MVA '96 IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

A Method to Register Multiple Range Images from Unknown Viewing Directions

Ikuko SHIMIZU and Koichiro DEGUCHI *
Faculty of Engineering
University of Tokyo

Abstract

This paper presents a total system to register multiple range images to obtain complete object shape in the case where their relative viewing positions are unknown. To determine the relative position of each viewpoint with respect to a common world coordinate system, the correspondences of the object's planar faces are used. For non-polyhedral objects, we produce a three dimensional convex hull of partial shape of the object observed in each range image. In the every two consecutive images, corresponding faces are found by matching adjacency graphs of their large faces. Then, the relative rotations between all viewpoints are calculated simultaneously by least squares fitting with a constraint that the rotation matrix should be orthonormal with unit determinant. Relative translations are obtained by simple least squares fitting.

1 Introduction

In order to measure complete 3D shape of an object using range finder, we have to obtain range images from multiple viewpoints and integrate them. Especially for complex objects containing many concaves, we have to arrange many viewpoints in 3D space. However, it is difficult to set them up accurately. In this paper, a method is proposed to overcome this problem. Instead of the accurate viewpoints arrangements, we obtain range images from sufficient number of arbitrary viewpoints and then recover their accurate relative positions from data set themselves.

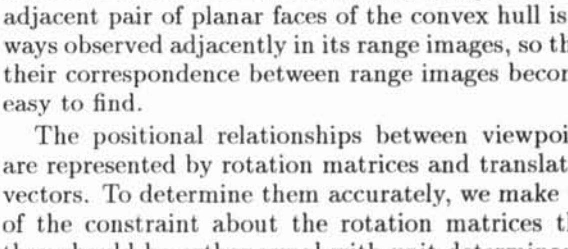
Many methods to register multiple range images obtained from unknown viewpoints have been proposed in the literatures[1][2][3]. Almost of them dealt with rather simple shape or assumed that the positions of viewpoints were approximately known. Our method does not assume both of them. We deal with not only simple polyhedra but also complex objects such as human hand shape.

To register the object's range images, we have to find correspondences of some kinds of object features observed in common in respective images and calculate their positional relationships. For these features, some methods[3] used planar faces of polyhedral object. The advantage of using the planar face is its ease to find correspondences between respective images. It is also useful that the normal vector of the plane is invariant to parallel translation.

To treat complex objects rather than polyhedra, we produce a three dimensional convex hull of partial shape of the object observed in each range image. It is always a part of the convex hull of the whole object. Hence, some faces of the convex hull are always observed in common from different viewpoints. An adjacent pair of planar faces of the convex hull is always observed adjacently in its range images, so that, their correspondence between range images becomes easy to find.

The positional relationships between viewpoints are represented by rotation matrices and translation vectors. To determine them accurately, we make use of the constraint about the rotation matrices that they should be orthonormal with unit determinant.

Section 2 shows the details of the process in the system described in this paper. For the explanation, we use an example model of human hand shown in Fig.1(a). Figure 1(b) shows two of its partial shapes observed in the range images obtained from different viewpoints. We first produce three dimensional convex hull of the partial shape of the object observed in each range image (Section 2.1). Then, we find corresponding face pairs between them (Section 2.2). Finally we determine relative positions of all the viewpoints and express each partial shape of the object in a common world coordinate system (Section 2.3). Experimental results are shown in Section



(a) (b)

Figure 1. (a)A wooden hand model used as an example object and (b)its partial shapes observed in the range images obtained form different viewpoints.

*Address: 7-3-1 Hongo, Bunkyo-ku, Tokyo 113 Japan. E-mail: {ikuko, deguchi}@meip7.t.u-tokyo.ac.jp

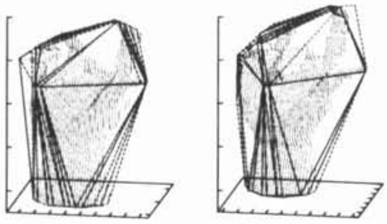


Figure 2. The resulting 3D convex hull of two range images of a hand model.

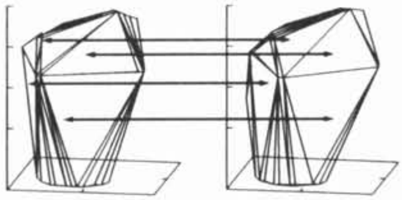


Figure 3. The correspondences of planar face pairs of the convex hull.

3. All the range images used in the experiments were obtained with a range finder which projects spatially coded pattern on an object and calculates the range from its image based on the triangulation principle.

2 System Configuration

2.1 Producing of the Convex Hull

To produce a three dimensional convex hull of the partial shape of the object observed in each range image, we apply so-called "gift-wrapping" method[4]. This method simulates wrapping of an object by a sheet. We find a plane supported by three object points such that all the object points except them lie in one side of the plane. This plane will be a face of the convex hull. However, because of the measurement error, it is possible that these faces of the resulting convex hull may be divided into many small triangles. To avoid such the division of the face, we employ an eight way perturbation technique which is an analogy of SoS(Simulation of Simplicity)[5]. In this technique, the object points are displaced with small amounts in an outward direction. The vertices supporting the face move relatively large amount and the points on and near the face are forced to be placed inside the face.

The convex hulls resulted for the two range images of Fig.1(b) are shown in Fig.2.

2.2 Finding Corresponding Face Pairs

Next, we find corresponding face pairs between consecutive two images in the series of range images.

To obtain sufficiently accurate relative position, we find correspondences only between large faces.

We build adjacency graphs of large faces within the each respective convex hull, and match the graphs by evaluating areas of the faces.

The resulted correspondences for a pair of range images of Fig.2 is shown in Fig.3.

2.3 Determining Relative Position of Viewpoints

Finally, we determine relative positions of all the viewpoints simultaneously and express each partial shape of the object in a common world coordinate system.

2.3.1 Notation

The positional relationship between i-th($i = 1, \ldots, N$) camera coordinate system and the world coordinate system is represented by a rotation matrix $\boldsymbol{R}_i$ and a translation vector $\boldsymbol{t}_i$. A point $\boldsymbol{x}$ in the world coordinate system is expressed in the i-th camera coordinate system with $\boldsymbol{R}_i$ and $\boldsymbol{t}_i$ as

$$\boldsymbol{x}_i = \boldsymbol{R}_i \boldsymbol{x} + \boldsymbol{t}_i. \qquad (1)$$

The j-th($j = 1, \ldots, M$) planar face is represented in the world coordinate system by a normal vector $\boldsymbol{n}_j$ and distance d_j to the origin, while, in the i-th coordinate system, it is represented by $\boldsymbol{n}_j^i$ and d_j^i. The relation between $\boldsymbol{n}_j$ and $\boldsymbol{n}_j^i$ is given as

$$\boldsymbol{n}_j^i = \boldsymbol{R}_i \boldsymbol{n}_j, \qquad (2)$$

and the relation between d_j and d_j^i is given as

$$d_j^i = (\boldsymbol{R}_i \boldsymbol{n}_j)^T \boldsymbol{t}_i + d_j. \qquad (3)$$

From Eq.(2), we determine $\boldsymbol{R}_i$'s and $\boldsymbol{n}_j$'s first by using $\boldsymbol{n}_j^i$'s calculated from observed range images.

2.3.2 Rotation Matrix

Now, a matrix $\boldsymbol{P}$ is defined as

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{n}_1^1 & \cdots & \boldsymbol{n}_j^1 & \cdots & \boldsymbol{n}_M^1 \\ \ddots & \vdots & \ddots & \vdots & \ddots \\ \boldsymbol{n}_1^i & \cdots & \boldsymbol{n}_j^i & \cdots & \boldsymbol{n}_M^i \\ \ddots & \vdots & \ddots & \vdots & \ddots \\ \boldsymbol{n}_1^N & \cdots & \boldsymbol{n}_j^N & \cdots & \boldsymbol{n}_M^N \end{bmatrix}. \qquad (4)$$

Substituting Eq.(2) into Eq.(4), we have

$$\boldsymbol{P} = \boldsymbol{R}\boldsymbol{N}, \qquad (5)$$

where

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_1 \\ \boldsymbol{R}_2 \\ \vdots \\ \boldsymbol{R}_N \end{bmatrix} \text{ and } \boldsymbol{N} = [\boldsymbol{n}_1, \boldsymbol{n}_2, \cdots, \boldsymbol{n}_M]. \qquad (6)$$

Hence, $\boldsymbol{R}_i$ and $\boldsymbol{n}_j$ are determined by decomposing $\boldsymbol{P}$ into $3N \times 3$ matrix $\boldsymbol{R}$ and $3 \times M$ matrix $\boldsymbol{N}$.

However, many of components of the $\boldsymbol{P}$ cannot be obtained because every viewpoint has some invisible faces. Then, we find the best decomposition with respect to the observed components $\boldsymbol{n}_j^i$ of $\boldsymbol{P}$. We define a criterion function J_1[3] to be minimized as

$$J_1 = \sum_i \sum_j \gamma_{ij} \| \boldsymbol{n}_j^i - \boldsymbol{R}_i \boldsymbol{n}_j \|^2 \qquad (7)$$

where γ_{ij} equals to the reliability of the observation λ_{ij} if $\boldsymbol{n}_j^i$ is observed, and otherwise it equals to 0. The value λ_{ij} depends on the accuracy of the observation of j-th face from the i-th viewpoints. In the following experiments, all λ_{ij}'s are set to be 1.

As described before, we make use of the constraint about the rotation matrices that they should be orthonormal with unit determinant. The rotation matrices which satisfy this condition are determined by an iterative algorithm. In the algorithm, when correct $\boldsymbol{R}_i$'s are given, $\boldsymbol{n}_j$'s minimizing J_1 is easily obtained by standard least squares fitting, and when correct $\boldsymbol{n}_j$'s are given, each $\boldsymbol{R}_i$ which minimize J_1 is obtained according to the theory described in [6].

If we assume that correct $\boldsymbol{R}_i$'s have been given, $\boldsymbol{n}_j$'s which minimize J_1 are obtained as the solution of

$$\boldsymbol{w} = \boldsymbol{G}\boldsymbol{n}, \tag{8}$$

where

$$\boldsymbol{n} = \begin{bmatrix} \boldsymbol{n}_1 \\ \boldsymbol{n}_2 \\ \vdots \\ \boldsymbol{n}_M \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} \gamma_{11}\boldsymbol{R}_1 & & \\ \vdots & & \\ \gamma_{N1}\boldsymbol{R}_N & & \\ & \vdots & \\ & & \gamma_{1M}\boldsymbol{R}_1 \\ & & \vdots \\ & & \gamma_{NM}\boldsymbol{R}_N \end{bmatrix}, \tag{9}$$

and

$$\boldsymbol{w}^T = \left[\gamma_{11}\boldsymbol{n}_1^{1T}, \ldots, \gamma_{N1}\boldsymbol{n}_1^{NT}, \ldots, \gamma_{NM}\boldsymbol{n}_M^{NT}\right]^T. \tag{10}$$

On the other hand, if we assume that correct $\boldsymbol{n}_j$'s have been given, each $\boldsymbol{R}_i$ which minimize J_1 with constraint that it should be orthonormal with unit determinant is obtained by conditional minimization of J_1. It is realized by introducing Lagrange multipliers. The solution has been given in [6]. According to it, $\boldsymbol{R}_i$ is given as follows. Let

$$\boldsymbol{W}_i = [\gamma_{i1}\boldsymbol{n}_1^i \ \ \gamma_{i2}\boldsymbol{n}_2^i \ \ \ldots \ \ \gamma_{iM}\boldsymbol{n}_M^i]. \tag{11}$$

Then, by the singular value decomposition, we have

$$\boldsymbol{W}_i\boldsymbol{N}^T = \boldsymbol{U}_i\boldsymbol{D}_i\boldsymbol{V}_i^T, \tag{12}$$

where $\boldsymbol{U}_i$ and $\boldsymbol{V}_i$ are column orthogonal matrices and $\boldsymbol{D}_i$ is a diagonal matrix. Then, the solution is given as

$$\boldsymbol{R}_i = \boldsymbol{U}_i\boldsymbol{S}_i\boldsymbol{V}_i^T, \tag{13}$$

where

$$\boldsymbol{S}_i = \begin{cases} \boldsymbol{I}, & \text{if } \det \boldsymbol{W}_i\boldsymbol{N}^T > 0 \\ \operatorname{diag}[1, 1, -1], & \text{if } \det \boldsymbol{W}_i\boldsymbol{N}^T < 0 \end{cases}. \tag{14}$$

Thus, the algorithm is:

1. Initialize $\boldsymbol{R}_i$'s for all i.
2. Update $\boldsymbol{n}_j$'s to minimize J_1 assuming that correct $\boldsymbol{R}_i$'s have been given.
3. Update each $\boldsymbol{R}_i$ to minimize J_1 assuming that correct $\boldsymbol{n}_j$'s have been given.
4. Stop if converged, or go back to 2.

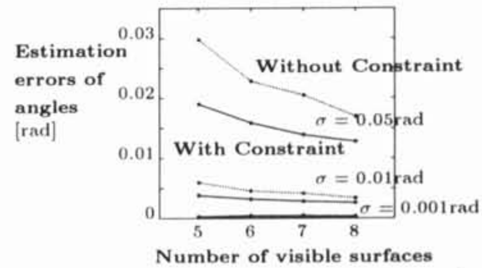
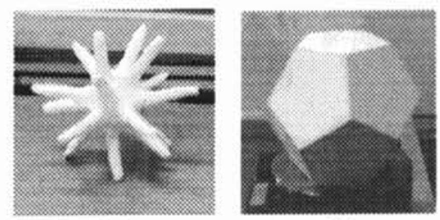


Figure 4. Absolute estimation errors of rotation angles for the number of visible surfaces.



(a) (b)

Figure 5. The models used for the experiment. (a)Non-polyhedral object and (b) polyhedron.

2.3.3 Translation Vector

Once $\boldsymbol{R}_i$'s and $\boldsymbol{n}_j$'s are determined, $\boldsymbol{t}_i$'s and d_j's are obtained by standard least squares fitting. To do this, we define a criterion function J_2 as

$$J_2 = \sum_i \sum_j \gamma_{ij} \| d_j^i - \{(\boldsymbol{R}_i\boldsymbol{n}_j)^T\boldsymbol{t}_i + d_j\} \|^2. \tag{15}$$

$\boldsymbol{t}_i$'s and d_j's which minimize J_2 are given as the solution of

$$\boldsymbol{v} = \boldsymbol{F}\boldsymbol{p} \tag{16}$$

where

$$\boldsymbol{p} = \begin{bmatrix} \boldsymbol{t}_1 \\ \boldsymbol{t}_2 \\ \vdots \\ \boldsymbol{t}_N \\ d_1 \\ d_2 \\ \vdots \\ d_M \end{bmatrix}, \quad \boldsymbol{F} = \begin{bmatrix} (\boldsymbol{R}_1\boldsymbol{n}_1)^T & & 1 & \\ \vdots & & & \ddots \\ (\boldsymbol{R}_1\boldsymbol{n}_M)^T & & & 1 \\ & \vdots & & \\ & (\boldsymbol{R}_N\boldsymbol{n}_1)^T & 1 & \\ & \vdots & & \ddots \\ & (\boldsymbol{R}_N\boldsymbol{n}_M)^T & & 1 \end{bmatrix}, \tag{17}$$

and

$$\boldsymbol{v}^T = \left[d_1^1, \ldots, d_M^1, \ldots, d_1^N, \ldots, d_M^N\right]^T. \tag{18}$$

3 Experimental Results

3.1 Experimental Results of Determining Rotation Matrices

To examine the performance of the algorithm for determining rotation matrices in Section 2.3.2, the following experiments were done. In the experiments, instead of observing the object from multiple viewpoints, we observe it from fixed viewpoint by rotating it.

First, we applied our method to a set of synthetic normal vectors. We generated twelve unit vectors randomly in 3D space. And, we rotated the set of these vectors by $\frac{1}{6}\pi n$ radians ($n = 1, \ldots, 12$). Furthermore, we added Gaussian noise to each observation in each set. The experiments were done for three cases where the standard deviation of the added noise was $0.001, 0.01$ and 0.05 radians respectively. Using from 5 to 8 vectors for each set, we estimated the rotation angle by the proposed algorithm. Figure 4 shows the estimation errors of the angles. It also shows errors of angles estimated by the method without the constraint of the rotation matrices.

As shown in Fig.4, for all cases, better results were obtained by the proposed method than by the method without the constraint. The improvement became distinct as the given noise was large. In addition, better results were obtained with number of the observed vectors.

Next, we applied our algorithm to real range images of a complex object and a simple dodecahedron shown in Fig.5. We observed each objects from 18 viewpoints respectively, by rotating them by 20 degrees. Table 1 shows estimation errors of the angles. It shows the results obtained by the proposed method along with the results by the following two methods:

(1) obtaining all rotation matrices simultaneously without constraint,

(2) obtaining rotation matrices among two viewpoints successively with constraint.

As shown in Table1, the best results are obtained by the proposed method.

Finally, we show an experimental result about a convergence of the iteration in the algorithm. We observed the object shown in Fig.5(a) from 18 viewpoints by rotating the object by 20 degrees. The convex hull of the object had 14 faces. The number of observable face from each viewpoint was 5 or 6. Figure 6 shows the values of $J_1 = \sum_i \sum_j \gamma_{ij} \|\boldsymbol{n}_j^i - \boldsymbol{R}_i \boldsymbol{n}_j\|^2$ in Eq.(7) with respect to the number of iterations in the algorithm. They are averages of every 100 trials. In each trial, $\boldsymbol{R}_i$'s were initialized randomly. It also shows the results obtained by the Gauss-Newton method without constraint of the rotation matrix. The results obtained by the proposed algorithm converged as well as the those by Gauss-Newton algorithm.

3.2 Experimental Result for the Hand Model

Figure 7 shows a reconstructed shape of the hand model shown in Fig.1(a). We observed this hand model from 6 viewpoints. Two of them were shown in Fig.1(b). We produced first three dimensional convex hull of partial shape of the object observed in each range images(Fig.2), found correspondence of faces(Fig.3), determined relative position of viewpoints, and expressed each partial object in the world coordinate system.

Table 1. Absolute estimation errors of rotation angles for real range images shown in Fig.6. These values are shown in degrees.

	Proposed Method	(1)	(2)
Complex object	0.47	1.02	0.96
Polyhedron	0.26	0.43	0.47

4 Conclusion

A method to register multiple range images has been presented. To deal with non-polyhedral objects, the 3D convex hull of the object was introduced. Then the correspondences of the faces of convex hull are found in such the way that we build first adjacency graph of comparatively large faces in the convex hull, then match them based on those face areas. Using this face correspondences, the relative positions of respective viewpoints are calculated by the algorithm which takes account of the constraint of the rotation matrix that it should be orthonormal with unit determinant. Using this method, we obtained complete shape of a complex target object.

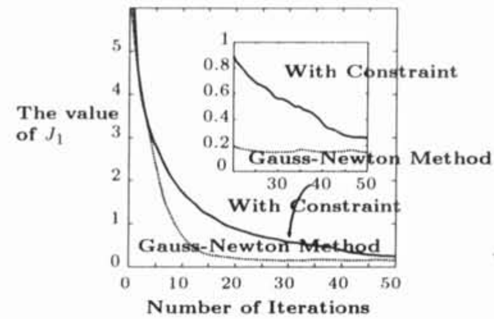


Figure 6. The value of J_1 with respect to the number of iterations.

Figure 7. The reconstructed shape of the model of a hand.

References

[1] Y.Chen and G.Medioni. Object Modeling by Registration of Multiple Range Images. *Proc. IEEE Int. Conf. RA*, pp.2724–2729, 1991.

[2] Y.Kawai, T.Ueshiba, T.Yoshimi and M.Oshima. Reconstruction of 3D Objects by Integration of Multiple Range Data. *Proc. ICPR*, pp. A154–157, 1992.

[3] H.Shum, K.Ikeuchi and R.Reddy. Principal Component Analysis with Missing Data and Its Application to Polyhedral Object Modeling. *IEEE Trans. PAMI*, Vol. 17, No. 9, pp. 854–867, 1995.

[4] D.R.Chand and S,S.Kapur. An Algorithm for Convex Polytopes. *J. ACM*, Vol. 17, No. 1, pp. 78–86, 1970.

[5] T.Ono. An Eight-Way Perturbation Technique for the Three-Dimensional Convex Hull. *IPSJ SIG Notes Algorithm*, Vol. 39, No. 5, pp. 33–39, 1994.

[6] S.Umeyama. Least-Square Estimation of Transformation Parameters Between Two Point Patterns. *IEEE Trans. PAMI*, Vol. 13, No. 4, pp. 376–380, 1991.