IMAGE DISTORTION CORRECTION FOR BIPRISM-BASED SINGLE-LENS STEREOVISION SYSTEM

QIAN BEIBEI

A THESIS SUBMITTED FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2016

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in this entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Qian Beibei

25 July, 2016

ACKNOWLEDGEMENT

I would like to take this opportunity to express my deep gratitude to my supervisor, Professor Lim Kah Bin, who impressed me with his integral views on research, untiring support, stimulating suggestions and encouragement. Besides, his personalities like having a good sense of humor and being a man of principle also make him my role model. It is a great pleasure for me to pursue my PhD degree under his supervision. For my colleagues: Kee Weiloon, Bai Yading, Wang Daolei, and Zhao Meijun, I am thankful for their discussions and advices on my research. I also want to thank the lab technicians, Mrs Ooi-Toh Chew Hoey and Miss Hamidah Bte Jasman, for their administrative and daily assistance during my PhD program.

Finally yet importantly, I want to thank my family and friends, especially my wife Zhao Mengyu. With your supports and company, my life in Singapore becomes colorful.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D/2-D	Two-dimension
3D/3-D	Three-dimension
FOV	Field of View
BSS	Biprism-based Single-lens Stereovision
GABDC	Geometrical Analysis Based Distortion Correction
FC-MFDC	Fully Constrained and Model-free Distortion Correction
BDMC	Biprism Distortion Modelling and Correction
BD-Model	Biprism Distortion Model
RD-Model	Radial Distortion Model
CCD	Charge-Coupled Devices
FRs	Fundamental Rays
SVD	Singular Value Decomposition

LIST OF SYMBOLS

The optical center of the camera:	O
Effective real camera focal length:	f
The angle of the biprism:	α
Rotation matrix:	R
The distance between the apex of the biprism to the optical center of the camera:	D
The horizontal length of the back plane of the biprism:	T
The distorted point on the image plane w.r.t. the camera coordinates:	P
The undistorted point on the image plane w.r.t. the camera coordinates:	P'
The distorted point on the image plane w.r.t. the image coordinates:	q^d
The undistorted point on the image plane w.r.t. the image coordinates:	q^p
The refractive index of the biprism glass material:	n
The sensor size of the real camera:	I

LIST OF APPENDICES

ABSTRACT

Single-lens stereovision system allows perceiving depth information with only one camera, by employing some extra optical devices or some known cues. This system has several advantages over the conventional two or multi-camera stereovision system, including the low cost, simple setup, compactness and fewer system parameters. Thus, many research works have been done to develop a single-lens stereovision system to replace the conventional two-camera system. In this research, a single-lens stereovision system using a biprism was introduced. The essential part of the system is the biprism in front of the camera, which splits the incoming scene into two sub-scenes along different view angles and generates stereo image pair within one image. However, the extra optical device will induce unexpected distortions in the image. The aim of this thesis is to study and correct the image distortion induced by the biprism. Three different distortion correction methods are proposed in this thesis.

The first method proposes a geometrical based analysis of the distortion formation. With the knowledge of all the system parameters, including the camera parameters, biprism parameters and the geometrical relationship between the camera and the biprism, the projection procedure of the incident lights could be calculated using Snell's law. Hence, the real (distorted) projection could be determined. Meanwhile, assumptions for an ideal (undistorted) projection are also formulated. Therefore, with the correspondence between the distorted and undistorted projections, the distortion correction maps could be estimated. Experiments with synthetic (computer generated) data are first conducted to verify the effects of the system parameters to the distortion formation by varying different system parameters and calculating the amount of distortion in the image. Then, real experimental

images with different levels of distortions are generated and corrected using the corresponding distortion correction maps. The results show that the proposed method is effective in removing the distortions for different system setups.

The second method is a fully constrained and model-free distortion correction (FC-MFDC) method which is able to correct the distortions in the image caused by both biprism and camera lens. Comparing with the first method, FC-MFDC requires no system parameters, but employs all the projective invariants of a planar checkerboard template as the correction constraints, which include straight lines, cross-ratio and convergence at vanishing point, along with the distortion-free reference point (principle point) as an additional constraint from the system. The extracted sample points are corrected by minimizing the total cost function formed by all these constraints. With both sets of distorted and corrected points, and the intermediate points interpolated by a local transformation, the correction maps could be determined. Thereafter, all the subsequent images could be distortion-corrected by the correction maps. This method performs well on the distorted image data captured by the system, and shows improvements in accuracy and robustness comparing with other correction methods.

The third method is to develop a parametric distortion model which could be integrated into the existing camera calibration methods. In this case, the distortion could be calibrated and corrected along with the calibration procedure. With the distortion formation analysis proposed in the first method, two different strategies are applied to build the distortion model: one of them is to employ a general bivariate polynomial function which is only associated with the image coordinates. In this way, the complicated 3D analysis of the projection deviation could be avoided. This general polynomial could be further refined based

on the distortion characteristics. The other way of modelling is to simplify the projection deviation analysis from 3D space to 2D planes but in different directions. Therefore, the distortions could be represented by algebraic functions, which could be further expanded into a polynomial functions using Taylor series. Generally, these two ways of modelling processes should lead to a similar model since they are used to describe the same distortion. This final distortion model is called biprism distortion model (BD-Model) in this thesis. Finally, this model could be integrated into the camera calibration procedure for parametric determination and enhance the calibration and depth recovery accuracy of the virtual cameras.

To summarize, the main contribution of this thesis is the development of biprism distortion correction technique. All the efforts are made to reduce the image distortion in the system and to improve the accuracy and robustness of the stereo camera calibration and the depth recovery using the biprism based single-lens stereovision system. For our future works, some recommendations on improving these distortion correction methods are also provided.

Chapter 1. INTRODUCTION

1.1 STEREOVISION SYSTEM

Stereovision has the ability to extract 3D information from digital images, obtained by a CCD camera. Normally, stereovision system has two or more cameras capturing the same scene from different viewpoints. By detecting and comparing the scene from these images, 3D information could be extracted by computing the relative positions of objects in different image planes. In the last decades, stereovision has been evolving into a multi-disciplinary subject and comprehensive researches have been done on both software and hardware developments for various applications, such as robotics [1, 2], medical devices [3, 4], video games [5, 6], pattern recognition [7, 8], and artificial intelligence [7, 9].

To obtain the 3D information of the scene, several preparatory works need to be done for the stereovision system, such as camera calibration, stereo image pair rectification, and correspondence searching. Typical stereovision system employs two horizontally displaced cameras to obtain two different views of a scene, which is in a manner similar to human binocular vision. When the cameras have been setup properly, the following and unavoidable step is the camera calibration, which refers to the determination of intrinsic and extrinsic parameters of the cameras, which include the optical center, focal length, and relative position and orientation of the two cameras. However, in some cases, such as the use of low cost cameras, lens distortion may exist and will reduce the accuracy of calibration; subsequently it reduces the efficiency and accuracy of correspondence searching. Thus, the calibration always comes along with the distortion correction.

Regarding the stereovision system, instead of using two or multiple cameras, single-lens stereovision system has been studied and developed since 1987 [10]. A lot of research works involving hardware and software development have been done to enhance and expand its applications. As a comparison, single-lens stereovision system has some advantages. It has a more compact setup, lower cost, simpler implementation process, easier camera synchronization since only one camera is used, and also simultaneous image capturing since the stereo image pairs are captured within one single image plane at the same time, etc.

This project presents our investigation on the distortion phenomenon and its correction of prism-based single-lens stereovision system, which was proposed and developed by K. B. Lim [11] who modelled and modified the previous mirror based stereovision system to the current prism-based stereovision system. It has the ability to obtain stereo images, by employing a biprism placed in front of the camera. The incoming scene was equally split into two sub-scenes projecting at different angles on the two halves of the sensor. These two sub-images could be considered as a stereo image pairs. To facilitate the understanding of the system, these image pairs could be considered as being captured by the virtual cameras from two different viewpoints and therefor this system was also called virtual stereovision system.

Several research aspects of our system has been studied, including the virtual camera concept [11], geometrical ray calibration method [11-13], stereo image pair rectification [14, 15], correspondence searching [13, 16, 17], and the occlusion problem [18, 19]. However, one significant but has always been easily ignored issue, the distortion induced by the biprism, attracts a great deal of attention recently. The proposed virtual camera concept and geometrical ray approach in [11] consider that the rays are lying on the 2D plane which is perpendicular to the biprism, and the apex line will split the incoming scene equally. However,

the rays may not be perpendicular to the apex line of the biprism and their projections may result in errors after two refractions through the biprism. The details of the errors caused by the distortion will be further discussed in §3.1. Besides, another type of distortion caused by the misalignment of the biprism may result in additional and more complex distorted image. This would make the task of distortion correction for the system even more challenging.

1.2 OBJECTIVES

This project aims to develop novel methods for distortion correction to improve the quality of the distorted images for the biprism-based single-lens stereovision system. These methods would be able to handle the distortion correction task for different system setups. Three methods have been proposed in this thesis.

The first one is a geometrical analysis based distortion correction (GABDC) method which is able to deduce the distortion formation for the given system setup and the biprism properties. The basic idea of this method is to mathematically express the real (distorted) projection steps of the incident rays from 3D world coordinates to the 2D image coordinates. Based on the ray projection analysis, assumptions to define the ideal rays are also made to establish the ideal (undistorted) projections. Hence, the distortion characteristics correlating with the system parameters could be determined. Therefore, as long as the system parameters, including the biprism properties and camera parameters, are known, the distortion can be determined and corrected. GABDC could also be used to analyze the effects of the system parameters to the formation of the distortion.

The second one is a fully constrained and model-free distortion correction (FC-MFDC) algorithm which is insensitive to system setup and biprism properties. The basis of this

method is that both virtual cameras generated by the biprism can be modelled as pin-hole cameras and hence obey the perspective projection principle. Hence, the principle of perspective projection can be employed as the correction constraints to generate the undistorted projections from the distorted sample points. These constraints include the straightness, cross-ratio, and vanishing point. In addition, an extra constraint of the distortion-free reference point is also applied to improve the robustness of the correction method.

The last one is the biprism distortion modelling and correction (BDMC) method in which a biprism distortion model has been developed according to the distortion characteristics in the system. Two different methodologies are used to build this model: one of them is to establish a bivariate polynomial function which is refined according to the distortion maps generated by GABDC. The other one is to deduce the distortion formation using a geometrical approach. In this way, the distortion maps could be expressed by algebraic functions, which could be further expanded to polynomial functions using Taylor series expansion. At last, this distortion model could be incorporated to the camera calibration procedure (such as Zhang's method [20]), as a replacement of the typical lens distortion models.

Among these methods, it is difficult to judge which one is the best since each of them is suitable for different scenarios. Besides, given the fact that there are no standard criteria to define a good correction, these methods are compared from various aspects based on their advantages and disadvantages, which will be further discussed in the experiments.

A single-lens biprism stereovision system has been setup and several experiments have been carried out to test the effectiveness of the single-lens binocular stereovision systems and to verify the efficiency of the proposed distortion correction methods. Results from the experiments are compared with the conventional approaches to confirm its accuracy and

effectiveness. We believe that most of the work presented in this thesis are novel and practically useful in scientific and industrial areas.

1.3 OUTLINE OF THE THESIS

In this thesis, details and recent progress of our work are presented. The structure of the thesis is organized as follows: Chapter 2 reviews the development history of single-lens stereovision system, different forms of image distortion and the correction algorithms; The proposed geometrical analysis based distortion correction technique is described in Chapter 3. This method also provides a quantitative way to measure the effects of the system parameters to the formation of the distortion. In Chapter 4, a fully constrained and model-free distortion correction method is introduced. This method requires no prior knowledges of the system parameters but employs the perspective projection invariances and the distortion-free reference point. Chapter 5 presents the third correction method by modelling the special distortion in the system. With this particular model, the distortion can be corrected together with the camera calibration procedure. The experiments and results are also presented in Chapter 3, Chapter 4 and Chapter 5. Last but not least, the summary of the current work and the future works are given in Chapter 6.

Chapter 2. LITERATURE REVIEW

In this chapter, a review on the development history of single-lens stereovision system is presented. As a comparison, all the advantages and drawbacks of these systems would be briefly discussed in Section 2.1. Besides, to better understand the biprism-based single-lens stereovision system, it gives the detailed introduction of the formation of the virtual cameras in this section. The subsequent section presents the occurrence of image distortion in the system and different types of distortions and correction methods are studied in Section 2.2 and Section 2.3, respectively. A brief summary on the system and distortion correction methods is presented in Section 2.4.

2.1 SINGLE-LENS STEREOVISION SYSTEM

In most cases, a stereovision system employs two or multiple cameras set up at different viewpoints to capture the same scene. The relative difference in position on image plane leads to the capability of depth recovery. There are outstanding features of this kind of system, but they also face some drawbacks due to the system setup. In the past few decades, researchers have always been seeking for an alternative solution for the conventional stereovision system, by focusing on the investigation of single-lens stereovision system [10, 21-31].

Back in the 1980's, the prototype of single-lens stereovision system was proposed, and has been improved and redefined over the years. To introduce its development history, it is better to categorize the single-lens stereovision techniques into two groups, based on the basic features of the system. One of them is by exploiting some extra optical devices, such as mirrors, glasses, or prisms; the other is by using some known cues either from the system

itself, such as the camera movement, or from its environment, or known illumination conditions.

2.1.1 Single-lens Stereovision System Using Optical Devices

The first prototype of single-lens stereovision system with optical devices was proposed by W. Teoh and X. D. Zhang [21]. As shown in Fig. 2.1, the system employs three mirrors with two of them fixed at 45° to the image plane at the top and bottom, and with one being rotatable in the middle. To generate the stereo image pair, two shots need to be taken with the third mirror aligned to be parallel to the fixed ones, successively.

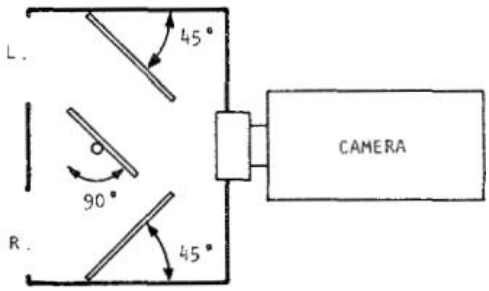


Fig. 2.1 A single-lens stereovision system using three mirrors

Similarly, instead of using multiple mirrors, Nishimoto and Shirai [10] proposed their single-lens stereovision system by placing a rotatable glass plate in front of the camera as shown in Fig. 2.2. By rotating the glass plate, the incident lights would be slightly shifted because of its refraction. Compared with the previous system, this system requires less optical devices, but sacrifices the disparity since the shift by the refraction of the glass plate is relatively small.

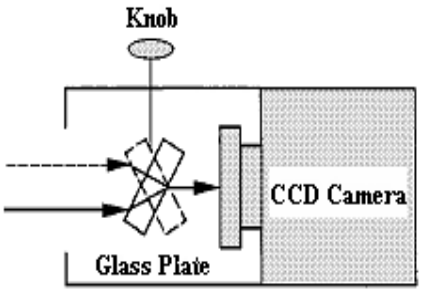


Fig. 2.2 A single-lens stereovision system using a glass plate

As both of the mentioned systems require the camera to take two separate shots to obtain one pair of stereo images, their applications are limited to a static scene or slowly changing environment. To overcome the drawbacks, Gosthasby and Gruver [23] proposed another mirror-based single-lens stereovision system which captured image pairs by the reflectance from the two mirrors as shown in Fig. 2.3. Dissimilar with the system in Fig. 2.1, the images acquired here are inverted and need to be transformed before carrying out the correspondence searching and depth recovery.

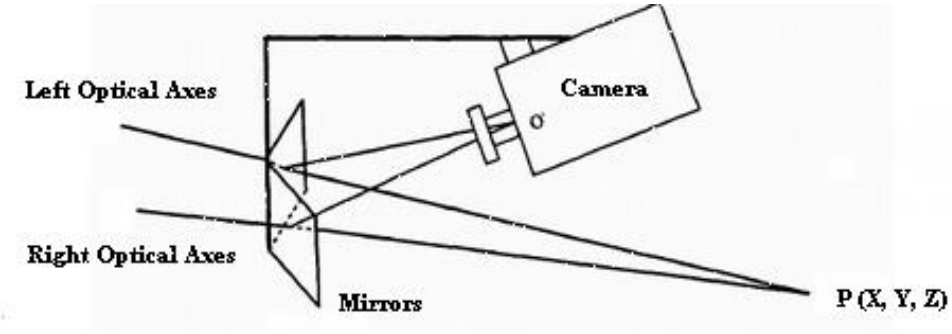


Fig. 2.3 A single-lens stereovision system using two mirrors

Almost at the same time, Inaba *et al.* [29] introduced another mirror-based stereovision system which could control its field of view (FOV) by equipping with a set of moveable mirrors, as shown in Fig. 2.4. The servo module *S* controls the outside mirrors (M_1 and M_4) through the links to control the stereo view.

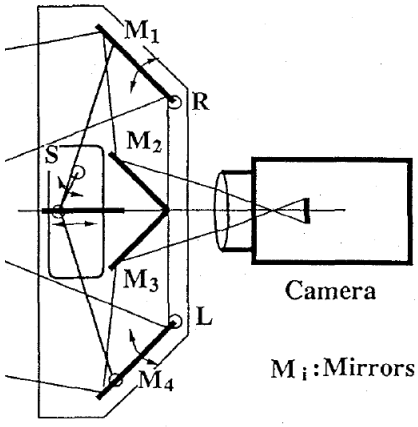


Fig. 2.4 A single-lens stereovision system using moveable mirrors

In the later years, mirror-based systems were further refined. Nene and Nayar [30] performed further analysis on this kind of system by exploring its performance using not only planar mirrors, but also hyperbolical, ellipsoidal, and paraboloidal mirrors. By using such non-planar reflecting surfaces, a wide FOV could be obtained. Later, researchers also considered that the mirror-based single-lens stereovision system should not be limited by capturing the images only from the mirrors, but also directly from the object. Motivated by this, Zhang and Tsui [22] proposed such kind of system by placing an planar mirror next to the object and capturing both the object and the reflecting image from the mirror. Francois *et al.* [24] further refined the concepts of stereovision from a single perspective of a mirror symmetric scene and concluded that a mirror symmetric scene is equivalent to observing the scene with two cameras, and all traditional analysis tools of binocular stereovision can be applied.

2.1.2 Single-lens Stereovision System Using Known Cues

Efforts were also made to the investigation of single-lens stereovision system using known cues, such as illuminations, known geometry of objects, etc. The system designed by Segan and Kumar[27] employs a point light source to track user's hand in 3D space. The light source and camera were fixed and the calibration for both light source and camera needed to be

done in advance. Then, the 3D position and orientation of the hands or fingers could be determined from the projections of hand and its shadow. Moore and Hayes [31] introduced a single-lens stereovision system which could track the position and orientation of objects by exploiting the perspective projection model. Three coplanar points on the subject were identified and their distances from the camera lens were measured, which were used as the cues for estimating the location and orientation of the objects.

Single-lens stereovision system could also be realized by only using one camera with its known movement. LeGrand and Luo [26] proposed a technique which retained the nonlinear camera dynamics and provided an accurate 3D position estimation of selected targets within the environment. This technique was practically used in robotic manipulation, navigation, and localization. Besides, it could also provide probabilistic information for making intelligent and informed decisions.

Adelson and Wang [28] described a novel camera for performing the analysis of how the light striking the lens aperture. By noticing the fact that the light striking a given sub-region of the lens aperture is structured somewhat differently than the light striking an adjacent sub-region, they exploited an extra lenticular array placed at the sensor plane to form a plenoptic camera, and to achieve the depth information by the proposed method.

However, these systems using known cues highly depend on the external resources, which would limit the range of their applications. For example, the systems employing the light sources may not effectively work in an outdoor environment due to the interference from the sun lights.

2.1.3 Bi-prism-based Single-lens Stereovisions (BSS) System

The prototype of biprism-based single-lens stereovision system was proposed by Lee and Kweon [25], who employed a biprism in front of a camera to generate the pair of stereo images. The advantages of the system are obvious on cost saving, compactness, and lesser system parameters. They utilized the geometry of biprism to describe the forming process of stereo image pairs. These image pairs were captured on the left and right halves of the image plane due to the refraction of the light rays through the prism. They also introduced the concept of virtual points to solve the corresponding problem. By simplifying the biprism structure and the refraction process, they proposed a simple method to calculate the disparities of the virtual points. However, these simplifications are only valid in ideal cases and with strict constraints. In real experiments, the errors would become significant when the structure of the biprism changed, such as increasing the angle of the biprism.

Based on Lee's work, Lim and Xiao [11, 12] improved the single-lens stereovision system and extended the biprism study to multi-face prism. Different from Lee's work, they introduced the concept of virtual cameras instead of using the virtual points. Meanwhile, they presented a geometrical ray approach to calibrate the virtual cameras. Basically, the biprism in front of the CCD camera with its apex bisecting the CCD image plane would refract the incident light rays. By back projecting these incident light rays, left and right virtual camera positions and orientations could be estimated. They also demonstrated that the focal lengths of virtual cameras were nearly equal to that of the real camera. Fig. 2.5 shows a 3-D schematic diagram of the left and right virtual cameras generated by the biprism in front of a CCD camera. Using this system for depth recovery requires the knowledge of the disparities of corresponding points in the two images [11]. Image distortion will affect adversely the accuracy of the

disparities, and hence the recovered 3D depth. More details on the formation of the virtual cameras and the common view zone will be reviewed in the following section.

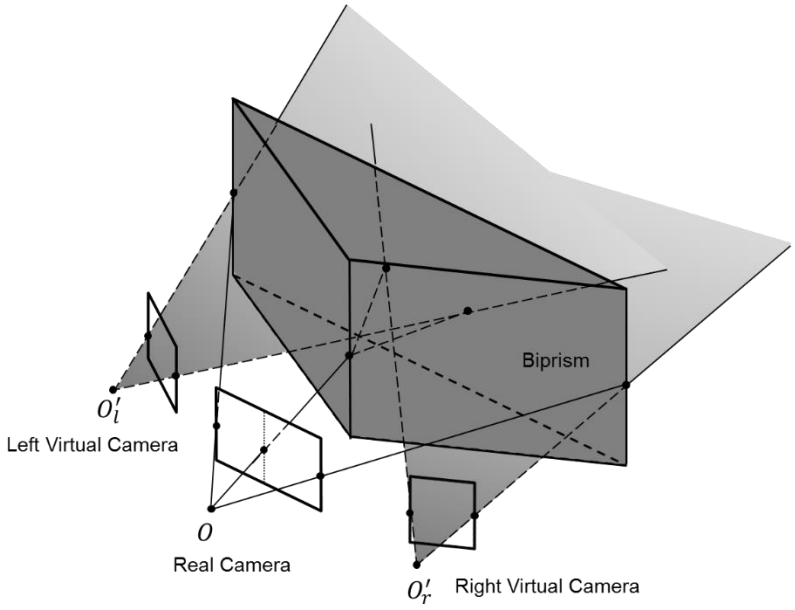


Fig. 2.5 3-D schematic diagram of biprism-based single-lens stereovision

Li and Wang [32] further studied the nonlinear magnification and color aberration of the biprism stereo adapter in the application of stereoscopic. Genovese *et al*. [33] pointed out the limitations of the biprism-based single-lens stereovision system and adopted a robust generalized stereo-system calibration framework to improve the accuracy of the system. However, this framework requires complex mechanical fixture assembly which demands high accuracy, which may not be available for most of the applications.

2.1.4 Formation of the Virtual Cameras

The system presented in this thesis achieves stereovision effect with the aid of a CCD camera and a biprism positioned in front of the camera. This system, as schematically presented in Fig. 2.6, is known as biprism-based single-lens stereovision (BSS) system. As mentioned previously in Section 2.1.3, this set-up is able to simultaneously capture two views of the same

scene to form one stereo image pair. To demonstrate the formation of the virtual cameras and the corresponding field of views (FOVs), a simplified demonstration of the virtual cameras of the BSS system is shown in Fig. 2.6.

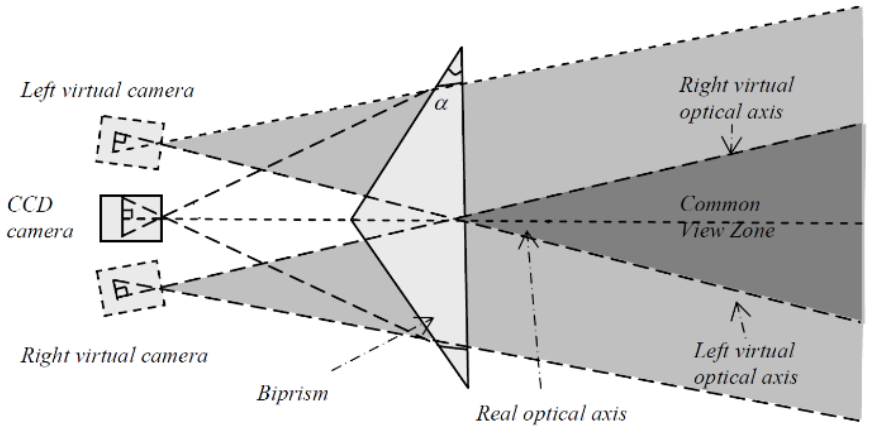
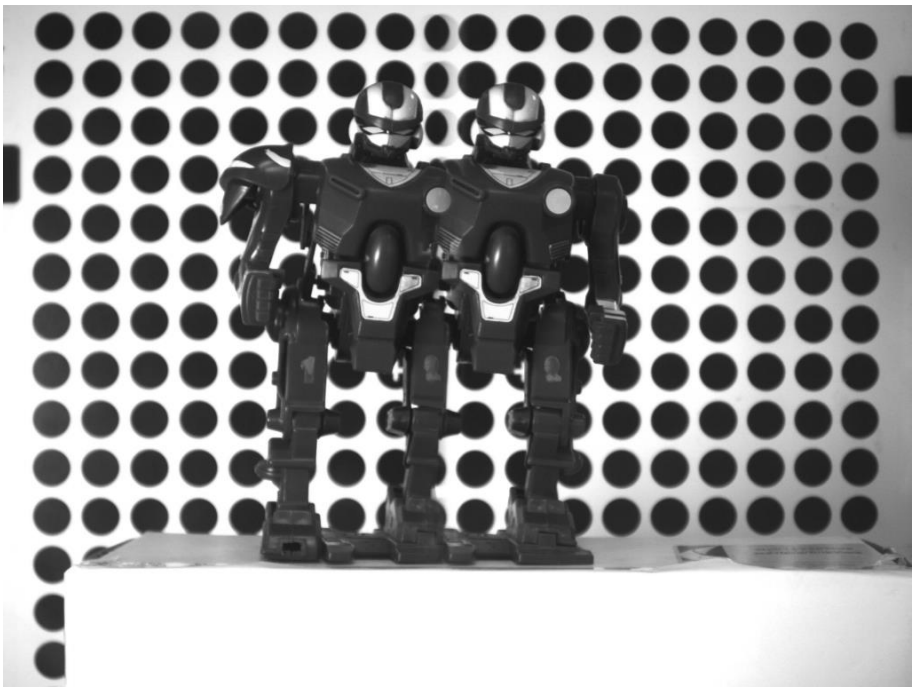


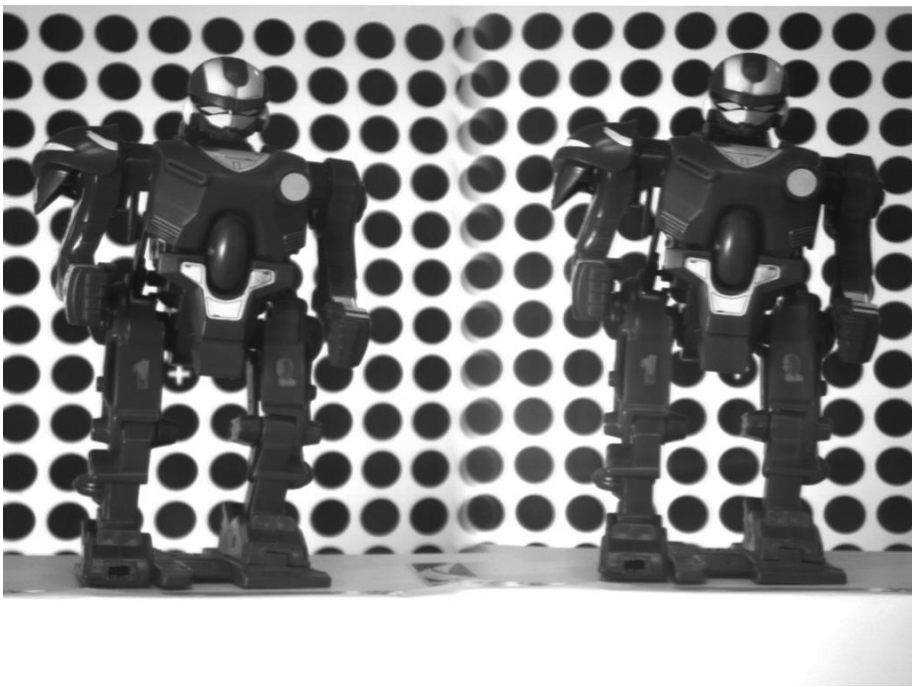
Fig. 2.6 Virtual cameras and the corresponding FOVs of the BSS system (top view)

As shown in Fig. 2.6, for each virtual camera, there exists its own FOV which is constrained by two boundary lines: one is the back-projection of the real camera's FOV boundary refracted by one side of the biprism; the other one is the back-projection line of the real camera's optical axis refracted by the same side of the biprism, and it acts as the optical axis of the virtual camera as well. The overlap region between the left and right virtual camera FOVs is known as the common FOV. This region defines the area in which an object can be 'captured' by both the left and right virtual cameras and thus is fully projected on the real camera image plane. It is also a region in which an object should be placed in order for correspondence searching to be possible. As the viewpoints of these two virtual cameras are different, there exists difference in the coordinates in the left and right images of the same object in the common FOV. Fig. 2.7 shows two examples of stereo image pairs using biprisms with different prism angles, in which the "robot" are located at different positions in the left and right

images. This difference in image location is the disparity of two common points on the scene objects. The disparity is necessary in the determination of the 3D depth of the object in question. Therefore, with this setup, the depth information of the scene can be recovered,



(a) Stereo image pair using 6.4° biprism



(b) Stereo image pair using 22° biprism

Fig. 2.7 Two examples of stereo-image pair taken by the BSS system using (a) 6.4° biprism; (b) 22° biprism

Several studies and developments based on this system have been proposed, which included system construction [34, 35], calibration [13-15], and stereo matching or correspondence search [36-38]. These studies were based on the assumptions that the system is ideally constructed and the image distortions introduced by the prism are negligible. However, unlike standard stereovision systems using independent cameras whose distortions might be negligible, or could be evaluated and eliminated easily using existing method [20], the distortion induced by the biprism in front of the camera will always exist, even under ideal system setup. Specifically, while being refracted by the biprism, some incident lights from different projection angles may be twisted. In addition, the lens distortion from the CCD camera itself may also further distort the projections. In consequence, they appear as unexpected distortions on the image plane. Taking the stereo image pairs in Fig. 2.7 as the examples to illustrate the distortion situation by the biprism, we notice that the bigger the biprism angle was, the more obvious the distortion would be.

In the next sections, a review on the typical lens distortions and the atypical distortion from the biprism will be studied, followed by the distortion correction methods proposed in the recent years.

2.2 LENS DISTORTION

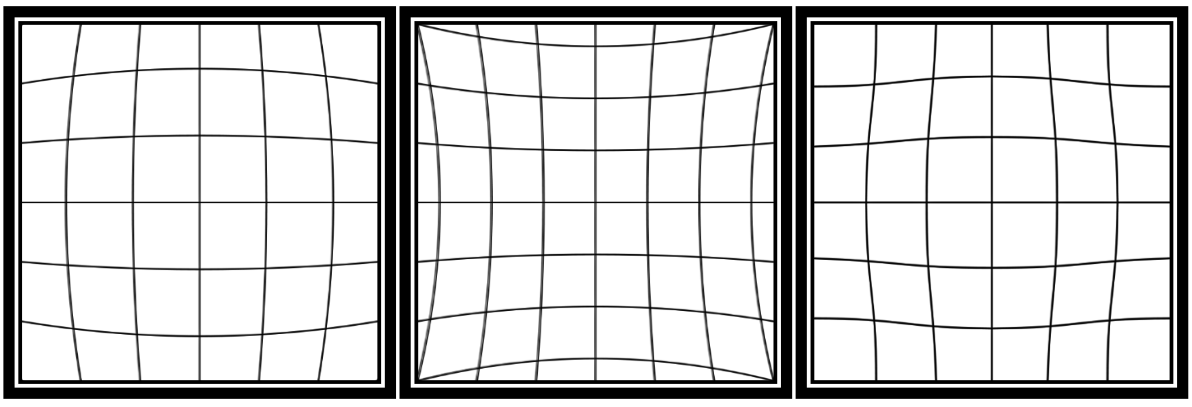
In geometric optics, distortion refers to a deviation from rectilinear projection. For example, the straight lines in a scene would remain straight after the rectilinear projection. But with distortion, the lines may not be straight in the image plane. It is a form of optical aberration. Part of the reasons for the image distortion is the lens geometry. In principle, the lens with a simple convex shape, like spherical surface, will be sufficient for photographing. However, this simple lens would cause many optical aberrations, such as image distortions. In practice,

camera lens is made up of several optical lens elements to reduce the optical aberrations as much as possible. Unfortunately, however hard the manufacturers try to create flawless optics, there is simply no way to manufacture lens that does not suffer from distortions and aberrations to some degree.

2.2.1 Typical Distortions

The most commonly encountered distortions, as well as the most severe ones, are the radial distortion, decentering distortion, and thin prism distortion [39]. These types of geometrical distortions have been well recognized and dealt with to different extent [40, 41]:

- **Radial Distortion:** this type of distortion is a point-symmetric at the optical center of the lens and causes an inward or outward shift, or a mixture of both shifts of image points from their initial perspective projection. The extent of this displacement increases in image points further away from the optical center. The examples of the radial distortion in image are shown in Fig. 2.8.



(a) Outward shift (b) inward shift (c) a mixture type

Fig. 2.8 Different types of radial distortions.

The radial distortion of a perfectly center-aligned lens (δ_r) can be expressed by the following expression [39]:

$$\delta_r = K_1 r^3 + K_2 r^5 + K_3 r^7 + \cdots \quad \textit{Eq. 2.1}$$

where r is the radial distance from the distortion center which is assumed to be the principle center, and $K_1, K_2, K_3, \ldots$ are the radial distortion coefficients.

For each image point represented by polar coordinates (r, φ), the radial distortion corresponds to the distortion along the radial (r) direction. Hence, the image point can also be expressed in terms of the Cartesian coordinates (x, y) with

$$\begin{cases} x = r\cos(\varphi) \\ y = r\sin(\varphi) \end{cases} \quad \textit{Eq. 2.2}$$

and the amount of distortion along each of the Cartesian image coordinates can be represented by

$$\begin{cases} \delta_{xr} = K_1 x(x^2 + y^2) + O[(x, y)^5] \\ \delta_{yr} = K_1 v(x^2 + y^2) + O[(x, y)^5] \end{cases} \quad \textit{Eq. 2.3}$$

$O[(x, y)^5]$ is the high order term of the radial distortion. Practically, the radial distortion is primarily dominated by the low order radial components [42]. Hence, the high order components could generally be ignored in the distortion calibration and correction procedure.

- **Decentering Distortion:** as the camera lens is made up of a number of optical elements, their optical centers may not align well. This misalignment of the optical centers in the sensor induces a decentering distortion which has both radial and tangential components, and can be described analytically by the following expressions [39]:

$$\begin{aligned} \delta_{rd} &= 3(p_1 r^2 + p_2 r^4 + \cdots)\sin(\varphi - \varphi_0) \\ \delta_{td} &= (p_1 r^2 + p_2 r^4 + \cdots)\cos(\varphi - \varphi_0) \end{aligned} \quad \textit{Eq. 2.4}$$

where the quantity φ_0 is the angle between the positive X axis and a line of reference known as the axis of maximum tangential distortion as shown in Fig. 2.9.

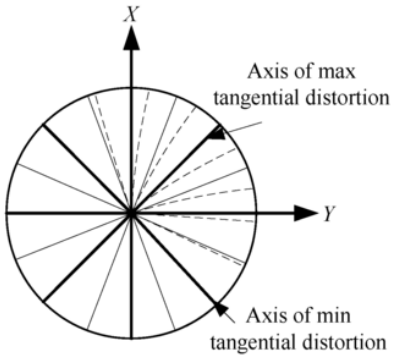


Fig. 2.9 Effect of tangential distortion. Solid lines: distortion-free; dashed lines: with tangential distortion

The resulting amount of this distortion along the X and Y axes are given in terms of δ_{rd} and δ_{td} by

$$\begin{pmatrix}\delta_{xd}\\ \delta_{yd}\end{pmatrix}=\begin{pmatrix}\cos(\varphi) & -\sin(\varphi)\\ \sin(\varphi) & \cos(\varphi)\end{pmatrix}\begin{pmatrix}\delta_{rd}\\ \delta_{td}\end{pmatrix} \qquad \textit{Eq. 2.5}$$

Given the relationships of $\cos(\varphi)={}^{x}/_{r}$ and $\sin(\varphi)={}^{y}/_{r}$, and letting $P_1=-p_1\sin(\varphi_0)$ and $P_2=p_1\cos(\varphi_0)$, the mathematical expression of the decentering distortion could be expressed by:

$$\begin{cases}\delta_{xd}=P_1(3x^2+y^2)+2P_2xy+O[(x,y)^4]\\ \delta_{yd}=2P_1xy+P_2(x^2+3y^2)+O[(x,y)^4]\end{cases} \qquad \textit{Eq. 2.6}$$

- **Thin Prism Distortion:** this kind of distortion is mainly caused by the manufacturing imperfections of lens elements and misalignment of CCD sensor arrays from their ideal, perpendicular orientation to the optical axis. Thin prism distortion could also be represented as additional radial and tangential distortions, but with different forms of expression:

$$\begin{aligned}\delta_{rd}&=(s_1r^2+s_2r^4+\cdots)\sin(\varphi-\varphi_1)\\ \delta_{rd}&=(s_1r^2+s_2r^4+\cdots)\cos(\varphi-\varphi_1)\end{aligned} \qquad \textit{Eq. 2.7}$$

where φ_1 is the angle between the positive X axis and the axis of maximum tangential distortion shown in Fig. 2.9. Letting $S_1=-s_1\sin(\varphi_1)$ and $S_2=s_1\cos(\varphi_1)$, the thin prism distortion along the X and Y axes could be expressed in the following forms:

$$\begin{cases}\delta_{xp} = S_1(x^2 + y^2) + O[(x, y)^4] \\ \delta_{yp} = S_2(x^2 + y^2) + O[(x, y)^4]\end{cases} \quad \textit{Eq. 2.8}$$

- **Total Distortion:** Until now, three types of lens distortions have been discussed and described. Although the decentering distortion and thin prism distortion are similar in the forms of coefficients shown in Eq. 2.4 and Eq. 2.7, respectively, they model different types of distortion, and in particular, each may have a different axis of maximum tangential distortion. When all the above distortions are present in the image, the effective distortion can be modeled by the addition of the corresponding distortion expressions. Generally, the high orders in these expressions are considered negligible and can be omitted. Hence, the total amount of distortion along the X and Y axes can be obtained by combining Eq. 2.3, Eq. 2.6 and Eq. 2.8:

$$\begin{cases}\delta_x = K_1 x(x^2 + y^2) + P_1(3x^2 + y^2) + 2P_2 xy + S_1(x^2 + y^2) \\ \delta_y = K_1 y(x^2 + y^2) + 2P_1 xy + P_2(x^2 + 3y^2) + S_2(x^2 + y^2)\end{cases} \quad \textit{Eq. 2.9}$$

2.2.2 Distortions in BSS System

In the BSS system, the extra optical device, the biprism in front of the camera, could be considered as an additional optical element of the camera lens. The manufacturing imperfections and misalignments of the biprism would also induce unexpected distortions to the projections. Furthermore, unlike the camera lens with spherical surfaces, the plane surfaces of the biprism lead to a different form of distortions. In other words, the image distortion caused by the biprism would be dissimilar with the typical radial or tangential distortions. Fig. 2.10 shows one example of the distorted checkerboard image captured by our single-lens stereovision system. The blurry area in the image is the intersection of the left and right sub-images formed by the left and right faces of the biprism. As can be seen from the image, the unexpected displacement of the pixels in the vertical direction is much more

serious than the one in the horizontal direction, since the vertical lines appear more “curvy” after the projection, but the horizontal lines remain unchanged or with unobservable distortion. This distortion phenomenon is different from the typical lens distortion introduced in § 2.2.1. Thus, correcting the distortion caused by the biprism becomes a challenge because existing lens distortion correction methods may not be applicable.

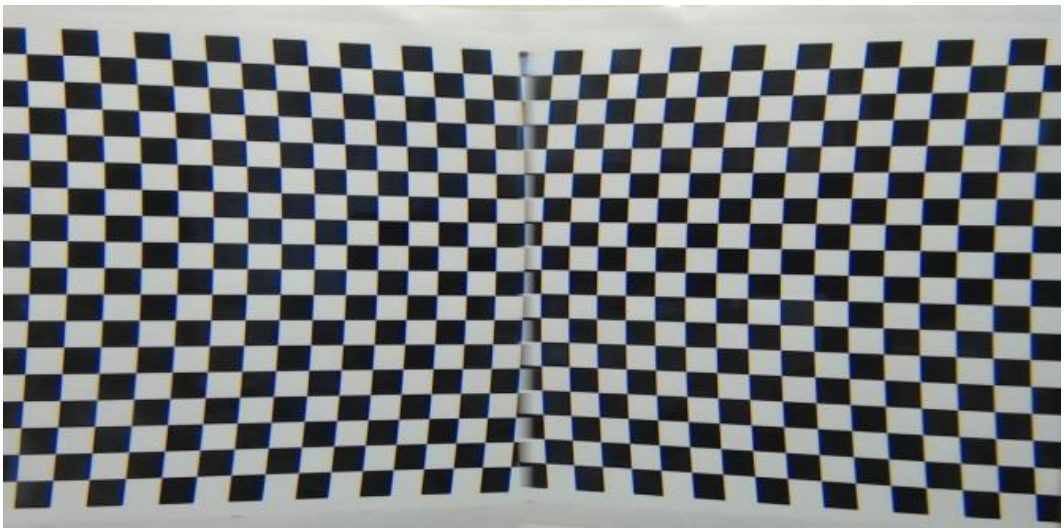


Fig. 2.10 Distortions in biprism-based single-lens stereovision system

2.3 LENS DISTORTION CORRECTION METHODS

For camera lens distortion correction, large amounts of relevant literatures have been produced. The approaches could be basically classified into two groups: model-based approaches and model-free ones. Model-based approaches correct the entire image by defining a global rule; model-free methods adjust the image based on the local correction of the distortion.

2.3.1 Model-based Distortion Correction Methods

The early and popular model-based distortion correction method was proposed by Brown [43, 44], who combined the radial, the decentering and the thin prism distortions together as the three major components of the regular lens distortion. Practically, the approach was to correct the whole image by a parametric nonlinear polynomial function of the distance between the generic point and the distortion center. The distortion parameters could be estimated together with the pin-hole camera model through camera calibration methods [20, 39, 45], or calculated according to an idealized pin-hole camera after the camera was calibrated. However, the distortion was most often coupled with intrinsic and extrinsic parameters of the system, thus the estimation depended on the pin-hole model would cause high errors [39]. To avoid this, Hartley [46] proposed a non-parametric model in which only the radial distortion was considered. Although the radial distortion was predominant in most cases, the decentering and the thin prism distortions should still be taken into account since they were always coupled. Some other methods used correspondences between points in different images from multiple viewpoints to estimate the distortion parameters [47-49], or calibrated the high-order correlations in frequency domain [50]. In [51], a more detailed survey was presented, which gave a quantitative comparison performances of the most frequently used calibration and model-based distortion correction techniques.

Recent distortion correction methods use the correspondences between two or several images, which require no knowledge of any camera information. The main tool they use is slackened epipolar constraints, which incorporate lens distortion into the epipolar geometry. Several iterative [52-54] or non-iterative methods [55-59] are used to estimate the distortion and to correct it. These methods are used with a low order parametric distortion model and therefore cannot achieve high precision.

2.3.2 Model-free Distortion Correction Methods

In contrast, another family of model-free approaches has been presented in recent years. Instead of correcting the whole image with a fixed parametric function, locally adjusting the sub-areas of the image according to the regional correction of the distortion seems to be more accurate than model-based methods, especially for dealing more accurately with complex imaging systems like stereo optical microscopes [60], scanning electron microscopes [61, 62] and medical endoscopes [63-65]. One outstanding example of using mode-free method to correct the lens distortion was proposed in [66]. More recently, Genovese *et al*.[33] proposed a model-free optimization-based procedure to calibrate the distortion for a particular single-lens system by mapping the distorted images with the undistorted ones captured from the standard stereovision system. This method relied on the complex system and might induce unexpected errors if the reference points were coupled with distortions.

In addition, some methods were proposed which took advantage of some image features like straight lines [65, 67-72], vanishing points[73-75], cross-ratio [76], or the image of a sphere [77]. Ricolfe *et al.*[78, 79] presented a robust model-free distortion correction approach (MLDC) combining the cross-ratio and line constraints of a chess board, which could isolate the distortion correction from the camera calibration process under stable conditions. It was claimed in his work [79] that the corrected points using MLDC method are not a uniquely right solution but a better one. However, although it satisfies the straight line and cross-ratio constraints, this method may fail to meet the vanishing point constraint which is indeed one of the checkerboard template constraints. Besides, MLDC is sensitive to the position of the distortion center which is assumed to be around the center of the image. Otherwise, the correction may fail, or at least not accurate enough.

2.4 SUMMARY

In this chapter, various types of single-lens stereovision systems have been reviewed. In addition to dealing with the different forms of distortion in the system, different image correction algorithms have also been studied. Based on the knowledge gained from the literature, different ideas will be gathered to formulate novel approaches in distortion corrections, especially for our BSS system. These approaches will be presented in details in the following chapters of this thesis.

Chapter 3. A GEOMETRICAL ANALYSIS BASED DISTORTION CORRECTION (GABDC) METHOD

To better understand the distortion in the BSS system presented in § 2.1.3, the first step is to analyze the formation of the distortion due to the presence of a biprism. Benefitting from the simple system setup, the parameters of the system, including the camera parameters and the biprism properties, could be obtained either from the measurement or the technical documents. In this manner, it is feasible to express the distortion mathematically by applying a geometrical analysis based method to study the projection of the rays in the system. A similar geometrical method has been proposed and studied in our earlier works, either for camera calibration [11], correspondence searching [13, 37], or quantization error analysis of the system [80]. However, all of them ignore the distortion caused by the biprism.

Specifically, both distorted and undistorted projections from 3D coordinates to 2D image coordinates through the biprism and pin-hole camera could be calculated based on the given geometrical information of the system. Here we need to point out that the undistorted projections refer to the distortion-free (or ideal) projections whose definitions will be given in §3.2. To calculate both distorted and undistorted projections, a 3D geometrical analysis based method is proposed first to determine the distorted back-projection rays. This geometrical analysis applies three fundamentals: Snell's law, the concept of virtual camera, and ray tracing in which all the pertinent points, lines, and planes are expressed in the 3D camera coordinates. Thereafter, the corresponding undistorted rays are determined based on the undistorted projection assumptions which will be presented in §3.2. Hence, the correlation between the distorted and undistorted image points can be calculated. With this

correlation, the distortion correction maps can then be determined and used for the distortion correction for the given BSS system.

3.1 DISTORTION FORMATION ANALYSIS

Ideally, with a pin-hole camera model, the function of the biprism is only to bisect the incoming scene along the apex line into two sub-scenes without inducing any distortion, as illustrated in Fig. 2.6. In this case, the refraction is assumed to occur on a horizontal plane. However, when considering the imaging procedure in a 3D sense, as schematically illustrated in Fig. 3.1, the projection error will always be introduced by the biprism, even when the real camera is modelled as an ideal pin-hole camera. In this case, the refraction does not occur on a horizontal plane. This projection error is considered as the distortion caused by the biprism in this system. Specifically, since the biprism has certain thickness, the projection in both X_C and Y_C directions (w.r.t. camera coordinates) will be "twisted" according to the projecting angle due to refraction. Taking the projection in the Y_C direction shown in Fig. 3.1 as an example, the refraction on either side of the biprism may lead to a projection deviation in the Y_C direction between the real (distorted) ray and the ideal (undistorted) ray. Meanwhile, it can be deduced that this deviation will vanish when the incident ray is close to the Z_C axis, namely, the projection angle $\beta \to 0$. In our system, this deviation is considered as the distortion caused by the biprism. Fig. 3.2 shows an example of the images captured with and without biprism placed in front of the camera. As can be seen, the real camera is distortion-free, but the distortion induced by the biprism is quite obvious, in both horizontal and vertical directions.

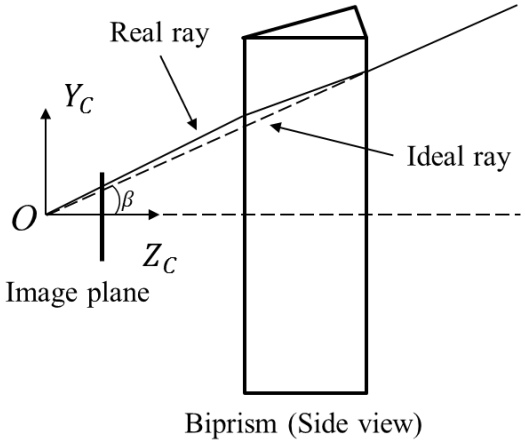


Fig. 3.1 Side view of the refraction deviation between the ideal and real incident rays.

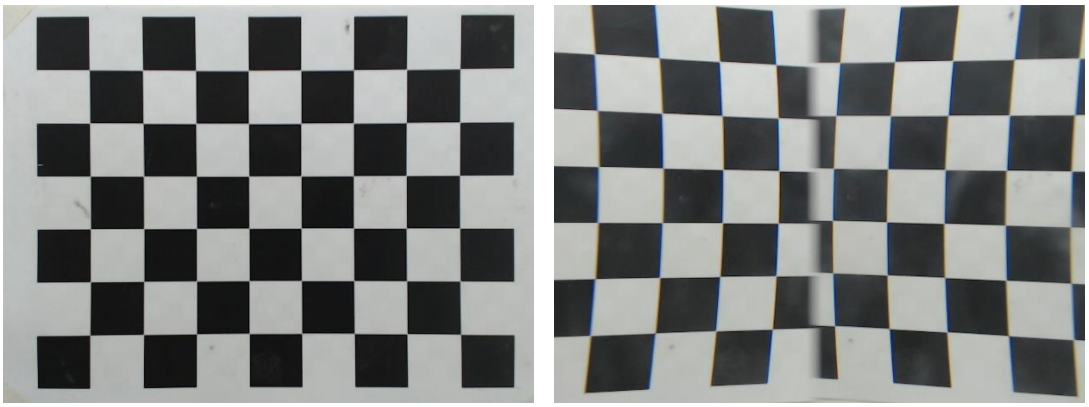


Fig. 3.2 Images captured without (left) and with (right) biprism placed in front of the camera

This simple demonstration of the biprism distortion formation only gives a qualitative description the biprism distortion. To mathematically express the formation of the biprism distortion, and hence to build the distortion correction maps for the system, 2D view of the refraction may not be sufficient because of the complex refraction through the biprism. Therefore, in the following sections, a more detailed 3D refraction analysis based on the back-projection method and the geometrical relationship between the camera and the biprism are discussed. Meanwhile, the assumptions used to define the ideal projections are also proposed.

As a result of the analysis, both distorted and undistorted projections can be determined, and then the distortion correction maps could be built and evaluated consequently.

3.2 DISTORTED AND UNDISTORTED PROJECTION ANALYSIS

To simplify the determination of the distorted and undistorted projections, a back-projection analysis [11, 13] (Assume that rays are traced from the image plane to the 3D target) is applied instead of the natural projection way (Rays are traced from the 3D target to the image plane). There are two advantages of using back-projection for the analysis. One advantage is it guarantees that the projection is located within the camera FOV, since all the selected image points are within the image boundaries. The other advantage is that the calculation of back-projection is simpler and more straightforward than that of the natural projection . Compared with the natural projection way, it avoids the difficulty of determining which incident light from the 3D target is the one projected on the image plane, since the projection through the biprism is more complicated than a pin-hole model.

For the discussion of the back-projection in this thesis, definitions of the real (distorted) and ideal (undistorted) projections are proposed. According to distortion phenomenon described in §3.1, the real rays refracted by the biprism will induce distortions in the system. Thus, these projections which obey the law of refraction are defined as the distorted projections. In contrast, the undistorted rays are defined as the projections to the virtual cameras without deviation by the biprism in both the X_C and Y_C directions. In other words, they may not follow the law of refraction. To realize that, three assumptions of the ideal projection are imposed to the projection procedure to define the undistorted rays:

1. The rays through the optical centre of the real camera are distortion-free, and are named as “Fundamental Rays (FRs)”;

2. The refractions on the front and back faces of the biprism only lead to rays' rotations about the Y_C axis towards the FRs;
3. The rotation angles in Assumption 2 are the same with the ones in FRs at each face, respectively.

As mentioned in §2.1.4, FRs are also the fundamental rays which are used to form the virtual cameras, and their refractions at both faces are only in the X_CZ_C plane. The second assumption eliminates the deviation in the Y_C direction as illustrated in Fig. 3.1. The last assumption is to guarantee that the projection angle of the ideal rays w.r.t. the FRs is consistent before and after the refraction by the biprism. Specifically, the projection angle between any incident ray and the FRs will always be the same before and after the passing through the biprism, as illustrated in Eq. 3.3. In this figure, θ denotes the angle between $\overrightarrow{R1_l'}$ and $\overrightarrow{R1_l}$ at the optical center O; γ_1 and γ_2 are the angles between the emergent rays ($\overrightarrow{R2_l'}$ and $\overrightarrow{R3_l'}$) and $\overrightarrow{R1_l'}$ of FRs, respectively; β_1 and β_2 are the angles between the emergent rays ($\overrightarrow{R2_l}$ and $\overrightarrow{R3_l}$) and $\overrightarrow{R1_l}$ of the ideal (undistorted) ray, respectively. According to the third assumption of the ideal projection, we can get $\gamma_1 = \beta_1$ and $\gamma_2 = \beta_2$. In this way, the projection angles between the ideal ray and the FRs after each refraction remain unchanged.

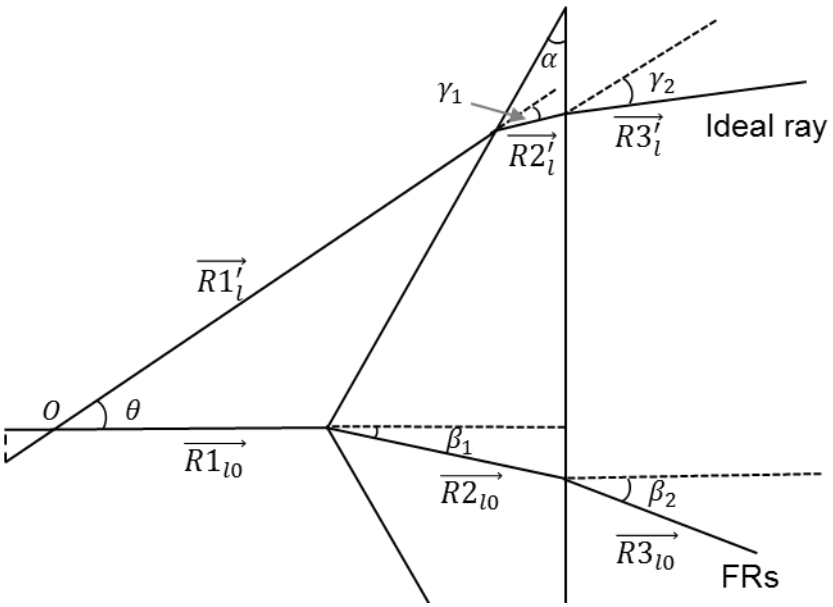


Fig. 3.3 A top view of the ideal back-projection the FRs.

A detailed demonstration of the distortion formation on the left side of the biprism is illustrated in Fig. 3.4. We considered two back-projected rays: the ray $\overrightarrow{R1_{l0}}$ starting from O is passing through the principle point (image centre p_{l0}) and acting as the reference ray without distortion (FRs); the other ray $\overrightarrow{R1_l}$ also from O is passing through a random sample point p_l on the image plane, whose real back-projection $\overrightarrow{R3_l}$ deviates from the ideal one $\overrightarrow{R3_l}'$. All these refracted rays $\overrightarrow{R3_{l0}}, \overrightarrow{R3_l}, \text{and } \overrightarrow{R3_l}'$ should intersect at the same point O_l, which is the optical centre of the left virtual camera. According to the definition of the ideal (undistorted) projection, the angle between $\overrightarrow{R3_l}'$ and $\overrightarrow{R3_{l0}}$ should be the same as the one between $\overrightarrow{R1_l}$ and $\overrightarrow{R1_{l0}}$. The analysis in the following sections focuses on the determination of the mathematical expressions of these rays w.r.t. the real camera coordinates $\{O \rightarrow X_C, Y_C, Z_C\}$. They are then used to build the relationship between the distorted and undistorted projections of the image. To carry out this geometrical analysis, we assume that the following system setup parameters are known, either by measurement or from technical data:

1. α, the corner angle of the biprism;

2. n, the refractive index of the biprism material;
3. D, the distance from the real camera pinhole to the apex of biprism;
4. T, the length of biprism in horizontal direction (perpendicular to the apex line);
5. f, the focal length of real camera, and;
6. I, the real camera image width.

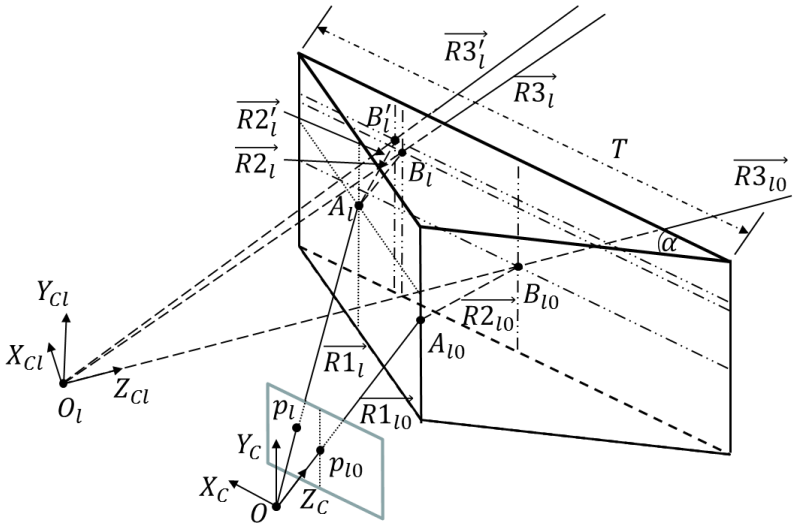


Fig. 3.4 Refraction ray trace of the incident lights in 3D space using back-projection.

3.2.1 Distorted Back-projection

We first analyze the distorted projection, which totally obeys Snell's law in 3D space. Based on the back-projection concept, the ray tracing of the real projection in Fig. 3.4 could be interpreted in the following way: a ray starts from the camera origin O, and passes through the sample pixel p_l (ray $\overrightarrow{R1_l}$). It hits the front face of the biprism at A_l.It is then refracted along the ray $\overrightarrow{R2_l}$ and leaves the back face of the prism at B_l. After the second refraction, the ray $\overrightarrow{R3_l}$ goes through the target in 3D space. We shall determine the rays $\overrightarrow{R1_l}$, $\overrightarrow{R2_l}$ and $\overrightarrow{R3_l}$, and the intersection points A_l and B_l according to the geometrical relationships.

Snell's law in vector form is defined as

$$\overrightarrow{V_{refract}} = r\vec{l} + \left(rc - \sqrt{1 - r^2(1 - c^2)}\right)\vec{n} \qquad \textit{Eq. 3.1}$$

in which $r = n_1/n_2$ is the ratio of the refractive indices in the air and glass; $\vec{l}$ is the incident ray in vector form; $\vec{n}$ is the normalized plane normal vector, and $c = -\vec{n} \cdot \vec{l}$.

Suppose that p_l is located at (x_l, y_l, f) w.r.t. the camera coordinates $\{O \to X_C, Y_C, Z_C\}$. Then the first ray $\overrightarrow{R1_l}$ could be written in the vector form:

$$\overrightarrow{R1_l} = (r1_1, r1_2, r1_3) = norm(x_l, y_l, f) \quad \textit{Eq. 3.2}$$

The function $norm(\cdot)$ refers to the normalization of the input vector which makes the length of the vector 1 (e.g. $\left\|\overrightarrow{R1_l}\right\| = 1$). For convenience, all the vectors in the following sections in this chapter are represented in the normalized form. The position of the intersection point A_l (in the form of $\left(X_{A_l}, Y_{A_l}, Z_{A_l}\right)$) on the front face could be determined (Appendix D):

$$\begin{cases} X_{A_l} = \dfrac{D\ x_l/f}{1 - \tan(\alpha)\ x_l/f} \\ Y_{A_l} = \dfrac{D\ y_l/f}{1 - \tan(\alpha)\, x_l/f} \\ Z_{A_l} = \dfrac{D}{1 - \tan(\alpha)\, x_l/f} \end{cases} \quad \textit{Eq. 3.3}$$

Then, we need to calculate the normal vector on A_l. It is easy to note that this normal vector $\overrightarrow{n_{A_l}}$ only involves the prism angle α:

$$\overrightarrow{n_{A_l}} = (\sin(\alpha), 0, -\cos(\alpha)) \quad \textit{Eq. 3.4}$$

Substituting Eq. 3.2 and Eq. 3.4 into Eq. 3.1, we could get the ray vector $\overrightarrow{R2_l}$:

$$\overrightarrow{R2_l} = (r2_1, r2_2, r2_3) = \frac{1}{n} \cdot \overrightarrow{R1_l} + \left(-\frac{1}{n} \cdot \overrightarrow{n_{A_l}} \cdot \overrightarrow{R1_l} - \sqrt{1 - \frac{1}{n^2}\left(1 - \left(\overrightarrow{n_{A_l}} \cdot \overrightarrow{R1_l}\right)^2\right)} \right) \cdot \overrightarrow{n_{A_l}} \quad \textit{Eq. 3.5}$$

with its initial point A_l. After that, $\overrightarrow{R2_l}$ exits from the back plane of the biprism at B_l. The normal vector at point B_l is

$$\overrightarrow{n_{B_l}} = (0, 0, -1) \quad \textit{Eq. 3.6}$$

Thus, the emergent ray $\overrightarrow{R3_l}$ could be calculated by substituting Eq. 3.5 and Eq. 3.6 to Eq. 3.1 and using $\overrightarrow{R2_l}$ as the incident ray:

$$\overrightarrow{R3_l} = (r3_1, r3_2, r3_3) = n \cdot \overrightarrow{R2_l} + \left(-n \cdot \overrightarrow{n_{B_l}} \cdot \overrightarrow{R2_l} - \sqrt{1 - n^2\left(1 - \left(\overrightarrow{n_{B_l}} \cdot \overrightarrow{R2_l}\right)^2\right)}\right) \cdot \overrightarrow{n_{B_l}} \quad \textit{Eq. 3.7}$$

with the point $B_l = \left(X_{B_l}, Y_{B_l}, Z_{B_l}\right)$, which could be calculated from A_l and $\overrightarrow{R2_l}$ (Appendix D):

$$B_l = A_l + \frac{\overrightarrow{R2_l}}{r2_3} \cdot \tan(\alpha) \cdot \left(\frac{T}{2} - X_{A_l}\right) \quad \textit{Eq. 3.8}$$

Hence, all the rays in different directions could be determined. With the known variables (f, D, T, α, n), all the expressions from Eq. 3.2 to Eq. 3.8 are the functions of (x_l, y_l), they could be used to express the ray $\overrightarrow{R3_l}$ and its initial point B_l. This group of functions is named as "Distorted Back-projection Formulas" (DBPFs) for convenience.

$$\left[\overrightarrow{R3_l}, B_l\right] = DBPFs(x_l, y_l) \quad \textit{Eq. 3.9}$$

3.2.2 Undistorted Back-projection

Based on the three conditions of the ideal (undistorted) projection presented in §3.2, the Law of Refraction is only used to calculate the FRs, while the other emergent rays are determined according to the initial orientation w.r.t. the FRs. At the beginning, these FRs, such as $\overrightarrow{R1_{l0}} = (r01_1, r01_2, r01_3)$, $\overrightarrow{R2_{l0}} = (r02_1, r02_2, r02_3)$ and $\overrightarrow{R3_{l0}} = (r03_1, r03_2, r03_3)$ for the left virtual camera, can be determined by the DBPFs with $\overrightarrow{R1_{l0}} = (0,0,1)$ as the initial incident ray, as shown in Fig. 3.3. We can see that the FRs are on the $X_C Z_C$ plane, and rotated about the Y axis on each intersection point. Besides, according to the constraints of the ideal projections, the generation of the ideally projected rays could be simplified by rotating the emergent rays about the Y axis with the same degrees as the corresponding FRs at each intersection point.

First, the rotation matrix about the Y_C axis could be defined as:

$$R_y(\vartheta) = \begin{bmatrix} \cos(\vartheta) & 0 & \sin(\vartheta) \\ 0 & 1 & 0 \\ -\sin(\vartheta) & 0 & \cos(\vartheta) \end{bmatrix} \quad \textit{Eq. 3.10}$$

in which ϑ is the rotation angle about the Y_C axis. Then, with the known FRs, the rotation angles β_1 (by $\overrightarrow{R1_{l0}}$ and $\overrightarrow{R2_{l0}}$) and β_2 (by $\overrightarrow{R1_{l0}}$ and $\overrightarrow{R3_{l0}}$) are:

$$\beta_1 = \operatorname{acos}\left(\frac{\overrightarrow{R1_{l0}} \cdot \overrightarrow{R2_{l0}}}{\left\|\overrightarrow{R1_{l0}}\right\| \left\|\overrightarrow{R2_{l0}}\right\|}\right);\ \beta_2 = \operatorname{acos}\left(\frac{\overrightarrow{R1_{l0}} \cdot \overrightarrow{R3_{l0}}}{\left\|\overrightarrow{R1_{l0}}\right\| \left\|\overrightarrow{R3_{l0}}\right\|}\right) \quad \textit{Eq. 3.11}$$

Since $\overrightarrow{R1_{l0}}$, $\overrightarrow{R2_{l0}}$ and $\overrightarrow{R3_{l0}}$ are on the *X-Z* plane and $\overrightarrow{R1_{l0}} = (0,0,1)$, this calculation could be simplified as

$$\beta_1 = \operatorname{atan}\left(\frac{r02_1}{r02_3}\right);\ \beta_2 = \operatorname{atan}\left(\frac{r03_1}{r03_3}\right) \quad \textit{Eq. 3.12}$$

Hence, the undistorted rays $\overrightarrow{R2_l{}'}$ and $\overrightarrow{R3_l{}'}$ and the intersection point $B_l{}'$ on the back face could be calculated:

$$\overrightarrow{R2_l{}'} = \left(r2_1', r2_2', r2_3'\right) = \overrightarrow{R1_l} \cdot R_y(\beta_1) \quad \textit{Eq. 3.13}$$

$$\overrightarrow{R3_l{}'} = \left(r3_1', r3_2', r3_3'\right) = \overrightarrow{R1_l} \cdot R_y(\beta_2) \quad \textit{Eq. 3.14}$$

$$B_l' = A_l + \frac{\overrightarrow{R2_l{}'}}{r2_3'} \cdot \tan(\alpha) \cdot \left(\frac{T}{2} - X_{A_l}\right) \quad \textit{Eq. 3.15}$$

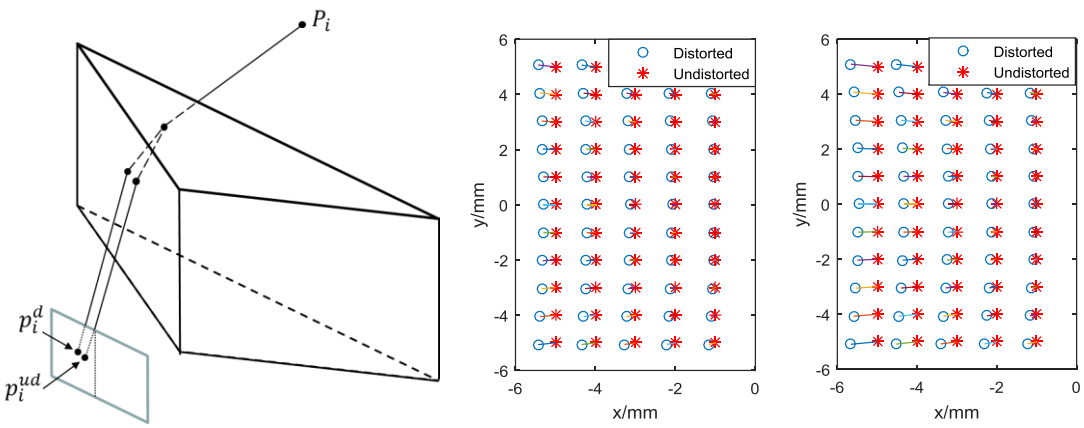
These expressions to determine the undistorted rays is named as "Undistorted Back-projection Formulas" (uDBPFs) in this thesis for convenience.

$$\left[\overrightarrow{R3_l{}'}, B_l'\right] = uDBPFs(x_l, y_l) \quad \textit{Eq. 3.16}$$

Note that the above DBPFs and uDBPFs are all based on the camera coordinates. Besides, according to the geometrical property of the biprism, the projection analysis for the right image is similar to the one for the left image, but symmetrical w.r.t. the apex line of the biprism.

3.2.3 Distorted and Undistorted Image Points

The above back-projection analysis is to determine the incident rays hitting the same point on the image plane with and without distortion. However, for the purpose of understanding the distortion correction with respect to the image coordinates, we would like to determine the projection relationship between the distorted and undistorted image points of the same 3D point, such as p_i^{ud} and p_i^d shown in Fig. 3.5(a), instead of the corresponding rays from the same image point. Here p_i^{ud} and p_i^d are the ideal (undistorted) and real (distorted) projection points from the same 3D point P_i, respectively. Generally, the determination of p_i^{ud} and p_i^d requires the inverse derivation of DBPFs and uDBPFs, since both of the formulas could only be initialized by the image points. However, according to the above geometric analysis in §3.2.1 and §3.2.2, the inverse derivations for both DBPFs and uDBPFs could be quite complex and tedious.



(a) (b) (c)

Fig. 3.5 (a) Ray tracing for distorted projection p_i^d and undistorted projection p_i^{ud} of the same 3D point P_i; (b) and (c) are the sample points with (o) and without (∗) distortion on the image plane using thin and thick biprisms, respectively.

To avoid this inverse derivation problem, a nonlinear optimization (or nonlinear data-fitting) method, Levenberg-Marquardt algorithm [81], was applied to find the projections of the 3D

points with and without the biprism distortion. Specifically, with the given positions of the 3D points (w.r.t. the camera coordinates) as the outputs of the DBPFs and uDBPFs, the inputs (distorted and undistorted 2D image points) could be determined, respectively. Given the fact that there always exists only one distorted point and one undistorted point on the image plane corresponding to the given 3D point within the camera FOV, a unique solution to this nonlinear optimization problem can be guaranteed.

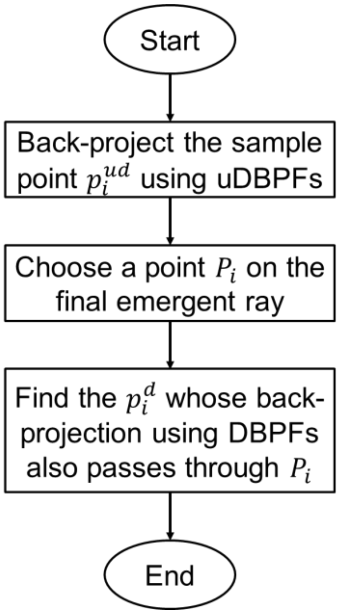


Fig. 3.6Flow Chart of the determination of the distorted and undistorted projections illustrated in Fig. 3.5 (a).

To generate the distorted and undistorted image data sets, three procedures are involved here, as illustrated in Fig. 3.6. The first step is to back-project all the ray traces with the given ith undistorted point p_i^{ud} using uDBPFs (shown in *Fig. 3.5* (a)). The second one is to pick one point P_i on the final emergent ray generated from the first step, which is the ray coming out from the back plane of the prism; then apply the nonlinear optimization method to the DBPFs using the P_i as the output. In this way, the input, the corresponding distorted point p_i^d on the

image sensor, can be determined. This second part can be represented as a minimization problem for the following term:

$$p_i^d = \underset{p_i \in \mathbb{R}^2}{arg\min}(\|uDBPFs(p_i) - P_i\|^2) \qquad \textit{Eq. 3.17}$$

With the undistorted point p_i^{ud} as the initial guess for the input p_i. By solving this minimization problem, each distorted sample point p_i^d can correlate to its corresponding undistorted point p_i^{ud}. Fig. 3.5 (b) and (c) illustrate two examples generated using biprisms with different angles, namely thin and thick biprisms. Note that the initial undistorted image point sets in the two examples are the same and the only difference between these two system setups is the biprism angle.

3.3 Distortion Correction Maps

Given the correlation between the undistorted image points $p^{ud} = (x_{ud}, y_{ud})$ and the corresponding distorted points $p^d = (x_d, y_d)$, the distortion $\Delta = (dx, dy)$ could be regarded as the coordinate adjustment from the distorted points to the undistorted ones, which could be expressed as

$$\Delta = p^d - p^{ud} \qquad \textit{Eq. 3.18}$$

A more specific way to express the distortion in both X and Y axes are given in Eq. 3.19:

$$\begin{cases} dx = x_d - x_{ud} \\ dy = y_d - y_{ud} \end{cases} \qquad \textit{Eq. 3.19}$$

Note that for each point p^d, Δ is different. To better represent the distortions, the dx and dy are plotted separately in 3D plots, with the image coordinates (x_{ud}, y_{ud}) as the X and Y axes and the corresponding dx or dy on the vertical or Z axis. These 3D maps (both dx and dy) are called distortion maps, or correction maps since the distortion could be corrected or removed if these maps are known for the given system.

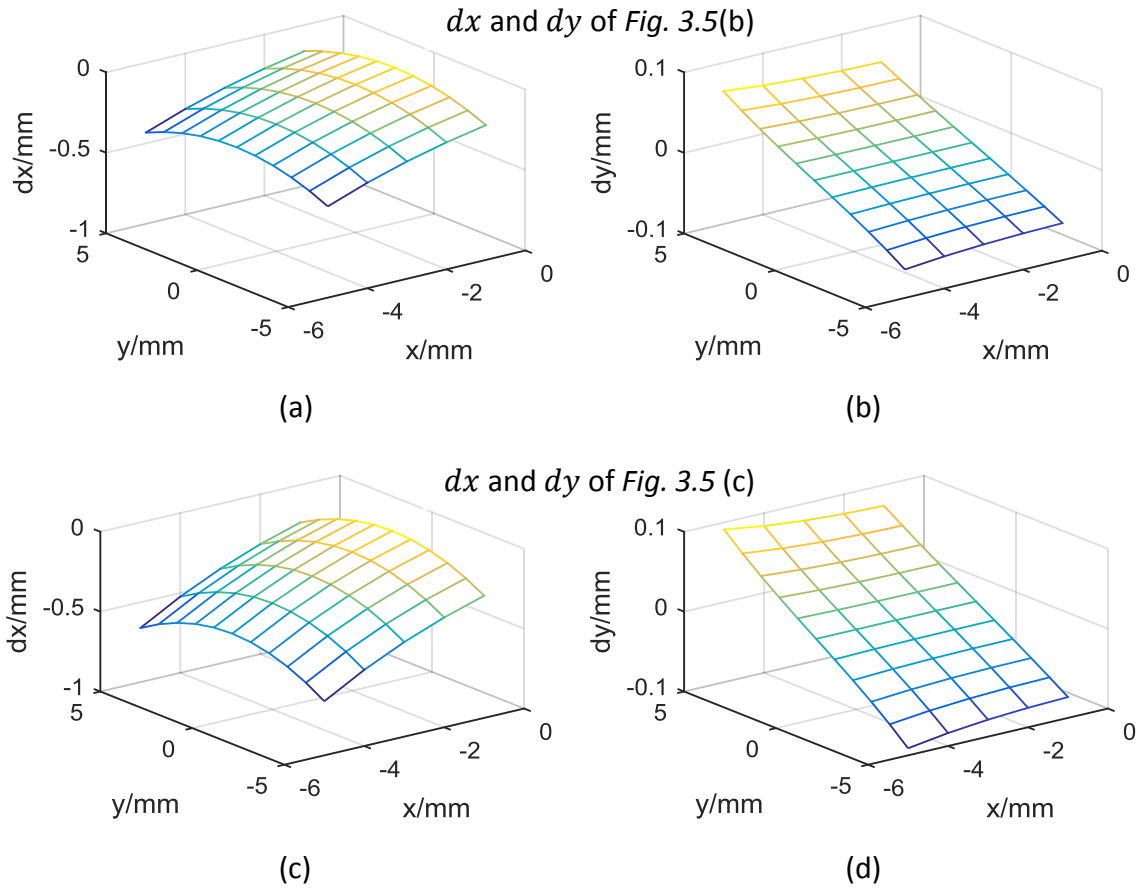


Fig. 3.7. The distortion maps dx (a) and dy (b) from Fig. 3.5 (b), and the distortion maps dx (c) and dy (d) from Fig. 3.5 (c).

Two examples of the distortion correction maps are illustrated in Fig. 3.7. The figures at the first row (Fig. 3.7 (a) and (b)) are the dx and dy distribution of the thin biprism distortion in Fig. 3.5 (b); the ones on the second row (Fig. 3.7 (c) and (d)) are for that of the thick biprism shown in Fig. 3.5 (c). As can be seen, the amount of distortion of dx is much higher than the that of dy, even with small biprism angle. Besides, the biprism angle has less effects on dy comparing to the obvious changes of dx in both magnitude and complexity. These observations reveal that this kind of distortion is different from the typical distortion, which indicates that the existing distortion models (radial, tangential or thin-prism distortions) may not be suitable for modelling and correcting the distortion in this system. Take Zhang's camera calibration method [20] as an example, the determinations of the camera and distortion parameters are always coupled due to the optimization procedure in the method. As a result

of using improper distortion models, the calibration accuracy decreases. In contrast, with the calculated distortion maps, the distortion could be corrected directly without doing the camera calibration.

However, the determination of the proposed distortion correction maps in this section requires the full knowledge of the system parameters, which may lead to average correction results if these parameters contain some errors. This influence of the system parameters on the correction maps will be further studied in the experiments.

3.4 BSS System Setup

The experimental setup of the BSS system is shown in Fig. 3.8, which includes a custom made stable metal stand, a set of fixtures with high precision adjustment possibility, a CCD camera and a biprism. More details on these devices are introduced as follows:

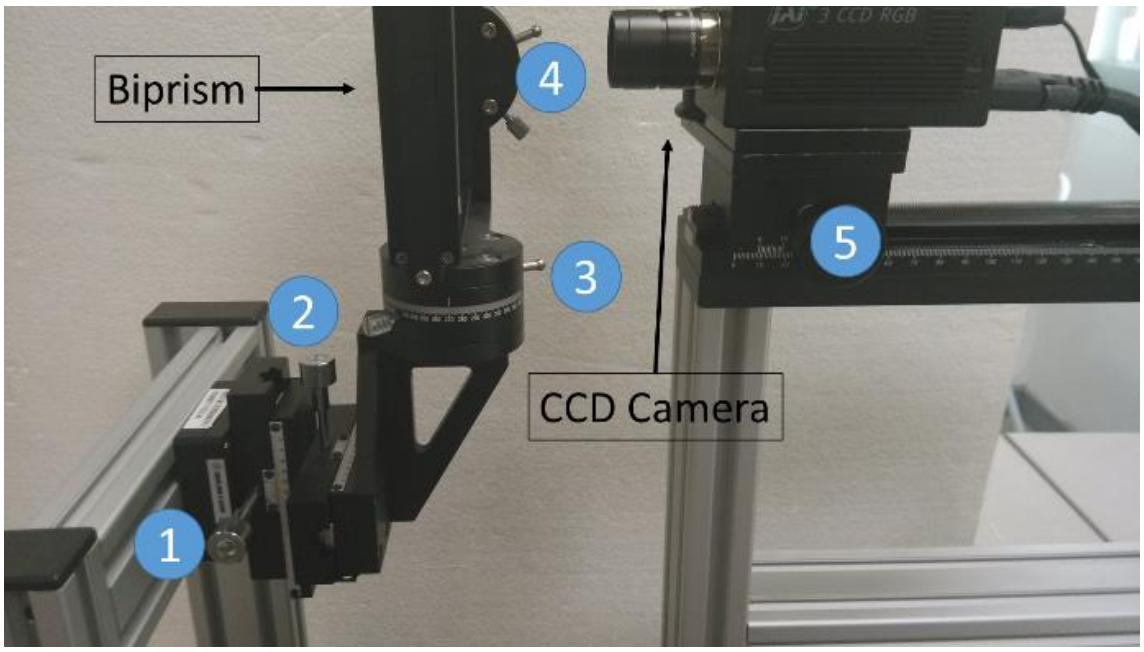


Fig. 3.8 Experimental setup of the biprism-based single-lens stereovision system

- A custom-made mechanical stand: the main function of this stand is to stabilize and secure the relative position and orientation between the camera and biprism.

- A set of fixtures: they consist of 3 vernier calipers labeled as fixtures 1,2 and 5 in Fig. 3.8, and 2 rotating stages labeled as fixtures 3 and 4. The vernier calipers are used to adjust the positions of the biprism and camera, and the rotating stages are for setting the orientations of the biprism. With the combined use of these fixtures, the system setup could be adjusted close to the ideal configuration.
- JAI-CV-M9CL CCD Camera: it is mounted on fixture 5, whose position could be adjusted to be in the forward or backward direction. The specifications of the camera are provided in the Table 3.1.

Table 3.1 Specification of the JAI CV-M9CL camera

Camera Parameters	**JAI-CV-M9CL**
Image width $U/pixels$	1024
Image height $V/pixels$	768
Focal length f/mm	8
Pixel size δ_x/mm	0.00465
Pixel size δ_y/mm	0.00465
Center of the image $(u_c, v_c)/pixels$	(512,384)

- Custom made optical biprism: it is mounted on the frame customized for the biprism. This frame is attached to fixture 3 and fixture 4, so that the orientation of the biprism could be adjusted. Specifically, the biprism should be properly aligned so that the back plane is parallel to the camera image plane and the apex of the biprism should bisect the image plane. One of the biprisms used in our experiments is shown in Fig. 3.9. In our system setup, the biprism frame is fixed, but could be adapted to the biprisms

with different angles but with the same back plane size. In this way, different stereovision system types could be realized by changing the biprisms with different biprism angles.

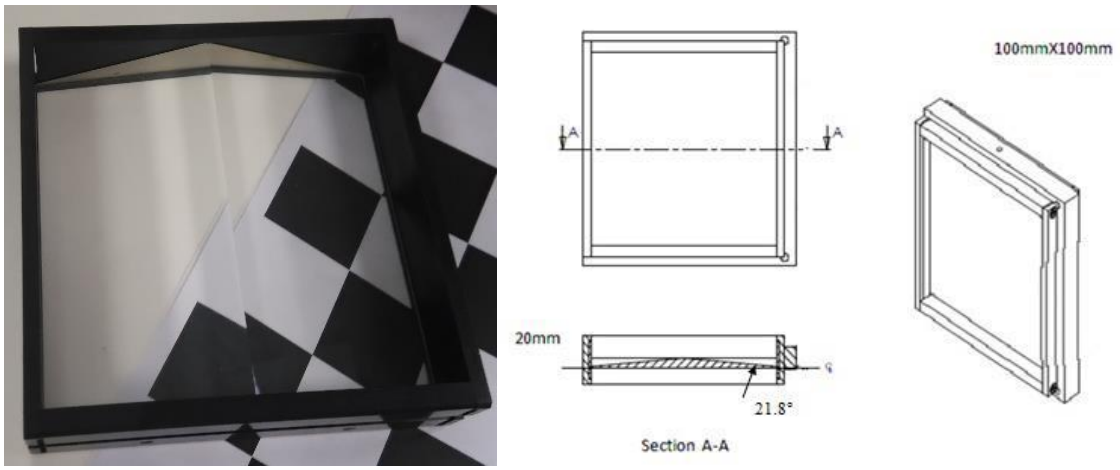


Fig. 3.9 One of the optical biprism used in our experiments

3.5 EXPERIMENTAL RESULTS AND DISCUSSIONS

In the experiments, both simulated and real data are utilized to verify the feasibility of the proposed distortion correction method using the correction maps. Specifically, the simulated data is used to investigate the effects of the system parameters to the correction maps by varying these parameters and measuring the amount of resulting distortion. It helps us to identify which parameter is more crucial to the distortion formation. The real data test is to apply the correction maps to the original distorted images, hence to test the feasibility of the proposed correction method by measuring the stereo camera calibration accuracy and depth recovery accuracy.

3.5.1 Effects of System Parameters on Correction Maps

Recalling the back-projection analysis in §3.2, the determinations of DBPFs and uDBPFs are based on the assumption that all the system parameters are known, either by measurement or from technical data. These parameters are listed in Table 3.2. Note that there also exist other parameters which could affect the distortion correction maps (such as the orientation of the biprism), but the evaluations on them are out of the scope of our investigation in this section.

Table 3.2 System parameters involved in the distortion correction maps determination

	Name	Parameter
Camera	Image width	I
	Focal length	f
Biprism	Corner angle	α
	Back plane width	T
	Index of refraction	n
	Distance to the camera	D

From Table 3.2, we notice that there are 6 parameters involved. If all of them are varied at the same time, the evaluation data could be of 7-Dimension (6 inputs and 1 output), which is very difficult to analyze. Given the real circumstance, in which not all these parameters could be varied, we will focus on the f, α and D in this section. A similar analysis of these three parameters on the depth recovery accuracy has also been proposed in [82]. The difference of our study in this section is to qualitatively analyze the influences of these parameters to the distortion in the system.

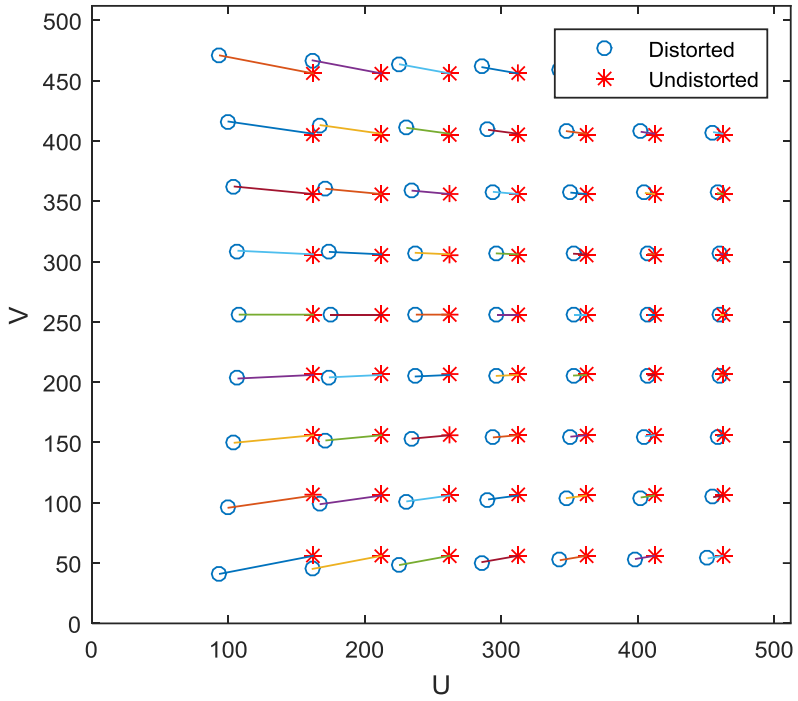
The evaluation measurement used in this section is to calculate the displacement between the distorted and undistorted points on the image plane, which is named as distortion displacement (E_d) in this section:

$$E_d = \left\| p^d - p^{ud} \right\| \quad \text{Eq. 3.20}$$

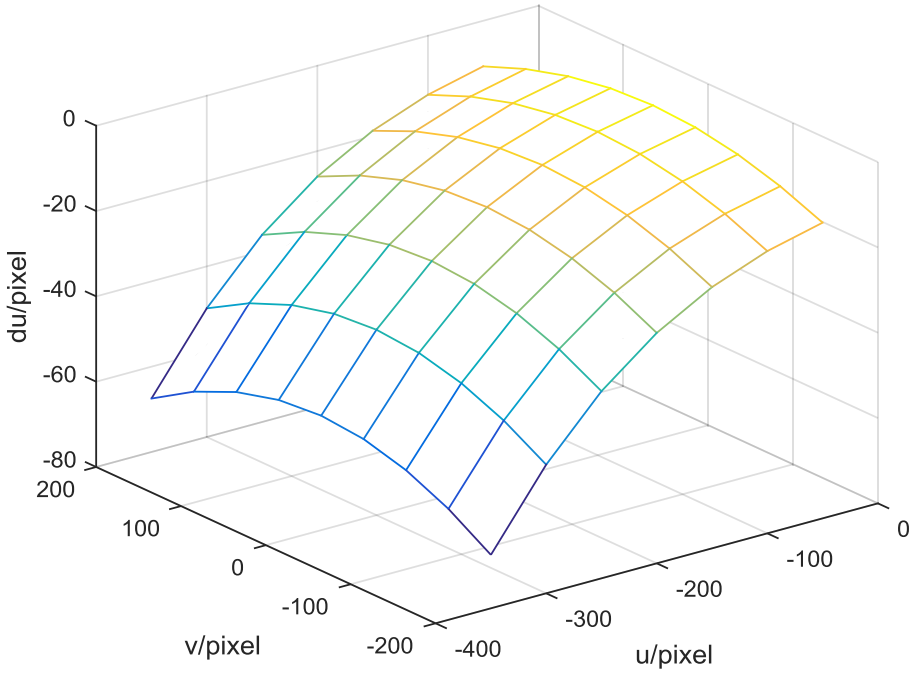
Besides, the displacements in both X and Y directions are calculated as well. The system parameters are list in Table 3.3. System parameters, f, α and D are variables within the given range; and n is fixed. For the simulation, the undistorted sample points on the image plane (resolution of $512 \times 768\ pixels$) are created with a dimension of 9×7 squares (8×6 corner points) and a distance of $50\ pixels$ between neighbor points. They are fixed during the evaluation. In other words, while changing the system parameters, the undistorted sample points remain the same, but their corresponding distorted points vary. Fig. 3.10 shows an example of the distorted and undistorted sample points on the left image plane. The correlation between these two sets of points are labeled with the lines connected between corresponding points.

Table 3.3. System parameters for simulation

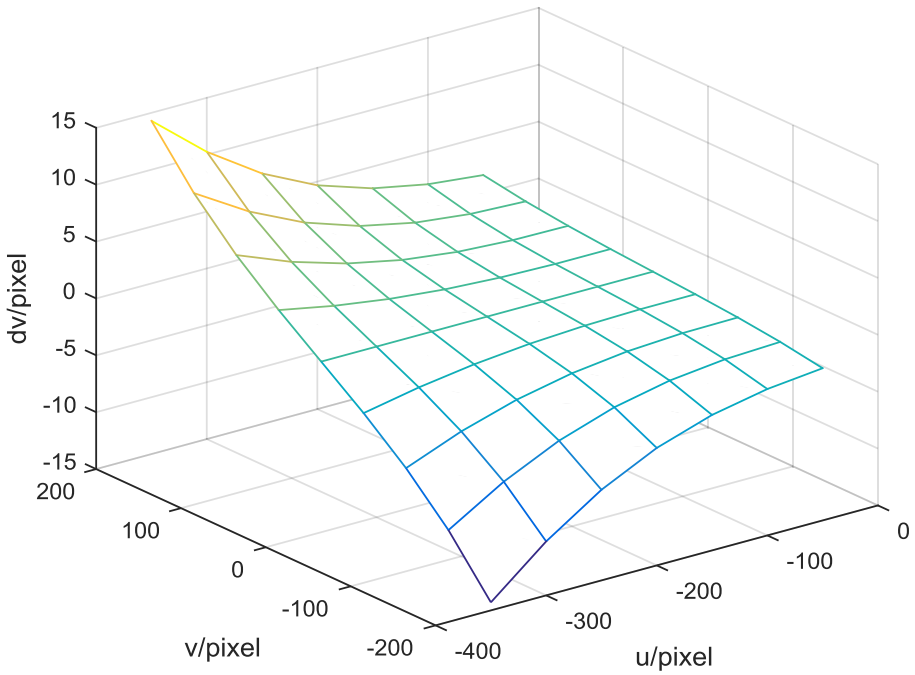
f	D	α	n
$4 \sim 25\ mm$	$100 \sim 200\ mm$	$11^\circ \sim 27^\circ$	1.48



(a) Undistorted and distorted points



(b) Distortion in the X direction



(c) Distortion in the Y direction

Fig. 3.10 Simulated distorted and undistorted sample points on the left image plane using f of 4 mm, D of 150 mm and α of 15 degrees.

3.5.1.1 Variable Focal Length, f

Fig. 3.11 shows the distortion displacement in the system using D of 150 mm, biprism angle α of 15 $degrees$ with different focal lengths varying from 4 to 25 mm. The distortions in the X and Y directions are plotted separately, along with the total distortion. On each box in the figure, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extending to the bottom and top ends indicate the minimum and maximum values, respectively. The upcoming boxplot figures in this section follow the same criteria.

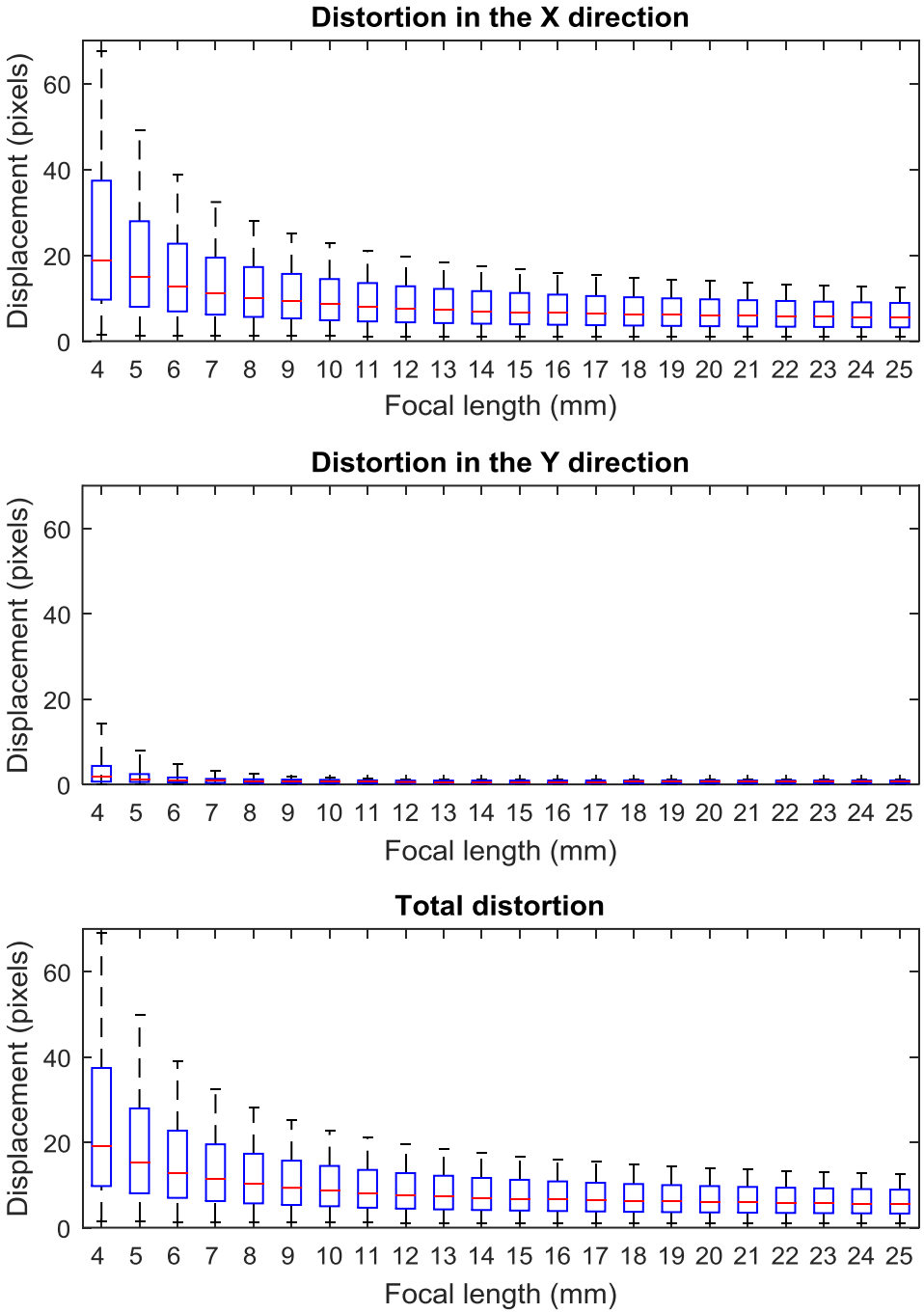


Fig. 3.11 Distortion error using D of 150 mm, α of 15 degrees with different focal lengths

As can be seen from Fig. 3.11, the total distortion could be reduced by increasing the focal length, and it tends toward a stable low value at larger focal length comparing with the dramatic changes at low focal lengths. Moreover, while comparing the distortions in the X and Y directions, it is obvious that the distortions along the X axis is much more severe than

those along the Y direction. In other words, the total distortion in the image is mainly caused by the refraction along the X direction of the biprism This performance is consistent with the analysis in §3.3. In addition, since the image sensor size remains the same, increasing the focal length means decreasing the FOV of the real camera. The results from Fig. 3.11 can also be interpreted as the distortion performance in relation to different FOVs, i.e., the wider the real camera FOV is, the more severe the distortion will be. Hence, using a small FOV camera will help decrease the amount of distortion in the system. However, from the perspective of practical usage, small FOV cameras may not be feasible for some applications.

3.5.1.2 Variable Biprism Angle, α

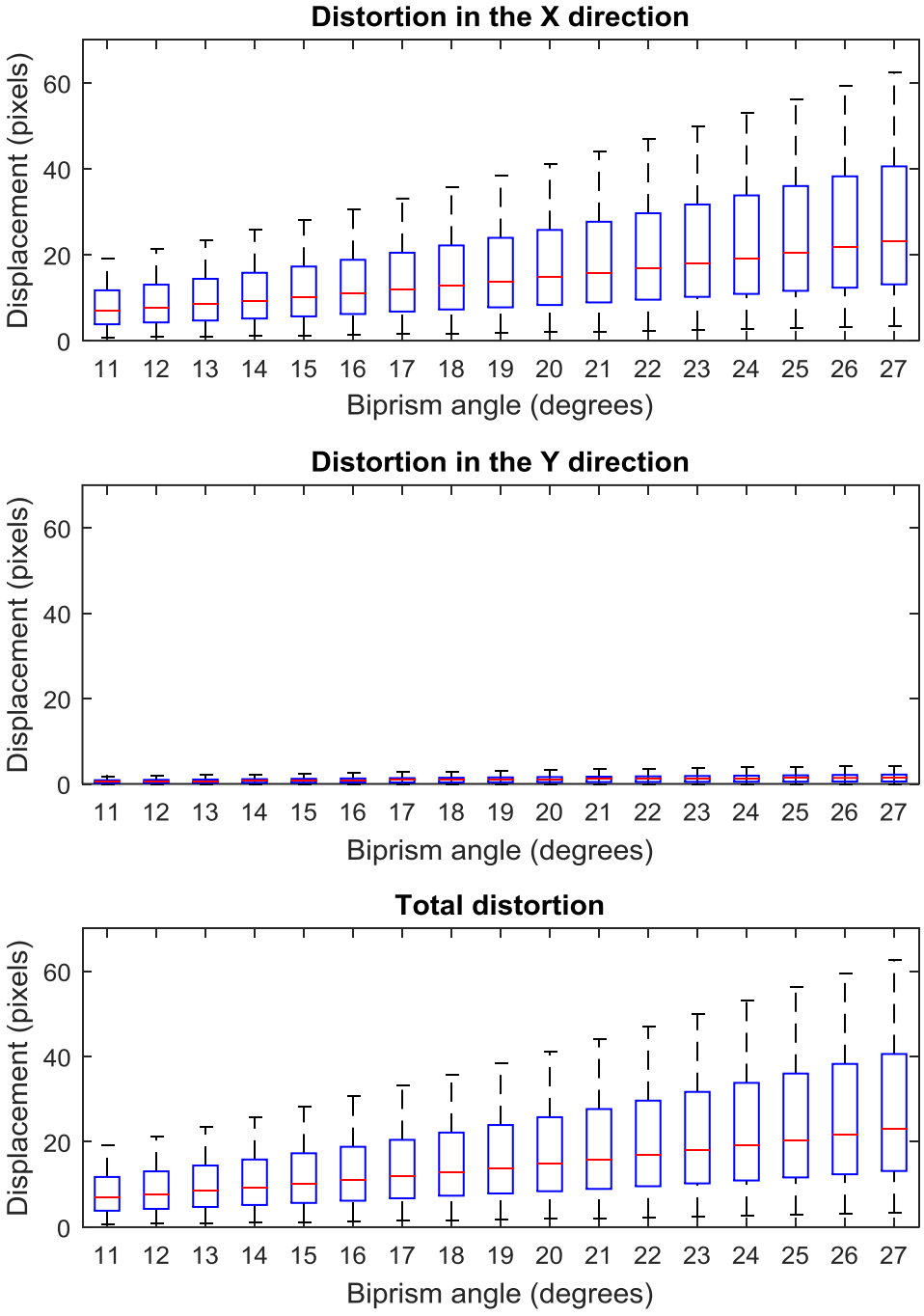


Fig. 3.12 Distortion error using D of 150 mm, f of 8 mm with different biprism angles

Fig. 3.12 shows the distortion characteristics using D of 150 mm, f of 8 mm, and with different biprism angles varying from 11 to 27 $degrees$. As can be seen, the total amount of distortion increases while raising the biprism angle. Moreover, since the variation of the biprism angle mainly affects the refraction in the X direction, the distortion in the Y direction

barely changes and stays at a very low level. Comparing with the distortion in the X direction, the distortion along the Y axis could be considered as negligible. From this experiment, we can deduce that small biprism angle is beneficial if we are trying to reduce the image distortions in the image. However, smaller biprism angle will lead to a smaller stereo disparity and a higher depth recovery error [82].

3.5.1.3 Variable Biprism-Camera Distance, D

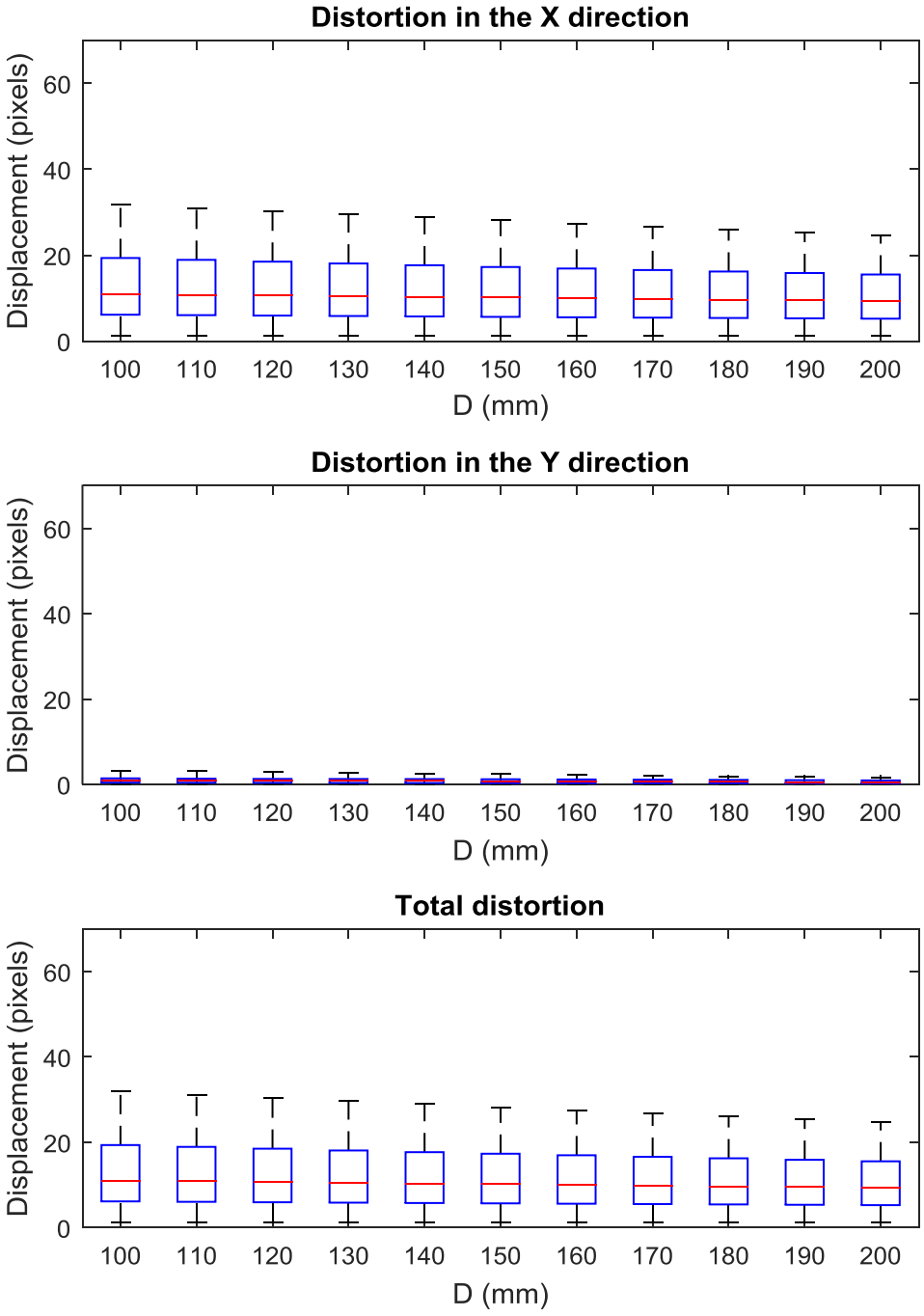


Fig. 3.13 Distortion error using f of 8 mm, α of 15 degrees with different biprism-camera distances

Fig. 3.13 shows the distortion characteristics while using f of 8 mm, α of 15 degrees with different biprism-camera distances. As can be seen in the figure, the total amount of distortion slowly decreases while raising the D value. In other words, the distortion in the image is insensitive to the biprism-camera distance. Besides, the distortion amount in the X

direction is much more than the one in the Y direction. The latter one could even be considered negligible in this setup, which is slightly different with the distortion characteristics while varying f and α. This is because the thickness of the biprism in the Y direction is consistent, which means the refractions in the Y direction remain the same even the distance D is changed.

Therefore, four observations on the distortion characteristics can be listed:

1) Regarding the distortion performance in different directions, the distortion displacement in the X direction is much more severe than that in the Y direction, regardless of the changes in the system parameters. In most cases, the distortions in the Y direction could be considered negligible;
2) Pertaining to the distortion performance associating with the focal length f, the distortion displacement decreases dramatically while increasing f from a small value, but the rate of decline of the distortion slows down at larger values of f;
3) On the distortion performance associating with the biprism angle α, the distortion displacement increases with increasing α.
4) Regarding the distortion performance associating with the biprism-camera distance D, the distortion displacement decreases with increasing D. However, the rate of decline is quite small. In other words, the effect of D to the distortion is quite limited.

3.5.2 Distortion Correction for Real Image Data

In real experiments, different image data sets have been acquired by varying the biprism angle α and the camera focal length f. All the system parameters required for the determination of the distortion correction maps are listed in Table 3.4. The image captured by the real CCD camera can be split into left and right sub-images which are regarded as the images from the

left and right virtual cameras. These sub-images could be corrected using the corresponding distortion maps. Here we only consider the left sub-images in our experiment in distortion correction. Fig. 3.14 shows the sample images of the original distorted images and Fig. 3.15 shows the corrected results using the corresponding distortion correction maps. The correction procedure using the distortion correction maps have been introduced in §3.3.

Table 3.4. System parameters for the experimental setups

Image resolution ($pixels$)	Principal point ($pixels$)	Focal length (mm)	Biprism angle ($degrees$)	Biprism back plane size (mm)	Index of refraction
1024×768	$(512,384)$	$4, 8, 25$	$15.5, 21.8, 35$	100×100	1.48

As can be seen in Fig. 3.14, sample images are arranged according to the system parameters f and α, in order to provide a distinct comparison of the amount of distortion between different system setups. Specifically, regarding the distorted images in Fig. 3.14 with different α but the same f (e.g. images in rows), the amount of distortion increases significantly, especially at a smaller f. However, the amount of distortion decreases when f increases (e.g. images in columns). This performance on the changes of the amount of distortion along different f and α are consistent with the conclusions in §3.5.1.

Fig. 3.15 shows the corrected images corresponding to the distorted images in Fig. 3.14. They are corrected using the pre-determined distortion maps (e.g *Fig. 3.10*(a) and (b)) according to the given system parameters. As can be seen, images with different amounts of distortion can be corrected properly, in which the lines formed by the edges of the checkerboard are straight, instead of curves in Fig. 3.14. Besides, the blank area in the corrected image (*Fig. 3.15*) indicates the correction displacement for the pixels at the edge, which also indicates the maximum amount of the distortion corrected.

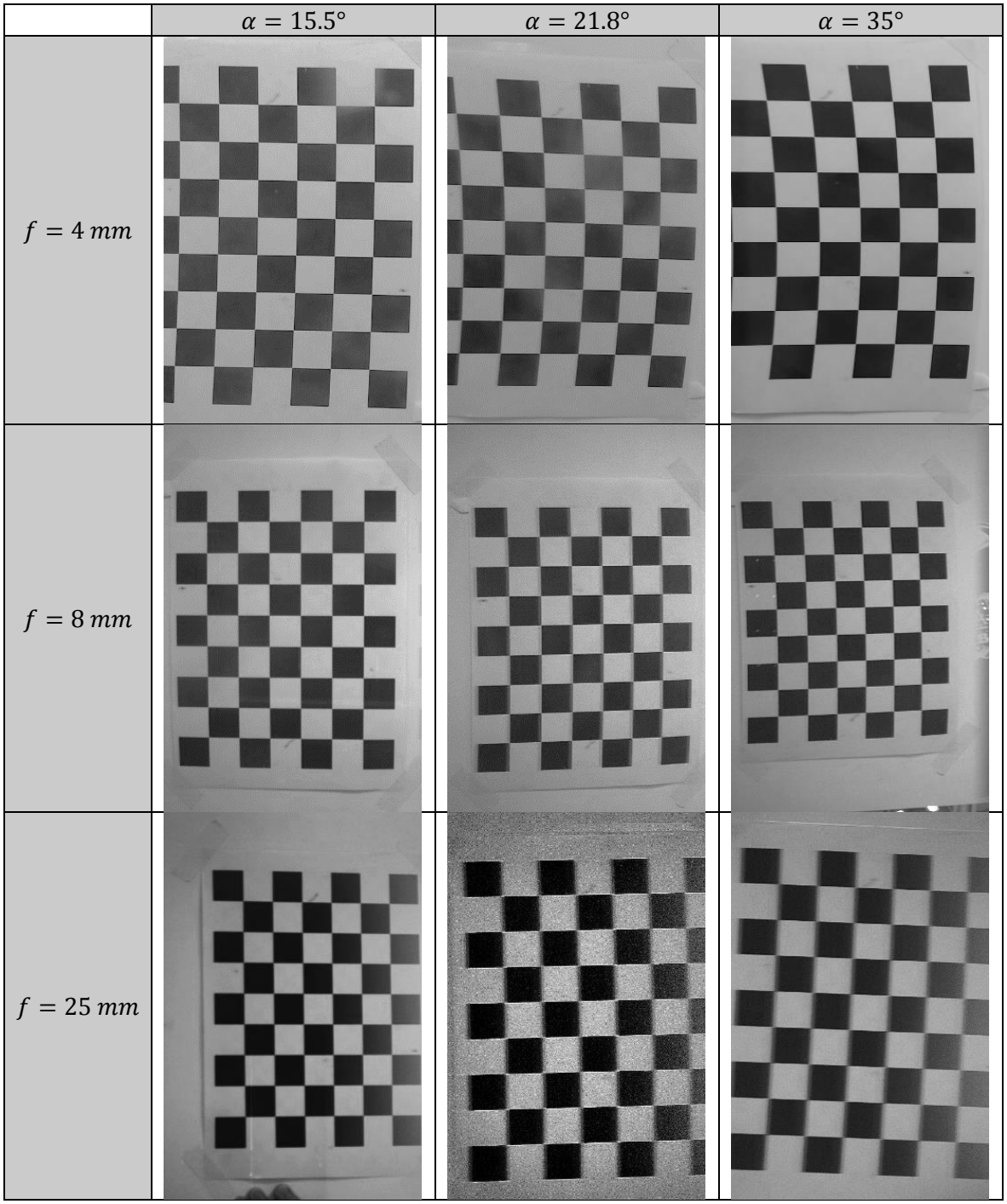


Fig. 3.14 Original distorted images from the systems using different camera focal lengths and biprism angles.

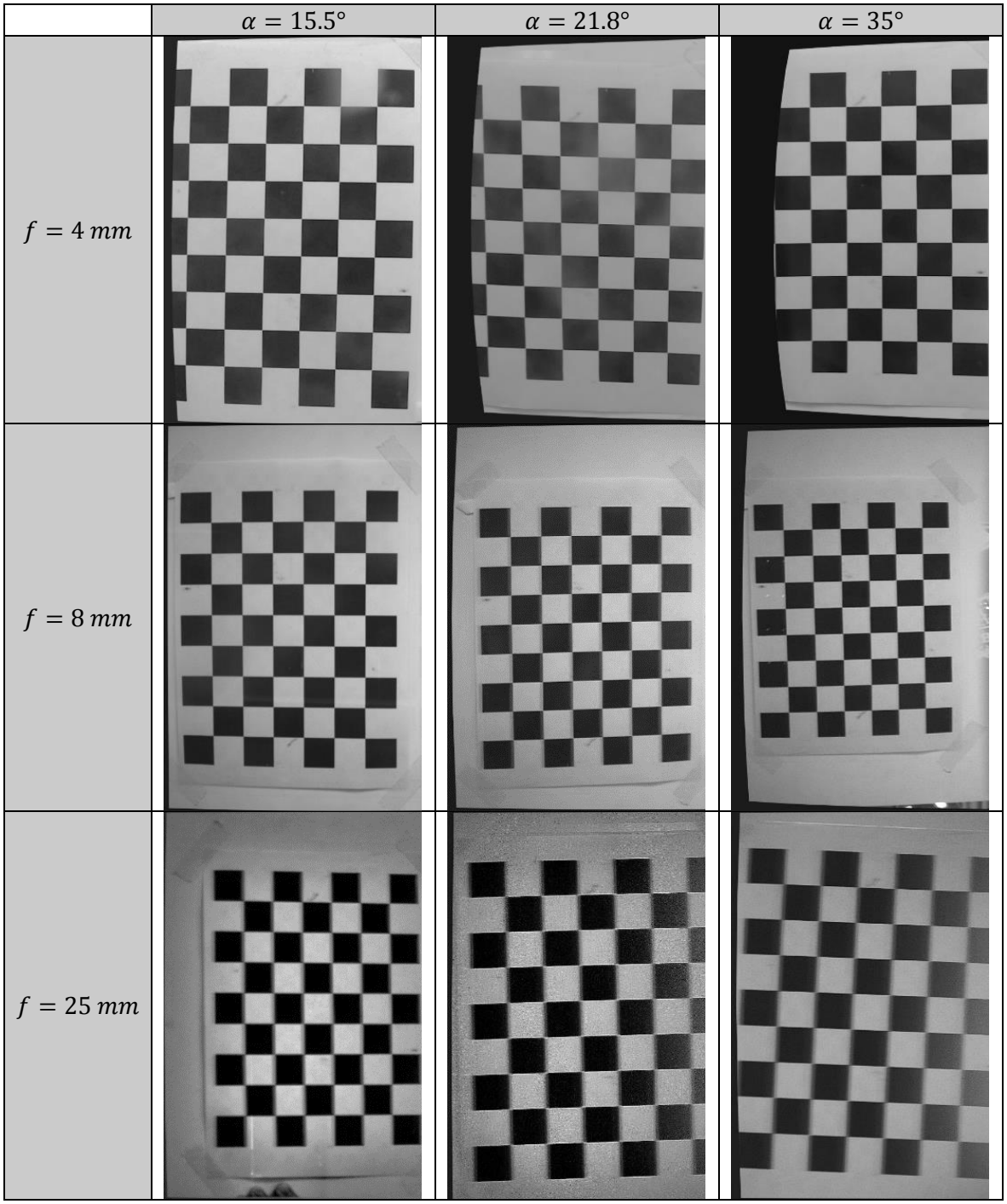


Fig. 3.15 Corrected images from the systems using different camera focal lengths and biprism angles.

3.5.2.1 Results of the Camera Calibration

We shall now define a calibration error (e_d), which is used in this thesis to denote errors in measurement. It is defined as an average in discrepancy between pixel coordinates $p_i(u_i, v_i)$ and $p_i'(u_i', v_i')$, where (u_i, v_i) are the pixel coordinates of the ith sample point measured directly from image planes, and (u_i', v_i') are the coordinates of the ith point re-projected from the checkerboard to the image planes.

$$e_d = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\left(u_i - u_i^{'}\right)^2 + \left(v_i - v_i^{'}\right)^2} \qquad \textit{Eq. 3.21}$$

As the re-projection process will necessarily involve the system intrinsic and extrinsic parameters. Distortion coefficients in the mathematical models representing distortion (e.g. $K1$ and $K2$ in Eq. 2.1) might also be involved when distortion is considered during calibration.

To further validate the proposed distortion correction method, this calibration error is used in the following sections of the thesis. Specifically, for each system setup listed in Table 3.4, 15 images of a checkerboard with 9×7 squares (8×6 corner points) have been acquired at different positions and orientations, as illustrated in Fig. 3.14. In our case, the calibration error can also be used to measure the performance of distortion correction. The calibration errors using the distorted and corrected images are shown in Table 3.5 and Table 3.6, respectively.

Table 3.5 Calibration errors (pixels) using the original distorted images

Focal length	Left virtual camera			Right virtual camera		
	$\alpha = 15.5°$	$\alpha = 21.8°$	$\alpha = 35°$	$\alpha = 15.5°$	$\alpha = 21.8°$	$\alpha = 35°$
$f = 4\ mm$	3.5487	13.4217	19.2245	4.4471	15.5447	19.5654
$f = 8\ mm$	1.6548	6.3474	8.1456	2.4157	5.6587	9.6687
$f = 25\ mm$	1.0354	2.5784	4.2158	2.2246	2.8879	4.6387

Table 3.6 Calibration errors (pixels) using the corrected images

Focal length	Left virtual camera			Right virtual camera		
	$\alpha = 15.5°$	$\alpha = 21.8°$	$\alpha = 35°$	$\alpha = 15.5°$	$\alpha = 21.8°$	$\alpha = 35°$
$f = 4\ mm$	0.5474	0.6224	0.9478	0.5627	0.7547	1.2147
$f = 8\ mm$	0.4574	0.5447	0.7514	0.5474	0.5874	0.8547
$f = 25\ mm$	0.3447	0.5424	0.6357	0.3325	0.5597	0.8547

As can be seen from Table 3.5, the overall mean calibration error is higher than 1 $pixel$, which indicates a poor calibration caused by the distortion in the image. Regarding the amount of distortion for the different system setups, we can observe from the table that the more severe the distortion is, the bigger the calibration error will be. For example, for the image set captured by the system using $f = 4\ mm$ and $\alpha = 35°$ whose amount of distortion is the most severe one in the left and right virtual cameras, respectively, the calibration error is as high as 20 $pixels$, which is much higher than that of other image sets.

However, as shown in Table 3.6, the calibration error decreases significantly after the distortion correction using the proposed method. The overall calibration error is lower than 1 $pixel$, which indicates a good calibration result. We also notice that even for the images with severe distortion ($f = 4\ mm$ and $\alpha = 35°$ for both left and right virtual cameras), the calibration error is reduced from 20 $pixels$ to around 1 $pixel$.

3.5.2.2 Results of the Depth Recovery

In addition, both original and corrected images are used to recover the depth of the sample points. The overall mean depth error of the sample points within the same image is calculated using the difference between the recovered depth and the actual depth. The actual depth data is obtained using a precise laser sensor with accuracy up to 0.001mm. The recovered

depth is calculated using a geometrical based depth recovery method proposed in [83]. This method only requires the real camera and biprism parameters but ignores the distortion caused by the biprism. Since the real camera and biprism parameters can be obtained with high accuracy, either by calibrating the real camera or directly from the technical documents, the error of the recovered depth is mainly affected by the distortion. In other words, the depth recovered by this method could be directly used to verify the distortion correction method without going through the virtual camera calibration. The depth recovery errors of the image data from different system setups are shown in Fig. 3.16.

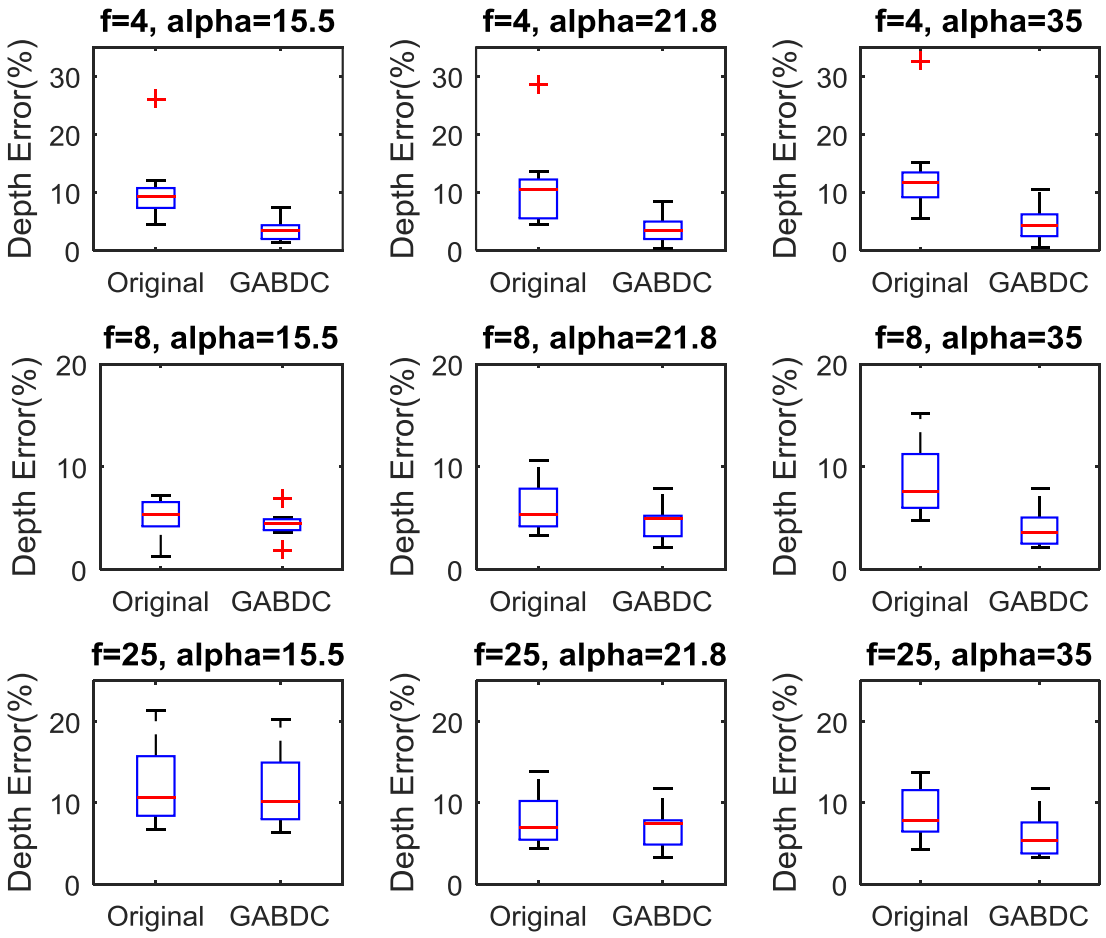


Fig. 3.16 Recovered depth errors of different system setups using the original (distorted) and the corrected images

As can be seen in Fig. 3.16, the overall depth error drops after the correction by GABDC for each system setup, especially for the system with the highest amount of distortion ($f = 4\ mm, \alpha = 35°$), whose average error decreases from the initial 12% to around 5% after the distortion correction. Although some image data with severe distortions lead to quite high depth recovery error (such as the errors with $f = 4\ mm$), the GABDC can handle the distortion properly and improve the depth recovery accuracy significantly. On the other hand, for the systems with larger f but lower α, the improvements on the depth recovery accuracy are not that obvious since the amount of the distortions in these systems is quite minor. Meanwhile, we also notice that the system with the least amount of distortion ($f = 25\ mm, \alpha = 15.5°$) achieves relatively high depth errors. This is because the common FOV in this system is very small and the targets are put far away from the camera, which causes higher error in the depth recovery [37].

3.6 SUMMARY

In this chapter, a geometrical analysis based distortion correction method is proposed, which requires no specific correction pattern (such as the checkerboard) but the full knowledge of the system parameters. This method applies the back-projection method to trace the refraction of the rays in a 3D space. Based on this back-projection analysis of the real (distorted) rays, the assumptions to define the ideal (undistorted) rays are also proposed and verified. Hence, with both real (distorted) and ideal (undistorted) rays, the correlation between the distorted and undistorted image points can be determined. This correlation can be further used to develop the distortion correction maps in the image coordinates, which contain the displacement between the distorted and undistorted points on the image plane in both X and Y axes. These distortion maps can be directly applied to the distorted images,

and hence to remove the distortions as long as the system setup remains unchanged. Meanwhile, this geometrical approach can also be used to analyze the effects of the system parameters to the formation of the distortion. The proposed method proves to be efficient in distortion correction by improving the accuracy of the stereo camera calibration and depth recovery.

In the next chapter, a fully constrained and model-free distortion correction method will be presented. It requires no prior knowledge of the system parameters but employs all the perspective projection invariances and the distortion-free reference points. In this manner, different types of distortions could be corrected.

Chapter 4. FULLY CONSTRAINED AND MODEL-FREE DISTORTION CORRECTION (FC-MFDC) METHOD

As introduced in the last chapter, the BSS system induces particular distortion characteristics in the image. In this chapter, a fully constrained and model-free distortion correction method is proposed. It employs all the projective invariants of a planar checkerboard template as the correction constraints, including straight lines, cross-ratio and convergence at vanishing point, along with the distortion-free reference point as an additional constraint from the system. The method from [78] has proposed a similar idea by using cross-ratio and straight line constraints to correct the distortion. However, as mentioned in § 2.3.2, one of the drawbacks of this method appears to be insufficient restriction on straight lines, since no further constraints for the orientation of these lines are considered. Another drawback is that the undistorted points by this method are highly correlated with the distorted ones, which may result in improper correction for the atypical distortions. As a consequence, the correction result might not be the right solution for our biprism based single-lens stereovision system. In contrast, the FC-MFDC proposed in this chapter overcomes these drawbacks by imposing more constraints to the correction algorithm. Experimental results using both synthetic and real image data show the improvements of the proposed method on the accuracy of camera calibration and depth recovery.

4.1 DISTORTION CORRECTION CONSTRAINTS

In this section, we are seeking a fully constrained and model-free image distortion correction (FC-MFDC) method which requires no specific distortion models. The corrected image by this method will always satisfy the basic projective invariants, including straight lines, vanishing points and cross-ratio constraints. In addition, a principle point is imposed as the distortion correction reference point to further constrain the optimization procedure in the method.

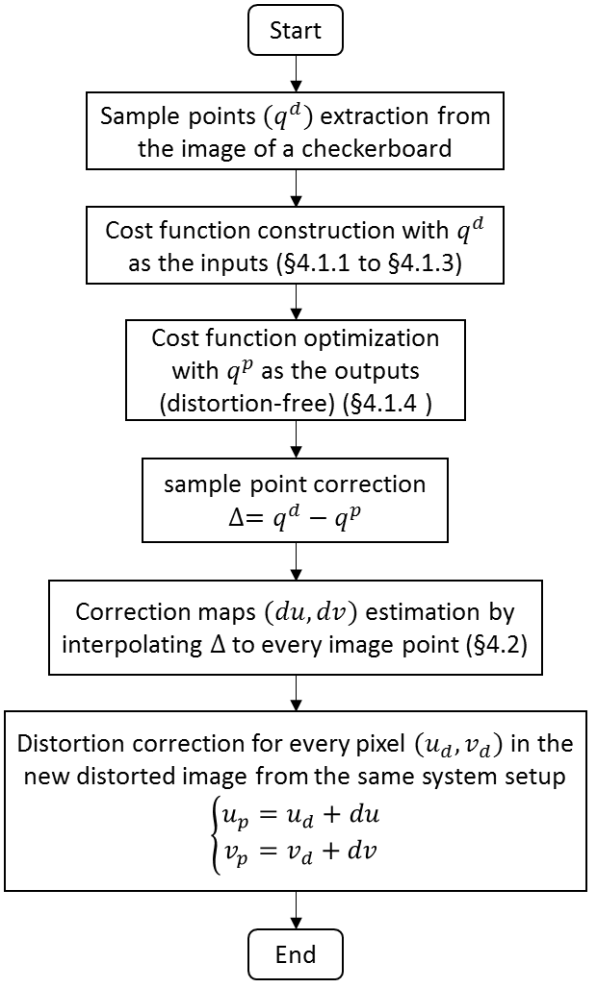


Fig. 4.1 Flow chart of the distortion correction procedure using proposed method

The distortion correction process is summarized in Fig. 4.1. Generally, a planer checkerboard (Fig. 4.2 shows an example) is used as the correction target. The corner points of the squares in the image of this checkerboard are extracted using the method in [84]. They are used as the input sample points q^d for the correction, in which the superscript d indicates the

distorted points. Then, the cost functions are constructed by applying the 3 projection invariances to q^d (introduced in §4.1.1 and §4.1.2), along with the correction reference point as an additional optimization constraint (§4.1.3). Next, these cost functions and constraints are combined to form the total cost function, which is optimized to generate the distortion corrected sample points q^p (§4.1.4). Here the superscript p represents the corrected points. With the distorted and undistorted sample points, the correction maps du (in the X direction) and dv (in the Y direction) for every image point can be estimated by interpolation based on the displacements ($\Delta = q^d - q^p$) of the sample points (§4.2). Finally, every distorted images from the same system setup can be directly corrected by shifting the image points according to the correction maps. Besides, as we know that the distortion in the system is hardware dependent, the distortion correction maps (du, dv) only need to be calculated once as long as the system setup remains unchanged.

4.1.1 Straight Lines Converging at Vanishing Point

Lines representing the 3D parallel lines under the perspective projection will intersect at one point in the image. This point of convergence is called the Vanishing Point (VP). For the planar template of a 'checkerboard' used for calibration (as shown in Fig. 4.2), there are two sets of parallel lines which are orthogonal with each other. Thus, there always exists a VP for each of the vertical and horizontal directions. The detection of the VP is not difficult due to the simple geometry of the checkerboard template. Even with distortion or noise in the extracted points, VP estimation could be done by first detecting the lines within the areas where the distortion or noise level is not severe, such as the areas close to the optical center; then by calculating

the intersections of these lines. Some constraints on the VP detection will be further discussed in §4.1.2.

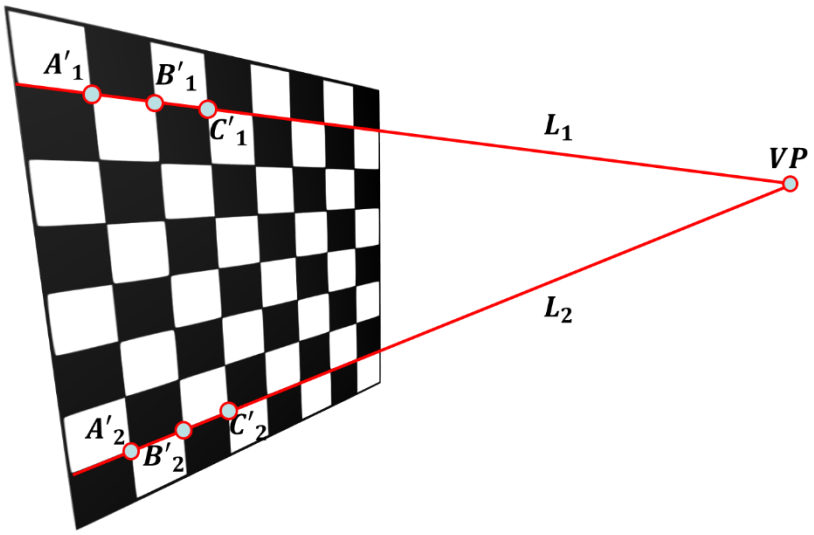


Fig. 4.2 Parallel lines on the checkerboard converge at VP after the projection

According to the definition of VP, all points on a line should be collinear with the corresponding VP. In other words, the points $q_{k,l} = (u_{k,l}, v_{k,l})$ and the VP $q_{vp} = (u_{vp}, v_{vp})$ with respect to (w.r.t.) the image coordinates (u, v) should be on the same line l, in which $k = 1,2, \dots, m$ and $l = 1,2, \dots, n$, where there are n lines with m points on each line in the image. Thus, all these collinear points must fulfill the line equation $au + bv + c = 0$, where a, b and c are line coefficients. Transforming these data points into matrix form, we can get

$$\begin{bmatrix} u_{1,l} & v_{1,l} & 1 \\ u_{2,l} & v_{2,l} & 1 \\ u_{3,l} & v_{3,l} & 1 \\ & \vdots & \\ u_{m,l} & v_{m.l} & 1 \\ u_{vp} & v_{vp} & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \qquad \textit{Eq. 4.1}$$

Eq. 4.1 is in the form of $Ax = 0$, and the Singular Value Decomposation (SVD) could be used to analyse the collinearity of the points:

$$SVD(A) = U \begin{bmatrix} \delta_1 & 0 & 0 \\ 0 & \delta_2 & 0 \\ 0 & 0 & \delta_3 \\ & \vdots & \end{bmatrix} V^T \qquad \textit{Eq. 4.2}$$

where U and V are unitary matrixes, $\delta_i (i = 1,2,3)$ are the non-negative singular values of A. Since all the points of $q_{k,l}$ $(k = 1,2, \dots, m)$ and q_{vp} lie on the same line, the rank of the matrix A should be 2 ($\mathrm{rank}(A) = 2$). So after the SVD, the last singular value of A should be 0 ($\delta_3 = 0$). However, due to the distortion, these points may not be collinear with the VP, or even not with each other, in consequence $\delta_3 > 0$. Hence, we use the following function to represent this last non-negative singular value for the points on line l and the corresponding VP, which is given by:

$$\delta_{3,l} = D\left(q_{1,l}, q_{2,l}, \dots, q_{m,l}, q_{vp}\right) \qquad \textit{Eq. 4.3}$$

Considering this non-negative value as the "cost" of the collinearity, we could form the cost function of the collinearity associated with the points $q_{k,l}$:

$$J\_VP = \sum_{l=1}^{n} D\left(q_{1,l}, q_{2,l}, \dots, q_{m,l}, q_{vp}\right) \qquad \textit{Eq. 4.4}$$

Alternatively, according to the collinearity of points, we can search three points at a time. Thus, the Eq. 4.4 can be re-written as:

$$J_{VP}(q_{k,l}) = \sum_{l=1}^{n} \sum_{k=1}^{m-2} D\left(q_{k,l}, q_{k+1,l}, q_{k+2,l}, q_{vp}\right) \qquad \textit{Eq. 4.5}$$

Here, n is the total line number which includes all the vertical and horizontal lines, and q_{vp} is the corresponding VP coordinates, i.e. q_{vp} is different for different line directions. By

minimizing the cost function $J_{VP}(q_{k,l})$ with distorted points $q_{k,l}^{d}$ as the initial inputs, we could determine the corresponding ideal points $q_{k,l}^{p}$ which fulfill the straight line constraint. Of course, the constraint of collinearity may not be enough to restrict the distortion correction. Other properties should be explored and combined, such as the cross-ratio invariability.

4.1.2 Cross-ratio Invariability

Under projective transformations, there are many geometric dimensions or properties (distances or ratio of distances) which do not remain invariant. But the Cross-ratio, the ratio of ratio of distances, also called 'double ratio', is the simplest property that stays invariant under projective transformations. As illustrated in Fig. 4.3, four points A, B, C and D are lying on the same line and the corresponding image points are denoted as $A', B', C'$and D'. The mutual distances of the target points are represented as AC, AD, BC and BD and the corresponding pixel distances are $A'C', A'D', B'C'$and $B'D'$. The cross-ratios for the four points on the lines L and L_1 are respectively defined as:

$$CR(A,B,C,D) = \frac{AC}{AD} / \frac{BC}{BD} \qquad \textit{Eq. 4.6}$$

$$CR(A',B',C',D') = \frac{A'C'}{A'D'} / \frac{B'C'}{B'D'} \qquad \textit{Eq. 4.7}$$

Assuming that the image coordinates of points $A', B', C'$and D' are $q_1^p = \left(u_1^p, v_1^p\right), q_2^p = \left(u_2^p, v_2^p\right), q_3^p = (u_3^p, v_3^p)$ and $q_4^p = (u_4^p, v_4^p)$, respectively, there exists the following equation based on the cross-ratio invariability.

$$CR(A,B,C,D) = CR(A',B',C',D') = \frac{s_{13} \cdot s_{24}}{s_{14} \cdot s_{23}} \qquad \textit{Eq. 4.8}$$

where s_{ij} represents the distance between the point q_{ip} and q_{jp}, which is defined as ${s_{ij}}^2 = \left(u_i^p - u_j^p\right)^2 + \left(v_i^p - v_j^p\right)^2$. Since the checkerboard template has a regular distribution of all

intersection points, $CR(A,B,C,D)$ is equal to all four collinear neighbor points on the checkerboard; and their corresponding points in the image should have the same property.

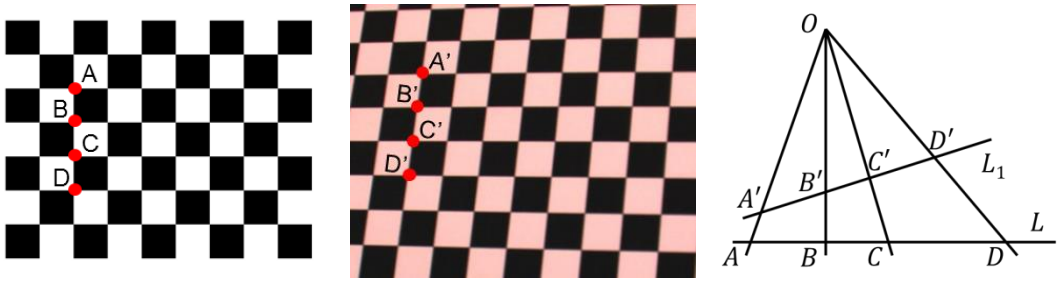


Fig. 4.3 (a) checkerboard template, (b) image of the checkerboard, (c) principle of cross-ratio invariability under perspective projection.

Let us now consider the cross-ratio when one of the four points lies at infinity. If we put $D = \infty$ (Fig. 4.4), then the cross-ratio of the four points will be

$$CR(A,B,C,\infty) = AC/BC \qquad \textit{Eq. 4.9}$$

in which the distances involved D at infinity are eliminated (AD and BD). Meanwhile, the corresponding image point D' is the VP of the line, which is in most cases, finite and fixed. Thus, another equation based on VP is

$$CR(A,B,C,\infty) = {}^{AC}/_{BC} = CR(A',B',C',VP) \qquad \textit{Eq. 4.10}$$

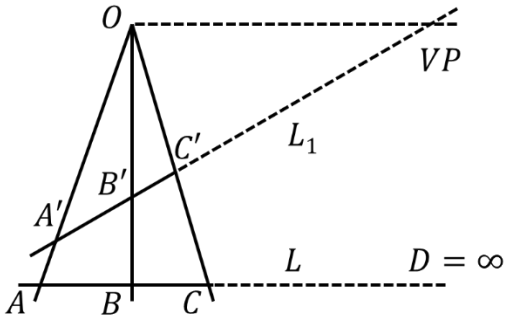


Fig. 4.4 Cross-ratio invariability for the point at infinity

Eq. 4.10 may not hold for all points in the image on account of the distortion. Therefore, another cost function of the cross-ratio invariability could be formed:

$$J_{CR}(q_{k,l}) = \sum_{l=1}^{n}\sum_{k=1}^{m-2}\left(CR(q_{k,l}, q_{k+1,l}, q_{k+2,l}, q_{vp}) - CR(p_1, p_2, p_3, \infty)\right)^2 \qquad \textit{Eq. 4.11}$$

where p_1, p_2 and p_3 are the corresponding 3D points on the checkerboard template, and $CR(p_1, p_2, p_3, \infty)$ is pre-determined when the checkerboard template is designed. There may be some special cases such as when the image plane is parallel to the calibration board, say $L_1 \parallel L$. In this case, the projected parallel lines should still be parallel to each other, and the VP will be at infinity ($q_{vp} = \infty$). Through the definition of the cross ratio, we can verify that the Eq. 4.10 still holds in this case.

Until now, we notice that both straight line and cross-ratio constraints apply the VP in the calculation. However, when the image was distorted, the detection of the VPs would contain some errors due to the distortion. In this case, three constraints are proposed to minimize the influence on the VP estimation:

1) The checkerboard should cover the whole camera FOV;

2) The checkerboard should be placed parallel to the image plane as much as possible;

3) Only the points around the distortion-free reference point (which will be introduced in the next section) are used for the VP detection.

The first constraint is to ensure that the sample points are evenly affected by the distortion. The second one is to make the VP far away from the image points, so that $s_{14} \approx s_{24} \gg s_{13} \approx s_{23}$ if D' is the VP in Eq. 4.8. Hence, the cross-ratio value is close to s_{13}/s_{23} and the errors in straight line and cross-ratio calculations caused by the VP estimation error could be minimized. The third constraint is to directly reduce the error of the chosen points for the VP estimation. Note that these three constraints are only for the training data to better represent and then correct the distortions.

4.1.3 Distortion-free Reference Point

Generally, there exists at least one point that is, or assumed to be, distortion-free in the image plane. This point is normally the principal point or the image center. In other words, the principal point will remain unchanged before and after the correction. The estimation of the principal points on both left and right virtual cameras will be introduced in details in §4.4.1 To incorporate this distortion-free point into the correction procedure, it is assumed that the straightness and cross-ratio cost functions at this point are always minimized to be zero.

$$J_{vp}(q_{c,l}) = \sum_{l=1}^{2} D(q_{c-1,l}, q_{c,l}, q_{c+1,l}, q_{vp}) = 0$$

$$J_{CR}(q_{c,l}) = \sum_{l=1}^{2} \left(CR(q_{c-1,l}, q_{c,l}, q_{c+1,l}, q_{vp}) - CR(p_1, p_2, p_3, \infty) \right)^2 = 0 \qquad \textit{Eq. 4.12}$$

Here $q_{c,l}$ denotes the principal point and J_{VP} and J_{CR} are calculated in both vertical and horizontal directions.

4.1.4 Total Cost Function Formation and Optimization

According to the properties of VP, cross-ratio and distortion correction reference constraints mentioned in the above sections (§4.1.1, §4.1.2 and §4.1.3), we can interpret these restrictions in a simple way: VP constraint is only to correct the distorted points onto the line which is passing through the VP, and cross-ratio constraint is only to correct the ratio of distances between neighboring points. In addition, the principal point as the reference point will globally restrict the positions of the undistorted points. In this manner, the combination of all the constraints will correct both the positions of the distorted points and the

orientations of the lines formed by these points in a global sense. Thus, the total cost function (TCF) could be represented by combining Eq. 4.5, Eq. 4.11 and Eq. 4.12 as shown below:

$$J_{TC}(q_{k,l}) = \sum_{l=1}^{n}\sum_{k=1}^{m-2}\left(\begin{array}{c} D(q_{k,l}, q_{k+1,l}, q_{k+2,l}, q_{vp}) + \\ \left(CR(q_{k,l}, q_{k+1,l}, q_{k+2,l}, q_{vp}) - CR(p_1, p_2, p_3, \infty)\right)^2 \end{array}\right) \quad \textit{Eq. 4.13}$$

$$with\ J_{TC}(q_{c,l}) = 0$$

The ideal points $q_{k,l}^{p}$ fulfilling the straight line, cross-ratio, vanishing point, and correction reference constraints could be calculated by minimizing J_{TC}:

$$q_{k,l}^{p} = \underset{q_{k,l}\in\mathbb{R}^2}{arg\min} J_{TC}\left(q_{k,l}\right) with\ q_{c,l}^{p} = q_{c,l} \quad \textit{Eq. 4.14}$$

This nonlinear optimization problem could be solved by using Levenberg-Marquardt algorithm [81] with the distorted points $q_{k,l}^{d}$ as the initial inputs. The cross-ratio values of the points on the template $CR(p_1, p_2, p_3, \infty)$ are fixed since they are uniformly distributed, and q_{vp} is pre-estimated when the image was captured. In addition, taking into account that the chosen checkerboard template has straight lines in the vertical and horizontal directions, this equation must be true for both the horizontal and vertical lines and so as the corresponding VPs. The principal point $q_{c,l}$ could either be set as the center of the image (captured by conventional cameras) or the estimated principal point (virtual cameras) using the geometrical approach in our single-lens stereovision system [11].

By minimizing the total cost function (Eq. 4.13), we are able to find the undistorted sample points q^{p} corresponding to the distorted sample points q^{d}. However, the undistorted points for the other distorted points in the image remain unknown. Hence, the next step is to estimate these undistorted points and, therefore, to build the distortion correction maps of the given system setup.

4.2 CORRECTION MAPS ESTIMATION

With the known displacements ($\Delta = q^d - q^p$) of the sample points, the correction displacements of all the other image points can be estimated through the interpolation based on Δ. Specifically, since Δ is a 2D vector composed of du (X direction) and dv (Y direction), each of them (du and dv) could be constructed into a 3D map separately, with their corresponding image coordinates u_d and v_d plotted as the X and Y axes and du or dv values in the Z axis, as illustrated in Fig. 4.5 with the star points ($*$). Using the sample points as the reference, the interpolation method can be applied to estimate the du and dv for every other image point (represented by the surfaces in Fig. 4.5). The 3D maps constructed by the interpolated du and dv are named as the correction maps (or correction displacement maps). Thereafter, the distortion correction can be done by shifting every distorted point (u_d, v_d) to the corrected point (u_p, v_p) according to its corresponding (du, dv) on the correction maps:

$$\begin{cases} u_p = u_d - du \\ v_p = v_d - dv \end{cases} \qquad \textit{Eq. 4.15}$$

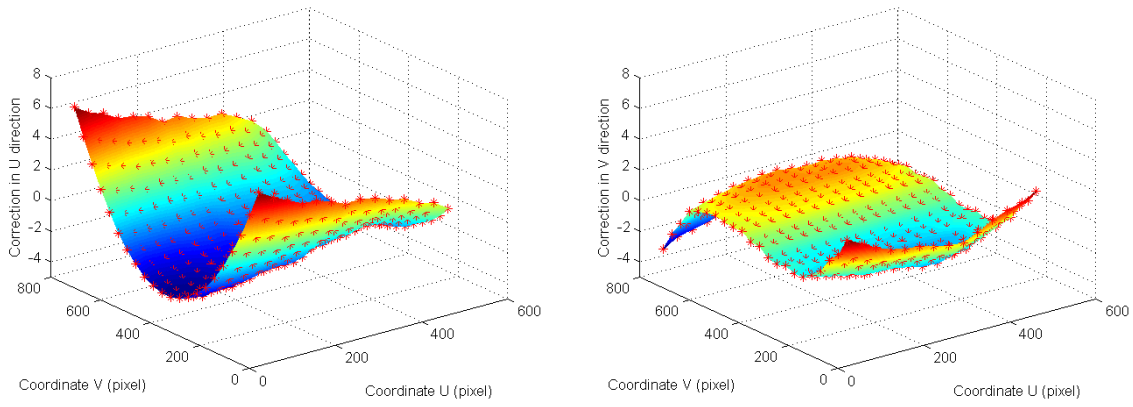


Fig. 4.5 The correction map $(\mathbf{du}, \mathbf{dv})$ between the original distorted points and the undistorted ones in $\mathbf{u}$ and $\mathbf{v}$ directions. Red dots are the sample points and the correction for the rest of the pixels in the image could be interpolated by the surrounding data.

While implementing the interpolation, we can use either parametric function as the global matching, or local transformation [79] which only involves the surrounding data for the local deformation. However, given the fact that the distortion model of our system remains unknown, applying the global matching by a certain polynomial function may not be feasible. Besides, local transformation could achieve higher fitting accuracy when there are some strong local gradients in the maps. Therefore, in our system, the local transformation could be a better choice for the estimation of the distortion maps.

4.3 LIMITATIONS OF THE CORRECTION TECHNIQUE

For the proposed distortion correction technique presented in this chapter, the quality of the sample points $q_{k,l}$ is more important than the quantity. The quality of the sample points is given by the distribution of the feature points defined by the corners of the squares on the checkerboard, such as A', B', C' and D' in Fig. 4.3(b); and quantity refers to the density (or number) of the sample points in the image required for distortion correction. Specifically, the main limitations of the proposed method are presented in the following three aspects:

(1) The checkerboard used for correction should cover the whole image plane as much as possible. The proposed correction procedure depends highly on the sample points from the image. Although these sample points could be precisely corrected, the remaining pixels in the image need to be interpolated by applying some local transformation methods. In this case, correction results are mostly interpolated through the sample points. If these sample points only cover part of the Field of View (FOV), this method may fail to interpolate or recover the remaining part of the image.

(2) For the training data used to generate the correction maps, the checkerboard should be placed properly so that it would be parallel to the image plane and at least one of the sample points must be close to the principle point (usually, it is the center of the image plane). Thus, the VP and the distortion reference point constraints could be applied to the correction procedure properly. Combining with the first limitation, we could find that the position and orientation of the checkerboard were limited in a certain range in order to generate a good correction maps.

(3) According to the TCF (Eq. 4.13), the minimal number of the sample points should be 9 and distributed in the form of a 3×3 matrix. However, taking into account the warping process, this minimum requirement would cause large errors in the final correction results, since the distortions may not distribute uniformly. On the other hand, raising the number of sample points means reducing the size of the square on the checkerboard, and this would introduce inaccuracy during the sample point detection and inefficiency in the nonlinear optimization.

4.4 EXPERIMENTAL RESULTS AND DISCUSSIONS

To evaluate the validity of FC-MFDC method for our BSS system, the method has been tested and compared with another existing model-free correction method on the system. The latter one is the nonlinear lens distortion correction (NLLDC) method presented by Ricolfe *et al*. in [78], which is based on straight line and cross ratio invariance. Both methods apply no typical distortion models, but the projective invariants. Besides, they both need only one checkerboard template as the training data for the correction. The advantage of FC-MFDC over NLLDC is that the former includes two global constraints, VP convergence and distortion-free reference point, which could improve the robustness and accuracy of the distortion correction, especially when dealing with the image distortion of our biprism single-lens

stereovision system presented in this thesis. To evaluate the accuracy of the correction, we develop a set of measurements based on the perspective projection invariance. In addition, the corrected cameras are calibrated using Zhang's method [20] and hence to compare the calibration accuracy and depth recovery error using different distortion correction methods. Both manually generated data and real image data have been tested and evluated in this work.

4.4.1 Synthetic Data

The synthetic image data is generated by determining the data points using the virtual cameras. These virtual cameras are generated based on the simulated BSS system whose parameters are listed in Table 4.1. Note that the biprism placed in front of the camera bisects the camera FOV to generate left and right virtual cameras, thus, each virtual camera has an effective resolution of 512×768 pixels. Here we treat the left and right virtual cameras as independent cameras. These two virtual cameras have their own image coordinate systems.

Table 4.1. System parameters for simulation

Image resolution $(pixels)$	Principal point $(pixels)$	Focal length $(pixels)$	Biprism angle $(degrees)$	Biprism back plane size (mm)	Index of refraction
1024×768	$(512,384)$	$f_x = 1024, f_y = 1024$	25	100×100	1.48

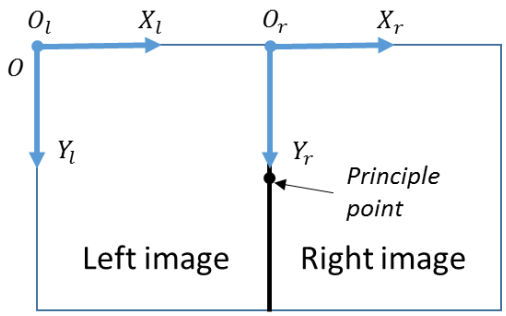


Fig. 4.6 Coordinate systems of the left and right virtual image planes (with $\mathbf{O_l}$ and $\mathbf{O_r}$ as the origins) and the real image plane ($\mathbf{O}$ as the origin)

Fig. 4.6 shows the relationship between the image coordinate systems of the real image plane and the two virtual image planes. The real image plane resolution is $(1024,768)$ with O being the origin. O_l and O_r are the origins of the left and right virtual image planes, respectively. According to the geometrical calibration method in [11], the principle point of the left virtual camera image is at $(512,384)$ w.r.t. O_l. Similarly, that of the right virtual camera image is at $(0,384)$ w.r.t. O_r.

The generated image data used in the experiment are created in the following way. Beginning with an undistorted image of the checkerboard. The distortion is manually generated by imposing a parametric distortion function (radial and/or tangential distortions). Alternatively, the points in the undistorted image are shifted by vectors of different pre-determined magnitudes and directions. For each virtual camera, 20 sets of image points with different angles and positions are generated and distorted according to the calculation in §3.2.1. Within these data sets, 10 data sets are chosen as the training data sets for the distortion correction. These 10 images should be properly captured so that the checkerboard are almost parallel to the image plane and covered the whole FOV as much as possible. In addition, within these sample points in each training image, there should be at least one sample point which is located around the principle point. Each of the training data set will generate their own correction maps and the final correction maps are the average of these 10 correction maps. The rest of the 10 sets of points are used as the testing data and are corrected using the correction maps from the training data described above.

Fig. 4.7 illustrates the distorted and undistorted points of one of the training data sets on the left and right image planes using NLLDC and FC-MFDC. As can be seen, the differences between the undistorted points using NLLDC and FC-MFDC are quite obvious. Overall, FC-

MFDC tries to correct the distorted image towards the principal point, but NLLDC corrects the points closer to the original distorted ones. Specifically, regarding the distortion-free reference point constraint (§4.1.3), FC-MFDC obeys it closely since the distortion correction around the principal point is negligible. In contrast, the NLLDC leads to a correction which produces a distinct shift around the principal point. Besides, comparing the original distorted points and the corrected ones by NLLDC, the points with less amount of distortions are grouped around the middle of the upper and lower halves of the image. This correction result indicates that the principle point is not distortion-free, which is inconsistent with the constraint in §4.1.3. On the contrary, the corrected points with less amount of distortion from FC-MFDC are aligned on the same horizontal line with the principal point. Therefore, we can deduce that the distortion at the principal point using FC-MFDC is also minimized, which indicates that using the sample point close to the principal point could also fulfill the correction constraint in §4.1.3. In addition, another observed difference between the NLLDC and FC-MFDC is the orientation of the corrected lines. Since FC-MFDC imposes the VP constraint, more constraints are applied to the orientations of the lines, and therefore considered more accurate and robust than NLLDC.

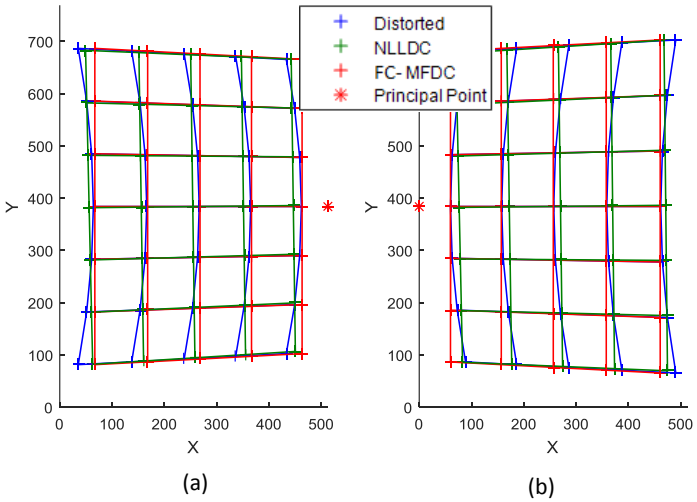


Fig. 4.7 Distortion correction of the training data for both left (a) and right (b) cameras using NLLDC and FC-MFDC

To further evaluate the performance of the correction methods, the corrected testing data sets by NLLDC and FC-MFDC are utilized in the camera calibration. The calibration accuracy of the principal point (u_c, v_c) and re-projection error (e_d) are computed and compared in Table 4.2. Here e_d is the overall mean error of the testing data. Notice that the reference values of the virtual camera parameters in the table are obtained using the geometrical calibration approach in [11], which is insensitive to the image distortion. As can be seen, although both methods lead to good calibration results in general, but FC-MFDC outperforms NLLDC with lower errors. Particularly, regarding the position of the principal point, NLLDC leads to an error of around 30 pixels in the horizontal direction, but FC-MFDC obtains as small as 3 pixels error in average. This advantage mainly benefits from the distortion-free reference point constraint in the FC-MFDC. On the other hand, we notice that the vertical position v_c is calibrated with high accuracy for both methods. This is because the distribution of the training data set covers both upper and lower sides of the principal point, hence the correction in the vertical direction is efficient for both methods. Besides, in regard to the error e_d which could represent the calibration accuracy in general terms, FC-MFDC achieves a lower error than NLLDC for both cameras, which indicates a more accurate calibration by FC-MFDC.

Table 4.2. Calibration results of the testing data sets corrected by NLLDC and FC-MFDC

Parameters/$pixel$	Left camera			Right camera		
	Reference	*NLLDC*	*FC-MFDC*	Reference	*NLLDC*	*FC-MFDC*
u_c	512	482.34	515.25	0	31.38	3.16
v_c	384	392.22	387.16	384	382.05	389.25
f_x	1024	890.45	987.58	1024	902.14	1008.22
f_y	1024	958.54	998.24	1024	965.54	1010.54
e_d	$<$ 1e-4	0.43	0.21	$<$ 1e-4	0.51	0.22

Another evaluation on the distortion correction performance is to measure the error of the recovered depth of the sample points in the testing data. Here, the recovered depth could be calculated using the method in [83] and then compared with the actual depth. The results of the depth errors are shown in Fig. 4.8. As can be seen, the proposed FC-MFDC generally outperforms NLLDC in the depth recovery accuracy with a lower depth error than both the correction results by NLLDC and the original data. Specifically, the median depth error using the corrected results by FC-MFDC is around 4%, which is much better than that by NLLDC (10%) or the original distorted data (12%).

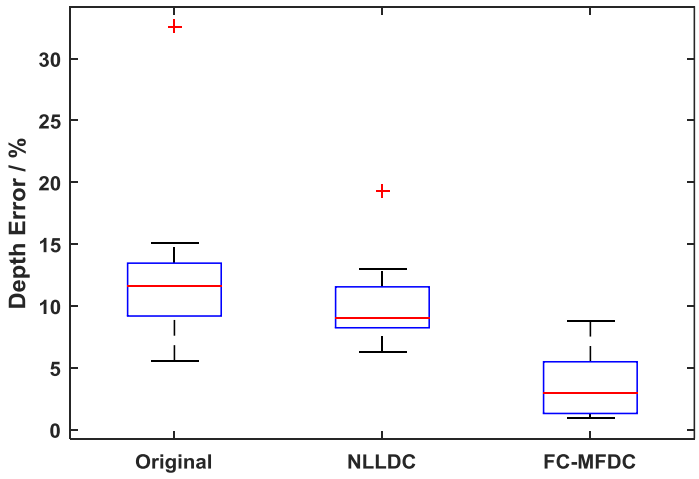
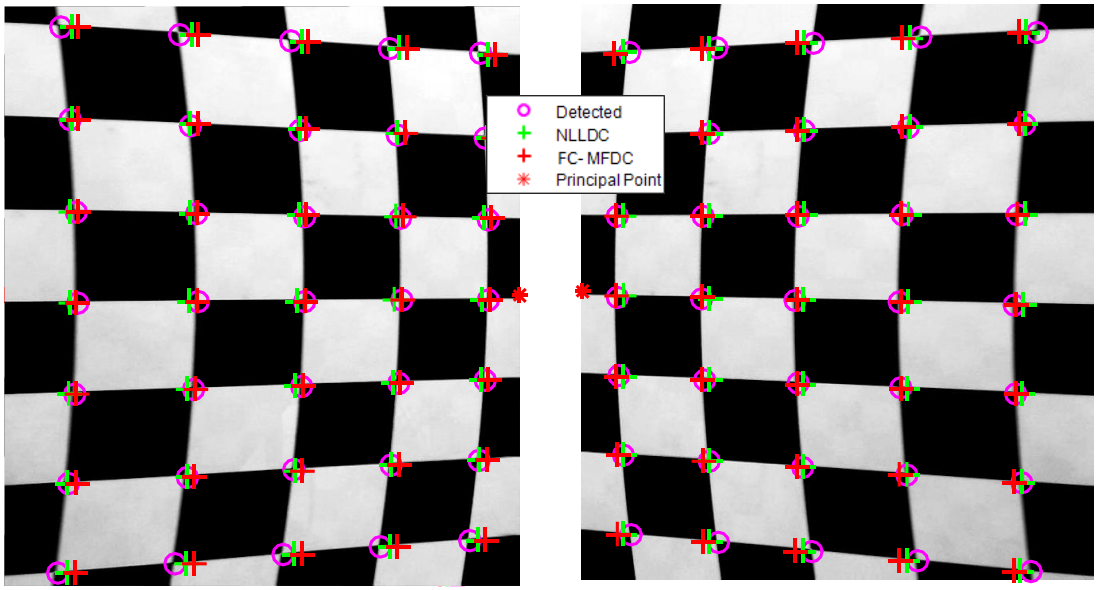


Fig. 4.8 Comparison of the depth error using different distortion correction methods for the synthetic testing data.

4.4.2 Experiments on the Single-lens Stereovision System

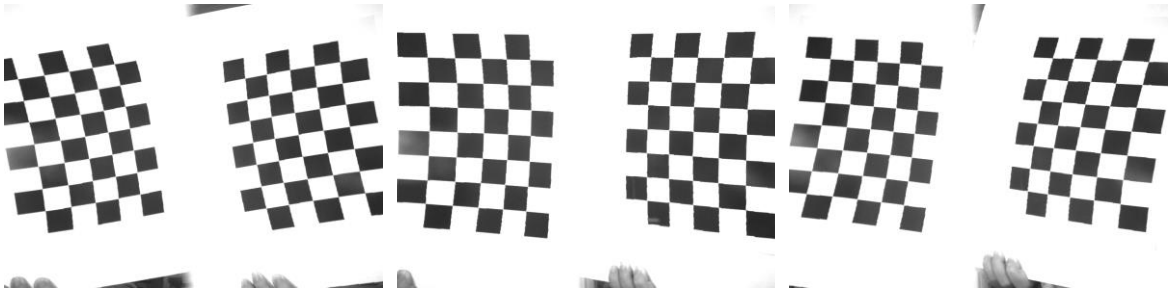
The system is equipped with a CM-140 MCL camera which is designed for Machine Vision systems and is calibrated in advance to obtain the camera parameters and to eliminate the lens distortion [20] as well. The biprism has a corner angle of $21.8\ degrees$, refraction index of 1.48, and thickness of $20\ mm$. The size of the captured image is 1024×768 pixels. The biprism is well aligned with the camera by using the fixtures. Hence, both virtual cameras bisect the real image plane into two halves and form their own image coordinate systems.



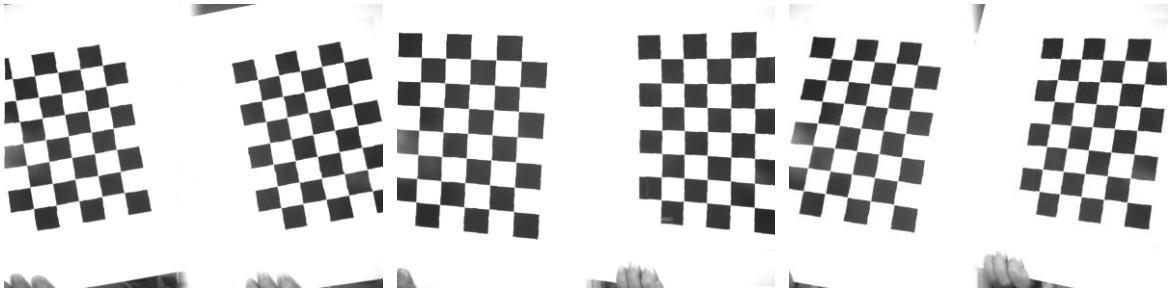
(a) (b)

Fig. 4.9 Training image data extraction (labelled by 'O') and correction (labelled by '+') for both left (a) and right (b) sub-images using NLLDC and FC-MFDC. The principle points are marked by '*'.

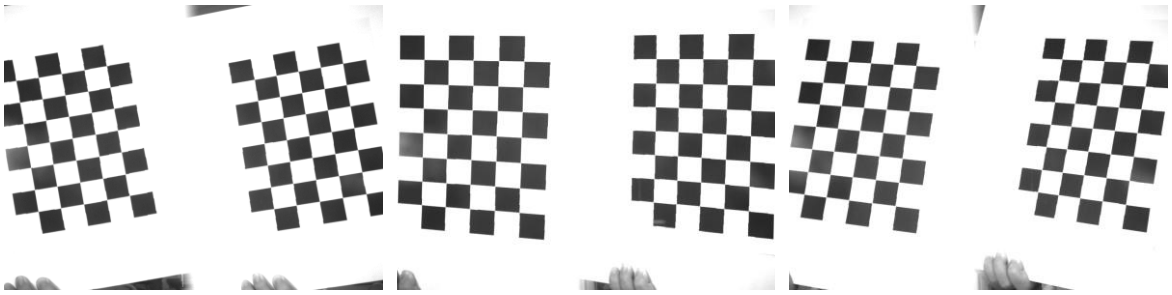
For the correction target, we choose a checkerboard with 5×7 cross points, as shown in Fig. 4.9. 10 sets of training images have been captured with different rotations and distances w.r.t. the camera. Notice that these training images are properly captured following the constraints mentioned in §4.1.2. One example of this training image and its correction results in both left and right cameras are shown in Fig. 4.9. The corner points in the images are first extracted as the correction sample points, and the principal point in each side of the image is determined using the geometrical approach. Then, both NLLDC and FC-MFDC are employed to correct these sample points, and therefore to generate the distortion correction maps, respectively. At last, the testing images are corrected based on these correction maps, as mentioned in §3.3. For further comparison of different correction methods, we use the same testing images in §3.5.2.



(a) Original distorted testing images



(b) Distortion corrected images using NLLDC



(c) Distortion corrected images using FC-MFDC

Fig. 4.10 (a) Testing images, and their corresponding distortion corrected results using(b) NLLDC and (c) FC-MFDC

The correction results of the testing images using NLLDC's and FC-MFDC's correction maps are illustrated and compared in Fig. 4.10, which contains the original distorted testing images (Fig. 4.10 (a)), the corrected images by NLLDC (Fig. 4.10 (b)), and the corrected ones by FC-MFDC (Fig. 4.10 (c)). As can be seen, both left and right sub-images are significantly corrected by both NLLDC and FC-MFDC, compared with the original distorted images. From the corrected images, it may be ambiguous to define which correction method is better for the distortion correction for this system. Hence, further evaluation on the calibration results using the undistorted testing data sets is implemented and compared in Table 4.3. Note that the

'Reference' data in this table is obtained by calibrating the real CCD camera without placing the biprism in front of it.

Table 4.3. Calibration results of the undistorted testing images corrected by NLLDC and FC-MFDC

Parameters/$pixel$	Left camera			Right camera		
	Reference	*NLLDC*	*FC-MFDC*	Reference	*NLLDC*	*FC-MFDC*
u_c	516.46	501.61	515.07	6.89	25.15	9.56
v_c	388.39	392.41	386.26	388.51	384.87	387.02
f_x	1005.47	881.42	987.58	1011.54	942.69	1025.08
f_y	998.89	961.34	995.44	1024	975.49	1011.57
e_d	0.22	0.61	0.48	0.22	0.79	0.42

Within these calibration parameters, the reference data is obtained from the real camera, which has been calibrated in advance. In particular, the principal point of the real camera is at $(519.51,387.49)$, which is slightly shifted w.r.t. the image center which is at $(512,384)$. Hence, the principal points of the left and right cameras are shifted according to the position of the real camera's principle point. From these calibration results, we can see that the calibrated principal points from FC-MFDC are closer to the reference with only about 2 pixels error in average, while the ones from NLLDC are as high as 17 pixels error. Besides, although both distortion correction methods do not yield better re-projection error than the real camera, the accuracy of both cameras are still below 1 pixel which is considered as good correction results. Moreover, FC-MFDC leads to a better accuracy over the NLLDC in the calibration.

At last, the distortion corrected testing images by NLLDC and FC-MFDC are used to recover the depth of the sample points, and their recovered depth errors are shown in Fig. 4.11. As can be seen, both NLLDC and FC-MFDC can improve the depth recovery accuracy by reducing the distortions in the image for the system setups with different biprisms. However, FC-MFDC performs better than NLLDC with a lower average error, especially for the system with severe distortions, such as the systems with $f = 4\,mm$. This is because with the distortion-free

reference point constraint, the positions of the corrected points by FC-MFDC are less affected by the original positions of the distorted points comparing with NLLDC. Although NLLDC can still correct the distortions by straightening the lines and correcting the cross-ratio values, the corrected image shifts away from their ideal undistorted position. Therefore, while using these corrected results for the depth recovery, the error arises. On the other hand, with the additional constraints of VP and distortion-free reference point, FC-MFDC achieves a better distortion correction results and hence improves the depth recovery accuracy.

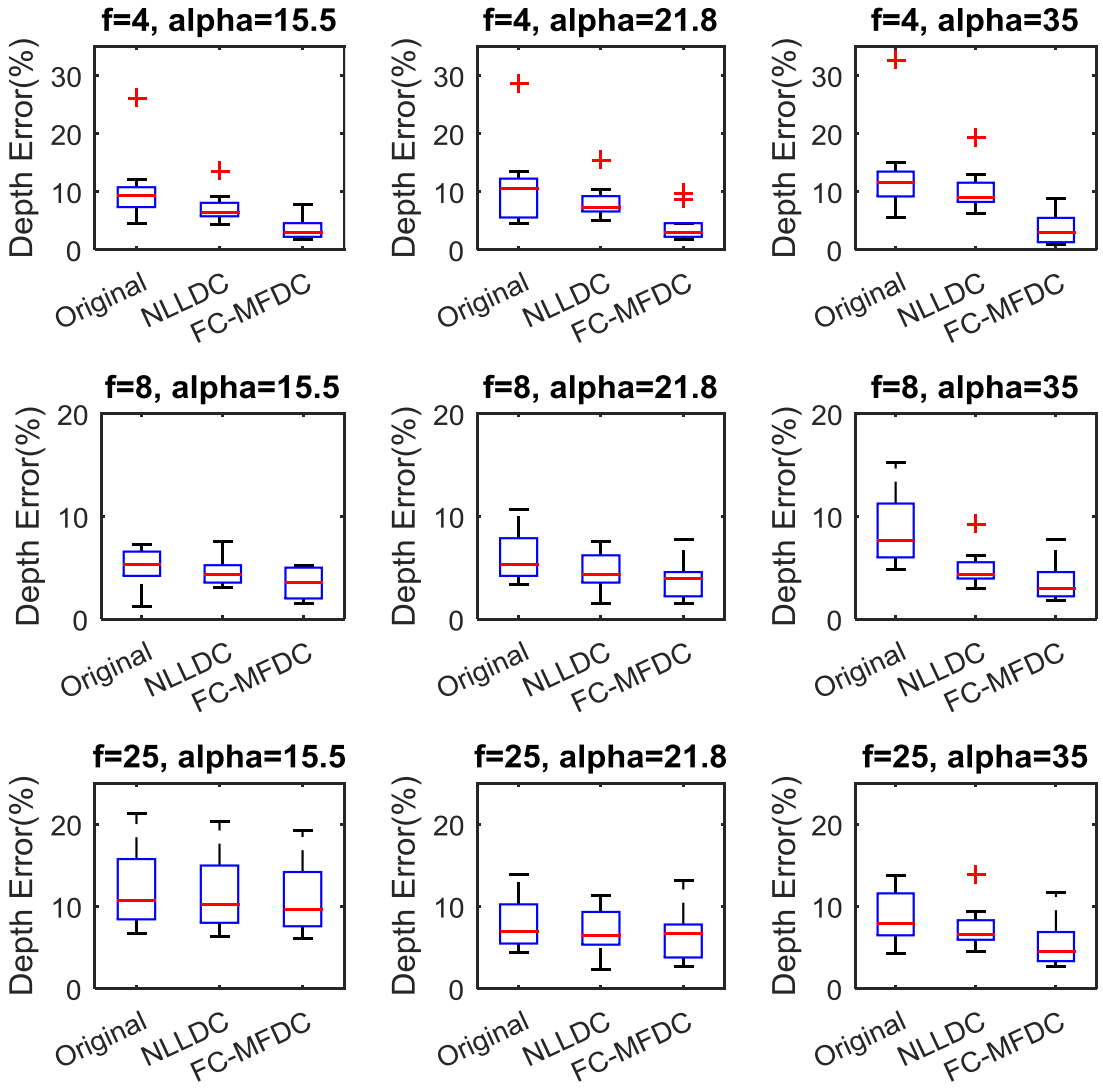


Fig. 4.11 Comparison of the depth error using different distortion correction methods for the real testing data from different system setups.

4.5 SUMMARY

In this chapter, we proposed a fully constrained and model-free distortion correction method for the single-lens stereo vision system using a biprism. In view of the distortion induced by the thick biprism in front of the camera, typical distortion calibration or correction methods are not effective or applicable. The proposed method requires no a-priori knowledge of the camera parameters, but employs all the projective constraints including the cross-ratio, straight line, and convergence to the vanishing points on the planar checkerboard, together with the distortion-free reference point. By minimizing the total cost resulting from all the constraints, ideal undistorted points in the distorted images can be retrieved. Therefore, the correction maps for the distorted images are estimated. Thus, all the subsequent images can be corrected by these correction maps, as long as the system setup remains unchanged. Compared with another model-free correction method, the proposed method makes considerable improvements in distortion correction accuracy, and hence achieves a better calibration of the stereo cameras and a better depth recovery accuracy.

In the next chapter, a model-based distortion correction method will be presented. This distortion model is developed in two different ways, and could be applied to the existing camera calibration methods for parameter determination.

Chapter 5. BIPRISM DISTORTION MODELING AND CORRECTION (BDMC) METHOD

The radial and tangential lens distortions could be modeled by some parametric functions based on the properties of these distortions, which have been reviewed in §2.2.1. Generally, these models were developed by analyzing the refraction deviations by the lens. These deviations are caused either by the manufactory error by the lens or the misalignment between the image sensor and the lens. The analysis is to mathematically explain the causes of these deviations and hence to model the distortions in general forms. These distortion models or patterns could be applied to the camera calibration procedure to enhance the accuracy of camera parameter determination.

Inspired by the development of these lens distortion models, the biprism distortion model is built and evaluated in this chapter. However, different from the lens distortion modeling which is based on a simplified pin-hole camera model, the modeling for the biprism distortions in our system is more complex. This is because although the virtual cameras could be modeled as the pin-hole cameras, the projection procedures for both the left and right virtual cameras are more complicated than those in a simplified pin-hole camera. Specifically, the determination of the real (distorted) and ideal (undistorted) projections rely on the complex calculations proposed in Chapter 4. In the biprism distortion model development, we proposed two ways to build the distortion model: one of them is to employ a general bivariate polynomial function which is only associated with the image coordinates. In this way, the complicated 3D analysis of the projection deviation could be avoided. This general polynomial could be further refined by the distortion properties. The other way of modelling

is to fully express the distortion formation based on the projection steps through the biprism in 3D space. Therefore, the distortions could be expressed by algebraic functions, which could then be further expanded into polynomial functions using Taylor series. Generally, these two ways of modelling should lead to the same model since they are used to describe the same distortion phenomenal. This final distortion model is called biprism distortion model (BD-Model) in this thesis. Lastly, this model could be integrated into the camera calibration procedure for parametric determination and then to enhance the calibration accuracy of the virtual cameras.

The parametric determination for the distortion model is usually done together with the camera calibration (such as Zhang's method [20]). This camera calibration method should be suitable for the calibration of our unique biprism based single-lens stereovision system. Meanwhile, it should also be able to handle the distortion model which might be different from the typical ones such as radial and/or tangential distortions. Therefore, before carrying out the distortion model development, we would like to give a brief introduction of the calibration techniques.

5.1 Introduction of Virtual Camera Calibration Techniques

Camera calibration is defined as the formulation of the projection equations which link the known coordinates of a set of 3D points and their projections on the camera image plane w.r.t. the image coordinates. This relationship is described by the intrinsic parameters (such as the camera focal length, principle point, and lens distortion parameters) and the extrinsic parameters (such as the relative position and orientation of the cameras w.r.t. the pre-defined world coordinate system). A simple calibration process was proposed by Tasi [85] to determine the extrinsic and intrinsic parameters which correlate the 3D world coordinates

with the image plane pixel coordinates as shown in Fig. 5.1. A more detailed illustration of the relationship between the coordinate systems is shown in Fig. 5.2.

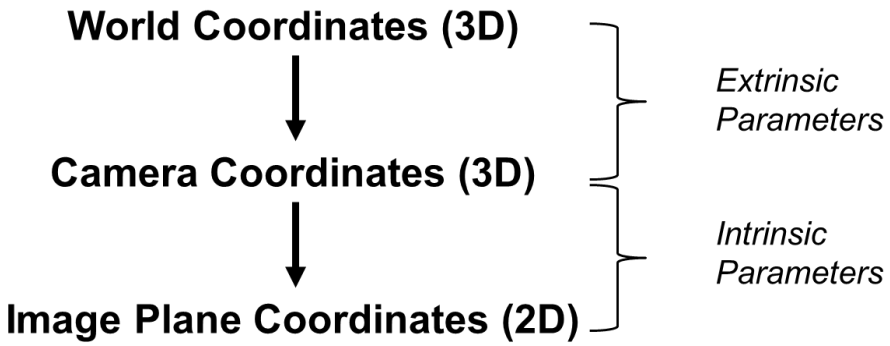


Fig. 5.1 Transformation of 3-D world coordinates to camera image plane coordinates

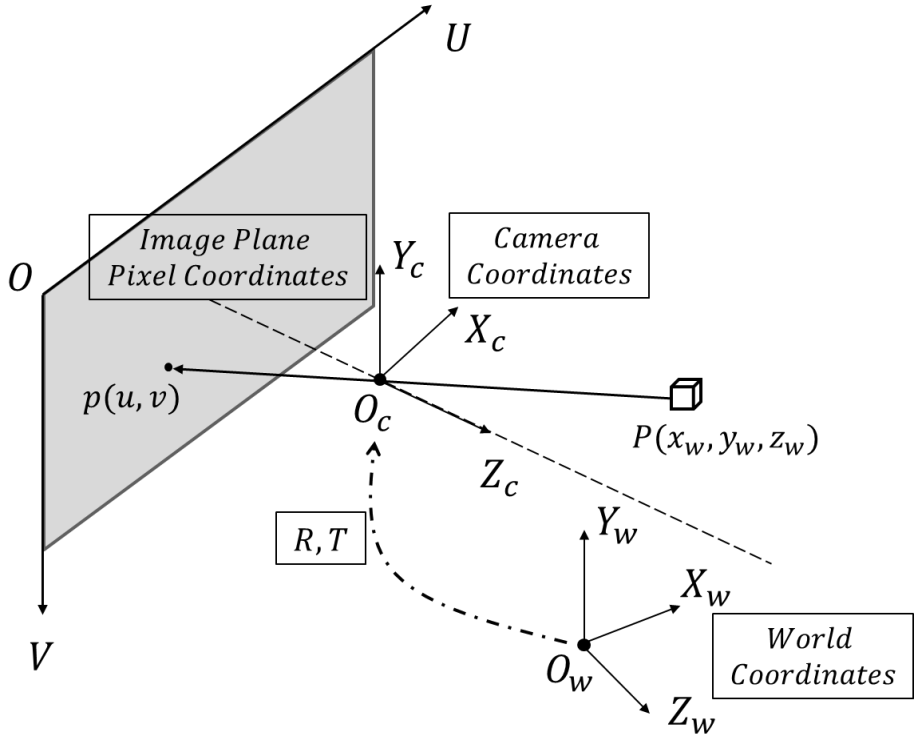


Fig. 5.2 Illustrations of the coordinate systems involved in the camera calibration

As shown in Fig. 5.2, the projection of the 3D target point P on the image plane coordinates could be expressed as

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} \boldsymbol{R} & \boldsymbol{T} \\ \boldsymbol{0} & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \qquad \textit{Eq. 5.1}$$

The coordinate systems in the calibration process are defined as follows:

(X_w, Y_w, Z_w): World coordinate system

(X_c, Y_c, Z_c): Camera coordinate system

(U, V): Image plane coordinate system

The world coordinate system is arbitrarily assigned. The camera coordinates have their origin located at the camera optical center O_c with its Z_c axis perpendicular to the image plane. Here, R and T are the rotational and translational matrices from the world coordinates to the camera coordinates, respectively. $\boldsymbol{A}$ denotes the intrinsic matrix which includes the focal length, principle point offset, axis skew and distortion [20].

The objective of the calibration process is to recover the parameters in both intrinsic and extrinsic matrices. With the known parameters, the coordinates of a world point $P(x_w, y_w, z_w)$ in the 3D world coordinate system could be recovered according to its corresponding points $p(u, v)$ on the image plane coordinates. A literature review on the calibration techniques for the Biprism based single-lens stereovision system will be presented in the following sections.

5.1.1 Conventional Camera Calibration Methods

The accuracy of camera calibration to recover the intrinsic and extrinsic parameters will directly affect the performance of a stereovision system, since almost all the stereovision applications (such as depth recovery, motion detection, 3D reconstruction, etc.) utilize these parameters. Therefore, a great deal of effort has been spent to enhance the calibration

accuracy. Based on the techniques used, camera calibration methods can be classified into 3 categories:

1. *Linear transformation methods.* In this category, the objective is to linearize the relationship between the intrinsic and extrinsic parameters [86, 87]. Therefore, the parameters can be recovered by solving these linear equations using the least squares method. The advantage is the simplicity of the camera model which promotes a simple and rapid calibration. One drawback is the linear techniques are useless for lens distortion modeling, entailing a low accuracy of the system.

2. *Nonlinear optimization methods.* A calibration method becomes nonlinear when any kind of lens distortion is included in the camera model. In this case, the intrinsic and extrinsic parameters are usually obtained through iteration with the constraint of minimizing an objective function. This function is usually formed by the distance between the imaged points and the modeled projections obtained by iteration. The advantage of these iterating techniques is that almost any model with lens distortion can be calibrated and accuracy usually increases by increasing the number of iterations up to convergence. However, they generally require a good initial guess in order to guarantee the convergence. These methods are only used in the classical calibration techniques [88, 89].

3. *Hybrid methods.* These methods make use of the advantages of those in the two previous categories. Generally, they are comprised of two steps: the first step involves solving most of the camera parameters in linear equations; the second step is to iteratively compute the remaining parameters. These methods are able to solve different camera models with different lens distortion models. Therefore, they are widely studied and used in recent works [41, 90, 91].

Comparing with the above methods, Zhang [20] proposed a more flexible calibration technique which only requires the camera to observe a planar pattern which is captured at different orientations. It allows either the camera or the planar pattern to be freely moved without knowing the motions, which is easier to use comparing with other methods. Of course, this method is also suitable for the camera calibration with different distortion models.

5.1.2 Virtual Camera Calibration Technique Using Geometrical Approach

For the prism-based single-lens stereovision system, Lim and Xiao [11, 12] proposed a novel technique to calibrate the virtual cameras generated by the system. In their works, they proved that their proposed calibration technique outperforms the classical methods for this particular single-lens stereovision system. A brief description of this calibration technique will be given in the following sections.

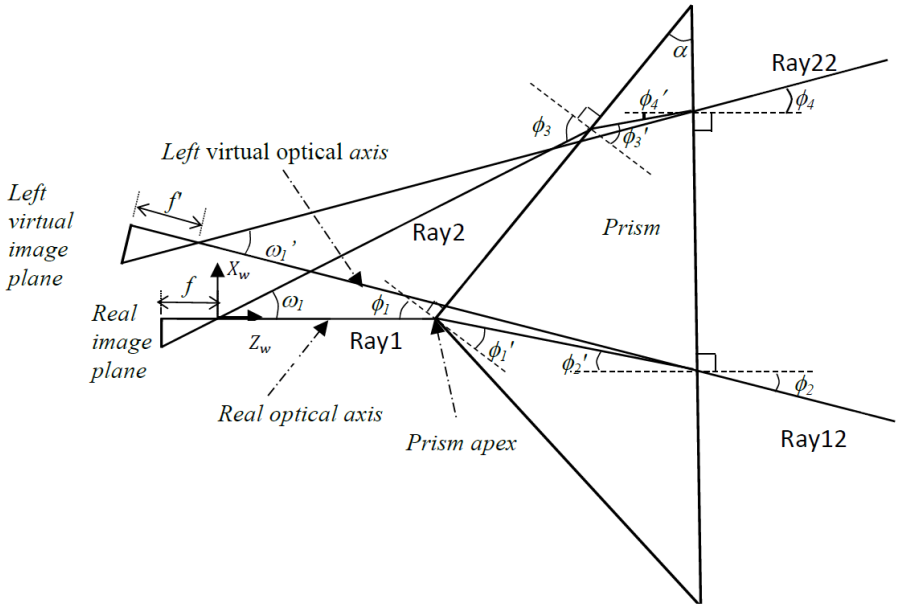


Fig. 5.3 Formation of virtual cameras using biprism and the geometrical relationship between the real and virtual image plane.

As shown in Fig. 5.3, the formation of the left virtual camera is achieved by the two particular rays in the system: one is Ray1 which is formed by the image center point and the optical center; the other one is Ray2 which starts from the point at the image boundary and passes

through the optical center. The relationship between these two rays and the camera FOV has been introduced in §2.1.4. Here, it emphasizes that the virtual camera parameters, namely the intrinsic parameters, could be calculated using the geometrical relationship between the real camera and the biprism. Moreover, all the intermediate parameters, such as the incident and refracted angles $\emptyset_i$ and $\emptyset_i', i = 1,2,3,4$, can be determined. In their methods, they assume that all the rays and points are in 2D ($X_w Z_w$) plane.

This technique is simple since all the geometrical and mathematical derivations are carried out in 2D space. However, there are two main limitations using this method. The first one is that it requires the full knowledge of the system parameters, including the real camera parameters. The second one is that it only analyzes the 2D refraction on the $X_w Z_w$ plane, and no distortion in the Y_w direction is taken into consideration. Therefore, the distortion in this system will result in lower accuracy during the calibration process, even if the system setup up was close to ideal condition.

Although this geometrical analysis based virtual camera calibration method could not be directly applied to the distortion correction, the idea of the geometrical analysis might help to identify the formation of the distortion in the system, and hence to build the distortion model with mathematical expressions. This model could then be applied to the camera calibration techniques to enhance the calibration accuracy.

In the next sections, we propose two different methods to develop a biprism distortion model for the distortion induced by the biprism in the system. One is an empirical way by refining a general bivariate polynomial function according to the distortion properties. The other one is to derive the model from the deviation of the distorted and undistorted projections using a

geometrical analysis. Different from the estimation of the correction maps using nonlinear optimization in Chapter 4, all the derivations here are algebraic-based.

5.2 BIPRISM DISTORTION MODEL DEVELOPMENT

5.2.1 Empirical Biprism Distortion Model

A distortion modeling process could be interpreted as finding a parametric function which could fit the distortion maps. Therefore, with the known distortion maps from Chapter 4, the distortion modeling problem becomes a curve fitting problem. Initially, we assume that both of the distortion data in the X and Y directions can be fitted by the functions f_x and f_y with x_{ud} and y_{ud} as the variables:

$$\begin{cases} dx = f_x(x_{ud}, y_{ud}) \\ dy = f_y(x_{ud}, y_{ud}) \end{cases} \quad \textit{Eq. 5.2}$$

Without loss of generality, a bivariate polynomial function with constant coefficients is employed to build the distortion models, which could be expressed in the following forms:

$$\begin{cases} f_x(x_{ud}, y_{ud}) = \sum_{0 \le i+j \le M_x} p_{ij}^x x_{ud}^i y_{ud}^j \\ f_y(x_{ud}, y_{ud}) = \sum_{0 \le i+j \le M_y} p_{ij}^y x_{ud}^i y_{ud}^j \end{cases} \quad \textit{Eq. 5.3}$$

Here M_x and M_y are the degree of polynomial for dx and dy, respectively; p^x and p^y are the fitting coefficients, also known as distortion coefficients, for dx and dy, respectively. The subscripts/superscripts x and y refer to the x and y axes. The subscripts i and j for p^x and p^y are the indices, and the superscripts i and j for x_{ud} and y_{ud} are the power. Both i and j are non-negative integers. Different from the distortion models from Brown [43, 44], p^x and p^y may not be the same or correlated.

Note that all the above calculation is w.r.t. the camera coordinates. Therefore, the general form of the distortion model with reference to the image coordinates (u, v) could be written as

$$\begin{cases} f_u(u_{ud}, v_{ud}) = \sum_{0 \le i+j \le M_x} p_{ij}^u u_{ud}^i v_{ud}^j \\ f_v(u_{ud}, v_{ud}) = \sum_{0 \le i+j \le M_y} p_{ij}^v u_{ud}^i v_{ud}^j \end{cases} \quad \textit{Eq. 5.4}$$

where f_u and f_v are the distortion functions of the center shifted distorted points $(u_{ud}, v_{ud}) = (u, v) - (u_0, v_0)$ in u and v axes, respectively. Here (u_0, v_0) is the principal point in pixel unit, and p^u and p^v are the corresponding distortion coefficients.

However, it may not be feasible to apply the model in Eq. 5.4 to the calibration procedure directly. Some refinements need to be carried out to extract the essential terms, and hence to improve the efficiency of the model. These refinement conditions could be explored from the properties of the biprism distortion formation.

The first property is the vertical-distortion-free line. It was mentioned in §3.1 (and shown in *Fig. 3.1*) that the vertical distortion induced by the biprism depends on the incident angle in the vertical direction. In other words, the projections along the $y = 0$ (or $v = v_0$) line are vertical-distortion-free, since all the projections passing that line are projected without vertical components. Moreover, it is also distortion-free at the principal point, hence, the constant term in $f_u(u_{ud}, v_{ud})$ are abandoned. Therefore, the terms in the distortion model $f_v(u_{ud}, v_{ud})$ could be refined as:

$$\begin{cases} f_u(u_{ud}, v_{ud}) = \sum_{1 \le i+j \le M_x} p_{ij}^u u_{ud}^i v_{ud}^j \\ f_v(u_{ud}, v_{ud}) = \sum_{0 \le i+j \le M_x - 1} p_{ij}^v u_{ud}^i v_{ud}^{j+1} \end{cases} \quad \textit{Eq. 5.5}$$

in which all the terms without v_{ud} are abandoned.

The second property is the distortion symmetry. Based on the analysis, the distortion is symmetrical w.r.t. the distortion-free line $y = 0$ ($v = v_0$). Specifically, for the horizontal distortion function $f_u(u_{ud}, v_{ud})$, the odd-order terms of v_{ud} could be abandoned. In contrast, the even-order terms of v_{ud} could be omitted in the vertical distortion function $f_v(u_{ud}, v_{ud})$. Therefore, the distortion functions could be further refined as:

$$\begin{cases} f_u(u_{ud}, v_{ud}) = \sum\limits_{\substack{1 \le i+j \le M_x \\ j=2n}} p_{ij}^u u_{ud}^i v_{ud}^j \\ f_v(u_{ud}, v_{ud}) = \sum\limits_{\substack{0 \le i+j \le M_x-1 \\ j=2n}} p_{ij}^v u_{ud}^i v_{ud}^{j+1} \end{cases} \quad \textit{Eq. 5.6}$$

where $n = 0,1,2, \dots, m$.

So far, the redundant terms have been removed according to the geometrical properties of the biprism distortion. We name this distortion model as the "Empirical Biprism Distortion Model (EBD-Model)" which is developed from a general bivariate polynomial function and only involves the distortions induced by the biprism.

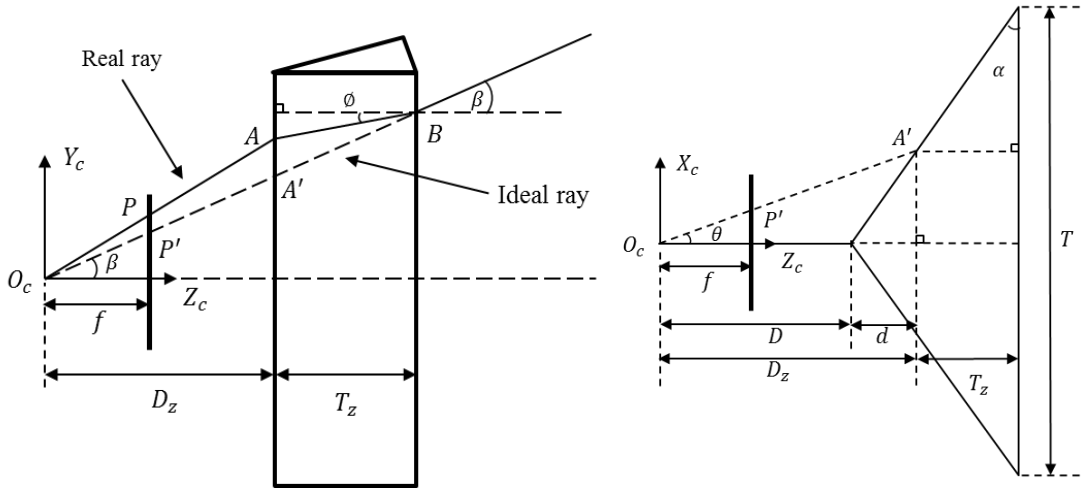
5.2.2 Theoretical Biprism Distortion Model

Instead of building the distortion model using empirical polynomial functions, a more theoretical development on the formation of the distoriton is proposed in this section. Generally, this development is based on a geometrical analysis of the projections of the rays, which is similar with the one presented in Chapter 3. The difference here is that we are trying to derive algebraic expressions representing the projection deviations between the ideal rays and distorted rays, instead of using the nonlinear optimization methods. In this way, the mathematical expressions of the distortions w.r.t. the image plane coordinates could be

determined. Of course, the main purpose is not to fully determine the distortions maps, but to construct the distortion model similar to Eq. 5.6.

As we know that the projections and the distortions in the X and Y directions are quite different because of the unique structure of the biprism. Therefore, they are studied along the X and Y axes seperately.

5.2.2.1 Distortion in the $\mathbf{Y}$ axis



(a) Side view (b) Top view

Fig. 5.4 Distortion analysis in the Y direction from (a) side view, and its corresponding top view (b)

We first start with the analysis along the Y axis whose distortoin formation is simpler than the one along the X direction. A side view of the real (distorted) and ideal (undistorted) projections of an incident ray is presented in $Fig.5.4$ (a). In this figure, β and $\emptyset$ are the incident and refracted angles w.r.t. the Z_c axis in the Y_cZ_c plane at the intersection point B, respectively. A and A' are the intersections of the real and ideal rays at the front plane, respectively, and $P(x_d, y_d)$ and $P'(x_{ud}, y_{ud})$ are the corresponding real and ideal image points. Note that ideal and real image points refer to the undistorted and distorted points, respectively. Some system parameters are also labelled here, that include the camera focal

length f, distance D_z from the camera origin to intersection A' in the Z direction, and T_z as the distance from the ray entering point to the back plane of the biprism in the Z direction. Note that D_z and T_z are not the same as D and T when the ideal ray has a projection angle θ w.r.t. the Z_c axis in the X_cZ_c plane, as illustrated in Fig. 5.4 (b). This difference is named as d:

$$d = D_z - D = \frac{T}{2} \cdot tan(\alpha) - T_z \qquad \textit{Eq. 5.7}$$

According to the geometrical analysis given in Appendix A, the distortion in the Y axis $dy = y_d - y_{ud}$ could be expressed with the system parameters and the point coordinates:

$$dy = \arctan\left(\frac{y_{ud}}{f}\right) \cdot \left(1 - \frac{1}{n}\right) \cdot \left(\frac{T}{2}\tan(\alpha) - \frac{x_{ud} D \tan(\alpha)}{f - x_{ud}\tan(\alpha)}\right) \cdot \frac{f - x_{ud}\tan(\alpha)}{D} \qquad \textit{Eq. 5.8}$$

From the perspective of distortion modelling and calibration, the above expression of dy would be too complicated to be useful (such as the camera calibration using the distortion model). A polynomial function like Eq. 5.6 would be easier to apply. Therefore, some transformation to the terms involving with x_{ud} and y_{ud} needs to be performed.

First, the $arctan$ function could be expanded using the Taylor series:

$$\arctan\left(\frac{y_{ud}}{f}\right) = \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1}\left(\frac{y_{ud}}{f}\right)^{2i+1} \qquad \textit{Eq. 5.9}$$

with the condition of $\left|\frac{y_{ud}}{f}\right| \leq 1$. This condition could be satisfied when the vertical FOV of the camera is less than 90°, which is feasible for most of the cameras. To simplify the expressions, all the constant values or the coefficients of the terms associating with y_{ud} and x_{ud} are represented by C_i^k in the following sections. Here the superscript k denotes the index of different Taylor expansions, and the subscript i is the index of the coefficients in the current expansion. Hence, Eq. 5.9 can be simplified as

$$\beta = \sum_{i=0}^{\infty} C_i^1 \cdot {y_{ud}}^{2i+1} \qquad \textit{Eq. 5.10}$$

Second, the term $\frac{x_{ud} D \tan\alpha}{f - x_{ud}\tan\alpha}$ with x_{ud} in the denominator posses inconvenience in the distortion modeling. Hence, it is rewritten using the Taylor series:

$$\frac{x_{ud} D \tan\alpha}{f - x_{ud}\tan\alpha} = \frac{x_{ud} D \tan\alpha}{f} \cdot \frac{1}{1 - \frac{\tan\alpha}{f} x_{ud}} = \frac{x_{ud} D \tan\alpha}{f} \cdot \sum_{i=0}^{\infty} \left(\frac{\tan\alpha}{f} x_{ud}\right)^i \qquad \textit{Eq. 5.11}$$

The condition for convergence of this Taylor expansion is $\left|\frac{\tan\alpha}{f} x_{ud}\right| < 1$, which is also generally feasible for most of the system setups. This expansion can be simplified as

$$\frac{x_{ud} D \tan\alpha}{f - x_{ud}\tan\alpha} = \sum_{i=0}^{\infty} C_i^2 \cdot {x_{ud}}^{i+1} \qquad \textit{Eq. 5.12}$$

With Eq. 5.12, the terms associated with x_{ud} in Eq. 5.8 could be further combined and represented by:

$$\left(1 - \frac{1}{n}\right) \cdot \left(\frac{T}{2}\tan(\alpha) - \frac{x_{ud} D \tan(\alpha)}{f - x_{ud}\tan(\alpha)}\right) \cdot \frac{f - x_{ud}\tan(\alpha)}{D} = \left(\sum_{i=0}^{\infty} C_i^3 \cdot {x_{ud}}^i + C^4\right) \qquad \textit{Eq. 5.13}$$

Substitute Eq. 5.10 and Eq. 5.13 to Eq. 5.8, we can obtain the final theoretical distortion model in the Y direction:

$$dy = \sum_{i=0}^{\infty} C_i^1 \cdot {y_{ud}}^{2i+1} \cdot \left(\sum_{i=0}^{\infty} C_i^3 \cdot {x_{ud}}^i + C^4\right) \; for \; \left|\frac{y_{ud}}{f}\right| \le 1 \; and \left|\frac{\tan\alpha}{f} x_{ud}\right| < 1 \qquad \textit{Eq. 5.14}$$

5.2.2.2 Distortion in the X axis

Similar to the geometrical analysis of the distortion in the Y direction, the distortion due to refraction in the X direction is illustrated in Fig. 5.5, which is a more detailed version of Fig. 5.4 (b). Here γ is the projection angle of the FRs. It is a constant value with the given biprism; θ is the ideal (undistorted) ray's projection angle w.r.t. the FRs in the $X_c Z_c$ plane.

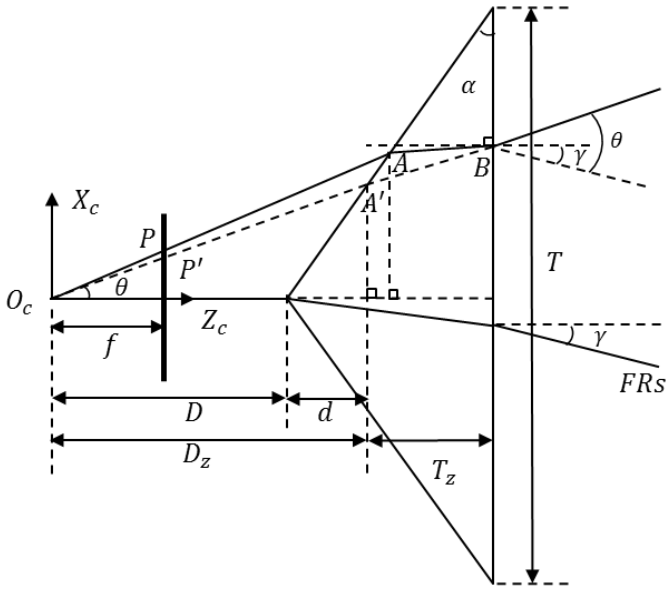


Fig. 5.5 Distortion analysis in the X direction from the top view

The final theoretical distortion model in the X direction could be expressed by (Appendix B)

$$dx = \frac{f\Delta X}{D_z} = \left(1 - \frac{1}{n}\right) \cdot \left(\arctan\left(\frac{x_{ud}}{f}\right) - \gamma\right) \cdot \left(\frac{T}{2}\tan(\alpha) - \frac{x_{ud} D \tan(\alpha)}{f - x_{ud}\tan(\alpha)}\right) \cdot \frac{f - x_{ud}\tan(\alpha)}{D} \cdot \sqrt{1 + \left(\frac{y_{ud}}{f}\right)^2} \quad \textit{Eq. 5.15}$$

Comparing dy in Eq. 5.8 and dx in Eq. 5.15, we note that there are common terms associated with x_{ud}. Thence, substituting Eq. 5.13 to Eq. 5.15, the dx expression can be simplified as:

$$dx = \left(\arctan\left(\frac{x_{ud}}{f}\right) - \gamma\right) \cdot \left(\sum_{i=0}^{\infty} C_i^3 \cdot {x_{ud}}^i + C^4\right) \cdot \sqrt{1 + \left(\frac{y_{ud}}{f}\right)^2} \quad for \left|\frac{\tan\alpha}{f} x_{ud}\right| < 1 \quad \textit{Eq. 5.16}$$

Here, we shall use Taylor expansions to represent $\arctan\left(\frac{x_{ud}}{f}\right)$ and $\sqrt{1 + \left(\frac{y_{ud}}{f}\right)^2}$:

$$\arctan\left(\frac{x_{ud}}{f}\right) = \sum_{i=0}^{\infty} C_i^1 \cdot {x_{ud}}^{2i+1} \; for \; \left|\frac{x_{ud}}{f}\right| \le 1 \quad \textit{Eq. 5.17}$$

$$\sqrt{1+\left(\frac{y_{ud}}{f}\right)^2}=\sum_{i=0}^{\infty}C_i^5\cdot {y_{ud}}^{2i}\ \ for\ \left|\frac{\tan\alpha}{f}y_{ud}\right|<1 \qquad \textit{Eq. 5.18}$$

Substitute Eq. 5.17 and Eq. 5.18 to Eq. 5.16, we can obtain the final theoretical biprism distortion model in the X direction:

$$\begin{aligned} dx &= \left(\sum_{i=0}^{\infty}C_i^1\cdot {x_{ud}}^{2i+1}-\gamma\right)\cdot\left(\sum_{i=0}^{\infty}C_i^3\cdot {x_{ud}}^{i}+C^4\right)\cdot\sum_{i=0}^{\infty}C_i^5\cdot {y_{ud}}^{2i} \\ &= \left(\sum_{i=0}^{\infty}C_i^6\cdot {x_{ud}}^{i}+C^7\right)\cdot\sum_{i=0}^{\infty}C_i^5\cdot {y_{ud}}^{2i} \end{aligned} \qquad \textit{Eq. 5.19}$$

with the constraints of $\left|\frac{\tan\alpha}{f}x_{ud}\right|<1$, $\left|\frac{x_{ud}}{f}\right|\le 1$ and $\left|\frac{\tan\alpha}{f}y_{ud}\right|<1$. Here we need to point out that γ is a constant value for the given system as mentioned above in this section.

Until now, both dx and dy can be fully expressed in image plane coordinates. Combining Eq. 5.14 and Eq. 5.19, the full form of the biprism distortion model could be represented by

$$\begin{cases} dx=\left(\sum_{i=0}^{\infty}C_i^6\cdot {x_{ud}}^{i}+C^7\right)\cdot\sum_{i=0}^{\infty}C_i^5\cdot {y_{ud}}^{2i} \\ dy=\sum_{i=0}^{\infty}C_i^1\cdot {y_{ud}}^{2i+1}\cdot\left(\sum_{i=0}^{\infty}C_i^3\cdot {x_{ud}}^{i}+C^4\right) \end{cases} \qquad \textit{Eq. 5.20}$$
$$for\ \left|\frac{\tan\alpha}{f}x_{ud}\right|<1,\left|\frac{x_{ud}}{f}\right|\le 1,\left|\frac{y_{ud}}{f}\right|\le 1,and\ \left|\frac{\tan\alpha}{f}y_{ud}\right|<1$$

or

$$\begin{cases} f_u(u_{ud},v_{ud})=\left(\sum_{i=0}^{\infty}C_i^6\cdot {u_{ud}}^{i}+C^7\right)\cdot\sum_{i=0}^{\infty}C_i^5\cdot {v_{ud}}^{2i} \\ f_v(u_{ud},v_{ud})=\sum_{i=0}^{\infty}C_i^1\cdot {u_{ud}}^{2i+1}\cdot\left(\sum_{i=0}^{\infty}C_i^3\cdot {v_{ud}}^{i}+C^4\right) \end{cases} \qquad \textit{Eq. 5.21}$$
$$for\ \left|\frac{\tan\alpha}{f}x_{ud}\right|<1,\left|\frac{x_{ud}}{f}\right|\le 1,\left|\frac{y_{ud}}{f}\right|\le 1,and\ \left|\frac{\tan\alpha}{f}y_{ud}\right|<1$$

in the image coordinates (u,v) in pixels.

Practically, these conditions in the model can be generally fulfilled by the system whose real camera's horizontal and vertical FOVs are under $90^{\circ}(\left|\frac{x_{ud}}{f}\right| \leq 1, \left|\frac{y_{ud}}{f}\right| \leq 1)$, and biprism's angle is less than 45° $(\tan\alpha < 1)$. This fully expanded distortion model is named as "Theoretical Biprism Distortion Model (TBD-Model)".

While comparing the EBD-Model in Eq. 5.6 and the TBD-Model in Eq. 5.21, we can observe that that they are surprisingly similar if we only consider the terms associated with u_{ud} and v_{ud}. The next step is to find the essential terms in the EBD-Model or TBD-Model, which can represent the distortions more effectively than other terms in the model.

From the perspective of camera calibration, terms in the distortion models with high order are generally not considered, in order to avoid overfitting the noise. Referring to the examples in Fig. 3.7, cubic polynomial function (degree of 3) for our distortion model might be sufficient based on the empirical data from experiments. Therefore, the expansion of the refined model could be:

$$\begin{cases} f_u(u_{ud}, v_{ud}) = p_{10}^u u_{ud} + p_{20}^u u_{ud}^2 + p_{02}^u v_{ud}^2 + p_{12}^u u_{ud} v_{ud}^2 + p_{30}^u u_{ud}^3 \\ f_v(u_{ud}, v_{ud}) = p_{01}^v v_{ud} + p_{11}^v u_{ud} v_{ud} + p_{21}^v u_{ud}^2 v_{ud} + p_{03}^v v_{ud}^3 \end{cases} \quad \textit{Eq. 5.22}$$

This simplified distortion model in Eq. 5.22 is named as "Biprism Distortion Model (BD-Model)". Further verification on this BD-Model will be carried out in §5.4.

5.3 VIRTUAL CAMERA CALIBRATION WITH BD-MODEL

As introduced in §5.1, several camera calibration techniques could be applied to determine the camera parameters. Given the unique biprism distortion model developed in §5.2, a hybrid calibration method (such as Zhang's method [20]) could be a proper way to calibrate the virtual cameras with the biprism distortions in our system.

According to Zhang's calibration method, the BD-Model coefficients, along with other projection parameters, could be first estimated using the linear transformation methods (see Appendix D). Then, they could be refined through maximum likelihood inference using a nonlinear optimization method. Specifically, with the given n 2D-3D correspondences $m_i = (u_i, v_i) \leftrightarrow M_i = (X_i, Y_i, Z_i)$, the maximum likelihood estimation can be obtained by minimizing the distances between the image points and their predicted positions on the image plane:

$$\min_{\boldsymbol{A},\boldsymbol{R},\boldsymbol{t},\boldsymbol{P}} \sum_i \|m_i - \emptyset(\boldsymbol{A}, \boldsymbol{R}, \boldsymbol{T}, \boldsymbol{P}, M_i)\|^2 \qquad \textit{Eq. 5.23}$$

where $\emptyset(\boldsymbol{A}, \boldsymbol{R}, \boldsymbol{t}, \boldsymbol{P}, M_i)$ is the projection of M_i onto the image according to Eq. 5.1, in addition to the biprism distortion model whose parameters are $\boldsymbol{P}$.

This is a nonlinear minimization problem, which can be solved with the Levenberg-Marquardt [81]. It requires an initial guess of $\boldsymbol{A}, \boldsymbol{R}$ and $\boldsymbol{T}$ which can be obtained using the linear transformation described in Appendix D. The distortion parameter $\boldsymbol{P}$ can be initialized with $\boldsymbol{0}$. Note that only the BD-Model is included in above minimization procedure, because we assume that the typical distortions (radial and tangential distortions) are corrected in advance for the real camera.

5.4 Experimental Results and Discussions

In this section, both synthetic/simulated and real image data under different system setups are carried out to test the proposed BD-Model represented by Eq. 5.22. In the experiments with simulated data, we first concentrate on the BD-Model evaluation by fitting it to the simulated distortion data in the following manner. These distortion data only consists of the biprism distortion data, and no other types of distortions (such as radial and/or tangential

distortions). In the fitting procedure, different combinations of the BD-Model based on its polynomial order are evaluated and refined according to the goodness of fit on different distortion data. Secondly, with the refined BD-Model, camera calibration is implemented to test this model's accuracy and robustness by evaluating the calibration accuracy and efficiency. In the actual experiments, since the actual distortion maps are normally not available, the refined BD-Model is tested together with the camera calibration. Moreover, a comparison between the conventional distortion model (radial distortion) and the refined BD-Model are also accomplished by measuring the camera calibration accuracy using different biprisms.

5.4.1 Biprism Distortion Model Evaluation with the Simulated Data

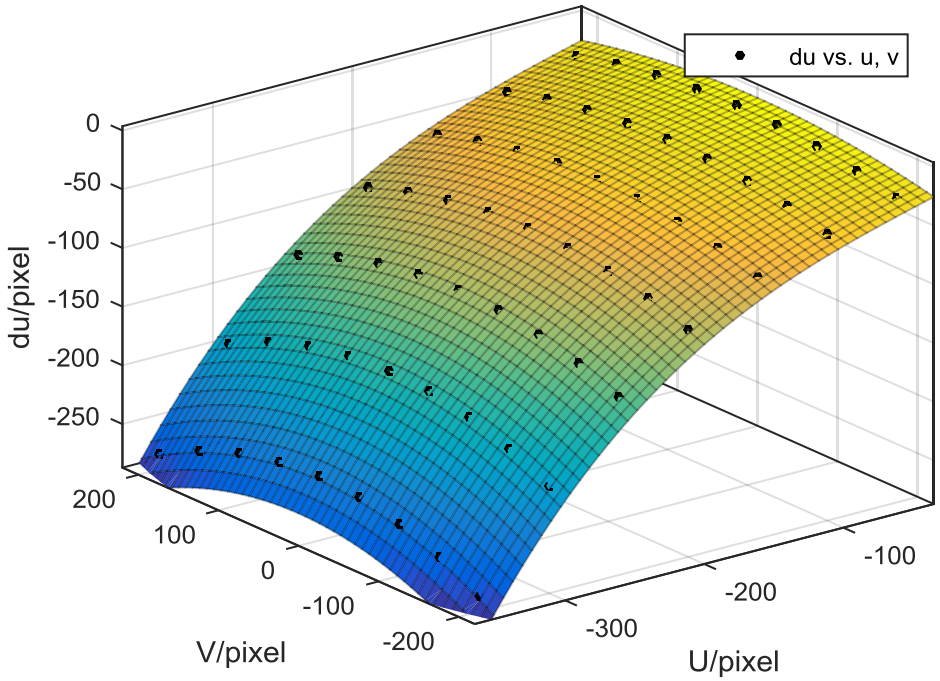
In this experiment with simulated data, two kinds of data sets are generated. One of them uses back-projection to obtain the distorted image points with the given undistorted ones, to get the pure distortion data for model fitting evaluation. The other one applies the natural projection (see §3.2) to achieve the distorted projection data of a checkerboard for the calibration. The simulated system is built with the setup parameters in Table 5.1. The biprism angle is set as 15 and 40 $degrees$ to generate soft (thin biprism with small prism angle) and hard (thick biprism with large prism angle) distortions, respectively.

Table 5.1. System parameters for simulation

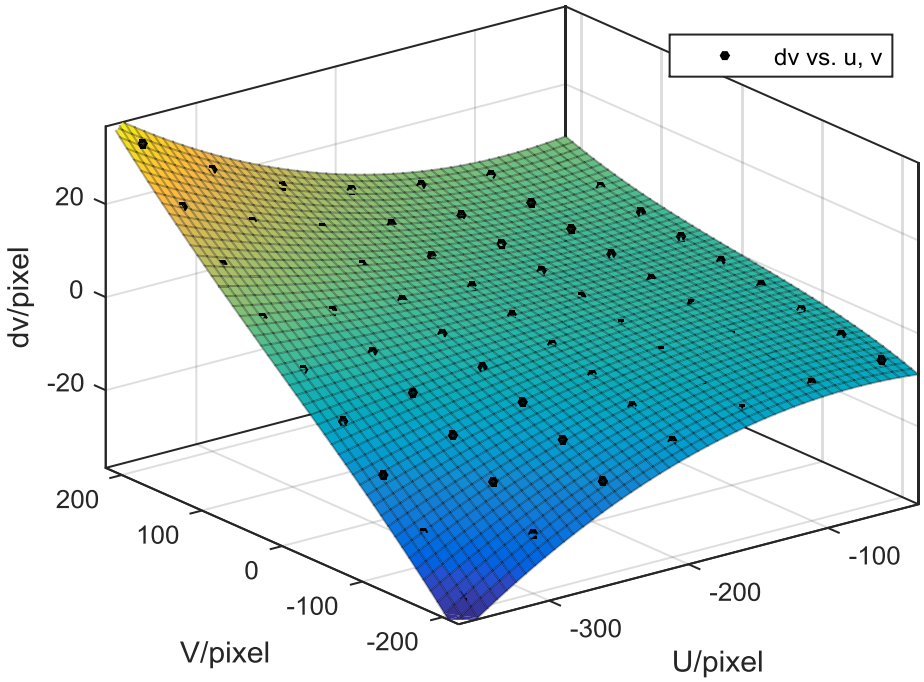
Image resolution	Principal point	Focal length	Distance from camera to biprism	Biprism back plane size	Index of refraction
1024×768	$(512,384)$	$f_x = 1024, f_y = 1024$	$35\ mm$	$100 \times 100\ mm$	1.48

5.4.1.1 Distortion Model Fitting

In this experiment, the undistorted image points $p^{ud}(u_{ud}, v_{ud})$ on each virtual image plane are first created with a dimension of 9×7 squares (8×6 corner points) and a distance of $50\ pixels$ between neighbour points. Then, the corresponding distorted points $p^{d}(u_d, v_d)$ are determined by the geometrical method proposed in §3.2. Thereafter, with both distorted and undistorted points, the distortion data sets represented by $(du, dv) = (u_{ud}, v_{ud}) - (u_d, v_d)$ can be calculated and then fitted to the BD-Model. For model fitting, du and dv are separately constructed into a 3D map with u_{ud} and v_{ud} in the X and Y axes, and du or dv on the Z axis, as shown in Fig. 5.6. This figure illustrates an example of the fit performance of the left image distortion data sets at $\alpha = 40\ \mathrm{dregee}s$ using BD-Model. Note that the U and V axes are the X and Y directions in the image coordinates, respectively, and are center shifted with respect to the principal point. In addition to verify the essential terms for different levels of distortions, combinations with different order terms in BD-Model are tested and measured according to their fit qualities.



(a)



(b)

Fig. 5.6 Distortion displacement maps in the X (a) and Y (b) directions ($\alpha = 40$ dregees). Dots are the image points and fitted using BD-Model.

To evaluate the goodness of fit, or fit quality, the degrees of freedom adjusted R-Square (Adjusted R-Square) and the root mean squared error (RMSE) are utilized.

$$Adjusted\ R-Square = 1-\frac{\sum_{i=1}^{n-1}(y_i-\hat{y}_i)}{\sum_{i=1}^{n-m}(y_i-\bar{y})}$$

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)}{n-m}} \qquad \textit{Eq. 5.24}$$

where y_i is the observation data; $\hat{y}_i$ is the response value from the fit and $\bar{y}$ is the mean of the observation data. n is the number of data sets, and m is the number of fitted coefficients. The adjusted R-square statistic can take on any value less than or equal to 1, with a value closer to 1 indicating a better fit. A *RMSE* value closer to 0 indicates a fit that is more useful for prediction. Both of the measurements can be obtained from the MATLAB Curve Fitting Toolbox.

The statistics of the fit using different combinations of the terms in BD-Model are shown in Table 5.2 and Table 5.3, respectively for du and dv. These combinations (named as Combo in Table 5.2 and Table 5.3) are ordered according to the degree of the polynomials. Here we need to point out that $f_u(u_{ud}, v_{ud})$ and $f_v(u_{ud}, v_{ud})$ are independent since their coefficients are not correlated.

Table 5.2. Fit quality of soft and hard distortions with different fitting terms in $\mathrm{f_u(u_{ud}, v_{ud})}$

Combo	$f_u(u_{ud}, v_{ud})$	Soft Distortion		Hard Distortion	
		$Adjusted\ R-Square$	$RMSE$	$Adjusted\ R-Square$	$RMSE$
1	u_{ud}, v_{ud}^2	0.890	3.563	0.832	32.53
2	$u_{ud}, u_{ud}^2, v_{ud}^2$	0.999	0.359	0.994	6.606
3	$u_{ud}, u_{ud}^2, v_{ud}^2, u_{ud}v_{ud}^2$	0.999	0.349	0.994	6.043
4	$u_{ud}, u_{ud}^2, v_{ud}^2, u_{ud}^3$	0.999	0.230	0.998	3.284
5	$u_{ud}, u_{ud}^2, v_{ud}^2, u_{ud}v_{ud}^2, u_{ud}^3$	1.000	0.047	0.999	0.762

Table 5.3. Fit quality of soft and hard distortions with different fitting terms in $f_v(u_{ud}, v_{ud})$

Combo	$f_v(u_{ud}, v_{ud})$	Soft Distortion		Hard Distortion	
		$Adjusted\ R-Square$	$RMSE$	$Adjusted\ R-Square$	$RMSE$
1	$v_{ud}, u_{ud}v_{ud}$	0.987	0.237	0.965	2.091
2	$v_{ud}, u_{ud}v_{ud}, v_{ud}^3$	0.989	0.222	0.967	2.052
3	$v_{ud}, u_{ud}v_{ud}, u_{ud}^2v_{ud}$	0.998	0.101	0.997	0.651
4	$v_{ud}, u_{ud}v_{ud}, u_{ud}^2v_{ud}, v_{ud}^3$	0.999	0.049	0.999	0.436

As can be seen from both Table 5.2 and Table 5.3, the fitting on the soft distortion data performs much better than the one on the hard distortion data. Almost all the model combinations could achieve good fit qualities, except the Combo 1 in Table 5.2 which obtains a relative low adjusted R-Square and a much higher *RMSE* than the others. The overall good fitting performance on the soft distortion data is because it suffers less influence by the biprism with small prism angle, and hence has smaller amount of distortions and less data variance. These properties of the soft distortion reduce the complexity of the distortion model. Therefore, for the soft distortion, Combo 2 in $f_u(u_{ud}, v_{ud})$ and Combo 1 in $f_v(u_{ud}, v_{ud})$ are good enough for the distortion modelling.

In contrast, when increasing the biprism angle, the choice of the distortion model becomes more crucial. For the hard distortion data in Table 5.2, Combo 1 gives the poorest performance among all combinations, which is unsurprising since it is even unable to properly fit the soft distortion data that has less deviation. For the rest of the combinations of $f_u(u_{ud}, v_{ud})$, although their *RMSE* values vary from 6.6 to less than 1, they are relatively small comparing with the one by Combo 1. Besides, all their adjusted R-Square values are higher than 0.99, which are still considered as good fitting models. Therefore, Combo 2 of $f_u(u_{ud}, v_{ud})$ should be sufficient for the hard distortion fitting. On the other hand, for the hard distortion fitting in Table 5.3, all the combinations of $f_v(u_{ud}, v_{ud})$ can achieve good fit

qualities. However, when comparing the fitting performance between each Combo, there is a significant improvement from Combo 2 to Combo 3 in both adjusted R-Square and *RMSE*. This indicates that the term $u_{ud}^2 v_{ud}$ in $f_v(u_{ud}, v_{ud})$ is an essential term to fit the hard distortion in the Y axis.

In summary, a further refined BD-Model according to the fit quality for different amount of distortions could be defined as:

$$\begin{cases} f_u(u_{ud}, v_{ud}) = p_{10}^u u_{ud} + p_{20}^u u_{ud}^2 + p_{02}^u v_{ud}^2 \left[+p_{30}^u u_{ud}^3 \left[+p_{12}^u u_{ud} v_{ud}^2\right]\right] \\ f_v(u_{ud}, v_{ud}) = p_{01}^v v_{ud} + p_{11}^v u_{ud} v_{ud} + p_{21}^v u_{ud}^2 v_{ud} \left[+p_{03}^v v_{ud}^3\right] \end{cases} \quad \textit{Eq. 5.25}$$

Here the terms $p_{30}^u u_{ud}^3$ and $p_{12}^u u_{ud} v_{ud}^2$ of the horizontal distortion and the term $p_{03}^v v_{ud}^3$ of the vertical distortion are optional according to the amount of distortion observed and the accuracy and efficiency required. To simplify the presentation of the model combinations, here we use $[du\#_1, dv\#_2]$ to indicate the number of terms used in the model. For example, $[du4, dv3]$ means the first 4 terms in $f_u(u_{ud}, v_{ud})$ and first 3 terms in $f_v(u_{ud}, v_{ud})$ are used. Generally, more terms can yield higher accuracy and can better deal with severer distortions, but doing so would reduce the efficiency in the distortion modelling. Practically, the choice of the distortion model may affect the calibration efficiency and accuracy while applying the proposed distortion model in the calibration procedure.

5.4.1.2 Camera Calibration with the BD-Model

As BD-Model has been tested to be efficient to fit the biprism distortion. This model is then tested in the calibration procedure to replace the typical radial and/or tangential distortion models. Different from the distortion data sets for the model fitting, the distorted image points for the calibration are generated based on the natural projection (§3.2) of a checkerboard. To achieve this, the checkerboard with a dimension of 9×7 squares (8×6

corner points) and square size of 25 mm was chosen. By placing this checkerboard at 20 different positions and orientations within the camera FOV, image points of the corner points of the squares on this checkboard can be extracted. Zhang's calibration method [20] is then applied to these data sets, with BD-Model as a replacement of the radial and tangential distortion models. Both soft (small biprism angle) and hard (big biprism angle) distorted image data sets are used for the calibration and evaluation.

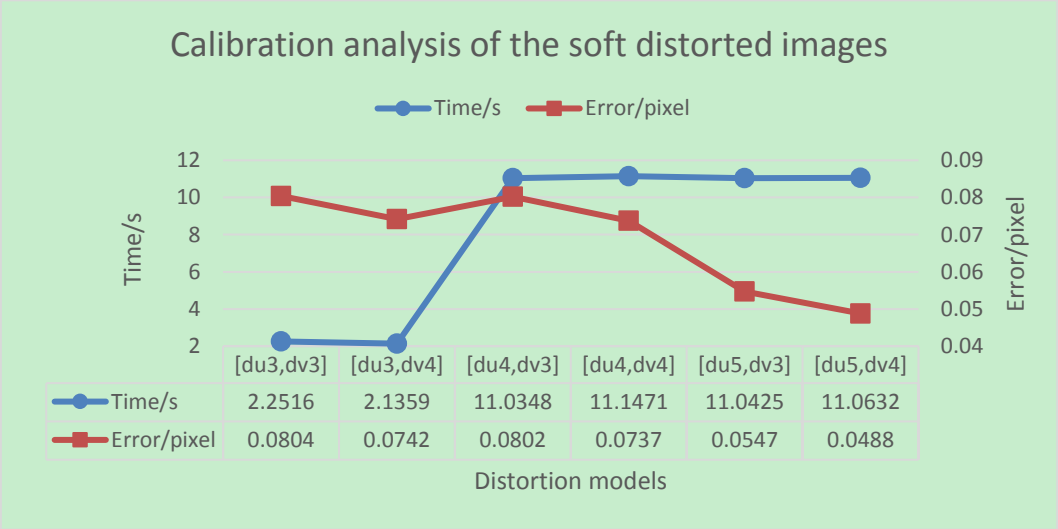


Fig. 5.7 Calibration time and accuracy of the soft distorted data with different models

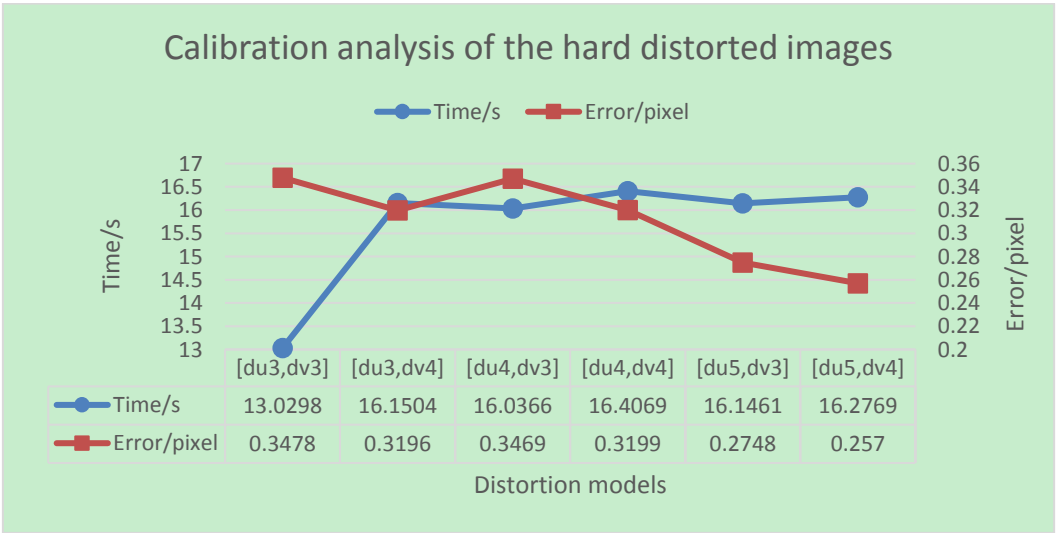


Fig. 5.8 Calibration time and accuracy of the hard distorted data with different models

The measurements include the calibration time and accuracy. Generally, the calibration accuracy, or calibration error (§3.5.2.1), can be computed by determining the discrepancy between the re-projected image points using the estimated system parameters through calibration and the observed distorted image points [92]. The calibration time is measured by running the calibration method with different distortion models on the same computer. The results for the soft and hard distorted cameras are shown in Fig. 5.7 and Fig. 5.8, respectively. As can be seen from both of the calibration statistics, the overall calibration time is relatively small when using the $[du3, dv3]$ terms comparing with the others, while the corresponding calibration errors are still acceptable (under $1\ pixel$) and are considered as good calibration result. We also notice that the calibration time increases by a large amount when using $[du4, dv3]$ for the soft distorted images and $[du3, dv4]$ for the hard distorted images. This dramatic change may be due to the optimization procedure in the calibration method, whose regression speed is sensitive to the number of coefficients. Meanwhile, the calibration accuracy improves while increasing the coefficients of the distortion model. This trend of the calibration accuracy in relation to the number of coefficients is consistent with the performance in the model fitting evaluation in *Table 5.2* and *Table 5.3*.

In summary, both of the simulated soft and hard distortions are well corrected with the calibration errors lower than $1\ pixel$, which indicates that BD-Model is effective for the biprism distortion modelling. Furthermore, using more coefficients (or terms) in the distortion model should improve the calibration. However, it may cause overfitting to the noise in the image. In the next section, more tests on this distortion model will be carried out with the real image data.

5.4.2 Biprism Distortion Model Evaluation with the Real Data

To generate the data sets for the calibration, 20 images of a checkerboard with 9×7 squares (8×6 corner points) are acquired at different positions and orientations within the virtual camera FOV. By changing the biprism in front of the camera and repeating the capturing process, image data sets with different levels of distortions can be generated. These image data sets are utilized for the calibration of the intrinsic parameters of the camera and the BD-Model coefficients. Thereafter, the distortion maps can be calculated using this BD-Model and its coefficients. At last, the images can be finally corrected with the distortion maps.

Table 5.4. Calibration results using BD-Model for different system setups

Parameters	Real camera		Left virtual camera			Right virtual camera		
	RD-Model	BD-Model	$\alpha = 15.5°$	$\alpha = 21.8°$	$\alpha = 35°$	$\alpha = 15.5°$	$\alpha = 21.8°$	$\alpha = 35°$
$u_c/pixels$	649.39	668.46	681.92	740.67	729.74	-28.23	-97.02	-77.37
$v_c/pixels$	363.28	365.07	357.78	346.99	336.35	365.64	349.48	334.89
$f_x/pixels$	935.46	898.60	941.99	982.03	941.52	934.08	954.79	923.50
$f_y/pixels$	939.27	903.61	914.78	952.73	914.87	914.52	956.34	940.73
p_{10}^u/mm^{-1}	-	0.0104	0.1610	0.0949	-0.0323	-0.1060	0.0112	0.2168
p_{20}^u/mm^{-2}	-	0.0406	0.4843	0.4646	0.2526	0.3459	0.2840	-0.0580
p_{02}^u/mm^{-2}	-	0.0623	0.3297	0.8122	1.2621	0.2847	0.8494	1.4198
p_{30}^u/mm^{-3}	-	-0.0397	0.3875	0.2430	-0.0283	-0.2705	-0.1487	0.2658
p_{12}^u/mm^{-3}	-	0.0353	0.5012	0.8889	1.6689	-0.404	-0.9085	-2.0452
p_{01}^v/mm^{-2}	-	-0.0045	-0.0202	-0.0039	-0.0122	0.0002	-0.0149	0.0126
p_{11}^v/mm^{-2}	-	0.0274	-0.0146	0.1919	0.1799	-0.0198	-0.0713	-0.3640
p_{21}^v/mm^{-3}	-	0.1193	0.0314	0.2376	0.3939	0.0182	0.0661	0.6717
p_{03}^v/mm^{-3}	-	0.0111	-0.0274	0.2407	-0.0127	-0.0231	0.2444	0.0299
$\boldsymbol{e_d/pixels}$	**0.21**	**0.29**	**0.24**	**0.31**	**0.35**	**0.18**	**0.32**	**0.39**

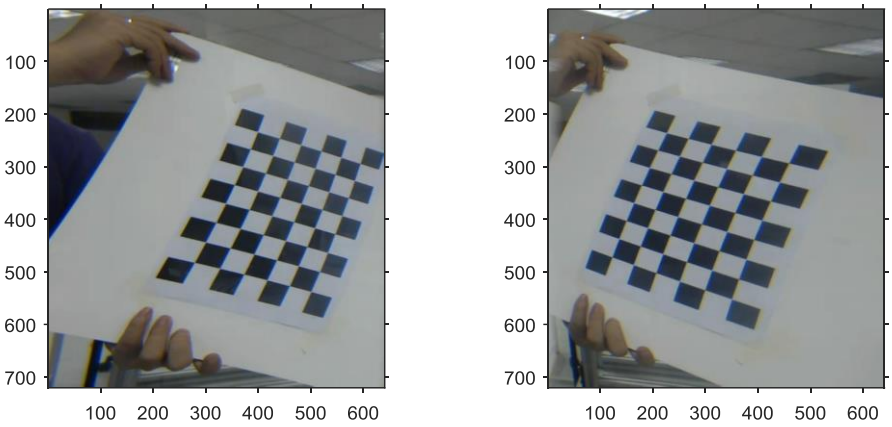
Detailed results of the camera calibration are shown in Table 5.4, which includes the principal point (u_c, v_c), the camera focal length (f_x, f_y), the distortion coefficients (p^u, p^v), and the calibration error (e_d). Regarding the real camera without using the biprism, the camera parameters are calibrated in advance using the radial distortion model (RD-Model), and are used as the reference data, it can obtain a good calibration accuracy of 0.21 $pixel$ error.

As a comparison, the real camera is also calibrated using the same images with BD-Model. It can be seen in *Table 5.4* that the proposed BD-Model can also give a good camera parameter estimation for the real camera with a low calibration error comparing with the calibration results using RD-Model. For the calibration results of the virtual cameras, the calibration error reveals that BD-Model is effective to correct the biprism distortion in the system, by obtaining low calibration errors for different system setups. While comparing u_c and v_c in the left and right virtual cameras, respectively, the deviations in u_c are obviously higher than v_c. This is because the distortion in the X_c direction becomes much more severe than that in the Y_c direction when the biprism angle increases. More distortion along X_c direction will therefore reduce the accuracy of the principle point estimation in the horizontal direction, since the process of determining the principle point is always coupled with the distortion parameter estimation in our calibration procedure.

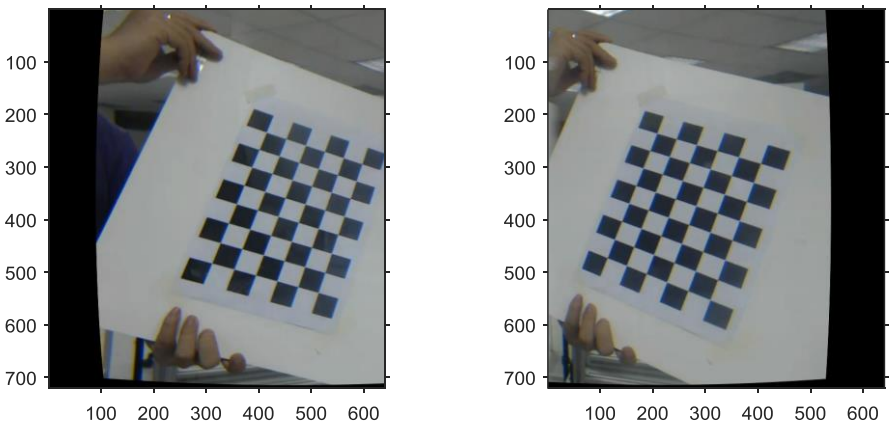
When we compare the calibration errors in the left and right virtual cameras, we notice that the calibration error increases with a larger biprism angle. The reasons behind this may include: 1) the severe distortion may intensify the image noise as well, which causes more errors in the corner point detection; 2) as the amount of distortion increases, it may require higher order (or more coefficients) for the BD-Model to better model the distortion. However,

from a practical point of view, the current calibration accuracies by the proposed BD-Model are generally good enough for the future stereo applications.

Fig. 5.9 illustrates an example of the original and distortion corrected images from the system using $35°$ biprism. The original distorted image is separated into left and right sub-images (shown in Fig. 5.9 (a)) and then corrected using BD-Model (shown in Fig. 5.9 (b)), respectively. As can be seen, observable severe distortions in both the left and right virtual cameras are corrected significantly. Besides, the distortion along the X direction is much more serious than the Y direction, which is consistent with the analysis in §3.2.3.



(a) Original distorted sub-images from the left and right virtual cameras



(b) Corrected sub-images for the left and right virtual cameras

Fig. 5.9 Distorted and corrected images for both left and right virtual cameras.

In addition, to further evaluate the distortion correction using the BD-Model, systems with different setups (as illustrated in §3.5.2) are calibrated. Thereafter, the testing image data from §3.5.2 are corrected and used for depth recovery using the same depth recovery method. The recovered depth errors before and after the correction are shown in Fig. 5.10. As can be seen, the overall depth recovery accuracy is improved after the correction using BDMC, which indicates that the BD-Model is suitable for modelling the distortions in all the system setups. In particular, the model can recover the depth with high accuracy for system known to have severe distortions (such as the system using $f = 4\ mm, \alpha = 35°$) . Besides, for those systems with less amount of distortions, BDMC is still able to calibrate the system properly and achieve a good depth recovery accuracy. In other words, the BD-Model proposed in this section can model the distortion adequately and lead to a better calibration and depth recovery accuracy.

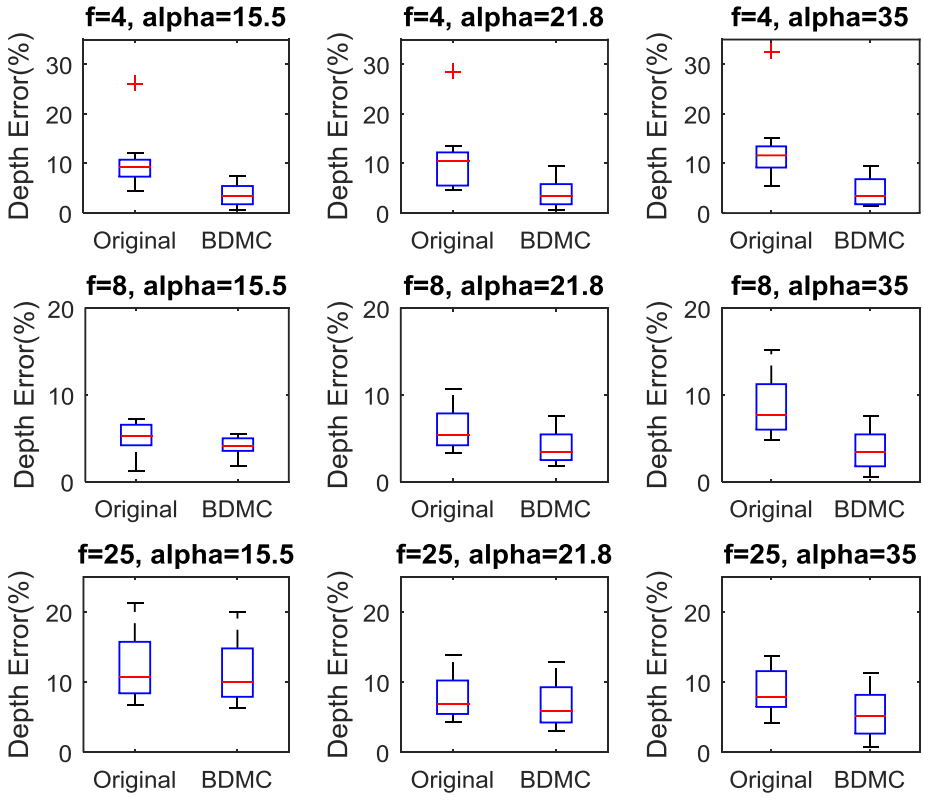


Fig. 5.10 Comparison of the depth error before and after the correction by BDMC for the real testing data from different system setups.

5.5 SUMMARY

In summary, the biprism distortion model (BD-Model) is proposed using two different methods in this chapter. The first one is an empirical or experimental method which is to develop the distortion model by curve-fitting the given distortion maps using a general bivariate polynomial function. This function is further refined by the biprism properties. The second one is a result of the geometrical expression of the distortion formation. It fully represents the biprism distortion by transforming the algebraic expressions to polynomial expressions using the Taylor expansions. Although the derivations of these two models are totally different, they are found to be similar as far as the terms associating with the image coordinates are concerned. This modelling process helps us to better understand the formation of the biprism distortion in the system, and provides a robust way of correcting the distortions. Furthermore, the experiments show that the proposed BD-Model can fit the distortion data well and improve the accuracy of calibration and depth recovery of the system.

Chapter 6. CONCLUSIONS AND FUTURE WORKS

This work studied the formation and correction of distortions in the images captured by a single-lens biprism-based stereovision system. Three different methods were proposed: (1) geometrical analysis based distortion correction (GABDC) method, (2) fully constrained and model-free distortion correction (FC-MFDC) method, and (3) biprism distortion modelling and correction (BDMC) method. Through experiments, all of them have been shown to be efficient, to different degrees, in removing the distortions and improved the accuracy in system calibration and depth recovery of the said stereovision system.

6.1 CONTRIBUTIONS OF THE THESIS

The contributions of this thesis include the exploration of the formation of the biprism distortion in the system and the development of three distortion correction methods:

(a) Geometrical analysis based distortion correction (GABDC) method

The aim of GABDC is to determine the distortion formation during the refractions through the biprism. To achieve this, both the real (distorted) and ideal (undistorted) projections are first defined and calculated. Then, for the given system parameters, the distortion maps can be obtained from the differences between the real (distorted) projection and the ideal (undistorted) projection. Therefore, the images can be corrected using these distortion maps as long as the system parameters remain unchanged. This method can also be used to analyze the effects of the system parameters to the formation of the distortion. Both synthetic data and experimental data are used to evaluate the effectiveness of the method and the results are encouraging. They show that this method is effective in reducing the amount of distortion

in image and improve the accuracy of the camera calibration and depth recovery. However, although no captured images are required for the determination of the distortion maps, GABDC depends highly on the system parameters, especially the focal length of the camera and the biprism angle. The errors on the measurement of the system parameters will directly affect the accuracy of the distortion correction maps, and would therefore raise the risk of failure in correction.

(b) Fully constrained and model-free distortion correction (FC-MFDC) method

The goal of this method is to determine the distortion corrected points for the given distorted points, based only on the perspective projection constraints. These constraints include the straight line, cross-ratio, and vanishing point constraints. Based on the special formation of the virtual cameras, another constrain of distortion-free reference point is also introduced to enhance the accuracy of correction. With both the distorted and corrected points, the distortion maps can be built and is then used to correct the captured images. Experiments have been carried out with both synthetic and real image data. Comparing with another model-free method (MLDC), FC-MFDC outperforms in both calibration and depth recovery. The advantage of this method is that it requires no prior knowledge of the system parameters, but only the perspective projection constraints. However, FC-MFDC also suffers from some limitations including the restrictions on the acquisition of the training data and the errors of the vanishing points estimation.

(c) Biprism distortion modelling and correction (BDMC) method

Both GABDC and FC-MFDC try to isolate the distortion from the camera calibration by determining the distortion maps and removing the distortion directly from the image. BDMC provides another understanding of the distortion formation by using a polynomial function to

describe the distortion, which is named as biprism distortion model (BD-Model). This model could be further applied to the camera calibration procedure for distortion correction and system parameter determination. Two different methods are proposed to develop the BD-Model. The first method is to fit the distortion maps generated from GABDC using a bivariate polynomial function. This function is further refined according to the properties of distortion present. The second method is to mathematically express the formation of the distortion by geometrically analyzing the distortions during each projection steps through the biprism. Although these two methods are different in the modelling process, their distortion functions are found to be similar if we only consider the terms associating with the function inputs. Comparing between the experimental results of the stereovision system calibration using Zhang's method [20] and the BD-Model show that the latter could effectively represent the biprism distortion and enhance the accuracy of the system calibration and depth recovery.

Table 6.1 Advantages and disadvantages of the proposed distortion correction methods

Method	Advantages	Disadvantages
GABDC	• Geometrical analysis based • No training images or calibration procedure required • Quantitatively analyze the effects of the system parameters to the distortion	• Require full knowledge of the system parameters • Sensitive to the error of the system parameters
FC-MFDC	• Require no system parameters • Insensitive to the system misalignment • Insensitive to the image noise • Adapt to different forms of distortions	• Strict requirements on the capturing of the training image data
BDMC	• Provide mathematical expression of the distortion • Compatible with the existing calibration methods • Insensitive to the system misalignment	• Time consuming

In summary, the three above proposed distortion correction methods could effectively remove the distortion in the system. Each of them has their own advantages and disadvantages, which are summarized and listed in Table 6.1.

6.2 FUTURE WORK

In this thesis, the developed algorithms for correcting the image distortion in a biprism-based single-lens stereovision system have been shown to be efficient with the improvement of the accuracy of the system calibration and depth recovery. Besides, these methods provide a new understanding of the biprism distortion and would help the future development of the single-lens stereovision system. However, there are still some limitations. The overcoming of them will form the future endeavors in this research. Some recommendations are suggested as follows:

1) There is room for improving the proposed correction methods by taking into consideration of the system alignment errors, and light refraction and diffraction errors. Besides, the effect of the distortion will be even more severe when using the wide angle cameras, such as the one used in endoscopic applications. The contributions of these possible sources or errors should be studied in details.

2) In GABDC, we have pointed out that the main parameters which would affect the distortions are the focal length and the biprism angle. The latter could be measured accurately by benefiting from its simple geometry. However, the measurement accuracy of the focal length would be more critical to the distortion especially when the camera has a small focal length. In addition, camera with small focal length is always accompanied by the severe

distortions which shall be removed properly before the determination of the biprism distortion. Besides, the effects of the biprism alignment errors to the distortion could also be evaluated using GABDC, by introducing the position (mainly the horizontal movement) and orientation parameters of the biprism to the calculation.

3) FC-MFDC imports all the constraints from both the perspective projection invariances and the system setup. However, within these constraints, the determination of the VP would be critical when there is severe distortion in the system. The error on the VP estimation should be further studied. Meanwhile, although some conditions on the training image capturing have been provided, the evaluation on how the errors are incurred during the training image acquisition would affect the correction should be carried out in the future. Moreover, given the geometrical relationship between the left and right virtual cameras, introducing the co-dependency of the distortions to the correction procedure could be a good way to improve the robustness and accuracy of the distortion correction. However, this co-dependency of the distortions depends on the alignment conditions of the biprism, which should be further discussed in details.

4) For the BDMC method proposed in Chapter 5, the main purpose is to find the distortion model which is corresponding to the image coordinates (x, y). Thus, all the constant terms in the model are simplified. In the future, more specific analysis of the relationship between the amount of distortions and the system parameters (such as focal length, image size, biprism angle, distance between the camera and the biprism, etc.) could be carried out by fully expressing the distortion model in terms of these parameters.

LIST OF PUBLICATIONS

1. B. B. Qian, and K. B. Lim, Image distortion correction for single-lens stereo vision system employing a bi-prism. In *Journal of Electronic Imaging*, (accepted on 15 July 2016).
2. K. B. Lim, and B. B. Qian, Biprism Distortion Modelling and Calibration for Single-lens Stereovision System. In *Journal of the Optical Society of America A*, (under review since 16 May 2016).

REFERENCE

[1] D. Murray and J. J. Little, "Using real-time stereo vision for mobile robot navigation," *Autonomous Robots,* vol. 8(2), pp. 161-171, 2000.

[2] N. Ayache, *Artificial vision for mobile robots: stereo vision and multisensory perception*: MIT Press, 1991.

[3] C. Bert, K. G. Metheany, K. Doppke, and G. T. Chen, "A phantom evaluation of a stereo-vision surface imaging system for radiotherapy patient setup," *Medical physics,* vol. 32(9), pp. 2753-2762, 2005.

[4] W. Birkfellner, M. Figl, C. Matula, J. Hummel, R. Hanel, H. Imhof, *et al.*, "Computer-enhanced stereoscopic vision in a head-mounted operating binocular," *Physics in medicine and biology,* vol. 48(3), p. N49, 2003.

[5] D. O. Gorodnichy and G. Roth, "Nouse 'use your nose as a mouse'perceptual vision technology for hands-free games and interfaces," *Image and Vision Computing,* vol. 22(12), pp. 931-942, 2004.

[6] J. P. Wachs, M. Kölsch, H. Stern, and Y. Edan, "Vision-based hand-gesture applications," *Communications of the ACM,* vol. 54(2), pp. 60-71, 2011.

[7] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis, and machine vision* vol. 3: Thomson Toronto, 2008.

[8] Y. Matsumoto and A. Zelinsky, "An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement," in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pp. 499-504, 2000.

[9] M. Agrawal and K. Konolige, "Real-time Localization in Outdoor Environments using Stereo Vision and Inexpensive GPS," in *ICPR (3)*, pp. 1063-1068, 2006.

[10] Y. Nishimoto and Y. Shirai, "A feature-based stereo model using small disparities," in *Proc. Computer Vision and Pattern Recognition*, pp. 192-196, 1987.

[11] K. B. Lim and Y. Xiao, "Virtual stereovision system: new understanding on single-lens stereovision using a biprism," *Journal of Electronic Imaging,* vol. 14(4), pp. 043020-1-043020-11, 2005.

[12] Y. Xiao and K. B. Lim, "A prism-based single-lens stereovision system: From trinocular to multi-ocular," *Image and Vision Computing,* vol. 25(11), pp. 1725-1736, 2007.

[13] K. B. Lim, W. L. Kee, and D. Wang, "Virtual camera calibration and stereo correspondence of single-lens bi-prism stereovision system using geometrical approach," *Signal Processing: Image Communication,* vol. 28(9), pp. 1059-1071, 2013.

[14] K. B. Lim, D. Wang, and W. L. Kee, "Virtual camera rectification with geometrical approach on single-lens stereovision using a biprism," *Journal of Electronic Imaging,* vol. 21(2), pp. 023003-1-023003-9, 2012.

[15] D. Wang, K. B. Lim, and W. L. Kee, "Geometrical approach for rectification of single-lens stereovision system with a triprism," *Machine vision and applications,* vol. 24(4), pp. 821-833, 2013.

[16] D. Wang and K. B. Lim, "Obtaining depth map from segment-based stereo matching using graph cuts," *Journal of Visual Communication and Image Representation,* vol. 22(4), pp. 325-331, 2011.

[17] M. Zhao, K. B. Lim, and W. L. Kee, "Geometrical-Analysis-Based Algorithm for Stereo Matching of Single-Lens Binocular and Multi-Ocular Stereovision System," *Journal of Electronic Science and Technology,* vol. 10(2), pp. 107-112, 2012.

[18] K. B. Lim and J. Y. Wu, "Recognition of occluded objects by reducing feature interactions," *Image and Vision Computing,* vol. 30(11), pp. 906-914, 2012.

[19] J. Y. Wu, K. B. Lim, and X. Cui, "Recognition of occluded objects by feature interactions," in *Robotics Automation and Mechatronics (RAM), 2010 IEEE Conference on*, pp. 374-379, 2010.

[20] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 22(11), pp. 1330-1334, 2000.

[21] W. Teoh and X. Zhang, "An inexpensive stereoscopic vision system for robots," in *Robotics and Automation. Proceedings. 1984 IEEE International Conference on*, pp. 186-189, 1984.

[22] Z.-Y. Zhang and H.-T. Tsui, "3D reconstruction from a single view of an object and its image in a plane mirror," in *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on*, pp. 1174-1176, 1998.

[23] A. Goshtasby and W. A. Gruver, "Design of a single-lens stereo camera system," *Pattern Recognition,* vol. 26(6), pp. 923-937, 1993.

[24] A. R. François, G. G. Medioni, and R. Waupotitsch, "Mirror symmetry⇒ 2-view stereo geometry," *Image and Vision Computing,* vol. 21(2), pp. 137-143, 2003.

[25] D. Lee and I. Kweon, "A novel stereo camera system by a biprism," *Robotics and Automation, IEEE Transactions on,* vol. 16(5), pp. 528-541, 2000.

[26] R. LeGrand and R. C. Luo, "Position estimation of selected targets," in *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, pp. 1714-1719, 1996.

[27] J. Segen and S. Kumar, "Shadow gestures: 3D hand pose estimation using a single camera," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, 1999.

[28] E. H. Adelson and J. Y. Wang, "Single lens stereo with a plenoptic camera," *IEEE transactions on pattern analysis and machine intelligence,* vol. 14(2), pp. 99-106, 1992.

[29] M. Inaba, T. Hara, and H. Inoue, "A stereo viewer based on a single camera with view-control mechanisms," in *Intelligent Robots and Systems' 93, IROS'93. Proceedings of the 1993 IEEE/RSJ International Conference on*, pp. 1857-1865, 1993.

[30] S. A. Nene and S. K. Nayar, "Stereo with mirrors," in *Computer Vision, 1998. Sixth International Conference on*, pp. 1087-1094, 1998.

[31] D. Moore and M. Hayes, "Tracking 3d position and orientation from 2D sequences using simple geometry," in *Signals, Systems and Computers, 1996. Conference Record of the Thirtieth Asilomar Conference on*, pp. 125-129, 1996.

[32] X. Li and R. Wang, "Analysis and optimization of the stereo system with a biprism adapter," in *International Conference on Optical Instrumentation and Technology*, pp. 75061V-75061V-8, 2009.

[33] K. Genovese, L. Casaletto, J. Rayas, V. Flores, and A. Martinez, "Stereo-Digital Image Correlation (DIC) measurements with a single camera using a biprism," *Optics and Lasers in Engineering,* vol. 51(3), pp. 278-285, 2013.

[34] X. Cui, K. B. Lim, Q. Guo, and D. Wang, "Accurate geometrical optics model for single-lens stereovision system using a prism," *JOSA A,* vol. 29(9), pp. 1828-1837, 2012.

[35] X. Cui and K. B. Lim, "Position Estimation Method for Prism Based Stereovision System," in *AASRI Procedia*, pp. 481-487, 2012.

[36] M. Zhao and K. B. Lim, "Stereo matching of single-lens bi-prism based stereovision system," in *Proceedings of the 2010 International Conference on Future Information Technology*, pp. 179-183, 2010.

[37] M. Zhao, K. B. Lim, and W. L. Kee, "Geometrical-analysis-based algorithm for stereo matching of single-lens binocular and multi-ocular stereovision system," *J. Electron. Sci. Technol.*, vol. 10(2), pp. 107-112, 2012.

[38] W. L. Kee, K. B. Lim, and D.-L. Wang, "Virtual epipolar line construction of single-lens bi-prism stereovision system," *J. Electron. Sci. Technol*, vol. 10(2), pp. 97-101, 2012.

[39] J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14(10), pp. 965-980, 1992.

[40] S. S. Beauchemin and R. Bajcsy, "Modelling and removing radial and tangential distortions in spherical lenses," in *Multi-Image Analysis*, ed: Springer, 2001, pp. 1-21.

[41] S. Wang, Y. Geng, and R. Jin, "A Novel Error Model of Optical Systems and an On-Orbit Calibration Method for Star Sensors," *Sensors*, vol. 15(12), pp. 31428-31441, 2015.

[42] J. P. De Villiers, F. W. Leuschner, and R. Geldenhuys, "Centi-pixel accurate real-time inverse distortion correction," in *International Symposium on Optomechatronic Technologies*, pp. 726611-726611-8, 2008.

[43] D. C. Brown, "Decentering distortion of lenses," *Photometric Engineering*, vol. 32(3), pp. 444-462, 1966.

[44] D. C. Brown, "Close-range camera calibration," *Photogrammetric engineering*, vol. 37(8), pp. 855-866, 1971.

[45] 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*, vol. 3(4), pp. 323-344, 1987.

[46] R. I. Hartley and S. B. Kang, "Parameter-free radial distortion correction with centre of distortion estimation," in *Tenth IEEE International Conference on Computer Vision, 2005. ICCV 2005.* , pp. 1834-1841, 2005.

[47] Z. Zhang, "On the epipolar geometry between two images with lens distortion," in *Proceedings of the 13th International Conference on Pattern Recognition*, pp. 407-411, 1996.

[48] G. P. Stein, "Lens distortion calibration using point correspondences," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 602-608, 1997.

[49] A. W. Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. I-125-I-132 vol. 1, 2001.

[50] H. Farid and A. C. Popescu, "Blind removal of lens distortion," *JOSA A,* vol. 18(9), pp. 2072-2078, 2001.

[51] J. Salvi, X. Armangué, and J. Batlle, "A comparative review of camera calibrating methods with accuracy evaluation," *Pattern recognition,* vol. 35(7), pp. 1617-1635, 2002.

[52] G. P. Stein, "Lens distortion calibration using point correspondences," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pp. 602-608, 1997.

[53] Z. Zhang, "On the epipolar geometry between two images with lens distortion," in *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, pp. 407-411, 1996.

[54] D. Gonzalez-Aguilera, J. Gomez-Lahoz, and P. Rodriguez-Gonzalvez, "An automatic approach for radial lens distortion correction from a single image," *IEEE Sensors journal,* vol. 11(4), pp. 956-965, 2011.

[55] H. Li and R. Hartley, "A non-iterative method for correcting lens distortion from nine point correspondences," presented at the OMNIVIS 2005, 2005.

[56] B. Mičušík and T. Pajdla, "Estimation of omnidirectional camera model from epipolar geometry," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, pp. I-485-I-490 vol. 1, 2003.

[57] A. W. Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, pp. I-125-I-132 vol. 1, 2001.

[58] J. P. Barreto and K. Daniilidis, "Fundamental matrix for cameras with radial distortion," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, pp. 625-632, 2005.

[59] D. Claus and A. W. Fitzgibbon, "A rational function lens distortion model for general cameras," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on,* pp. 213-219, 2005.

[60] H. Schreier, D. Garcia, and M. Sutton, "Advances in light microscope stereo vision," *Experimental mechanics,* vol. 44(3), pp. 278-288, 2004.

[61] M. Sutton, N. Li, D. Garcia, N. Cornille, J. Orteu, S. McNeill, *et al.*, "Scanning electron microscopy for quantitative small and large deformation measurements Part II: experimental validation for magnifications from 200 to 10,000," *Experimental mechanics,* vol. 47(6), pp. 789-804, 2007.

[62] M. Neu, N. Moll, L. Gross, G. Meyer, F. J. Giessibl, and J. Repp, "Image correction for atomic force microscopy images with functionalized tips," *Physical Review B,* vol. 89(20), p. 205407, 2014.

[63] R. Melo, J. P. Barreto, and G. Falcao, "A new solution for camera calibration and real-time image distortion correction in medical endoscopy–initial technical evaluation," *IEEE Transactions on Biomedical Engineering,* vol. 59(3), pp. 634-644, 2012.

[64] T.-Y. Lee, T.-S. Chang, C.-H. Wei, S.-H. Lai, K.-C. Liu, and H.-S. Wu, "Automatic distortion correction of endoscopic images captured with wide-angle zoom lens," *IEEE Transactions on Biomedical Engineering,* vol. 60(9), pp. 2603-2613, 2013.

[65] J. P. Helferty, C. Zhang, G. McLennan, and W. E. Higgins, "Videoendoscopic distortion correction and its application to virtual guidance of endoscopy," *IEEE transactions on medical imaging,* vol. 20(7), pp. 605-617, 2001.

[66] R. Hartley and S. B. Kang, "Parameter-free radial distortion correction with center of distortion estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 29(8), pp. 1309-1321, 2007.

[67] M. Ahmed and A. Farag, "Nonmetric calibration of camera lens distortion: differential methods and robust estimation," *IEEE Transactions on Image Processing,* vol. 14(8), pp. 1215-1230, 2005.

[68] T. Thormählen, H. Broszio, and I. Wassermann, "Robust line-based calibration of lens distortion from a single view," in *Proceedings of MIRAGE,* pp. 105-112, 2003.

[69] F. Zhou, Y. Cui, H. Gao, and Y. Wang, "Line-based camera calibration with lens distortion correction from a single image," *Optics and lasers in Engineering,* vol. 51(12), pp. 1332-1343, 2013.

[70] M. Alemán-Flores, L. Alvarez, L. Gomez, and D. Santana-Cedrés, "Line detection in images showing significant lens distortion and application to distortion correction," *Pattern Recognition Letters,* vol. 36(pp. 261-271, 2014.

[71] C. Brauer-Burchardt and K. Voss, "A new algorithm to correct fish-eye-and strong wide-angle-lens-distortion from single images," in *Image Processing, 2001. Proceedings. 2001 International Conference on*, pp. 225-228, 2001.

[72] G.-Y. Song and J.-W. Lee, "Correction of radial distortion based on line-fitting," *International Journal of Control, Automation and Systems,* vol. 8(3), pp. 615-621, 2010.

[73] S. Becker and V. Bove, "Semi-automatic 3D model extraction from uncalibrated 2D camera views," in *Proceedings of Visual Data Exploration and Analysis II 2*, pp. 447-461, 1995.

[74] R. Orghidan, J. Salvi, M. Gordan, C. Florea, and J. Batlle, "Structured light self-calibration with vanishing points," *Machine vision and applications,* vol. 25(2), pp. 489-500, 2014.

[75] H. Wildenauer and B. Micusik, "Closed form solution for radial distortion estimation from a single vanishing point," in *BMVC*, 2013.

[76] G. Zhang, J. He, and X. Yang, "Calibrating camera radial distortion with cross-ratio invariability," *Optics & Laser Technology,* vol. 35(6), pp. 457-461, 2003.

[77] M. A. Penna, "Camera calibration: a quick and easy way to determine the scale factor," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 13(12), pp. 1240-1245, 1991.

[78] C. Ricolfe-Viala and A.-J. Sánchez-Salmerón, "Correcting non-linear lens distortion in cameras without using a model," *Optics & Laser Technology,* vol. 42(4), pp. 628-639, 2010.

[79] C. Ricolfe-Viala and A.-J. Sánchez-Salmerón, "Using the camera pin-hole model restrictions to calibrate the lens distortion model," *Optics & Laser Technology,* vol. 43(6), pp. 996-1005, 2011.

[80] W. L. Kee, Y. Bai, and K. B. Lim, "Parameter error analysis of single-lens prism-based stereovision system," *JOSA A,* vol. 32(3), pp. 367-373, 2015.

[81] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," in *Numerical analysis*, ed: Springer, 1978, pp. 105-116.

[82] K. W. LOON, "Depth recovery and parameter analysis using single-lens prism based stereovision system," PhD Thesis, National University of Singapore, 2014.

[83] W. L. Kee, K. B. Lim, and D.-L. Wang, "Virtual epipolar line construction of single-lens bi-prism stereovision system," *J. Electron. Sci. Technol.,* vol. 10(2), pp. 97-101, 2012.

[84] A. Geiger, F. Moosmann, Ö. Car, and B. Schuster, "Automatic camera and range sensor calibration using a single shot," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pp. 3936-3943, 2012.

[85] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *Robotics and Automation, IEEE Journal of,* vol. 3(4), pp. 323-344, 1987.

[86] S. Ganapathy, "Decomposition of transformation matrices for robot vision," in *Robotics and Automation. Proceedings. 1984 IEEE International Conference on*, pp. 130-139, 1984.

[87] O. D. Faugeras and G. Toscani, "The calibration problem for stereo," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 15-20, 1986.

[88] K. W. Wong, "Mathematical formulation and digital analysis in close-range photogrammetry," *Photogrammetric Engineering and Remote Sensing,* vol. 44(11), 1975.

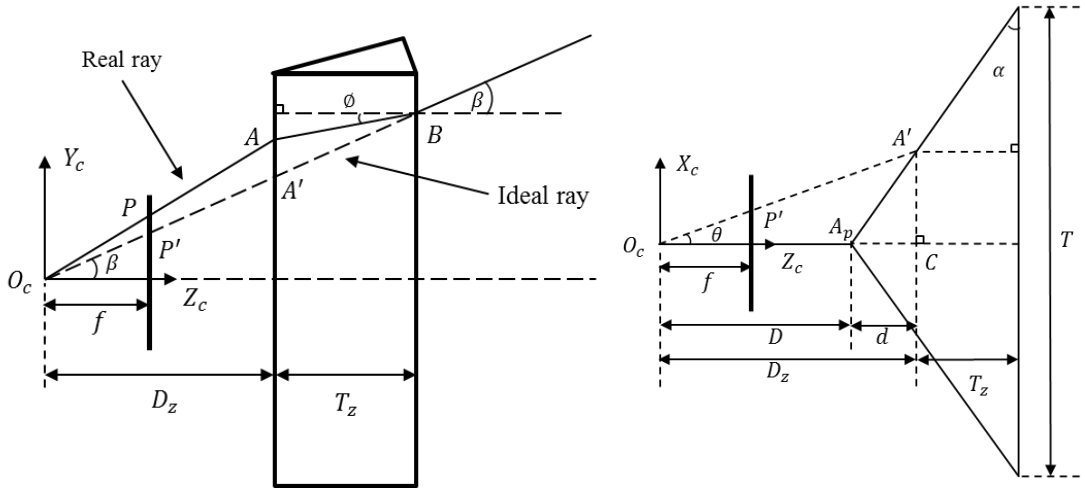
[89] W. Faig, "Calibration of close-range photogrammetric systems: Mathematical formulation," *Photogrammetric engineering and remote sensing,* vol. 41(12), 1975.

[90] H. Gao, C. Wu, L. Gao, and B. Li, "An improved two-stage camera calibration method," in *Intelligent Control and Automation, 2006. WCICA 2006. The Sixth World Congress on*, pp. 9514-9518, 2006.

[91] H. Bacakoglu and M. S. Kamel, "A three-step camera calibration method," *Instrumentation and Measurement, IEEE Transactions on*, vol. 46(5), pp. 1165-1172, 1997.

[92] C. Ricolfe-Viala and A.-J. Sánchez-Salmerón, "Robust metric calibration of non-linear camera lens distortion," *Pattern Recognition,* vol. 43(4), pp. 1688-1699, 2010.

APPENDIX A. GEOMETRICAL DERIVATION OF THE DISTORTION dy



(a) Side view (b) Top view

Fig. A1 Distortion analysis in the Y direction from (a) side view, and its corresponding top view (b)

Fig. A1(a) shows the side view (in the plane of $\{O_c \rightarrow Y_c Z_c\}$) of the system, in which the real (distorted) and ideal (undistorted) rays are plotted. Here we assume that the vertical distance between A and A' is ΔY, and the vertical distance between $P(x_d, y_d)$ and $P'(x_{ud}, y_{ud})$ is dy which is regarded as the distortion in the Y axis. Since triangle $\Delta O_c PP'$ and $\Delta O_c AA'$ are similar, it can be deduced that:

$$\frac{dy}{\Delta Y} = \frac{f}{D_z} \qquad \textit{Eq. A1}$$

From the triangle $\Delta ABA'$ and the incident angle β and the refracted angle $\emptyset$, ΔY could be represented in another way:

$$\Delta Y = T_z(\tan(\beta) - \tan(\emptyset)) \qquad \textit{Eq. A2}$$

According to Snell's law and the small-angle approximation, we can have:

$$\tan(\beta) \approx \beta$$
$$\tan(\emptyset) \approx \emptyset \qquad \textit{Eq. A3}$$

$$\frac{\sin(\beta)}{\sin(\emptyset)} = n \approx \frac{\beta}{\emptyset}$$

Substitute Eq. A3 to Eq. A2, ΔY could be represented as:

$$\Delta Y = \left(1 - \frac{1}{n}\right)\beta T_z \qquad \textit{Eq. A4}$$

Therefore, by substituting Eq. A4 to Eq. A1, the vertical distortion dy could be expressed by:

$$dy = \frac{f\Delta Y}{D_z} = \left(1 - \frac{1}{n}\right) \cdot \beta \cdot T_z \cdot \frac{f}{D_z} \qquad \textit{Eq. A5}$$

Here, the components β, T_z and D_z are unknown, but could be expressed with the known parameters and the point coordinates.

Since β is the vertical projection angle of the undistorted incident ray w.r.t. the Z axis in the YZ plane, it has the relationship with the focal length f and the vertical position y_{ud} of P':

$$\beta = \arctan\left(\frac{y_{ud}}{f}\right) \qquad \textit{Eq. A6}$$

From *Fig. A1*(b), which shows the corresponding top view (in the plane of $\{O_c \rightarrow X_c Z_c\}$) of the system, T_z could be represented by

$$T_z = \frac{T}{2} \cdot tan(\alpha) - d \qquad \textit{Eq. A7}$$

Here d could be further determined using the triangles $\Delta A' A_p C$ and $\Delta A' O_c C$:

$$\left.\begin{array}{l} \dfrac{d}{A'C} = \tan(\alpha) \\ \dfrac{A'C}{D+d} = \tan(\theta) \end{array}\right\} \Longrightarrow d = \frac{D\tan(\alpha)\tan(\theta)}{1 - \tan(\alpha)\tan(\theta)} \qquad \textit{Eq. A8}$$

in which θ is the projection angle of the undistorted ray w.r.t. the X axis in the XZ plane. Thus, we have:

$$\tan(\theta) = \frac{x_{ud}}{f} \qquad \textit{Eq. A9}$$

By substituting Eq. A9 to Eq. A8, d could be written as:

$$d = \frac{x_{ud}\, D \tan(\alpha)}{f - x_{ud} \tan(\alpha)} \qquad \textit{Eq. A10}$$

Therefore, with Eq. A10 and Eq. A7, T_z could be finally represented as:

$$T_z = \frac{T}{2} \cdot tan(\alpha) - \frac{x_{ud}\, D \tan(\alpha)}{f - x_{ud} \tan(\alpha)} \qquad \textit{Eq. A11}$$

For D_z, we have:

$$D_z = D + d = D + \frac{x_{ud}\, D \tan(\alpha)}{f - x_{ud} \tan(\alpha)} = \frac{fD}{f - x_{ud} \tan(\alpha)} \qquad \textit{Eq. A12}$$

Finally, substitute Eq. A6, Eq. A11, and Eq. A12 to Eq. A5, the vertical distortion could be expressed by:

$$dy = \arctan\left(\frac{y_{ud}}{f}\right) \cdot \left(1 - \frac{1}{n}\right) \cdot \left(\frac{T}{2}\tan(\alpha) - \frac{x_{ud} D \tan(\alpha)}{f - x_{ud} \tan(\alpha)}\right) \cdot \frac{f - x_{ud} \tan(\alpha)}{D} \qquad \textit{Eq. A13}$$

APPENDIX B. GEOMETRICAL DERIVATION OF THE DISTORTION dx

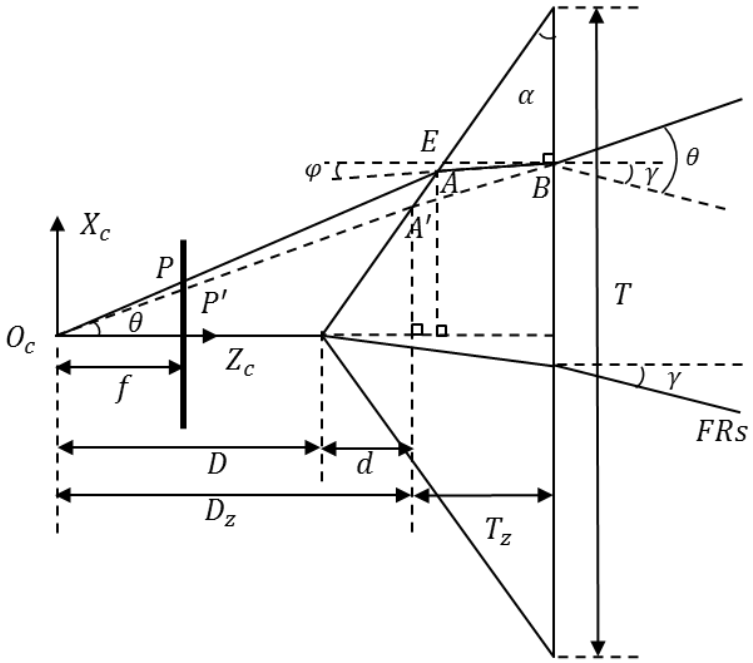


Fig. B1 Distortion analysis in the X direction from the top view

Fig. B1 shows the corresponding top view (in the plane of $\{O_c \rightarrow X_c Z_c\}$) of the system with real and ideal rays. The distortion in the X direction ($dx = x_d - x_{ud}$) could be expressed using the relationship between $\Delta O_c PP'$ and $\Delta O_c AA'$:

$$\frac{dy}{\Delta X} = \frac{f}{D_z} \qquad \textit{Eq. B1}$$

which is similar with Eq. A1 by assuming $\Delta O_c PP' \sim \Delta O_c AA'$. Here ΔX is the horizontal distance between A and A'. Eq. B1 can hold only if $\angle P O_c P'$ and α are both small angles. According to the analysis and experiments in Chapter 3, small α will lead to small amount of distortion, namely, small $\angle P O_c P'$. Hence, for the given system setup (α is relatively small), Eq. B1 still holds. Similarly, since the distances between E, A and A' are very small, we can have:

$$\tan(\varphi) = \frac{EA}{T_z'}$$
$$\tan(\theta - \gamma) = \frac{EA'}{T_z'} \quad \textit{Eq. B2}$$

Thus,

$$\Delta Y = EA' - EA = T_z'(\tan(\theta - \gamma) - \tan(\varphi)) \quad \textit{Eq. B3}$$

Here γ is the incident angle of the FRs, which is fixed for the given system; T_z' is the length of the ideal ray passing through the biprism on the plane of $\{O_c \rightarrow Y_c Z_c\}$ (such as the $A'B$ in *Fig. A1*(a)); φ is the refracted angle of the sample ray with the incident angle of $(\theta - \gamma)$.

Here T_z' may not equal to T_z because of the incident angle β of the sample ray in the Y direction. Instead, we have:

$$T_z' = \frac{T_z}{\cos(\beta)} \quad \textit{Eq. B4}$$

Moreover,

$$\cos(\beta) = \frac{f}{\sqrt{f^2 + y_{ud}^2}} = \frac{1}{\sqrt{1 + \left(\frac{y_{ud}}{f}\right)^2}} \quad \textit{Eq. B5}$$

So, T_z' can be further expressed as:

$$T_z' = T_z \sqrt{1 + \left(\frac{y_{ud}}{f}\right)^2} \quad \textit{Eq. B6}$$

For the refracted angle φ, we have

$$n = \frac{\theta - \gamma}{\varphi} \quad \textit{Eq. B7}$$

Using the small-angle approximation and substituting Eq. B3 and Eq. B7 to Eq. B1, dy could be expressed by:

$$dx = \frac{f\Delta X}{D_z} = \left(1 - \frac{1}{n}\right) \cdot (\theta - \gamma) \cdot T_z \cdot \frac{f}{D_z} \cdot \sqrt{1 + \left(\frac{y_{ud}}{f}\right)^2} \qquad \textit{Eq. B8}$$

within this equation, θ could be achieved from Eq. A9, and D_z has been determined in Eq. A12. Therefore, the final expression of dx is

$$dx = \frac{f\Delta X}{D_z} = \left(1 - \frac{1}{n}\right) \cdot \left(\arctan\left(\frac{x_{ud}}{f}\right) - \gamma\right) \cdot \left(\frac{T}{2}\tan(\alpha) - \frac{x_{ud} D \tan(\alpha)}{f - x_{ud}\tan(\alpha)}\right) \cdot \frac{f - x_{ud}\tan(\alpha)}{D} \cdot \sqrt{1 + \left(\frac{y_{ud}}{f}\right)^2} \qquad \textit{Eq. B9}$$

APPENDIX C. TAYLOR SERIES

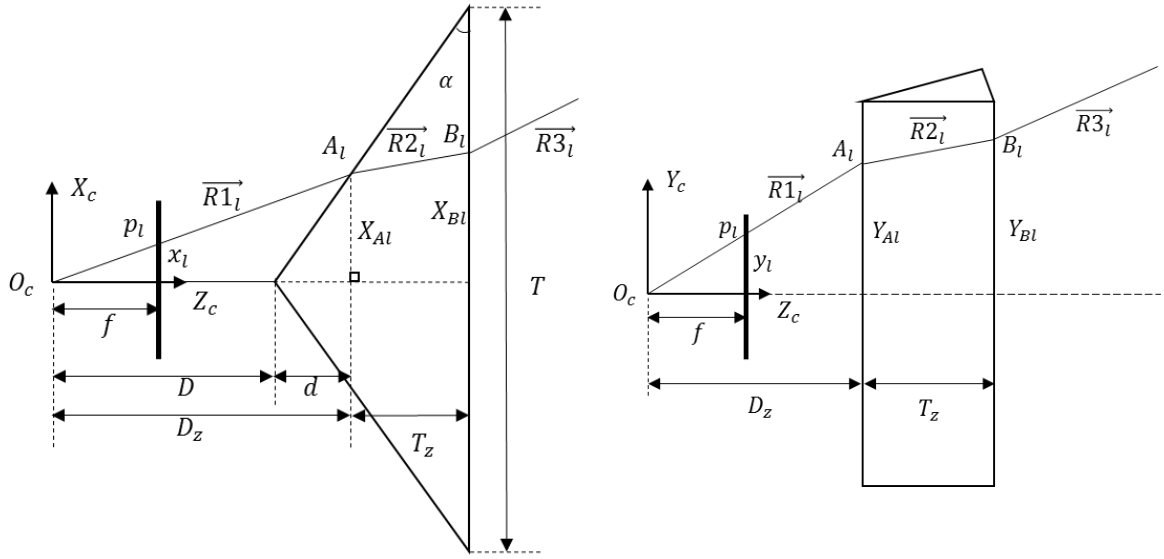
Some of the common functions could be expanded using Taylor series, with the arguments x:

$$\frac{1}{1-x} = \sum_{i=0}^{\infty} x^i \ for\ |x| < 1 \qquad \textit{Eq. C1}$$

$$(1+x)^{\alpha} = \sum_{i=0}^{\infty} C_i x^i \ for\ all\ |x| < 1, C_i = \prod_{k=1}^{i} \frac{\alpha - k + 1}{k} \qquad \textit{Eq. C2}$$

Here C_i is a constant value if α is given.

APPENDIX D. DETERMINATION OF THE POSITIONS OF THE REFRACTION POINTS



(a) Top view (b) Side view

Fig. D1 Determination of the positions of the refraction points A_l and B_l according to the (a) top view and (b) side view.

As shown in Fig. D1, with the given system parameters (f, D, x_l, y_l, T and α), the positions of A_l and B_l could be calculated.

1. To determine the position of A_l

From the top view, we have

$$\frac{X_{Al}}{x_l} = \frac{D+d}{f} \qquad \textit{Eq. D1}$$

We also have

$$\frac{d}{X_{Al}} = \tan(\alpha) \qquad \textit{Eq. D2}$$

Combining Eq. D1 and Eq. D2, we can obtain:

$$X_{Al} = \frac{D\ x_l/f}{1 - \tan(\alpha)\ x_l/f} \qquad \textit{Eq. D3}$$

and

$$Z_{Al} = D_z = D + d = \frac{D}{1 - \tan(\alpha)\ x_l/f} \qquad \textit{Eq. D4}$$

From the side view, Y_{Al} could be obtained by

$$Y_{Al} = \frac{D_z}{f} y_l = \frac{D\ y_l/f}{1 - \tan(\alpha)\ x_l/f} \qquad \textit{Eq. D5}$$

Therefore, the position of the refraction point A_l w.r.t. the camera coordinate $\{O_c \rightarrow X_c, Y_c, Z_c\}$ is

$$\begin{cases} X_{A_l} = \dfrac{D\ x_l/f}{1 - \tan(\alpha)\ x_l/f} \\ Y_{A_l} = \dfrac{D\ y_l/f}{1 - \tan(\alpha)\, x_l/f} \\ Z_{A_l} = \dfrac{D}{1 - \tan(\alpha)\, x_l/f} \end{cases} \qquad \textit{Eq. D6}$$

2. To determine the position of B_l

As we know that $\overrightarrow{R2_l} = (r2_1, r2_2, r2_3)$ is a normalized vector, and B_l could be calculated by

$$B_l = A_l + \overrightarrow{R2_l}\, t \qquad \textit{Eq. D7}$$

here t is the distance between A_l and B_l.

The horizontal distance between A_l and B_l is T_z:

$$T_z = \tan(\alpha) \cdot \left(\frac{T}{2} - X_{A_l}\right) \qquad \textit{Eq. D8}$$

The relationship between t and T_z is

$$\frac{T_z}{t} = \frac{1}{r2_3} \qquad \textit{Eq. D9}$$

Therefore, by substituting Eq. D8 and Eq. D9 to Eq. D7, we have

$$B_l = A_l + \frac{\overrightarrow{R2_l}}{r2_3} \cdot \tan(\alpha) \cdot \left(\frac{T}{2} - X_{A_l}\right) \qquad \textit{Eq. D10}$$