K}^*\mathbf{r}_1) && (14)\\
&\sim (\mathbf{K}^*((C_z\mathbf{r}_1 - C_x\mathbf{r}_3))) \times (\mathbf{K}^*\mathbf{r}_1) && (15)\\
&\sim \mathbf{K}(C_z\mathbf{r}_1 \times \mathbf{r}_1 - C_x\mathbf{r}_3 \times \mathbf{r}_1) && (16)\\
&\sim C_x\mathbf{K}\mathbf{r}_2 && (17)
\end{aligned}$$

where "$*$" denotes the matrix of cofactors, and $C_{x,z}$ are the coordinates of $\mathbf{C}$. Algebraically, $C_x = 0$ implies that the line $\tilde{\mathbf{l}}$ is identical to the vanishing line $\boldsymbol{\omega}\tilde{\mathbf{v}}_x$, in which case our method will degenerate to the method [5] and the pole-polar relationship $\tilde{\mathbf{l}} \sim \boldsymbol{\omega}\tilde{\mathbf{v}}_x$ offers two independent constraints on the camera internal parameters per view. However, the configuration $C_x = 0$ indicates that the camera center lies on the plane of symmetry $\mathbf{\Pi}$ in Fig. 1, which is rare to happen in practice.

V. Experimental Results

We first show the synthetic simulations for the 1-D object (the 2-D case is similar). We then show the results for some real images.

A. Computer Simulation

The simulated camera has a focal length of $f = 1020$, unit aspect ratio, zero skew, and the principal point at $(\,316 \;\; 243\,)$. The image resolution is 640×480. We observed the 1-D objects randomly at seven positions [14]. The 3-D world points are $\mathbf{M}_1 = [\,-75 \;\; 0 \;\; 0\,]^T$, $\mathbf{M}_2 = [\,-45 \;\; -150 \;\; 0\,]^T$, $\mathbf{M}'_1 = [\,75 \;\; 0 \;\; 0\,]^T$, $\mathbf{M}'_2 = [\,45 \;\; -150 \;\; 0\,]^T$. For each observation, we switched the 1-D object between two positions of the 3-D points given above.

1) Performance Versus Noise Level: We used five image pairs with different positions and orientations of the camera. The estimated camera intrinsic and extrinsic parameters were then compared with the ground truth. We added a zero-mean Gaussian noise with standard derivation varying from 0.1 pixels to 1.5 pixels, and run our algorithm over 1000 independent trials. The averaged results are shown in Fig. 3(a)–(d). For noise level of 1.5 pixels, the relative error for the focal length f is 1.97%. The maximum relative error of principal points is around 0.9%. Good performance is achieved for all extrinsic parameters, i.e., relative errors less than 0.22% for t_x and 0.31% for t_y, absolute errors less than 0.82° for θ_x, less than 0.33° for θ_y, and less than 0.09° for θ_z. The error in θ_x is larger than that in θ_z. The

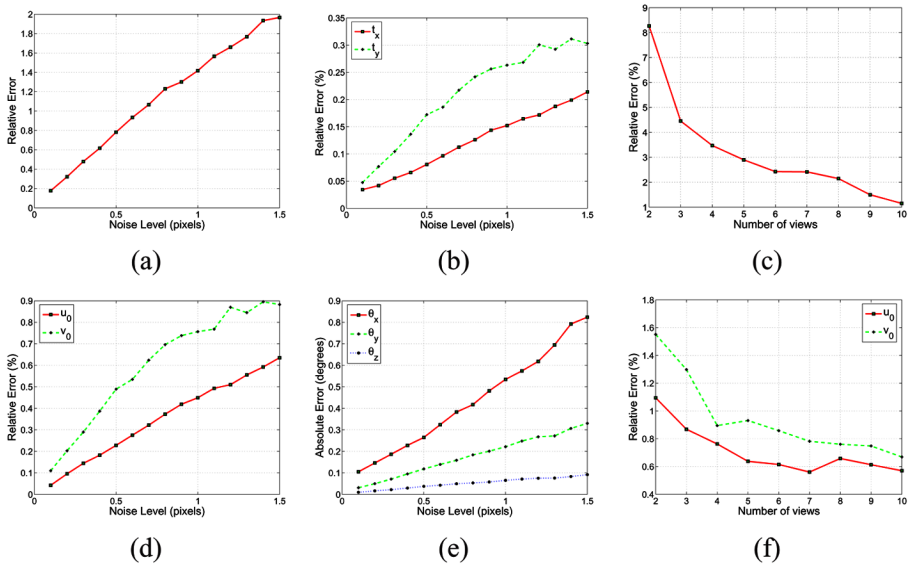


Fig. 3. Performance versus noise (in pixels): (a), (d) relative errors in focal length and the principal point; (b), (e) relative errors in translations and the absolute errors in rotations; (c), (f) relative errors in the focal length and the principal point as a function of the number of images.

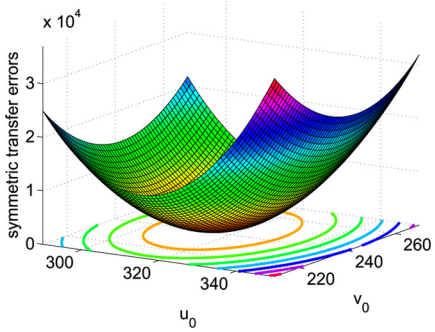


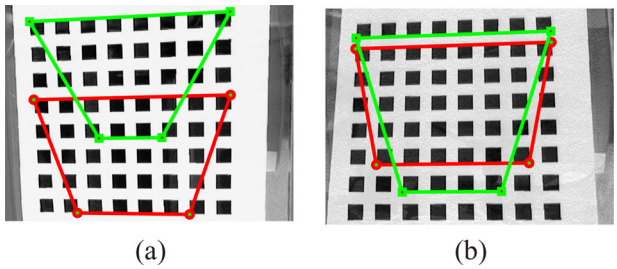
Fig. 4. Error surface over the space of sampled principal points (71×71 with 1.0 pixel intervals for both u_0 and v_0), with added random noise of 1.5 pixels.

main reason is due to the accumulated errors toward the computation of θ_x as detailed in [14]. For example, θ_x is computed from θ_y, θ_z, and all the camera internal parameters, while θ_z depends only on the principal point. Compared to our previous minimization approach [14] of sampling exhaustively the space of principal points, the nonlinear minimization gives us very similar solutions over reasonably big number of trials (> 50). For instance, with a noise level of 1.0 pixel, the relative error is 1.42% using nonlinear solution, compared to 1.33% using exhaustively search. To illustrate the trend in the cost function, we show in Fig. 4 the error surface over the space of sampled principal points.

2) Performance Versus Number of Viewpoints: In this experiment, we varied the number of available viewpoints from 2 to 10. Results are shown in Fig. 3(e) and (f). For this experimentation, the noise level was kept at 1.5 pixels, and the results were averaged over 100 independent trials. From 2 to 4, the errors decrease significantly. The more viewpoints we have, the more accurate camera calibration will be in practice.

B. Real Data

We applied our method to both 2-D and 1-D objects. For 2-D objects, we adopted the data set (five images) used by Zhang [2] with no knowledge about the 3-D coordinates of the calibration grid. The radial distortion was removed according to Zhang's experimental results. We used four trapezoids (i.e., 16 corners) as shown in Fig. 5 for gathering



(a) (b)

Fig. 5. Four trapezoids projected to two images.

TABLE I
CAMERA INTRINSIC PARAMETERS FOR ZHANG'S DATA

	(1234)	(1235)	(1245)	(1345)	(2345)	mean	dev
f [2]	831.82	832.10	837.53	829.91	833.11	832.90	2.84
f	**833.55**	**834.29**	**835.96**	**832.33**	**834.28**	**834.08**	**1.32**
u_0 [2]	304.53	304.32	304.57	303.95	303.53	304.18	0.44
$\mathbf{u_0}$	**303.50**	**304.25**	**303.25**	**304.00**	**304.50**	**303.90**	**0.52**
v_0 [2]	206.79	206.23	207.30	207.16	206.33	206.76	0.48
$\mathbf{v_0}$	**217.25**	**213.25**	**212.25**	**205.25**	**208.50**	**211.30**	**4.60**

TABLE II
CAMERA EXTRINSIC PARAMETERS FOR ZHANG'S REAL DATA

	Rotation angles $(\theta_x, \theta_y, \theta_z)$		$(t_x/t_z, t_y/t_z)\%$	
Img	[2]	Ours	[2]	Ours
1^{st}	(7.4,-16.2,9.5)	(7.0, -16.2,10.2)	(-30,29)	(-30,28)
2^{nd}	(-9.2,-13.6,10.8)	(-9.5,-13.6,11.4)	(-28,26)	(-28,28)
3^{rd}	(8.3, -33.1,8.3)	(7.9,-33.0,9.3)	(-21,27)	(-21,26)
4^{th}	(6.8, -0.1,10.4)	(6.5,-0.1,10.9)	(-27,29)	(-27,29)
5^{th}	N/A	N/A	(-28,22)	(-28,20)

Fig. 6. Three collinear points (marked in red) along a TV antenna.

TABLE III
CAMERA INTRINSIC PARAMETERS FOR 1-D TV ANTENNA

img set	f (rel. err.(%))	u_0 (rel. err.(%))	v_0 (rel. err.(%))
1^{st}	2443.22 (-0.61)	1137.00 (0.12)	838.67 (-0.02)
2^{nd}	2604.29 (5.94)	1143.00 (0.36)	844.00 (0.20)
3^{rd}	2510.88 (2.14)	1125.33 (-0.36)	838.00 (-0.05)
4^{th}	2347.97 (-4.48)	1137.83 (0.15)	838.00 (-0.05)
5^{th}	2472.31 (0.57)	1131.00 (-0.13)	838.00 (-0.05)
6^{th}	2411.91 (-1.88)	1137.00 (0.12)	838.00 (-0.05)
7^{th}	2456.22 (-0.08)	1131.00 (-0.13)	840.67 (0.06)
8^{th}	2418.83 (-1.60)	1131.00 (-0.13)	838.00 (-0.05)

the error statistics. For each quadruples of images, we applied our algorithm independently and the results are shown in Tables I and II. In order to evaluate our results (in bold), we use Zhang's results as the ground truth. Note that we evaluate our focal length f using Zhang's focal length β along the image y axis due to the fact in his data the aspect ratio is practically equal to one. We also used an approach similar to [2] based on estimating the uncertainty of the results using the standard deviation of the estimated internal parameters f and u_0, v_0 . The largest error of the estimated focal length is around 3 pixels, and for the principal point around 11 pixels. For the rotation angles in Table II, we use the relative rotation between the current image and the fifth image, and $\mathbf{R} = \mathbf{R}_z\mathbf{R}_y\mathbf{R}_x$. For the translation parameters, we use the ratios between the components since we can only recover them up to a scale ambiguity.

We used the antenna of a home TV as a 1-D object and took 16 images with 8 different viewpoints. One pair of images observed from the same viewpoint is shown in Fig. 6. Three points along the antenna are chosen to generate three isosceles trapezoids. The calculated intrinsic parameters using these images are listed in Table III. We used the sample mean as the ground truth and show the relative difference with respect to the mean value of f. The largest relative error of f, in our case, is less than 6%.

C. Application to 3-D Reconstruction

To demonstrate the effectiveness of the method, we show an example [20] for reconstructing a partially viewed symmetric object given two views obtained by an unknown camera from unknown arbitrary positions and orientations. Fig. 7(a) and (b) shows the real images of a symmetric object, the printer. We first calibrate the camera as described before, and then apply the optimal triangulation method [17, p. 318] with special consideration of visibility to recover both visible and occluded parts. Some recovered 3-D coordinates (in the format $[X, Y, Z]$) are shown in Fig. 7(a). In order to evaluate the results of reconstruction, we manually measured 3-D distances between randomly selected ten pairs of features points [the black points in Fig. 7(a)], and compared them to our computed one. The statistical results are

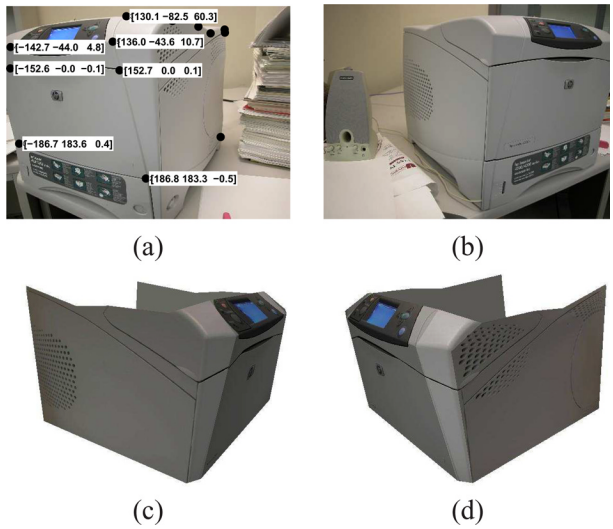


Fig. 7. (a), (b) Two real images of a partially viewed symmetric object; (c), (d) snapshots of the reconstructed 3-D model including the occluded left and right portions.

TABLE IV
ERRORS IN DISTANCES BETWEEN TEN PAIRS OF POINTS

	absolute min	absolute max	mean	std
Error (cm)	0.25	1.50	0.02	0.88

shown in Table IV. Considering the dimensions of the printer in cm as $(w \times h \times d) = (41.2 \times 36.6 \times 43.9)$, the average reconstruction error is around less than 5%. Finally, two snapshots of a piecewise planar model with mapped texture are shown in Fig. 7(c) and (d).

VI. CONCLUSION

We extend the current state-of-the-art calibration methods to situations where only one vanishing point per view is available. The proposed technique is based on deriving the interimage homography as a function of only the principal point, and requires no knowledge of the 3-D coordinates of the calibration object. The method uses images of objects with mirror symmetry, which are frequently found in both indoor and outdoor environments, e.g., a chair, a vase, eye glasses, a face, a car, an airplane, etc. The fact that prior knowledge of the 3-D coordinates of the calibration object is not required, makes the method a versatile tool that can be used without requiring a precisely machined calibration rig (e.g., grids), and also makes calibration possible when the object is not accessible for measurements, e.g., in remote sensing, image-based rendering, or simply when the images are taken by other people as demonstrated in our experimentation with Zhang's real data.

REFERENCES

[1] R. 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, Apr. 1987.
[2] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.
[3] P. Sturm and S. Maybank, "On plane-based camera calibration: A general algorithm, singularities, applications," in *Proc. IEEE CVPR*, 1999, pp. 432–437.
[4] L. Lucchese, "Geometric calibration of digital cameras through multiview rectification," *Image Vis. Comput.*, vol. 23, no. 5, pp. 517–539, 2005.

[5] K. -Y. Wong, R. Mendonça, and R. Cipolla, "Camera calibration from surfaces of revolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 147–161, Feb. 2003.

[6] K. Wong and R. Cipolla, "Reconstruction of sculpture from its profiles with unknown camera positions," *IEEE Trans. Image Process.*, vol. 13, no. 3, pp. 381–389, Mar. 2004.

[7] C. Colombo, A. Bimbo, and F. Pernici, "Metric 3D reconstruction and texture acquisition of surfaces of revolution from a single uncalibrated view," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 1, pp. 99–114, Jan. 2005.

[8] Z. Zhang, "Camera calibration with one-dimensional objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 7, pp. 892–899, Jul. 2004.

[9] O. Faugeras and Q. Luong, *The Geometry of Multiple Images*. Cambridge, NJ: MIT Press, 2001.

[10] B. Triggs, "Autocalibration and the absolute quadric," in *Proc. IEEE CVPR*, 1997, pp. 609–614.

[11] Y. Shan and G. Boon, "Camera self-calibration from video sequences with changing focal length," in *Proc. IEEE Int. Conf. Image Processing*, 1998, pp. 176–180.

[12] M. Pollefeys, R. Koch, and L. V. Gool, "Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters," *Int. J. Comput. Vis.*, vol. 32, no. 1, pp. 7–25, 1999.

[13] W. Lao, Z. Cheng, A. Kam, T. Tan, and A. Kassim, "Focal length self-calibration based on degenerated kruppa's equations: Method and evaluation," in *Proc. IEEE Int. Conf. Image Processing*, 2004, pp. 3391–3394.

[14] X. Cao and H. Foroosh, "Camera calibration without metric information using 1D objects," in *Proc. IEEE Int. Conf. Image Processing*, 2004, pp. 1349–1352.

[15] M. Ahmed and A. Farag, "Nonmetric calibration of camera lens distortion: Differential methods and robust estimation," *IEEE Trans. Image Process.*, vol. 14, no. 8, pp. 1215–1230, Aug. 2005.

[16] X. Cao and H. Foroosh, "Simple calibration without metric information using an isosceles trapezoid," in *Proc. ICPR*, 2004, pp. 104–107.

[17] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[18] G. Jiang, H. T. Tsui, L. Quan, and A. Zisserman, "Single axis geometry by fitting conics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1343–1348, Oct. 2003.

[19] J. More, *The Levenberg-Marquardt Algorithm, Implementation, and Theory*. New York: Springer-Verlag, 1977.

[20] H. Foroosh, M. Balci, and X. Cao, "Self-calibrated reconstruction of partially viewed symmetric objects," in *Proc. IEEE ICASSP*, 2005, pp. 869–872.