Camera Calibration from Symmetry

Kwan-Yee K. Wong, Paulo R. S. Mendonça, and Roberto Cipolla

University of Cambridge, Department of Engineering,
Trumpington Street, Cambridge, CB2 1PZ, UK

Summary. This paper addresses the problem of calibrating a pinhole camera from images of a surface of revolution. Camera calibration is the process of determining the intrinsic or internal parameters (i.e. aspect ratio, focal length and principal point) of a camera, and is important for both motion estimation and metric reconstruction of 3D models. In this paper, a novel and simple calibration technique has been introduced which is based on the symmetry of images of surfaces of revolution. Traditional techniques for camera calibration involve taking images of some precisely machined calibration pattern (such as a calibration grid). The use of surfaces of revolution, which are commonly found in daily life (e.g. bowls and vases), makes the process easier as a result of the reduced cost and increased accessibility of the calibration objects. In this paper, it is shown that 2 images of surface of revolution will provide enough information for determining the aspect ratio, focal length and principal point of a camera. An analytical error model is developed, providing variances and confidence intervals of the parameters estimated. The techniques presented in this paper have been implemented and tested with both synthetic and real data. Experiment results show that the camera calibration method presented here is both practical and accurate.

1 Introduction

An essential step for motion estimation and 3D Euclidean reconstruction, two important tasks in computer vision, is the determination of the intrinsic parameters of cameras. This process, known as *camera calibration*, usually involves taking images of some special patterns with known geometry (see [15,10,8] and [6, Chapter 3]). Such methods do not require direct mechanical measurements on the cameras, and often produce very good results. Nevertheless, they involve the design and use of highly accurate tailor-made calibration patterns, which are both difficult and expensive to manufacture.

In this paper a novel technique for camera calibration is introduced. It relates the ideas from [2,11] for calibration from vanishing points to symmetry properties of images of surfaces of revolution [16,13,17,5,12]. The method presented here allows the camera to be calibrated from two or more images of surfaces of revolution, which are commonly found in daily life (bowls, vases etc.). The use of such objects has the advantage of easy accessibility and low cost, in contrast to the traditional calibration pattern.

Section 2 shows how the symmetry that appears in the images of surfaces of revolution provides information about vanishing points related to a set of three mutually orthogonal directions. By extending the techniques for calibration from vanishing

points, such information can be used in the development of practical algorithms for camera calibration. These algorithms, detailed in Section 3, are capable of dealing with both known and unknown aspect ratio, whereas previous techniques for calibration based on vanishing points can only handle the former case. The error model is described in Section 4. Section 5 first presents results of experiments conducted on synthetic data, which are used to perform an evaluation of the robustness of the algorithm in the presence of noise. Experiments on real data show the usefulness of the proposed method. Finally, conclusions are presented in Section 6.

2   Theoretical Background

The major contributions of [2] and [17] are briefly reviewed in this section, which provides the mathematical background for the algorithms developed in this paper.

In [17] it has been shown that the perspective image of a surface of revolution exhibits a special symmetry which can be expressed in terms of a transformation known as a *harmonic homology* (see details in [14, Chapter IX]). Consider a surface of revolution S. The image of S taken by a pinhole camera $\mathbf{P}$ is a curve ε. Let $\mathbf{l}_\mathrm{s}$ be the image of the axis of revolution of S_r in the camera $\mathbf{P}$. The optical center of $\mathbf{P}$ and the axis of revolution define a plane Π, whose normal direction is $\mathbf{n}_x$. The image of the point at infinity in the direction $\mathbf{n}_x$ is the vanishing point $\mathbf{v}_x$.

If $\mathbf{v}_x$ and $\mathbf{l}_\mathrm{s}$ are represented in homogeneous coordinates as $\mathbf{v}_x = [u\ v\ 1]^\mathrm{T}$ and $\mathbf{l}_\mathrm{s} = [\cos\theta\ \sin\theta\ -d]^\mathrm{T}$, the 2D collineation $\mathbf{W}$ given by

$$\mathbf{W} = \mathbb{I} - 2\frac{\mathbf{v}_x\mathbf{l}_\mathrm{s}^\mathrm{T}}{\mathbf{v}_x^\mathrm{T}\mathbf{l}_\mathrm{s}} \tag{1}$$

is a harmonic homology. The profile ε will be invariant to this transformation, which simply maps one side of the profile (with respect to the image of the axis of rotation) to the other

Consider now any two vectors $\mathbf{n}_y$ and $\mathbf{n}_z$ parallel to Π and orthogonal to each other, which together with $\mathbf{n}_x$ form a set of three mutually orthogonal directions. By construction, the vanishing points corresponding to the directions of $\mathbf{n}_y$ and $\mathbf{n}_z$ will lie on $\mathbf{l}_\mathrm{s}$ (hereafter, referred to as *axis of revolution*).

These three vanishing points can be used to determine the focal length and the principal point of the camera $\mathbf{P}$ assuming that it has zero skew and aspect ratio 1, as shown in [2,3]. In that paper it is proved that the principal point will coincide with the orthocenter of the triangle with vertices given by the vanishing points, and it follows that the square root of the product of the distances from the orthocenter to any vertex and to the opposite side will give the focal length (see Fig. 1(**a**)). As a result, given a harmonic homology $\mathbf{W}$ defined by the vanishing point $\mathbf{v}_x$ and the axis of revolution $\mathbf{l}_\mathrm{s}$, the principal point of the camera $\mathbf{P}$ will lie on a line $\mathbf{l}_x$ passing through $\mathbf{v}_x$ and perpendicular to $\mathbf{l}_\mathrm{s}$. The product of the distances from the principal point to $\mathbf{v}_x$ and to $\mathbf{l}_\mathrm{s}$ will give the square of the focal length (see Fig. 1(**b**)).
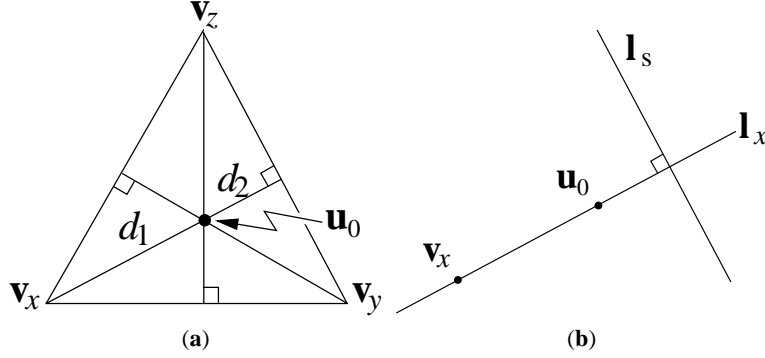
Fig. 1. (**a**) The principal point $\mathbf{u}_0$ of the camera coincides with the orthocenter of the triangle with vertices given by the vanishing points $\mathbf{v}_x$, $\mathbf{v}_y$ and $\mathbf{v}_z$, and the product of the distances d_1 and d_2 is equal to the square of the focal length. (**b**) The vanishing point $\mathbf{v}_x$ and the axis of revolution $\mathbf{l}_s$ define a line $\mathbf{l}_x$ along which the principal point must lie

3   Algorithms and Implementation

3.1   Estimation of the Harmonic Homology W

The profile ε of each surface of revolution is extracted from the image by applying a Canny edge detector (see Fig. 2). The harmonic homology $\mathbf{W}$ that maps each side of the profile ε to its symmetrical counterpart is then estimated by minimizing the geometrical distances between the original profile and its transformed version. This can be done by sampling N evenly spaced points $\mathbf{x}_i$ along the profile ε and optimizing the cost function

$$Cost_{\mathbf{W}}(\mathbf{v}_x, \mathbf{l}_s) = \sum_{i=1}^{N} \mathrm{dist}(\varepsilon, \mathbf{W}(\mathbf{v}_x, \mathbf{l}_s)\mathbf{x}_i)^2, \tag{2}$$

where $\mathrm{dist}(\varepsilon, \mathbf{W}(\mathbf{v}_x, \mathbf{l}_s)\mathbf{x}_i)$ is the distance between the original profile ε and the transformed sample point $\mathbf{W}(\mathbf{v}_x, \mathbf{l}_s)\mathbf{x}_i$. The 4 parameters for the optimization are θ and d, which define the axis of revolution $\mathbf{l}_s$, and ρ and r, which define the vanishing point $\mathbf{v}_x$ in the x-direction (see Section 2).

The success of most non-linear optimization problems requires a good initialization so as to avoid convergence to a local minimum. This is achieved here by using bitangents of the profile [16]. Two points near a bitangent are selected and a polynomial is fitted to the profile in the neighbourhood of each point. The bitangent and the bitangent points can then be obtained analytically from the two polynomials. Given two bitangents $\mathbf{l}(\mathbf{p}_1, \mathbf{p}_2)$ and $\mathbf{l}(\mathbf{q}_1, \mathbf{q}_2)$ on the two sides of the profile ε with bitangent points $\mathbf{p}_1, \mathbf{p}_2$ and $\mathbf{q}_1, \mathbf{q}_2$ respectively (see Fig. 3), the intersection of the two bitangents $(\mathbf{l}(\mathbf{p}_1, \mathbf{p}_2), \mathbf{l}(\mathbf{q}_1, \mathbf{q}_2))$ and the intersection of the diagonals $(\mathbf{l}(\mathbf{p}_1, \mathbf{q}_2), \mathbf{l}(\mathbf{q}_1, \mathbf{p}_2))$ give two points which define a line for an estimate of $\mathbf{l}_s$. An estimate for the vanishing point $\mathbf{v}_x$ is given by the point of intersection of the lines
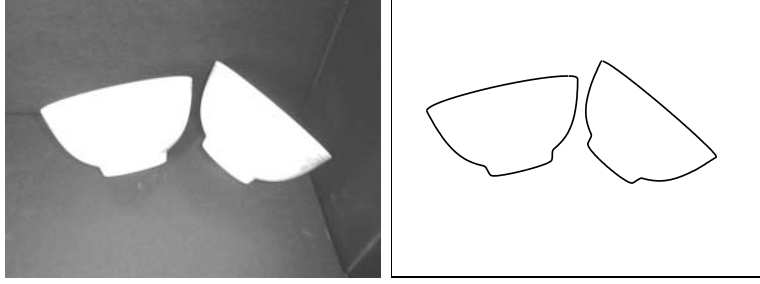
Fig. 2. The profiles of the surfaces of revolution (bowls) are extracted by Canny edge detector

$l(\mathbf{p}_1, \mathbf{q}_1)$ and $l(\mathbf{p}_2, \mathbf{q}_2)$. The initialization of $\mathbf{l}_s$ and $\mathbf{v}_x$ from bitangents often provides an excellent initial guess for the optimization problem. This is generally good enough to avoid any local minimum and allows convergence to the global minimum in a small number of iterations. The estimation of the harmonic homology $\mathbf{W}$ is summarized in Algorithm 1.
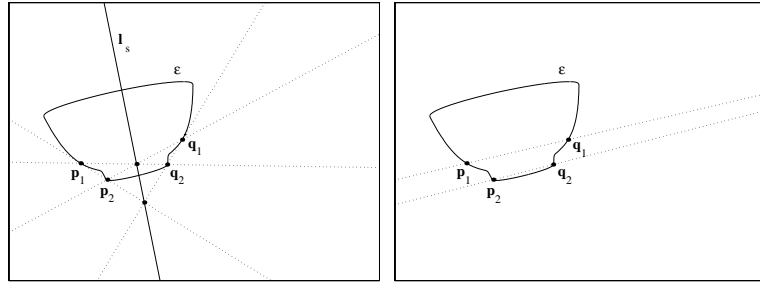

Fig. 3. Initialization of the optimization parameters $\mathbf{l}_s$ and $\mathbf{v}_x$ from the bitangents and lines formed by bitangent points

3.2   Estimation of Intrinsic Parameters

Known Aspect Ratio   When the aspect ratio of the camera is 1, the line $\mathbf{l}_x$ passing through the principal point (u_0, v_0) and the vanishing point $\mathbf{v}_x$ will be perpendicular to the axis of revolution $\mathbf{l}_s$ (see Section 2). As a result, $\mathbf{l}_x$ can be expressed in terms of $\mathbf{v}_x$ and $\mathbf{l}_s$, and is given by $\mathbf{l}_x = [l_2 \ \ -l_1 \ \ l_1v_2 - l_2v_1]$, where $\mathbf{v}_x$ and $\mathbf{l}_s$ are represented by $[v_1 \ v_2 \ 1]^\mathrm{T}$ and $[l_1 \ l_2 \ l_3]^\mathrm{T}$ respectively. Given two such lines $\mathbf{l}_{x1}$ and $\mathbf{l}_{x2}$, the principal point (u_0, v_0) will be given by the point of intersection of the two

---

Algorithm 1 Estimation of the harmonic homology $\mathbf{W}$

---

extract the profile ε of the surface of revolution by using a Canny edge detector;
sample N evenly spaced points $\mathbf{x}_i$ along ε;
initialize the axis of revolution $\mathbf{l}_s$ and the vanishing point $\mathbf{v}_x$ by identifying bitangents in ε;
while not converged **do**
   transform each point $\mathbf{x}_i$ by $\mathbf{W}$;
   compute the distances between ε and the transformed points $\mathbf{W}\mathbf{x}_i$;
   update $\mathbf{l}_s$ and $\mathbf{v}_x$ to minimize the function in (2);
end while

---

lines. When more than two lines are available, the principal point (u_0, v_0) can be estimated by linear least-squares method from

$$
\begin{pmatrix} \mathbf{l}_{x1}^{\mathrm{T}} \\ \mathbf{l}_{x2}^{\mathrm{T}} \\ \vdots \\ \mathbf{l}_{xN}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} su_0 \\ sv_0 \\ s \end{pmatrix} = \mathbf{0},
\tag{3}
$$

where $N \geq 2$ is the total number of lines (i.e. number of profiles) and s is a scale factor. The estimated principal point (u_0, v_0) is then projected on each line $\mathbf{l}_{xi}$ orthogonally as $\mathbf{u}_{0i}$, and the focal length f will be given by

$$
f = \frac{1}{N} \sum_{i=1}^{N} \sqrt{\mathrm{dist}(\mathbf{u}_{0i}, \mathbf{v}_{xi}) \times \mathrm{dist}(\mathbf{u}_{0i}, \mathbf{l}_{si})},
\tag{4}
$$

where $\mathrm{dist}(\mathbf{u}_{0i}, \mathbf{v}_{xi})$ is the distance between the vanishing point $\mathbf{v}_{xi}$ and the projected principal point $\mathbf{u}_{0i}$, and $\mathrm{dist}(\mathbf{u}_{0i}, \mathbf{l}_{si})$ is the distance between the axis of revolution $\mathbf{l}_{si}$ and the projected principal point $\mathbf{u}_{0i}$. Note that the terms for summation are actually the focal lengths estimated from each pair of $\mathbf{v}_{xi}$ and $\mathbf{l}_{si}$ with the estimated principal point projected onto the corresponding $\mathbf{l}_{xi}$ (see Section 2), and the focal length f is taken to be the mean of these estimated values. When the aspect ratio a is known but not 1, there exists a homography $\mathbf{T}(a)$ given by

$$
\mathbf{T}(a) = \begin{pmatrix} \frac{1}{a} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\tag{5}
$$

that transforms the image coordinate system such that the resulting camera would have aspect ratio 1 with the original focal length preserved. After transforming the camera to one with aspect ratio 1, the algorithm presented above can be applied in the same way and the principle point will be given by (au_0, v_0), where (u_0, v_0) are obtained from (3).

Unknown Aspect Ratio   When the aspect ratio a of the camera is unknown, the principal point (u_0, v_0) and focal length f cannot be obtained directly from the vanishing points $\mathbf{v}_{xi}$ and axes of revolution $\mathbf{l}_{si}$. However, a search for a can be performed by optimizing the cost function

$$Cost_a(a) = \text{VAR}(\{\text{dist}(\mathbf{u}'_{0i}, \mathbf{T}(a)\mathbf{v}_{xi})$$
$$\times \text{dist}(\mathbf{u}'_{0i}, \mathbf{T}^{-\text{T}}(a)\mathbf{l}_{si})\}_{i=1}^{N}), \tag{6}$$

where VAR is the variance of the data set. $\mathbf{u}'_{0i}$ is the projection of the principal point $\mathbf{u}'_0$ onto the line passing through $\mathbf{T}(a)\mathbf{v}_{xi}$ and orthogonal to $\mathbf{T}^{-1}(a)\mathbf{l}_{si}$. $\text{dist}(\mathbf{u}'_{0i}, \mathbf{T}(a)\mathbf{v}_{xi})$ is the distance between the transformed vanishing point $\mathbf{T}(a)\mathbf{v}_{xi}$ and the projected principal point $\mathbf{u}'_{0i}$, and $\text{dist}(\mathbf{p}'_{0i}, \mathbf{T}^{-1}(a)\mathbf{l}_{si})$ is the distance between the transformed axis of revolution $\mathbf{T}^{-1}(a)\mathbf{l}_{si}$ and the projected principal point $\mathbf{u}'_{0i}$.

Since the search space is only one-dimensional, instead of performing an extensive search, one can just plot the cost against typical range of a and obtain the result directly (see Fig. 4).
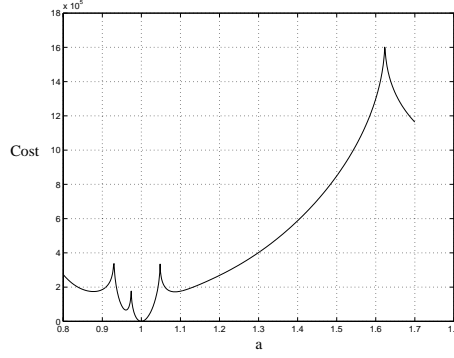

Fig. 4. Plot of the cost function in (6) against a from 0.8 to 1.7 with step size 0.001 for a camera with aspect ratio 1

After convergence, the aspect ratio a obtained can then be used to transform the vanishing points $\mathbf{v}_{xi}$ and axes of revolution $\mathbf{l}_{si}$ using the homography $\mathbf{T}(a)$, and the principal point (u_0, v_0) and focal length f can be recovered as in the case of known aspect ratio.

4   Error Model

An important part of any estimation problem is the determination of the uncertainties of the parameters estimated. It is not enough to state the value of a focal length after calibrating a camera: a confidence interval for the value estimated must also be

provided. A comprehensive introduction to this and related problems can be found in [9].

4.1 Error Propagation

Consider the cost function

$$\xi^2(\boldsymbol{\theta}) = \sum_{i=1}^{N}(\mathbf{y}_i - \mathbf{f}(\mathbf{x}_i; \boldsymbol{\theta}))^{\mathrm{T}}(\mathbf{y}_i - \mathbf{f}(\mathbf{x}_i; \boldsymbol{\theta})). \tag{7}$$

The value $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ that minimizes (7) is a *maximum likelihood estimator* (MLE) for the parameter $\boldsymbol{\theta}$ of the parametric model $\mathbf{y} = \mathbf{f}(\mathbf{x}; \boldsymbol{\theta})$. For noise free-data, i.e., $(\mathbf{x}_i, \mathbf{y}_i)$ exactly satisfying the model, the *residual* $\xi^2(\hat{\boldsymbol{\theta}})$ in (7) will be zero. However, due to the presence of noise, the data will not perfectly fit the model, and $\xi^2(\hat{\boldsymbol{\theta}}) \geq 0$. It is possible to propagate this residual and obtain the corresponding uncertainty for the parameter $\hat{\boldsymbol{\theta}}$ by projecting the points $(2\xi^2(\hat{\boldsymbol{\theta}}), \boldsymbol{\theta}) \cap (\xi^2(\boldsymbol{\theta}), \boldsymbol{\theta})$ onto the plane $(0, \boldsymbol{\theta})$ (see Fig. 5). If $\mathbf{H}|_{\hat{\boldsymbol{\theta}}}$ is the Hessian of $\xi^2(\boldsymbol{\theta})$ computed in $\hat{\boldsymbol{\theta}}$, it can be shown that the covariance matrix $\mathbf{C}(\hat{\boldsymbol{\theta}})$ of $\hat{\boldsymbol{\theta}}$ will be given by $\mathbf{C}(\hat{\boldsymbol{\theta}}) = 2\xi^2\mathbf{H}|_{\hat{\boldsymbol{\theta}}}^{-1}$.
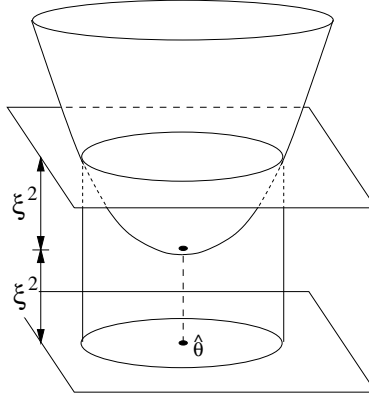

Fig. 5. The covariance matrix of the parameter $\hat{\boldsymbol{\theta}}$ can be found by projecting points of the manifold $(2\xi^2(\hat{\boldsymbol{\theta}}), \boldsymbol{\theta}) \cap (\xi^2(\boldsymbol{\theta}), \boldsymbol{\theta})$ onto the plane $(0, \boldsymbol{\theta})$

Analysis of several covariance matrices for the parameters of $\mathbf{l}_s$ and $\mathbf{v}_x$ showed that the uncertainty is essentially in $\mathbf{v}_x$. This was expected, since the vanishing point $\mathbf{v}_x$ corresponds to the intersection of lines which are nearly parallel. Thus it is justified, in order to simplify the analysis, to ignore the uncertainty in the parameters of $\mathbf{l}_s$. The covariance matrix of u and v in $\mathbf{v}_x$ will be henceforth denoted by $\mathbf{C}(\mathbf{v}_x)$.

4.2   Covariance Matrix of the Principal Point

This error propagation technique can be directly used to determine the covariance matrix $\mathbf{C}(\mathbf{u}_0)$ of the principal point $\mathbf{u}_0$. The computation of the hessian of the function (2) is carried out during the optimization process.

After the computation of the homologies, $\mathbf{u}_0$ is recovered from (3), which can be rewritten as

$$\mathbf{L}_x \begin{bmatrix} \mathbf{u}_0 \\ 1 \end{bmatrix} = \mathbf{0}. \tag{8}$$

If $\mathbf{L}_x = [\mathbf{N}_x - \mathbf{d}]$, then the least squares solution of (3) will be given by $\hat{\mathbf{u}}_0 = \mathbf{N}_x^+ \mathbf{d}$, where $\mathbf{N}_x^+$ is the Moore-Penrose inverse of $\mathbf{N}_x$. Therefore,

$$\mathbf{C}(\hat{\mathbf{u}}_0) = \mathbf{N}_x^+ \mathrm{diag}(\sigma^2(\mathbf{d})) \mathbf{N}_x^{+T}, \tag{9}$$

where $\mathrm{diag}(\sigma^2(\mathbf{d}))$ corresponds to the diagonal matrix whose entries are $\sigma^2(d_i)$.

4.3   Variance of the Focal Length

The computation of the variance of the focal length as estimated by (4) is more elaborated. Each term in the summation in (4) is a separate (but not independent) estimate f_i of f, given by $f_i = \sqrt{d_{i1}}\sqrt{d_{i2}}$. In order to simplify the calculations some hypothesis will be assumed: (a) $\sqrt{d_{i2}}$ is constant; since $d_{i2} = \mathbf{v}_x^{\mathrm{T}} \mathbf{l}_s$, d_{i2} is not constant; nevertheless, the vanishing point $\mathbf{v}_x$ is, in general, tens of thousands of pixels away from $\mathbf{l}_s$; the covariance matrix of $\mathbf{v}_x$, in all the experiments made, indicates an uncertainty of less than 50 pixels for this parameter, justifying the approximation; (b)f_i is independent of f_j if $i \neq j$; in fact, they all depend on $\mathbf{u}_0$, but the influence of $\mathbf{l}_s$ and $\mathbf{v}_x$ is dominant. Under this hypothesis it can be shown that

$$\mathrm{VAR}(f) \approx \frac{1}{N^2} \sum_{i=1}^{N} \frac{1}{4d_{i1}} \mathrm{VAR}(d_{i1}) d_{i2}. \tag{10}$$

Despite all the approximations, (10) and (9) produce good results, as shown by the results of the experiments.

5   Experimental Results

Experiments on both synthetic and real data have been performed, and the results will be presented in the following subsections.

5.1   Synthetic Data

Generation of Data   The experimental setup consists of a surface of revolution which is made up of two spheres intersecting each other. The intrinsic parameters of the synthetic camera are given by the calibration matrix $\mathbf{K}$ as in [6], where focal

length $f = 700$, aspect ratio $a = 1$ and with principal point at $(320, 240)$, and the images have a dimension of 640×480. Each sphere $\mathbf{Q}$ will be projected as a conic $\mathbf{C}$ on the image plane and the projection is given by [4] $\mathbf{C} = (\mathbf{PQ}^{-1}\mathbf{P}^T)^{-1}$, where $\mathbf{P}$ is the 3×4 projection matrix of the camera, $\mathbf{Q}$ is a 4×4 matrix representing the sphere and $\mathbf{C}$ is a 3×3 matrix representing the conic. The profile in each image is found by projecting each sphere into the image as a conic and finding points on each conic that lie outside the other conic. The profiles formed by the three cameras are shown in Fig. 6.
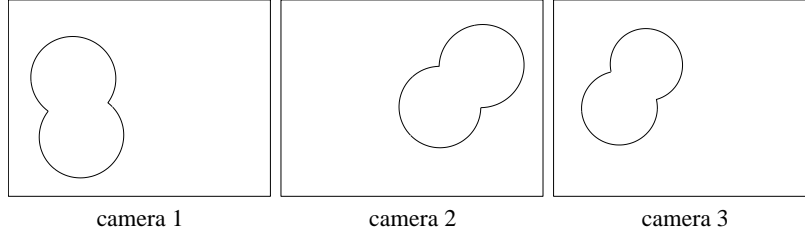

| camera 1 | camera 2 | camera 3 |

Fig. 6. Profiles of the surface of revolution in the images taken by the synthetic camera at the three different positions

In order to evaluate the robustness of the algorithms described in Section 3, different levels of noise have been added along the normal direction of each pixel on the profile. Nevertheless, whereas for corner features the assumption of uncorrelated noise is acceptable, the noise affecting adjacent edgels of profiles exhibits a strong correlation. This is due to the fact that one of the steps of edge detection is a smoothing of the image, which introduces correlation between adjacent pixels. To simulate this effect, uncorrelated uniform noise was convolved with the same Gaussian kernel used to smooth the image, and the output of this filtering was used as to disturb the edgels of the profile.

Experiments and Results Experiments on noise-free data (see Fig. 6) and data with six different noise levels have been performed. The six noise levels are 0.5, 0.7, 1.0, 1.2 and 1.5 pixel respectively. For each noise level, 10 experiments have been conducted using the algorithms described in Section 3 for unknown aspect ratio. One hundred evenly spaced points have been sampled from each profile for the estimation of the harmonic homology $\mathbf{W}$.

The results of the experiments are shown in Table 1 , Table 2 and Table 3. The estimated parameters being shown for each noise level are the mean values over the 10 experiments. The errors being listed are the percentage errors of the mean value of each parameter relative to the ground truth focal length. As the noise level increases, both the relative errors and standard deviations increase. For a noise level of 1.0 pixel, the error for the focal length and principal point is less than 5%.

noise level	mean(f)	std(f)	error(f)
ground truth	700.00	0.00	0.00%
0.0	698.18	0.00	-0.26%
0.5	686.34	13.21	-1.95%
0.7	674.65	23.50	-3.62%
1.0	676.05	28.15	-3.42%
1.2	685.69	7.70	-2.04%
1.5	638.67	30.02	-8.76%

Table 1. Results of calibration of focal length from profiles of surface of revolution using synthetic data with different noise levels

noise level	mean(a)	std(a)	error(a)
ground truth	1.0000	0.0000	0.00%
0.0	1.0000	0.0000	0.00%
0.5	1.0031	0.0063	0.31%
0.7	1.0059	0.0078	0.59%
1.0	1.0014	0.0086	0.14%
1.2	0.9950	0.0027	-0.50%
1.5	1.0080	0.0069	0.80%

Table 2. Results of calibration of the aspect ratio from profiles of surface of revolution using synthetic data with different noise levels

noise level	mean(u_0, v_0)	std(u_0, v_0)	error(u_0, v_0)
ground truth	(320.00,240.00)	(0.00,0.00)	(0.00%,0.00%)
0.0	(321.05,240.06)	(0.00,0.00)	(0.15%,0.01%)
0.5	(318.54,247.18)	(12.73,8.19)	(-0.21,%1.03%)
0.7	(311.87,258.64)	(18.00,13.06)	(-1.16%,2.66%)
1.0	(322.20,256.65)	(22.06,16.96)	(0.31%,2.38%)
1.2	(339.01,247.17)	(7.70,4.09)	(2.72%,1.02%)
1.5	(316.64,253.99)	(17.62,14.94)	(-0.48%,2.00%)

Table 3. Results of calibration of the principal point. Errors are relative to the focal length

5.2   Real Data

Ground Truth The camera used in the experiments for real data is a FUJI MX-700 digital camera and the image size is chosen to be 640×480. The ground truth for the camera's intrinsic parameters is obtained by using a calibration grid.

11 images of the calibration grid have been taken with the camera at different orientations. Corners are extracted from each image using a Canny edge detector and line fitting techniques. For each image, a linear least-squares method [1] is first used to obtain estimates for the 11 parameters of the projection matrix. These estimates

are then used to initialize the optimization for the projection matrix elements which minimize the re-projection errors [7] while enforcing aspect ratio 1 and zero skew.

Experiments and Results A set of real images have been used for the calibration of the digital camera, consisting of 3 images of two bowls, providing 4 profiles of surface of revolution (see Fig. 7). The results of calibration from the images using the algorithm presented in Section 3 for unknown aspect ratio are shown in Table 4. The errors shown are the percentage error of the each parameter relative to the ground truth focal length $f = 685.00$ estimated from the calibration grid.
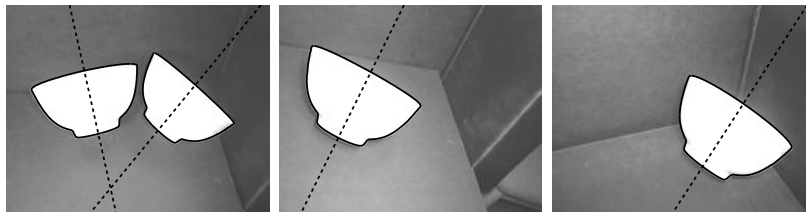

Fig. 7. 3 images of the bowls with the extracted profiles and estimated axes of revolution represented by solid and dash lines respectively

parameter	f	a	u_0	v_0
MLE	697.53	1.0100	318.36	244.40
ground truth	685.00	1.0008	322.60	232.15
error	1.83%	0.9245%	-0.62%	1.79%
predicted std	10.37	-	5.19	3.41

Table 4. Results of calibration using the profiles extracted from the images of the bowls

6   Conclusions

By exploiting the symmetry of surfaces of revolution and properties of vanishing points, a practical technique for camera calibration has been developed. The use of a surface of revolution makes the calibration process easier, in not requiring the use of any precisely adjusted device such as a calibration grid. Note that a surface of revolution can be generated by rotating any arbitrarily shaped object on a turntable [12] and thus the method can be seamlessly incorporated into reconstruction systems based on turntable sequences. The proposed method is promising, as demonstrated by experimental results on synthetic and real data. The focal lengths

are estimated with high accuracy, with an error smaller than 4% for realistic levels of noise.

The experiments with synthetic data have demonstrated that the calibration algorithm is remarkably robust, performing well for any practical level of noise. The estimation of the focal length seems to be biased, and the focal length is consistently underestimated. Nevertheless, this trend was not confirmed by the real data experiments, which produced a nearly 2% overestimation of f. This suggests that the error model for the synthetic data experiments must be improved.

The error analysis developed in Section 4 is accurate, except in the prediction of the range of v_0. This parameter, however, presented instability even for the calibration obtained from the grid, which produced results in the interval 225.69 – 239.7245. It also has a standard deviation larger then all the other parameters, computed from the 11 values found by using the calibration grid. A natural extension of this ideas is to use the error analysis not only to predict uncertainty bounds, but also to improve the quality of the estimation.

Acknowledgements

References

1. Y. I. Abdel-Aziz and H. M. Karara. Direct linear transformation into object space coordinates in close-range photogrammetry. In *Proc. Symp. Close-Range Photogrammetry*, pages 1–18, University of Illinois, Urbana, 1971.
2. B. Caprile and V. Torre. Vanishing points for camera calibration. *Int. Journal of Computer Vision*, 4:127–139, 1990.
3. R. Cipolla, T. Drummond, and D. Robertson. Camera calibration from vanishing points in images of architectural scenes. In *Proc. British Machine Vision Conference*, pages 382–391, Nottingham, UK, 1999.
4. G. Cross and A. Zisserman. Quadric reconstruction from dual-space geometry. In *Proc. 6th Int. Conf. on Computer Vision*, pages 25–31, 1998.
5. R.W. Curwen, C.V. Stewart, and J.L. Mundy. Recognition of plane projective symmetry. In *Proc. 6th Int. Conf. on Computer Vision*, pages 1115–1122, 1998.
6. O. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.
7. O.D. Faugeras and G. Toscani. The calibration problem for stereo. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 15–20, 1986.
8. O.D. Faugeras and G. Toscani. Camera calibration for 3D computer vision. In *Proc. Inter. Workshop Mach. Vision Mach. Intell.*, pages 25–34, 1987.
9. K. Kanatani. *Statistical Optimization for Geometric Computation: Theory and Practice*. Elsevier Science, Amsterdam, 1996.
10. R. K. Lenz and R. Y. Tsai. Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology. *IEEE Trans. Pattern Analysis and Machine Intell.*, 10(5):713–720, 1988.

11. D. Liebowitz and A. Zisserman. Combining scene and auto-calibration constraints. In *Proc. 7th Int. Conf. on Computer Vision*, pages 293–300, 1999.

12. P. R. S. Mendonça, K-Y. K. Wong, and R. Cipolla. Circular motion recovery from image profiles. In B. Triggs, R. Szeliski, and A. Zisserman, editors, *ICCV Vision and Algorithms Workshop: Theory and Practice*, Corfu, Greece, 21–22 September 1999. Springer–Verlag.

13. J. L. Mundy and A. Zisserman. Repeated structures: Image correspondence constraints and 3d structure recovery. In J. L. Mundy, A. Zisserman, and D. Forsyth, editors, *Applications of Invariance in Computer Vision*, volume 825 of *Lecture Notes in Computer Science*, pages 89–106. Springer–Verlag, 1994.

14. J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1952.

15. 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):323–344, 1987.

16. A. Zisserman, D. Forsyth, J. Mundy, and C. A. Rothwell. Recognizing general curved objects efficiently. In J.L Mundy and A. Zisserman, editors, *Geometric Invariance in Computer Vision*, chapter 11, pages 228–251. MIT Press, Cambridge, Mass., 1992.

17. A. Zisserman, J. L. Mundy, D. A. Forsyth, J. Liu, N. Pillow, C. Rothwell, and S. Utcke. Class-based grouping in perspective images. In *Proc. 5th Int. Conf. on Computer Vision*, pages 183–188, 1995.