A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion

Davide Scaramuzza[1], Agostino Martinelli, Roland Siegwart

Swiss Federal Institute of Technology Lausanne (EPFL)

CH-1015 Lausanne, Switzerland

{davide.scaramuzza, agostino.martinelli, roland.siegwart}@epfl.ch

Abstract

In this paper, we present a flexible new technique for single viewpoint omnidirectional camera calibration. The proposed method only requires the camera to observe a planar pattern shown at a few different orientations. Either the camera or the planar pattern can be freely moved. No a priori knowledge of the motion is required, nor a specific model of the omnidirectional sensor. The only assumption is that the image projection function can be described by a Taylor series expansion whose coefficients are estimated by solving a two-step least-squares linear minimization problem. To test the proposed technique, we calibrated a panoramic camera having a field of view greater than 200° in the vertical direction, and we obtained very good results. To investigate the accuracy of the calibration, we also used the estimated omni-camera model in a structure from motion experiment. We obtained a 3D metric reconstruction of a scene from two highly distorted omnidirectional images by using image correspondences only. Compared with classical techniques, which rely on a specific parametric model of the omnidirectional camera, the proposed procedure is independent of the sensor, easy to use, and flexible.

1. Introduction

Accurate calibration of a vision system is necessary for any computer vision task requiring extracting metric information of the environment from 2D images, like in ego-motion estimation and structure from motion. While a number of methods have been developed concerning planar camera calibration [19, 20, 21], little work on omnidirectional cameras has been done, and the primary focus has been on particular sensor types.

For omnidirectional camera is usually intended a vision system providing a 360° panoramic view of the scene. Such an enhanced field of view can be achieved by either using catadioptric systems, obtained by opportunely combining mirrors and conventional cameras, or employing purely dioptric fish-eye lenses [13]. As noted in [1, 3, 11], it is highly desirable that such imaging systems have a single viewpoint [4, 6]. That is, there exists a single center of projection, so that, every pixel in the sensed images measures the irradiance of the light passing through the same viewpoint in one particular direction. The reason a single viewpoint is so desirable is that it permits the generation of geometrically correct perspective images from the pictures captured by the omnidirectional camera. Moreover, it allows applying the known theory of epipolar geometry, which easily permits to perform ego-motion estimation and structure-from-motion from image correspondences only.

Previous works on omnidirectional camera calibration can be classified into two different categories. The first one includes methods which exploit prior knowledge about the scene, such as the presence of calibration patterns [5, 7] or plumb lines [8]. The second group covers techniques that do not use this knowledge. This includes calibration methods from pure rotation [7] or planar motion of the camera [9], and self-calibration procedures, which are performed from point correspondences and epipolar constraint through minimizing an objective function [10, 11].

All mentioned techniques allow obtaining accurate calibration results, but primarily focus on particular sensor types (e.g. hyperbolic and parabolic mirrors or fish-eye lenses). Moreover, some of them require special setting of the scene and expensive equipment [7, 9]. For instance, in [7], a fish-eye lens with a 183° field of view is used as an omnidirectional sensor.

COMPUTER SOCIETY

Here, the calibration is performed by using a half-cylindrical calibration pattern perpendicular to the camera sensor, which rotates on a turntable.

In [8, 10], the authors treat the case of a parabolic mirror. In [8] it is shown that vanishing points lie on a conic section which encodes the entire calibration information. Thus, projections of two sets of parallel lines suffice for intrinsic calibration. However, this property does not apply to non-parabolic mirrors. Therefore, the proposed technique cannot be easily generalized to other kinds of sensors.

Conversely, the methods described in [10, 11, 12, 14] fall in the self-calibration category. These methods require no calibration pattern, nor a priori knowledge about the scene. The only assumption is the capability to automatically find point correspondences in a set of panoramic images of the same scene. Then, calibration is directly performed by epipolar geometry by minimizing an objective function. In [10, 12], this is done by employing a parabolic mirror, while in [11, 14] a fish-eye lens with a view angle greater than 180° is used. However, besides focusing on particular sensor types, the mentioned self-calibration techniques may suffer in case of tracking difficulties and of a small number of features points [16].

All previous calibration procedures focus on particular sensor types, such as parabolic and hyperbolic mirrors or fish-eye lenses. Furthermore, they are strongly dependent on the omnidirectional sensor model they use, which is suitable only when the single effective viewpoint property is satisfied. Although several panoramic vision systems exist directly manufactured to have this property, for a catadioptric system this requires to accurately align the camera and the mirror axes. In addition, the focus point of the mirror has to coincide with the camera optical center. Since it is very difficult to avoid camera-mirror misalignments, an incorrectly aligned catadioptric sensor can lead to a quasi single-viewpoint optical system [2]. As a result, the sensor model used by the mentioned techniques could be suboptimal. In the case of fish-eye lenses the discussion above is analogue.

Motivated by this observation, we propose a calibration procedure which uses a generalized parametric model of the sensor, which is suitable to different kinds of omnidirectional vision systems, both catadioptric and dioptric. The proposed method requires the camera to observe a planar pattern shown at a few different locations. Either the camera or the planar pattern can be freely moved. No a-priori knowledge of the motion is required, nor a specific model of the omnidirectional sensor. The developed procedure is based on the assumption that the circular external boundary of the mirror or of the fish-eye lens (respec-

tively in the catadioptric and dioptric case) is visible in the image. Moreover, we assume that the image formation function, which manages the projection of a 3D real point onto a pixel of the image plane, can be described by a Taylor series expansion. The expansion coefficients, which constitute our calibration parameters, are estimated by solving a two-step least-squares linear minimization problem. Finally, the order of the series is determined by minimizing the reprojection error of the calibration points.

The proposed procedure does not require any expensive equipment. Moreover, it is very fast and completely automatic, as the user is only requested to collect a few images of the calibration pattern. The method was applied to a KAIDAN 360° One VR single-viewpoint mirror mounted on a CCD camera. The system has a vertical view angle greater than 200° and the image size is 900x1200 pixels. After calibration, we obtained an average reprojection error of 1 pixel. In order to test the accuracy of the method, we used the estimated model in a structure from motion problem, and we obtained a 3D metric reconstruction of a scene from two highly distorted omnidirectional images, by using image correspondences only.

The structure of the paper is the following. The omnidirectional camera model and calibration are described in Sec. 2 and 3. The results of the calibration of a real system are given in Sec. 4. Finally, the 3D structure from motion experiment and its accuracy are shown and discussed in Sec. 5.

2. Omnidirectional Camera Model

We want to generalize our procedure to different kinds of single-viewpoint omnidirectional vision systems, both catadioptric and dioptric. In this section we will use the notation given in [11].

In the general omnidirectional camera model, we identify two distinct references: the camera image plane (u', v') and the sensor plane (u'', v''). In Fig. 1 the two reference planes are shown in the case of a catadioptric system. In the dioptric case, the sign of u'' would be reversed because of the absence of a reflective surface. All coordinates will be expressed in the coordinate system placed in O, with the z axis aligned with the sensor axis (see Fig. 1a).

Let X be a scene point. Then, assume $u'' = [u'', v'']^T$ be the projection of X onto the sensor plane, and $u' = [u', v']^T$ its image in the camera plane (Fig. 1b and 1c). As observed in [11], the two systems are related by an affine transformation, which incorporates the digitizing process and small axes misalignments; thus $u'' = Au' + t$, where $A \in \Re^{2x2}$ and $t \in \Re^{2x1}$. Then, let us

introduce the image projection function $\boldsymbol{g}$, which captures the relationship between a point $\boldsymbol{u''}$, in the sensor plane, and the vector $\boldsymbol{p}$ emanating from the viewpoint O to a scene point $\boldsymbol{X}$ (see Fig. 1a). By doing so, the complete model of an omnidirectional camera is

$$\lambda \cdot p = \lambda \cdot \boldsymbol{g}(\boldsymbol{u''}) = \lambda \cdot \boldsymbol{g}(\mathrm{A}\boldsymbol{u'}+\mathrm{t}) = \mathrm{P}\boldsymbol{X}, \quad \lambda > 0, \quad (1)$$

where $\boldsymbol{X} \in \Re^4$ is expressed in homogeneous coordinates; $\mathrm{P} \in \Re^{3\times4}$ is the perspective projection matrix. By calibration of the omnidirectional camera we mean the estimation of the matrices A and t, and the non-linear function $\boldsymbol{g}$, so that all vectors $\boldsymbol{g}(\mathrm{A}\boldsymbol{u'}+\mathrm{t})$ satisfy the projection equation (1). This means that, once the omnidirectional camera is calibrated, we are able to reconstruct, from each pixel, the direction of the corresponding scene point in the real world. We assume for $\boldsymbol{g}$ the following expression

$$\boldsymbol{g}(u'',v'') = (u'', v'', f(u'',v''))^{\mathrm{T}}, \quad (2)$$

where f is rotationally symmetric with respect to the sensor axis. For instance, in the catadioptric case, this corresponds to assume that the mirror is perfectly symmetric with respect to its axis. In general, such an assumption is highly reasonable because both mirror profiles and fish-eye lenses are manufactured with micrometric precision.

(a) (b) (c)
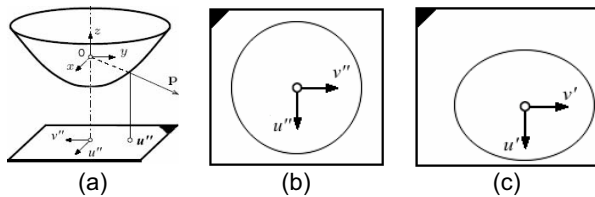
Figure 1. (a) Coordinate system in the catadioptric case. (b) Sensor plane, in metric coordinates. (c) Camera image plane, expressed in pixel coordinates. (b) and (c) are related by an affine transformation.

Function f can have various forms related to the mirror or the lens construction [12, 13, 14]. As mentioned in the introduction, we want to apply a generalized parametric model of f, which is suitable to different kinds of sensors. Moreover, we want this model to compensate for any misalignment between the focus point of the mirror (or the fish-eye lens) and the camera optical center. We propose the following polynomial form for f

$$f(u'',v'') = a_0 + a_1\rho'' + a_2\rho''^2 + ... + a_N\rho''^N, \quad (3)$$

where the coefficients a_i, $i = 0, 1, 2, ...N$, and the polynomial degree N are the model parameters to be determined by the calibration; $\rho'' > 0$ is the metric distance from the sensor axis. Thus, (1) can be rewritten as

$$\lambda \cdot \begin{bmatrix} u'' \\ v'' \\ w'' \end{bmatrix} = \lambda \cdot \boldsymbol{g}(\mathrm{A}\boldsymbol{u'}+\mathrm{t}) = \lambda \cdot \begin{bmatrix} (\mathrm{A}\boldsymbol{u'}+\mathrm{t}) \\ f(u'',v'') \end{bmatrix} = \mathrm{P} \cdot \boldsymbol{X}, \quad \lambda > 0 \ (4).$$

3. Camera Calibration

By calibration of an omnidirectional camera we mean the estimation of the parameters $[\mathrm{A}, \mathrm{t}, a_0, a_1, a_2, ..., a_N]$ so that all vectors $\boldsymbol{g}(\mathrm{A}\boldsymbol{u'}+\mathrm{t})$ satisfy the equation (4). In order to reduce the number of parameters to be estimated, we compute the matrices A and t, up to a scale factor α, by transforming the view field ellipse (see Fig. 1c) into a circle centered on the ellipse center. This transformation is calculated automatically by using an ellipse detector if the circular external boundary of the sensor is visible in the image. After performing the affine transformation, an image point $\boldsymbol{u'}$ is related to the corresponding point on the sensor plane $\boldsymbol{u''}$ by $\boldsymbol{u''} = \alpha \cdot \boldsymbol{u'}$. Thus, by substituting this relation in (4) and using (3), we have the following projection equation

$$(5)$$

$$\lambda \cdot \begin{bmatrix} u' \\ v' \\ w'' \end{bmatrix} = \lambda \cdot \boldsymbol{g}(\alpha \cdot \boldsymbol{u'}) = \lambda \cdot \begin{bmatrix} \alpha \cdot u' \\ \alpha \cdot v' \\ f(\alpha \cdot \rho') \end{bmatrix} = \lambda \cdot \alpha \cdot \begin{bmatrix} u' \\ v' \\ a_0 + ... + a_N\rho'^N \end{bmatrix} = \mathrm{P} \cdot \boldsymbol{X},$$

$$\lambda, \alpha > 0$$

where now u' and v' are the pixel coordinates of an image point with respect to the circle center, and ρ' is the Euclidean distance. Also, note that the factor α can be directly integrated in the depth factor λ; thus, only $N+1$ parameters ($a_0, a_1, a_2, ..., a_N$) need to be estimated.

During the calibration procedure, a planar pattern of known geometry is shown at different unknown positions, which are related to the sensor coordinate system by a rotation matrix $R = [\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{r_3}]$ and a translation $\boldsymbol{t}$, called extrinsic parameters. Let I^i be an observed image of the calibration pattern, $\boldsymbol{M}_{ij} = [X_{ij}, Y_{ij}, Z_{ij}]$ the 3D coordinate of its points in the pattern coordinate system, and $\boldsymbol{m}_{ij} = [u_{ij}, v_{ij}]^{\mathrm{T}}$ the

correspondent pixel coordinates in the image plane. Since we assumed the pattern to be planar, without loss of generality we have $Z_{ij} = 0$. Then, equation (5) becomes

$$(6)$$

$$\lambda_{ij} \cdot \boldsymbol{p_{ij}} = \lambda_{ij} \cdot \begin{bmatrix} u_{ij} \\ v_{ij} \\ a_0 + ... + a_N \rho_{ij}^N \end{bmatrix} = \mathrm{P}^i \cdot \boldsymbol{X} =$$

$$= \begin{bmatrix} \boldsymbol{r_1^i} & \boldsymbol{r_2^i} & \boldsymbol{r_3^i} & \boldsymbol{t^i} \end{bmatrix} \cdot \begin{bmatrix} X_{ij} \\ Y_{ij} \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{r_1^i} & \boldsymbol{r_2^i} & \boldsymbol{t^i} \end{bmatrix} \cdot \begin{bmatrix} X_{ij} \\ Y_{ij} \\ 1 \end{bmatrix}$$

Therefore, in order to solve for camera calibration, the extrinsic parameters have to be determined for each pose of the calibration pattern.

3.1. Solving for camera extrinsic parameters

Before describing how to determine the extrinsic parameters, let us eliminate the dependence from the depth scale λ_{ij}. This can be done by multiplying both sides of equation (6) vectorially by $\boldsymbol{p_{ij}}$

$$\lambda_{ij} \cdot \boldsymbol{p_{ij}} \wedge \boldsymbol{p_{ij}} = \boldsymbol{p_{ij}} \wedge \begin{bmatrix} \boldsymbol{r_1^i} & \boldsymbol{r_2^i} & \boldsymbol{t^i} \end{bmatrix} \cdot \begin{bmatrix} X_{ij} \\ Y_{ij} \\ 1 \end{bmatrix} = 0 \Rightarrow$$

$$\Rightarrow \begin{bmatrix} u_{ij} \\ v_{ij} \\ a_0 + ... + a_N \rho_{ij}^N \end{bmatrix} \wedge \begin{bmatrix} \boldsymbol{r_1^i} & \boldsymbol{r_2^i} & \boldsymbol{t^i} \end{bmatrix} \cdot \begin{bmatrix} X_{ij} \\ Y_{ij} \\ 1 \end{bmatrix} = 0 \qquad (7)$$

Now, let us focus on a particular observation of the calibration pattern. From (7), we have that each point p_j on the pattern contributes three homogeneous equations

$$v_j \cdot (r_{31}X_j + r_{32}Y_j + t_3) - f(\rho_j) \cdot (r_{21}X_j + r_{22}Y_j + t_2) = 0 \quad (8.1)$$

$$f(\rho_j) \cdot (r_{11}X_j + r_{12}Y_j + t_1) - u_j \cdot (r_{31}X_j + r_{32}Y_j + t_3) = 0 \quad (8.2)$$

$$u_j \cdot (r_{21}X_j + r_{22}Y_j + t_2) - v_j \cdot (r_{11}X_j + r_{12}Y_j + t_1) = 0 \qquad (8.3)$$

Here X_j, Y_j and Z_j are known, and so are u_j, v_j. Also, observe that only (8.3) is linear in the unknown

$r_{11}, r_{12}, r_{21}, r_{22}, t_1, t_2$. Thus, by stacking all the unknown entries of (8.3) into a vector, we rewrite the equation (8.3) for L points of the calibration pattern as a system of linear equations

$$M \cdot \mathrm{H} = 0 , \qquad (9)$$

where

$$H = [r_{11}, r_{12}, r_{21}, r_{22}, t_1, t_2]^{\mathrm{T}} ,$$

and

$$M = \begin{bmatrix} -v_1 X_1 & -v_1 Y_1 & u_1 X_1 & u_1 Y_1 & -v_1 & u_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -v_L X_L & -v_L Y_L & u_L X_L & u_L Y_L & -v_L & u_L \end{bmatrix}$$

A linear estimate of H can be obtained by minimizing the least-squares criterion $\min \| M \cdot \mathrm{H} \|^2$, subject to $\| \mathrm{H} \|^2 = 1$. This is accomplished by using the SVD. The solution of (9) is known up to a scale factor, which can be determined uniquely since vectors r_1, r_2 are orthonormal. Because of the orthonormality, the unknown entries r_{31}, r_{32} can also be computed uniquely.

To resume, the first calibration step allows finding the extrinsic parameters $r_{11}, r_{12}, r_{21}, r_{22}, r_{31}, r_{32}, t_1, t_2$ for each pose of the calibration pattern, except for the translation parameter t_3. This parameter will be computed in the next step, which concerns the estimation of the image projection function.

3.2. Solving for camera intrinsic parameters

In the previous step, we exploited equation (8.3) to find the camera extrinsic parameters. Now, we substitute the estimated values in the equations (8.1) and (8.2), and solve for the camera intrinsic parameters $a_0, a_1, a_2, ..., a_N$ that describe the shape of the image projection function $\boldsymbol{g}$. At the same time, we also compute the unknown t_3^i for each pose of the calibration pattern. As done above, we stack all the unknown entries of (8.1) and (8.2) into a vector and rewrite the equations as a system of linear equations. But now, we incorporate all K observations of the calibration rig. We obtain the following system

$$(10)$$

$$\begin{bmatrix} A_1 & A_1\rho_1 & .. & A_1\rho_1^N & -v_1 & 0 & .. & 0 \\ C_1 & C_1\rho_1 & .. & C_1\rho_1^N & .-u_1 & 0 & .. & 0 \\ \vdots & \vdots & .. & \vdots & \vdots & \vdots & .. & \vdots \\ A_K & A_K\rho_K & .. & A_K\rho_K^N & 0 & 0 & .. & -v_K \\ C_K & C_K\rho_K & .. & C_K\rho_K^N & 0 & 0 & .. & -u_K \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ \vdots \\ a_N \\ t_3^1 \\ \vdots \\ t_3^K \end{bmatrix} = \begin{bmatrix} B_1 \\ D_1 \\ \vdots \\ B_K \\ D_K \end{bmatrix}$$

where

$$A_i = r_{21}^i X^i + r_{22}^i Y^i + t_2^i , \quad B_i = v_i \cdot (r_{31}^i X^i + r_{32}^i Y^i) ,$$
$$C_i = r_{11}^i X^i + r_{12}^i Y^i + t_1^i \quad \text{and} \quad D_i = u^i \cdot (r_{31}^i X^i + r_{32}^i X^i) .$$

Finally, the least-squares solution of the overdetermined system is obtained by using the pseudoinverse. Thus, the intrinsic parameters $a_0, a_1, a_2, ..., a_N$, which describe the model, are now available. In order to compute the best polynomial degree N, we actually start from $N=2$. Then, we increase N by unitary steps and we compute the average value of the reprojection error of all calibration points. The procedure stops when a minimum error is found.

4. Experimental Results

The calibration algorithm presented in the previous sections was tested on real data. The omnidirectional sensor to be calibrated is a catadioptric system composed of a KAIDAN 360° One VR hyperbolic mirror and a SONY CCD camera having a resolution of 900x1200 pixels. The calibration rig is a checker pattern containing 9x7 squares, so there are 48 corners (calibration points) (see Fig. 4). The size of the pattern is 24.3cm x 18.9 cm. Eleven images of the plane under different orientations were taken, some of which are shown in Fig. 2.

Figure 2. Some images of the calibration pattern taken under different orientations

4.1. Performance with respect to the number of planes and the polynomial degree

This experiment investigates the performance of our technique with respect to the number of images of the planar pattern, for a given polynomial degree. We vary the number of pictures from 2 to 11, and for each set we perform the calibration. Next, according to the estimated extrinsic parameters, we reproject the 3D

calibration points onto the images. Then, we compute the Root of Mean Squared Distances (RMS), in pixels, between the detected image points and the reprojected ones. The calculated RMS values versus the number of images are plotted in Fig. 3 for different polynomial degrees. Note that the error decreases when more images are used. Moreover, by using a 4th order polynomial to fit the model, we obtain the minimum RMS value, that is of about 1.2 pixels. A 3rd order polynomial also provides a similar performance if more than four images are taken. Conversely, by using a 2nd order expansion, the RMS remains above 2 pixels. Thus, for our applications we used a 4th order expansion. As a result, the RMS error of all reprojected calibration points is 1.2 pixels. This value is very good if we consider that the image resolution is 900x1200 pixels, and that corner detection is less precise on omnidirectional images than on conventional perspective pictures. In Fig. 4 you can see several corner points used to perform the calibration, and the same points reprojected onto the image according to the intrinsic and extrinsic parameters estimated by the calibration.
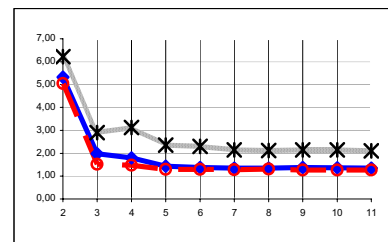

Figure 3. RMS error versus the number of images of the pattern. The RMS values are computed for different polynomial degrees: 2nd order (black ×), 3rd order (blue •) and 4th order (red ∘).
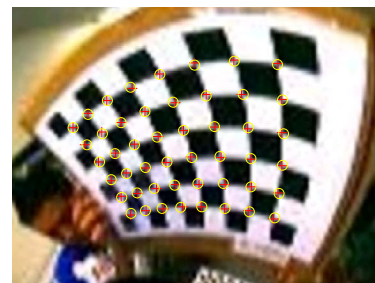

Figure 4. The corner points used for calibration (red crosses) and the reprojected ones (yellow rounds) after calibration.

4.2. Performance with respect to the noise level

In this experiment, we study the robustness of our calibration technique in case of inaccuracy in detecting the calibration points. At this end, Gaussian noise with mean 0 and standard deviation σ is added to the input calibration points. We vary the noise level from 0.1 pixels to 1.5 pixels. For each level, we perform the calibration and we compute the RMS error of the re-projected points. The results obtained using a 4th order polynomial are shown in Fig. 5. As it can be seen, the RMS values remain under 2 pixels.
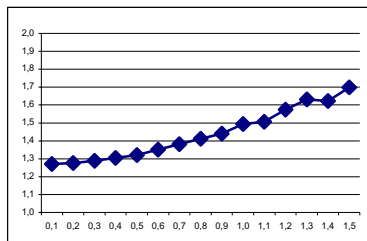

Figure 5. RMS error versus the noise level.

Figure 6. A sample image before rectification.

Figure 7. The sample image of Fig. 6 after rectification. Now the edges (highlighted) appear straight.

4.3. Performance with respect to image rectification

In this experiment, we test the accuracy of the estimated sensor model by rectifying all calibration images. Rectification determines a transformation of the original distorted image such that the new image appears as taken by a conventional perspective camera.

In general, is impossible to rectify the whole omnidirectional image because of a view field larger than 180°. However, it is possible to perform rectification on image regions which cover a smaller field of view. As a result, linearity is preserved in the rectified image. As you can see in Fig. 6, curved edges of a sample pattern in the original image (Fig. 6) appear straight after rectification (Fig. 7).

5. Application to Structure from Motion

Our work on omnidirectional camera calibration is motivated by the use of panoramic vision sensors for structure from motion and 3D reconstruction. In this section, we perform a 3D metric reconstruction of a real object from two omnidirectional images, by using the sensor model estimated by our calibration procedure. In order to compare the reconstruction results with a ground truth, we exploited a trihedral object composed of three orthogonal checker patterns of known geometry (see Fig. 8).


Figure 8. The sample trihedron used for the 3D reconstruction experiment.

Two images of the trihedron were taken by positioning our calibrated camera at two unknown different locations (see Fig. 9). Then, several point matches were picked manually from both views of the object and the *eight point algorithm* [17] was applied. In order to obtain good reconstruction results, more than eight points (actually 135) were extracted. Then, the coordinates of the correspondent 3D vectors, back-projected into the space, were normalized according to the non-uniform mirror resolution. The results of the reconstruction are shown in Fig. 10, where we used checker patches to fit the reconstructed 3D points (red rounds). In order to compare the results with the ground truth, we computed the angles between the three planes fitting the reconstructed points. We found the following values: 94.6°, 86.8° and 85.3°. Moreover, the average distances of these points from the fitted planes were respectively 0.05 cm, 0.75 cm and 0.07 cm. Finally, since we knew the size of each checker to be 6.0 cm x 6.0 cm, we also calculated the dimension of every reconstructed checker, and we found an average error of 0.29 cm.
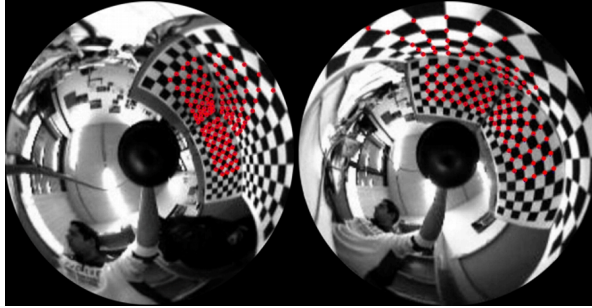
Figure 9. Two pictures of the trihedron taken by the omnidirectional camera. The points used for the 3D reconstruction are indicated by red dots.
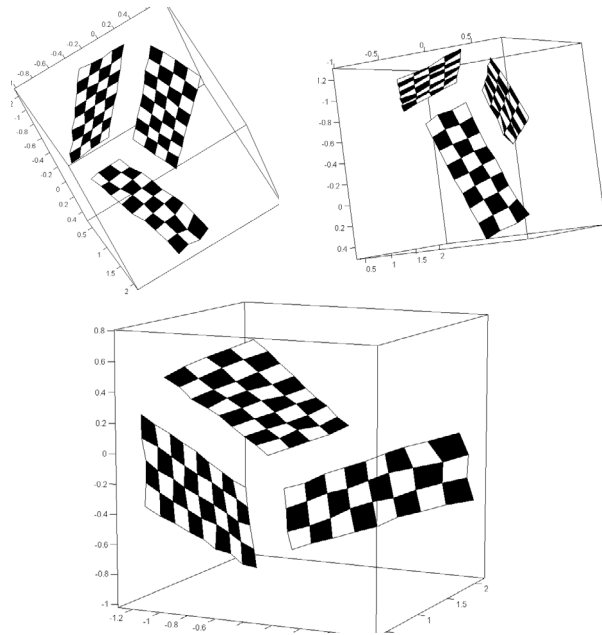
Figure 10. Three rendered views of the reconstructed trihedron. Note that the object was reconstructed only from two highly distorted omnidirectional images (as in Fig. 9).

6. Conclusions

In this paper, we presented a flexible new technique for single-viewpoint omnidirectional camera calibration. The proposed method only requires the camera to observe a planar pattern shown at a few different orientations. No a-priori knowledge of the motion is required, nor a specific model of the omnidirectional sensor. The only assumption is that the image projection function can be described by a Taylor series expansion whose coefficients are estimated by solving a two-step least-squares linear minimization problem. To test the proposed technique, we calibrated a panoramic camera having a field of view greater than 200° in the vertical direction, and we obtained very good results. To investigate the accuracy of the calibration, we also used the estimated omni-camera model in a structure from motion experiment. We obtained a 3D metric reconstruction of a real object from two omnidirectional images, by using image correspondences only. The reconstruction results were also compared with the ground truth. With respect to classical techniques, which rely on a specific parametric model of the omnidirectional camera, the proposed procedure is independent of the sensor, easy to use and flexible.

7. References

[1] Baker, S. and Nayar, S.K. "A theory of catadioptric image formation". In Proceedings of the IEEE International Conference on Computer Vision (ICCV'98), Bombay, India, 1998, pp. 35–42.

[2] R. Swaminathan, M. D. Grossberg, and N. S. K. "Caustics of catadioptric cameras". In Proceedings of the IEEE International Conference on Computer Vision (ICCV'01), Vancouver, Canada, 2001.

[3] B.Micusik, T.Pajdla. "Autocalibration & 3D Reconstruction with Non-central Catadioptric Cameras". In Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR'04), Washington US, 2004.

[4] S. Baker and S. Nayar. "A theory of single-viewpoint catadioptric image formation". International Journal of Computer Vision, 35(2), November 1999, pp. 175–196.

[5] C. Cauchois, E. Brassart, L. Delahoche, and T. Delhommelle. "Reconstruction with the calibrated syclop sensor". In Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS'00), Takamatsu, Japan, 2000, pp. 1493–1498.

[6] T. Svoboda, T. Pajdla, and V. Hlavac. "Central panoramic cameras: Geometry and design". Research report. Czech Technical University - Center for Machine Perception, Praha, Czech Republic, December 1997.

[7] H. Bakstein and T. Pajdla. "Panoramic mosaicing with a 180◦ field of view lens". In Proceedings of the IEEE Workshop on Omnidirectional Vision, 2002, pp. 60–67.

[8] C. Geyer and K. Daniilidis. "Paracatadioptric camera calibration". PAMI, 24(5), May 2002, pp. 687-695.

[9] J. Gluckman and S. K. Nayar. "Ego-motion and omnidirectional cameras". In Proceedings of the IEEE International Conference on Computer Vision (ICCV'98), Bombay, India, 1998, pp. 999-1005.

[10] S. B. Kang. "Catadioptric self-calibration". (CVPR'00), 2000, pp. 201-207.

[11] B. Micusik and T. Pajdla. "Estimation of omnidirectional camera model from epipolar geometry". In Proc. of CVPR'03, 2003, pp. 485.490.

[12] B.Micusik, T.Pajdla. "Para-catadioptric Camera Auto-calibration from Epipolar Geometry". ACCV 2004, Korea, January 2004.

[13] J. Kumler and M. Bauer. "Fisheye lens designs and their relative performance".

[14] B.Micusik, D.Martinec, T.Pajdla. "3D Metric Reconstruction from Uncalibrated Omnidirectional Images". ACCV'04, Korea, (January 2004).

[15] T. Svoboda, T.Pajdla. "Epipolar Geometry for Central Catadioptric Cameras". IJCV, 49(1), Kluwer, August 2002, pp. 23-37.

[16] S. Bougnoux. "From projective to Euclidean space under any practical situation, a criticism of self-calibration". In Proceedings of the 6th International Conference on Computer Vision, Jan. 1998, pp. 790–796.

[17] H.C. Longuet-Higgins. "A computer algorithm for reconstructing a scene from two projections". Nature, Sept 1981, 293:133–135.

[18] R. I. Hartley. "In defence of the 8-point algorithm". In Proceedings of the IEEE International Conference on Computer Vision (ICCV'95), 1995.

[19] Y. Ma, S. Soatto, J. Kosecka, S. Sastry, "An invitation to 3D vision, from images to geometric models models", Springer Verlag, ISBN-0-387-00893-4. 2003.

[20] Q.-T. Luong and O. Faugeras. "Self-calibration of a moving camera from point correspondences and fundamental matrices". The International Journal of Computer Vision, 22(3), 1997, pp. 261–289.

[21] Zhengyou Zhang. "A Flexible New Technique for Camera Calibration", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 22, Issue 11, November 2000, pp.: 1330 – 1334.