

Table 1: Summaries of the performances of three different camera calibration approaches under selected criteria.

Approaches	Number of Image Required	Pre-process	Complexity	Dimensionality	Initial Value Dependence	Error Sensitivity	Estimated R	Estimated ω
Point-based	one	identifying projections of 3D points	nonlinear form e.g. sine, fractions and square roots	14 unknowns $[\mathbf{R} \mathbf{T}]^*$ unavoidable	high	exponential growing	actual	actual
Image Correspondence	two	searching image corresponding points	linear form	four unknowns $[\mathbf{R} \mathbf{T}]^*$ not required	ignorable	exponential growing	ratio	actual
Parallel-line-based	one	identifying projections of lines	linear form	three unknowns $[\mathbf{R} \mathbf{T}]^*$ avoidable	ignorable	linear growing	actual	actual

* $[\mathbf{R}|\mathbf{T}]$ is a matrix that characterizes camera extrinsic parameters, where \mathbf{R} is a rotation matrix and \mathbf{T} is a translation vector.

formalizes the basic projective geometry of forming a cylindrical panorama, but evades the complexity of optical systems. This is justified because the optics of a line sensor is calibrated at production sites of line-based cameras. The main purpose of the camera model is the specification of the projection geometry of a cylindrical panorama assuming an ideal line sensor without optical distortion.

Our camera model has multiple projection centers and a cylindrical image surface. The geometry of the camera model is illustrated in Fig. 1, where \mathbf{C} (with subscripts i) denotes the different projection centers. These are uniformly distributed on a circle called *base circle*, drawn as a bold dashed circle. The plane incident with the base circle is called *base plane*. Here, \mathbf{O} denotes the center of the base circle and the off-axis distance R describes the radius of the base circle. The center of the base circle coincides with the center of the cylindrical image surface.

An image is partitioned into *image columns* of equal width which are parallel to the rotation axis. Figure 1 shows two adjacent image columns. The number of image columns is the *width* of a cylindrical panorama, and denoted by W , which is equal to the number of projection centers on the base circle. There is a one-to-one ordered mapping between those image columns and the projection centers. The distance between a projection center and its associated image column is called the *effective focal length* of a panoramic camera, and is denoted as f (cf. Fig. 1). The principal angle ω is the angle between a projection ray, which is emitting from \mathbf{C}_i and incident with the base plane, and the normal vector of the base circle at point \mathbf{C}_i (cf. Fig. 1).

The four parameters, R , f , ω , and W are the defining parameters of our camera model. The values of these parameters characterize how a panoramic image is acquired. For a panoramic image $E_{\mathcal{P}}$ we write $E_{\mathcal{P}}(R, f, \omega, W)$ to specify its camera parameters. Actually, $E_{\mathcal{P}}$ defines a functional into the set of all panoramic images assuming a fixed

rotation axis in 3D space and a fixed center \mathbf{O} on this axis.

3. Calibration

This section presents methods for calibrating off-axis distance R and principal angle ω . These two parameters are characterizing the ‘non-linear component’ of a panorama, and their calibration is the new challenge. The other two parameters, effective focal length f_{μ} (in pixels) and principal row¹ v_c , which characterize the ‘linear component’ of a panorama, are pre-calibrated. These two parameters are calibrated using the same calibration method as for the traditional pinhole cameras [12, 1].

The proofs of all four theorems of this section are given in [3].

3.1. Point-based Approach

A straightforward (traditional) camera calibration approach is to minimize the difference between ideal projections and actual projections of known 3D points, such as on calibration objects, or localized 3D scene points. The same concept is applied in this subsection for the calibration of off-axis distance R and principal angle ω .

Theorem 3.1. *Given a set of known 3D points $(X_{wi}, Y_{wi}, Z_{wi})^T$ in world coordinates and their actual projections $(\hat{u}_i, \hat{v}_i)^T$ in image coordinates, where $i = 1, 2, \dots, n$. The values of R and ω can be estimated by solving the following minimization:*

$$\min\{\|(A_1, A_2, \dots, A_n)^T\|^2 + \|(B_1, B_2, \dots, B_n)^T\|^2\}, \quad (1)$$

where we have

$$A_i = \sin\left(\frac{2\hat{u}_i\pi}{W} + \omega\right) - \frac{X_{oi}\sqrt{X_{oi}^2 + Z_{oi}^2 - R^2 \sin^2 \omega} + Z_{oi}R \sin \omega}{X_{oi}^2 + Z_{oi}^2},$$

$$B_i = \hat{v}_i + v_c - \frac{f_{\mu} Y_{oi}}{\sqrt{X_{oi}^2 + Z_{oi}^2 - R^2 \sin^2 \omega} - R \cos \omega},$$

¹The image row where the panorama intersects with the base plane.

and

$$\begin{pmatrix} X_{oi} \\ Y_{oi} \\ Z_{oi} \end{pmatrix} = \begin{pmatrix} X_{wi}t_{11} + Y_{wi}t_{12} + Z_{wi}t_{13} + t_{14} \\ X_{wi}t_{21} + Y_{wi}t_{22} + Z_{wi}t_{23} + t_{24} \\ X_{wi}t_{31} + Y_{wi}t_{32} + Z_{wi}t_{33} + t_{34} \end{pmatrix}.$$

The parameters f_μ and v_c are assumed to be pre-calibrated. Therefore, there are 14 parameters in total to be estimated using a nonlinear least square optimization method. These 14 parameters consist of the targeted parameters R , ω , and the twelve unknowns in the transformation matrix, namely t_{jk} , where $j = 1, 2, 3$ and $k = 1, 2, 3, 4$.

The objective function in Eqn. (1) is rather complicated. The parameters to be estimated are enclosed in sine functions and square roots involved in both numerator and denominator of the fractions. The dimensionality is high due to the fact that the extrinsic parameters are unavoidable in this approach. Hence, a large set of 3D points is needed for a reasonably accurate estimation.

The quality of calibration results following this approach highly depends on the given initial values for parameter estimation. Our error sensitivity analysis indicates exponential growth. Several assessments for this approach are summarized in Tab. 1.

3.2. Image Correspondence Approach

In this subsection, we investigate the possibility of calibrating R and ω using the information of corresponding image points in two panoramas. Since this approach requires neither scene measures nor a calibration object, thus avoiding a dependency from camera extrinsic parameters in the calibration process, it is surely of interest to see to what extent the camera parameters can be calibrated, and how its performance compares to other approaches.

The general idea of this approach is similar to a recovery of epipolar geometry for a panoramic pair. Epipolar curve equations are used to match provided image points, and the parameter estimation is based on such pairs of corresponding points.

For the camera calibration, we should choose a geometrical representation to be as simple as possible for describing the relation between two panoramas such that a more stable estimation can be obtained. We choose the geometry of a concentric panoramic pair² for detailing an image-correspondence based approach.

The concentric panoramic pair can be acquired in various ways. The authors obtained that the configuration that consists of different off-axis distances, say R_1 and R_2 , and the same principal angle ω , gives the best accuracy for image-correspondence based calibration. Hence we address this

²A pair of panoramas whose rotation axes and associated rotation centers coincide is called *concentric panoramic pair* [4, 5, 10].

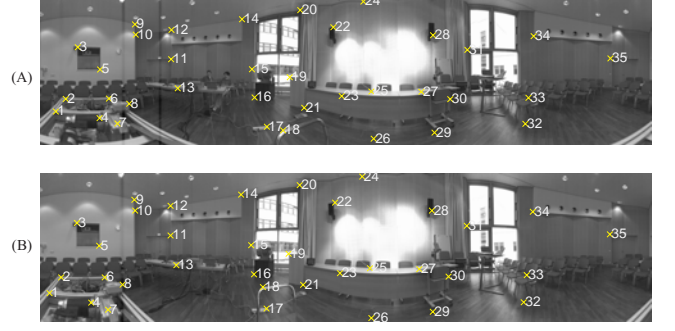


Figure 2: A concentric pair with 35 corresponding image points for the calibration of the camera parameters, off-axis distance R and principal angle ω .

configuration in the following theorem. Note that the effective focal length and the image width are assumed to be identical for both images.

Theorem 3.2. *Given n pairs of corresponding image points (x_{1i}, y_{1i}) and (x_{2i}, y_{2i}) , where $i = 1, 2, \dots, n$, in a concentric pair of panoramas $E_{\mathcal{P}_1}(R_1, f, \omega, W)$ and $E_{\mathcal{P}_2}(R_2, f, \omega, W)$. The ratio $R_1 : R_2$ and ω can be calibrated by minimizing the following cost function:*

$$\min \sum_{i=1}^n (y_{i2} \sin \sigma_i X_1 + (y_{i2} \cos \sigma_i - y_{i1}) X_2 - y_{i1} \sin \sigma_i X_3 + (y_{i1} \cos \sigma_i - y_{i2}) X_4)^2, \quad (2)$$

subject to the equality constraint $X_1 X_4 = X_2 X_3$, where $\sigma_i = \left(\frac{2\pi}{W\mu} (x_{1i} - x_{2i}) \right)$, μ is the pixel size, and $X_1 = R_2 \cos \omega$, $X_2 = R_2 \sin \omega$, $X_3 = R_1 \cos \omega$, and $X_4 = R_1 \sin \omega$.

Once the values of X_1 , X_2 , X_3 , and X_4 are obtained, $\frac{R_1}{R_2}$ and ω can be calculated by

$$\frac{R_1}{R_2} = \frac{\sqrt{X_3^2 + X_4^2}}{\sqrt{X_1^2 + X_2^2}} \quad \text{and} \quad \omega = \arccos \left(\frac{X_1}{\sqrt{X_1^2 + X_2^2}} \right).$$

The objective function of this correspondence-based approach is in linear form and there are only four unknowns

Table 2: Error sensitivity results of the camera calibration based on the image correspondence approach.

Input data error in pixel	Estimated ω error in (%)
0.0	0.00
0.5	0.17
1.0	0.52
1.5	2.78
2.0	10.34
3.0	28.14
4.0	68.25
5.0	155.62

to be estimated, namely $X_1, X_2, X_3,$ and X_4 in Eqn. 2. In this case, at least four pairs of corresponding image points are required. This can be considered to be a great improvement compared to the point-based approach. However, the results of estimated values for real scene data remain to be very poor, i.e. far from the known parameter values. An experiment using a real scene and a concentric panoramic pair is illustrated in Fig. 2, and there are 35 pairs of corresponding image points identified manually, marked by crosses and indexed by numbers. We use the optimization method of sequential quadratic programming [2] for estimating $\frac{R_1}{R_2}$ and ω . This experimental result stimulated further steps into the analysis of error sensitivity.

The authors analyzed the error sensitivity by simulation using synthetic data. The ground-truth data are generated in correspondence to a real case, and the errors are simulated by additive random noise in normal distribution, perturbing the coordinates of ideal corresponding image point-pairs. Calibration results after adding errors are shown in Tab. 3.2. We see that the estimated result is rather sensitive to these errors. The errors of the estimated parameters increase exponentially with respect to the input errors. These results may serve as guidance for studies of error behavior of calibration results for real image and scene data.

Despite the error problem, the correspondence-based approach is unable to recover the actual value of R , which is (of course) one of the main intentions of the intended calibration process. The assessments of this approach is also summarized in Tab. 1.

3.3. Parallel-line-based Approach

In this subsection, we discuss another possible approach that has been widely used for planar pinhole camera calibration, which we call *parallel-line-based* approach. This approach is presented by exploring the related geometric properties such as distances, lengths, orthogonalities of a few straight lines, and transforming them into constraints for estimating the camera parameters. 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.

Constraint 1: Distance

We assume that there are more than two straight line segments in the captured real scene (e.g. straight edges of objects), which are parallel to the rotation axis. The lengths of those line segments are known. All straight line segments are denoted as \mathcal{L} and indexed by a subscript for the distinction of multiple lines. The length of segment \mathcal{L}_i is denoted as H_i . The length of the projection of a line segment on an image column u can be determined from the

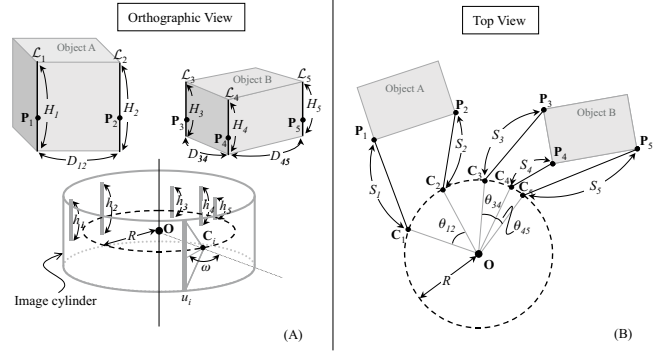


Figure 3: Geometrical interpretations of the parallel-line-based camera calibration approach.

input image, denoted as h in pixels. Examples of H_i and its corresponding h_i values are depicted in Fig. 3(A) where $i = 1, 2, \dots, 5$.

The distance between two lines \mathcal{L}_i and \mathcal{L}_j is the length of a line segment that connects and is perpendicular to both lines \mathcal{L}_i and \mathcal{L}_j . The distance is denoted as D_{ij} . If the distance between two straight lines is measured (in the 3D scene), then we say that both lines form a *line pair*. One line may be paired up with more than one other line. Figure 3(A) shows examples of three line pairs, namely $(\mathcal{L}_1, \mathcal{L}_2)$, $(\mathcal{L}_3, \mathcal{L}_4)$, and $(\mathcal{L}_4, \mathcal{L}_5)$.

Consider two straight lines \mathcal{L}_i and \mathcal{L}_j in 3D space and the image columns of their projections, denoted as u_i and u_j respectively, on a panoramic image. The camera optical centers associated with image columns u_i and u_j , respectively, are denoted as C_i and C_j . The angular distance of two associated image columns of lines \mathcal{L}_i and \mathcal{L}_j is the angle defined by line segments $\overline{C_i O}$ and $\overline{C_j O}$. We denote the angular distance of a line pair $(\mathcal{L}_i, \mathcal{L}_j)$ as θ_{ij} . Examples of angular distances for some line pairs are given in Fig. 3(B). The angular distance θ_{ij} can be calculated in terms of d_{ij} , that is $\theta_{ij} = \frac{2\pi|u_i - u_j|}{W}$.

The distance between a line \mathcal{L}_i and the associated camera optical center (which ‘sees’ the line \mathcal{L}_i) is defined by the length of a line segment starting from the optical center and ending at one point on \mathcal{L}_i such that the line segment is perpendicular to the line \mathcal{L}_i . The distance is denoted as S_i . We can infer the distance S_i by $S_i = \frac{f_{\mu} H_i}{h_i}$.

Theorem 3.3. Given n pairs of $(\mathcal{L}_{it}, \mathcal{L}_{jt})$, where $t = 1, 2, \dots, n$. The values of R and ω can be estimated by solving the following minimization:

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

subject to the equality constraint $X_1 = X_2^2 + X_3^2$, where $K_{st}, s = 1, 2, 3, 4,$ are coefficients, and $X_s, s = 1, 2, 3$ are



Figure 4: Line-based panoramic camera at the Institute of Space Sensor Technology and Planetary Exploration, German Aerospace Center (DLR), Berlin.

three linearly independent variables. We have $X_1 = R^2$, $X_2 = R \cos \omega$, and $X_3 = R \sin \omega$. Moreover, we have

$$\begin{aligned} K_{1t} &= 1 - \cos \theta_{ijt}, \\ K_{2t} &= (S_{it} + S_{jt})(1 - \cos \theta_{ijt}), \\ K_{3t} &= -(S_{it} - S_{jt}) \sin \theta_{ijt}, \quad \text{and} \\ K_{4t} &= \frac{S_{it}^2 + S_{jt}^2 - D_{ijt}^2}{2} - S_{it}S_{jt} \cos \theta_{ijt}, \end{aligned}$$

which can be calculated based on the measurements from real scenes and the image.

The values of R and ω can be found uniquely by

$$R = \sqrt{X_1} \quad \text{and} \quad \omega = \arccos \left(\frac{X_2}{\sqrt{X_1}} \right).$$

Constraint 2: Orthogonality

We say that an ordered triple $(\mathcal{L}_i, \mathcal{L}_j, \mathcal{L}_k)$ of three parallel line segments is *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. It follows that the line \mathcal{L}_j is the intersection of these two planes. For example, in Fig. 3(A), $(\mathcal{L}_3, \mathcal{L}_4, \mathcal{L}_5)$ is an orthogonal triple. Under the condition that $(\mathcal{L}_i, \mathcal{L}_j, \mathcal{L}_k)$ is an orthogonal triple, we may conclude the following theorem without the knowledge of the distance between any line pair. In other words, the values of D_{ij} and D_{jk} can be unknown.

Theorem 3.4. *For any orthogonal triple $(\mathcal{L}_i, \mathcal{L}_j, \mathcal{L}_k)$, we may derive a linear relation which is the same as in the distance-based approach except that the expressions of the four coefficients are different. Hence, the minimization of Eqn. 3 and the calculations of R and ω in the distance-based approach also apply to this modified approach.*

DLR Berlin-Adlershof provided the line camera WAAC, see Fig. 4, for experiments with real images, scenes and

panoramic images. The camera was mounted on a turntable supporting an extension arm with values of R up to 1.0 m. The value of R was set to be 10 cm in our experiments.

Figure 5 shows one of the panoramic images taken in a seminar room of the DLR-Institute of Space Sensor Technology and Planetary Exploration at Berlin. The size of the seminar room is about 120 m^2 . The image has a resolution of $5,184 \times 21,388$ pixels. The pairs of lines (eight pairs in total) are highlighted and indexed. The lengths of those lines are also manually measured, with an expected error of no more than 0.5% of their readings. The data of these sample lines used for the camera calibration are summarized in Tab. 3.3.

The results are summarized as follows: when all pairs are used, we obtain $R = 10.32 \text{ cm}$ and $\omega = 161.68^\circ$. If we select pairs $\{2,3,4,7,8\}$, we have $R = 10.87 \text{ cm}$ and $\omega = 151.88^\circ$. If we only use the pairs $\{2,4,8\}$, then $R = 10.83 \text{ cm}$ and $\omega = 157.21^\circ$. This indicates influences of sample selections and of the quality of sample data onto the calibration results; further investigations of the selection and quality criteria will follow.

We also tested the error sensitivity for both constraints to analyze the impact of errors onto the estimated parameters. Synthetic data are generated in correspondence with the previously used values for real data (i.e. $R = 10 \text{ cm}$ and $\omega = 155^\circ$). Errors in values of S_i , D_{ij} , and θ_{ij} are introduced into the synthetic data independently with a maximum of 5% additive random noise in normal distribution. The range of S_i is from 1 m to 8 m, and the range of θ_{ij} is from 4° to 35° . The sample size is eight. The average results of 100 trials are shown in Fig. 6. The results suggest that estimated parameters using the orthogonality constraint are more sensitive to errors than in the case of using the distance constraint. The errors of the estimated parameters increase linearly with respect to the input errors for both cases.

4. Conclusion and Future Work

This paper discussed three different approaches, namely point-based, image correspondence, and parallel-line-

Table 3: Parallel-line-based panoramic camera calibration measurements associate with the panorama shown in Fig. 5.

Index	$H_i = H_j$ (m)	h_i (pixel)	h_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



Figure 5: A test panorama image (a seminar room at DLR Berlin-Adlershof) with indexed line-pairs.

based, for calibrating line-based panoramic cameras. The experiments allowed to compare these approaches with respect to the performance criteria specified in Tab. 1. As a result of the comparison shown in Tab. 1, we conclude that the parallel-line-based approach allows the most accurate calibration results as well as the best numerical stability among these three studied approaches.

In future work, we are interested in looking into utilizing different geometric features (e.g. conic), properties, or ‘hybrid’ for avoiding the strict assumption that the rotation axis must be parallel to the calibration lines in the parallel-line-based approach. Also, in the image correspondence approach, we will investigate other possible configurations of panoramic pairs such that the actual value of off-axial distance R can be calibrated.

References

- [1] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*, The MIT Press, London, England, 1993.

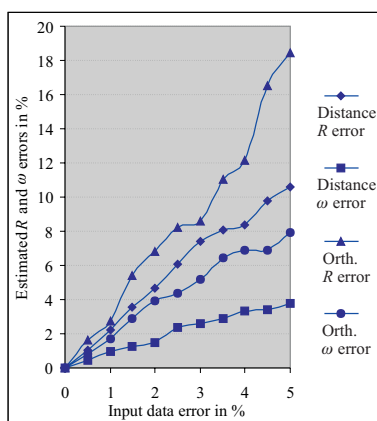


Figure 6: Error sensitivity results of parallel-line-based approach.

- [2] P.E. Gill, W. Murray, and M.H. Wright. *Practical Optimization*, Academic Press, London, 1981.
- [3] F. Huang. *Epipolar Geometry and Camera Calibration of Cylindrical Panoramas*, PhD Thesis, CITR, Computer Science Department, The University of Auckland, Auckland, New Zealand, 2002.
- [4] F. Huang and T. Pajdla. “Epipolar geometry in concentric panoramas”. Research Report CTU–CMP–2000–07, Center for Machine Perception, Czech Technical University, Prague, Czech Republic, March 2000.
- [5] F. Huang, S.-K. Wei, and R. Klette. “Geometrical Fundamentals of Polycentric Panoramas”. *ICCV’01*, pages I:560–565, Vancouver, Canada, July, 2001.
- [6] H. Ishiguro, M. Yamamoto, and S. Tsuji. “Omni-directional stereo”. *PAMI*, 14(2): 257–262, 1992.
- [7] D.W. Murray. “Recovering range using virtual multicamera stereo”. *CVIU*, 61 (2): 285–291, 1995.
- [8] S. Peleg, Y. Pritch, and M. Ben-Ezra. “Cameras for stereo panoramic imaging”. *CVPR’00*, pages 208–214, Hilton Head Island, June, 2000.
- [9] R. Petty, M. Robinson, and J. Evans. “3D measurement using rotating line-scan sensors”. *Measurement Science and Technology*, 9(3): 339–346, 1998.
- [10] H.-Y. Shum and L.-W. He. “Rendering with concentric mosaics”. *SIGGRAPH’99*, pages 299–306, Los Angeles, California, USA, August, 1999.
- [11] H.-Y. Shum and R. Szeliski. “Stereo reconstruction from multiperspective panoramas”. *ICCV’99*, pages 14–21, Korfú, Greece, September, 1999.
- [12] 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*, 3(4):323–344, 1987.
- [13] S.-K. Wei, F. Huang, and R. Klette. “The Design of a Stereo Panorama Camera for Scenes of Dynamic Range”. *ICPR’2002*, pages 635–638, Quebec, Canada, August, 2002.