A B-SPLINE MODEL FOR CAMERA CALIBRATION

By

YIMING XU

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2012

I dedicate this to my parents.

ACKNOWLEDGMENTS

I would like to thank all my friends, family, and colleagues for their support, which made this work possible.

I owe my greatest gratitude to my advisor Carl Crane for his guidance and encouragement.

I am thankful for the constructive feedback from Warren Dixon, John Schueller, Prabir Barooah, and Mrinal Kumar that helped make this a stronger thesis.

Many thanks to my friends and colleagues for the numerous helpful discussions and suggestions: Alan Hamlet, Darsan Patel, Joseph Osentowski, Karl Brandt, Ke Huo, Olugbenga Moses Anubi, Sujin Jang, Taeho Kim, Yilun Liu, Youngjin Moon, Yungsheng Chang, and others.

This is dedicated to my parents for their invaluable love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

A B-SPLINE MODEL FOR CAMERA CALIBRATION

By

Yiming Xu

May 2012

Chair: Carl Crane
Major: Mechanical Engineering

This thesis presents a general B-spline surface model for camera calibration. The model implicitly characterizes lens distortion by directly relating positions on the image plane to lines of sight in the world reference frame. Resolution for the model can be adjusted by the numbers of control vertices. Orders for the B-spline surfaces can also be chosen for improvement of accuracy without increasing degrees of freedom in the model. A calibration grid with known position and orientation is used for data acquisition, and calibration results are compared with actual measurements. An auxiliary method is provided to recover the relationship between the camera model and the world reference frame after the camera is moved. Both numerical and practical tests are conducted for demonstration purposes.

CHAPTER 1
INTRODUCTION

1.1 Existing Calibration Methods

Camera calibration has always been an important procedure for extraction of accurate metric information from images, obtaining optical characteristics of lens along with the relative position and orientation of cameras with respect to some global reference frame. Over the years, considerable ingenuity has been devoted to this subject, and an extensive number of calibration methods and camera models were developed to achieve satisfactory accuracy with manageable computation. A detailed review can be seen from [21][23][27][7].

Calibration methods can generally be characterized by the camera models they adopt. Some of the models, usually earlier ones, ignore lens distortion so that a simple solution for camera parameters can be rapidly generated. The direct linear transformation (DLT), developed by Abdel-Aziz and Karara [1], uses a 3-by-3 transformation matrix to describe the perspective projection from the world coordinate frame to the image plane. A similar calibration model is presented by Hall et al. [14], that uses a 3-by-4 matrix to transform the homogeneous coordinates between the frames. Further development and application of the method can be seen from [10]. Another example for non-distortion camera models is the two plane method [18], where the rays of light corresponding to positions on the image plane are uniquely determined by their intersection points with the two imaginary planes that are placed within view of the camera. In these calibration methods, it is assumed that the perspective projection relationship is linear, therefore only simple, linear equations need to be solved for the camera model. However, for applications where greater precision is required, inclusion of lens distortion is necessary and thus nonlinear projection models need to be formulated.

Models that include lens distortion may in turn fall into two categories, namely, explicit and implicit ones. Calibration methods with explicit models utilize explicit physically interpretable parameters for description of the camera's optical characteristics. In computer vision, camera models for popular calibration methods are usually derived from the pinhole model and compensate lens distortions with additional terms in the model. For example in Tsai's work [27], a second-order radial distortion is modeled based on disparities of the linear projection with the real one. The method developed by Heikkila and Silven [15] first obtains initial estimates with the DLT calculation, and then obtains the coefficients for radial and decentering distortion using a nonlinear least squares technique. Zhang's method [30] considers third and fifth order radial distortion, whereas Weng et al. [28] consider radial distortion, decentering distortion, and thin prism distortion in his model. These calibration methods, by using a handful of parameters in the camera models, are proven to be efficient (or even automatic) ways to achieve adequate precision in calibrating off-the-shelf cameras. Some other examples can be found in [2][26].

Another group of explicit camera models appears in calibration of catadioptric imaging systems, where images are formed through a combination of refraction and reflection via lenses (dioptrics) and curved mirrors (catoptrics). Depending on design intention and assembly accuracy, the rays of light entering such systems may not necessarily intersect at a single point, but rather pass through a locus in three dimensional space, referred as a caustic [26]. Study for such imaging systems can be found in [25][24][29], and will not be discussed extensively here.

Generally, calibration with explicit models requires some knowledge of the geometry and optical characteristics of the cameras in order to obtain an accurate outcome. Although high precision and some automation may have been achieved with these approaches, they are restricted for certain imaging systems and are inherently limited in their abilities to correct irregular or local distortions. On the contrary, implicit calibration

methods consider the imaging systems as a 'black box' and formulate the camera models without explicitly computing their physical and optical parameters [23]. Therefore models of this type are useless for extraction of camera parameters. However, exclusion of explicit physical parameters also relaxes the requirement for prior knowledge of the cameras, and allows for more flexibility and generality of the calibration methods. Implicit models seek to directly relate input and output of the imaging systems, which are essentially the rays of light in three dimensional space and the two dimensional positions on the image plane. A generic imaging model developed by Grossberg and Nayar [13] allocates optical properties to every pixel on the image plane, so that light rays can be determined from positions and directions of a caustic surface. Due to the fact that explicit models being 'black boxes' relaxes the constraints in camera's optical properties, the extra flexibility could make it more difficult to achieve automation in the calibration process. Despite this, Ramalingam et al. [20] presented an autonomous calibration method for central or slightly non-central cameras that is able to associate projection rays directly with image pixel indices, and to estimate positions and orientations of overlapping calibration grids with bundle adjustment. These two examples for implicit camera calibration methods both relate rays of light with discrete pixels on the image plane, which may be helpful for rapid computation and for cameras with discontinuous properties (for example a compound camera). However, for continuous image systems, since the target points from image processing might have sub-pixel level positions on the image plane, it might be favorable to adopt a continuous model to exploit better accuracy from interpolation.

1.2 Development of B-spline Model

The original two plane method can be regarded as an implicit model that does not rely on knowledge of optical lens characteristics, but it is limited in accuracy due to linear mapping. The method is extended by Gremban et al. [12] using quadratic transformation or triangular patches, in order to carry distortion information. In [5] and [6]

Champleboux et al. propose a mathematical model called n-planes B-spline (NPBS) using bi-cubic B-spline surface functions, which relate two dimensional coordinates on the image plane with three dimensional coordinates in the world frame. The NPBS method solves for coefficients for the cubic spline components of the surfaces by minimizing a linearized cost function that has one term to interpolate the data and the other to smooth the surface. Thanks to flexibility in bi-cubic B-spline surfaces, the model can theoretically compensate any continuous lens distortion up to the number of subdivisions chosen in its surface functions. Since it only accounts for correspondence between image positions and rays of light, the imaging system is not restricted to have a single viewpoint.

This thesis presents a model using general B-spline surfaces, which have tunable parameters such as the number of vertices and the order of the surfaces, in order that the model can be more flexible without the necessity to increase the degrees of freedom.

The contents of this thesis are presented as follows. Chapter 1 serves as an introduction. Chapter 2 describes the general B-spline camera model. Experimental test for camera calibration is conducted in Chapter 3. An auxiliary method for obtaining the position and orientation of the camera is presented in Chapter 4, together with a numerical test in Chapter 4 and a practical test in Chapter 5 to verify the practicality of the method. Chapter 6 presents a conclusion and further discussion.

CHAPTER 2
CAMERA MODEL

2.1 B-Spline Surfaces

Similar to the NPBS method, multiple surfaces are introduced that intersect with rays of light into the camera, and the points of intersection are associated with positions on the image plane with bi-parameter surface functions. The positions and shapes of the surfaces can be chosen arbitrarly, as long as they cover the field of view of the camera. For any position on the image plane, a point can be located on each of the surfaces, so that the points from multiple surfaces together can determine the rays of light cooresponding to the image positions. In this approach, the functions are general B-spline surfaces defined as [22][19]

$$\boldsymbol{Q}(u, w) = \sum_{i=1}^{n+1} \sum_{j=1}^{m+1} \boldsymbol{B}_{i,j} N_{i,k}(u) M_{j,l}(w), \tag{2–1}$$

that maps parameter pairs u and w to coordinates of points on the surfaces. The basis functions $N_{i,k}(u)$ and $M_{j,l}(w)$ are defined as

$$\begin{aligned}
N_{i,1}(u) &= \begin{cases} 1 & \text{if } x_i \leq u < x_{i+1} \\ 0 & \text{otherwise} \end{cases} \\
N_{i,k}(u) &= \frac{(u - x_i) N_{i,k-1}(u)}{x_{i+k-1} - x_i} + \frac{(x_{i+k} - u) N_{i+1,k-1}(u)}{x_{i+k} - x_{i+1}} \\
M_{j,1}(w) &= \begin{cases} 1 & \text{if } y_j \leq w < y_{j+1} \\ 0 & \text{otherwise} \end{cases} \\
M_{j,l}(w) &= \frac{(w - y_j) M_{j,l-1}(w)}{y_{j+l-1} - y_j} + \frac{(y_{j+l} - w) M_{j+1,l-1}(w)}{y_{j+l} - y_{j+1}}
\end{aligned} \tag{2–2}$$

where x_i and y_j are elements of uniform knot vectors which can be readily determined from the orders n and m of the surfaces, and range of the parameters [22]. If orders are set to be $k = l = 4$, then the model becomes bi-cubic B-spline surfaces as in the NPBS method. Intuitively, an extension to a general model brings more flexibility, which can be beneficial. In a later experiment test, it will be shown that it is sometimes necessary

to choose different orders for basis functions, so that better accuracy can be obtained, without increasing the degree of freedom of the model.

2.2 Surface Fitting

As mentioned earlier, the surfaces are to represent the mapping from position (u, w) on the image plane to point (x, y, z) in the world frame. For any given position (u, w), at least two B-spline surfaces are necessary to locate the corresponding line of sight in the world reference frame, by fitting a straight line to the points of intersection on the surfaces. The remaining task before the surfaces are available requires finding the coordinates for each $\boldsymbol{B}_{i,j}$, vertex of the polygonal control net of a B-spline surface. Since the basis functions are scalars, the three components in the coordinates can be treated separately. For example, the x-coordinate of vertex $x_{i,j}$ can be found using a linear least square approximation [16][4] by minimizing

$$x_{i,j} = \operatorname{argmin} \sum_{p} \left(\hat{x}_p - \sum_{i=1}^{n+1} \sum_{j=1}^{m+1} x_{i,j} N_{i,k}(u_p) M_{i,l}(w_p) \right)^2 , \tag{2–3}$$

where $\hat{x}_p$ is x-coordinate for the p-th data point from measurement, and u_p and w_p are corresponding image positions. It should be noted that there is no smoothing constraints in the cost function, so that the calculation is linear and that the surface can be fitted as closely to the data points as possible. Consequently, sufficient data points are required for the approximation to be solvable, that is, at least one data point should be registered within the control region of every vertex.

2.3 Formulation of B-spline Camera Model

With the help of a calibration grid, much efficiency can be achieved by simultaneously registering an array of points. For each position and orientation of the calibration grid, a B-spline surface can be fitted to the data. After at least two surfaces are defined, the lines of sight for any position (u, w) on the image can be located from the intersection points (x, y, z) from each of the surfaces. If there are more than two B-spline surfaces in the model, integrating them into two that uniquely determines the lines can reduce the

workload for online computation. Since coordinates of points on the surfaces are linear functions in terms of coordinates of the control vertices, for an ideal camera model with no error, vertex points at the same control knot in different surfaces should be collinear. Therefore, one can fit straight lines to the corresponding vertices, intersect them with two B-spline surfaces, and formulate the two surface models that uniquely defines the line of sight for any image positions (Figure 2-1). Here a parametric expression for a straight line is adopted to avoid scaling errors in between coordinate components.

$$x = a\,t + b, \quad y = c\,t + d, \quad z = e\,t + f. \tag{2–4}$$

Each line corresponds to the trace of a control vertex. Coordinates of vertices for the two resulting B-spline surfaces can be found by specifying two distinct parameters t_1 and t_2. Again, linear least square approximation is useful for fitting the straight lines.

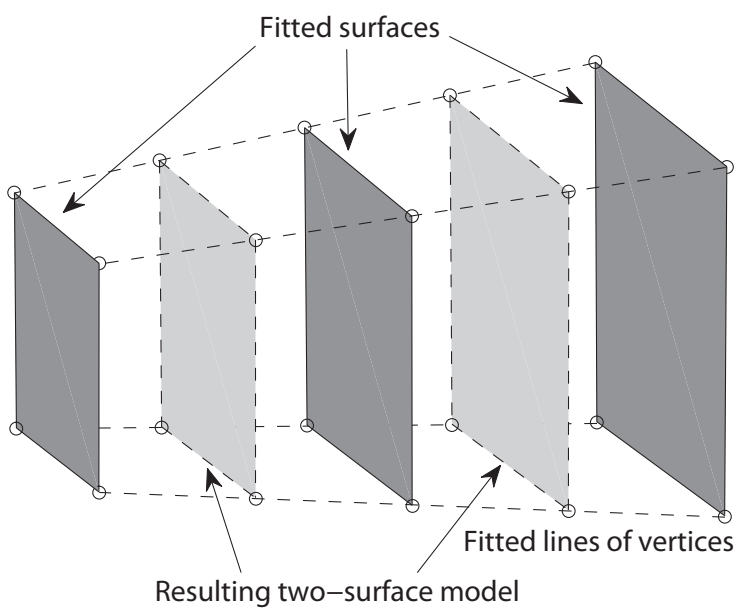


Figure 2-1. Formulation of a two-surface model by fitting lines to vertices.

CHAPTER 3
CALIBRATION TEST

An experiment test is carried out using a calibration grid and a slideway. B-spline surfaces of different orders are fitted to the data points, and the order of surfaces with the smallest error is chosen to formulate the camera model. The calibration model is compared with data point measurements that are not used in the modeling process. At the end of this chapter, the calibration model is then used to rectify one of the captured images.

3.1 Data Acquisition

In this calibration test, a CMOS digital video camera was used that transmits VGA (640×480) graphics. A 50×50 mm grid was printed on an A0 paper (1189×841 mm) which was pasted onto a wooden drawing board. The board was then installed onto a calibrated slideway for perpendicular movements (ranging from 800 to 1500 mm in front of the camera), as illustrated in Figure 3-1. The world coordinate frame was aligned to the device for convenience, and eight images were captured in accordance, as the calibration board traveled to eight different positions. Intersections of the grid were detected from the image as data points [9] (Figure 3-2). Obtained data points are summarized in Table 3-1.

Table 3-1. Summary of data points in the captured images

Image index	Z distance (mm)	Number of data points
1	579	330
2	659	325
3	739	302
4	819	257
5	899	224
6	979	188
7	1059	165
8	1139	130

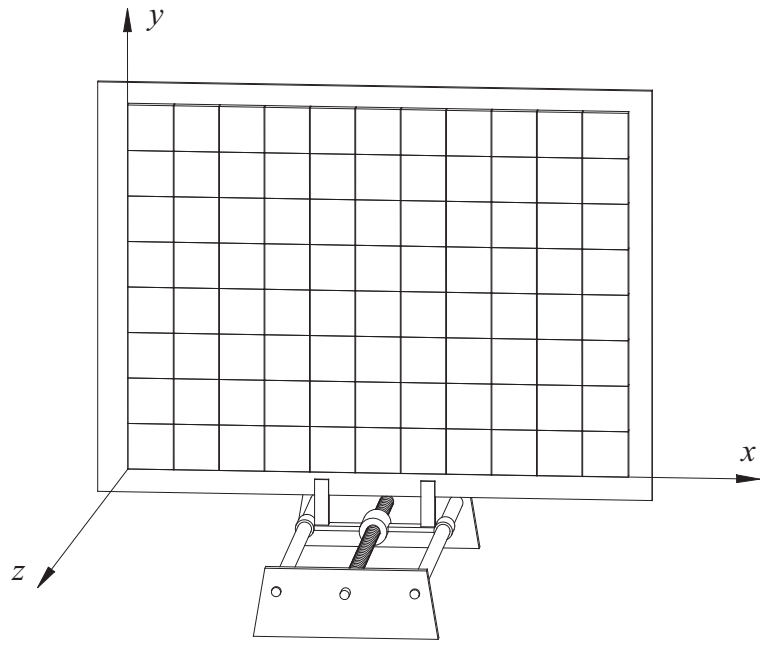
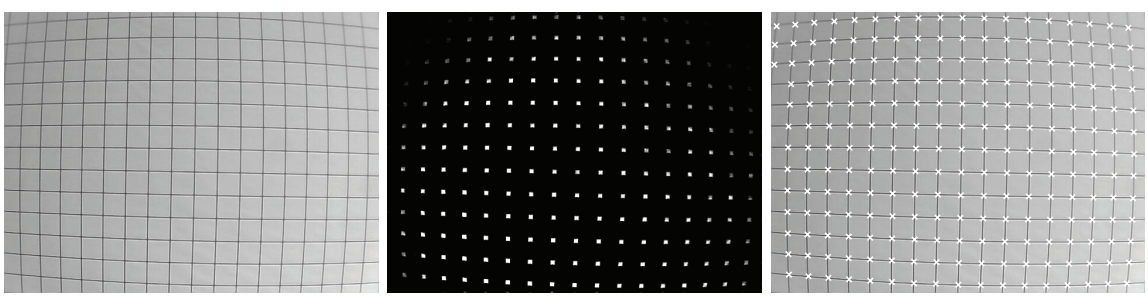


Figure 3-1. Schematic diagram of the calibration grid setup.



A Original image B Corner metric C Detected corner points

Figure 3-2. Process of corner detection (marked as white crosses) with image 5. (Photo courtesy of Yiming Xu.)

3.2 Model Formulation

B-spline surfaces were subsequently fitted to the acquired data points. In order to decide the best order for the surfaces, a fitting test was carried out with all the 8 images using different orders. The number of vertices was defined to be

$$(n+1) \times (m+1) = 7 \times 6. \tag{3–1}$$

Table 3-2. Surface fitting error of image 2 versus surface order

Order	Maximum error (mm)	Average error (mm)	Error variance (mm^2)
3	1.988	0.481	0.099
4	1.830	0.464	0.092
5	1.801	0.463	0.091
6	1.779	0.461	0.090

Table 3-3. Surface fitting error of image 5 versus surface order

Order	Maximum error (mm)	Average error (mm)	Error variance (mm^2)
3	1.427	0.312	0.046
4	1.318	0.297	0.041
5	1.253	0.298	0.040
6	1.213	0.298	0.038

The fitting error of image 2 and image 5 with different orders are displayed in Table 3-2 and Table 3-3 respectively.

In this particular case, the orders of 4 and 5 were found to have least fitting errors. By comparison between the two, a majority of 5th order surfaces had less average fitting error (for image 2, 3, 4, 7, and 8), and less maximum error (for image 2, 5, 6, 7, and 8). However, 4th order surfaces had the most significant error decrease from its neighboring lower order, which is greater than that of 5th order surfaces. Since the error differences between order 4 and 5 are comparatively small, an order of 4 would be used for constructing the B-spline camera model for less computation. Nevertheless, it should be noted that although this model had the same order as the NPBS method, it is always beneficial to have more choices for the model. In fact, based on different requirement and preference, one may model this camera using a different order.

Thus, except for image 3 and 6 that were reserved for error analysis, 4th order B-spline surfaces were fitted to the data from each of the remaining pictures: $k = l = 4$. The knot vectors are

$$\begin{aligned} x_i &: [0,\ 0,\ 0,\ 0,\ 160,\ 320,\ 480,\ 640,\ 640,\ 640,\ 640], \\ y_j &: [0,\ 0,\ 0,\ 0,\ 160,\ 320,\ 480,\ 480,\ 480,\ 480]. \end{aligned} \tag{3–2}$$

The fitting result of image 5 and its error vectors are shown in Figure 3-3 and 3-4. Results for the six images are in shown in Figure 3-5, and the fitting error is summarized in Table 3-4.

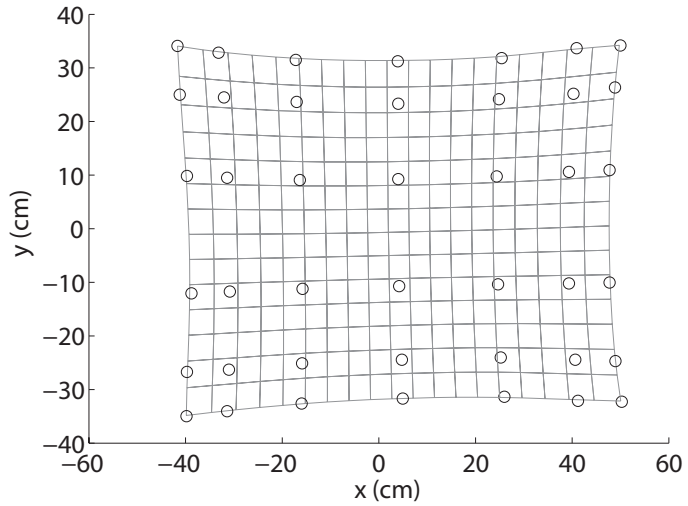


Figure 3-3. B-spline surface fitted to data points in image 5. Circles represent control vertices of the B-spline surface.

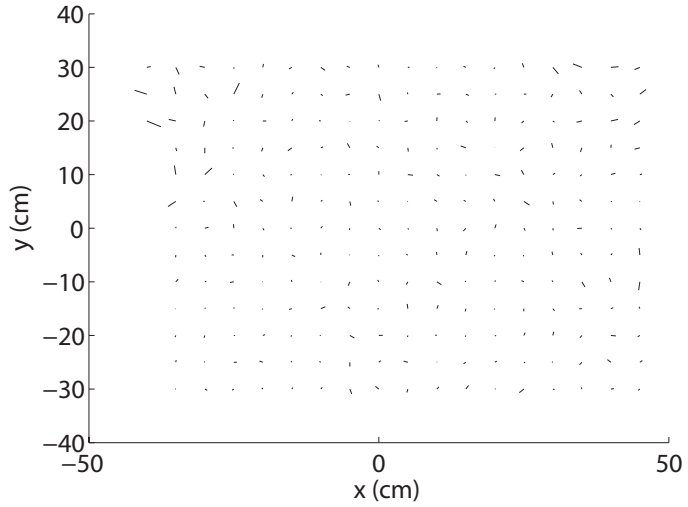


Figure 3-4. Surface fitting error in image 5, with 20 times of original magnitudes.

From the resulting six B-spline surfaces, a straight line was fitted to vertices at each corresponding control knot. Intercepting the lines at the very two planes that contain data points from the reserved images (image 3 and 6), such that the resulted

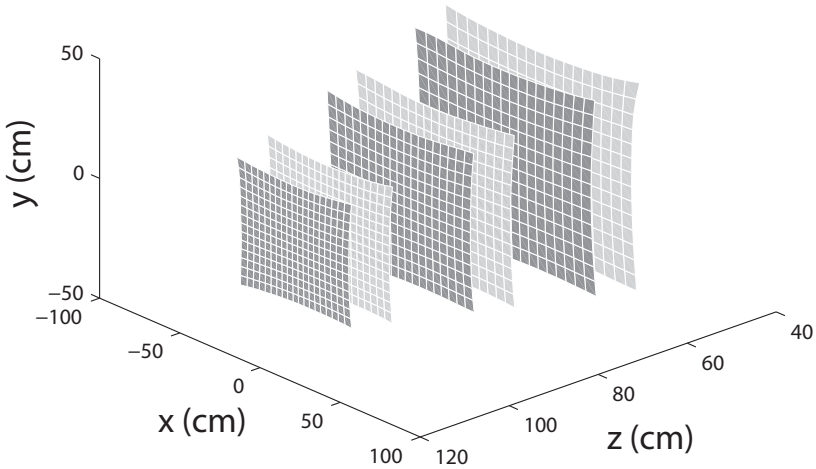


Figure 3-5. B-spline surfaces fitted for image 1, 2, 4, 5, 7, and 8, from right to left.

Table 3-4. Surface fitting error summary

Image index	Maximum error (mm)	Average error (mm)	Error variance (mm^2)
1	1.742	0.475	0.083
2	1.830	0.464	0.093
4	2.086	0.347	0.076
5	1.318	0.297	0.041
7	0.882	0.226	0.020
8	0.413	0.169	0.008

two surface camera model (Figure 3-6) could be compared directly to the measured data in image 3 and 6.

3.3 Error Analysis

Since the two surfaces are so placed that they coincide with the planes of data points in image 3 and 6, the calibration error could be evaluated by comparing the coordinates of the measured data points on the grid with the coordinates for the detected corner points in the images, which can be obtained by substituting the corresponding image positions into the B-spline surface functions.

Calibration error is defined by vectors from the measured positions of the corner points to the calculated ones using the camera model. Here in Figure 3-7 and 3-8, these

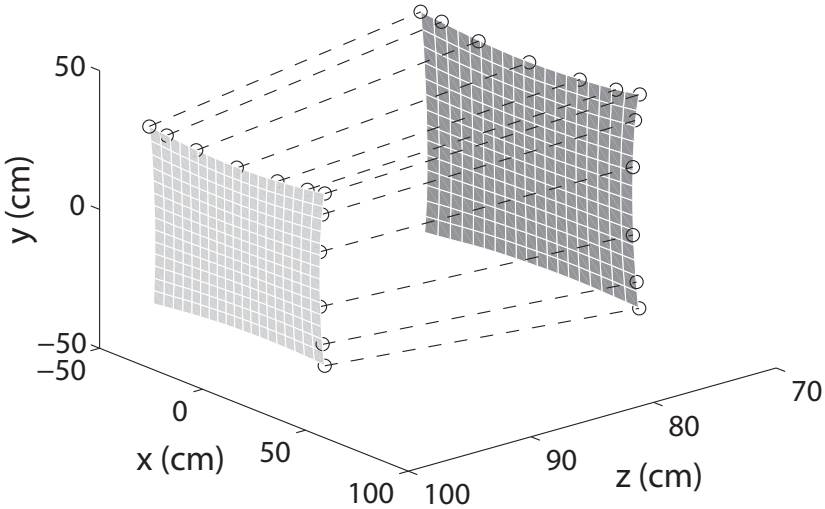


Figure 3-6. Two surface camera model resulted from fitting lines to control vertices.

Table 3-5. Calibration error summary

Picture index	Maximum error (mm)	Average error (mm)	Error variance (mm^2)
3	3.16	0.73	0.11
6	2.21	0.90	0.19

vectors are displayed with 20 times their original magnitudes. The lengths of the error vectors, or scalar calibration error, are presented in Table 3-5.

There are numerous sources of error that may contribute to the disparities between outcomes from calibration and measurement. First, the number of control vertices and orders of the surfaces in the camera model may be chosen differently to yield better accuracy, which is unlikely because the fitting error (Table 3-4) does not appear to be comparatively significant. Second, resolution of the camera as well as precision of the corner detection techniques may affect accuracy. As one can see in Figure 3-2, lines of the calibration grid at corners of the image appear fuzzy and the extracted corner metric at the corresponding places is vague. These all bring difficulties and of course inaccuracy for image processing. Last but not least, the quality of the camera model heavily depends on the accuracy of the acquired data. Flatness of the calibration board,

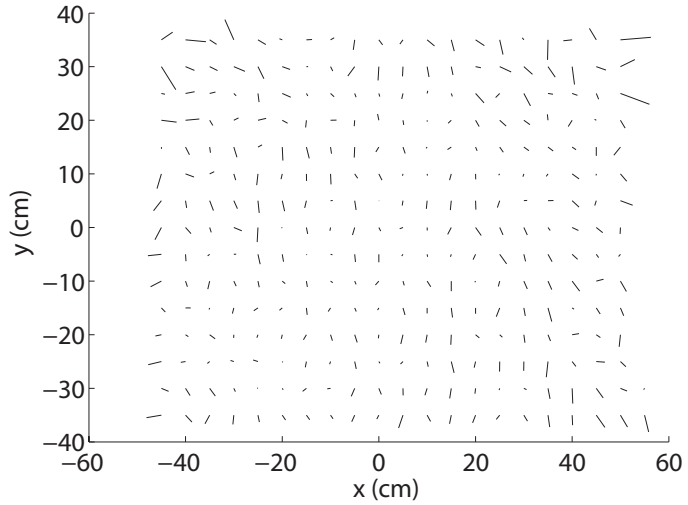


Figure 3-7. Vectors of calibration error for image 3, with 20 times of original magnitudes.

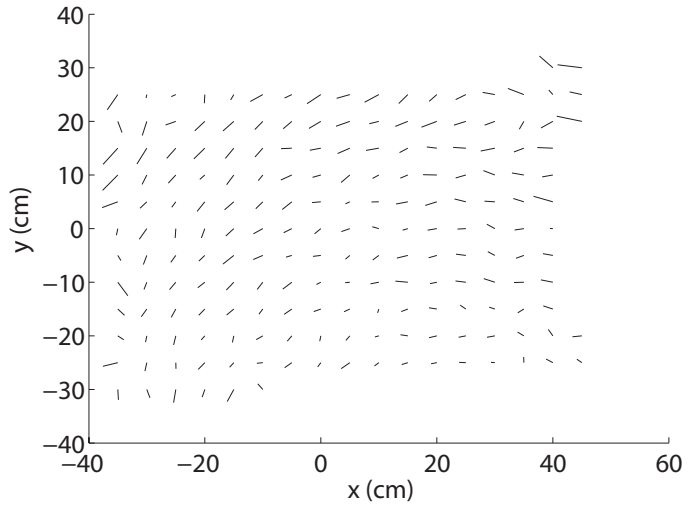


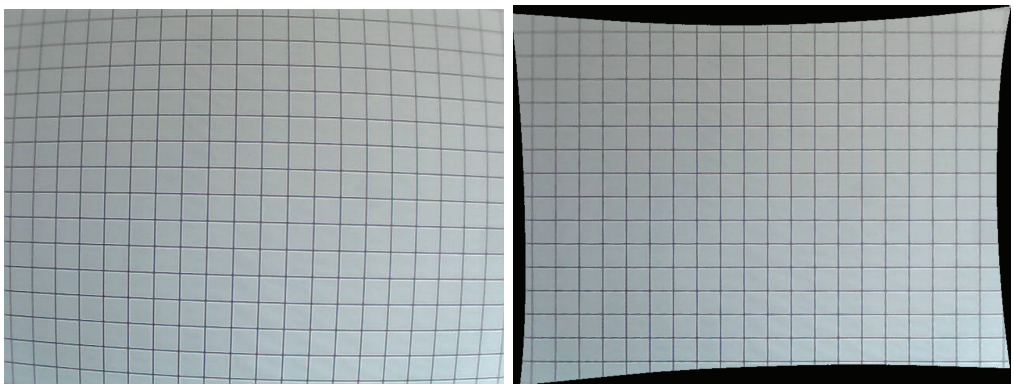
Figure 3-8. Vectors of calibration error for image 6, with 20 times of original magnitudes.

printing precision of the grid, and accuracy of coordinate measurements, to name a few, are key factors for the overall performance of the calibration method. Unfortunately, the device available for the experiment test has only limited accuracy. Considerable error was introduced due to the fact that the grid board could not be rigidly positioned on the slideway, therefore the planes of data points were not as precisely parallel to each other as they were registered to be.

A rough estimation of the overall error due to imprecise measurements would be around 0.5 mm. In comparison, the calibration error is adequately small for objects within the calibrated range. Predictably, a better camera model requires a better slideway and calibration board.

3.4 Image Rectification

With this camera model, the original image can be projected along the lines of sight onto any surface, and form a new image. If the surface is defined to be a flat plane, the resulting image will be devoid of the distortions compensated by the camera model. To demonstrate this, image 3 was thus projected onto the corresponding plane of the calibration grid where it is originated from, and a rectified image was generated as shown in Figure 3-9.



A Original image B Rectified image

Figure 3-9. Rectification of image 3 with B-spline camera model. (Photo courtesy of Yiming Xu.)

CHAPTER 4
CAMERA POSE ESTIMATION

Up to this point, a B-spline camera model has been formulated with respect to some fixed reference frame. The reference frame of the camera model (denoted as Frame C) coincides to the fixed world frame (denoted as Frame W), if the camera remains fixed after its calibration. However, as the camera moves, Frame C would be moved away from Frame W. A possible way to recover the calibration result is to find the transformation matrix [8] from Frame C to the current Frame W, which can be accomplished by using the camera model and a characteristic object with known geometry and known position and orientation in Frame W.

4.1 Least Squares Pose Estimation

Suppose the object has n $(n \geq 3)$ non-collinear target points that have known positions in Frame W, and are easily recognizable from the image. Based on image position (u_k, w_k) of the k-th target point, the corresponding line of sight with respect to Frame C can be located using the camera model. Since the target points are located on the cooresponding lines of sight, their coordinates in Frame C can be written in parametric expressions of the lines defined in (2–4). The position vector for the k-th target point in Frame C, ${}^{C}\boldsymbol{P}_k$, can be written in an expression with known parameters a_k through f_k and an unknown variable t_k. The position vector for the point with respect to Frame W, ${}^{W}\boldsymbol{P}_k$, is known as assumed.

$$ {}^{C}\boldsymbol{P}_k = \begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \begin{bmatrix} a_k t_k + b_k \\ c_k t_k + d_k \\ e_k t_k + f_k \end{bmatrix}, \quad {}^{W}\boldsymbol{P}_k = \begin{bmatrix} p_k \\ q_k \\ r_k \end{bmatrix}. \tag{4–1} $$

Since distances between any combination of the target points should be identical in both coordinate frames, parameter t_k can be determined from minimizing the following cost

function with nonlinear least square approximation [17].

$$t_k = \operatorname{argmin} \frac{1}{2} \sum_{i \neq j} \left(\| {}^{C}\boldsymbol{V}_{ij} \|^2 - \| {}^{W}\boldsymbol{V}_{ij} \|^2 \right), \tag{4–2}$$

where vectors ${}^{C}\boldsymbol{V}_{ij}$ and ${}^{W}\boldsymbol{V}_{ij}$ are vectors between Point i and Point j:

$$ {}^{C}\boldsymbol{V}_{ij} = {}^{C}\boldsymbol{P}_i - {}^{C}\boldsymbol{P}_j, \quad {}^{W}\boldsymbol{V}_{ij} = {}^{W}\boldsymbol{P}_i - {}^{W}\boldsymbol{P}_j. \tag{4–3}$$

Substituting t_k into (4–1) gives coordinates for the k-th target point in Frame C. Since coordinates of the target points in both frames are obtained, the transformation matrix between Frame W and Frame C can be obtained by linear least square techniques. The rotation matrix ${}^{C}_{W}\boldsymbol{R}$ is estimated using vectors between the target points.

$$ {}^{C}_{W}\boldsymbol{R} = \operatorname{argmin} \frac{1}{2} \sum_{i \neq j} \| {}^{C}\boldsymbol{V}_{ij} - {}^{C}_{W}\boldsymbol{R}\, {}^{W}\boldsymbol{V}_{ij} \|^2. \tag{4–4}$$

Vector from the origin of Frame W to the origin of Frame C can be calculated as

$$ {}^{C}_{W}\boldsymbol{V} = \frac{1}{n} \sum_{k} {}^{C}\boldsymbol{P}_k - {}^{C}_{W}\boldsymbol{R} \frac{1}{n} \sum_{k} {}^{W}\boldsymbol{P}_k. \tag{4–5}$$

Thus the transformation matrix can be written as

$$ {}^{C}_{W}\boldsymbol{T} = \begin{bmatrix} {}^{C}_{W}\boldsymbol{R} & {}^{C}_{W}\boldsymbol{V} \\ \boldsymbol{0}^{\top} & 1 \end{bmatrix}. \tag{4–6}$$

One should note that this algorithm of finding the transformation matrix does not in itself guarantee orthogonal properties of the rotation matrix, but for the overall result to be accurate, input data should at least be accurate, which meantime ensures validity of the outcome properties.

4.2 Numerical Test

To evaluate the algorithm of finding transformation matrices, a numerical test was conducted using the data points from the reserved image 3 and 6. Three points were randomly selected as target points from each image, as shown in Figure 4-1, and their

coordinates (in Frame C) were transformed to a new reference frame (denoted as Frame W) with a transformation matrix defined as

$$
{}^{C}_{W}\boldsymbol{T}_0 = \begin{bmatrix} 0.7036 & 0.7036 & -0.0998 & 20.0000 \\ -0.6547 & 0.6964 & 0.2940 & 18.0000 \\ 0.2764 & -0.1415 & 0.9506 & -5.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}; \tag{4–7}
$$

which represents an XYZ rotation from Frame W to Frame C by angles of $17.2°$, $5.73°$, and $45.0°$, followed by a translation of $[20,\ 18,\ -5]^{\top}$.

Corresponding positions on the images were substituted into the B-spline camera model to locate lines of sight with respect to Frame C. Thus, using the aforementioned method, coordinates of the target points with respect to Frame C were estimated with nonlinear least square approximation. A hybrid algorithm of Levenberg-Marquardt (L-M) and Quasi-Newton (Q-N) was used, for fast convergence and globally robust iterations [17]. Due to the fact that the rays of light are likely to be in similar directions (or roughly parallel), an ambiguous result from the estimation is possible. Depending on the initial guesses for t_k, one may obtain a true estimate, or a false estimate that 'mirrors' the true one along the rays of light, as shown in Figure 4-2 and 4-3 respectively.

It is true in this case, that a good initial guess for parameter t_k is important. However, a false result can be easily identified and disambiguation is possible. Essentially, the false result represents a local minimum of the cost function, whose value is significantly greater than that for the true estimate (Table 4-1). Therefore, even if a good initial guess is not available, one can always try another 'mirror' guess and compare the results.

Table 4-1. Comparison between the true and false estimation results

	Maximum error (mm)	Average error (mm)	Error variance (mm^2)	Cost function
1	3.5802e+2	2.3593e+2	1.4775e+4	62.9069
2	1.0899	0.6643	0.0593	9.2942e-4

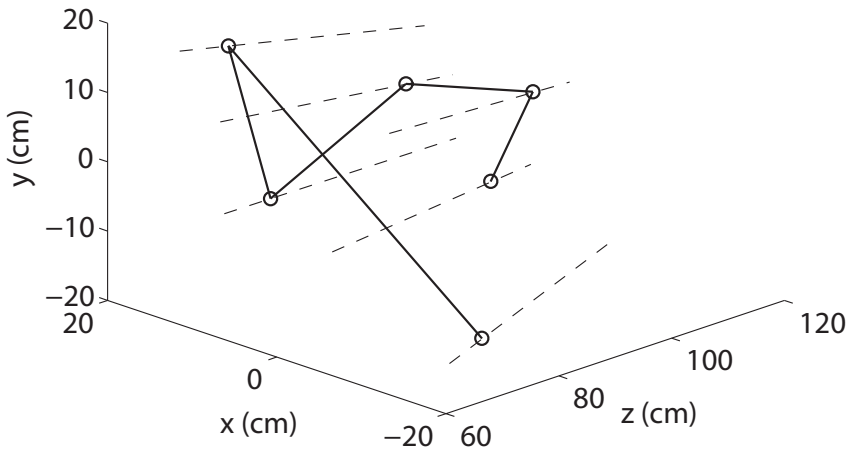


Figure 4-1. Target points sampled from image 3 and 6.

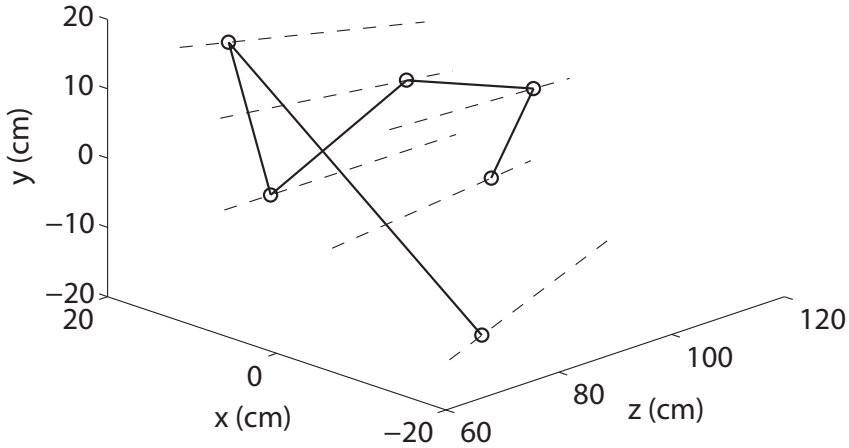


Figure 4-2. A true estimation of target points.

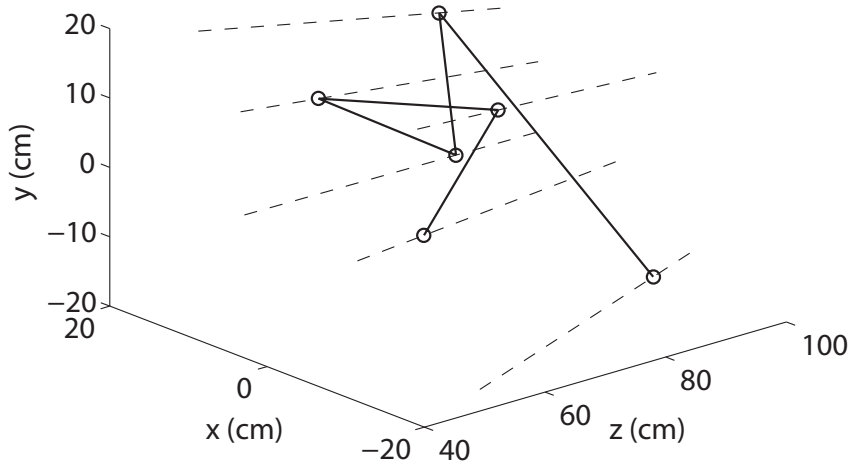


Figure 4-3. A false estimation of target points.

The transformation matrix was subsequently computed based on the global result.

$$
{}^{C}_{W}\boldsymbol{T} = \begin{bmatrix} 0.7015 & 0.7058 & -0.0991 & 19.9941 \\ -0.6563 & 0.6959 & 0.2918 & 18.1914 \\ 0.2743 & -0.1400 & 0.9513 & -5.0534 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}, \tag{4–8}
$$

which represents an XYZ rotation from Frame W to Frame C by angles of $17.1°$, $5.69°$, and $45.2°$, followed by a translation of $[19.9,\ 18.2,\ -5.1]^{\top}$.

The transformation matrix was a good estimate of the one in (4–7), considering that modeling error was also introduced in the calculations. If an ideal camera model (with no calibration error), the method could yield an estimation result with an error on the level of floating point relative accuracy.

CHAPTER 5
TARGET EXTRACTION

In this chapter, image processing techniques are demonstrated for extraction of target points from a characteristic object, for completeness of the calibration method. A remote control vehicle was used as the characteristic object, as shown in Figure 5-1.

Color segmentation was performed in the first place to obtain the target from the input image. The image was transformed into HSV space, in which color region for the target is more regular than that in RGB space. The color region is shown in Figure 5-2 and Figure 5-3, and can be described by the following functions.

$$
\begin{aligned}
&|h - 0.04| \leq 0.09, \\
&s \geq 0.03, \\
&v \geq 0.1467\, s^2 - 0.873\, s + 0.9413.
\end{aligned}
\tag{5–1}
$$

The largest connected component in the image that satisfies the description was considered to be the target vehicle.

Figure 5-1. Characteristic object for feature extraction. (Photo courtesy of Yiming Xu.)

With binary morphological operations [3][11], parts of the vehicle such as the windscreen, engine cover, and side windows, can be extracted from the main body and

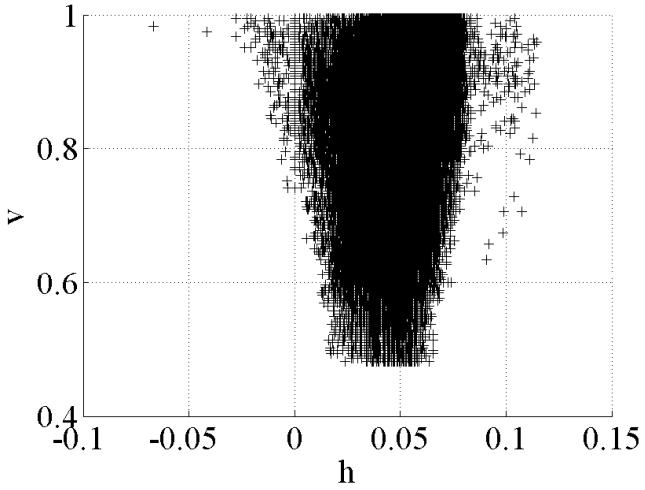


Figure 5-2. Color range of target on H-V plane.

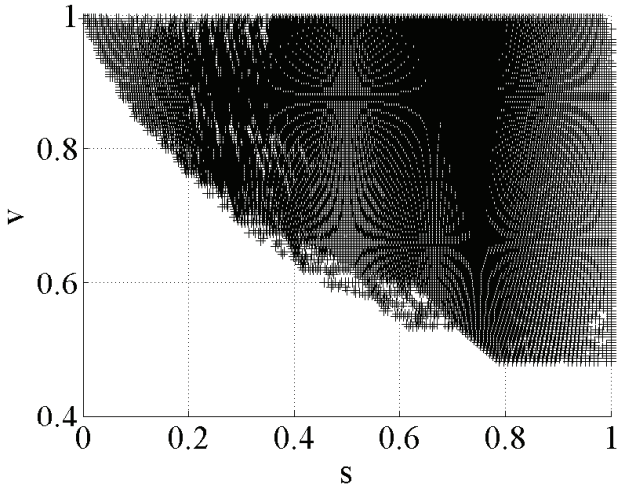


Figure 5-3. Color range of target on S-V plane.

separated by their morphological features. Image positions for these target components were represented with a polar coordinate system that originates from the centroid of main body. Thus, relative position and orientation of these components in the polar reference frame are useful to further determine the direction of the target vehicle. Then each individual component can be recognized based on its relative location. For this test, six feature points were chosen as target points, namely, four corner points on the

windscreen and two corner points on the engine cover. Image processing result for the original image (Figure 5-1) is shown in Figure 5-4.

Figure 5-4. Feature extraction result for the target vehicle.

This feature extraction process requires that the background has a different color than the target vehicle and that the camera maintains a downward-looking perspective. Surely the image processing strategy should accommodate the application requirements, and the one used here has proved to have an adequate robustness. However, a precise measurement of the feature points on the vehicle is not available at the moment, therefore it is not possible to integrate this with the pose estimation algorithm and form an error analysis for the entire calibration process.

CHAPTER 6
CONCLUSION

In this thesis, a general B-spline camera model is defined, which implicitly embodies lens distortions and directly relates points on the image plane to lines of sight in the world reference frame, without prior knowledge of optical characteristics of the camera. It extends the camera model in NPBS method to one with freedom to choose orders for the B-spline surfaces, which introduce more flexibility without necessarily increasing degrees of freedom in the model. Accuracy of the calibration depends on measurement and a good choice for the vertex number and the orders of B-spline surfaces. The experimental tests have shown adequate accuracy, given that the precision of input data is limited by the device that was used. For relocated cameras, an auxiliary method is also presented, to recover the transformation relationships of the camera frame and the world frame, with the help of a characteristic object with known geometry and pose in the scene.

Although high accuracy can be achieved with the B-spline camera model, calculations with B-spline surfaces are certainly not advantageous over linear camera models. If computing speed is rendered of overwhelming importance, it is necessary to pre-calculate the lines of sight for an adequate amount of sample positions, thus saving the need for online calculation of B-spline surfaces. Furthermore, generality and flexibility of the camera model allows for less constraint on the imaging systems, which also bring difficulties for developing an autonomous calibration technique. Even though the auxiliary method is provided to recover the camera model, so that calibration only need to be perform once. The calibration process is still cumbersome and requires high accuracy measurements. Therefore, the B-spline camera model needs further development for an easier, more automatic calibration method.

REFERENCES

[1] Abdel-Aziz, Y. I. and Karara, H. M. "Direct linear transformation from comparator coordinates into object space coordinates." *Proceedings of American Society for Photogrammetry and Remote Sensing* 1 (1971).

[2] Basu, A. and Ravi, K. "Active camera calibration using pan, tilt, and roll." *Proceedings of IEEE International Conference on Robotics and Automation* 3 (1995): 2961–2967.

[3] Born, M. and Wolf, E. *Principles of optics*. Permagon Press, 1965.

[4] Buchanan, J. and Turner, P. *Numerical methods and analysis*. McGraw-Hill, 1992.

[5] Champleboux, G., Lavallée, S., Sautot, P., and Cinquin, P. "Accurate calibration of cameras and range imaging sensor: The NPBS method." *Proceedings of IEEE International Conference on Robotics Automation* (1992): 1552–1557.

[6] Champleboux, G., Lavallée, S., Szeliski, R., and Brunie, L. "From accurate range imaging sensor calibration to accurate model-based 3-D object localization." *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (1992): 83–89.

[7] Clarke, T. A. and Fryer, J. G. "The development of camera calibration methods and models." *The Photogrammetric Record* 16 (1998).91: 51–66.

[8] Crane, C. and Duffy, J. *Kinematic Analysis of Robot Manipulators*. Cambridge University Press, 2008.

[9] Davies, E. R. *Machine vision: theory, algorithms, practicalities*. Morgan Kaufmann, 2005, 3rd ed.

[10] Faugeras, O. D. and Toscani, G. "The calibration problem for stereo." *Proceedings of IEEE Computer Vision and Pattern Recognition* (1986): 15–20.

[11] Gonzalez, R. and Woods, R. *Digital image processing*. Prentice Hall, 2007, 3rd ed.

[12] Gremban, K. D., Thorpe, C. E., and Kanade, T. "Geometric camera calibration using system of linear equation." *Proceedings of IEEE of International Conference on Robotics and Automation* 1 (1988): 947– 951.

[13] Grossberg, M. D. and Nayar, S. K. "A general imaging model and a method for finding its parameters." *Proceedings of the 8th International Conference on Computer Vision* 2 (2001): 108–115.

[14] Hall, E. L., Tio, J. B. K., McPherson, C. A., and Sadjadi, F. A. "Measuring curved surfaces for robot vision." *Computer* 15 (1982).12: 42–54.

[15] Heikkilä, J. and Silvén, O. “A four-step camera calibration procedure with implicit image correction.” *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (1997): 1106C1112.

[16] Hoffman, J. *Numerical methods for engineers and scientists*. Marcel Dekker, 2001, 2nd ed.

[17] Madsen, K., Nielsen, H. B., and Tingleff, O. *Methods for nonlinear least squares problems*. Informatics and Mathematical Modeling, Technical University of Denmark, 2004.

[18] Martins, H., Birk, J., and Kelley, R. “Camera models based on data from two calibration planes.” *Computer Graphics and Image Processing* 17 (1981): 173–180.

[19] Piegl, L. A. and Tiller, W. *The NURBS book*. Springer, 1997.

[20] Ramalingam, S., Sturm, P., and Lodha, S.K. “Towards complete generic camera calibration.” *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1 (2005): 1093–1098.

[21] Remondino, F. and Fraser, C. “Digital camera calibration methods: considerations and comparisons.” *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (2006).5: 266–272.

[22] Rogers, D. *An introduction to NURBS: with historical perspective*. Morgan Kaufmann, 2001.

[23] Salvi, J., Armangué, X., and Batlle, J. “A comparative review of camera calibrating methods with accuracy evaluation.” *Pattern Recognition* 35 (2002).7: 1617–1635.

[24] Sturm, P. “Multiview geometry for general camera models.” *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2005): 206–212.

[25] Swaminathan, R., Grossberg, M. D., and Nayar, S. K. “Non-single viewpoint catadioptric cameras: geometry and analysis.” *International Journal of Computer Vision* 66 (2006).3: 211–229.

[26] Triggs, B. “Autocalibration from planar scenes.” *Proceedings of the 5th European Conference on Computer Vision* (1998): 89–105.

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

[28] Weng, J., Cohen, P., and Herniou, M. “Camera calibration with distortion models and accuracy evaluation.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1992): 965–980.

[29] Ying, X. and Hu, Z. “Catadioptric camera calibration using geometric invariants.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26 (2004).10: 1260–1271.

[30] Zhang, Z. “A flexible new technique for camera calibration.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (2000).11: 1330–1334.

BIOGRAPHICAL SKETCH

Yiming Xu received his B.S. degree in 2010 from the Department of Control Science and Engineering at Zhejiang University. He is currently studying as a graduate student in Department of Mechanical and Aerospace Engineering at University of Florida. His research interest includes camera calibration, image processing, and dynamic system control.