# A Study on the Image Based 3D Modeling by Weighted Bi-directional Registration

Hyo Sung Kim[1], Yong Gi Park[1], Ki Gon Nam [1], Sung Wook Seol[1], Jae Heum Joo[2]
[1]Department of Electronics Engineering, Pusan National University, Pusan, Korea
[2]School of Information Engineering, Catholic University of Pusan, Pusan, Korea
ygpark@vision3d.ee.pusan.ac.kr

## Abstract

The image-based 3D modeling is the technique of generating a 3D graphic model from images acquired from cameras. It is being researched as an alternative technique for the expensive 3D scanner. In this paper, we propose the image-based 3D modeling system using calibrated stereo camera and the weighted bi-directional method that is robust against error in the 3D registration. The proposed algorithm for rendering 3D model consists of three steps; camera calibration, 3D reconstruction, and 3D registration step. In the camera calibration step, we estimate the camera matrix for the image acquisition camera. In the 3D reconstruction step, we estimate 3D coordinates using triangulation from corresponding points of stereo image. In the 3D registration step, we estimate the transformation matrix that transform individually reconstructed 3D coordinates to the reference coordinate to render the single 3D mode by proposed algorithm. As shown in the result, we generated relatively accurate 3D model that is robust against error.

**Keywords**: 3D Reconstruction, Camera Calibration, Registration, 3D Modeling

## 1 Introduction

Recently, due to the rapid development of personal computers and wide-spreading high speed internet, we desire a 3D multimedia than a 2D one. To build a virtual world such as a 3D animation, a 3D game and a virtual shopping mall, we should build a 3D model.

The technique of the 3D modeling is classified into an active type and a passive type by the sensing method [1]. The active sensing is the technique that extracts depth information by directly projecting a laser beam or a pattern on to the object. It is used the 3D scanner normally. This method has high accuracy, but the facilities are very expensive. So this is difficult to be popularized and has several restrictions to adopt the object and place. The passive sensing is the technique that extracts 3D information from acquired images from cameras. This method has low accuracy, but it has studied hard in the field of the computer vision due to being realized with low price equipments.

The passive sensing is classified into the calibrated and uncalibrated reconstruction [2]. The 3D reconstruction from calibrated images is reconstructed to Euclidean coordinate, but it is impossible for the 3D modeling to be reconstructed from uncalibrated images. The 3D reconstruction from uncalibrated images does not need any calibrating process with camera, but the 3D modeling could not be reconstructed to complete Euclidean coordinate and very sensitive to the input information (matching points).

In this paper, we propose the low-cost image based 3D modeling system with the robust algorithm against errors in the 3D registration. To build the 3D modeling similar to the real object, we render the 3D reconstruction from calibrated images. Our system is composed of the calibration camera step, the 3D reconstruction step and the 3D registration step. In camera calibration step, the camera matrix is obtained by using the calibration pattern. The 3D coordinates are reconstructed by the triangulation from stereo matching points in the 3D reconstruction step. In the 3D registration step, the 3D coordinates are mapped from other views to the reference view. The method of the registration is mostly used in the 2D image mosaicing [3] and is applied to the 3D modelling. The weighted bi-directional registration is used to solve the accumulated errors in this paper. We perform the texture mapping and generate the realistic 3D modeling.

The paper is organized as follows. Section 2 introduces the image-based 3D modeling system. Section 3 describes the camera calibration. Section 4 describes the reconstruction of 3D coordinate. Section 5 describes the registration of 3D coordinate. Section 6 gives experimental results. Finally, section 7 concludes the paper.

## 2 Image-Based 3D Modeling System

The block diagram of the proposed image based 3D modeling system is shown in Fig. 1. The system is composed of the camera calibration, the 3D reconstruction and the 3D registration step. The camera matrices from the left and right cameras are obtained by using the calibration patterns in the camera calibration step. In the 3D reconstruction step, the 3D coordinates which are reconstructed by using the triangulation from the stereo matching points and the calculated camera matrices are converted to Euclidean coordinate. As transforming the 3D coordinates which are reconstructed from each stereo image to the reference coordinates, the final single 3D modeling is generated in the 3D registration step. In the 3D registration step, the final single 3D modeling is generated through transforming the 3D coordinates that are reconstructed from each stereo image to the reference coordinate.

## 3 Camera Calibration

### 3.1 Projective Projection Camera Model

Generally, the CCD camera is a projective or a pinhole camera [4]. The point $\mathbf{X} = [X\ Y\ Z\ 1]^T$ in the 3D coordinate is projected to the point $\mathbf{x} = [u\ v\ 1]^T$ in the image plane through the camera matrix $\mathbf{P}$ [5] and it is represented by

$$\mathbf{x} \cong \mathbf{K}[\mathbf{R} \mid \mathbf{T}]\mathbf{X} = \mathbf{P}\mathbf{X} \qquad (1)$$

where $\mathbf{P}$ is 3x4 matrix and it is composed of the internal parameter matrix $\mathbf{K}$, the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{T}$. $\cong$ is equal in homogeneous coordinate.

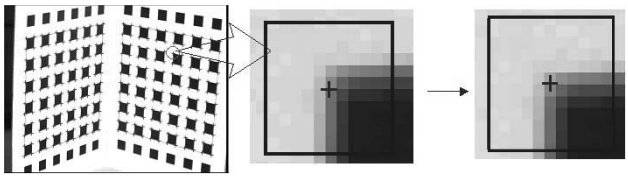
### 3.2 Feature Extraction and Assignment of 3D Coordinate

To obtain $\mathbf{P}$ matrix, the used calibration patterns are composed of rectangular patterns, which are orthogonal each other. The Fig.2(a) shows the calibration pattern image which is acquired from the camera. The corner points of each pattern are selected as the feature. For extracting the corner points, the Harris corner detector [6] is used. According to the second-order differential values from intensity in the image, the Harris corner detector determines corners. To subdivide the pixel unit, the central point of the second-order differential is estimated using Eq(2). The Fig 2 (c) shows that the position of the initial feature point in Fig 2 (b) moves to the corner of the rectangular by Eq(2).

$$\begin{bmatrix} \sum_i\sum_j g_{xx}(i,j) & \sum_i\sum_j g_{xy}(i,j) \\ \sum_i\sum_j g_{xy}(i,j) & \sum_i\sum_j g_{yy}(i,j) \end{bmatrix} \begin{bmatrix} c_x \\ c_y \end{bmatrix} = \begin{bmatrix} \sum_i [g_{xx}(i,j)u_i] + \sum_j [g_{xy}(i,j)v_i] \\ \sum_i [g_{xy}(i,j)u_i] + \sum_j [g_{yy}(i,j)v_i] \end{bmatrix} \qquad (2)$$

The extracted feature point is assigned to the 3D point correspond to the corner position like Fig 3. The Fig 3 (a) shows the assignment of X, Z coordinate in 3D

where $Z_I$ is $45^\circ$, L is 13mm, and $Z_I$ is 500mm. The Fig 3 (b) shows the assignment of Y coordinate in 3D.



(a) (b) (c)

**Fig. 2.** Feature extraction.
(a) Calibrated pattern image.
(b) Initial position of feature.
(c) Refined position of feature.

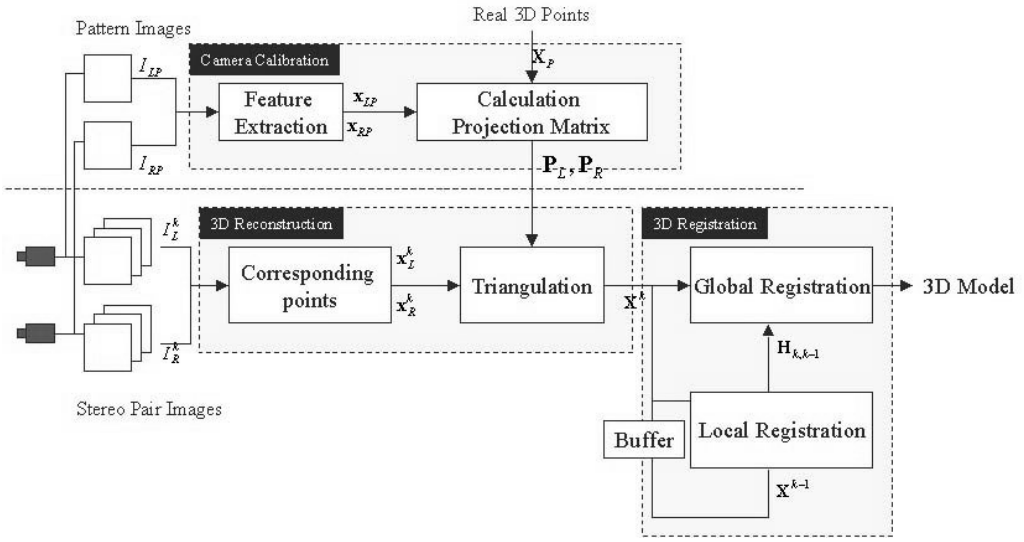
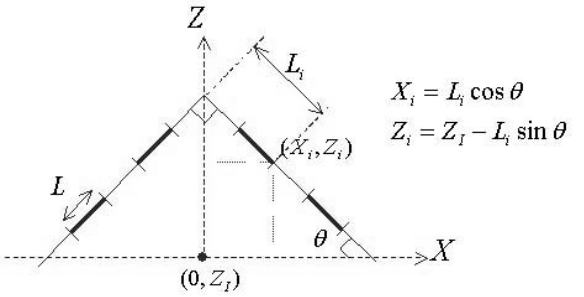
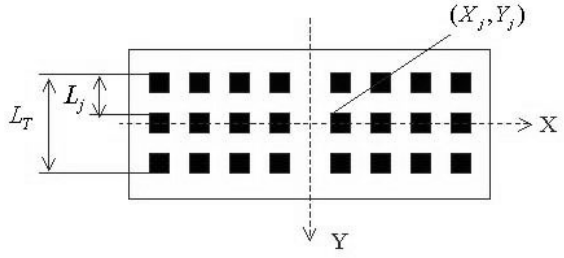



**Fig.1.** Block diagram of image-based 3D.




(a) Assignment of X, Z coordinate.




(b) Assignment of Y coordinate.

**Fig. 3.** Assignment of 3D coordinate.

### 3.3 Calculating Camera Matrix

The camera matrix of the assigned and is solved as follows. The Eq(1) is expressed as the Eq(3) by its elements. When dividing the Eq(3), two simultaneous equations are obtained like the Eq(4). Because has no concern with scale and we could set to 1, it is expressed from the n matching point pairs like the Eq(5). Because the number of the unknown is 11 and two equations come from one point, we can calculate from at least 6 matching point pairs.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3}$$

$$u = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}$$
$$v = \frac{p_{21}X + p_{22}Y + p_{23}Z + p_{24}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}} \tag{4}$$

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1X_1 & -u_1Y_1 & -u_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1X_1 & -v_1Y_1 & -v_1Z_1 \\ X_2 & Y_2 & Z_2 & 1 & 0 & 0 & 0 & 0 & -u_2X_2 & -u_2Y_2 & -u_2Z_2 \\ 0 & 0 & 0 & 0 & X_2 & Y_2 & Z_2 & 1 & -v_2X_2 & -v_2Y_2 & -v_2Z_2 \\ \vdots & & & & \vdots & & & & \vdots & & \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -u_nX_n & -u_nY_n & -u_nZ_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -v_nX_n & -v_nY_n & -v_nZ_n \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ \vdots \\ p_{32} \\ p_{33} \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ \vdots \\ u_n \\ v_n \end{bmatrix} \tag{5}$$

## 4 Reconstruction of 3D Coordinate

The reconstruction of 3D coordinate from the matching points in the stereo images becomes accomplished by the triangulation [7]. The triangulation is computed by the estimated right-left camera matrix $\mathbf{P_L}$ and $\mathbf{P_R}$ in the chapter 3.3 and matching points $\mathbf{x_L}$ and $\mathbf{x_R}$ like the Fig. 4. The Eq(1) of the left-right camera is expressed as the Eq(6) and to remove the scale factor, the cross product is used and it can be express with the Eq(7). When the Eq(7) is divided, it could be induced three independent simultaneous equations from the each camera like the Eq(8). The two simultaneous equations among three simultaneous equations are independent each other and it can be expressed with single equation form to calculate $\mathbf{X}$ like the Eq(9). The $\mathbf{P^{iT}}$ is i-th row of $\mathbf{P}$ and $\mathbf{X}$ is found in the null-space of $\mathbf{A}$ [8].

$$\mathbf{x}_L \cong \mathbf{P}_L\mathbf{X}, \quad \mathbf{x}_R \cong \mathbf{P}_R\mathbf{X} \tag{6}$$

$$\mathbf{x}_L \times \{\mathbf{P}_L\mathbf{X}\} = 0, \quad \mathbf{x}_R \times \{\mathbf{P}_R\mathbf{X}\} = 0 \tag{7}$$

$$\begin{cases} u_L(\mathbf{P}_L^{3T}\mathbf{X}) - (\mathbf{P}_L^{1T}\mathbf{X}) = \mathbf{0} \\ v_L(\mathbf{P}_L^{3T}\mathbf{X}) - (\mathbf{P}_L^{2T}\mathbf{X}) = \mathbf{0} \\ u_L(\mathbf{P}_L^{3T}\mathbf{X}) - v_L(\mathbf{P}_L^{1T}\mathbf{X}) = \mathbf{0} \end{cases} \begin{cases} u_R(\mathbf{P}_R^{3T}\mathbf{X}) - (\mathbf{P}_R^{1T}\mathbf{X}) = \mathbf{0} \\ v_R(\mathbf{P}_R^{3T}\mathbf{X}) - (\mathbf{P}_R^{2T}\mathbf{X}) = \mathbf{0} \\ u_R(\mathbf{P}_R^{3T}\mathbf{X}) - v_R(\mathbf{P}_R^{1T}\mathbf{X}) = \mathbf{0} \end{cases} \tag{8}$$

$$\begin{bmatrix} u_L\mathbf{P}_L^{3T} - \mathbf{P}_L^{1T} \\ v_L\mathbf{P}_L^{3T} - \mathbf{P}_L^{2T} \\ u_R\mathbf{P}_R^{3T} - \mathbf{P}_R^{1T} \\ v_R\mathbf{P}_R^{3T} - \mathbf{P}_R^{2T} \end{bmatrix} \mathbf{X} = \mathbf{A}\mathbf{X} = \mathbf{0} \tag{9}$$

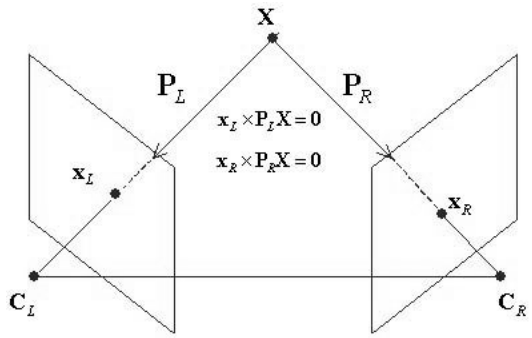



**Fig. 4.** Reconstruction of 3D coordinate using triangulation.

## 5 Registration of 3D coordinate

The reconstructed 3D coordinates from each stereo image must be converted to the reference coordinate. The registration is divided with the local registration and the global registration. The local registration is the mapping relation of each coordinate and it is expressed as $\mathbf{H_{ji}}$ in the Fig 5. The $\mathbf{H_{ji}}$ is a transformation matrix from $\mathbf{X_j}$ to $\mathbf{X_i}$. The global registration is the mapping relation between the reference coordinate and the current coordinate and it is expressed as $\mathbf{H_j}$ in the Fig 5. The $\mathbf{H_j}$ is a transformation matrix from $\mathbf{X_j}$ to $\mathbf{X_1}$.

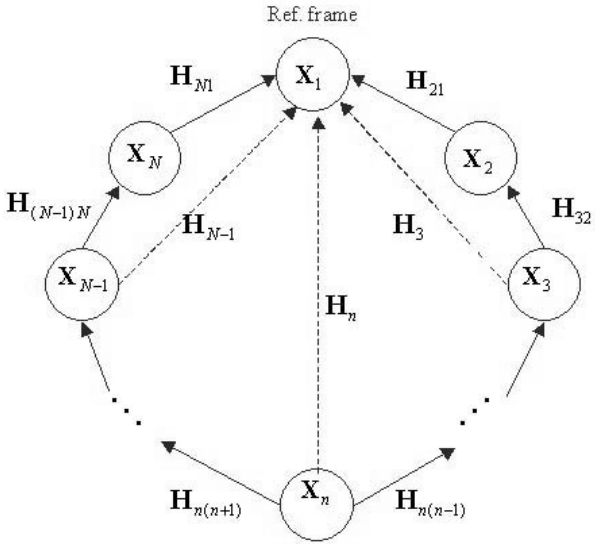



**Fig. 5.** Registration of 3D coordinate.

### 5.1 Local Registration

The transformation relation of the reconstructed 3D coordinates which is neighboured each other is represented by the Eq(10). The $\mathbf{H}_{ji}$ is a 4×4 matrix. Because the reconstructed 3D is the Euclidean coordinate, the transformation from $\mathbf{X}_j$ to $\mathbf{X}_i$ could be accomplished by the affined transformation, so $\mathbf{H}_{ji}$ is expressed as the Eq(11). The solution of $\mathbf{H}_{ji}$ is similar to the chapter 3.3 and it is solved using the Eq(12). The value of m should be 4 at least.

$$\mathbf{X}_i \cong \mathbf{H}_{ji}\mathbf{X}_j \tag{10}$$

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_j \\ Y_j \\ Z_j \\ 1 \end{bmatrix} \tag{11}$$

$$\begin{bmatrix} X_j^1 & Y_j^1 & Z_j^1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & X_j^1 & Y_j^1 & Z_j^1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X_j^1 & Y_j^1 & Z_j^1 & 1 \\ X_j^2 & Y_j^2 & Z_j^2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & X_j^2 & Y_j^2 & Z_j^2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X_j^2 & Y_j^2 & Z_j^2 & 1 \\ & & \vdots & & & \vdots & & & \vdots & & & \vdots \\ X_j^m & Y_j^m & Z_j^m & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & X_j^m & Y_j^m & Z_j^m & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X_j^m & Y_j^m & Z_j^m & 1 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{14} \\ h_{21} \\ h_{22} \\ \vdots \\ h_{32} \\ h_{33} \\ h_{34} \end{bmatrix} = \begin{bmatrix} X_i^1 \\ Y_i^1 \\ Z_i^1 \\ X_i^2 \\ Y_i^2 \\ Z_i^2 \\ \vdots \\ X_i^m \\ Y_i^m \\ Z_i^m \end{bmatrix} \tag{12}$$

### 5.2 Global Registration

It is the global registration to find $\mathbf{H}_j$, which converted into the reference coordinate directly using $\mathbf{H}_{ji}$ calculated from the local registration. The simplest method to compute $\mathbf{H}_j$ is $\mathbf{H}_{ji}^T\mathbf{H}_i^T - \mathbf{H}_j^T = \mathbf{0}$, which is continuously multiplied by $\mathbf{H}_{ji}$ directly. It has the weak point of the accumulated error that occurs from the local registration. So the generated error should be dispersed without accumulating it. Davis [3] adopted UDM(Uni-Directional Method) which computes the Eq(14) using the Eq(13) in 2D image mosaicing. The form of the Eq(14) is $\mathbf{B} = \mathbf{AX}$ and it is calculated by $\mathbf{X} = (\mathbf{A}^{-1}\mathbf{A})\mathbf{A}^{\mathrm{T}}\mathbf{B}$. However it generate the errors seriously at the registration, because the UDM is computed by the uni-directional mapping and has no connection the first with the last coordinate. To solve the problem, we suggest the BDM(Bi-Directional Method) and the WBDM(Weighted Bi-Directional Method). The BDM is expressed in Eq(15) and Eq(16), so the accumulated error is diminished. The BDM that is given the weight by the Eq(19) according to generated error from the local registration is transformed into the WBDM, so it diminishes both the accumulated and the probability error of miscalculating from generated error in the local registration.

Method 1 (UDM: Uni-Directional Method):

$$\mathbf{H}_{ji}^T\mathbf{H}_i^T - \mathbf{H}_j^T = \mathbf{0} \tag{13}$$

$$\begin{bmatrix} \mathbf{I} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{21}^T & -\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & & & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_{j(j-1)}^T & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \vdots & & & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{H}_{N(N-1)}^T & -\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{H}_1^T \\ \mathbf{H}_2^T \\ \vdots \\ \mathbf{H}_j^T \\ \vdots \\ \mathbf{H}_N^T \end{bmatrix} \tag{14}$$

Method 2 (BDM: Bi-Directional Method):

$$\mathbf{H}_{ji}^T\mathbf{H}_i^T + \mathbf{H}_{jk}^T\mathbf{H}_k^T - 2\mathbf{H}_j^T = \mathbf{0} \tag{15}$$

$$\begin{bmatrix} \mathbf{I} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{21}^T & -2\mathbf{I} & \mathbf{H}_{23}^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & & & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_{j(j-1)}^T & -2\mathbf{I} & \mathbf{H}_{j(j+1)}^T & \mathbf{0} \\ \vdots & & & \vdots & & \vdots \\ \mathbf{H}_{N1}^T & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{H}_{N(N-1)}^T & -2\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{H}_1^T \\ \mathbf{H}_2^T \\ \vdots \\ \mathbf{H}_j^T \\ \vdots \\ \mathbf{H}_N^T \end{bmatrix} \tag{16}$$

Method 3 (WBDM: Weighted Bi-Directional Method):

$$w_i\mathbf{H}_{ji}^T\mathbf{H}_i^T + w_k\mathbf{H}_{jk}^T\mathbf{H}_k^T - (w_i + w_k)\mathbf{H}_j^T = \mathbf{0} \tag{17}$$

$$\begin{bmatrix} \mathbf{I} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ w_1\mathbf{H}_{21}^T & -(w_1+w_3)\mathbf{I} & w_3\mathbf{H}_{23}^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & & & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & w_{j-1}\mathbf{H}_{j(j-1)}^T & -(w_{j-1}+w_{j+1})\mathbf{I} & w_{j+1}\mathbf{H}_{j(j+1)}^T & \mathbf{0} \\ \vdots & & & \vdots & & \vdots \\ w_1\mathbf{H}_{N1}^T & \mathbf{0} & \mathbf{0} & \mathbf{0} & w_{N-1}\mathbf{H}_{N(N-1)}^T & -(w_1+w_{N-1})\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{H}_1^T \\ \mathbf{H}_2^T \\ \vdots \\ \mathbf{H}_j^T \\ \vdots \\ \mathbf{H}_N^T \end{bmatrix} \tag{18}$$

$$w_i = e^{-\frac{E_i - E_{min}}{E_{max} - E_{min}} * E_{std}^2}$$

$$E_i = \frac{1}{M_{m-1}} \left\| \mathbf{X}_i^m - \mathbf{H}_{ji} \mathbf{X}_j^m \right\| \quad E_{std} = \sqrt{\frac{1}{N_{i-1}} (E_i - E_{avg})^2}$$

$$\left( = E_{max} = \min\{E_i\} \quad = E_{max} = \max\{E_i\} \quad E_{avg} = \frac{1}{N} \sum_{i=1}^{N} E_i \right) \qquad (19)$$

## 6 Experiment and Result

The camera to capture images in the image-based 3D modeling system is the XC-55 grey level video module. In order to obtain stereo images, we use two camera of same type. The acquired image has the grey intensity and the resolution of 640 by 480. We display the 3D modeling using the OpenGL [9].

The left-right calibration pattern images for the camera calibration is shown in the Fig. 6. After calibrating the camera, we rotate the object for the 3D modeling and obtain the stereo images at the same time. The object for the experiment is the miniature computer. The four among the total eight stereo images are showed in the Fig. 7.

The Fig. 8 shows the selected feature points and the planes for 3D modeling primitives and texture mappings from k=1 stereo images. The feature points are manually selected with the mouse. The Fig 9 is the result of reconstructed 3D modeling from the input images in the Fig 7. The Fig 10 shows the result that is combined two 3D modeling into a 3D modeling by the local registration. The Fig 11 shows the 3D modeling mesh by the global registration.

When there is no the miscalculating from the errors in the local registration, the errors by the result of the BDM and the WBDM are distributed in the whole region. The graph in Fig. 12 shows the standard deviation of 3D reconstruction coordinates by the quantitative analysis. The standard deviation is the deviation of the single coordinate generated by averaging the 3D coordinates using the Eq(20). In the ideal case, the value of the standard deviation is zero.

$$Error_{std} = \frac{1}{m} \sum_{k=1}^{m} \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\mathbf{H}_i \mathbf{X}_i^k - \overline{\mathbf{X}}^k)^2} \qquad (20)$$

$$\text{where, } \overline{\mathbf{X}}^k = \frac{1}{n} \sum_{i=1}^{n} \mathbf{H}_i \mathbf{X}_i^k$$

The standard deviation of the WBDM is the lowest, if there is the miscalculating according to the generated errors in local registration. The graph in Fig.13 shows the standard deviation of 3D reconstruction coordinates by the quantitative analysis.

The Fig. 14 shows the 3D modeling of the single coordinate by averaging the 3D coordinate generated by the WBDM. The Fig 15 shows the final 3D modeling from the various angles. The proposed 3D modeling system produces the realistic 3D modeling using only input stereo images.

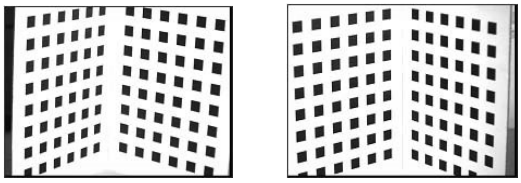


**Fig. 6.** Left-Right images for camera calibration.

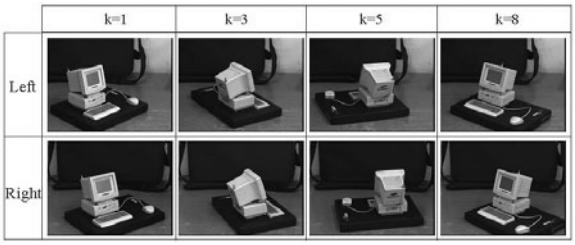



**Fig. 7.** Input stereo images.

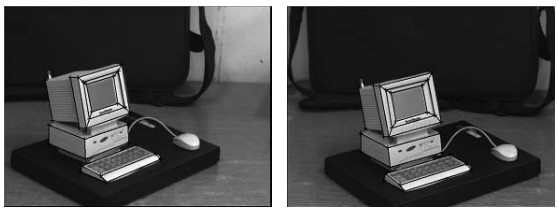


**Fig. 8.** Detected feature points and primitives.

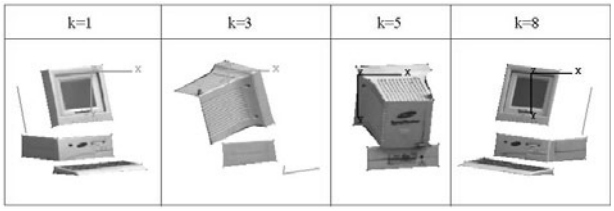



**Fig. 9.** 3D model images for each stereo image.

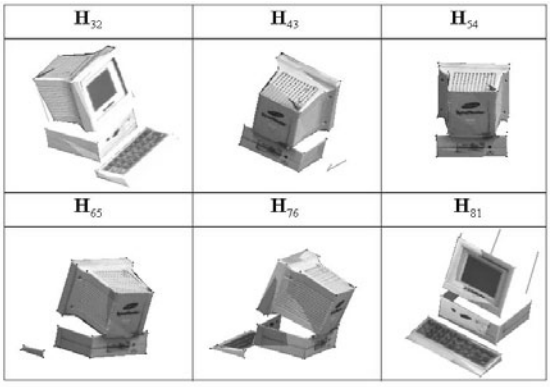



**Fig. 10.** 3D model images after local registration.

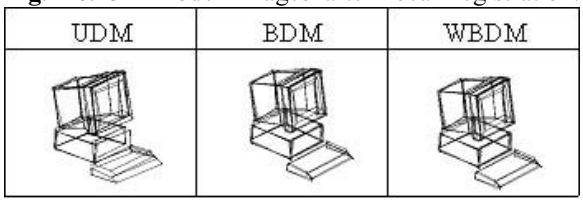



**Fig. 11.** 3D model mesh for each method of global registration

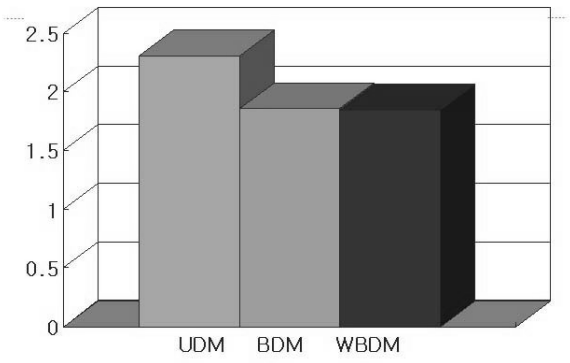



**Fig. 12.** Comparison graph of standard deviation.

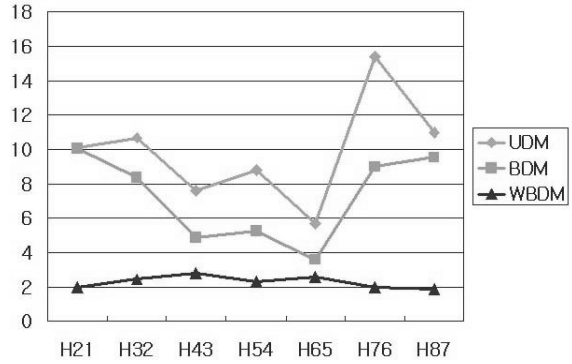



**Fig 13.** Comparison graph of standard deviation.

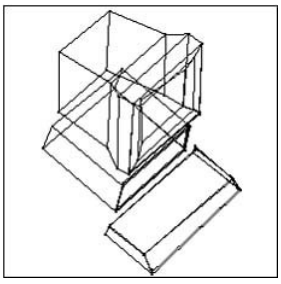


**Fig. 14.** Mesh image after assignment of single 3D coordinate.

**Fig. 15.** Final 3D model images in the multi-view point.

## 7 References

[1] B. Girod, G. Greiner, H. Niemann, Principles of 3D Image Analysis and Synthesis, Kluwer Academic Publishers, 2000.

[2] Richard Hartley and Andrew Zisserman, Multiple View Geometry in Computer Vision, cambridge univ. press, 2000.

[3] Davis, J., "Mosaics of Scenes with Moving Objects," IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR98), 1998.

[4] Z. Zhang, “A Flexible New Technique for Camera Calibration,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.

[5] O. Faugeras. Three-Dimensional Computer Vision: a Geometric Viewpoint. MIT Press, 1993.

[6] C. J. Harris and M. Stephens, “A combined corner and edge detector,”In Proc. 4th Alvey Vision Conf., Manchester, pages 147-151, 1988.

[7] R. I. Hartley and P. Sturm, “Triangulation,” Computer Vision and Image Understanding, vol. 68, no. 2, pp. 146-157, 1997.

[8] W. H. Press, Numercial Recipe in C : the art of scientific computing(2nd ed.), UK: Cambridge Press, 1992.

[9] Richard S. Wright, Jr. and Michael Sweet, OpenGL Superbible, 1997.

# Accuracy Improvement in Camera Calibration

FaJie Li, Qi Zang and Reinhard Klette
CITR, Computer Science Department
The University of Auckland
Tamaki Campus, Auckland, New Zealand
fli006, qzan001@ec.auckland.ac.nz
r.klette@auckland.ac.nz

## Abstract

Camera calibration is a necessary and critical step in 3D object analysis. The accuracy of calibration results will affect the object's position in world coordinates, especially for 3D object tracking. In this paper, we present a new camera calibration approach, and discuss its accuracy. We use 3D marks instead of 2D marks for calibration. Our experimental results show that our approach has the potential to improve the calibration accuracy.

**Keywords**: 3D reconstruction, camera calibration.

## 1 Introduction

In the context of three-dimensional machine vision, camera calibration is a process for determining the internal geometric and optical camera characteristics (intrinsic parameters), and the 3D position and orientation data of the camera frame relative to a defined world coordinate system (extrinsic parameters). A calibration technique is based on known 3D coordinates of geometrically configured points. Here the 3D coordinates are usually referenced to the world coordinate system. The configured points are commonly referred to as *calibration points* which are physically realized by *calibration marks* on a *calibration object*.

Common calibration methods are DLT (Direct Linear Transform) and Tsai's method [1, 2]. The latter one also models lens distortion coefficients and the mapping of sensor elements to an image buffer matrix; it requires at least seven accurately detected calibration points in an arbitrary but known geometric configuration. Using Tsai's calibration method, we can transfer 3D world coordinates into image coordinates. But even using the same calibration method, differences in image acquisition environments may affect the accuracy, such as distances between camera and object, the size and number of calibration marks, or the size of calibration objects. We evaluated Tsai's method with respect to such variations, where 3D marks are used instead of the common 2D calibration marks. Section 2 reports about improvements in calibration accuracy; Section 3 specifies this new alteration of Tsai's method; Section 4 presents further experimental results, and Section 5 contains our conclusions.

## 2 Accuracy Improvement

Calibration accuracy will affect calculations of object's position and tracking parameters. We summarize four methods for evaluating camera calibration accuracy:

1. *Image Coordinates Error Statistics*: This measurement method analyzes the distorted error statistics on the image plane. Steps are: convert world coordinates into camera coordinates, then into undistorted sensor plane coordinates, then into distorted sensor plane coordinates, then into image coordinates. After all these conversions, determine errors between ideal image coordinates and actual locations of data points.

2. *Undistorted Image Plane Error Statistics*: This measurement method analyzes the undistorted error statistics on the image plane. Steps are: convert world coordinates into camera coordinates, then into undistorted sensor plane coordinates; convert from 2D image coordinates into distorted sensor plane coordinates, then into undistorted sensor plane coordinates. After converting, determine the error between ideal and actual location of the data point.

3. *Camera Coordinates Error Statistics*: This measurement method analyzes the error statistics in object space. Steps are: convert world coordinates into camera coordinates; convert from 2D image coordinates into distorted sensor plane coordinates, then into undistorted sensor plane coordinates, then into 3D camera coordinates. After converting, the error is defined by distances of closest points to ideal projection rays in 3D camera coordinates space.