MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28–30, 2000, The University of Tokyo, Japan

13—8 Registration of Range Images Considering Depth Measurement Error

Kang-Sik Yoon[†], Byung-Uk Lee[††], Rae-Hong Park[†], and Woo-Shik Kim[†]
[†] Dept. of Electronic Eng., Sogang University, C.P.O. Box 1142, Seoul 100-611, Korea
[††]Dept. of Information Electronics, Ewha Womans University, Seoul, Korea
E-mail: kws@eevision1.sogang.ac.kr

Abstract

To enhance the accuracy of rotation and translation parameters of 3-D positions calculated from two depth images, we present an algorithm to calculate the registration error. First, we derive the camera error from triangulation, which is represented in a covariance matrix. Then we find orientation and position parameters of an object. We verify that inclusion of the camera sensitivity enhances the accuracy of registration from computer simulations and experiments.

1 Introduction

Since range finders with high accuracy are becoming commercially available, applications of depth data to modeling a three-dimensional (3-D) object or to medical research area have been greatly increased. A range image provides a depth or surface information from a front view. Therefore, it is necessary to merge several range images from a side or rear view of the same object to build a complete 3-D model. To fulfill the requirement, we need to estimate the translation and rotation parameters of the range data between the images. Many works have been done to find registration parameters: iterative closest point (ICP) algorithm on two range data viewed from the different directions on an object [1-3], a method using statistical criterion function for the accuracy of initial values of the ICP algorithm [4], a method applying the ICP algorithm to free form objects [5], a method using the line correspondence [6], and a method to segment outliers from range data with least mean squares (LMS) and using the ICP algorithm to estimate registration parameters [7].

To achieve accurate estimation, the precise measurement of depth is essential. But the depth information in real absolute world coordinate cannot be known and depends on several factors. Image noise and quantization errors are typical factors. Especially these factors are deeply related to camera setting parameters. Therefore to calculate more precise relative registration parameters, the camera sensitivity, which represents how sensitive the coordinate transformation between world coordinate and image coordinate is, is to be compensated in the process of registration. The camera sensitivity can be interpreted with 3-D data measurement error. The analysis of 3-D data measurement error has been an important research area: a method to analyze a triangularization method assuming that location of 3-D points has uniform probability distribution within unit volume [8], a method to use data redundancy after acquiring more than two range data of an object from the various locations of a camera [9], a method considering the system parameter of a camera related to error such as the separation between sensor elements, the camera lens focal length, and the sensor array dimension [10], and a method considering noise for the estimation of registration parameters of a moving object [11]. Also there have been various works on error analysis in measuring the location of 3-D points [12-14].

We present the error (camera sensitivity) in 3-D position measurement to improve the accuracy of the registration. The estimation using a covariance matrix based on the camera sensitivity as error weights is proposed. This paper is structured as follows. In Section 2, we present a method acquiring range data. In Section 3, we present an expression for the camera sensitivity in the range measurement. In Section 4, we estimate registration parameters with the proposed error weights from the camera sensitivity. In Section 5, we present experimental results comparing the estimated results obtained with and without error weights. Finally, we conclude in Section 6.

2 Measurement of Range Data Using Triangulation

Fig. 1 illustrates the principle of a depth measurement method based on triangulation. It has a structured light which divides a space into 2^n regions from n sets of images [15]. From the image, we can calculate the distance from the camera to the surface of the object.

After camera calibration, we obtain a transformation matrix ($\mathbf{T}_c$) between the world (C_w) and camera coordinates (C_1), and a transformation ($\mathbf{T}_p$) between the world (C_w) and the projector coordinates (C_2). The 3-D coordinate values of the object are calculated according to the following procedures:

$$\begin{pmatrix} x_w & y_w & z_w & 1 \end{pmatrix} \begin{pmatrix} t_{c11} & t_{c12} & t_{c13} \\ t_{c21} & t_{c22} & t_{c23} \\ t_{c31} & t_{c32} & t_{c33} \\ t_{c41} & t_{c42} & t_{c43} \end{pmatrix} = \begin{pmatrix} h_c u & h_c v & h_c \end{pmatrix} \quad (1)$$

$$\begin{pmatrix} x_w & y_w & z_w & 1 \end{pmatrix} \begin{pmatrix} t_{p11} & t_{p12} \\ t_{p21} & t_{p22} \\ t_{p31} & t_{p32} \\ t_{p41} & t_{p42} \end{pmatrix} = \begin{pmatrix} h_p w & h_p \end{pmatrix} \quad (2)$$

where the indices c of h_c and p of h_p denote camera and projector, respectively, and u, v, and w represent image planes of the camera and projector coordinate, respectively. The parameter h represents the homogeneous coordinates. After removing h_c and h_p from (1) and (2), **Q**, **V**, and **F** are derived as follows:

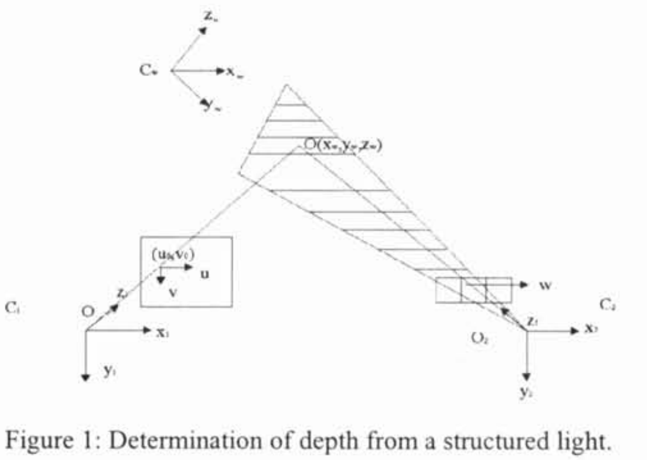


Figure 1: Determination of depth from a structured light.

$$\mathbf{Q} = \begin{pmatrix} t_{c11} - t_{c13}u & t_{c21} - t_{c23}u & t_{c31} - t_{c33}u \\ t_{c12} - t_{c13}v & t_{c22} - t_{c23}v & t_{c32} - t_{c33}v \\ t_{p11} - t_{p12}w & t_{p21} - t_{p22}w & t_{p31} - t_{p32}w \end{pmatrix} \quad (3)$$

$$\mathbf{V} = \begin{pmatrix} x_w & y_w & z_w \end{pmatrix}^t \quad (4)$$

$$\mathbf{F} = \begin{pmatrix} t_{c43}u - t_{c41} \\ t_{c43}v - t_{c42} \\ t_{p42}w - t_{p41} \end{pmatrix} \quad (5)$$

$$\mathbf{QV} = \mathbf{F} \quad (6)$$

$$\mathbf{V} = \mathbf{Q}^{-1}\mathbf{F}. \quad (7)$$

Therefore, 3-D coordinate values **V** are obtained by (7). In other words, with two camera parameter matrices, $\mathbf{T}_c$ and $\mathbf{T}_p$, together with image coordinates (u, v) and the space code w, we can calculate a 3-D position (x_w, y_w, z_w).

3 Estimation of the Camera Sensitivity

The camera position u, v and the space code w have error due to the image noise and quantization error. In this section, we analyze the camera sensitivity with the variance of u, v, and w on the calculated 3-D position (x,y,z).

Error of a 3-D coordinate position can be expressed as

$$\begin{aligned} \Delta x &= \frac{\partial x}{\partial u}\delta u + \frac{\partial x}{\partial v}\delta v + \frac{\partial x}{\partial w}\delta w \\ \Delta y &= \frac{\partial y}{\partial u}\delta u + \frac{\partial y}{\partial v}\delta v + \frac{\partial y}{\partial w}\delta w \\ \Delta z &= \frac{\partial z}{\partial u}\delta u + \frac{\partial z}{\partial v}\delta v + \frac{\partial z}{\partial w}\delta w \end{aligned} \quad (8)$$

and (8) is expressed in matrix form:

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{bmatrix} \begin{bmatrix} \delta u \\ \delta v \\ \delta w \end{bmatrix} \quad (9)$$

$$\Delta\mathbf{P} = \mathbf{S}\Delta\mathbf{m}. \quad (10)$$

E (Expectation) of the square of 3-D coordinates becomes a covariance matrix as defined by

$$\begin{aligned} E\left(\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}\begin{bmatrix} \Delta x & \Delta y & \Delta z \end{bmatrix}\right) &= E((\Delta\mathbf{P})(\Delta\mathbf{P})') \\ &= E((\mathbf{S}\Delta\mathbf{m})(\mathbf{S}\Delta\mathbf{m})' \\ &= E(\mathbf{S}\Delta\mathbf{m}\Delta\mathbf{m}'\mathbf{S}'). \end{aligned} \quad (11)$$

If we assume identical independent distribution (i.i.d.) on $\Delta\mathbf{m}$, then we can obtain the following results:

$$E(\Delta\mathbf{m}\Delta\mathbf{m}') = \begin{pmatrix} \sigma_u^2 & 0.0 & 0.0 \\ 0.0 & \sigma_v^2 & 0.0 \\ 0.0 & 0.0 & \sigma_w^2 \end{pmatrix}. \quad (12)$$

Since **S** and $\mathbf{S}'$ are deterministic, the covariance matrix becomes

$$E\left(\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}\begin{bmatrix} \Delta x & \Delta y & \Delta z \end{bmatrix}\right) = \sigma^2 \boldsymbol{SS}' \quad (13)$$

after assuming that $\sigma_u^2 = \sigma_v^2 = \sigma_w^2 = \sigma^2$. We will employ the covariance matrix as error weights for estimation of registration parameters.

4 Registration with the Estimated Position Error Using the Camera Sensitivity

We present the rotation of an object between the two range images with a unit quaternion [16] that has the advantages of the reduced computational complexity, and estimation of registration parameters is done with a nonlinear numerical optimization, Levenberg-Marquardt method [17].

4.1 Representation of Rotation

In general, rotation transformation is expressed as a matrix. It has a simple form and easy to understand. However, it has six constraints to satisfy the orthonormality of the unitary matrix. On the other hand, orientation angle representation has only three parameters. However its trigonometric nature makes it cumbersome for numerical analysis. Unit quaternion representation of orientation consists of four real numbers, and a constraint which requires that the norm of four real components should be unity. Also the amount of calculation is greatly reduced. The quaternion $\boldsymbol{q}$ is expressed with the complex notation:

$$\boldsymbol{q} = q_0 + iq_x + jq_y + kq_z. \quad (14)$$

A point in the 3-D space is expressed as the following quaternion $\boldsymbol{t}$ having only imaginary parts,

$$\boldsymbol{t} = 0 + it_x + jt_y + kt_z. \quad (15)$$

A point, $\boldsymbol{t}$, rotated by a unit quaternion $\boldsymbol{r}$ is expressed as

$$\boldsymbol{t}' = \boldsymbol{rtr}^* \quad (16)$$

and rotation $\boldsymbol{r}$ is as follows:

$$\boldsymbol{r} = \cos(\frac{\theta}{2}) + \sin(\frac{\theta}{2})\boldsymbol{\omega} = r_0 + ir_x + jr_y + kr_z \quad (17)$$

where rotation angle and rotation axis are θ and $\boldsymbol{\omega}$, respectively.

Unit quaternion constraint requires $r_0^2+r_x^2+r_y^2+r_z^2=1$. Therefore the rotation angle and the axis are calculated from r_0, r_x, r_y, and r_z.

4.2 Levenberg-Marquardt Method

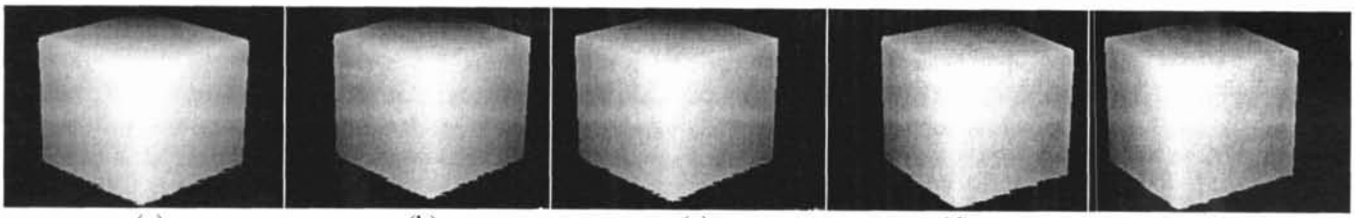
When a model is nonlinearly related to undetermined parameters a_k, $1 \le k \le M$, the χ^2 error function is defined and parameters are searched to minimize the error function. Due to nonlinear dependency, minimization procedure is iterated until the parameter value converges to a local minimum. When a model containing parameters is

$$y = y(x;\mathbf{a}), \quad (18)$$

the χ^2 merit function is written as

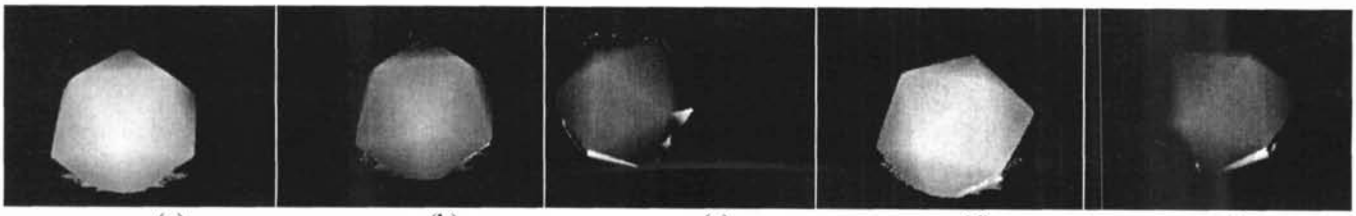
$$\chi^2(\mathbf{a}) = \sum_{i=1}^{N}\left[\frac{y_i - y(x_i;\mathbf{a})}{\sigma_i}\right]^2 \quad (19)$$

where N is the number of data used for the estimation of



(a) (b) (c) (d) (e)

Figure 2: Images used in experiments (Cube).



(a) (b) (c) (d) (e)

Figure 3: Images used in experiments (Polyhedron).

model parameters. For our experiments, the equation can be modified as

$$\chi^2(\mathbf{T}) = \sum_{i=1}^{N} [y_i - y(x_i;\mathbf{T})] \Sigma^{-1} [y_i - y(x_i;\mathbf{T})]. \quad (20)$$

The error weight $\mathbf{\Sigma}$ is the error covariance matrix of $[y_i - y(x_i;\mathbf{T})]$. Here we assume that we have N corresponding points between reference image 1 and image 2. A point y_i in image 1 corresponds to a point x_i in image 2 transformed by $\mathbf{T}$.

The 3-D position measurement error at the reference position 1, that in image 2, and those error values after transforming the point in image 2 to the reference image 1 are represented as $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_2'$, respectively. Then we can derive the error weighting matrix $\mathbf{\Sigma}$ as

$$E[\mathbf{e}_1 \mathbf{e}_1^t] = \mathbf{\Sigma}_1 \quad (21)$$

$$E[\mathbf{e}_2 \mathbf{e}_2^t] = \mathbf{\Sigma}_2 \quad (22)$$

$$\mathbf{e}_2' = \mathbf{T}\,\mathbf{e}_1 \quad (23)$$

$$E[\mathbf{e}_2' \mathbf{e}_2'^t] = E[\mathbf{T}\,\mathbf{e}_1 \mathbf{e}_1^t \mathbf{T}] = \mathbf{T}\,\mathbf{\Sigma}_1 \mathbf{T}^t. \quad (24)$$

Therefore, the total error is defined by

$$(\mathbf{e}_2 - \mathbf{e}_2')(\mathbf{e}_2 - \mathbf{e}_2')^t = (\mathbf{e}_2 - \mathbf{e}_2')(\mathbf{e}_2^t - \mathbf{e}_2'^t)$$
$$= \mathbf{e}_2 \mathbf{e}_2^t - \mathbf{e}_2 \mathbf{e}_2'^t - \mathbf{e}_2' \mathbf{e}_2^t + \mathbf{e}_2' \mathbf{e}_2'^t. \quad (25)$$

The expectation of (25) becomes

$$E[(\mathbf{e}_2 - \mathbf{e}_2')(\mathbf{e}_2 - \mathbf{e}_2')^t] = \mathbf{\Sigma}_2 + \mathbf{T}\,\mathbf{\Sigma}_1 \mathbf{T}^t \quad (26)$$

and the error weight matrix, $\mathbf{\Sigma}$, is expressed as $\mathbf{\Sigma}_2 + \mathbf{T}\,\mathbf{\Sigma}_1 \mathbf{T}^t$. For simplicity, we can assume that the rotation, $\mathbf{R}$, or $\mathbf{T}$ is small, then $\mathbf{R}$ can be approximated as the identity matrix $\mathbf{I}$. Then it follows that $\mathbf{\Sigma}_1 \cong \mathbf{\Sigma}_2$. Hence we can use the approximation $\mathbf{\Sigma} = 2\mathbf{\Sigma}_2$.

5 Experimental Results and Discussions

Fig. 2 shows range images of a cube used in experiments, where pixels nearer to the camera are brighter. Space encoded light patterns are projected through an LCD, and the images are taken from a CCD camera. The range data is obtained using (7). The image (a) is the object at the reference position. The images (b) and (c) represent the object of (a) translated by 2 *cm* in the x coordinate and the object of (a) translated by 2 *cm* in the y direction. The image (d) represents the object (a) rotated by 10° with respect to the z coordinate. The image (e) is obtained after translating the object (d) by -3 *cm* in the x direction and by 3 *cm* in the y direction. The object was translated and rotated on a ruled paper. We assume that the point correspondence of the vertices has been established. Fig. 3 is a set of range images of a polyhedron.

The error weight from the camera noise sensitivity has been applied in estimating the registration parameters. The camera sensitivity is calculated from the variations of x, y, and z with respect to u, v, and w at point in Fig. 2(c), and then the covariance matrix, $\mathbf{\Sigma} = \sigma^2 \mathbf{S}\mathbf{S}^t$, is used as error weights which has the following value:

$$\mathbf{\Sigma}_1 = \sigma^2 SS^t = \begin{pmatrix} 1432.2678 & 1651.9968 & -680.4394 \\ 1651.9968 & 1953.6555 & -796.5092 \\ -680.4394 & -796.5092 & 346.5571 \end{pmatrix}. \quad (27)$$

On the other hand, the covariance matrix, $\mathbf{\Sigma}$, without considering error weights has the following value:

$$\mathbf{\Sigma} = \begin{pmatrix} 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{pmatrix}. \quad (28)$$

The scale factor on the error weight $\mathbf{\Sigma}$ has no effects on the rotation parameters.

Tables 1 and 2 show the comparison of estimated registration parameters with and without error weights using the point matching for Cube and Polyhedron, respectively. It shows that rotation parameters are closer to the exact values than translation parameters, regardless of the positions of an object in Fig. 2.

Especially the rotation parameters considering error weights are much closer to the exact values than rotation parameters ignoring measurement error weights. The convergence of estimation with error weights is faster than that of estimation without error weights. Root mean square error (RMSE) is defined as the square root of mean value of squares of the difference of distance between 3-D coordinate values transformed by exact registration parameters and 3-D coordinate values transformed by estimated registration parameters. It is proportional to the Euclidean distance from the exact position to the estimated one.

Table 3 compares the RMSE with and without error weights. We can verify that the RMSE with error weights

Table 1: Comparison of registration parameters estimated with and without error weights **Σ** (Cube).

			exact values		without **Σ**		with **Σ**	
			r	***t***	***r***	***t***	***r***	***t***
b	r_0 r_x r_y r_z	t_x t_y t_z	1.0 0.0 0.0 0.0	2.0 0.0 0.0	0.9998 -0.0811 0.0421 0.0006	3.82 -1.28 0.53	0.9999 0.0321 0.0001 0.0002	2.36 -0.07 0.72
c	r_0 r_x r_y r_z	t_x t_y t_z	1.0 0.0 0.0 0.0	0.0 2.0 0.0	0.9998 0.02 -0.065 0.009	1.25 3.51 2.11	0.9999 0.01 -0.022 0.009	0.12 2.42 0.05
d	r_0 r_x r_y r_z	t_x t_y t_z	0.996 -0.03 0.033 0.071	0.0 0.0 0.0	0.9971 0.0528 -0.035 -0.067	0.91 -2.01 0.06	0.9965 - 0.0312 0.041 -0.079	0.02 -0.79 0.05
e	r_0 r_x r_y r_z	t_x t_y t_z	0.996 -0.03 0.033 0.071	-3.0 3.0 0.0	0.9979 0.0008 -0.0071 0.0349	-4.79 3.41 -0.04	0.9959 0.012 - 0.0031 0.089	-3.15 3.89 1 -0.05

Table 2: Comparison of registration parameters estimated with and without error weights **Σ** (Polyhedron).

			exact values		without **Σ**		with **Σ**	
			r	***t***	***r***	***t***	***r***	***t***
b	r_0 r_x r_y r_z	t_x t_y t_z	1.0 0.0 0.0 0.0	3.7 0.0 0.0	0.9986 0.092 -0.045 0.0009	5.9 1.3 -0.9	0.9999 0.089 -0.031 0.003	4.7 0.9 -0.5
c	r_0 r_x r_y r_z	t_x t_y t_z	1.0 0.0 0.0 0.0	3.7 6.0 0.0	0.9981 0.077 -0.068 0.001	6.1 8.7 -0.2	0.9990 0.061 0.059 - 0.0005	4.3 7.2 -0.09
d	r_0 r_x r_y r_z	t_x t_y t_z	0.986 0.08 0.12 0.11	0.0 0.0 0.0	0.9755 0.095 0.15 0.13	-0.5 3.05 0.9	0.9802 0.091 0.098 0.131	-0.6 2.1 0.3
e	r_0 r_x r_y r_z	t_x t_y t_z	0.986 0.08 0.12 0.11	3.7 3.7 0.0	0.9701 0.099 0.2 0.12	6.1 5.2 -1.9	0.9811 0.081 0.11 0.139	4.1 5.5 -0.3

is smaller than that ignoring the error weights. We observe that the estimation of registration parameters based on error weights from the camera sensitivity is more accurate than that without error weights.

6 Conclusions

The sensitivity of 3-D position error caused by 2-D camera position is calculated, and the obtained error covariance matrix is applied to estimation of registration parameters. It results in better estimation of registration parameters. Experiments show that our proposed method results in more exact registration parameters.

References

[1] D. W. Eggert, A. W. Fitzgibbon, and R. B. Fisher, "Simultaneous registration of multiple range views for use in reverse engineering," in *Proc. Int. Conf. Pattern Recognition '96*, vol. I, pp. 243-247, Vienna, Austria, 1996.

Table 3: Comparison of the RMSE.

	RMSE			
	cube		polyhedron	
	w/o **Σ**	with **Σ**	w/o **Σ**	with **Σ**
b	242.2123	81.3454	479.4291	325.7183
c	194.7499	72.5226	672.3852	526.2594
d	441.1082	345.5902	822.9819	721.6679
e	158.0073	153.3775	839.1146	766.5133

[2] T. Masuda, K. Sakaue, and N. Yokoya, "Registration and integration of multiple range images for 3-D model construction," in *Proc. Int. Conf. Pattern Recognition '96,* vol. I, pp. 879-883, Vienna, Austria, 1996.

[3] A. J. Stoddart and A. Hilton, "Registration of multiple point sets," in *Proc. Int. Conf. Pattern Recognition '96,* vol. I, pp. 40-44, Vienna, Austria, 1996.

[4] V. Charvillat and B. Thiesse, "Registration of stereo-based 3D maps for object modeling: A stochastic yet intelligent solution," in *Proc. Int. Conf. Pattern Recognition '96,* vol. I, pp. 780-785, Vienna, Austria, 1996.

[5] D. Brujic and M. Ristic, "Analysis of free form surface registration," in *Proc. Int. Conf. Image Processing '96,* vol. II, pp. 393-396, Lausanne, Switzerland, 1996.

[6] E. Salari and C. M. Jong, "Calculation of the structure and motion parameters from line correspondence," in *Proc. SPIE Conf. Intelligent Robot and Computer Vision VIII: Algorithms and Techniques*, vol. 1192, pp. 789-797, Philadelphia, PA, USA, 1989.

[7] T. Masuda, "A robust method for registration and segmentation of multiple range images," *Computer Vision and Image Understanding*, vol. 61, no. 3, pp. 295-307, 1995.

[8] S. D. Blostein and T. S. Huang, "Error analysis in stereo determination of 3-D point positions," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no. 6, pp. 752-765, 1997.

[9] J. Weng, T. S. Huang, and N. Ahuja, "Motion and structure from two perspective views: Algorithms, error analysis, and error," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-11, no. 5, pp. 451-476, 1989.

[10] E. S. Mcvey and J. W. Lee, "Some accuracy and resolution aspects of computer vision distance measurements," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-4, no. 6, pp. 646-649, 1982.

[11] J. Weng, N. Ahuja, and T. S. Huang, "Optimal motion and structure estimation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-15, no. 9, pp. 864-884, 1993.

[12] Y. Yasumoto and G. Medioni, "Robust estimation of three-dimensional motion parameters from a sequence of image frames using regularization," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, no. 4, pp. 464-471, 1986.

[13] T. Broida and R. Chellappa, "Estimation of object motion parameters from noisy images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, no. 1, pp. 90-99, 1986.

[14] J. Weng, T. S. Huang, and N. Ahuja, "3-D motion estimation, understanding, and prediction from noisy image sequences," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no. 3, pp. 370-389, 1987.

[15] S. Inoguchi, K. Sato, and F. Matsuda, "Range-imaging system for 3-D object recognition," in *Proc. 7th Int. Conf. Pattern Recognition*, pp. 806-808, Montreal, Canada, 1984.

[16] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Am. A*, vol. 4, no. 4, pp. 629-642, 1987.

[17] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge Univ. Press, pp. 683-688, New York, USA, 1992.