Computer Science Department of The University of Auckland
CITR at Tamaki Campus (http://www.citr.auckland.ac.nz/)

CITR-TR-108 December 2001

Calibration of Panoramic Cameras Using 3D Scene Information

Fay Huang, Shou Kang Wei, and Reinhard Klette[1]

Abstract

This paper proposes a novel approach for the calibration of a panoramic camera using geometric information available in real scenes. A panoramic camera possesses some flexibility in acquiring different types of panoramas, such as single-center (e.g. as assumed for QTVR), symmetric stereo, concentric or polycentric panoramas. Panoramic camera are based on the use of line sensors rotating around an axis, and are of increasing value for various applications in computer vision, computer graphics or robotics. Previously developed camera calibration methods (for `standard' camera architectures) are not applicable due to the non-linearity of the panoramic camera, defined by the existence of multiple (nonlinear) optical points and a cylindrical image surface. This article addresses the calibration subject of panoramic cameras for the first time. The paper focuses on the calibration of two dominant parameters that characterize the camera model and provide flexibility in selecting different types of panoramas. It elaborates selected geometric constraints (for increasing numerical stability), experiments with captured image data, an error-sensitivity analysis by simulation, and a discussion why other approaches (designed for `standard' camera architectures) would fail.

[1] Center for Image Technology and Robotics Tamaki Campus, The University of Auckland, Auckland, New Zealand. fay@citr.auckland.ac.nz, shoukang@citr.auckland.ac.nz and r.klette@auckland.ac.nz

Calibration of Panoramic Cameras Using 3D Scene Information

Abstract

This paper proposes a novel approach for the calibration of a panoramic camera using geometric information available in real scenes. A panoramic camera possesses some flexibility in acquiring different types of panoramas, such as single-center (e.g. as assumed for QTVR), symmetric stereo, concentric or polycentric panoramas. Panoramic camera are based on the use of line sensors rotating around an axis, and are of increasing value for various applications in computer vision, computer graphics or robotics. Previously developed camera calibration methods (for 'standard' camera architectures) are not applicable due to the non-linearity of the panoramic camera, defined by the existence of multiple (nonlinear) optical points and a cylindrical image surface. This article addresses the calibration subject of panoramic cameras for the first time. The paper focuses on the calibration of two dominant parameters that characterize the camera model and provide flexibility in selecting different types of panoramas. It elaborates selected geometric constraints (for increasing numerical stability), experiments with captured image data, an error-sensitivity analysis by simulation, and a discussion why other approaches (designed for 'standard' camera architectures) would fail.

1 Introduction

An example of a prototype of a panoramic camera is depicted in Fig. 1. The architecture of the camera allows the acquisition of different panoramas for various applications, such as single-center panoramas for QTVR, route-planning, or robot localization [1, 10]; symmetric panoramas for stereo reconstruction and/or visualization [4, 7, 9]; and concentric or polycentric panoramas for image-based rendering, novel view synthesis, or walk-through animations [3, 5, 6].

Unfortunately such a widely applicable panoramic camera cannot be calibrated using traditional camera calibration methods due to its non-linearity: it possesses multiple (nonlinear) optical points and a cylindrical image surface,

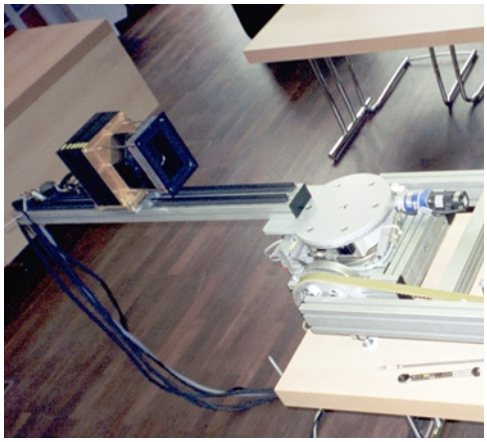

Figure 1. A panoramic camera build at the space sensory institute of DLR (German Aerospace Center).

see Fig. 2. This camera requires a new methodology for the design of a calibration method.

Technically, it is possible to calibrate the camera using production-site facilities. But they are not available for the end user. This paper addresses an on-site camera calibration method utilizing geometric properties available in real scenes.

1.1 Camera Model

A panoramic camera can be geometrically described as follows: let **O** be the rotation center and **C** the optical center, see Fig. 2(a). The camera-viewing angle is defined by the angle ω between optical axis and the line passing through rotation center **O** and optical center **C**. A line sensor is rotated with respect to **O** at the distance R away from the rotation axis (i.e. distance between **O** and **C**). It captures one slit (line) image for subsequent constant-size angular increments during rotation. (Note: each slit image contributes to one column of a panoramic image.) Ideally the parameters remain constant throughout one acquisition process (we neglect minor deviations during one rotation in this article). Different types of panoramas are acquired depending on

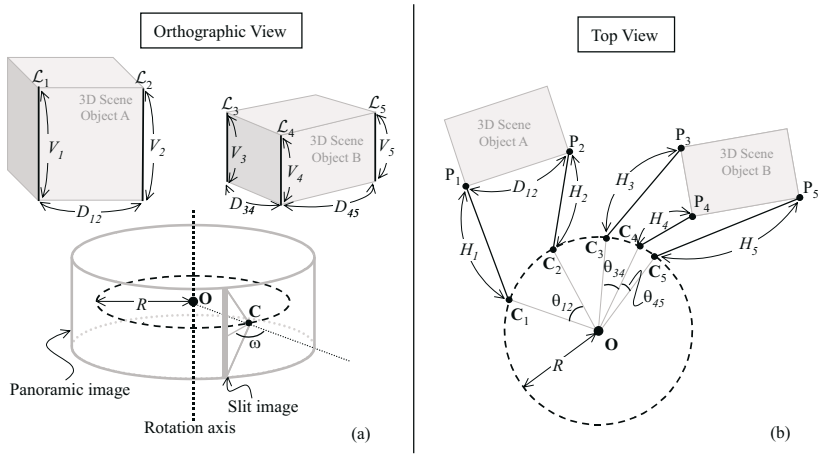

Figure 2. Geometrical relation among parameters.

the specification of the camera parameters R and ω. Failure in measuring these two parameters results in unsatisfactory/unstable outcomes in applications, e.g. an oversimplified assumption of a nodal point perfectly on the rotation axis in the specification of a panorama stitching problem. The paper focuses on calibration of these two dominant parameters.

1.2 Traditional Scenario and Comparisons

Traditionally, if a camera model is available, a straightforward way in camera calibration is to minimize the difference between ideal projections and actual projections of known 3D points (such as calibration objects or localised 3D scene points). In the context of panorama camera calibration, given a set of known 3D points (X_w, Y_w, Z_w) (in world coordinates) and their projections (u, v) (in image coordinates), this means the following minimization:

$$min \sum_{i=1}^{n} \left(\frac{2u_i\,\pi}{W_P} + \omega - \arcsin\left(\frac{X_{io}\,A + Z_{io}\,R\sin\omega}{X_{io}^2 + Z_{io}^2} \right) \right)^2 + \left(\left(v_i - \frac{H_P}{2}\right)\mu - \frac{f Y_{io}}{A - R\cos\omega} \right)^2, \qquad (1)$$

	Dimensionality	Complexity	Initial Value Dependence	Error Sensitivity
Projection-difference approach	[**R**|**T**]* unavoidable	inverse sine and square root	high	exponential growing
Our approach – distance & orthogonality	[**R**|**T**]* avoidable under our assumption	linear form	ignorable	linear growing

* A matrix of the camera extrinsic parameters consists of a rotation matrix **R** and translation vector **T**.

Table 1. Comparisons between approaches.

where $A = \sqrt{X_{io}^2 + Z_{io}^2 - R^2\sin^2\omega}$ and $(X_{io}, Y_{io}, Z_{io})^{\mathrm{T}} = [\mathbf{R_{wo}}| - \mathbf{R_{wo}}\mathbf{T_{wo}}](X_{iw}, Y_{iw}, Z_{iw}, 1)^{\mathrm{T}}$. The matrix $\mathbf{R_{wo}}$ and the vector $\mathbf{T_{wo}}$ describe the rotation and translation of the panoramic camera coordinate system with respect to the world coordinate system. The parameters W_P and H_P are the width and height of a panoramic image in pixels, and μ is the size of one CCD cell.

The complexity of Eq. 1 is rather large involving inverse sine, square root etc. The dimensionality is high compared to our approach (to be specified in the sequel), because extrinsic parameters in $\mathbf{R_{wo}}$ and $\mathbf{T_{wo}}$ are unavoidable in this projection-difference approach. The quality of a calibration result using this approach highly depends on the given initial values, and the error sensitivity showed exponential-growing trends in our experiments. All this motivated us to explore linear geometric relations such that better (numerically more stable) calibration results can be obtained. Table 1 summarizes our comparison of a traditional projection-difference approach with our approach which will be detailed in the rest of the paper.

2 Problem Statement

The problem is to estimate the values of camera parameters R and ω from a given panoramic image based on some knowledge available from 3D scenes (such as distances, lengths, or orthogonalities) under the following assumptions:

We assume there are more than two straight lines in the captured real scene (e.g. a special object with straight edges) which are parallel to the rotation axis. Moreover, for each straight line we assume that: (1) we may identify two points on the line which are visible in the panorama; (2) either there exists another parallel straight line and the

distance between theses lines is known, or there exists two more parallel straight lines such that these three lines are orthogonal (note: see definition in the next section). The effective focal length of the camera and the CCD cell size are assumed to be pre-calibrated using a traditional method, e.g. Tsai's method [8].

The main intention is to find a single linear equation that links 3D geometric scene features to the camera model such that by providing sufficient scene measurements we are able to calibrate the values of R and ω with good accuracy.

3 Geometric Constraints

This section starts with some definitions, followed by describing two selected geometric constraints that are frequently observable in real scenes. Calibration results based on these two constraints are reported in the experimental section.

Note that even though all the measurements are defined in 3D space, the geometrical relation among those can be described on a plane (i.e. in 2D space) since all our lines are assumed to be parallel.

3.1 Definitions

Available straight lines in the 3D scene are denoted and indexed as $\mathcal{L}_i$. The distance, denoted as D_{ij}, between two lines $\mathcal{L}_i$ and $\mathcal{L}_j$ is the length of such a line segment which connects $\mathcal{L}_i$ and $\mathcal{L}_j$ and which is perpendicular to both lines.

If the distance between two straight lines is known, then we say that both lines form a *pair*. One line may be paired up with more than one other line.

We say three parallel lines $\mathcal{L}_i$, $\mathcal{L}_j$, and $\mathcal{L}_k$ are *orthogonal* iff the plane defined by lines $\mathcal{L}_i$ and $\mathcal{L}_j$ and the plane defined by lines $\mathcal{L}_j$ and $\mathcal{L}_k$ are orthogonal. (Note: line $\mathcal{L}_j$ is the intersection of these two planes.)

Figure 2 shows examples of pairs of lines and orthogonal lines. There are three pairs of lines, namely ($\mathcal{L}_1$, $\mathcal{L}_2$), ($\mathcal{L}_3$, $\mathcal{L}_4$), and ($\mathcal{L}_4$, $\mathcal{L}_5$). Lines $\mathcal{L}_3$, $\mathcal{L}_4$, and $\mathcal{L}_5$ are orthogonal lines.

The distance, denote as H_i, between the camera's optical point and the line $\mathcal{L}_i$ is the length of a line segment starting from the optical point and ending at one point on $\mathcal{L}_i$, which is perpendicular to line $\mathcal{L}_i$.

The angular distance, denoted as θ_{ij}, of a pair of lines $\mathcal{L}_i$ and $\mathcal{L}_j$ is the angle between the image columns of their projections on the panoramic image with respect to the rotation axis.

Figure 2(b) shows the geometrical relation between the camera parameters (R and ω) and the calibration parameters (H_i, H_j, D_{ij}, and θ_{ij}).

3.2 By Distance

Consider two lines $\mathcal{L}_i$ and $\mathcal{L}_j$ in 3D space and the image columns of their projections on the panoramic image, denoted as x_i and x_j respectively. Let $d_{ij} = |x_i - x_j|$ in pixel. The angular distance θ_{ij} between lines $\mathcal{L}_i$ and $\mathcal{L}_j$ can be calculated from the value of d_{ij}.

Select two points on each line that are visible in the panorama. The distance of these two points on line $\mathcal{L}_i$ ($\mathcal{L}_j$) is measured and denoted as V_i (V_j), which can also be interpreted as the length of line segment defined by those two points. The length, denoted as $v_i(v_j)$ in pixel, of projections of the corresponding line segment on image column x_i (x_j) can be determined from the image. We use all these information plus the effective camera focal length and the pixel size pre-calibrated to calculate the values of H_i and H_j.

Denote the camera's optical points as $\mathbf{C}_i$ and $\mathbf{C}_j$ associated to image columns x_i and x_j respectively. A 2D coordinate system is defined on the base-plane where all the camera's optical points lie on with the origin meets at the rotation center. The y-axis passes through the camera's optical point $\mathbf{C}_i$.

The position of $\mathbf{C}_i$ can be described by coordinates $(0, R)$ and the position $\mathbf{C}_j$ can be described by coordinates $(R\sin\theta_{ij}, R\cos\theta_{ij})$.

The intersection point of line $\mathcal{L}_i$ and the base-plane, denoted as $\mathbf{P}_i$, can be described by a sum vector of $\overrightarrow{\mathbf{OC}_i}$ and $\overrightarrow{\mathbf{C}_i\mathbf{P}_i}$. Thus, we have

$$\mathbf{P}_i = \begin{pmatrix} H_i \sin\omega \\ R + H_i\cos\omega \end{pmatrix}.$$

Similarly, the intersection point of line $\mathcal{L}_j$ and the base-plane, denoted as $\mathbf{P}_j$, can be described by a sum vector of $\overrightarrow{\mathbf{OC}_j}$ and $\overrightarrow{\mathbf{C}_j\mathbf{P}_j}$. We have

$$\mathbf{P}_i = \begin{pmatrix} R\sin\theta_{ij} + H_j\sin(\theta_{ij}+\omega) \\ R\cos\theta_{ij} + H_j\cos(\theta_{ij}+\omega) \end{pmatrix}.$$

The distance between points $\mathbf{P}_i$ and $\mathbf{P}_j$ is D_{ij}, thus we have the following equation

$$\begin{aligned} 0 &= (1-\cos\theta_{ij})R^2 \\ &+ (H_i+H_j)(1-\cos\theta_{ij})R\cos\omega \\ &- (H_i-H_j)\sin\theta_{ij}R\sin\omega \\ &+ \frac{H_i^2+H_j^2-D_{ij}^2}{2} - H_iH_j\cos\theta_{ij}. \end{aligned} \quad (2)$$

3.3 By Orthogonality

Consider three orthogonal lines $\mathcal{L}_i$, $\mathcal{L}_j$, and $\mathcal{L}_k$ in 3D space. The measures of H_i, H_j, H_k, θ_{ij}, and θ_{jk} are obtained similar to the distance-constraint case, and a 2D coordinate system is defined, also similar to this case.

The position $\mathbf{C}_j$ can be described by coordinates $(0, R)$, the position $\mathbf{C}_i$ by coordinates $(-R\sin\theta_{ij}, R\cos\theta_{ij})$, and the position $\mathbf{C}_k$ by coordinates $(R\sin\theta_{jk}, R\cos\theta_{jk})$.

The intersection points of lines $\mathcal{L}_i$, $\mathcal{L}_j$, and $\mathcal{L}_k$ with the base-plane are denoted as $\mathbf{P}_i$, $\mathbf{P}_j$, and $\mathbf{P}_k$, respectively. We have

$$\mathbf{P}_i = \begin{pmatrix} -R\sin\theta_{ij} + H_i\sin(\omega - \theta_{ij}) \\ R\cos\theta_{ij} + H_j\cos(\omega - \theta_{ij}) \end{pmatrix},$$

$$\mathbf{P}_j = \begin{pmatrix} H_j\sin\omega \\ R + H_j\cos\omega \end{pmatrix},$$

and

$$\mathbf{P}_k = \begin{pmatrix} R\sin\theta_{jk} + H_k\sin(\theta_{jk} + \omega) \\ R\cos\theta_{jk} + H_k\cos(\theta_{jk} + \omega) \end{pmatrix}.$$

Vectors $\overrightarrow{\mathbf{P}_i\mathbf{P}_j}$ and $\overrightarrow{\mathbf{P}_j\mathbf{P}_k}$ are orthogonal, thus we have the following equation

$$\begin{aligned} 0 &= (1 - \cos\theta_{ij} - \cos\theta_{jk} + \cos(\theta_{ij} + \theta_{jk}))R^2 \\ &+ (2H_j - (H_j + H_k)\cos\theta_{ij} - (H_i + H_j)\cos\theta_{jk} \\ &\quad + (H_i + H_k)\cos(\theta_{ij} + \theta_{jk}))R\cos\omega \\ &+ ((H_k - H_j)\sin\theta_{ij} + (H_j - H_i)\sin\theta_{jk} \\ &\quad + (H_i - H_k)\sin(\theta_{ij} + \theta_{jk}))R\sin\omega \\ &+ H_j^2 + H_iH_k\cos(\theta_{ij} + \theta_{jk}) \\ &\quad - H_iH_j\cos\theta_{ij} - H_jH_k\cos\theta_{jk}. \end{aligned} \quad (3)$$

4 Experiments

We describe two experiments: the first demonstrates the practicality of our approach using real image data, and the second is about error sensitivity in dependence of used constraints.

4.1 Real Image and Scene Data

A WAAC camera (see Fig. 1) is used, originally designed for space missions. The camera's effective focal length is $21.7mm$, and the CCD cell size is $0.007mm^2$. The camera mounts on a rotational rig supporting an extension arm up to $1m$. The value of R is manually set to be $10cm$. The camera's viewing angle ω is equal to $155°$ according to our definition (which is $25°$ in terms of WAAC specification).

Figure 3 is a panoramic image taken in a seminar room at DLR in Berlin. The size of the room is about $12m^2$. The image has a resolution of $5,184 \times 21,388$ pixels. Line-pairs (eight pairs in total) are highlighted and indexed, They are used the for estimating R and ω where only the distance constraint is applied in this case. Lengths are manually measured, with an expected error of no more than 0.5% of their readings. The data of these sample lines used for calibration are summarized in Tab. 2.

Index	$V_i = V_j$ (m)	v_i (pixel)	v_j (pixel)	D_{ij} (m)	d_{ij} (pixel)
1	0.0690	91.2	133.8	1.4000	1003.1
2	0.6320	600.8	683.0	1.0000	447.3
3	0.5725	351.4	367.4	1.5500	490.5
4	1.0860	1269.0	1337.6	0.6000	360.9
5	0.2180	273.0	273.6	0.2870	180.1
6	0.0690	81.8	104.2	1.4000	910.5
7	0.5725	318.0	292.0	1.5500	398.2
8	1.3300	831.2	859.4	1.3400	422.5

Table 2. Data used for calibrating the panoramic camera capturing Fig. 1.

We use optimization and sequential quadratic programming [2] for determining R and ω. We minimize the following:

$$min\sum_{i=1}^{n}\left(K_{3n}X_1 + K_{2n}X_2 + K_{1n}X_3 + K_{0n}\right)^2, \quad (4)$$

subject to the equality constraint $X_1 = X_2^2 + X_3^2$, where $K_{in}, i = 0,1,2,3$, are calculated based on the measurements of real scenes and the image, and $X_1 = R^2$, $X_2 = R\cos\omega$, and $X_3 = R\sin\omega$.

The results are summarized as follows. When all pairs are used, we obtain $R = 10.32cm$ and $\omega = 161.68°$. If we select pairs {2,3,4,7,8}, we have $R = 10.87cm$ and $\omega = 151.88°$. If we only use the pairs {2,4,8}, then $R = 10.83cm$ and $\omega = 157.21°$. We conclude that sample selections and the quality of sample data are influential to the calibration results.

4.2 Simulation of Error Sensitivity

Our error sensitivity test addresses for both constraints: the distance between two parallel lines, and the orthogonality of three lines.

Due to unknown error sources for real data, an error sensitivity analysis based on real data might not be reliable. We generated ground-truth data in correspondence to values of the real case (i.e. $R = 10cm$ and $\omega = 155°$) and simulated the errors of all values of H_i, D_{ij}, and θ_{ij}, independently with a maximum of 5% additive random noise (normal distribution). The range of H_i is from $1m$ to $8m$, and the range of θ_{ij} is from 4 to 35 degrees. The sample size is eight.

The mean results (100 trials) are shown in Fig. 4. The results suggest that estimated parameters using the orthogonality constraint are more sensitive to errors than in case of using the distance constraint. The errors of the estimated parameters increase linearly with respect to the input errors. The results may also serve as a reference for error behaviors of calibration results from real image and scene data.

Figure 3. A test panorama image (a seminar room at DLR in Berlin) with indexed line-pairs.

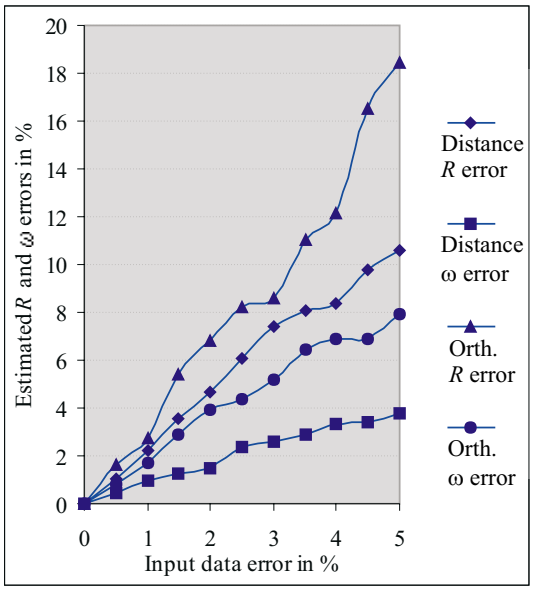

Figure 4. Error sensitivity results.

5 Conclusion and Future work

A novel approach of calibrating a panoramic camera is proposed, using real scene information. Two geometric constraints have been introduced. Simplified formulations of both constraints lead to identical algebraic forms allowing a mixed optimization in accordance with available 'scene structure'. Experiments with real image and scene data illustrate the method's practicability. Our results on error sensitivity analysis suggest that the use of the orthogonality constraint is more sensitive to input-errors than the distance constraint. A comparison between our approach and traditional projection-difference approaches suggests the importance of the exploration of geometric properties/constraints for the calibration of a panoramic camera. Our approach is based on the assumption that reference lines are parallel to the camera's rotation axis. Although the results show that our approach is tolerable to a certain degree of errors, we will attempt to relax this assumption in future studies.

References

[1] S. E. Chen. QuickTimeVR - an image-based approach to virtual environment navigation. In *Proc. SIGGRAPH'95*, pages 29–38, Los Angeles, California, USA, August 1995.

[2] P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, London, 1981.

[3] F. Huang, S. K. Wei, and R. Klette. Geometrical fundamentals of polycentric panoramas. In *Proc. ICCV'01*, pages 560–565, Vancouver, Canada, July 2001.

[4] H. Ishiguro, M. Yamamoto, and S. Tsuji. Omni-directional stereo. *PAMI*, 14(2):257–262, 1992.

[5] L. McMillan and G. Bishop. Plenoptic modeling: an image-based rendering system. In *Proc. SIGGRAPH'95*, pages 39–46, Los Angeles, California, USA, August 1995.

[6] H.-Y. Shum and L.-W. He. Rendering with concentric mosaics. In *Proc. SIGGRAPH'99*, pages 299–306, Los Angeles, California, USA, August 1999.

[7] H.-Y. Shum and R. Szeliski. Stereo reconstruction from multiperspective panoramas. In *Proc. ICCV'99*, pages 14–21, Korfu, Greece, September 1999.

[8] R. 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*, 3(4):323–344, 1987.

[9] S.-K. Wei, F. Huang, and R. Klette. Three-dimensional scene navigation through anaglyphic panorama visualization. In *Proc. CAIP'99 (LNCS 1689)*, pages 542–549, Ljubljana, Slovenia, September 1999.

[10] J.-Y. Zheng and S. Tsuji. Panoramic representation for route recognition by a mobile robot. *IJCV*, 9(1):55–76, 1992.