Calibration of Dynamic Stereo and Binocular Stereo for Large Scale Objects

by

Yongbao Zhang

Supervised by

Professor Reinhard Klette

A Thesis

Submitted in Partial Fulfillment of the

Requirements of the Degree of

Master of Science in Computer Science

Department of Computer Science, CITR Tamaki

The University of Auckland

May 2001

Abstract

This thesis mainly deals with large scale objects of height around 1.5 meters. It uses a dynamic stereo model or a binocular stereo model to calibrate and reconstruct a 3D large scale object.

This thesis briefly introduces two basic calibration techniques, i.e. DLT calibration technique and Tsai's calibration technique. It then discusses and evaluates the impact of different factors on the calibration system, especially the size of an object, the distance between the camera and the calibration object, the number of calibration points and its distribution pattern, the method for finding the center of a calibration mark, the covered area by a calibration object in an image and the calibration error distribution around the object height.

The thesis extends Tsai's calibration method and builds up two mathematical models. One model is the dynamic stereo system (one camera + one turntable system). The other is the binocular stereo system. It then demonstrates the corresponding calibration methods and evaluates the accuracies of these two systems in reconstructing a 3D object respectively. Problems of the dynamic stereo system are also discussed.

From the results of evaluation, the accuracy of the binocular stereo system is much better than the dynamic stereo system.

Acknowledgements

I would like to thank my supervisor, Professor Reinhard Klette, for his original ideas, comments and suggestions during the development of this thesis. His support, encouragement and patience were invaluable and essential in bringing it to completion.

I also like to thank Associate Professors Georgy Gimel'farb and Robert Marshall for their suggestions and comments during the development of this thesis.

Furthermore, I wish to thank Yen Chen and Yuan-sheng Tsai, our project members, for their cooperation and support.

Thanks to James Harper, technical staff of CITR, for his technical support.

Thanks also to Brian Gaby for his construction of the open calibration cube.

Finally, I would like to thank Yating Xu, my wife, and Peiyuan Zhang, my daughter for their support and encouragement.

Contents

CHAPTER 1

Introduction

Camera calibration in the context of three dimensional (3D) machine vision is the process of determining the internal camera geometric and optical characteristics (intrinsic parameters) and the 3D position and orientation of the camera frame relative to a certain world coordinate system (extrinsic parameters).

This thesis deals with large scale objects and calibration of two models. One model is a dynamic stereo system (one camera + one turntable system). The other model is a binocular stereo system.

1.1 Purpose of this Thesis

There are many camera calibration methods, of which the most famous are DLT (Direct Linear Transform) method and Tsai's calibration method. DLT method is the simplest camera calibration method whilst Tsai's method has better accuracy. The above two camera calibration methods can only transfer 3D world coordinates to image coordinates or transfer image coordinates to coplanar world coordinates.

For a camera calibration, the different working situations may also affect the accuracy of camera calibration results. One of the purposes of this thesis is to find out which impacts some factors may have on a camera calibration, such as the size of calibration object, the distance between a camera and a object, the number of calibration marks and its distribution pattern, the methods for finding the center of a calibration mark, the covered area by a calibration object and the calibration error distribution around the object height.

For a 3D object reconstruction, we need to reconstruct a 3D point in a world coordinate system. The above camera calibration method is not enough. We need to build up a stereo system and to extend Tsai's camera calibration method to calibrate a stereo system, not just a single camera, which is one of the purposes.

Another purpose of this thesis is to evaluate two stereo systems we build up and find out the accuracies of these stereo systems. The problems of these stereo system are also discussed.

1.2 Related Works

Images taken with wide-angle cameras tend to have severe distortions which pull points towards the optical center. Rahul [9] proposes a simple method for recovering the distortion parameters without calibration objects. Since distortions cause straight lines on the scene as curves in the image, the proposed algorithm seeks to find the distortion parameters that map the image curves to straight lines. The user selects a small set of points along the image curves. Recovery of the distortion parameters is formulated as the minimization of an objective function which is designed to explicitly account for noise in the selected image points. Experimental results are presented for synthetic data as well as real images. The paper [9] also presents the idea of a polycamera which is defined as a tightly packed camera cluster. Possible configurations are proposed to capture very large fields of view. Such camera clusters tend to have a nonsingle viewpoint. This paper therefore provides analysis of what the authors call the minimum working distance for such clusters. Finally, this paper presents results for a polycamera consisting of four wide-angle sensors having a minimum working distance of about 4 m. On undistorting the acquired images using the proposed technique, the real time high resolution panoramas are created.

Zhengyou [10] proposes a flexible new technique to easily calibrate a camera. It only requires the camera to observe a planar pattern shown at a few (at least two) different orientations. Either the camera or the planar pattern can be freely moved. The motion needs not to be known. Radial lens distortion is modeled. The proposed procedure consists of a closed-form solution, followed by a nonlinear refinement based on the maximum likelihood criterion. Both computer simulation and real data have been used to test the proposed technique and very good results have been obtained. Compared with classical techniques which use expensive equipment such as two or three orthogonal planes, the proposed technique is easy to use and flexible. It advances 3D computer vision one more step from laboratory environment to real world use.

Current calibration methods typically assume that the observations are unbiased, the only error is the zero-mean independent and identically distributed random noise in the observed image coordinates, and the camera model completely explains the mapping between the 3D coordinates and the image coordinates. In general, these conditions are not met, causing the calibration results to be less accurate than expected. Janne [11] proposes a calibration procedure for precise 3D computer vision applications. It introduces bias correction for circular control points and a nonrecursive method for reversing the distortion model. The accuracy analysis is presented and the error sources that can reduce the theoretical accuracy are discussed. The tests with synthetic images indicate improvements in the calibration results in limited error conditions. In real images, the suppression of external error sources becomes a prerequisite for successful calibration.

1.3 Thesis Organization

Chapter 2 briefly introduces the basic calibration techniques, i.e. Direct Linear Transformation calibration method and Tsai's calibration method. It also demonstrates the use of these calibration techniques.

Chapter 3 evaluates and discusses the impact of some factors in the calibration system, such as the size of the calibration object, the distance between the camera and the calibration object, the number of calibration marks and its distribution pattern, the different method to find the center of a calibration mark, the covered area by a calibration object and the calibration error distribution around the object height.

Chapter 4 builds up a new mathematical model for the dynamic stereo system (one camera + one turntable system). It also compares this new model to the aligned turntable model mentioned in the textbook [1]. Furthermore, this chapter demonstrates the use of method for this new calibration model.

Chapter 5 demonstrates the use of method for the evaluation of dynamic stereo system. It also evaluates this dynamic stereo system. According to the evaluation result, this chapter discusses the problems of this dynamic system.

Chapter 6 builds up a mathematical model for a binocular stereo system and also demonstrates the method for the reconstruction of a 3D object. Furthermore, this chapter evaluates the accuracy of this binocular stereo system.

Chapter 7 introduces the software we developed for both of these two calibration systems we build. It demonstrates the applications running environment and the structures of applications. It also demonstrates the use of method in the context PSM based sharp recovery for the dynamic stereo system and binocular stereo system.

Chapter 8 gives an overall conclusion for this thesis.

CHAPTER 2

Basic Calibration Techniques

A calibration technique is based on known 3D space coordinates (e.g. in the world coordinate system) of geometrically configured points (calibration points) which are physically realized by marks (calibration marks) on a certain calibration object.

There are many different calibration techniques. The simplest one is DLT (direct linear transform) method which is easily understood. Tsai's method is the most widely used calibration technique. These two methods are basic calibration techniques in the computer vision field.

2.1 Direct Linear Transform Method

The Direct Linear Transform Method (DLT method) was suggested by Y.I. Abdel-Aziz and H.M. Karara (1971)[4]. It has the advantage that it is the simplest model and is easily understood.

Figure 2.1 shows the relations between world coordinates, camera coordinates and image coordinates. If there is a point P that is (X_w, Y_w, Z_w) in the world coordinate system, it is (X_k, Y_k, Z_k) in the camera system. It is clear that a translation operation plus a rotation operation on (X_w, Y_w, Z_w) will transform it to (X_k, Y_k, Z_k) in the camera coordinate system. The trans-

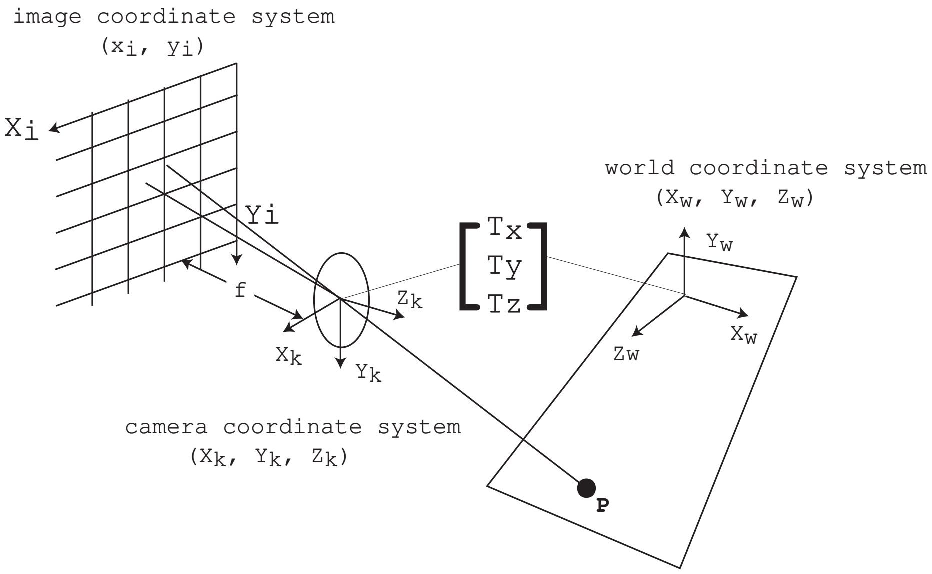

Figure 2.1: Relationships of world coordinates, camera coordinates and image coordinates

formation relations are as follows:

$$\begin{pmatrix} X_k \\ Y_k \\ Z_k \end{pmatrix} = \begin{pmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} \tag{2.1}$$

In this equation, r_i are the coefficients of the rotation matrix $R = R_x \cdot R_y \cdot R_z$ of the Euclidean transformation (world into camera coordinates). T_x, T_y and T_z are the coefficients of the translation vector T.

Central projection or orthogonal parallel projection can be assumed for the transformation of camera coordinates into image coordinates.

If the orthogonal parallel projection model is used, the relation is as follows:

$$\begin{pmatrix} x_i - c_x \\ y_i - c_y \end{pmatrix} = \begin{pmatrix} X_k \\ Y_k \end{pmatrix} \tag{2.2}$$

If the central projection model is used, the relation is then as follow:

$$\begin{pmatrix} x_i - c_x \\ y_i - c_y \end{pmatrix} = -\frac{f}{Z_k} \begin{pmatrix} X_k \\ Y_k \end{pmatrix} \tag{2.3}$$

From Equ. 2.1, we can obtain the following equations:

$$\begin{pmatrix} x_i - c_x \\ y_i - c_y \end{pmatrix} = \frac{-f}{r_7 X_w + r_8 Y_w + r_9 Z_w + T_z} \begin{pmatrix} r_1 X_w + r_2 Y_w + r_3 Z_w + T_x \\ r_4 X_w + r_5 Y_w + r_6 Z_w + T_y \end{pmatrix} \tag{2.4}$$

In the above equations, (c_x, c_y) is the coordinate of the principal point (i.e. the intersection point of the optical axis with the image plane) according to the image coordinate system. The parameter f is the effective focal length of the lens (the distance between the lens center and image plane).

In most of cases, the central project model should be used. The calibration procedure using DLT method is explained in the following subsections.

2.1.1 World Coordinates and Image Coordinates

A picture of the calibration object, i.e. the marks of the geometric configured calibration points is taken. The geometric configuration of these points has to be known. It can be done through constructing a world coordinate system on the calibration object, e.g. if the calibration object is an open cube, the three intersection lines of the three planes of the open cube can be used as the axes of the world coordinate system.

The device dependent row and column positions, i.e. the positions of the calibration points in the image coordinate system (x_b, y_b), are determined for every calibration point having a visible projected mark in the image.

2.1.2 Eleven Transformation Parameters

Equation 2.4 can be rewritten into the following form:

$$M \cdot \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_{11} \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \end{pmatrix} \tag{2.5}$$

where

$$M = \begin{pmatrix} X_w & Y_w & Z_w & 1 & 0 & 0 & 0 & 0 & -x_iX_w & -x_iY_w & -x_iZ_w \\ 0 & 0 & 0 & 0 & X_w & Y_w & Z_w & 1 & -x_iX_w & -x_iY_w & -x_iZ_w \end{pmatrix}$$

It follows that

$$\begin{array}{lll} L_1 = \frac{r_7c_x - fr_1}{T_z} & L_2 = \frac{r_8c_x - fr_2}{T_z} & L_3 = \frac{r_9c_x - fr_3}{T_z} \\ L_4 = \frac{T_zc_x - fT_x}{T_z} & L_5 = \frac{r_7c_x - fr_4}{T_z} & L_6 = \frac{r_8c_x - fr_5}{T_z} \\ L_7 = \frac{r_9c_x - fr_6}{T_z} & L_8 = \frac{T_zc_y - fT_y}{T_z} & L_9 = \frac{r_7}{T_z} \\ L_{10} = \frac{r_8}{T_z} & L_{11} = \frac{r_9}{T_z} & \end{array}$$

For a mathematically unambiguous solution, at least six calibration points are required to solve the above equation with respect to the eleven unknown transformation parameters $L_1, \ldots, L_{11}$. In general, six points or more lead to an over-determined system of linear equations. More points are also required with respect of stability of the calibrated solution.

The eleven parameters of transformation vector can be solved by using the pseudo-inverse technique [6].

2.1.3 Principal Point and T_z

The eleven parameters of transformation vector have been solved in the above step. These parameters "contain" the intrinsic and extrinsic parameters.

The image coordinates of principal point can be determined by following equations:

$$\begin{pmatrix} c_x \\ c_y \end{pmatrix} = \frac{1}{L_9^2 + L_{10}^2 + L_{11}^2} \begin{pmatrix} L_1 & L_2 & L_3 \\ L_5 & L_6 & L_7 \end{pmatrix} \begin{pmatrix} L_9 \\ L_{10} \\ L_{11} \end{pmatrix} \tag{2.6}$$

The Z component of translation vector T can be obtained from following equation:

$$T_z = \sqrt{\frac{1}{L_9^2 + L_{10}^2 + L_{11}^2}} \tag{2.7}$$

In most of cases, the origin of world coordinate system should be in front of the lens, i.e. T_z is positive. Otherwise, T_z would be negative.

2.1.4 Focal Length

The effective focal length f can be calculated by using any of the following two equations:

$$f = \sqrt{\frac{L_1^2 + L_2^2 + L_3^2}{L_9^2 + L_{10}^2 + L_{11}^2} - c_x^2} \tag{2.8}$$

$$f = \sqrt{\frac{L_5^2 + L_6^2 + L_7^2}{L_9^2 + L_{10}^2 + L_{11}^2} - c_y^2} \tag{2.9}$$

With respect to the accuracy of calibration, we should use the average of both equation results as the final value of focal length f.

2.1.5 Translation Vector and Rotation Matrix

T_z of translation vector has been calibrated. T_x and T_y can be calibrated from the following equation:

$$\begin{pmatrix} T_x \\ T_y \end{pmatrix} = \frac{T_z}{f} \begin{pmatrix} c_x - L_4 \\ c_y - L_8 \end{pmatrix} \tag{2.10}$$

The coefficients r_7, r_8 and r_9 of rotation matrix are determined by following equation:

$$\begin{pmatrix} r_7 \\ r_8 \\ r_9 \end{pmatrix} = T_z \begin{pmatrix} L_9 \\ L_{10} \\ L_{11} \end{pmatrix} \tag{2.11}$$

The rest of rotation matrix coefficients can then be calculated:

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{pmatrix} = \frac{1}{f} \begin{pmatrix} r_7 c_x - L_1 T_z \\ r_8 c_x - L_2 T_z \\ r_9 c_x - L_3 T_z \\ r_7 c_y - L_5 T_z \\ r_8 c_y - L_6 T_z \\ r_9 c_y - L_7 T_z \end{pmatrix} \tag{2.12}$$

2.2 Tsai's Method

There are a number of techniques for geometric calibration of CCD cameras, which are known from the computer vision literature. Tsai's calibration method was suggested by R.Y. Tsai in 1986 [5]. This calibration method is widely used because its accuracy is good enough for most applications and there is a fully developed implementation software provided on the internet [2].

Unlike the DLT method described in Section 2.1, the Tsai's calibration method also includes the determination of the lens distortion coefficients and mapping sensor elements to image buffer matrix. The method requires at least seven non-coplanar, accurately detected calibration points which were given in any arbitrary but known geometric configuration.

Radial lens distortion generally occurs with common cameras. Figure 2.2 shows the influence of this distortion on the acquired image. The effects of

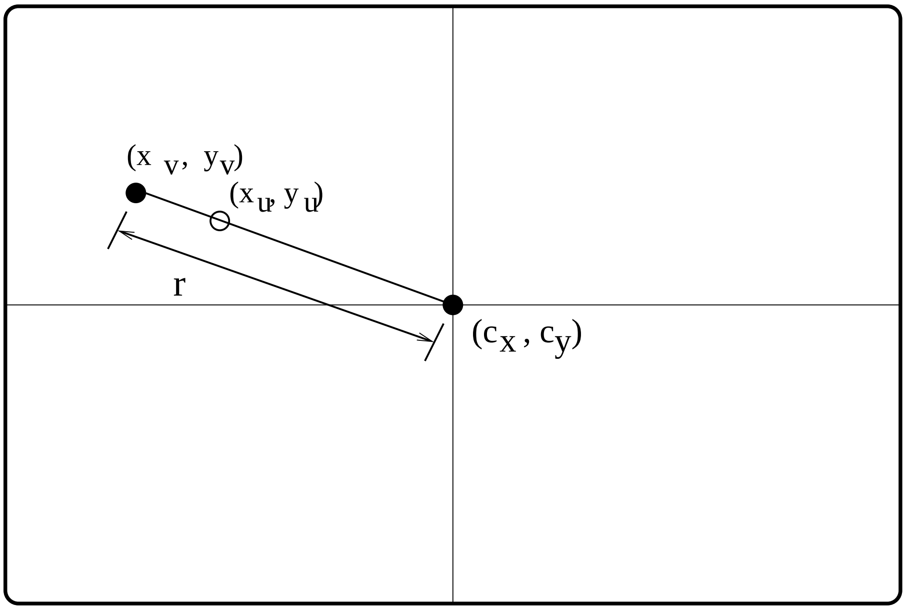

Figure 2.2: Deformation of ideal image by radial lens distortion.

the radial lens distortion can be described mathematically. However, an ideal modeling of the lens distortion leads to an infinite number of distortion coefficients. In practice, only the first two coefficients are needed to be considered because it is good enough for the approximation of these distortions. The valid, i.e. distorted projected image coordinates (x_v, y_v) can be determined from undistorted image coordinates (x_u, y_u) by using the following equations:

$$\begin{pmatrix} x_v \\ y_v \end{pmatrix} = \begin{pmatrix} x_u \\ y_u \end{pmatrix} - \begin{pmatrix} D_x \\ D_y \end{pmatrix} \tag{2.13}$$

with $D_x = x_v \cdot (k_1 r^2 + k_2 r^4)$, $D_y = y_v \cdot (k_1 r^2 + k_2 r^4)$ and $r = \sqrt{x_v^2 + y_v^2}$.

The sensors of CCD cameras are not perfect squares, so this has to be considered in the mapping of image coordinates to image buffer matrix. Fig-

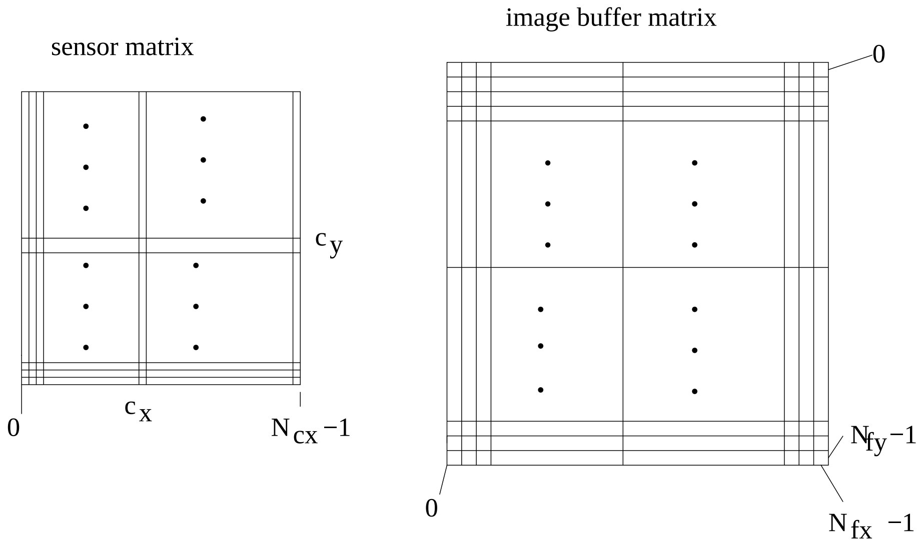

Figure 2.3: Conversion of principal point centered sensor coordinates into digital image buffer coordinates

ure 2.3 shows the following relations:

$$\begin{pmatrix} x_b \\ y_b \end{pmatrix} = \begin{pmatrix} \frac{s_x x_v}{d'_x} \\ \frac{y_v}{d_y} \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix} \tag{2.14}$$

In the above equations, the coordinates (c_x, c_y) specify the principal point. The parameter N_{cx} is the number of sensor elements in one image row. The parameter N_{fx} is the number of pixels in one row of the final digital image. The parameter $d'_x = d_x \frac{N_{cx}}{N_{fx}}$ is the horizontal distance between two image pixels. The parameter d_x is the horizontal distance between two sensor elements. The parameter d_y is the vertical distance between two sensor elements. The parameter s_x is the horizontal scaling factor.

The scaling factor s_x has to be calibrated. The parameters d_x, d_y, N_{cx}, and N_{fx} are normally included in the manufacturer's data sheets for the camera and the image digitization unit (frame grabber).

Nowadays the CCD sensors are more likely to be squares. So the conversion of the principal point centered sensor coordinates into digital image buffer coordinates is not a necessary consideration.

Tsai's calibration method can be explained as a procedure consisting of the following seven processing steps as explained in [1]. The following explanations are based on a non-coplanar case.

2.2.1 Sensor Coordinates

The position (0,0) is defined to be located in the top-left corner of the image and it is in a right-hand coordinate system. This ensures positive image buffer coordinates. The number of sensor elements in one image row is represented by N_{cx}. The number of pixels in one image row is represented by N_{fx}. The horizontal distance between two sensor elements d'_x and the vertical distance between two sensor elements d_y can be read in the manufacturer's data sheets.

The principal point (c_x, c_y) is assumed to be at the image center so it can be obtained from calculation. The corresponding sensor coordinates (x_s, y_s) are calculated for all (centers of) calibration marks by the following formulas:

$$x_s = \frac{(x_b - c_x) \cdot d_x\prime}{s_x} \quad (2.15)$$

$$y_s = (y_b - c_y) \cdot d_y \quad (2.16)$$

The scaling factor s_x is initially set to the value one and it will be determined explicitly later in Sec. 2.2.4.

2.2.2 Seven Transformation Parameters

In the camera coordinates system, the projection vector of the calibration points (X_k, Y_k, Z_k) onto the $X_k - Y_k$ plane parallel to the corresponding image vector (x_s, y_s). Their outer product should be zero. It holds:

$$(X_k, Y_k) \times (x_s, y_s) = 0 \tag{2.17}$$

Using Equ. 2.1, the projection of marks of calibration points (X_w, Y_w, Z_w) in the world coordinate system onto corresponding image points (x_s, y_s) in sensor coordinates is characterized by the following linear equation. It holds:

$$x_s = (y_s X_w, y_s Y_w, y_s Z_w, y_s, -x_s X_w, -x_s Y_w, -x_s Z_w) \cdot L \tag{2.18}$$

with $L = [\frac{r_1 s_x}{T_y}, \frac{r_2 s_x}{T_y}, \frac{r_3 s_x}{T_y}, \frac{T_x s_x}{T_y}, \frac{r_4}{T_y}, \frac{r_5}{T_y}, \frac{r_6}{T_y}]^T$ for $T_y \neq 0$.

The parameters T_x and T_y denote the components of the translation vector T, and the r_i values represent the elements of the rotation matrix R. The use of more than seven calibration points leads to an over-determined set of equations which can be solved for L using the pseudo-inverse technique (Moore-Penrose inverse) as following:

$$L = (M^T M)^{-1} M^T X \tag{2.19}$$

with

$$M = \begin{bmatrix} y_{s1}Xw1 & y_{s1}Y_{w1} & y_{s1}Z_{w1} & y_{s1} & -x_{s1}X_{w1} & -x_{s1}Y_{w1} & -x_{s1}Z_{w1} \\ y_{s2}Xw2 & y_{s2}Y_{w2} & y_{s2}Z_{w2} & y_{s2} & -x_{s2}X_{w2} & -x_{s2}Y_{w2} & -x_{s2}Z_{w2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{sn}Xwn & y_{sn}Y_{wn} & y_{sn}Z_{wn} & y_{sn} & -x_{sn}X_{wn} & -x_{sn}Y_{wn} & -x_{sn}Z_{wn} \end{bmatrix}$$

$and \quad X = (x_{s1}, x_{s2}, \cdots, x_{sn})^T,$ for n calibration points.

2.2.3 Y-component of Translation Vector

The components of the solution vector L will be described in the following by the abbreviations a_1 to a_7:

$$
\begin{aligned}
a_1 &= \frac{r_1 s_x}{T_y} \\
a_2 &= \frac{r_2 s_x}{T_y} \\
a_3 &= \frac{r_3 s_x}{T_y} \\
a_4 &= \frac{T_x s_x}{T_y} \\
a_5 &= \frac{r_4}{T_y} \\
a_6 &= \frac{r_5}{T_y} \\
a_7 &= \frac{r_6}{T_y}
\end{aligned}
\tag{2.20}
$$

All parameters a_1 to a_7 were already determined at Subsection 2.2.2. From the orthonormality property of R, the relation $r_4^2 + r_5^2 + r_6^2 = 1$ holds. Therefore:

$$(a_5^2 + a_6^2 + a_7^2) = T_y^{-2} \cdot (r_4^2 + r_5^2 + r_6^2) = T_y^{-2} \tag{2.21}$$

Thus, the value of T_y can be calculated by

$$\mid T_y \mid = \frac{1}{\sqrt{a_5^2 + a_6^2 + a_7^2}} \tag{2.22}$$

For determining the sign of T_y, an object point $P = (X_w, Y_w, Z_w)$ is chosen such that its image position (x_b, y_b) lies as far away from the principal point as possible. Firstly the sign of T_y is assumed to be positive. Assume s_x is one and using Equ. 2.20, the values of r_1, r_2, r_3, r_4, r_5, r_6 and T_x can be calculated. Then the position can also be calculated by the following equation:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \end{pmatrix} \tag{2.23}$$

Compare x and x_s as well as y and y_s, if they have the same signs, then T_y is positive, otherwise the sign of T_y is negative.

2.2.4 Scaling Factor

Because of the orthonormality property of R, $r_1^2 + r_2^2 + r_3^2 = 1$ holds. Using Equ. 2.20, the following relation can be obtained:

$$\sqrt{a_1^2 + a_2^2 + a_3^2} \cdot \mid T_y \mid = s_x \cdot \sqrt{r_1^2 + r_2^2 + r_3^2} = s_x \tag{2.24}$$

Thus, the scaling factor can be determined by

$$s_x = \sqrt{a_1^2 + a_2^2 + a_3^2} \cdot \mid T_y \mid \tag{2.25}$$

2.2.5 Rotation Matrix and X-component of Translation Vector

From Equ. 2.20, the components of rotation matrix $r1$ to $r6$ and the X-component of translation vector can be calculated by the following equations:

$$r_1 = a_1 \cdot \frac{T_y}{s_x} \tag{2.26}$$

$$r_2 = a_2 \cdot \frac{T_y}{s_x} \tag{2.27}$$

$$r_3 = a_3 \cdot \frac{T_y}{s_x} \tag{2.28}$$

$$r_4 = a_5 \cdot T_y \tag{2.29}$$

$$r_5 = a_6 \cdot T_y \tag{2.30}$$

$$r_6 = a_7 \cdot T_y \tag{2.31}$$
$$T_x = \frac{a_4 \cdot T_y}{s_x} \tag{2.32}$$

The still missing components of rotation matrix r_7, r_8 and r_9 are calculated with the inner vector product of the first two rows of the rotation matrix R.

$$r_7 = r_2 \cdot r_6 - r_3 \cdot r_5 \tag{2.33}$$
$$r_8 = r_3 \cdot r_4 - r_1 \cdot r_6 \tag{2.34}$$
$$r_9 = r_1 \cdot r_5 - r_2 \cdot r_4 \tag{2.35}$$

2.2.6 Approximation of Focal Length and Z-component of Translation Vector

If we ignore the radial lens distortion, i.e. let the coefficients k_1 and k_2 be as zero, then the distortion coordinates are the same as the undistorted coordinates. It holds:

$$y_s = \frac{fY_k}{Z_k} \tag{2.36}$$

Using Equ. 2.1, the linear equation

$$\begin{pmatrix} y & -y_s \end{pmatrix} \begin{pmatrix} f \\ T_z \end{pmatrix} = w \cdot y_s \tag{2.37}$$

is formulated for every calibration point with

$$\begin{pmatrix} y \\ w \end{pmatrix} = \begin{pmatrix} r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + \begin{pmatrix} T_y \\ 0 \end{pmatrix} \tag{2.38}$$

Assuming more than two calibration points can be used. For the unknown

f and T_z, using the pseudo-inverse technique, it holds:

$$\begin{pmatrix} f \\ T_z \end{pmatrix} = (M^T M)^{-1} M^T X \tag{2.39}$$

where M and X are calculated using n calibration points with

$$M = \begin{pmatrix} y_1 & -y_{s1} \\ y_2 & -y_{s2} \\ \vdots & \vdots \\ y_n & -y_{sn} \end{pmatrix}$$

$$X = \begin{pmatrix} w_1 y_{s1} \\ w_2 y_{s2} \\ \vdots \\ w_n y_{sn} \end{pmatrix}$$

2.2.7 Calculation of Focal Length, Z-component of Translation Vector and Radial Distortion Coefficients

The utilization of a standard optimization technique allows more accurate calculation of the camera constant f, of the depth T_z and of the distortion coefficients k_1 and k_2. The already determined approximations of f and T_z act as the starting values. Zero is assumed as the initial value for the radial distortion coefficients.

The steepest descent method is used for optimization. The equations

$$x_{u1} = f \cdot \frac{r_1 X_w + r_2 Y_w + r_3 Z_w + T_x}{r_7 X_w + r_8 Y_w + r_9 Z_w + T_z} \tag{2.40}$$

$$y_{u1} = f \cdot \frac{r_4 X_w + r_5 Y_w + r_6 Z_w + T_y}{r_7 X_w + r_8 Y_w + r_9 Z_w + T_z} \tag{2.41}$$

can be used for determination of the parameters. These equations describe the transformation of a world coordinate point into the camera-centered coordinate system assuming perspective projection in between.

Projected ideal image points can also be generated by a radial rectification of the actually projected, valid image points (x_v, y_v). Therefore:

$$x_{u2} = x_v \cdot (1 + k_1 r^2 + k_2 r^4) \tag{2.42}$$

$$y_{u2} = y_v \cdot (1 + k_1 r^2 + k_2 r^4) \tag{2.43}$$

where $r = \sqrt{x_v^2 + y_v^2}$

For a search in a four-dimensional vector space, the error function

$$\varepsilon(k_1, k_2, f, T_z) = \sqrt{(x_{u2} - x_{u1})^2 + (y_{u2} - y_{u1})^2} \tag{2.44}$$

arises from the Euclidean distance between a pair of ideal image points that were calculated in different ways. The gradient of ε is calculated with the partial derivatives:

$$\frac{\partial \varepsilon}{\partial k_1} = \frac{r^2 \cdot ((x_{u2} - xu1) \cdot x_v + (y_{u2} - y_{u1}) \cdot y_v)}{\varepsilon} \tag{2.45}$$

$$\frac{\partial \varepsilon}{\partial k_2} = \frac{r^4 \cdot ((x_{u2} - xu1) \cdot x_v + (y_{u2} - y_{u1}) \cdot y_v)}{\varepsilon} \tag{2.46}$$

$$\begin{aligned} \frac{\partial \varepsilon}{\partial f} &= -\frac{(x_{u2} - xu1) \cdot (r_1 X_w + r_2 Y_w + r_3 Z_w + T_x)}{\varepsilon \cdot (r_7 X_w + r_8 Y_w + r_9 Z_w + T_z)} \\ &\quad -\frac{(y_{u2} - yu1) \cdot (r_4 X_w + r_5 Y_w + r_6 Z_w + T_y)}{\varepsilon \cdot (r_7 X_w + r_8 Y_w + r_9 Z_w + T_z)} \end{aligned} \tag{2.47}$$

$$\begin{aligned} \frac{\partial \varepsilon}{\partial T_z} &= \frac{f \cdot (x_{u2} - xu1) \cdot (r_1 X_w + r_2 Y_w + r_3 Z_w + T_x)}{\varepsilon \cdot (r_7 X_w + r_8 Y_w + r_9 Z_w + T_z)^2} \\ &\quad +\frac{f \cdot (y_{u2} - yu1) \cdot (r_4 X_w + r_5 Y_w + r_6 Z_w + T_y)}{\varepsilon \cdot (r_7 X_w + r_8 Y_w + r_9 Z_w + T_z)^2} \end{aligned} \tag{2.48}$$

This gradient allows to specify the direction of the steepest increase of the error function ε. The optimum has to be analyzed contrary to the direction of the gradient because a minimum error is needed.

The determination of the principal point (c_x, c_y) is also possible with this optimization method if the image point in the sensor coordinates is replaced by

$$\begin{aligned} x_s &= \frac{d'_x \cdot (x_b - c_x)}{s_x} \\ y_s &= d_y \cdot (y_b - c_y) \end{aligned}$$

in this error function ε, and if the gradient $\nabla\varepsilon$ is adjusted accordingly.

CHAPTER 3

Camera Calibration for Large Scale Objects

There are many factors which may affect the camera calibration result. In principle, the calibration technique should consider every factor which may affect the calibration result, but it is impossible in practice. We have to ignore some factors which stay unchanged during the usage of calibration or which only cause a small error on the calibration result.

In fact, every known calibration technique only considers several limited factors. The calibration technique which considers more factors results in better accuracy than the one which considers less factors, e.g. Tsai's calibration method is more accurate than DLT's calibration method because Tsai's method does consider the lens distortion and the mapping of sensor coordinates to image buffer coordinates and DLT's method does not. This result is well-known and needs not be discussed anymore.

In this chapter, we will discuss some factors which are not considered in every calibration technique but do affect the calibration results. We will discuss their impacts and find out the best situation for calibration of large scale objects. The calibration object we deal with is 1.5 meter in height. The following discussions are based on Tsai's calibration method.

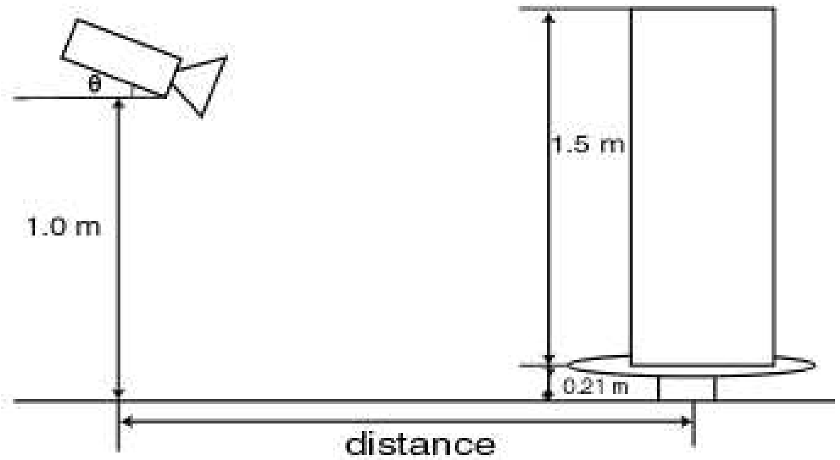

Figure 3.1: Setup of camera and calibration object

3.1 Setup of Working Environment

The camera we used is JVC KY-F55BE CCD camera and it is positioned somewhere as shown in Fig. 3.1. Figure 3.1 shows the related positions about camera, calibration object and turntable. The camera's specifications can be found in the web site [3]. It has $752(H) \times 582(V)$ pixels and $4.8mm \times 3.6mm$ CCD size. The frame grabber we used has $768(H) \times 576(V)$ pixels.

According to the above information, we set up the camera initial parameters as in Tab. 3.1.

Table 3.1 shows the initial values of our camera. N_{cx} is the number of sensor elements in one image row. N_{fx} is the number of pixels in one row of the frame grabber window. d_x is the the horizontal distance between two CCD sensor elements. d_y is the vertical distance between two CCD sensor elements. d_{px} is the horizontal distance between two image pixels. d_{py} is the

Camera Parameter	Value
N_{cx}	752 [sel]
N_{fx}	768 [pix]
d_x	4.8/752 [mm/sel]
d_y	3.6/582 [mm/sel]
d_{px}	$d_x \cdot N_{cx}/N_{fx}$ [mm/pix]
d_{py}	d_y [mm/pix]
C_x	768/2 [pix]
C_y	576/2 [pix]
S_x	1.0

Table 3.1: Initial camera parameters

vertical distance between two image pixels. C_x is the X-coordinate of the image center point. C_y is the Y-coordinate of the image center point. S_x is the horizontal scaling factor.

The video control panel is set up as in Fig. 3.2.

The world coordinate system used should be the right hand system. The world coordinate axes on the open calibration cube should be set up as Fig. 3.3.

The image coordinate system should also use the right hand system even though it only has two image coordinate axes. The image coordinate axes on the image window should be set up as Fig. 3.4.

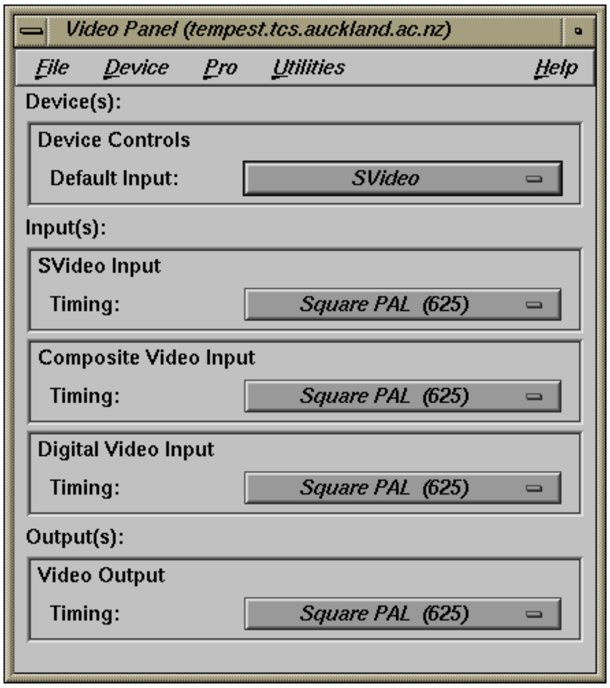

Figure 3.2: Camera control panel

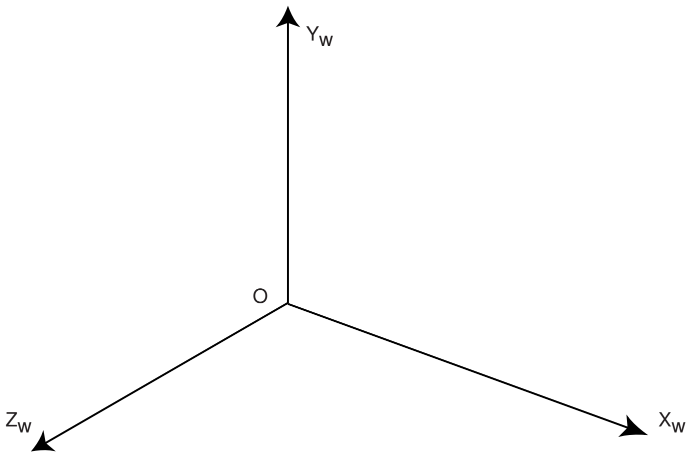

Figure 3.3: Setting of world coordinate system

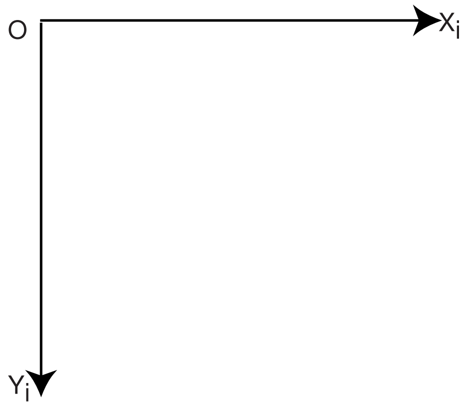

Figure 3.4: Setting of image coordinate system

3.2 Error on the Implementation Software of Tsai's Calibration Method

The implementation software of Tsai's calibration method we used can be downloaded from the website [2]. This software is a freeware and is widely used in the computer version field and robotics field.

There is a minor error in this software. In function *ncc_compute_Tx_and_Ty*() of file *cal_main.c*, T_x should not be calculated here. It should be calculated after the calibration of S_x and the formula should be Equ. 2.32 rather than $T_x = a_4 \cdot T_y$.

There is a corresponding error in the original paper of Tsai's calibration method. On line 2 of right column of page 333 in [8], the equation of T_x should be Equ. 2.32 and it misses S_x.

In order to correct this error, we calculate T_x after calculation of S_x and put it in function *ncc_compute_sx*() of file *cal_main.c*.

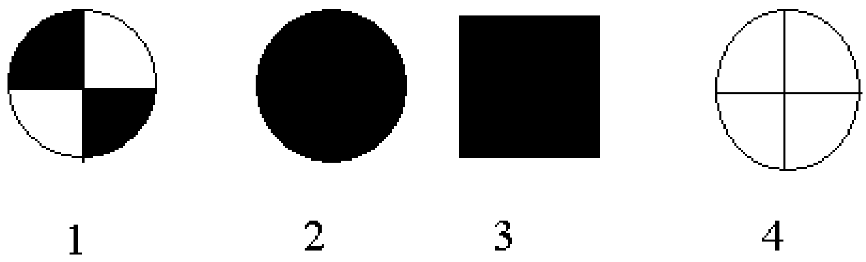

Figure 3.5: Examples of marks used in camera calibration

3.3 Different Methods for Finding Center of Calibration Mark

No matter what kind of calibration technique is used, we have to mark several marks on the calibration object and determine the center points of the marks on the projected image plane. The determination method will directly affect the correctness of the center point positions of the marks.

There are a lot of marks which can be used for camera calibration purpose. Figure 3.5 shows the examples of marks used in the camera calibration. These four marks are the most often used in the camera calibration. There are three ways often used to determine the calibration points.

The first method is to determine the center points of the marks by calculation of the moments of the mark's pixels. The formulas are as follow:

$$x = \frac{\Sigma x_i p_i}{\Sigma p_i} \tag{3.1}$$

$$y = \frac{\Sigma y_i p_i}{\Sigma p_i} \tag{3.2}$$

In the above formulas, x_i and y_i are the coordinates of every pixel in the calibration mark. p_i is the pixel value of (x_i, y_i).

From our experiences, the brightness of lighting will directly affect the accuracy of the result. If the image is too bright, the calculated moments are not accurate. If the image is too dark, it is hard to distinguish the mark and the background. The best situation of brightness is that the background pixel value is about $210 \sim 254$ in case of maximum pixel value is 255.

This method is the simplest determination method and has a good accuracy for most situations. It can use the mark 1, mark 2 and mark 3 of Fig. 3.5 as calibration marks.

The second method is to determine the vertex points of calibration marks by searching through the whole mark. It can use the mark 1 and mark 3 of Fig. 3.5 as the calibration marks. There are two weaknesses for this determination method. One is that the search algorithm is hard to define and the algorithm will not be applicable for every case. Another is the accuracy will be seriously affected by the threshold technique necessary used for preprocessing the image.

The third method is to determine the center points of calibration marks by determination of the cross point of two straight lines. It can use the mark 1 and mark 4 of Fig. 3.5 as the calibration marks. The weakness of this method is that it can only be used for the coplanar case.

Figure 3.6 shows an image of calibration marks after preprocessing. We can see that the marks on the bottom plane are seriously distorted. In this case, we can not find the straight line on these marks and of course are unable to find the center points by using the third method.

In conclusion, there are serious distortions in the 3D non-coplanar case. We suggest it is better to use mark 2 of Fig. 3.5 as the calibration marks and use the first method to determine the center point of a calibration mark.

Figure 3.6: Image marks after preprocessing

3.4 Evaluation Methods

Camera and calibration object are setup as in Fig. 3.1. The distance between the camera and the calibration object is 2.8 meters and the camera angle θ is 10 degree. The camera focal length is 6.4 mm. We use 32 calibration points uniformly distributed as in Fig. 3.7 to calibrate the camera.

After calibration, there are four methods to evaluate the accuracy of calibration result. The first method is to evaluate errors on the image plane. We project the calibration points from the world coordinate system to image plane and calculate the distances between these points and their corresponding true projected points on the image plane as the errors on the image plane.

Figure 3.8 shows the evaluation result by using the first method, i.e. errors on the image plane. The horizontal axis is the calibration mark which indicates the geometric position shown in Fig. 3.7. The vertical axis is the errors on the image plane.

The second method is to evaluate the closest errors on the object plane. It calculates the minimum distance from the true point in the world coordinate space to the straight line passing through the corresponding image point and the principal point. It has the same pattern with the first method.

The third method is to evaluate the normalized calibration error which is proposed by Weng [12]. It evaluates the mean of the ratio of the lateral triangulation error to the lateral standard deviation of the pixel digitization noise at the depth.

The fourth method is to evaluate the errors on the object plane. It calculates the distance between the true point and the intersection point of the back-projection line of the test point and the plane $z = z_i$ of the true point. It represents the error on the world coordinate system.

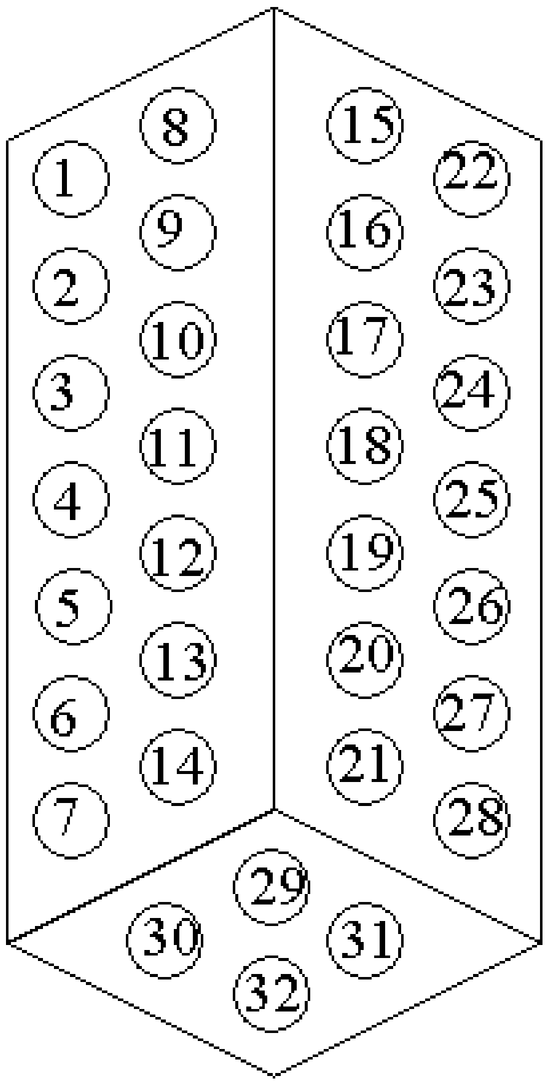

Figure 3.7: Calibration marks distribution

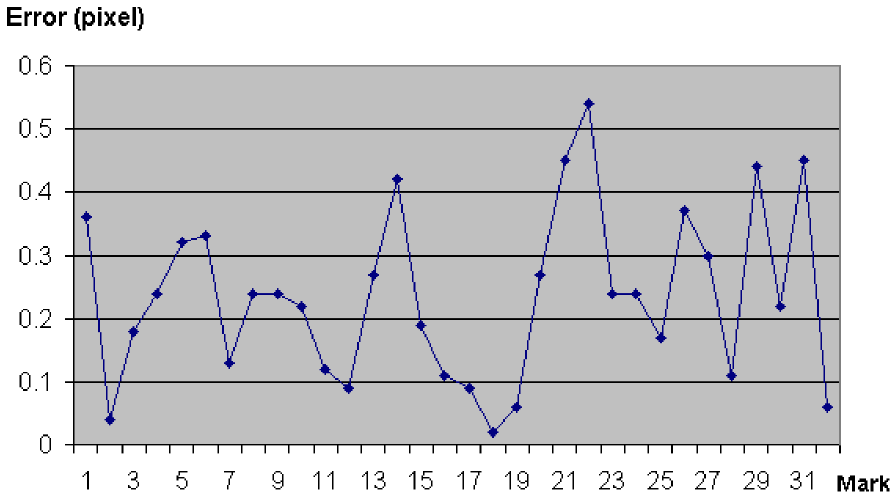

Figure 3.8: Errors on the image plane

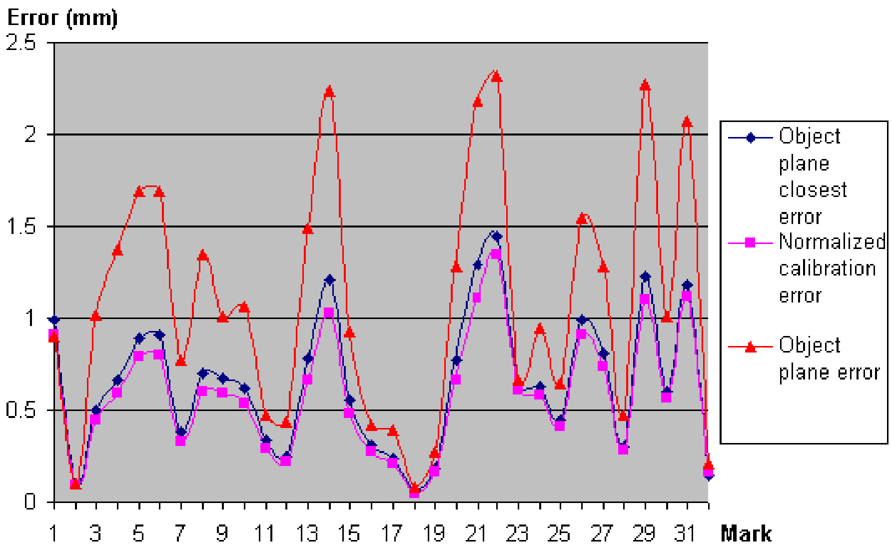

Figure 3.9: Errors on the object space

Figure 3.9 shows the evaluation result by using the second, third and forth methods, i.e. errors on the object space. The horizontal axis is the calibration mark which indicates the geometric position shown in Fig. 3.7. The vertical axis is the errors on the object space. We can see that these three methods have the similar error pattern and they just represent the different physical meaning.

From the above discussion, the four methods has similar error pattern but represent different physical meaning. In this chapter, camera calibration accuracy is evaluated. Therefore, the first evaluation method is chosen.

3.5 Calibration Points and Pattern

Tsai's calibration method requires at least seven calibration points. In theory, the more calibration points the less random error of calibration. In this section, we will find out the minimum calibration points and their distribution pattern. The camera and the calibration object are setup as in Fig. 3.1. The distance between the camera and the calibration object is 2.8 meters and the camera angle θ is 10 degree. The camera focal length is 6.4 mm.

Figure 3.10 shows the relationship between the number of calibration points on uniform distribution and the mean error on the image plane. The vertical axis is the mean error on the image plane and its unit is pixel. The horizontal axis is the number of calibration points used on uniform distribution. In this figure, it also shows how the calibration points distributed. The seven marks of number 7 are initial seven calibration points. The n calibration points are points with mark n and marks less than n. It is clear that the mean error will dramatically drop down when the number of calibration points increase to nine. It then drops down a little when the number of cal-

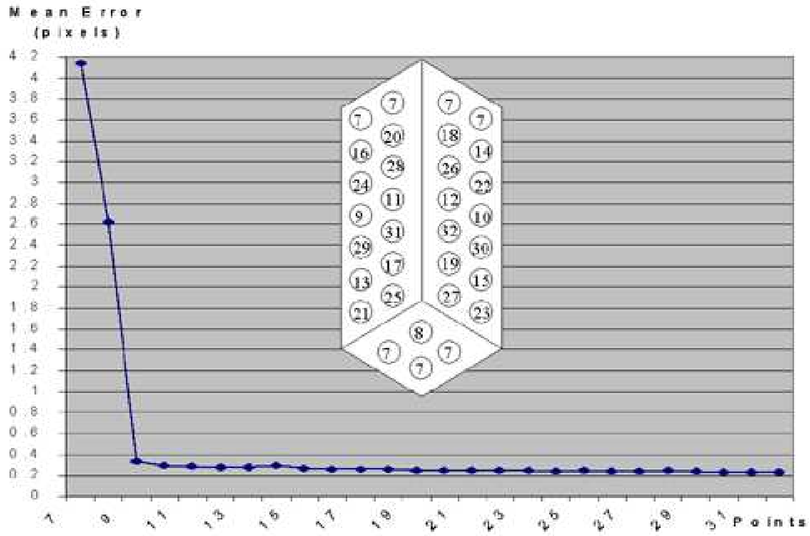

Figure 3.10: Relationship between number of calibration points on uniform distribution and mean error on image plane

ibration points increase. After the number of calibration reach to nineteen, it keeps stable when the number of calibration points increase.

Similar to Fig. 3.10, Figure 3.11 shows the relationship between the number of calibration points on non-uniform distribution and the mean error on the image plane. The vertical axis is the mean error on the image plane and its unit is pixel. The horizontal axis is the number of calibration points used on non-uniform distribution. Figure 3.11 also shows how the calibration points are distributed. The seven marks with number 7 are our initial seven calibration points. The n calibration points are points with mark n and marks less than n. We can see that the mean error will dramatically drop down when the number of calibration points increase to twelve. It then drops down a little when the number of calibration points increase. After the number of calibration reach to twenty-five, it keeps stable when the number

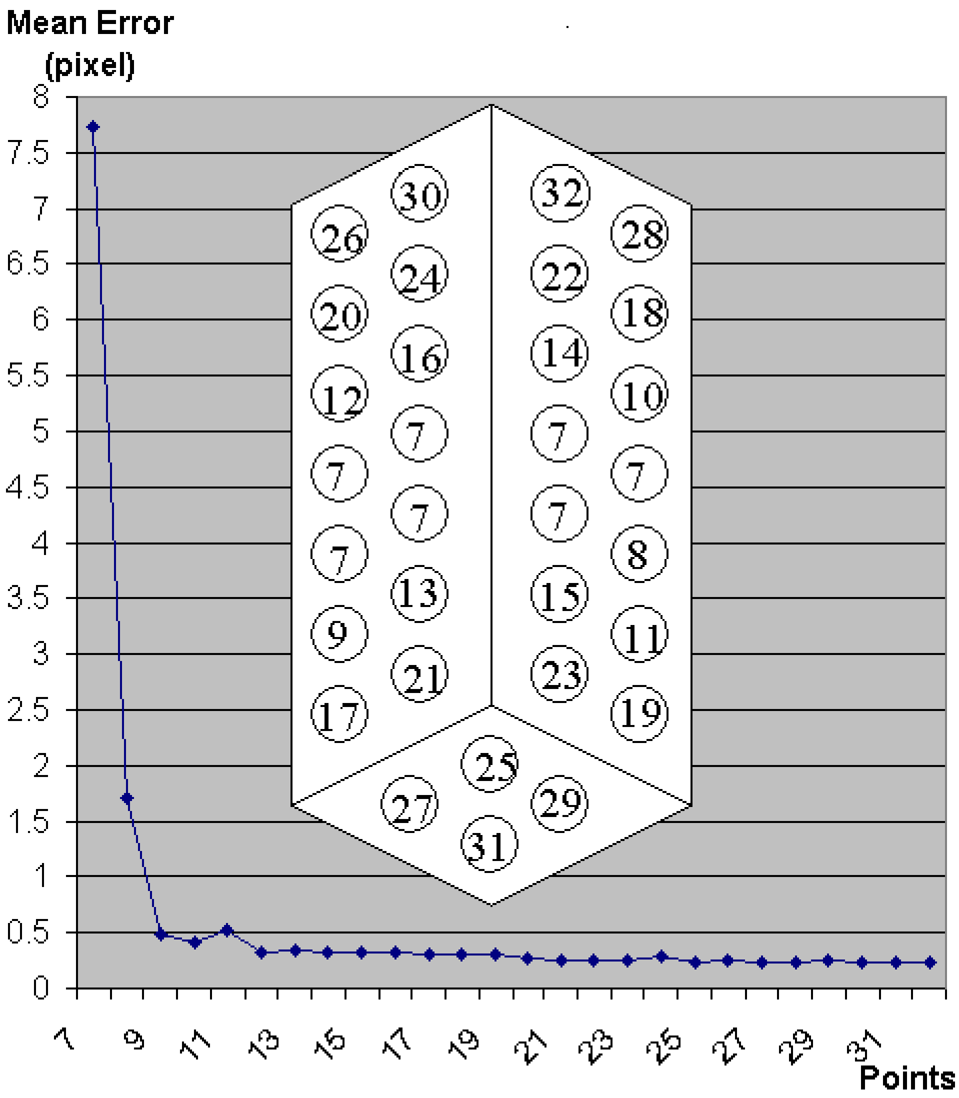

Figure 3.11: Relationship between number of calibration points on non-uniform distribution and mean error on image plane

of calibration points increase.

From the above discussion, it is clear that despite the fact that increasing the number of calibration points can reduce the mean error of calibration, it will be useless for the accuracy of calibration when the number of calibration points reaches a certain threshold. In conclusion, it is better to use twenty-seven calibration points uniformly distributed on three planes of an open cube, i.e. there are nine calibration points uniformly distributed on every plane.

For our calibration open cube, the bottom plane is quite smaller than the other two planes in order to reconstruct a object such as a mannequin. We use the pattern of uniform density distribution. There are 4 marks uniformly distributed on the bottom plane and there are 28 marks uniformly distributed in the same density as the bottom plane on the other two planes. Therefore, there are 32 calibration marks on our calibration open cube shown 3.7.

3.6 Covered Area by a Calibration Object

The camera and the calibration object are setup as in Fig. 3.1. The distance between the camera and the calibration object is 2.8 meters and the camera angle θ is 10 degree. The camera focal length is 6.4 mm.

In this section, we will discuss the impact of the covered area by the calibration object. We define covered area as the percentage of the calibration object height in the reconstruction object height.

$$covered\ area = \frac{the\ height\ of\ calibration\ object}{the\ height\ of\ reconstruction\ object}$$

We use 12 uniformly distributed calibration points to calibrate the camera and then use 32 points shown in Fig. 3.7 to evaluate the camera calibration

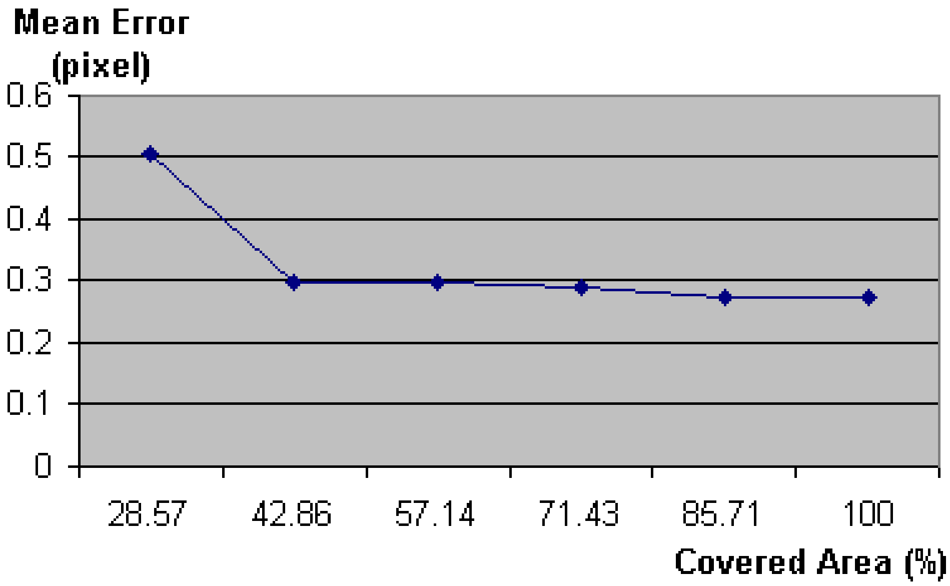

Figure 3.12: Relationship between the mean calibration error and the covered area by a calibration object

result. The 12 uniform distributed calibration points are chosen from 4 bottom points, 4 highest calibration points and 4 middle high calibration points, e.g. calibration object of cover area 71.43% we used are chosen the calibration points of mark 3, 10, 17, 24, 5, 12, 19, 26, 29, 30, 31, 32 shown in Fig. 3.7.

Figure 3.12 shows the relationship between the mean error on the image plane and the covered area by a calibration object. The vertical axis is the mean error on the image plane. The horizontal axis is the percentage of the covered area by a calibration object. It is clear that the mean calibration error will drop down when the percentage of the covered area increases. When the covered area is larger than 80%, the mean calibration error will keep stable.

In conclusion, we have to use a calibration object which will cover more

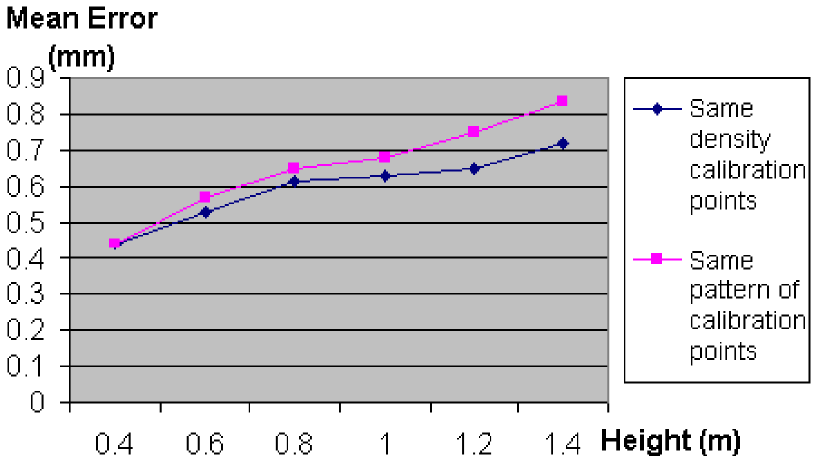

Figure 3.13: Relationship between calibration object size and mean error on the object space when the camera focal length is fixed

than 80% reconstruction area.

3.7 Size of Calibration Object

In this section, we will discuss the impact of the size of the calibration object. The camera and the calibration object are setup as in Fig. 3.1. In order to make that the calibration object is completely filled in the image window, there are two ways to adjust the object image when the calibration object size is changed.

The first way is to adjust the camera angle θ and the distance between the camera and the calibration object when the camera focal length is fixed. The camera focal length is 13 mm.

Figure 3.13 shows the relation between calibration object size and mean error on the object space when the camera focal length is fixed. The vertical

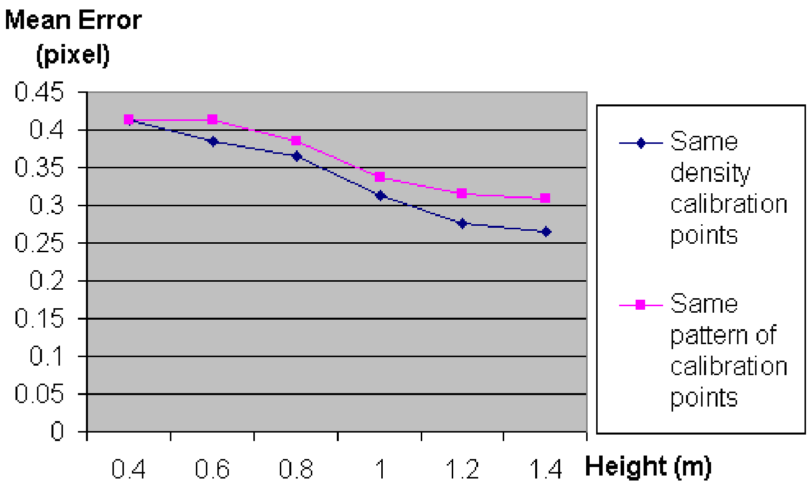

Figure 3.14: Relationship between the calibration object size and the mean error on the image plane when the camera focal length is fixed

axis is the mean error on the object space (closest error on the object space). The horizontal axis is the size of calibration object, i.e. the height of calibration object. One series called same pattern of calibration points uses 12 uniformly distributed calibration points, i.e. every 4 uniformly distributed points are distributed on every plane of 3 open cube planes. The other series called same density of calibration points uses all of calibration points under the height of calibration object shown in Fig. 3.7. When calibration object height (size) increases, the mean error on the object space increases. This is because the camera resolution for calibration object decrease when size of calibration object increase.

Figure 3.14 shows the relation between the calibration object size and the mean error on the image plane when the camera focal length is fixed. The vertical axis is the mean error on the image plane. The horizontal axis is the size of calibration object, i.e. the height of calibration object. One

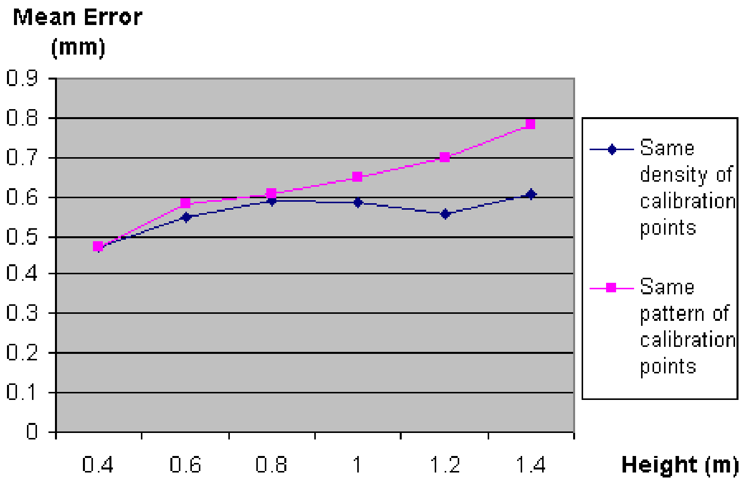

Figure 3.15: Relationship between the calibration object size and the mean error on the object space when the distance between the camera and the calibration object is fixed

series called same pattern of calibration points uses 12 uniformly distributed calibration points, i.e. every 4 uniformly distributed points are distributed on every plane of 3 open cube planes. The other series called same density of calibration points uses all of calibration points under the height of calibration object shown in Fig. 3.7. When calibration object height (size) increases, the mean error on the image plane decreases. This is because the camera resolution for calibration object decrease when size of calibration object increase and the errors of calibration marks is fixed.

The second way is to adjust the camera angle θ and the camera focal length when the distance between the camera and the calibration object is fixed. The distance between the camera and the calibration object is 2.8 meters.

Figure 3.15 shows the relation between the calibration object size and the

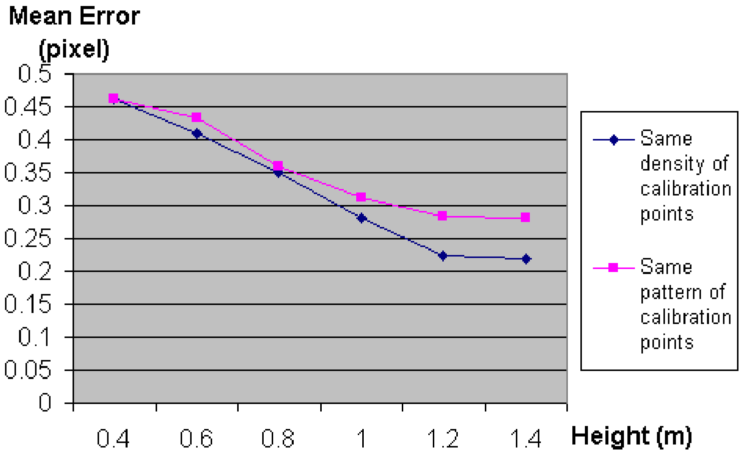

Figure 3.16: Relationship between the calibration object size and the mean error on the image plane when the distance between the camera and the calibration object is fixed

mean error on the object space when the distance between the camera and the calibration object is fixed. The vertical axis is the mean error on the object space (closest error on the object space). The horizontal axis is the size of calibration object, i.e. the height of calibration object. One series called same pattern of calibration points uses 12 uniformly distributed calibration points, i.e. every 4 uniformly distributed points are distributed on every plane of 3 open cube planes. The other series called same density of calibration points uses all of calibration points under the height of calibration object shown in Fig. 3.7. When the calibration object height (size) increases, the mean error on the object space increases. This is because the camera resolution for calibration object decrease when size of calibration object increase.

Figure 3.16 shows the relation between the calibration object size and the mean error on the image plane when the distance between the camera and the calibration object is fixed. The vertical axis is the mean error on the image plane. The horizontal axis is the size of calibration object, i.e. the height of calibration object. One series called same pattern of calibration points uses 12 uniformly distributed calibration points, i.e. every 4 uniformly distributed points are distributed on every plane of 3 open cube planes. The other series called same density of calibration points uses all of calibration points under the height of calibration object shown in Fig. 3.7. When the calibration object height (size) increases, the mean error on the image plane decreases. This is because the camera resolution for calibration object decrease when the size of calibration object increase and the errors of calibration marks is fixed.

In conclusion, for a certain camera, the camera resolution for calibration object decrease when the size of calibration object increase. The calibration error on object space will increase when the size of calibration object increase. On the other hand, the calibration error on image plan will decrease when the size of calibration object increase. This means that the small size of calibration object requires more accuracy calibration marks than the large size of calibration object.

3.8 Distance Between the Calibration Object and the Camera

The camera and the calibration object are setup as in Fig. 3.1. In this section, we will discuss the impact of the distance between the calibration object and the camera. We use a 1.5 meter height calibration object and keep its projected image as large as possible in the image window. In order to fill the calibration object image in the image window, we adjust the camera focal

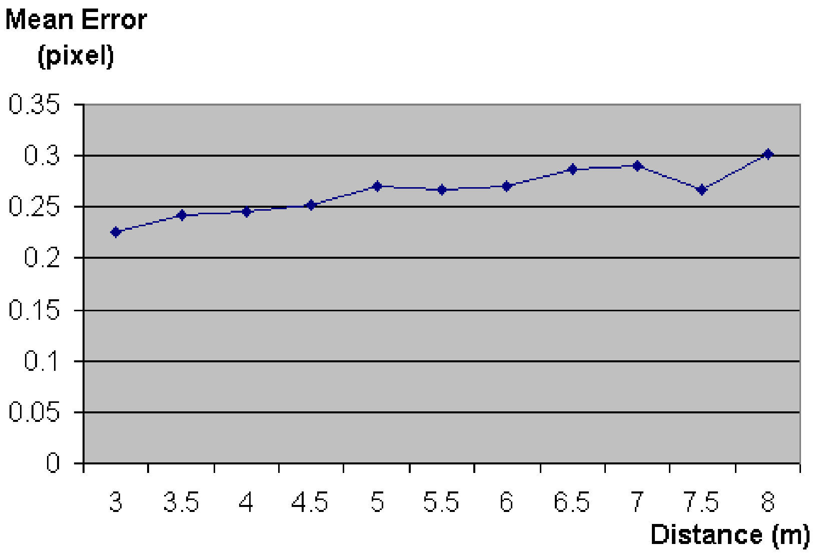

Figure 3.17: Relationship between the mean error on the image plane and the distance between the camera and the calibration object

length and the camera angle θ when the distance between the camera and the calibration object is changed.

Figure 3.17 shows the relation between the mean error on the image plane and the distance between the camera and the calibration object. The vertical axis is the mean error on the image plane. The horizontal axis is the distance between the camera and the calibration object. It is clear that the longer distance will result in the bigger calibration error.

In conclusion, it is better to keep the calibration object as closely to the camera as possible.

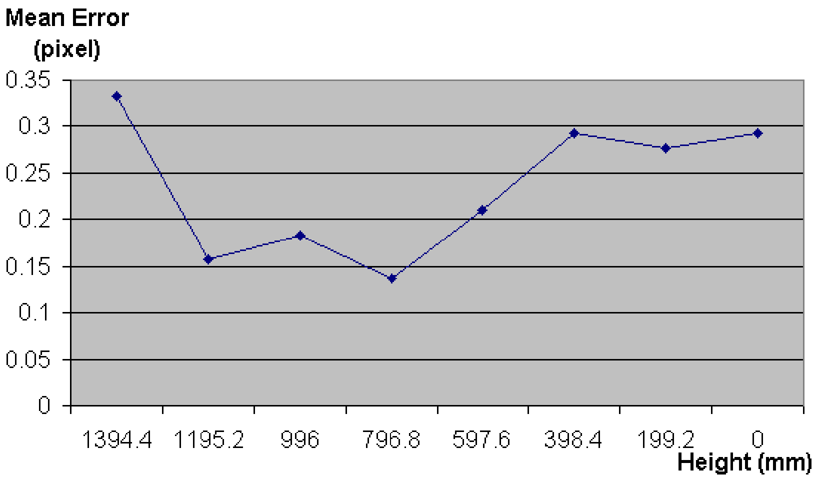

Figure 3.18: Distribution of the mean calibration error around the height of the calibration object

3.9 Distribution of the Calibration Error Around the Heigh of the Calibration Object

In this section, we will find out the pattern of the calibration error distribution around the height of the calibration object. The camera and the calibration object are setup as in Fig. 3.1. The distance between the camera and the calibration object is 2.8 meters and the camera angle θ is 10 degree. The camera focal length is 6.4 mm.

Figure 3.18 shows the distribution of the mean calibration error on the image plane around the height of the calibration object. The horizontal axis is the height of calibration object. The vertical axis is the mean calibration error of the same height of calibration points on the image plane.

From Fig. 3.18, we can get a clear pattern of calibration error distribution around the height of the calibration object. The mean calibration error in the middle height of calibration object is the smallest. The farther the calibration points from the middle height of calibration object, the bigger calibration mean error they have.

CHAPTER 4
Dynamic Stereo Calibration Technique

Our dynamic stereo system basically consists of one camera and one turntable. It is a simple solution for reconstruction of 3D objects. If the exact turning angle is known, the dynamic stereo system becomes a binocular stereo system in principle.

Our aim is to reconstruct a large scale object automatically. In order to calculate a 3D point, two different view images are needed. Our system uses one fixed camera and one turnable which can rotate object at any angle to get different view images. Practically it is hard to measure the turning angle accurately by hand. A calibration technique is needed to calibrate turning angle dynamically.

4.1 Mathematical Model

Our system, we refer it as the general turntable model, assumes that the central projection is used. One point in the camera image plane represents one straight line across this point and the principal point. One point which rotates a certain angle in the world coordinate system will result in a pair of points in the image coordinate system.

Figure 4.1 shows the relationship of dynamic stereo system.

The task in this chapter is to get the world coordinates from a pair of corresponding image points. Tsai's calibration method is chosen as a basic

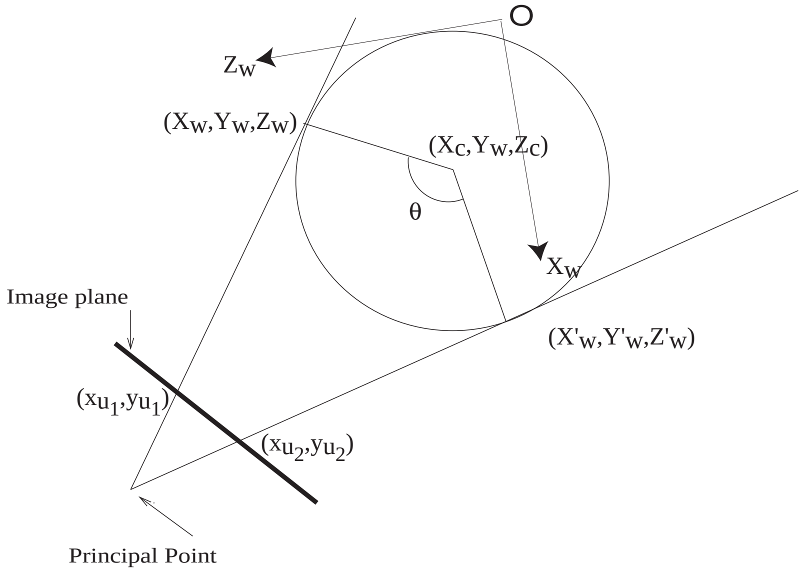

Figure 4.1: Dynamic stereo system

camera calibration technique. Assume a pair of points in image coordinates (x_{i1}, y_{i1}) and (x_{i2}, y_{i2}), the corresponding world coordinate points are $P = (X, Y, Z)$ and the point after rotation $P' = (X', Y', Z')$. The corresponding camera coordinate points are $C = (X_k, Y_k, Z_k)$ and the point after rotation $C' = (X'_k, Y'_k, Z'_k)$.

If the feature of CCD camera sensors and the radial lens distortion are concerned, the distorted image coordinates (x_{v1}, y_{v1}) and (x_{v2}, y_{v2}) can be calculated from Equ. 2.14.

$$x_{v1} = \frac{x_{i1}d'_x - c_x}{s_x} \quad (4.1)$$

$$y_{v1} = y_{i1}d_y - c_y \quad (4.2)$$

$$x_{v2} = \frac{x_{i2}d'_x - c_x}{s_x} \quad (4.3)$$

$$y_{v2} = y_{i2}d_y - c_y \quad (4.4)$$

The undistorted image coordinates (x_{u1}, y_{u1}) and (x_{u2}, y_{u2}) can be derived from Equ. 2.13:

$$x_{u1} = x_{v1} + D_{x1} \tag{4.5}$$

$$y_{u1} = y_{v1} + D_{y1} \tag{4.6}$$

$$x_{u2} = x_{v2} + D_{x2} \tag{4.7}$$

$$y_{u2} = y_{v2} + D_{y2} \tag{4.8}$$

where $D_{xi} = x_{vi} \cdot (k_1 r^2 + k_2 r^4)$, $D_{yi} = y_{vi} \cdot (k_1 r^2 + k_2 r^4)$ and $r = \sqrt{x_{vi}^2 + y_{vi}^2}$.

The undistorted image coordinates can be obtained from the corresponding camera coordinates by using central projection.

$$x_u = \frac{fX_k}{Z_k} \tag{4.9}$$

$$y_u = \frac{fY_k}{Z_k} \tag{4.10}$$

Therefore, the camera coordinates can be rewritten as the following form:

$$C = Z_k \cdot (\frac{x_{u1}}{f}, \frac{y_{u1}}{f}, 1)^T = Z_k \cdot E \tag{4.11}$$

$$C' = Z'_k \cdot (\frac{x_{u2}}{f}, \frac{y_{u2}}{f}, 1)^T = Z'_k \cdot E' \tag{4.12}$$

From Tsai's calibration method, we can get the rotation matrix R and translation vector T.

$$R = \begin{pmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{pmatrix}$$

$$T = (T_x, T_y, T_z)^T$$

The world coordinates can be transformed to the camera coordinates by translation and rotation operations using Equ. 2.1.

$$C = R \cdot P + T \tag{4.13}$$

$$C' = R \cdot P' + T \tag{4.14}$$

Using Equ. 4.11 and Equ. 4.12, we can get P and P':

$$\begin{aligned} P &= R^T \cdot (Z_k \cdot E - T) & (4.15) \\ P' &= R^T \cdot (Z'_k \cdot E' - T) & (4.16) \end{aligned}$$

Assumption 4.1.1. *The rotation axis is parallel to the Y axis of the world coordinate system.*

If our dynamic stereo system is satisfied by the above assumption, the rotation axis is across the point $T_c = (X_c, 0, Z_c)$. The values of X_c and Z_c can be calibrated during the calibration period. The point $P' = (X'_w, Y'_w, Z'_w)$ is the transformation point of $P = (X_w, Y_w, Z_w)$ which is rotated θ degree around the rotation axis. Figure 4.1 shows the relations between them. If we assume that our world coordinate system is right hand system, the rotation matrix R_θ is as follow:

$$R_\theta = \begin{pmatrix} cos\theta & o & sin\theta \\ 0 & 1 & 0 \\ -sin\theta & 0 & cos\theta \end{pmatrix} \quad (4.17)$$

The world coordinate points P' and P satisfy the following equation:

$$P' = T_c + R_\theta \cdot (P - T_c) \quad (4.18)$$

Using Equ. 4.15 and Equ. 4.16, the above equation can be rewritten as follow:

$$R^T \cdot (Z'_k \cdot E' - T) = T_c + R_\theta \cdot (R^T \cdot (Z_k \cdot E - T) - T_c) \quad (4.19)$$

The above vector equation contains three unknown variables θ, Z_k and Z'_k. It also generates three equations and can solve these three unknown variables.

From Equ. 4.19, we can get the following equation:

$$\alpha sin\theta + \beta cos\theta - \beta = 0 \tag{4.20}$$

The parameters α and β are constants if the corresponding two image points are known. They are as follows:

$$\begin{aligned} \alpha &= -a_2b_1c_1 + a_2c_1X_c - d_2 - d_1 - c_2Z_c + c_1Z_c - a_1a_2d_2 + a_1a_2d_1 \\ &\quad +a_1b_2c_2 - a_1c_2X_c & (4.21) \\ \beta &= -a_2c_1Z_c - b_2c_1 + c_1X_c - b_1c_2 + c_2X_c - a_1c_2Z_c & (4.22) \end{aligned}$$

where the parameters a_1, b_1, a_2, b_2 are constants related to x_{u1} and y_{u1} and the parameters a_3, b_3, a_4, b_4 are constants related to x_{u2} and y_{u2}.

$$\begin{aligned} a_1 &= \frac{(r_6r_8 - r_5r_9)x_{u1} + (r_2r_9 - r_3r_8)y_{u1} + (r_3r_5 - r_2r_6)f}{(r_5r_7 - r_4r_8)x_{u1} + (r_1r_8 - r_2r_7)y_{u1} + (r_2r_4 - r_1r_5)f} & (4.23) \\ b_1 &= \frac{(r_8T_y - r_5T_z)x_{u1} + (r_2T_z - r_8T_x)y_{u1} + (r_5T_x - r_2T_y)f}{(r_5r_7 - r_4r_8)x_{u1} + (r_1r_8 - r_2r_7)y_{u1} + (r_2r_4 - r_1r_5)f} & (4.24) \\ c_1 &= \frac{(r_4r_9 - r_6r_7)x_{u1} + (r_3r_7 - r_1r_9)y_{u1} + (r_1r_6 - r_3r_4)f}{(r_5r_7 - r_4r_8)x_{u1} + (r_1r_8 - r_2r_7)y_{u1} + (r_2r_4 - r_1r_5)f} & (4.25) \\ d_1 &= \frac{(r_4T_z - r_7T_y)x_{u1} + (r_7T_x - r_1T_z)y_{u1} + (r_1T_y - r_4T_x)f}{(r_5r_7 - r_4r_8)x_{u1} + (r_1r_8 - r_2r_7)y_{u1} + (r_2r_4 - r_1r_5)f} & (4.26) \\ a_2 &= \frac{(r_6r_8 - r_5r_9)x_{u2} + (r_2r_9 - r_3r_8)y_{u2} + (r_3r_5 - r_2r_6)f}{(r_5r_7 - r_4r_8)x_{u2} + (r_1r_8 - r_2r_7)y_{u2} + (r_2r_4 - r_1r_5)f} & (4.27) \\ b_2 &= \frac{(r_8T_y - r_5T_z)x_{u2} + (r_2T_z - r_8T_x)y_{u2} + (r_5T_x - r_2T_y)f}{(r_5r_7 - r_4r_8)x_{u2} + (r_1r_8 - r_2r_7)y_{u2} + (r_2r_4 - r_1r_5)f} & (4.28) \\ c_2 &= \frac{(r_4r_9 - r_6r_7)x_{u2} + (r_3r_7 - r_1r_9)y_{u2} + (r_1r_6 - r_3r_4)f}{(r_5r_7 - r_4r_8)x_{u2} + (r_1r_8 - r_2r_7)y_{u2} + (r_2r_4 - r_1r_5)f} & (4.29) \\ d_2 &= \frac{(r_4T_z - r_7T_y)x_{u2} + (r_7T_x - r_1T_z)y_{u2} + (r_1T_y - r_4T_x)f}{(r_5r_7 - r_4r_8)x_{u2} + (r_1r_8 - r_2r_7)y_{u2} + (r_2r_4 - r_1r_5)f} & (4.30) \end{aligned}$$

The following equation also holds:

$$sin^2\theta + cos^2\theta = 1 \tag{4.31}$$

From Equ. 4.20 and Equ. 4.31, the two solution of θ are usually obtained. One solution is $\theta = 0$. It is the principal point and should be thrown away.

Another solution is what we want. Therefore, the following solution can be obtained:

$$cos\theta = \frac{\beta^2 - \alpha^2}{\alpha^2 + \beta^2} \tag{4.32}$$

$$sin\theta = \frac{\beta}{\alpha}(1 - cos\theta) \tag{4.33}$$

The turning angle can be calculated through Equ. 4.32 and Equ. 4.33. A set of corresponding points with the same turning angle, if possible, should be used to calculate the average turning angle.

Once the turning angle has been calculated, Z_k and Z'_k can also be calculated from Equ. 4.19. The point P can be calculated by Equ. 4.15 or the following equation:

$$P = T_c + R_\theta^T \cdot (R^T \cdot (Z'_k \cdot E' - T) - T_c) \tag{4.34}$$

The Z_w of point P can be further derived from Equ. 4.15 and Equ. 4.34

$$Z_w = \frac{a_2d_2 - a_2d_1 - c_2b_2 + c_2X_c - c_2Z_csin\theta + (c_2b_1 - c_2X_c)cos\theta}{c_1a_2 - c_2sin\theta - a_1c_2cos\theta} \tag{4.35}$$

$$Z_w = \frac{d_2 - d_1 + c_2Z_c - (c_2b_1 - c_2X_c)sin\theta - c_2Z_ccos\theta}{c_1 + a_1c_2sin\theta - c_2cos\theta} \tag{4.36}$$

The results of Z_w calculated by using Equ. 4.35 or Equ. 4.36 should be same ideally. In practise, they are not the same. For our experience, the result of Equ. 4.36 is always closer to the true value than the result of Equ. 4.35. The difference of these two results can be used to evaluate the accuracy of the result.

After calculation of Z_w, the X_w and Y_w of world coordinates can be calculated by using the following equations:

$$X_w = a_1 \cdot Z_w + b_1 \tag{4.37}$$

$$Y_w = c_1 \cdot Z_w + d_1 \tag{4.38}$$

4.2 Comparison to Aligned Turntable Model

The dynamic stereo system (one camera + one turntable system) is also mentioned by the textbook [1]. In the textbook, it builds up another mathematical model to solve the calibration problem. We refer this mathematical model mentioned in the textbook as aligned turntable model.

In the textbook, there are two assumptions even if they are not clearly stated.

Assumption 4.2.1. *The rotation axis passes through the origin of world coordinate system.*

Assumption 4.2.2. *The rotation axis is an axis of the world coordinate system.*

According to Assumption 4.2.1, the point $P = (X_w, Y_w, Z_w)$ and the point after rotation $P' = (X'_w, Y'_w, Z'_w)$ satisfy the following equation:

$$P' = R_\delta \cdot P \tag{4.39}$$

In the above equation, R_δ is the rotation matrix. According to Assumption 4.2.2, we can get the following equation if we assume that the rotation axis is Y-axis of the world coordinate system and the coordinates system is a right-hand system.

$$R_\delta = \begin{pmatrix} cos\delta & 0 & sin\delta \\ 0 & 1 & 0 \\ -sin\delta & 0 & cos\delta \end{pmatrix} \tag{4.40}$$

It is clear that R_δ in this case is the same as the R_θ in the general turntable model describe in Sec. 4.1. There are two main differences between the general turntable model and the aligned turntable model. One is that the aligned turntable model is only for special case where Assumption 4.2.1 holds. In practise, it is very hard to hold Assumption 4.2.1. The other one is that the general turntable model is the special case where the Y-axis of world coordinate system is the rotation axis. Therefore, we can get $T_c = (0, 0, 0)$ if Assumption 4.2.1 holds for the general turntable model. And also assume that the rotation axis is Y-axis in aligned turntable mode, $R_\delta = R_\theta$ holds. Equation 4.18 is the same as Equ. 4.39. Except the above two differences, the two mathematical model use the same relation equations.

In conclusion, there are only two differences as described above between the general turntable model and the aligned turntable model. In despite of the above two differences, they are the same and can get the same solutions of world coordinates. The general turntable model is more practicable than the aligned turntable model.

4.3 Method of Reconstructing a 3D Point

In the implementation, we use Tsai's calibration method as a basic calibration method and extend it to deal with a dynamic stereo system. The general idea is as follows:

4.3.1 Calibration of Camera Parameters

The Tsai's calibration method is used to calibrate the camera system. An image with non-coplanar marks is taken and it is used as calibration image to get calibration data. It included the camera focal length, distortion coefficients k_1 and k_2, 3D rotation matrix R, translation vector T and the scaling factor s_x. This has already been described in Chapter 2.

4.3.2 Calibration of Rotation Axis

A mark is put on the turntable and it is rotated 180 degrees. The two images are taken respectively. There are two points on the camera image plane. From these two image points, we calculate the projected points on the turntable. In our system, the turntable plane is Y = -8 mm in the world coordinate system. We can calculate these two points projected on the turntable plane $(X_1, -8.0, Z_1)$ and $(X_2, -8.0, Z_2)$. The point which intersects with rotation axis and the plane Y = -8 mm is $(X_c, -8.0, Z_c)$, $X_c = (X_1 + X_2)/2$ and $Z_c = (Z_1 + Z_2)$. We can assume that Y axis is parallel to the rotation axis in the world coordinate system. The rotation axis is $X = X_c$ and $Z = Z_c$

The above calibration method is simple but the accuracy is not good enough. We can put a mark on the turntable and rotate the turntable for several rotation angle. More than three images are taken. From these images, we can calculate the corresponding projected points on the turntable. These projected points are $(X_i, -8.0, Z_i)$, $i = 1, 2, \cdots, n$. They satisfy the following equation:

$$(X_i - X_c)^2 + (Z_i - Z_c)^2 = r^2 \tag{4.41}$$

The above equation can be rewritten as follow:

$$(2X_i, 2Z_i, 1) \begin{pmatrix} X_c \\ Z_c \\ r^2 - X_c^2 - Z_c^2 \end{pmatrix} = X_i^2 + Z_i^2 \tag{4.42}$$

For more than three points, we can use the pseudo-inverse technique (*Moore-Penrose inverse*) to get the solutions of X_c and Z_c.

4.3.3 Calculation of Rotation Angle

The input arguments are a pair of corresponding image points. The camera calibration parameters of Tsai's calibration method is used. The parameters a_1, b_1, c_1, d_1, a_2, b_2, c_2 and d_2 can be calculated by Equ. 4.23, Equ. 4.24, Equ. 4.25, Equ. 4.26, Equ. 4.27, Equ. 4.28, Equ. 4.29, and Equ. 4.30.

The parameters α and β are then calculated by Equ. 4.21 and Equ. 4.22 respectively.

Finally, the values of $sin\theta$ and $cos\theta$ can be calculated by Equ. 4.33 and Equ. 4.32 respectively. The turning angle can be calculated by the values of $sin\theta$ and $cos\theta$.

In practices, the values of $sin\theta$ and $cos\theta$ can not be calculated exactly as the true value because of the system error and random error. If possible, a set of corresponding image points should be used to calculate the value of turning angle individually and then use the average of these turning angles as the value of calibration turning angle.

4.3.4 Calculation of World Coordinates

Using the calibrated turning angle θ, the $sin\theta$ and $cos\theta$ values can be calculated. The Z_w value then can be easily calculated through Equ. 4.35 or Equ. 4.36. The two solutions of Z_w can be obtained by these two equations. These two values should be equal ideally, but are not equal in practice. For our experiences, the result of Equ. 4.36 is always closer to true value than the result of Equ. 4.35. Therefore, Z_w can be calculated by using Equ. 4.36. The values of X_w and Y_w then can be calculated by Equ. 4.37 and Equ. 4.38 respectively.

CHAPTER 5

Evaluation of Dynamic Stereo System

We have already built up a mathematical model for our dynamic stereo system in Chapter 4. The accuracy of this dynamic stereo system needs to be evaluated. It also needs to find out the weakness of this system.

In this chapter, we show how to evaluate our dynamic stereo system and we will use a set of calibration data to reconstruct 3D points in a world coordinate system. These reconstructed 3D points are used to evaluate the accuracy of the dynamic stereo system. We will also discuss the problems of this system.

5.1 Evaluation Method

There are three steps for the evaluation of the dynamic stereo system. The evaluation method is demonstrated in the following subsections.

5.1.1 Obtain Calibration Data

The open cube is used as our calibration object. In order to deal with a large scale object, we build a open cube with 1.5 meter height, 0.5 meter width and 0.5 meter length. The calibration marks are needed and are put on the calibration open cube. It is impossible to build up a large scale calibration open cube by hand with a maximum error less than 1 mm. We use a plotter to print a calibration marks paper with 1 meter width and 2 meter length and then put it on the calibration open cube. It is shown in Fig. 5.1.

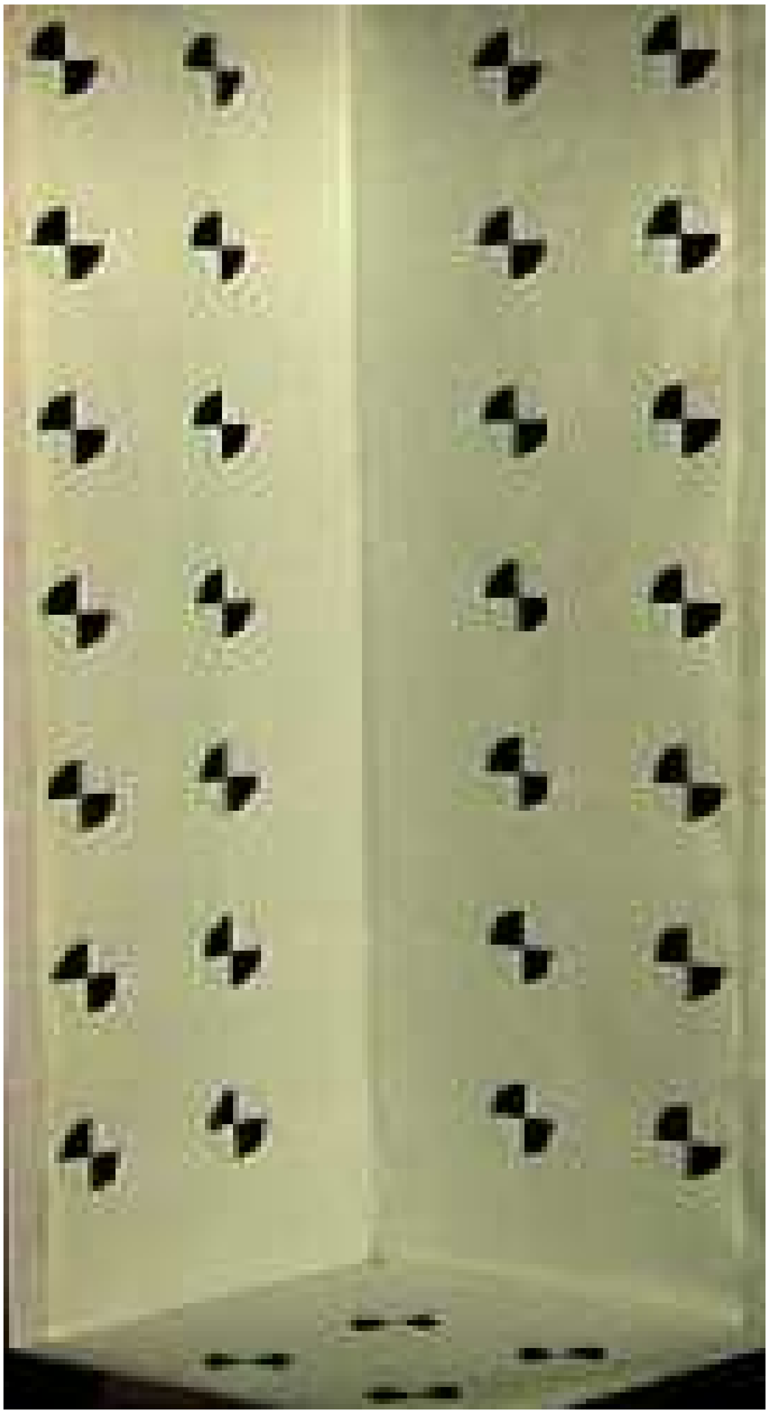

Figure 5.1: Calibration open cube

The open cube is put on the turntable. The three intersection lines of the planes of open cube are used as the axes of world coordinate system. The right hand coordinate system is used and the vertical direction is set as Y-axis of world coordinate system.

The distance between the camera and the calibration open cube is adjusted in order to fill the whole calibration object in the camera image window. The image of the calibration open cube is taken and then the turntable is rotated some degrees, e.g. about 30 degrees, anti-clockwise. Then another image of the calibration open cube is taken for further reconstruction of 3D points.

The images taken above are needed to be pre-processed. Firstly, the objects except the open cube on the image are cut off. Secondly, the values of image pixels are inverted. Finally, the threshold operation is applied to these images to deduct the noise on the images.

The program *fmc* is used with the file name of processed calibration open cube image as the parameter. It produces a data file called *ncc_cd.dat* and then we can rename this data file as what we need. The data file contains the world coordinates of every calibration mark and their corresponding image coordinates.

The image coordinate system must be a right hand system. It always has a virtual Z-axis from the image plane towards the calibration object. Therefore, the origin of image coordinates must be in the top-left of image. The X-axis is towards to right and Y-axis is towards to bottom.

5.1.2 Camera Parameters and Rotation Axis

Tsai's calibration method is used for the calibration of the camera parameters. The calibration data file produced by the first taken image should be named as *ncc_cd.dat*. The batch program *nccal* should be called. It produces the camera parameters file called *ncc_cpcc.dat*.

The calibration open cube is removed from the turntable and keep the turntable position unchanged. A calibration mark, e.g. the mark 2 in Fig. 3.5 is printed on a white paper, is put on the turnable. The image is taken and the other images are also taken after the turntable is rotated certain degrees.

These images are processed as the calibration open cube image. The program *findrc* is called. The name of camera parameters file produced during the last step is the first program parameter and these image file names are the other parameters of the program *findrc*. This program produces a data file called *tc.dat*. It contains the rotation axis coordinates X_c and Z_c. We assume that the rotation axis is parallel to Y-axis in the world coordinate system.

5.1.3 Reconstruction of 3D Points

The reconstruction of 3D points needs a pair of images which are taken for the same object and from the different view directions. The dynamic stereo system uses turntable to get the images with different view directions. For our evaluation, we use a pair of images taken at Sec. 5.1.1.

The program *ic2wc3d* is called. The file name of camera parameters is the first parameter of the program and the file name of rotation axis parameters is the second parameter of the program. It then prompts user to input a pair of image coordinates and produces a rotation angle and its 3D world coordinates.

The rotation angle may be various because of the noise on the image. If we know the exact rotation angle or the average of these rotation angles on the pair of images, the program *ic2wc3dw* can be used. The file name of camera parameters is also the first parameter of this program and the file name of rotation axis data is also the second parameter of this program. This program prompts user to input a pair of image coordinates and the rotation angle. It then produces the corresponding 3D world coordinates.

After we get these 3D world coordinates, we can evaluate the dynamic stereo system by comparing these 3D world coordinates to their true world coordinates.

5.2 Evaluation Result

The camera and the calibration object are setup as in Fig. 3.1. The distance between the camera and the calibration object is 2.8 meters and the camera angle θ is 10 degree. The camera focal length is 6.6 mm. The working environment setup as in Sec. 3.1.

We use the first image taken at rotation angle $\theta = 0$ to calibrate the camera. The 32 calibration marks are used to calibrate the camera. The camera parameters are calibrated shown in Tab. 5.1.

In Tab. 5.1, f is the focal length of the camera. k_1 is the camera lens distortion coefficient. T_x is the X component of the translation vector. T_y is the Y component of the translation vector. T_z is the Z component of the translation vector. R_x is the rotation angle around X-axis. R_y is the rotation angle around Y-axis. R_z is the rotation angle around Z-axis. C_x is the x-coordinate of the image center point. C_y is the y-coordinate of the image center point. S_x is the horizontal scaling factor. Using these camera calibration parameters, the camera calibration accuracy is that the mean error on the image plane is 0.26 pixel and the mean error on the object space (closest error on the object space) is 1.4 mm.

Camera Parameter	Value
f	6.582940 [mm]
k_1	0.005281476 [$1/mm^2$]
T_x	64.883741 [mm]
T_y	633.620886 [mm]
T_z	3273.513524 [mm]
R_x	-172.289584 [degree]
R_y	62.099022 [degree]
R_z	7.441918 [degree]
C_x	369.636984 [pixels]
C_y	296.529166 [pixels]
S_x	1.006039

Table 5.1: Camera calibration parameters

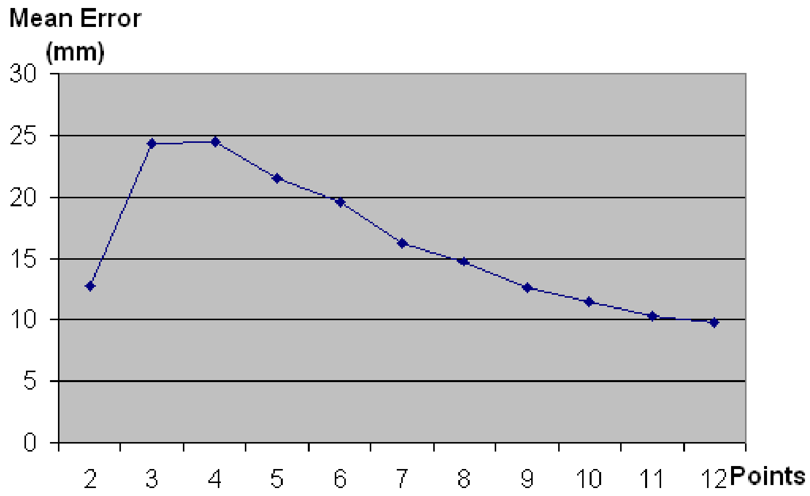

Figure 5.2: Relation between the mean error and the nonuniform distributed calibration points for calibrating the rotation axis

There are 32 3D points which are used to evaluate the calibration result. We firstly use the program without the requirement of rotation angle to reconstruct these 32 points in the world coordinates. These world coordinates are compared to the true values as the error on the object space.

In our dynamic stereo system, the calibration of the rotation axis is quite an important step. We should find out how many points are needed and how these points are distributed for calibration of the rotation axis.

Figure 5.2 shows the relation between the mean error on the object space and the number of nonuniform distributed calibration points for calibrating the rotation axis. The vertical axis is the mean error of the reconstructed points by using our dynamic stereo system without the requirement of the rotation angle. The horizontal axis is the number of nonuniformly distributed calibration points. These points are distributed as in Fig. 5.3. The two points are distributed as the mark 1 and the mark 12. The n ($n > 2$) points are

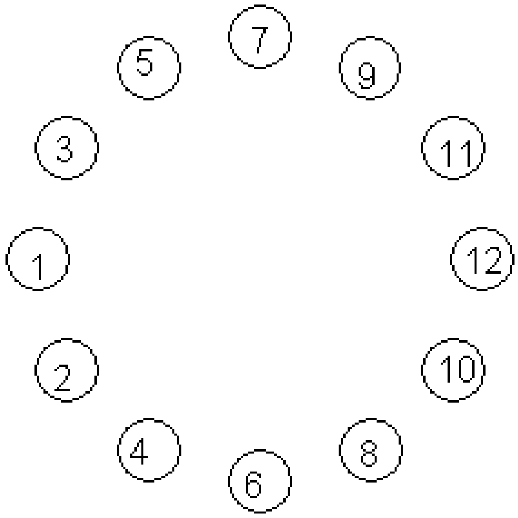

Figure 5.3: Nonuniform distribution of calibration marks for calibrating the rotation axis

distributed as the mark n and the mark m ($m < n$). It is clear that the uniform distribution is not accepted for calibration of the rotation axis.

Figure 5.4 shows the relation between the mean error on the object space and the number of uniformly distributed calibration points for calibrating the rotation axis. The vertical axis is the mean error of reconstructed points by using our dynamic stereo system without the requirement of the rotation angle. The horizontal axis is the number of uniform distributed calibration points for calibrating the rotation axis. These points are distributed as in Fig. 5.5. The n points are distributed as the mark n and the mark m ($m < n$). It is clear that the mean error will keep stable when the number of the uniformly distributed calibration marks reaches 8. Therefore, we should use at least 8 uniformly distributed points to calibrate the rotation axis.

From the above discussion, we know that the rotation axis is parallel to Y-axis and also passes through the point $(217.033362, -8.0, 249.33312)$ which

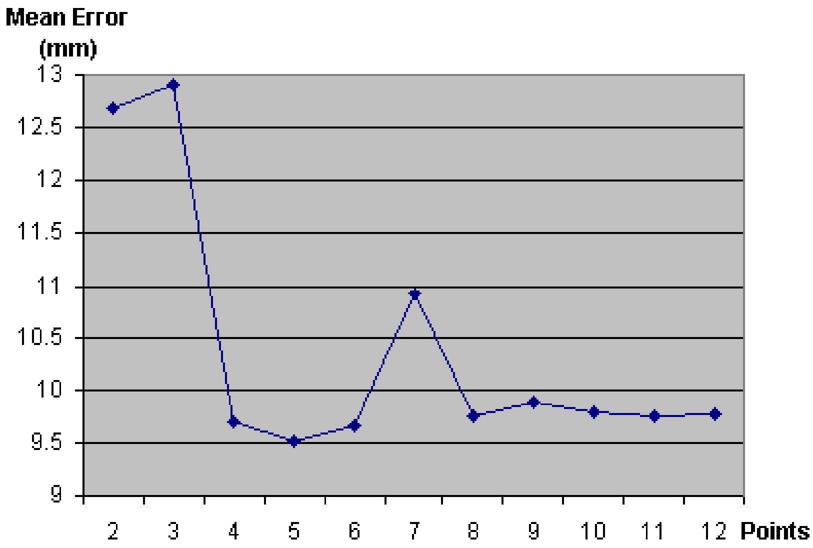

Figure 5.4: Relation between the mean error and the uniform distributed calibration points for calibrating the rotation axis

is calibrated in the above step. Using 32 pairs of image coordinates and the above calibration parameters, we can reconstruct 32 world coordinate points by using our dynamic stereo system.

Figure 5.6 shows how the errors are distributed in the geometric position by using the dynamic stereo reconstruction method. The vertical axis is the distant between the reconstructed point and its true point. The horizontal axis is the mark number of the reconstructed point whose geometric position is shown in Fig. 3.7. From Fig. 5.6, except the points of the mark 18 and the mark 11, the errors of other points are under 20 mm for the dynamically calibrating the rotation angle and are around 10 mm for the pre-calibrating the rotation angle. The mean error for the dynamically calibrated rotation angle is 9.78 mm and the mean error for the pre-calibrated rotation angle is 6.30 mm. The problem of reconstructing the two points of the mark 11 and

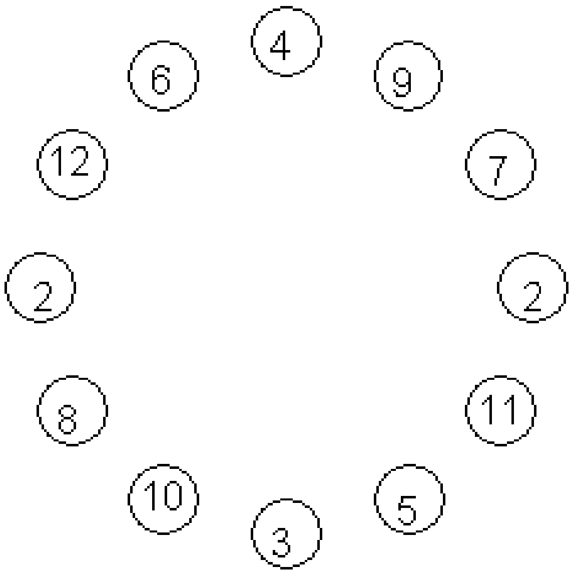

Figure 5.5: Uniform distribution of calibration marks for calibrating the rotation axis

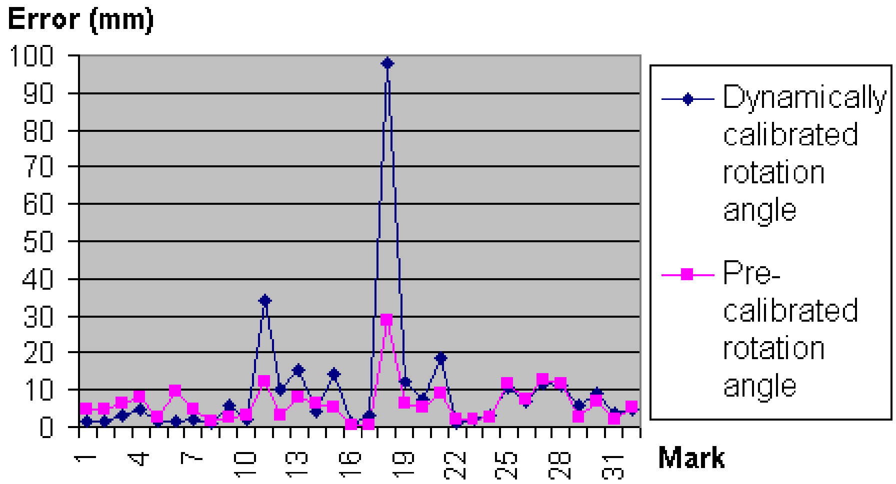

Figure 5.6: Error distribution

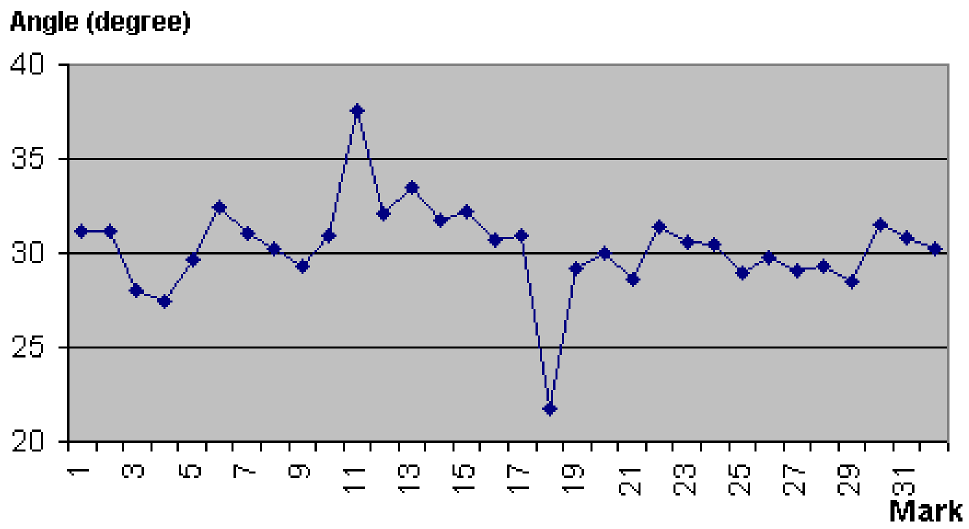

Figure 5.7: Distribution of the rotation angle

the mark 18 will be discussed in Sec. 5.3.

When we reconstruct these 32 points by using the program *ic2wc3d*, we also get the dynamic rotation angle. We know that the rotation angle is about 30 degrees. It is not the exact 30 degrees because we measure it by hand. Figure 5.7 shows how the rotation angle is distributed in the geometric position. The vertical axis is the rotation angle. The horizontal axis is the mark number of reconstructed point whose geometric position is shown in Fig. 3.7. The average rotation angle is 30.32 degrees. This average rotation angle is used as an input parameter in the program *ic2wc3dw* for the pre-calibrated rotation angle in Fig. 5.6. From Fig. 5.7 and Fig. 5.6, It is clear that the two points of the mark 18 and the mark 11 are far away from the true rotation angle and also have big errors. The problem of reconstructing these two points of the mark 11 and the mark 18 will be discussed in Sec. 5.3.

Figure 5.8 shows the error distribution around the height of calibration

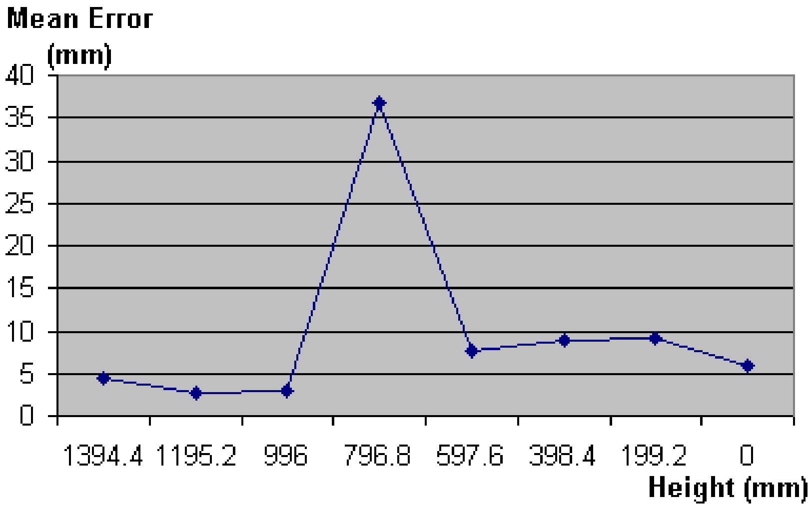

Figure 5.8: Error distribution around the height of calibration object for the model of dynamically calibrated rotation angle

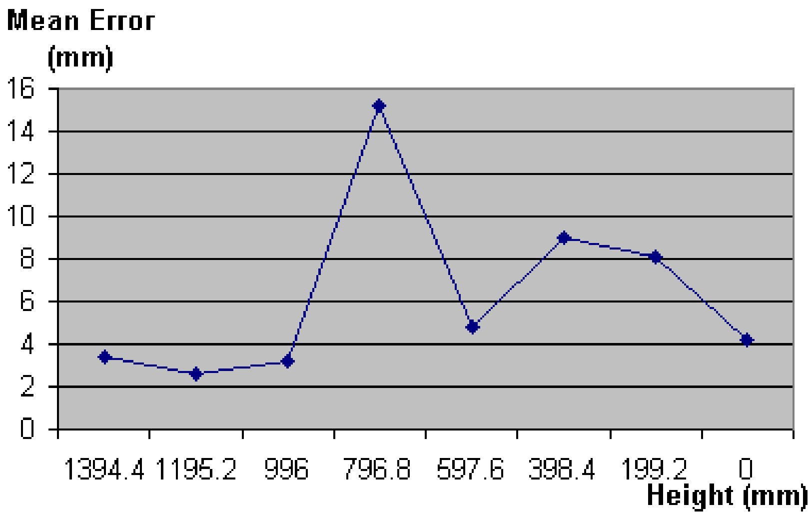

Figure 5.9: Error distribution around the height of calibration object for the model of pre-calibrated rotation angle

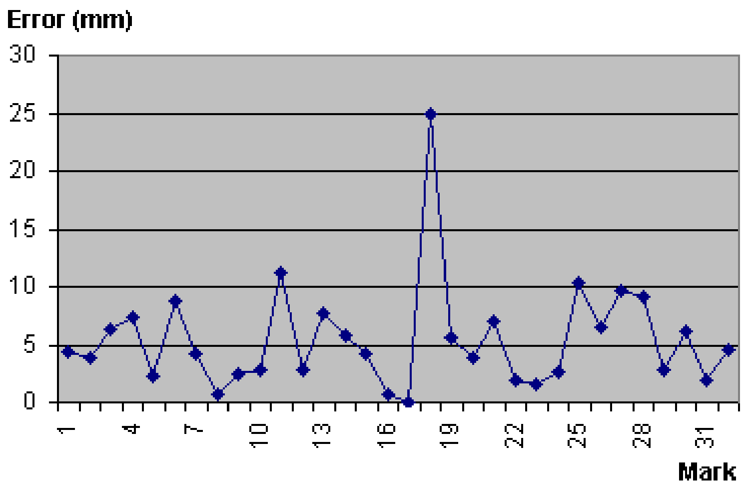

Figure 5.10: Errors on X-axis of world coordinate system

object for the model of dynamically calibrated rotation angle. Figure 5.9 shows the error distribution around the height of calibration object for the model of pre-calibrated rotation angle. Both of the horizontal axes are the height position of calibration object. Both of the vertical axes are the mean error, i.e. the average distance between the reconstructed points and their true points. It is clear that the mean error is higher in the middle of calibration object than in the low or high position of calibration object for both of the models. But the corresponding mean error of the dynamically calibrated rotation angle model is higher than the pre-calibrated rotation angle model. The reason that the error in the middle height of calibration object is higher than in the low or high position of calibration object will be discussed in Sec. 5.3.

Figure 5.10, Fig. 5.11 and Fig. 5.12 show the error of X coordinates, Y coordinates and Z coordinates distribution. All of the horizontal axes are the mark number of the reconstructed point whose geometric position shown in

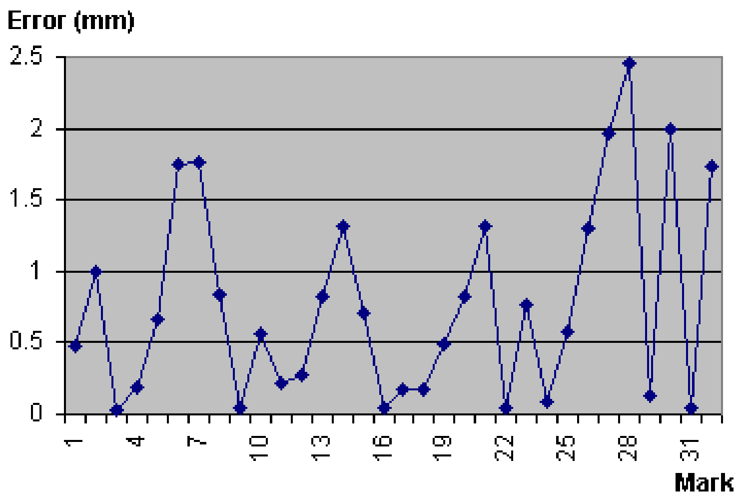

Figure 5.11: Errors on Y-axis of world coordinate system

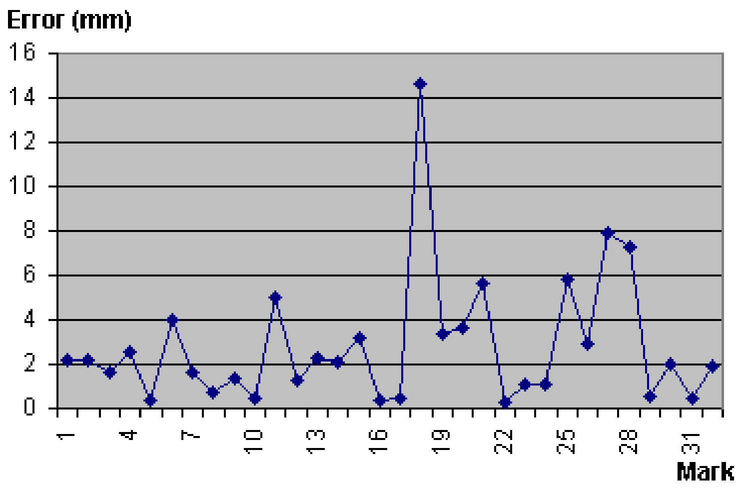

Figure 5.12: Errors on Z-axis of world coordinate system

Fig. 3.7. The vertical axes are the error of X coordinates, Y coordinates and Z coordinates respectively. From Fig. 5.11, the errors on the Y coordinates are quite small. The maximum error is 2.45 mm. The errors of mark 18 and mark 11 are 0.17 mm and 0.21 mm respectively. Most of the errors come from X coordinates and Y coordinates. From Fig. 5.10 and Fig. 5.12, the errors are quite small except two points of the mark 11 and the mark 18. The reason for that will be discussed in Sec. 5.3.

5.3 Problems of Dynamic Stereo System

From the above evaluation result, we know that there are some big errors for two points of the mark 18 and the mark 11. The most of errors of these two points come from X-axis and Z-axis, not Y-axis.

We use these two data sets to reconstruct these points by using binocular stereo model. The mean error is 1.2 mm and the maximum error is 2.42 mm. This is done in Chapter 6. The evaluation result is showed in Fig. 6.2. Therefore, the big error is only caused by the error of rotation angle and the error of rotation axis.

For the error of rotation angle, we can use a set of points with the same rotation angle to dynamic calibrate the rotation angle individually and using their average rotation angle as the true rotation angle. This method requires quite more points to get the true rotation angle. Another method is to use the rotation angle measured by hand. But it can not be used in some cases, e.g. there is not enough time to measure.

There is an assumption on our mathematical model. It assumes that the rotation axis is parallel to Y-axis. In fact, it is impossible to make the rotation axis absolutely parallel to Y-axis. Therefore, the error of rotation axis will increase when the height of calibration object increase for a large scale object.

From Fig. 3.7, we can see that the two points which have the biggest reconstruction errors are all at height of $Y = 796.8$ mm and are close to the rotation axis. From Fig. 3.1, it is clear that these two points are at the same height with the camera. Therefore, the two straight line across these points and their corresponding image points are parallel to the Y plane. This will increase freedom in our mathematical model. If there is a small error on rotation axis and also the points are close enough to the rotation axis, it will find a wrong point which is far away from the true point and generate quite a big error. This is why the two points of the mark 11 and the mark 18 have quite big errors. The more close to the middle height of the object the more freedom. This is why the mean error in the middle of the object height is higher than the low or high position of the object. Because there is not freedom increased in Y direction, the wrong point and the true point are always in the same Y plane or nearby. The most of error therefore come from X coordinate and Z coordinate.

In order to overcome the error of calibration axis, it is possible to calibrate the whole calibration axis and change our mathematical model including a rotation matrix. It is quite complicated and will not be discussed in our thesis.

In conclusion, the dynamic stereo system can reconstruct a 3D large scale object in acceptable accuracy. In our case, for a 1.5 meter high object, it can reach the accuracy characterized by a mean error of 6.3 mm and a maximum error of 28.93 mm.

CHAPTER 6

Binocular Stereo

As the hardware, such as camera and computer, becomes cheaper and more powerful nowadays. Binocular stereo can be widely used to reconstruct a 3D object.

This chapter demonstrates a mathematical model and its calibration implementation for a binocular stereo system. We also evaluate the accuracy of this binocular stereo system.

6.1 Mathematical Model

Binocular stereo system uses two cameras. It does not need a turntable and also it does not make any assumptions. This system needs more than one camera, which is not a problem nowaday because camera prices drop down.

Figure 6.1 shows the geometric illustration of a binocular stereo system. For every camera, we need to do camera calibration before it can be used. We use Tsai's calibration method to do camera calibration for the two cameras. Therefore, the two sets of cameras calibration parameters are obtained. They include the internal parameters and external parameters.

For any 3D point P in the world coordinate system, its world coordinates are (X_w, Y_w, Z_w). The projected point of P on the left camera image is $P_L(x_{i1}, y_{i1})$. The projected point of P on the right image is $P_R(x_{i2}, y_{i2})$.

From the derivation in Sec. 4.1, we know that one point in the image represents a straight line passing through that point and the principal point.

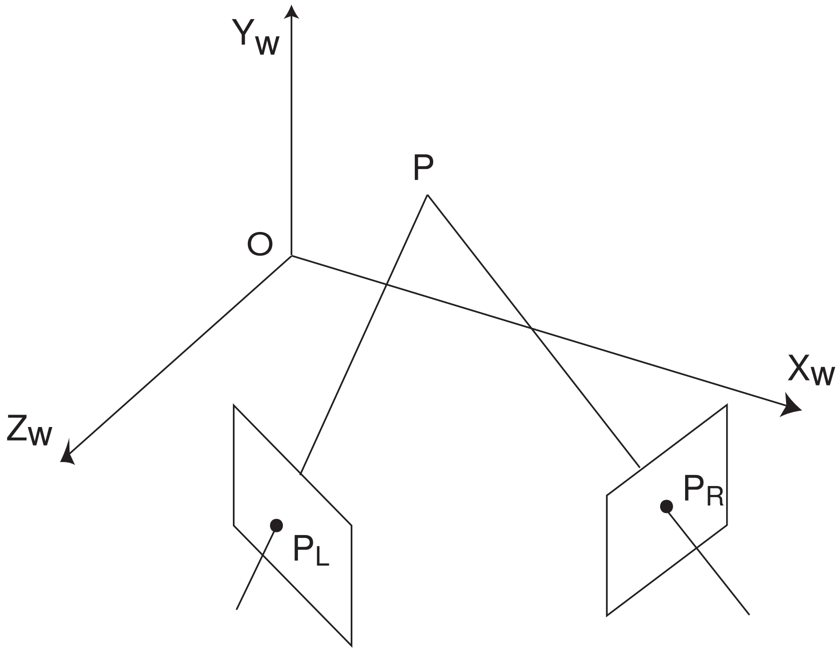

Figure 6.1: Binocular stereo system

This line can be represented as Equ. 4.37 and Equ. 4.38. For the convenience of reading, these two formulas are rewritten here:

$$\begin{aligned} X_w &= aZ_w + b \\ Y_w &= cZ_w + d \end{aligned}$$

The parameters a, b, c and d are only associated with the undistorted image coordinates if the camera is fixed. They are as follows:

$$a = \frac{(r_6r_8 - r_5r_9)x_u + (r_2r_9 - r_3r_8)y_u + (r_3r_5 - r_2r_6)f}{(r_5r_7 - r_4r_8)x_u + (r_1r_8 - r_2r_7)y_u + (r_2r_4 - r_1r_5)f} \tag{6.1}$$

$$b = \frac{(r_8T_y - r_5T_z)x_u + (r_2T_z - r_8T_x)y_u + (r_5T_x - r_2T_y)f}{(r_5r_7 - r_4r_8)x_u + (r_1r_8 - r_2r_7)y_u + (r_2r_4 - r_1r_5)f} \tag{6.2}$$

$$c = \frac{(r_4r_9 - r_6r_7)x_u + (r_3r_7 - r_1r_9)y_u + (r_1r_6 - r_3r_4)f}{(r_5r_7 - r_4r_8)x_u + (r_1r_8 - r_2r_7)y_u + (r_2r_4 - r_1r_5)f} \tag{6.3}$$

$$d = \frac{(r_4T_z - r_7T_y)x_u + (r_7T_x - r_1T_z)y_u + (r_1T_y - r_4T_x)f}{(r_5r_7 - r_4r_8)x_u + (r_1r_8 - r_2r_7)y_u + (r_2r_4 - r_1r_5)f} \tag{6.4}$$

For the two image points P_L and P_R, their corresponding undistorted image coordinates (x_{u1}, y_{u1}) and (x_{u2}, y_{u2}) can be calculated through Equ. 4.1 to Equ. 4.8. In the equations, the camera parameters should be a set of camera parameters which belong to their own camera.

For the left camera, $a1$, $b1$, $c1$ and $d1$ can be calculated through Equ. 6.1 to Equ. 6.4 by using the left camera calibration parameters and the undistorted image coordinates (x_{u1}, y_{u1}) of P_L.

For the right camera, $a2$, $b2$, $c2$ and $d2$ can be calculated through Equ. 6.1 to Equ. 6.4 by using the right camera calibration parameters and the undistorted image coordinates (x_{u2}, y_{u2}) of P_R.

The line $P_L P$ can be represented as follow:

$$X_w = a_1 \cdot Z_w + b_1 \tag{6.5}$$

$$Y_w = c_1 \cdot Z_w + d_1 \tag{6.6}$$

The line $P_R P$ can be represented as follow:

$$X_w = a_2 \cdot Z_w + b_2 \tag{6.7}$$

$$Y_w = c_2 \cdot Z_w + d_2 \tag{6.8}$$

The intersection point of these two lines is the point P. The coordinates of point P can be obtained by solving the above four equations Equ. 6.5, Equ. 6.6, Equ. 6.7 and Equ. 6.8.

From Equ. 6.5 and Equ. 6.7, we can get one solution of Z_w:

$$Z_w = \frac{b_2 - b_1}{a_1 - a_2} \tag{6.9}$$

From Equ. 6.6 and Equ. 6.8, we can get another solution of Z_w:

$$Z_w = \frac{d_2 - d_1}{c_1 - c_2} \tag{6.10}$$

From our experiences, the solution of Equ. 6.9 is always closer to the true value than the solution of Equ. 6.10. Furthermore, the solutions of Equ. 6.5 and Equ. 6.6 are also always closer to the true values than the solutions of Equ. 6.7 and Equ. 6.8 respectively.

After the calculation of Z_w, we can get the values of X_w and Y_w by using Equ. 6.5 and Equ. 6.6 respectively.

6.2 Method of Reconstructing a 3D Point

In this section, we demonstrate how to reconstruct a 3D point by using the binocular stereo system. The method of reconstructing a 3D point consists of the following three steps.

6.2.1 Calibration of Cameras

The binocular stereo system uses two cameras. Both of them need to be calibrated respectively. We use Tsai's calibration method to do camera calibration for these two cameras. The calibrated camera parameters must be carefully stored and can not be misused.

Both of world coordinate system and image coordinate system must be the same. In our experience, we used right hand system for both of the coordinate systems.

6.2.2 Calculation of Undistorted Image Coordinates

We refer to the left camera calibration parameters as 1 and the right camera calibration parameters as 2. Using the calibration parameters, we can calculate the distorted image coordinates (x_{v1}, y_{v1}) and (x_{v2}, y_{v2}) through the following equations:

$$
\begin{aligned}
x_{v1} &= \frac{x_{i1}d'_{x1} - c_{x1}}{s_{x1}} && (6.11) \\
y_{v1} &= y_{i1}d_{y1} - c_{y1} && (6.12) \\
x_{v2} &= \frac{x_{i2}d'_{x2} - c_{x2}}{s_(x2)} && (6.13) \\
y_{v2} &= y_{i2}d_{y2} - c_{y2} && (6.14)
\end{aligned}
$$

The undistorted image coordinates (x_{u1}, y_{u1}) and (x_{u2}, y_{u2}) can be calculated from the following equations:

$$
\begin{aligned}
x_{u1} &= x_{v1} + D_{x1} && (6.15) \\
y_{u1} &= y_{v1} + D_{y1} && (6.16) \\
x_{u2} &= x_{v2} + D_{x2} && (6.17) \\
y_{u2} &= y_{v2} + D_{y2} && (6.18)
\end{aligned}
$$

where $D_{xi} = x_{vi} \cdot (k_{1i}r^2 + k_{2i}r^4)$, $D_{yi} = y_{vi} \cdot (k_{1i}r^2 + k_{2i}r^4)$ and $r = \sqrt{x_{vi}^2 + y_{vi}^2}$.

6.2.3 Calculation of World Coordinates

After calculation of the undistorted image coordinates, we can calculate the parameters $a1$, $b1$, $c1$ and $d1$ of the left camera through Equ. 6.1 to Equ. 6.4 by using the left camera calibration parameters and its distorted image coordinates.

Furthermore, we can also calculate the parameters $a2$, $b2$, $c2$ and $d2$ of the right camera through Equ. 6.1 to Equ. 6.4 by using the right camera calibration parameters and its distorted image coordinates.

Finally, we can calculate the Z_w value by using Equ. 6.9 and calculate the X_w and Y_w by using Equ. 6.5 and Equ. 6.6 respectively.

6.3 Evaluation of Result

We only have one camera and one turntable. In order to evaluate the binocular stereo system, we use the turntable to rotate the calibration open cube certain degrees. The camera takes two images of calibration open cube at the different rotation angle. We can imagine that the calibration open cube is fixed and there are two cameras positioned on the two view directions.

The camera and the calibration object are setup as in Fig. 3.1. The distance between the camera and the calibration object is 2.8 meters and the camera angle θ is 10 degree. The working environment setup as in Sec. 3.1.

The two images taken can be used as the calibration images respectively. The world coordinates of calibration marks are the same for the two images taken. We use the first image data to calibrate the camera as the left camera. The first camera calibration parameters are as in Tab. 5.1. We use the second image data to calibrate the camera as the second virtual camera (right camera). The second virtual camera calibration parameters are as in Tab. 6.1.

In Tab. 6.1, f is the focal length of the camera. k_1 is the camera lens distortion coefficient. T_x is the X component of the translation vector. T_y is the Y component of the translation vector. T_z is the Z component of the translation vector. R_x is the rotation angle around X-axis. R_y is the rotation angle around Y-axis. R_z is the rotation angle around Z-axis. C_x is the x-coordinate of the image center point. C_y is the y-coordinate of the image center point. S_x is the horizontal scaling factor.

The reconstruction method mentioned in Sec. 6.2 is used. The 32 points of 3D world coordinates are reconstructed.

Figure 6.2 shows the evaluation result for the binocular stereo system. The horizontal axis represents the mark number of reconstructed point whose

Camera Parameter	Value
f	6.581203 [mm]
k_1	0.005649727 [$1/mm^2$]
T_x	-141.528361 [mm]
T_y	639.399258 [mm]
T_z	3285.625432 [mm]
R_x	-175.916652 [degree]
R_y	31.498625 [degree]
R_z	2.820074 [degree]
C_x	380.309823 [pixels]
C_y	293.730990 [pixels]
S_x	1.002447

Table 6.1: Second virtual camera calibration parameters

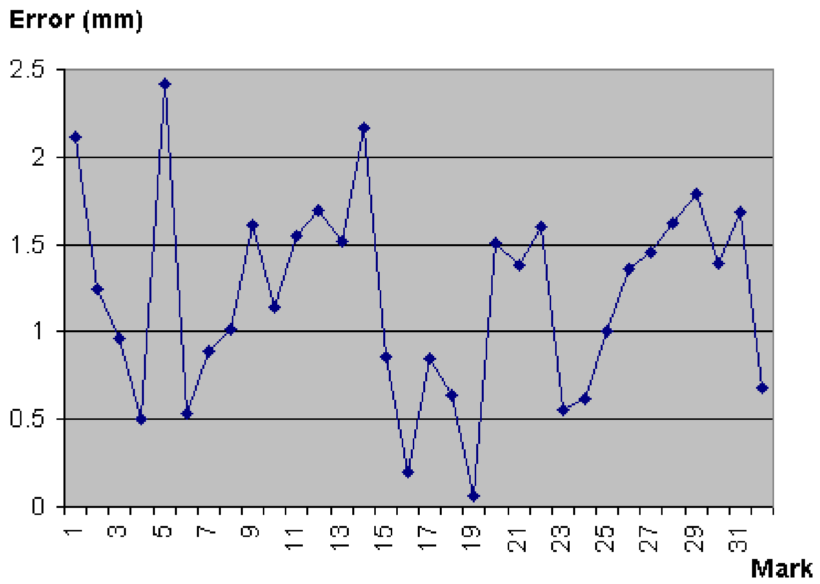

Figure 6.2: Evaluation result of the binocular stereo system

geometric position is shown in Fig. 3.7. The vertical axis is the distance between the reconstructed 3D point and its corresponding true point. The result is amazing. The maximum error is only 2.42 mm and the average error is only 1.21 mm. These two images are also used for the reconstruction by using the dynamic stereo method. Compared to the result of the dynamic stereo reconstruction method, the result of the binocular stereo reconstruction method is perfect.

Figure 6.3 shows the error distribution around the height of the calibration object. The horizontal axis is the height position of the calibration object. The vertical axis is the mean error, i.e. the average distance between the reconstructed points and their true points. It is clear that the mean error is smaller in the middle of the calibration object than in the low or high position of the calibration object.

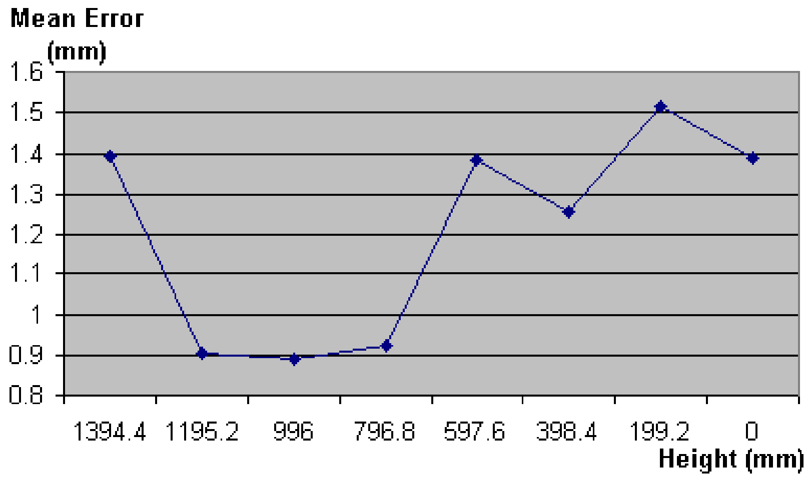

Figure 6.3: Mean error distribution around the height of calibration object

CHAPTER 7

Applications

The applications we developed are running in the Unix environment for a Silicon Graphics computer. The source codes are written in C language. It is developed based on the Tsai's calibration software which is provided by [2].

7.1 Application Running Environment

Our application are run under Unix operation system. It requires a Silicon Graphics computer. If the software needs to be moved to another platform, the programs *fmc* and *findrc* need to be rewritten because they use a image package in a Silicon Graphics computer. The functions dealing with RGB image format need to be removed and replaced with suitable functions.

The details of working environment setup is already described in Sec. 3.1.

7.2 Structures of Applications

The application we developed consisted of six functions. They can be used together or separately.

Figure 7.1 shows the flowchart of function fmc. The function fmc finds the world coordinates of the marks and also their corresponding image coordinates. It takes a RGB image file name as the only input parameter and outputs a data file named as *ncc_cd.dat*.

Figure 7.2 shows the flowchart of function $nccal\_fo$. The function $nccal\_fo$

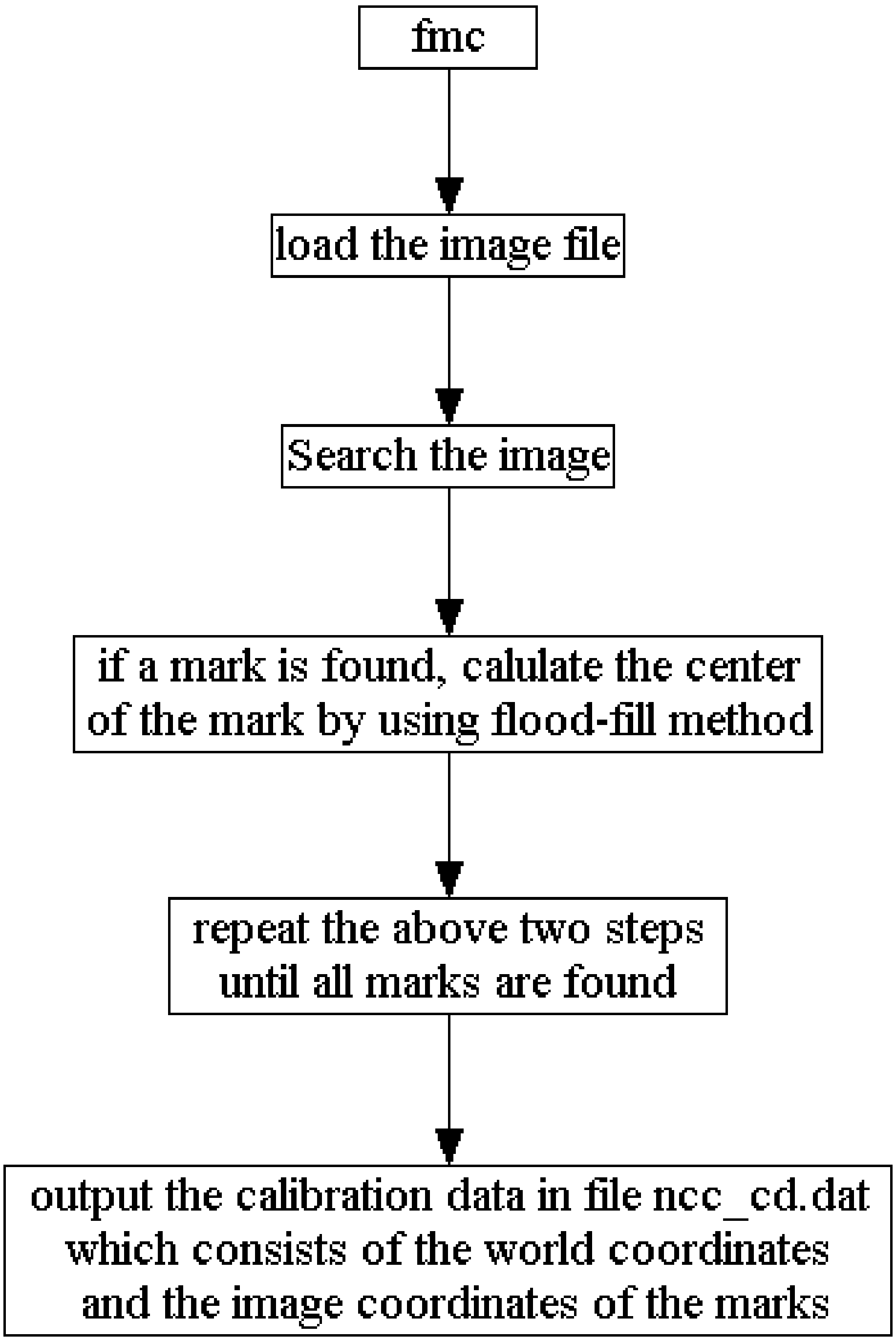

Figure 7.1: Flowchart of function fmc

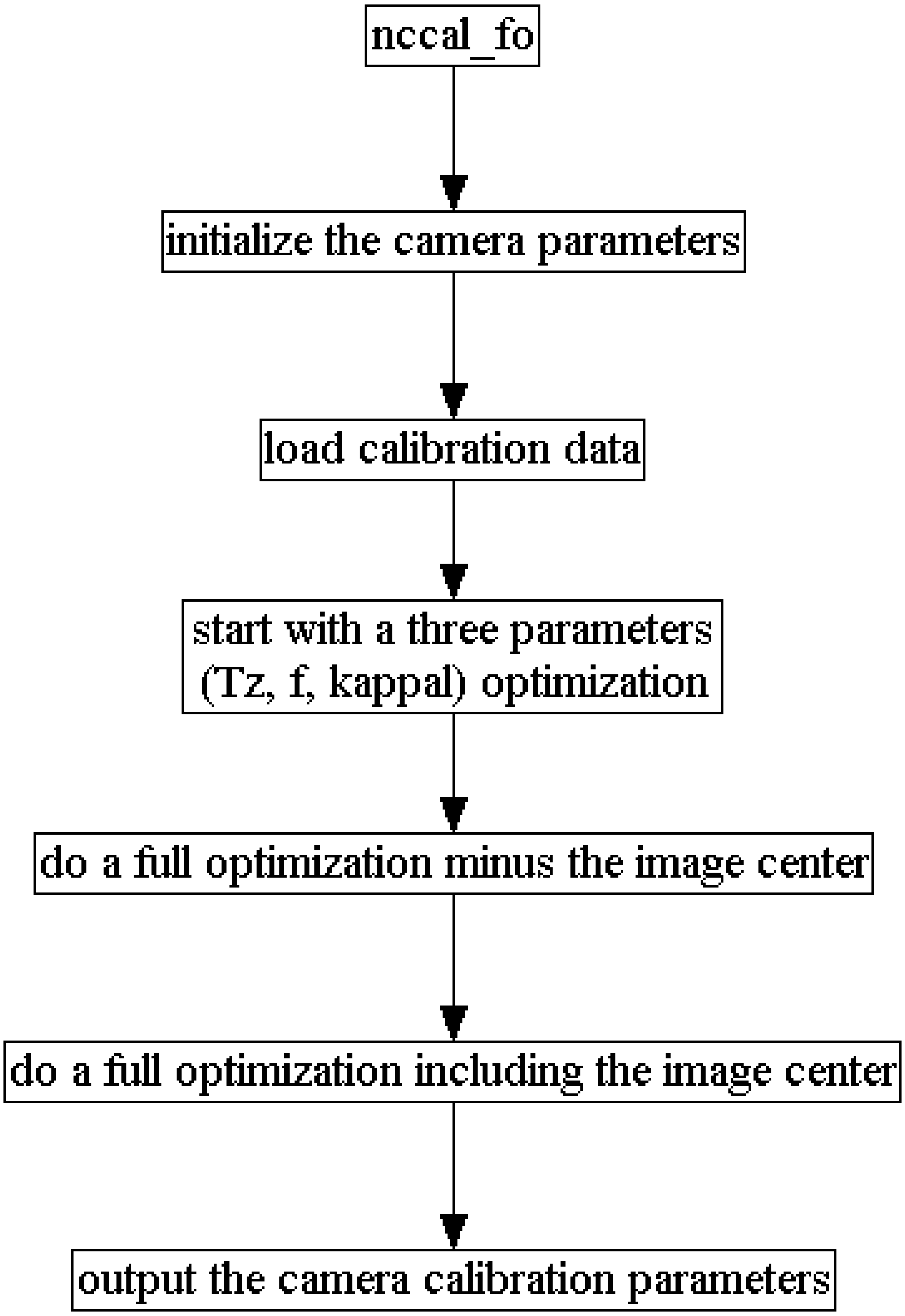

Figure 7.2: Flowchart of function $nccal\_fo$

calibrates the camera by using Tsai's calibration method. It takes a camera calibration data file name as the only input parameter and outputs the camera calibration parameters to screen (standard output).

Figure 7.3 shows the flowchart of function $findrc$. The function $findrc$ finds the rotation axis parameters. It detects the world coordinates of the intersection point of rotation axis and turntable. It takes the camera calibration parameters as the first input parameter and takes more than two RGB image file names as the rest input parameters. It then outputs a data file named as *tc.dat*.

Figure 7.4 shows the flowchart of function $ic2wc3d$. The function $ic2wc3d$ transfers a pair of image coordinates to their corresponding 3D world coordinates. It takes the camera parameters file name and the rotation parameters file name as the input parameters. It then prompts user to input a pair of image coordinates and then outputs the world coordinates to screen.

Figure 7.5 shows the flowchart of function $ic2wc3dw$. The function $ic2wc3dw$ transfers a pair of image coordinates to their corresponding 3D world coordinates. It takes the camera parameters file name and the rotation parameters file name as the input parameters. It then prompts user to input a pair of image coordinates and their rotation angle, and then outputs the world coordinates to screen.

Figure 7.6 shows the flowchart of function $biic2wc$. The function $biic2wc$ transfers a pair of image coordinates to their corresponding 3D world coordinates. It takes the first camera parameters file name and the second camera parameters file name as the input parameters. It then prompts user to input a pair of image coordinates and then outputs the world coordinates to screen.

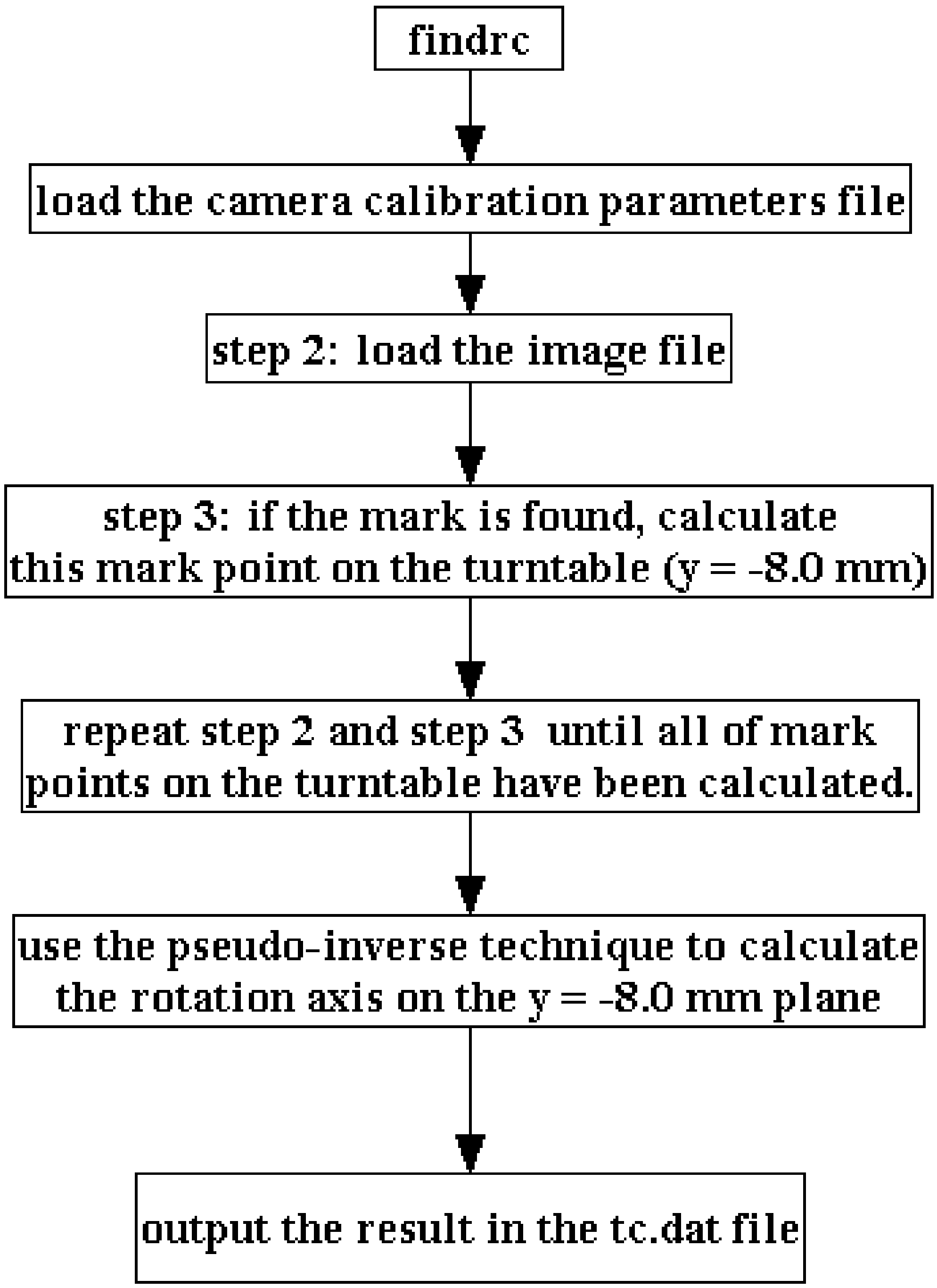

Figure 7.3: Flowchart of function $findrc$

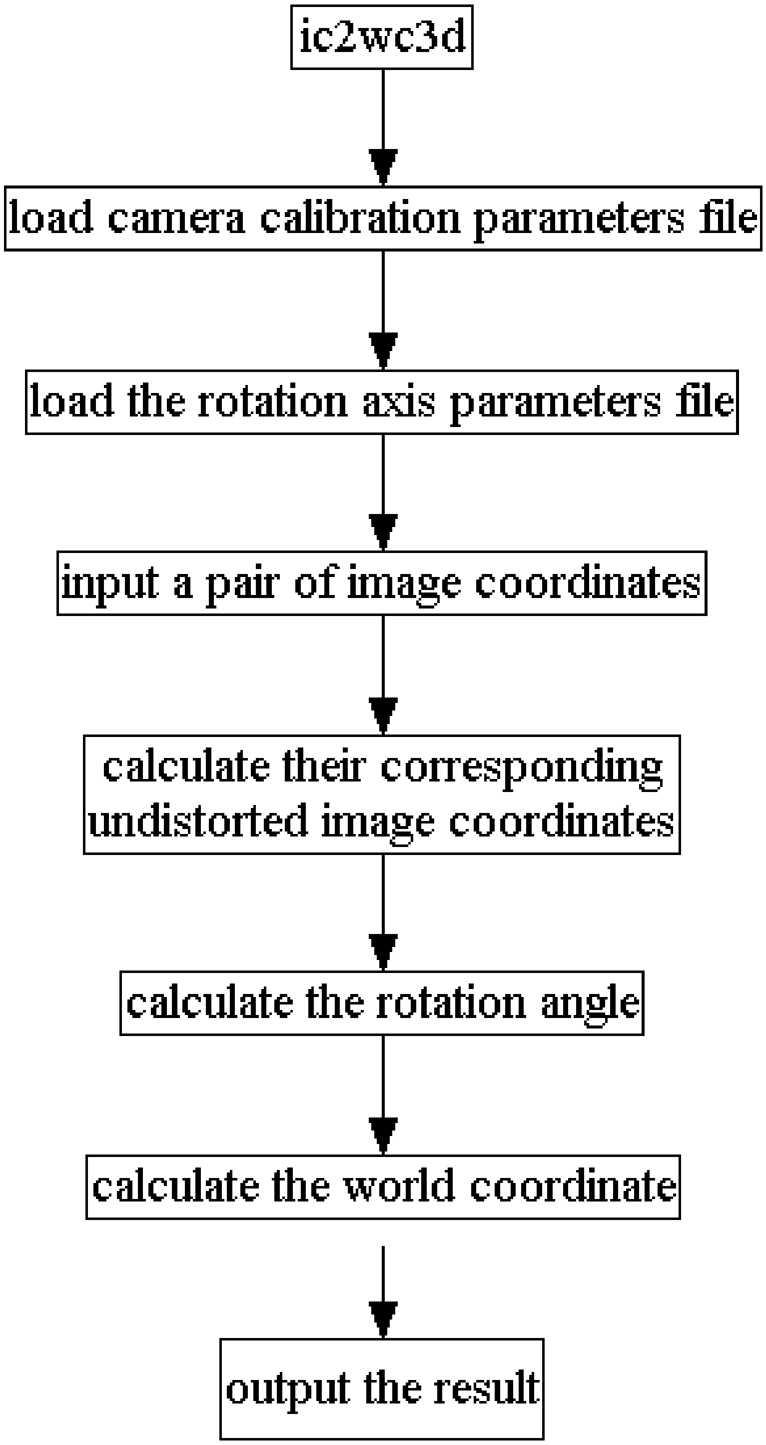

Figure 7.4: Flowchart of function *ic2wc3d*

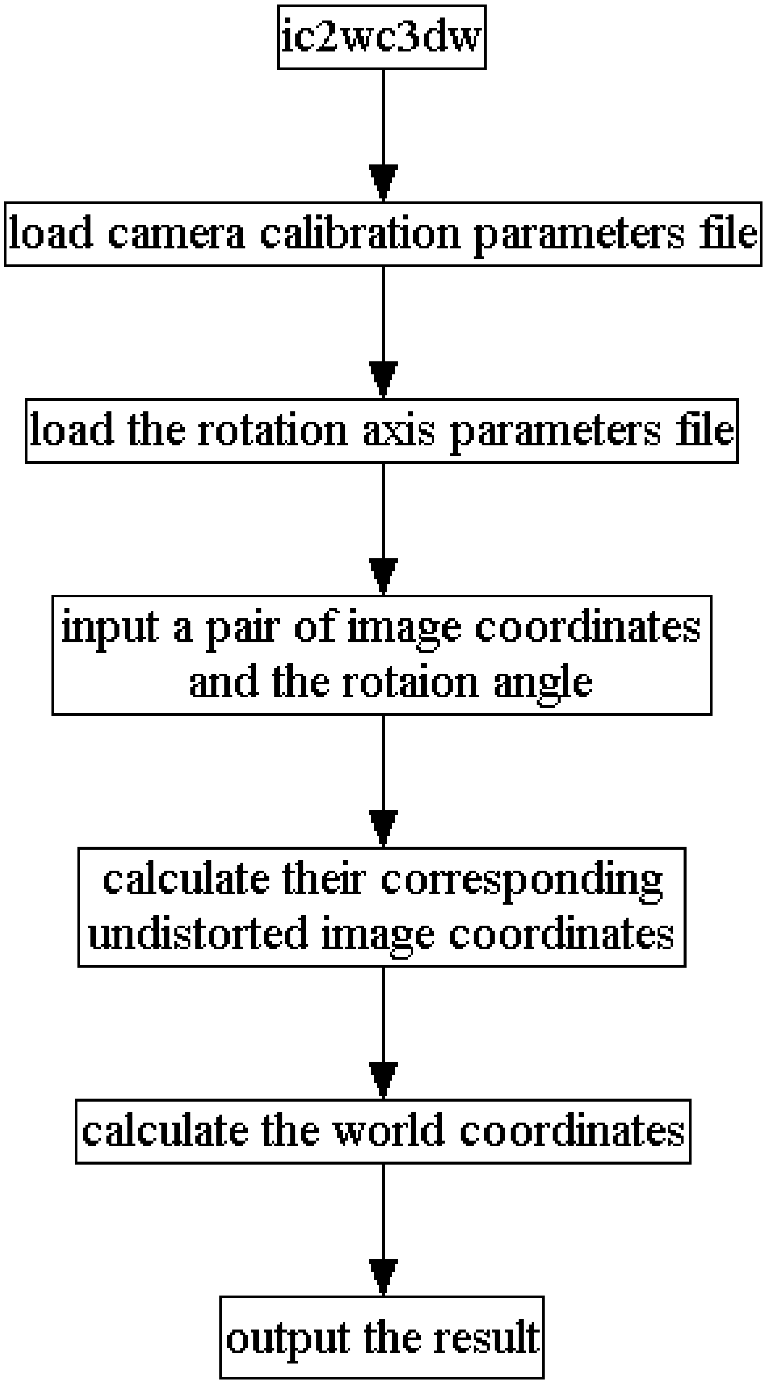

Figure 7.5: Flowchart of function *ic2wc3dw*

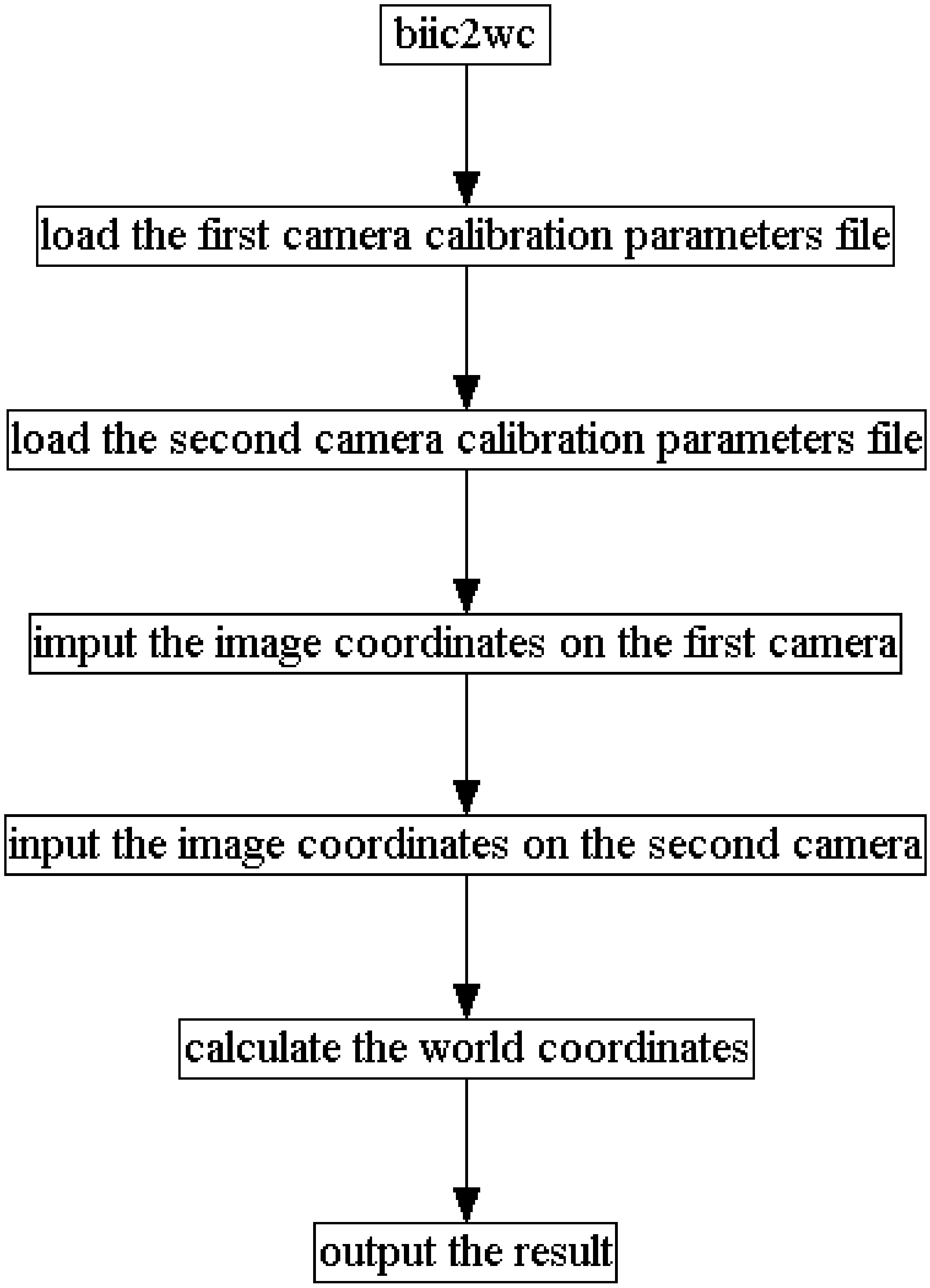

Figure 7.6: Flowchart of function *biic2wc*

7.3 Use of Method in the Context of PSM Based Shape Recovery

The application we developed can be used in the context of PSM based shape recovery. The use of method is different depending on what kind of model the PSM based shape recovery is using. This section demonstrates different usages for the PSM based shape recovery.

7.3.1 Dynamic Stereo Model

If the PSM based shape recovery uses the dynamic stereo model, the use of the calibration method is as follows:

Firstly, we need to set up the working environment as mentioned in Sec. 3.1. The positions of the turnable and the camera are adjusted in order to satisfy the requirements of the reconstructed object. The programs can be compiled by using the command:

make -f makefile.unx all

Secondly, the calibration open cube is put on the turntable. The image of calibration open cube is taken. The calibration open cube then can moved off the turntable. A paper with a calibration mark is put on the turntable. The image of this paper is taken. The turnable rotates certain degrees and more than two images of this paper at different rotation angle are taken. The mark on the paper should be uniformly rotated at the different angle.

The images taken above are preprocessed. The objects except the calibration marks are cut off from the images and also the threshold operation is applied to these images in order to deduct the noise on these images.

Thirdly, the function fmc is used to get calibration data. The command is as follow:

fmc imagefilename.rgb

imagefilename.rgb should be the name of the image file of calibration open cube. The file format should be RGB image format. It then produces the calibration data file called *ncc_cd.dat*.

Fourthly, the function *nccal_fo* is used to get calibration parameters. The command is as follow:

nccal_fo ncc_cd.dat > *ncc_cpcc.dat*

ncc_cd.dat is the name of camera calibration data file. *ncc_cpcc.dat* is the name of the camera calibration parameters file.

Fifthly, *findrc* is used to get the rotation axis parameters data. The command is as follow:

findrc ncc_cpcc.dat image1filename.rgb image2filename.rgb $\cdots$

image1filename.rgb, *image2filename.rgb* and etc. should be the names of the more than two image files of the calibration mark paper. The files format also should be RGB image format. *ncc_cpcc.dat* is the name of the camera calibration parameters file. It produces the rotation axis parameters file called *tc.dat*.

Finally, the function *ic2wc3d* is used to reconstruct a 3D point in the world coordinate system. The command is as follow:

ic2wc3d ncc_cpcc.dat tc.dat

ncc_cpcc.dat is the name of the camera calibration parameters file. *tc.dat* is the name of the rotation axis parameters file. It then prompts user to input a pair of image coordinates and then outputs the rotation angle and the corresponding world coordinates. It is very important that the first input image coordinates are from the image taken before the rotation and the other input image coordinates are from the image taken after the rotation.

If there are a lot of reconstructed points which are rotated at the same angle or the rotation angle is known, the function *ic2wc3dw* is used to reconstruct a 3D point in the world coordinate system. The command is as follow:

ic2wc3dw ncc_cpcc.dat tc.dat

ncc_cpcc.dat is the name of the camera calibration parameters file. *tc.dat* is the name of the rotation axis parameters file. It prompts user to input a pair of image coordinates and the rotation angle. It then outputs the corresponding world coordinates. The rotation angle is the average rotation angle or the known rotation angle. It is very important that the first input image coordinates are from the image taken before the rotation and the other input image coordinates are from the image taken after the rotation.

7.3.2 Binocular Stereo Model

If the PSM based shape recovery uses the binocular stereo model, the use of the calibration method is as follows:

Firstly, we need to set up the working environment as mentioned in Sec. 3.1. The two camera positions are adjusted in order to satisfy the requirements of the reconstructed object. The programs can be compiled by using the command:

make -f makefile.unx all

Secondly, the calibration open cube is put on the front of two cameras. The two images of calibration open cube from the two cameras are taken. These two images are preprocessed. The objects except the calibration marks are cut off from the two images and also the threshold operation is applied to these two images in order to deduct the noise on these two images.

Thirdly, the function fmc is used to get calibration data. The command is as follow:

fmc imagefilename.rgb

imagefilename.rgb should be the name of the image file of calibration open cube taken from the left camera. The file format should be RGB image format. It then produces the calibration data file called *ncc_cd.dat*.

The file *ncc_cd.dat* is renamed as *ncc_cd1.dat* and the above command is run again to process the image file taken from the right camera. The output camera calibration parameters is renamed as *ncc_cd2.dat*

Fourthly, the function *nccal_fo* is used to get calibration parameters. The command is as follow:

nccal_fo ncc_cd1.dat > *ncc_cpcc1.dat*

ncc_cd1.dat is the name of the left camera calibration data file. *ncc_cpcc1.dat* is the name of the left camera calibration parameters file.

The above command is run again to process the right camera calibration data. *ncc_cd1.dat* is replaced with *ncc_cd2.dat*. *ncc_cpcc1.dat* is replaced with *ncc_cpcc2.dat*.

Finally, the function *biic2wc* is used to reconstruct a 3D point in the world coordinate system. The command is as follow:

biic2wc ncc_cpcc1.dat ncc_cpcc2.dat

ncc_cpcc1.dat is the name of the left camera calibration parameters file. *ncc_cpcc2.dat* is the name of the right camera calibration parameters file. It then prompts user to input a pair of image coordinates and outputs the corresponding world coordinates. It is very important that the first input image coordinates are from the left camera image and the other input image coordinates are from the right camera.

CHAPTER 8

Conclusions

This thesis has built up two mathematical models for a dynamic stereo system (one camera + one turntable system) and a binocular stereo system. It deals with large scale objects such as human body. The thesis also evaluates these two calibration models and discusses the problems for these calibration models and the impact of some factors.

All of the models built up in this thesis use Tsai's calibration technique as the basic calibration technique and extend to their own calibration models.

There are many factors which may affect the calibration result. The following six factors can cause the different effects on the calibration result.

1. For a certain camera, the camera resolution for the calibration object decrease when the size of the calibration object increase. The calibration error on object space will increase when the size of the calibration object increase. On the other hand, the calibration error on image plan will decrease when the size of the calibration object increase. This means that the small size of calibration object requires more accurate calibration marks than the large size of calibration object.

2. The longer distance will increase the calibration error on the object space. It is better to keep the calibration object as close to the camera as possible.

3. Even thought increasing the number of calibration points can reduce the mean error of calibration, it will be useless for the accuracy of

calibration when the number of calibration points reach to a certain amount. It is better to use twenty-seven calibration points uniform distributed on three planes of open cube, i.e. there are nine calibration points uniformly distributed on every plane.

4. For a 3D calibration, there is a large distortion associated with the view direction. Mark 2 of Fig. 3.5 is the best choice as the calibration marks. We should use the first method mention in Sec. 3.3 to determine the center point of a mark.

5. We have to use a calibration object which will cover more than 80% of the reconstruction area.

6. The mean calibration error in the middle height of calibration object is the smallest. The farther the calibration points from the middle height of calibration object, the bigger calibration mean error they have.

For the dynamic stereo system, it assumes that the rotation axis is parallel to Y-axis. In fact, it is impossible to make the rotation axis absolutely parallel to Y-axis. Therefore, the higher object will result in a bigger error on the rotation axis. It will generate a little big reconstruction errors on the points around the intersection point of rotation axis and the plane parallel to turntable with $Y =$ camera height. For a rough reconstruction, the evaluation result for the dynamic stereo system shows that this model is quite acceptable for the reconstruction of a large scale object. In our case, for a 1.5 meter height object and the camera is 2.8 meters away from the object, it can reach the accuracy characterized by a mean error of 6.3 mm and a maximum error of 28.93 mm.

For the binocular stereo system, there is not any assumption at all. It suits any case. The result of evaluation is amazing. The camera is about

2.8 meters away from the calibration object which is about 1.5 meters in height. The maximum error on the calibration object is only 2.42 mm and the average error is only about 1.21 mm.

According to the evaluation result, for a large scale object, the accuracy of binocular stereo system is much better than dynamic stereo system. Nowaday the camera becomes much cheaper and also the computer's price always drops down. The computer networking speed increases and it is fast enough to transfer camera data to meet the requirement of reconstructing a 3D object. Therefore, we should abandon the camera + turntable system and use the binocular stereo system instead.

References

[1] R. Klette, K. Schlüns, A. Koschan: *Computer Vision - Three-Dimensional Data From Images.* Springer, Singapore, 1998.

[2] Tsai Camera Calibration Software, `http://almond.srv.cs.cmu.edu/afs/cs/usr/rgw/www/TsaiCode.html` (last visited: 15^{th} December, 2000)

[3] EHD 3 Chip RGB Color CCD Kamera KY-F55BE, `http://www.ehd.de/ky-f55.html` (last visited: 12^{th} April, 2000)

[4] Y. I. Abdel-Aziz, H. M. Karara: *Direct linear transformation into object space coordinates in close-range photogrammetry.* Proc. ASP Symposium on Close-Range Photogrammetry, Urbana, Illinois. USA, 1971, pp. 1-18.

[5] R. Y. Tsai: *An efficient and accurate camera calibration technique for 3D machine vision.* Proc. International Conference on Computer Vision and Pattern Recognition, Miami Beach Florida, USA, 1986, pp. 364-374.

[6] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery: *Numerical Recipes in C (FORTRAN, PASCAL).* 2nd Edition, Cambridge University Press, Cambridge, USA, 1992.

[7] N. Austin, Y. Chen, R. Klette, R. Marsh, Y.-s. Tsai, Y. Zhang: *A comparison of feature measurements for kinetic studies on human bodies.* Robot Vision, LNCS-1998, pp.43-51, 2001.

[8] 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. RA-3, No. 4, August 1987, pages 323-344.

[9] R. Swaminathan, S. K. Nayar: *Nonmetric Calibration of Wide-Angle Lenses and Polycameras.* IEEE Transactions on pattern analysis and and machine intelligence, Vol. 22, No. 10 October 2000, pages 1172-1178.

[10] Z. Zhang: *A Flexible New Technique for Camera Calibration.* IEEE Transactions on pattern analysis and and machine intelligence, Vol. 22, No. 11 November 2000, pages 1330-1334.

[11] J. Heikilä: *Geometric Camera Calibration Using Circular Control Points.* IEEE Transactions on pattern analysis and and machine intelligence, Vol. 22, No. 10 October 2000, pages 1066-1077.

[12] J. Weng, P. Cohen, M. Herniou: *Camera Calibration with Distortion Models and Accuracy Evaluation.* IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 14, No. 10, October 1992, pages 965-980.

[13] O. Faugeras: *Three Dimensional Computer vision: A geometric Viewpoint.* MIT Press, Cambridge, Massachusetts, USA, 1993.

[14] G. Xu, Z. Zhang: *Epipolar Geometry in Stereo, Motion and Object Recognition - A Unified Approach.* Kluwer, Dordrecht, 1996.

[15] L. Robert: *Camera calibration without feature extraction.* Computer Vision and Image Understanding 63 (1996), pp. 314-325.

[16] J. Heikkilä, O. Silvén: *A four-step camera calibration procedure with implicit image correction.* Proc. Int. Conf. on Computer Vision and Pattern Recognition 1997, San Juan, Puerto Rico, 1997, pp. 1106-1112.

[17] D. C. Brown: *Close-range camera calibration.* Photogrammetric Eng., Vol. 37, 1971, pp. 855-866.

[18] W. Faig: *Calibration of close-range photogrammetry systems: Mathematical formulation.* Photogrammetric Eng. Remote Sensing, Vol. 41, 1975, pp. 1479-1486.

[19] D. B. Gennery: *Stereo-camera calibration* in Proc. Image Understanding Workshop, 1979, pp.101-108.

[20] A. Isaguirre, P. Pu, J. Summers: *A new development in camera calibration: Calibration a pair of mobile cameras.* in Proc. Int. Conf. Robotics and Automation, 1985, pp. 74-79.

[21] J. Y. Luh, J. A. Klaasen: *A three-dimensional vision by off-shelf system with multi-cameras.* IEEE Trans. Pattern Anal. Machine Intell. Vol. PAMI-7, Jan. 1985, pp. 35-45.

[22] P. Grattoni, et al.: *Geometric camera calibration: a comparison of methods.* Proc. '91 ICAR-5th Internat. Conf. Advanced Robotics, Pisa, Jun. 1991, pp. 1775-1779.

[23] H. A. Martins, J. R. Birk, R. B. Kelley: *Camera Models Based on Data from Two Calibration Planes.* Computer Graphics and Image Processing, 17, 1981, pp. 173-180.

[24] I. Sobel: *On Calibrating Computer Controlled Cameras for Perceiving 3-D Scenes.* Artificial Intelligence, 5, 1974, pp. 185-198.

[25] Y. Yakimovsky, R. Cunningham: *A System for Extracting Three-Dimensional Measurements from a Stereo Pair of TV Cameras.* Computer Graphics and Image Processing, 7, 1986, pp. 195-210.

[26] R. K. Lenz, R. Y. Tsai: *Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology* Proc. IEEE Int. Conf. Robotics and Automation, Raleigh, NC, March 31-Apirl 3, 1987.

[27] M. Yachida, Y. Kitamura, M. Kimachi: *Trinocular vision: new approach for correspondence problem.* Proc. 8th International Conference on Pattern Recognition, Paris, France, 1986, pp.1041-1044.

List of Figures

List of Tables