Comparison of Calibrations from Different Camera Models with Fisheye Cameras

Guan Gui

SCPD student, guan.gui@kla-tencor.com

Fisheye cameras with large field-of-view (FOV) have numerous applications for vision-based motion estimation and 3D reconstruction, while its strong optical aberrations cause challenges in camera calibration. Here we review and summarize different camera models for wide FOV cameras from previous literatures. We also perform a camera calibration experiment with a clip-on fisheye lens and a homemade checkerboard. By summarizing the calibration and test results, we compare and discuss the performance of different camera models.

I. Introduction

Among various types of camera and imaging systems, fisheye cameras are special due to its capability to acquire visual information from a wide field of view (FOV). Large FOV is desired for several reasons: first, it is easier to capture more textured regions in the environment, which is required for stable vision-based motion estimation. Second, with a large field-of-view, large camera motions can be mapped to smaller pixel motions compared to cameras with a smaller field-of-view at the same resolution [1]. This ensures small optical flow between consecutive frames, which is particularly beneficial for direct methods. Fisheye cameras can produce hemisphere images that covers ~180 degree FOV, while strong optical aberrations are generated intentionally. With strong aberrations introduced, calibration and depth estimation with fisheye cameras pose extra challenges.

In this project, I will learn and summarize several camera models for cameras with large FOV from previous literatures. I will practice with camera calibration with experimental data taken from fisheye cameras, compare projection errors between different camera models. I will summarize the results and compare them with former literatures.

Fig. 1. A typical image from fisheye cameras. Strong optical distortion is presented at the edge of the image.

II. Previous work

Fisheye cameras can provide large FOV but suffer from significant optical aberration. The major aberration comes from the third-order distortion term, which causes changes in magnification (angular resolution) versus image height. As shown in Fig.1, the scale at the edge of the image is significantly different from the scale at the image center. Therefore, it's not surprising that a nonlinear mapping is needed to correct this distortion.

The most common camera model is the ***pinhole camera model***, as we learned in the class. In a pinhole camera model, the 3D point (x, y, z) in the world coordinate is projected to a 2D image plane (u, v) by

a simple geometric projection:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x \dfrac{x}{z} \\ f_y \dfrac{y}{z} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix}$$

where the f_x, f_y indicates the focal length and pixel size information, and c_x, c_y represents the camera center shift. Here we only include the intrinsic parameters by assuming that the z-axis in the world coordinate is aligned with the principle axis of the camera. If we consider the extrinsic parameters, we will have 6 more parameters, *i.e.,* 3 rotation parameters and 3 translation parameters.

The unified camera model (UCM) is a nonlinear projection that can be used for large FOV cameras. It's typically used for catadioptric cameras and can be applied to fisheye cameras [2]. The basic idea of UCM is to first project the 3D point (x, y, z) on a sphere surface and then a pinhole camera model. The projection can be formulized as:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x \dfrac{x}{\alpha d + (1-\alpha) z} \\ f_y \dfrac{y}{\alpha d + (1-\alpha) z} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix}$$
$$d = \sqrt{x^2 + y^2 + z^2}$$

where $\alpha \in [0,1]$ is an extra parameters to control the distance between the center of the sphere and the origin point, as illustrated in Fig.2 (a). When $\alpha = 0$, the model degrades to the pinhole camera model.

By generalizing the sphere in the UCM to an ellipsoid, the UCM can be generalized to ***the extended unified camera model (EUCM)*** [3]. The sphere is transformed to an ellipsoid by introducing another parameter β. The projection formula is similar to UCM and can be written as:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x \dfrac{x}{\alpha d + (1-\alpha) z} \\ f_y \dfrac{y}{\alpha d + (1-\alpha) z} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix}$$
$$d = \sqrt{\beta(x^2 + y^2) + z^2}$$

Note that the EUCM degrades to regular UCM when $\beta = 1$.

Another previous work proposes that ***Kannala-Brandt (KB) camera model*** fits well with the fisheye camera model [4]. The KB model introduced a nonlinear term of the projection angle $d(\theta)$ as the correction for wide FOV, where the nonlinear parameters can be tuned. The concept of this model is illustrated in Fig.2 (b) and the projection is defined as:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x d(\theta) \dfrac{x}{r} \\ f_y d(\theta) \dfrac{y}{r} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix}$$
$$r = \sqrt{x^2 + y^2}$$
$$\theta = \text{atan}\,(r/z)$$
$$d(\theta) = \theta + k_1\theta + k_2\theta^3 + \cdots$$

Note that if $k_1 = k_2 = \cdots = 0$, the KB model degrades to a linear model like the pinhole camera model.

Another previously proposed model is ***the field-of-view camera model*** [5]. The projection geometry is illustrated in Fig.2 (c), where the displacement of the projection from the optical center is proportional to the angle between the point and optical axis. The projection is written as:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x r_d \dfrac{x}{r} \\ f_y r_d \dfrac{y}{r} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix}$$
$$r = \sqrt{x^2 + y^2}$$
$$r_d = \frac{\text{atan}\,(2r\tan(w/2)/z)}{w}$$

Note that when $w \to 0$, the FOV model is close to the pinhole camera model.

The last camera model is ***the double-sphere (DS) camera model*** proposed by [1]. The DS camera model is similar to the UCM model but using two spheres during the projection process, as illustrated in Fig.2 (d). An extra parameter is introduced in the DS model and the projection is summarized as the follows:

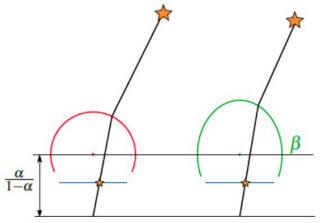
$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x \dfrac{x}{\alpha d_2 + (1-\alpha)(\xi d_1 + z)} \\ f_y \dfrac{y}{\alpha d_2 + (1-\alpha)(\xi d_1 + z)} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix}$$

$$d_1 = \sqrt{x^2 + y^2 + z^2}$$

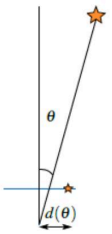
$$d_2 = \sqrt{x^2 + y^2 + (\xi d_1 + z)^2}$$

where ξ is the separation between the two sphere centers and the DS model is degraded to the UCM when $\xi = 0$.

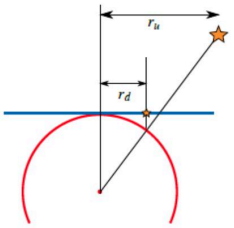
(a)



(b)



(c)



(d)

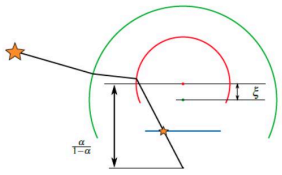


Fig. 2. Concept figure of different camera models. (a) pinhole model, (b) UCM, (c) EUCM, (d) KB model, (e) FOV model, (f) DS model.

III. Approach

To test these different camera models, we get a clip-on fisheye camera for my phone and take an image of a homemade checkerboard, as shown in Fig.3. We can see a strong barrel distortion at the edge of the image, indicating that this is good testbed for camera models with wide FOV. For camera calibration, we use the 14 red dots in Fig.3 to optimize the parameters in the camera models. Note that these camera models have 4-6 intrinsic parameters and 6 extrinsic parameters, while 14 red dots provide 14x2=28 constraints, which is enough for model parameter fitting. Once the camera model parameters are optimized, we use the 14 green dots to test the model. The loss function is defined as the averaged distances between model predictions and ground truth:

$$loss = \frac{1}{N} \sum_{i=1:N} \sqrt{(x_{i,pred} - x_{i,GT})^2 + (y_{i,pred} - y_{i,GT})^2}$$

where N is the number of test dots, $(x_{i,pred}, y_{i,pred})$ are predicted coordinates on the camera plane from the model, and $(x_{i,GT}, y_{i,GT})$ is the ground truth from the measurement.

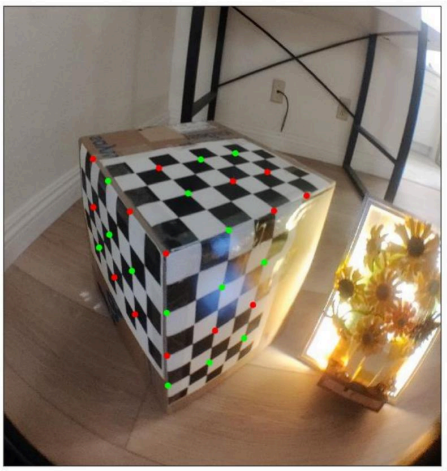


Fig. 3. Image of checkboard taken with clip-on fisheye lens on the phone camera. Red dots: used for calibration. Green dots: used for evaluation (loss function calculation).

IV. Results

The calibration and test results are summarized in Table.1 and Fig.4. The Pinhole camera model, which can be represented by a linear 3-by-4 matrix, shows a significant error (loss=23.46) when predicting the test data points. The red dots in Fig. 4(a) show the prediction from the pinhole camera model, and the green dots show the ground truth. Although most of the predictions are aligned with the ground truth, some predictions are significantly deviated, especially at the edge of the image, e.g., red dot in the bottom of Fig. 4(a), where the distortion is strong.

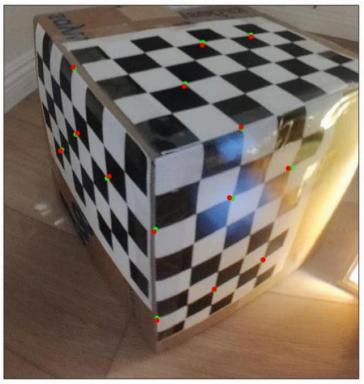
Camera Model	Loss
Pinhole camera model	23.46
UCM	10.67
EUCM	10.29
KB camera model (6 parameters)	10.73
FOV camera model	10.77
DS camera model	10.66

Table 1. Calibration results from different camera models. Loss: averaged distance between model prediction and ground truth.

UCM model add an extra parameter α to help correct the distortion. The loss of UCM model decreases to 10.67 with optimized $\alpha = 0.7$. Those dots that previously deviated in Fig. 4(a) are also much better overlapped in Fig. 4(b), including the red dot at the bottom of the image. This indicates that a nonlinear camera model can significantly improve calibration error for the fisheye camera. However, EUCM model has very marginal improvement from UCM model with optimized $\beta \approx 0$. Other nonlinear models, i.e., KB camera model, FOV camera model, DS camera model, also has limited improvements. We suspect that: (1) experimental errors might dominate in the nonlinear model calibration, e.g., homemade checkerboard not flat and inaccurate, point labeling error, etc. Improvement can be made by using a sharper and better labeled checkerboard. (2) Most of the calibration and test data points are still located in the middle area of the image, where the distortion is not strong enough. The advantage of nonlinear models might help further if these data points are located on the edge of the image.

(a) Pinhole model

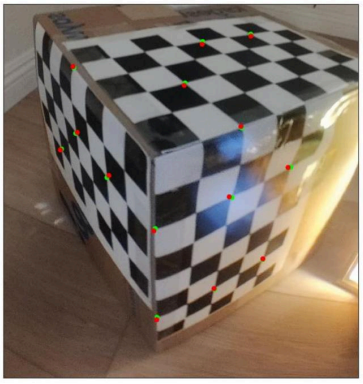
(b) UCM



(c) EUCM

(d) KB model

(e) FOV model



(f) DS model

Fig. 4. Calibration results from different camera models. (a) pinhole model, (b) UCM, (c) EUCM, (d) KB model, (e) FOV model, (f) DS model. Red dots: camera model predictions. Green dots: ground truth from measurements.

V. Conclusion

In the paper, we summarized several camera models for wide-FOV cameras from previous literatures. We performed a camera calibration experiment with a clip-on fisheye lens and a homemade checkerboard. Calibration results showed that significant improvement can be made from all tested nonlinear camera models, while the loss difference between these nonlinear camera models is marginal. We suspect the remaining error is dominant by the experimental measurement error, and nonlinear camera models might help further if sampling points are more located on the edge of image where distortion is stronger.

Acknowledgment

I hope to acknowledge Prof. Bohg and Prof. Savarese for great lectures on CS231A: Computer Vision, from 2D Perception to 2D reconstruction and beyond, which introduces me to many interesting topics. Special thanks to the teaching assistants, Krishnan, Congyue, and Tianyuan for supporting our

courses and projects, answering questions with patience, and sharing knowledge.

VI. Bibliography

1. Vladyslav Usenko et. al., " The Double Sphere Camera Model," arXiv:1807.08957v2 (2018).
2. X. Ying et al. “Can we consider central catadioptric cameras and fisheye cameras within a unified imaging model.” Computer Vision - ECCV pp.442–455 (2004)
3. B. Khomutenko, G. Garcia, and P. Martinet. “An enhanced unified camera model.” IEEE Robotics and Automation Letters, 1(1):137–144, (2016)
4. J. Kannala and S. S. Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. IEEE transactions on pattern analysis and machine intelligence, 28(8):1335–1340, (2006).
5. F. Devernay and O. Faugeras. Straight lines have to be straight. Machine vision and applications, 13(1):14–24, (2001)
6. Eunjin Son et. al., "Monpcular Dpeth Estimation from a Fisheye Camera Based on Knowledge Distillation", MDPI Sensors, 23(24), 9866 (2023)
7. Github, "Fisheye-Depth-Estimation", online source: https://github.com/shun74/Fisheye-Depth-Estimation