A COMPREHENSIVE OVERVIEW OF FISH-EYE CAMERA DISTORTION CORRECTION METHODS

Jian Xu
SUSTech
Southern University of Science and Technology
ShenZhen, GuangDong
hwllo772@gmail.com

De-Wei Han
SUSTech
Southern University of Science and Technology
ShenZhen, GuangDong
musthan@126.com

Kang Li
SUSTech
Southern University of Science and Technology
ShenZhen, GuangDong
likangqd@163.com

Jun-Jie Li
SUSTech
Southern University of Science and Technology
ShenZhen, GuangDong
lijj3@mail.sustech.edu.cn

Zhao-Yuan Ma
SUSTech
Southern University of Science and Technology
ShenZhen, GuangDong
mazy@sustech.edu.cn

ABSTRACT

The fisheye camera, with its unique wide field of view and other characteristics, has found extensive applications in various fields[1, 2]. However, the fisheye camera suffers from significant distortion compared to pinhole cameras, resulting in distorted images of captured objects. Fish-eye camera distortion is a common issue in digital image processing, requiring effective correction techniques to enhance image quality. This review provides a comprehensive overview of various methods used for fish-eye camera distortion correction[3]. The article explores the polynomial distortion model, which utilizes polynomial functions to model and correct radial distortions. Additionally, alternative approaches such as panorama mapping, grid mapping, direct methods, and deep learning-based methods are discussed. The review highlights the advantages, limitations, and recent advancements of each method, enabling readers to make informed decisions based on their specific needs.

***K*eywords** Fish-eye Camera · Distortion · Correction · Deep Learning · Panorama Mapping

1 Introduction

Fish-eye lenses have gained popularity in various fields, including photography[4], computer vision[5], robotics[6], and virtual reality[7], due to their wide field of view and unique visual effects. However, these lenses often introduce significant distortion to the captured images, which can distort the shapes of objects and degrade image quality. To overcome this challenge, fish-eye camera distortion correction methods have been developed to rectify the images and restore their original appearance.

The correction of fish-eye camera distortion is a crucial task in digital image processing. It involves the application of mathematical models and algorithms to compensate for the non-linear distortions introduced by fish-eye lenses. Correcting the distortion can improve the accuracy of measurements, facilitate accurate object recognition, and enhance overall image quality for various applications.

This review aims to provide a comprehensive overview of the different methods employed to correct fish-eye camera distortion. The review will cover both traditional and more recent approaches, discussing their underlying principles, advantages, limitations, and potential applications. By understanding the various methods available, researchers, professionals, and enthusiasts in the field can make informed decisions about the most suitable technique for their specific needs.

The following sections will delve into the polynomial distortion model, which is widely used for fish-eye camera distortion correction. Additionally, alternative methods such as panorama mapping, grid mapping, direct methods, and deep learning-based approaches will be explored. Each method will be examined in detail, highlighting their strengths and weaknesses and providing insights into their practical implementation.

Overall, this review aims to serve as a valuable resource for individuals interested in fish-eye camera distortion correction. By presenting a comprehensive overview of the available methods, this review aims to facilitate a deeper understanding of the techniques involved and foster further advancements in the field of digital image processing.

2 Camera Projection Models

The imaging process of a fisheye camera is commonly approximated as a unit sphere projection model. The imaging process of a fisheye camera can be decomposed into two steps: first, linearly projecting the 3D points in space onto a virtual unit sphere; and then projecting the points on the unit sphere onto the image plane, which is a nonlinear process. In the context of fisheye cameras, four common projection models are widely used: Equidistant Projection Model[8], Equiangular Projection Model[8], Orthographic Projection Model[9] and Stereographic Projection Model[10].

2.1 Equidistant Projection Model

The Equidistant Projection Model assumes that the rays of light passing through the lens and projecting onto the image sensor form equal angles with the optical axis. In this projection model, the mapping between 3D points (X, Y, Z) in the camera coordinate system and 2D image coordinates (u, v) can be expressed as follows:

$$\theta = \arctan(Y, X) \tag{1}$$

$$\varphi = \arctan(\sqrt{X^2 + Y^2}, Z) \tag{2}$$

$$r_d = f * \varphi \tag{3}$$

$$u = r_d * \cos(\theta) \tag{4}$$

$$v = r_d * \sin(\theta) \tag{5}$$

Here, (θ, ϕ) represents the spherical coordinates on the unit sphere, r is the radial distance from the optical center, (u, v) represents the normalized image coordinates, and f is the focal length of the fisheye lens.

2.2 Equiangular Projection Model

The Equiangular Projection Model is commonly used for capturing panoramic or 360-degree images with fisheye lenses. It involves mapping the 3D points on a unit sphere to 2D image coordinates using an equiangular grid. In this projection model, the mapping between 3D points (X, Y, Z) on the unit sphere and 2D image coordinates (u, v) can be expressed as follows:

$$\theta = \arctan(Y, X) \tag{6}$$

$$\varphi = \arctan(\sqrt{X^2 + Y^2}, Z) \tag{7}$$

$$u = \theta + \pi/(2\pi) \tag{8}$$

$$v = (\theta + \pi/2)/\pi \tag{9}$$

Here, (θ, ϕ) represents the spherical coordinates on the unit sphere, and (u, v) represents the normalized image coordinates, ranging from 0 to 1.

2.3 Orthographic Projection Model

The Orthographic Projection Model is a camera projection model that assumes the rays of light from the scene are parallel and perpendicular to the image plane. In this model, the 3D points are directly projected onto a 2D image without any perspective distortion. The mapping between the 3D points (X, Y, Z) in the camera coordinate system and the 2D image coordinates (u, v) can be expressed as follows:

$$u = X/scale_x + center_x \tag{10}$$

$$v = Y/scale_y + center_y \tag{11}$$

Here, (u, v) represents the image coordinates, (X, Y) represents the 3D points in the camera coordinate system.

2.4 Stereographic Projection Model

The characteristic of the Stereographic Projection Model is that it preserves angles, which is a desirable property in mathematics known as conformality. Preserving angles means that the angles formed by any intersecting lines remain unchanged after the transformation, even though the lines themselves may become curved. Under a conformal transformation, a circle still remains a circle (where a straight line can be considered a circle with an infinite diameter). Therefore, to some extent, a conformal transformation also preserves the "shape" of objects. In the simulated scenario below, all boundary lines on the surface of the cylinder are transformed into circular arcs, and all angles formed by intersecting lines remain unchanged at 90.

$$r_d = 2f * \tan(2 * \theta) \tag{12}$$

3 Distortion Correction Methods

Camera distortion is the alteration of an image's perspective caused by the camera's lens, sensor, or other factors. There are several types of distortion, including: Radial distortion, Tangential distortion, as well as Non-linear distortion. The purpose of camera distortion correction is to transform the distorted image captured by the camera into an image that resembles the ideal image produced by a pinhole camera. This correction aims to improve the accuracy of the image, enhance its visual quality, and meet the specific requirements of various applications.

The application of fisheye cameras in computer vision often requires advanced distortion correction methods to ensure accurate and reliable image analysis. Fisheye lenses introduce significant distortions that can impact the accuracy of measurements, object recognition, and scene understanding. In this section, we discuss various state-of-the-art methods for fisheye camera distortion correction, aiming to transform the distorted fisheye images into rectified images resembling those captured by ideal pinhole cameras.

3.1 Distortion Types

3.1.1 Symmetric Radial Distortions

Symmetric radial distortions are what are typically imagined when discussing image distortion. Often, this type of distortion will be characterized depending on if it is positive (pincushion) or negative (barrel) distortion.

3.1.2 Asymmetric Radial Distortions

Asymmetric radial distortions are radial distortion effects much like the above, but unlike symmetric radial distortion, asymmetric radial distortion characterizes distortion effects that are dependent both on the distance from the image centre as well as how far away the object being imaged is. Asymmetric radial effects are most pronounced in two scenarios:

Cameras with long focal lengths and very-short relative object distances. e.g. a very-near-field telephoto lens that is capturing many objects very close. Observing objects through a medium of high-refraction, or differing refractive indices. e.g. two objects underwater where one is near and one is far away. This type of distortion is typically tricky to visualize, as well as to quantify, because it is dependent on the environment. In most robotic and automated vehicle contexts, asymmetric radial distortion is not a great concern! Why? Well, the difference in distortions depends on the difference in distances between objects. This is usually because of some kind of refractive difference between two objects being imaged, or because the objects are out of focus of the camera (i.e. focal length is too large relative to the object distance).

Neither of the above two scenarios are typical; as such, asymmetric radial distortion is an important aspect of modeling the calibration in applications when these scenarios are encountered.

In most robotic contexts, the primary use for imaging and visual-odometry is done in relatively short ranges with cameras that have short focal lengths, and the primary medium for light to travel through is air. Since there doesn't tend to be big atmospheric variances between objects that are close, and since light is all traveling through the same medium, there isn't much of an asymmetric refractive effect to characterize or measure. As a result, this kind of radial distortion isn't common when calibrating cameras for these kinds of applications. If we can't measure it, we shouldn't try to model it!

3.1.3 Tangential (De-centering) Distortions

Tangential distortion is sometimes also called de-centering distortion, because the primary cause is due to the lens assembly not being centered over and parallel to the image plane. The geometric effect from tangential distortion is not purely along the radial axis. Instead, as can be seen in the figure above, it can perform a rotation and skew of the image plane that depends on the radius from the image centre!

3.2 Polynomial Distortion Model

The polynomial distortion model is one of the most commonly used methods for fish-eye camera distortion correction. It relies on a mathematical model that describes the radial distortion present in fish-eye images. Typically, this model uses polynomial functions to approximate the distortions and correct them. The correction process involves converting the pixel coordinates of the image to normalized coordinates and applying the polynomial functions to rectify the distortions. This method is widely adopted due to its simplicity and effectiveness in addressing radial distortions in fish-eye images. While there are more models than what is described here, the industry has largely standardized on the following two distortion models.

3.2.1 Brown-Conrady

Brown-Conrady distortion[11] is probably what most think of as the "standard" radial and tangential distortion model. It was first published in 1966 by Brown and Conrady, and has been used in many applications since.

This model characterizes radial distortion as a series of higher order polynomials:

$$r = \sqrt{x^2 + y^2} \tag{13}$$

$$\delta r = k_1 r^3 + k_2 r^5 + k_3 r^7 + ... + k_n r^{n+2} \tag{14}$$

In practice, only the k_1 through k_3 terms are typically used. For cameras with relatively simple lens assemblies (e.g. only contain one or two lenses in front of the CMOS/CCD sensor), it is often sufficient to just use the k_1 and k_2 terms.

To relate this back to our image coordinate system (i.e. x and y), we usually need to do some basic trigonometry:

$$\delta x_r = \sin(\varphi)\delta_r = x/r(k_1 e^3 + k_2 r^5 + k_3 r^7) \tag{15}$$

$$\delta y_r = \cos(\varphi)\delta_r = y/r(k_1 e^3 + k_2 r^5 + k_3 r^7) \tag{16}$$

Angential distortion, as characterized by the Brown-Conrady model, is often simplified into the following x and y components. We present these here first as they are probably what most are familiar with:

$$\delta x_t = 2p_2 xy + p_1(r^2 + 2x^2) \tag{17}$$

$$\delta y_r = p_2(r^2 + 2y^2) + 2p_1 xy \tag{18}$$

This actually derives from an even-power series much like the radial distortion is an odd-power series. The full formulation is a solution to the following:

$$\delta t = P(r)\cos(\varphi - \varphi_0) \tag{19}$$

Where p(r) is our de-centering distortion profile function, is the polar angle of the image plane coordinate, and is the angle to the axis of maximum tangential distortion (i.e. zero radial distortion). Expanding this into the general parameter set we use today is quite involved (read the original Brown paper!), however this will typically take the form:

$$\delta x_t = [p_1(r^2 + 2x^2) + 2p_2 xy](1 + p_3 r^2 + p_4 r^4 + p_5 r^6 + ...) \tag{20}$$

$$\delta y_r = [2p_1 xy + p_2(r^2 + 2y^2)](1 + p_3 r^2 + p_4 r^4 + p_5 r^6 + ...) \tag{21}$$

Because tangential distortion is usually small, we tend to approximate it using only the first two terms. It is rare for de-centering to be so extreme that our tangential distortion requires higher order terms because that would mean that our lens is greatly de-centered relative to our image plane. In most cases, one might ask if their lens should simply be re-attached in a more appropriate manner.

3.2.2 Kannala-Brandt

Almost a century later (2006, from the original Conrady paper in 1919), Juho Kannala and Sami Brandt[12] published their own paper on lens distortions. The main contribution of this paper adapts lens distortion modeling to be optimized for wide-angle, ultra wide-angle, and fish-eye lenses. Brown Conrady's modeling was largely founded on the physics of Seidel aberrations, which were first formulated around 1867 for standard lens physics of the time, which did not include ultra wide and fish-eye lenses.

The primary difference that most folks will notice using this model lies in symmetric radial distortion. Rather than characterizing radial distortion in terms of how far a point is from the image centre (the radius), Kannala-Brandt characterizes distortion as a function of the incidence angle of the light passing through the lens. This is done because the distortion function is smoother when parameterized with respect to this angle ( ), which makes it easier to model as a power-series:

$$\theta = \arctan(r, f) \tag{22}$$

$$\delta r = k_1\theta^2 + k_2\theta^4 + k_3\theta^6 + ... \tag{23}$$

Above, we've shown the formula for θ when using perspective projection, but the main advantage of the Kannala-Brandt model is that it can support different kinds of projection by swapping our formula for θ, which is what makes the distortion function smoother for wide-angle lenses. Kannala-Brandt also aims to characterize other radial (such as asymmetric) and tangential distortions. This is done with the following additional parameter sets:

$$\delta r_other = (l_1\theta + l_2\theta^3 + l_3\theta^5)(i_1 \cos\phi + i_2 \cos(2\phi) + i_4 \sin(2\phi) + ...) \tag{24}$$

$$\delta t = (m_1\theta + m_2\theta^3 + m_3\theta^5)(j_1 \cos\phi + j_2 \cos(2\phi) + j_4 \sin(2\phi) + ...) \tag{25}$$

Overall, this results in a 23 parameter model! This is admittedly overkill, and the original paper claims as much. These models, unlike the symmetric radial distortion, are an empirical model derived by fitting an N-term Fourier series to the data being calibrated. This is one way of characterizing it, but over-parameterizing our final model can lead to poor repeatability of our final estimated parameters. In practice, most systems will characterize Kannala-Brandt distortions purely in terms of the symmetric radial distortion, as that distortion is significantly larger in magnitude and will be the leading kind of distortion in wider-angle lenses.

3.3 Feature-based Methods

In the context of fisheye camera distortion correction, feature-based methods play a vital role by leveraging the characteristics of the fisheye image to infer and rectify the camera's distortion parameters. This section presents an overview of several feature-based methods commonly used for fisheye camera distortion correction.

3.3.1 Corner Detection and Rectification

Corner detection and rectification methods involve the detection of corners in the fisheye image, such as employing Harris corner detection[13] or Shi-Tomasi corner detection[14]. Subsequently, distortion rectification is performed by utilizing the relationships between the detected corners in terms of distances and angles to estimate the camera's distortion parameters. By employing these parameters, the entire image can be rectified to mitigate the fisheye distortion.

3.3.2 Feature Point Matching and Rectification

Feature point matching and rectification methods rely on the extraction and matching of feature points in the fisheye image. Popular techniques, such as SIFT feature points[15] or SURF feature points[16], are employed for feature point extraction. The detected feature points are then matched with their corresponding points in the rectified image, enabling the inference of the camera's distortion parameters and subsequent rectification.

3.3.3 Line Detection and Rectification

Line detection and rectification methods capitalize on the presence of straight lines in the fisheye image. By detecting lines using techniques like the Hough transform[17], the camera's distortion parameters can be inferred. The obtained parameters are subsequently utilized to rectify the entire image, ensuring that the straight lines preserve their linearity in the rectified image.

3.3.4 Optical Flow-based Methods

Optical flow-based methods exploit the information provided by the optical flow in the fisheye image for distortion correction. By calculating the pixel displacements in the fisheye image, the camera's distortion parameters can be inferred and used for rectification. Optical flow algorithms, such as the Lucas-Kanade method[18] or deep learning-based optical flow estimation, can be employed to estimate the pixel displacements.

3.4 Direct Methods

Direct methods for fish-eye camera distortion correction involve the detection and analysis of specific features or patterns in the image to estimate the distortion parameters. These methods typically rely on the relationships between the distorted image coordinates and the undistorted object coordinates. Techniques such as RANSAC (Random Sample Consensus) can be used to robustly estimate the distortion parameters from the feature correspondences. Direct methods are advantageous in scenarios where calibration data or a priori knowledge about the distortion model is not available.

3.4.1 Horizontal Expansion Method

The Horizontal Expansion Method is a technique used for fisheye image rectification and distortion correction. It aims to transform a distorted fisheye image into a rectilinear image, which has straight lines and a more natural perspective. The method involves expanding the horizontal field of view of the fisheye image and mapping the distorted pixels to their corresponding locations in the rectilinear image.

3.4.2 Latitude-Longitude Mapping Method

The Latitude-Longitude Mapping Method is a technique used for transforming a fisheye image into a panoramic or equirectangular projection. This method involves mapping the distorted fisheye pixels to their corresponding latitude and longitude coordinates on a spherical or cylindrical coordinate system, and then projecting them onto a 2D equirectangular grid.

3.4.3 Panorama Mapping Method

The Panorama Mapping Method is a technique used for fisheye image rectification and distortion correction. It aims to transform a distorted fisheye image into a rectilinear image, which has straight lines and a more natural perspective. The method involves mapping the distorted fisheye pixels to their corresponding locations in the rectilinear image.

3.5 Deep Learning-Based Methods

With the recent advancements in deep learning, neural network-based approaches have emerged for fish-eye camera distortion correction. These methods involve training a neural network to learn the mapping function between distorted and undistorted images. A large dataset of paired images with known distortions is used for training the network. Once trained, the network can perform distortion correction on new input images. Deep learning-based methods offer the advantage of learning complex distortion patterns and can handle a wide range of distortions effectively. However, they require a substantial amount of training data and computational resources. This section provides an overview of deep learning methods commonly used for fisheye camera distortion correction.

3.5.1 Converlutional Meural Metworks(CNNs)

Convolutional Neural Networks (CNNs) have been extensively employed for fisheye camera distortion correction. These networks consist of multiple convolutional layers that extract hierarchical features from the input images. By training CNNs on a large dataset of distorted and undistorted fisheye image pairs, they can learn the underlying patterns and relationships to predict the distortion-free version of a given fisheye image. (Rong et. al., 2016)[19] intends to employ CNNs (Convolutional Neural Networks), to achieve radial distortion correction. Inspired by the growing availability of image dataset with non-radial distortion (Rong et. al., 2016) propose a framework to address the issue by synthesizing images with radial distortion for CNNs. To this end (Lu et. al., 2017)[20] propose a fast level set model-based method for intensity inhomogeneity correction and a spectral properties-based color correction method to overcome these obstacles. In contrast to conventional approaches, the proposed model integrates a new signed energy force function that can detect contours at weak or blurred edges efficiently. (Borkar et. al., 2019)[21] evaluate the effect of image distortions like Gaussian blur and additive noise on the activations of pre-trained convolutional filters. (Borkar et. al., 2019) propose a metric to identify the most noise susceptible convolutional filters and rank them in order of the highest gain in classification accuracy upon correction. Radial lens distortion often exists in images taken by commercial cameras, which does not satisfy the assumption of pinhole camera model. They generated images with a large number of images of high variation of radial distortion, which can be well exploited by deep CNN with a high learning capacity, and reach the state-of-the-art results. (Shi et. al., 2018)[22] claim that a weight layer with inverted foveal models can be added to these existing CNNs methods for radial distortion correction. Convolutional neural networks (CNNs) have been widely used for road scene understanding in the last few years with great success. (Arsenali et. al., 2019)[23] propose RotInvMTL: a multi-task network (MTL) to perform joint semantic segmentation, boundary prediction, and object detection directly on raw fisheye images. An attempt to optimize a CNN-based detector for fisheye cameras was made, taking into consideration the barrel distortion, which complicates the object detection (Goodarzi et. al., 2019)[24]. The obtained result proves that fisheye augmentation can considerably advance a CNN-based detector's performance on fisheye images in spite of the distortion. (Sáez et. al., 2019)[25] present a methodology that provides real-time semantic segmentation on fisheye cameras leveraging only synthetic images. (Sáez et. al., 2019) propose some Convolutional Neural Networks(CNN) architectures based on Efficient Residual Factorized Network(ERFNet) that demonstrate notable skills handling distortion and a new training strategy that improves the segmentation on the image borders. (Vasiljevic et. al., 2020)[26] show that self-supervision can be used to learn accurate depth and ego-motion estimation without prior knowledge of the camera model. Inspired by the geometric model of Grossberg and Nayar, (Vasiljevic et. al., 2020) introduce Neural Ray Surfaces (NRS), convolutional networks that represent pixel-wise projection rays, approximating a wide range of cameras. The strong radial distortion breaks the translation invariance inductive bias of Convolutional Neural Networks. (Ramachandran et. al., 2022)[27] provide a detailed analysis on the competition which attracted the participation of 120 global teams and a total of 1492 submissions. The fisheye image has a severe geometric distortion which may interfere with the stage of image registration and stitching. In the stage of fisheye image correction (Hao et. al., 2023)[28] propose the Attention-based Nonlinear Activation Free Network (ANAFNet) to deblur fisheye images corrected by Zhang calibration method.

3.5.2 Generative Adversarial Networks(GANs)

Generative Adversarial Networks (GANs) have also been utilized for fisheye camera distortion correction. GANs consist of a generator network and a discriminator network, which are trained simultaneously in an adversarial manner. The generator network generates undistorted fisheye images, while the discriminator network aims to distinguish between the generated undistorted images and the real undistorted images. Through this adversarial training process, GANs can learn to generate high-quality undistorted fisheye images.

(Li et. al., 2012)[29] present a novel embedded real-time fisheye image distortion correction algorithm with application in IP network camera. A fast and simple distortion correction method is introduced based on Midpoint Circle Algorithm (MCA) which aims to determine the pixel positions along a circle circumference based on incremental calculation of decision parameters. Each lens is calibrated separately and interior/relative orientation parameters (IOPs and ROPs) of the camera are determined on the basis of designed calibration network on the central and side images captured by the aforementioned lenses (Aghayari et. al., 2017)[30]. Designed calibration network is considered as a free distortion grid and applied to the measured control points in the image space as correction terms by means of bilinear interpolation. Deep learning techniques have become popular for performing camera model identification. To expose weaknesses in these methods (Chen et. al., 2018)[31] propose a new anti-forensic framework that utilizes a generative adversarial network (GAN) to falsify an image's source camera model. (Nikonorov et. al., 2019)[32] present a new end-to-end framework applying two convolutional neural networks (CNNs) to reconstruct images captured with multilevel diffractive lenses (MDLs). The generative adversarial network (GAN) is first used to remove image-wise color distortion, while a patch-wise network is then used to apply chromatic deblur. Correction of the distortion of images is crucial in many computer vision applications. (Liao et. al., 2020)[33] present distortion rectification generative adversarial network (DR-GAN), a conditional generative adversarial network (GAN) for automatic radial DR. To the best of the knowledge, this is the first end-to-end trainable adversarial framework for radial distortion rectification. (Gallego et. al., 2020)[34] present a network architecture with parallel convolutional neural networks (CNN) for removing perspective distortion in images. While other works generate corrected images through the use of generative adversarial networks or encoder-decoder networks, (Gallego et. al., 2020) propose a method wherein three CNNs are trained in parallel, to predict a certain element pair in the 3*3 transformation matrix. Generative adversarial networks (GANs) have been implemented to convert IR images into RGB images for enriching semantic information. (Zhang et. al., 2021)[35] study wggan: a wavelet-guided generative adversarial network for thermal image translation. A wavelet-guided generative adversarial network (WGGAN) is proposed to address the problem. In order to improve the quality of low-light image (Zhang et. al., 2021) propose a Heterogenous low-light image enhancement method based on DenseNet generative adversarial network. Secondly, the feature map from low light image to normal light image is learned by using the generative adversarial network. (Thapa et. al., 2021)[36] present the distortion-guided network (DG-Net) for restoring distortion-free underwater images. (Thapa et. al., 2021) then use a generative adversarial network guided by the distortion map to restore the sharp distortion-free image. (Luo et. al., 2021)[37] propose an unsupervised deep convolutional network that takes rectified stereo image pairs as input and outputs corresponding dense disparity maps. Second, the left and right images, which are reconstructed based on the input image pair and corresponding disparities as well as the vertical correction maps, are regarded as the outputs of the generative term of the generative adversarial network (GAN).

3.5.3 Encoder-Decoder Architectures

Encoder-decoder architectures, such as U-Net and its variants, have proven effective for fisheye camera distortion correction. These architectures consist of an encoder network that captures the high-level features of the distorted fisheye image and a decoder network that reconstructs the undistorted image from the encoded features. The encoder-decoder structure allows for the preservation of spatial information during the distortion correction process.

4 Conclusion

Fish-eye camera distortion correction is a critical task in digital image processing, aimed at rectifying the distortions introduced by fish-eye lenses and improving image quality. In this review, we provided a comprehensive overview of various methods used for fish-eye camera distortion correction.

We discussed the polynomial distortion model, which utilizes polynomial functions to model and correct radial distortions. This method is widely adopted due to its simplicity and effectiveness. Additionally, alternative approaches such as panorama mapping, grid mapping, direct methods, and deep learning-based methods were explored. Each method has its strengths and limitations, and their suitability depends on specific requirements and constraints.

Through this review, researchers, professionals, and enthusiasts in the field of digital image processing gained a deeper understanding of the available techniques for fish-eye camera distortion correction. The review highlighted the underlying principles, advantages, limitations, and potential applications of each method, enabling informed decision-making.

To evaluate the performance of distortion correction methods, various experiments can be conducted, including synthetic data evaluation, calibration image evaluation, comparative studies, real-time performance evaluation, and application-specific evaluations. These experiments provide insights into the accuracy, computational efficiency, and applicability of the methods in different scenarios.

In conclusion, fish-eye camera distortion correction methods play a crucial role in enhancing image quality and enabling accurate analysis in various fields. By understanding the different techniques and conducting appropriate experiments, researchers can select the most suitable method for their specific needs, contributing to advancements in the field of digital image processing.

References

[1] David A. Ross, Jongwoo Lim, Ruei-Sung Lin, and Ming-Hsuan Yang. Incremental learning for robust visual tracking. *International Journal of Computer Vision*, 77(1–3):125–141, 2008.

[2] Eric Royer, Maxime Lhuillier, Michel Dhome, and Jean-Marc Lavest. Monocular vision for mobile robot localization and autonomous navigation. *International Journal of Computer Vision*, 74(3):237–260, 2007.

[3] C. Hughes, M. Glavin, E. Jones, and P. Denny. Review of geometric distortion compensation in fish-eye cameras. In *Proceedings of IET Irish Signals and Systems Conference*, pages 162–167, 2008.

[4] Mahdi Abbaspour Tehrani, Aditi Majumder, and M. Gopi. Correcting perceived perspective distortions using object specific planar transformations. In *2016 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10, May 2016.

[5] L F Posada, K K Narayanan, F Hoffmann, and T Bertram. Floor segmentation of omnidirectional images for mobile robot visual navigation. In *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, pages 804–809, 2010.

[6] Ivan Markovic, Francois Chaumette, and Ivan Petrovic. Moving object detection, tracking and following using an omnidirectional camera on a mobile robot. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 5630–5635, 2014.

[7] Yalin Xiong and Kenneth Turkowski. Creating image-based vr using a self-calibrating fisheye lens. In *Proceedings of IEEE computer society conference on computer vision and pattern recognition*, pages 237–243. IEEE, 1997.

[8] Rudolf Kingslake. *A history of the photographic lens*. Elsevier, 1989.

[9] Kenro Miyamoto. Fish Eye Lens. *Journal of the Optical Society of America*, 54(8):1060, 1964.

[10] D.E. Stevenson and M.M. Fleck. Nonparametric correction of distortion. In *Proceedings of IEEE Workshop on Applications of Computer Vision*, pages 214–219, 1996.

[11] C Brown Duane. Close-range camera calibration. *Photogramm. Eng*, 37(8):855–866, 1971.

[12] 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.

[13] Chris Harris, Mike Stephens, et al. A combined corner and edge detector. In *Alvey vision conference*, volume 15, pages 10–5244. Citeseer, 1988.

[14] Jianbo Shi et al. Good features to track. In *1994 Proceedings of IEEE conference on computer vision and pattern recognition*, pages 593–600. IEEE, 1994.

[15] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60:91–110, 2004.

[16] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9*, pages 404–417. Springer, 2006.

[17] Richard O Duda and Peter E Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972.

[18] Bruce D Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI'81: 7th international joint conference on Artificial intelligence*, volume 2, pages 674–679, 1981.

[19] Jiangpeng Rong, Shiyao Huang, Zeyu Shang, and Xianghua Ying. Radial lens distortion correction using convolutional neural networks trained with synthesized images. In *Computer Vision–ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part III 13*, pages 35–49. Springer, 2017.

[20] Huimin Lu, Bin Li, Junwu Zhu, Yujie Li, Yun Li, Xing Xu, Li He, Xin Li, Jianru Li, and Seiichi Serikawa. Wound intensity correction and segmentation with convolutional neural networks. *Concurrency and computation: practice and experience*, 29(6):e3927, 2017.

[21] Tejas S Borkar and Lina J Karam. Deepcorrect: Correcting dnn models against image distortions. *IEEE Transactions on Image Processing*, 28(12):6022–6034, 2019.

[22] Yongjie Shi, Danfeng Zhang, Jingsi Wen, Xin Tong, Xianghua Ying, and Hongbin Zha. Radial lens distortion correction by adding a weight layer with inverted foveal models to convolutional neural networks. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 1–6. IEEE, 2018.

[23] Bruno Arsenali, Prashanth Viswanath, and Jelena Novosel. Rotinvmtl: Rotation invariant multinet on fisheye images for autonomous driving applications. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pages 0–0, 2019.

[24] Payam Goodarzi, Martin Stellmacher, Michael Paetzold, Ahmed Hussein, and Elmar Matthes. Optimization of a cnn-based object detector for fisheye cameras. In *2019 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, pages 1–7. IEEE, 2019.

[25] Álvaro Sáez, Luis M Bergasa, Elena López-Guillén, Eduardo Romera, Miguel Tradacete, Carlos Gómez-Huélamo, and Javier Del Egido. Real-time semantic segmentation for fisheye urban driving images based on erfnet. *Sensors*, 19(3):503, 2019.

[26] Igor Vasiljevic, Vitor Guizilini, Rares Ambrus, Sudeep Pillai, Wolfram Burgard, Greg Shakhnarovich, and Adrien Gaidon. Neural ray surfaces for self-supervised learning of depth and ego-motion. In *2020 International Conference on 3D Vision (3DV)*, pages 1–11. IEEE, 2020.

[27] Saravanabalagi Ramachandran, Ganesh Sistu, Varun Ravi Kumar, John McDonald, and Senthil Yogamani. Woodscape fisheye object detection for autonomous driving–cvpr 2022 omnicv workshop challenge. *arXiv preprint arXiv:2206.12912*, 2022.

[28] Jing Hao, Jingming Xie, Jinyuan Zhang, and Moyun Liu. A stronger stitching algorithm for fisheye images based on deblurring and registration. *IEEE Sensors Letters*, 2023.

[29] Yongle Li, Maojun Zhang, Yu Liu, and Zhihui Xiong. Fish-eye distortion correction based on midpoint circle algorithm. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2224–2228. IEEE, 2012.

[30] S Aghayari, Mohammad Saadatseresht, Mohammad Omidalizarandi, and Ingo Neumann. Geometric calibration of full spherical panoramic ricoh-theta camera. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences; IV-1/W1*, 4:237–245, 2017.

[31] Chen Chen, Xinwei Zhao, and Matthew C Stamm. Mislgan: an anti-forensic camera model falsification framework using a generative adversarial network. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 535–539. IEEE, 2018.

[32] Artem Nikonorov, Viktoria Evdokimova, Maksim Petrov, Pavel Yakimov, Sergey Bibikov, Yuriy Yuzifovich, Roman Skidanov, and Nikolay Kazanskiy. Deep learning-based imaging using single-lens and multi-aperture diffractive optical systems. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pages 0–0, 2019.

[33] Kang Liao, Chunyu Lin, Yao Zhao, and Moncef Gabbouj. Dr-gan: Automatic radial distortion rectification using conditional gan in real-time. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(3):725–733, 2019.

[34] Neil Patrick Del Gallego, Joel Ilao, and Macario Cordel. Blind first-order perspective distortion correction using parallel convolutional neural networks. *Sensors*, 20(17):4898, 2020.

[35] Ran Zhang, Junchi Bin, Zheng Liu, and Erik Blasch. Wggan: A wavelet-guided generative adversarial network for thermal image translation. In *Generative Adversarial Networks for Image-to-Image Translation*, pages 313–327. Elsevier, 2021.

[36] Simron Thapa, Nianyi Li, and Jinwei Ye. Learning to remove refractive distortions from underwater images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5007–5016, 2021.

[37] Huoling Luo, Congcong Wang, Xingguang Duan, Hao Liu, Ping Wang, Qingmao Hu, and Fucang Jia. Unsupervised learning of depth estimation from imperfect rectified stereo laparoscopic images. *Computers in biology and medicine*, 140:105109, 2022.