2011 International Symposium on Nonlinear Theory and its Applications
NOLTA2011, Kobe, Japan, September 4-7, 2011

Real Time 3D Object Recognition and Tracking by Robust and Fast Alignment

Hironobu Fukai† and Gang Xu†

†College of Information Science & Engineering, Ritsumeikan University
1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577 Japan
Email: fukai@fc.ritsumei.ac.jp, xu@3dmedia.co.jp

Abstract—This paper proposes a alignment and tracking method of 3D point cloud. The problem of point cloud alignment can be roughly classified into two. The problem of estimation of the relation from unknown geometry position is coarse registration. Moreover, in case of roughly position is known, the problem of estimation of the relation from known initial position is fine registration. In this study, we solve the coarse registration problem by exhaustive search, and prepare the initial positions for fine registration. The fine registration problem is solved by the ICP algorithm. In this case, reduction of initial positions for the ICP algorithm by non-extremum suppression that uses distance between range data and model solves the problem of calculation cost. The distance evaluation function that robust for measurement error tackles the outlier problem. The problem of calculation cost is solved by registering the distance from the model beforehand by the distance field. The effectiveness of the proposed method is shown by a actual data.

1. Introduction

Recently, we can measure object 3D shape by stereo camera, infrared time of flight (TOF) camera, and so on. These measurements can get point clouds. This technology contributes to model of 3D shape by alignment of point clouds, to recognize a object, simultaneously localization and mapping (SLAM) compared to conventional 2D restoration methods. To achieve these applications, it is necessary to align the model and point cloud. In this study, we propose robust and fast alignment method.

One of the major alignment methods is iterative closest point (ICP) algorithm [1]. ICP algorithm estimates closest point between two point clouds, translation and rotation between model data and measured data. However, in general, before applying ICP algorithm, point clouds must be registered to roughly correct positions. Moreover, ICP has a outlier problem. To solve the outlier problem, many method are proposed [2, 3]. From these reports, the outlier problem is solve by weight coefficient. On the other hand, to solve the initial position problem, direction standardization method by feature point that is extracted by the curvature from the range image is proposed[4]. However, the method that uses the feature point has the problem that alignment cannot be done when the feature point cannot be extracted.

Therefore, motivation in this study is to propose a robust and fast alignment method that solves the outlier and initial position problems. “Robust” have 2 meanings. As for one, robustness for initial position and the second is robust for outlier. In the proposed method, we use the ICP algorithm for the alignment, however, local minimum problem happens according to the initial position, so we search initial positions by exhaustive matching. The calculation cost is reduced by non-extremum suppression for good initial position selection. Furthermore, in the ICP algorithm and exhaustive search evaluation function, we use Tukey weight function[7] to solve outlier problem. In addition, we realize fast processing by distance field (DF)[8]. DF is the look up table of the closest point and its distance. In this study, we use the proposed method for the 3D object recognition. Moreover, we use only ICP algorithm for object tracking.

2. Alignment

To estimate the relation of the geometry position of the range image and the model data is called alignment. The alignment becomes a problem of estimating the rotation matrix and the translation vector because we deal in the rigid body. This rotation matrix and the translation vector can be calculated by minimizing the least mean square error between the model data and the measured data. In this section, we show the problem in the alignment, rotation representation, and conventional alignment method.

2.1. The problem in the alignment

The relation of the geometry position estimation has 6 degree of freedom (DOF)(Fig.1). The model coordinate system is defined by (X_m, Y_m, Z_m), camera coordinate system is defined by (X_c, Y_c, Z_c). Moreover, 6 DOF parameter represent the rotation parameter (α, β, γ) and translation parameter (x, y, z).

The problem in the alignment, search space is very huge. This search space is not guaranteed that it is single peak space, and it is difficult to obtain the optimum solution in a simple search method such as hill climbing. Moreover, representation of the rotation is necessary. Furthermore, when model data and measured data are same shape and relation of the correspondence point is understood, the formulation is easy. However, in general, these correspondence point is unclear. In the alignment, we need to solve these problems.

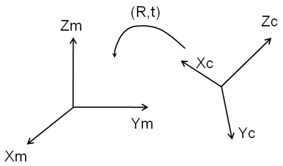

Figure 1: Representation of the geometry position in the alignment

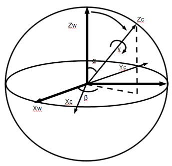

Figure 2: Representation of the rotation

2.2. Representation of the rotation

Mentioned above, representation of the rotation is one of the alignment problem. In this subsection, we describe the representation of the rotation. Rotation has 3 parameters (α, β, γ). Rotation representation is, in other words, how to represent the rotation matrix by these 3 parameters. In this study, we use latitude, longitude and camera rotation for representation of pose (Fig. 2). In addition, we describe rotation matrix $\boldsymbol{R}$ by 3D vector.The direction of a vector means rotation axis, norm of the vector indicates rotation.

The rotation matrix $\boldsymbol{R}$ is obtained by Rodrigues formula.

In this case, we define the rotation between Z_m and Z_c as the latitude α and the longitude β. Thus, rotation matrix is shown as follows.

$$\mathbf{R}(\alpha,\beta) = \mathbf{I} + \sin\alpha \begin{bmatrix} -\sin\beta \\ \cos\beta \\ 0 \end{bmatrix}_{\times} + (1-\cos\alpha) \begin{bmatrix} -\sin\beta \\ \cos\beta \\ 0 \end{bmatrix}_{\times}^{2} \tag{1}$$

We write rotation matrix $\mathbf{R}(\alpha,\beta)$ by 3D vector $\boldsymbol{r}$. Next, we explain rotation of the γ. This rotation means Z_c axis rotation.Then,

$$\mathbf{R}(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2}$$

Thus,

$$\mathbf{R} = \mathbf{R}(\alpha,\beta)\mathbf{R}(\gamma). \tag{3}$$

2.3. Alignment by least mean square

It has already been described to be able to formulate alignment as a minimization problem of the least mean square error when the correspondence of the point is already-known [5, 6]. However, correspondence of the point is unknown in almost case. Then, the ICP algorithm [1] was proposed that search associate points and estimate rotation and translation. However, in general, before applying the ICP algorithm, point clouds must be registered to roughly correct positions.

Therefore, in this study, we estimate global minimum by exhaustive search.

3. Alignment by exhaustive search

3.1. Exhaustive search

We search a optimal initial position for the ICP algorithm by exhaustive search. In conventional method, the optimize method that use the Levenberg-Marquardt algorithm from exhaustive search though it is not a range image is proposed [9]. Moreover, method that use exhaustive search and depth map is proposed [10]. The effectiveness of the exhaustive search is shown. Therefore, we use the exhaustive search. We search 6 DOF space roughly.This sampling late is defined by model shape.

Some cases that become the same local solutions even if the ICP algorithm is used respectively according to the initial positions exist. Then, we suppress initial position that is non-extremum point. This process calls non-extremum suppression, and is explained in next subsection.

3.2. Non-extremum suppression

All pose and position distances are too huge and overplus for initial positions of the ICP algorithm. Therefore, we reduce non-extremum initial positions by comparison of neighborhood distance value. Neighborhood position is defined by all neighbor positions $3\times3\times3$. However, neighborhood of pose definition is too difficult. In this study, we define the neighborhood of pose as less than constant value between 2 poses. We set the rotation matrix $\mathbf{R}$. Moreover, $\mathbf{R_1}$ means a rotation matrix of pose1, $\mathbf{R_2}$ means a rotation matrix of pose2, respectively. $\mathbf{R}$ can be obtained by following equation (Fig.3).

$$\mathbf{R} = \mathbf{R_1}\mathbf{R_2}^{T}. \tag{4}$$

If we set the rotation angle of $\mathbf{R}$ as a θ, it can be calculate by Rodrigues formula.

$$\theta = cos^{-1}\frac{trace(\mathbf{R}) - 1}{2} \tag{5}$$

In this study, the neighborhood is defined as double the sampling angle.

3.3. Alignment by the ICP algorithm

Precise alignment is achieved by the ICP algorithm. In this study, we use many initial position that is obtained by non-extremum suppression for the ICP algorithm. This image is shown in Fig. 4 by 1 dimension for simplicity. In this figure, horizontal axis means search space, vertical axis

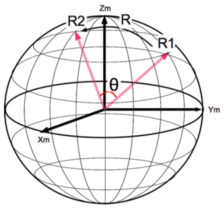

Figure 3: Calculation of pose neighborhood

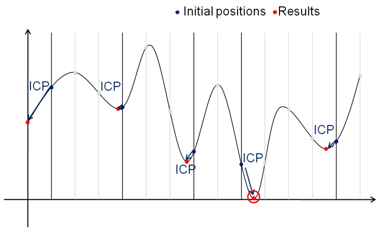

Figure 4: Image of the proposed method

means distance between the model data and the measured data. As you can see, to use some optimal initial positions, we can get global solution.

3.4. Tracking by the ICP algorithm

The proposed method track the recognized object. In tracking, we use only ICP algorithm for real time tracking.

4. Distance evaluation function and calculation

To evaluate a distance between model and measured data, it is necessary to calculate closest point and its distance. However, measured data has some noise and outlier. Then, distance evaluation of only associate point is necessary. In this section, we explain a evaluation function and distance field (DF) [8] that can calculate distance.

4.1. Evaluation function of exhaustive search

To evaluate the distance between model and measured data, we need to calculate the closest point and its distance. However, some measured point has noises and outliers. To solve this problems, there are many method has proposed [2, 3, 7, 11]. In this study, we use Tukey weight function for distance evaluation function.

$$\sum_{i=1}^{N} S(i) = \begin{cases} (1 - \frac{d(i)^2}{\tau_d^2})^2, & if(d^2 \leq \tau_d^2) \\ 0, & else \end{cases} \tag{6}$$

$$d(i)^2 = \|\mathbf{X_{mi}} - (\mathbf{R}\mathbf{X_{ci}} + \mathbf{t})\|^2 \tag{7}$$

where τ_d means a threshold of the distance. N is number of measured points,$d(i)$ is distance of each associate points, $\mathbf{Xc_i}$ means index number of point cloud that has N points.

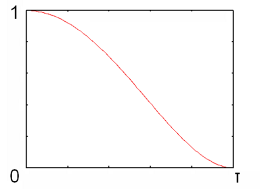

Figure 5: Graph of weight function

$\mathbf{R}$ is rotation matrix, $\mathbf{t}$ indicates translation vector. $\mathbf{Xm_i}$ is closest point of model from $\mathbf{RXc_i} + \mathbf{t}$. We can estimate robustly by using τ_d. This evaluation function is plotted in Fig.5.

From this figure, evaluation value becomes 0 when the distance over threshold τ. Then, we solve the outlier and mismatched part problem. In this study, τ is defined by half of translation search step width, because only a point very near the model wants to be acquired.

4.2. Evaluation function of the ICP algorithm

In ICP [1] algorithm, we use the evaluation function that use the Tukey weight function.

$$C_j = \sum_{i=1}^{N} w(i)\|\mathbf{Xm_i} - (\mathbf{R_{j+1}}\mathbf{Xc_{i,j}} + \mathbf{t_{j+1}})\|^2 \tag{8}$$

where N is number of point clouds, i is index number, j means iterative number. $\mathbf{Xc_{i,j}}$ is point of i number when j iterated. $\mathbf{R_j}$ indicates j rotation matrix, $\mathbf{t_j}$ indicates j translation vector. $\mathbf{Xm_i}$ is closest point from $\mathbf{R_{j+1}Xc_{i,j}} + \mathbf{t_{j+1}}$. Then, same as Fig.5, we reduce outlier and mismated point. In the ICP algorithm, threshold is defined by standard deviation in each iteration [11].

4.3. DF from 3D model

In order to evaluate distance between the model data and the measured data, it is necessary to calculate associate points and its distances, respectively. Thus, to reduce calculation costs, we employ DF[8] that is a look up table of the distances and associate points.

5. 3D object recognition

In this study, we apply the proposed method for the 3D object recognition. The relation of camera coordinate system and world coordinate system are shown in this section.

$$\mathbf{X_m} = \mathbf{R}(\mathbf{R_0}\mathbf{X_c} + \mathbf{t}). \tag{9}$$

where, $\mathbf{t}$ is translation vector, $\mathbf{R}$ means rotation vector, $\mathbf{R_0}$ means z axis invert rotation matrix. In this case, pose parameter ranges are $0 \leq \alpha \leq 180$, $0 \leq \beta < 360$, $0 \leq \gamma < 360$.

6. Experiments

In order to show the effectiveness of the proposed method, the experiment is done. The point cloud is measured by D-Imager that is Time of Flight (TOF) censer [12]. We use core-i7 2.6GHz CPU and 3GB memory. Moreover, maximum size of DF is set as 100. Furthermore, the size of DF allotted to the maximum size of the object was assumed to be 50. Sampling late of rotation is set as 20-30 degree and sampling late of translation is se as 8-10 voxels. These parameter is defined by model shape and experimentally.

6.1. Result

Table 1: Detail of 3D object recognition experimental result

Number of search	1M
Number of points	699
Number of initial position	33
Step width (mm)	165.31
DF voxel width (mm)	18.167
RMSE (mm)	36.28
Execution time (sec)	0.708

In this case, we use elevator hall model and box objects. These models are known data. The point clouds can get 19200 data (120×160), but we reduce there point for one-tenth. The result is shown in Table1. From this table, RMSE (root mean square error) is less than measuring error. It mean that this experiment can align precisely. From this result, we can realize the 3D object recognition. Moreover, we try to track the recognized object by only ICP algorithm. The system runs at over 30 frames per second.

7. Conclusions

In this study, we proposed the robust and fast alignment and tracking by exhaustive search. The proposed method use the ICP algorithm for fine registration and get the optimal initial positions for the ICP algorithm. Moreover, we use weight function for distance evaluation function. Then, we can estimate robustly. This method originality is not only robust but fast. The fast algorithm is realized by non-extremum suppression and DF. From this table, RMSE is less than measuring error. Then, we can estimate and track the 3D object.

In future works, we will apply the proposed method to various situations. Moreover, we realize SLAM, modeling, AR/MR, and so on by the proposed method.

Acknowledgments

This work was supported by the Private Univertigy Strategy Research Foundation of Business Development Support.

References

[1] Besl, P. J., and McKay, N. D. "A Method for Registration of 3-d Shapes," *IEEE Trans. on Pattern Analysia and Machine Intelligence*, Vol. 14, Num. 2, pp. 239–256, 1992.

[2] Z. Zhang, "Iterative Point Matching for Registration of Free-form Curves and Surfaces," *Int. Journal of Computer Vision*, Vol.13, No.2, pp.119–152, 1994.

[3] T. Masuda and N. Yokoya, "A Robust Method for Registration and Segmentation of Multiple Range Images," *CVIU*, Vol.61, No.3, pp.295–307, 1995.

[4] T. Takeguchi, S. Kaneko, T. Kondo, S. Igarashi "Robust Object Recognition Based on Depth Aspect Image Matching," IEICE, D-II, Vol.J84, No.8, pp.1710–1721, 2001.

[5] Tomasi, C. and Kanade, T., "Shape and Motion from image streams under orthography: A factorization method," *Int. Journal on Computer Vision*, No.9, Vol.2, pp.137–154, 1992.

[6] Xu, G. and Zhang, Z., "Epipolar Geometry in Stereo, Motion and Object Recognition –A Unified Approach," Kluwer, Dordrecht, 1996.

[7] Frederickmosteller, J. W. Tukey "Data Analysis and Regression," Addison-Wesley Publishing Company, Reading, MA, 1977.

[8] T. Kato, T. Hirata, T. Saito, and K. Kise "An Efficient Algorithm for the Duclidean Distance Transformation," IEICE, D-II, Vol. J78, No. 12, pp. 1750–1757, 1995.

[9] M. Ulrich, C. Wiedemann and C. Steger "CAD-Based Recognition of 3D Objects in Monocular Images," *Proc. of International Conference on Robotics and Automation*, pp.1191–1198, 2009.

[10] S. Kashihara and G. Xu "CAD-Based Recognition of 3D objects in Binocular Images," *Proc. of the 5th Joint Workshop on Machine Perception and Robotics*, Kyoto, 2009.

[11] MVTec software GmbH, Carsten Steger, Markus Ulrich, Christian Wiedemann "Machine Vision Algorithms and Applications," Linx, 2008.

[12] Panasonic, D-IMager, http://denko.panasonic.biz/Ebox/3DImageSensor/