PERGAMON

Pattern Recognition 36 (2003) 2041 – 2047

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Robust matching of 3D contours using iterative closest point algorithm improved by M-estimation

Shun'ichi Kaneko[a,*], Tomonori Kondo[b], Atsushi Miyamoto[c]

[a] Department of Control and Information Engineering, Graduate School of Engineering, Hokkaido University, Sapporo N13-W8, Japan
[b] Oki Data Co., Ltd., 4-1-22, Shibaura, Tokyo, Japan
[c] PERL Hitachi Co., Ltd., 292, Yoshida-cho, Yokohama, Japan

Received 7 June 2002; accepted 22 November 2002

Abstract

An extension of the iterative closest point matching by M-estimation is proposed for realization of robustness to non-overlapping data or outlying data in two sets of contour data or depth images for rigid bodies. An objective function which includes independent residual components for each of x, y and z coordinates is originally defined and proposed to evaluate the fitness, simultaneously dealing with a distribution of outlying gross noise. The proposed procedure is based on modified M-estimation iterations with bi-weighting coefficients for selecting corresponding points for optimization of estimating the transforms for matching. The transforms can be represented by 'quaternions' in the procedure to eliminate redundancy in representation of rotational degree of freedom by linear matrices. Optimization steps are performed by the simplex method because it does not need computation of differentiation. Some fundamental experiments utilizing real data of 2D and 3D measurement show effectiveness of the proposed method. When reasonable initial positions are given, the unique solution of position could be provided in spite of surplus point data in the objects. And then the outlying data could be filtered out from the normal ones by the proposed method.

Keywords: Depth image; Matching; Registration; Robustness; Outlier

1. Introduction

Recognition of three dimensional rigid objects and estimation of their positions and/or poses using sparse data or dense images of depth contours has been one of the important issues in many fields. The matching or registration techniques is utilized to obtain a rigid transformation consisting of rotation and translation parameters in order to perform the best fit between any contour data sets, which are well utilized in engineering fields, such as sensing in robotics [1], reverse engineering in manufacturing systems [2], and object modeling in advanced mixed or virtual reality [3]. There have been much requirement for reliable, effective and efficient methods in their real applications. Furthermore, it is expected to be robust for ill-conditions in data measured in real-world instrumentation, such as occlusion and deficiency. Depth data is generally classified into two categories: sparse depth data and dense depth images. For both kinds of depth data, it is essential to cope with insignificant or redundant portions, such as non-overlapping data, which are caused by the occlusion and deficiency of data derived from the difference in the field of view. Furthermore, it is important to handle isolated points from the significant parts in order to obtain consistent matching. These partial data can degrade matching to wrong solutions with inconsistent correspondence between points, therefore, these ill-conditioned data should be called outlying data. We have been looking for any efficient, effective and robust methods for these outlying data. Local curvatures were used for matching range images [4], but they are weak even for small noise observed in the real world data. A method using attributed depth

* Corresponding author. Tel.:/fax: +81-11-706-6436.
E-mail address: kanekos@coin.eng.hokudai.ac.jp (S. Kaneko).

0031-3203/03/$30.00 © 2003 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0031-3203(03)00050-5

data with other information such as color and intensity [5] have improved performance in computational convergence to a correct matching, but from different view directions it is generally difficult to have the consistent attributes which are dependent on illumination condition. Discriminating by coefficients which are dependent on magnitudes of depth residuals has been proposed for avoiding local minima in search [6]. However, it has a heuristic algorithm for coefficient modification and many parameters to be adjusted, and then it has no sufficient performance in convergence to global solutions. A method for robust registration of multiple range images utilizing the least median of squares [7] has been proposed for three dimensional object modeling. The least median of squares principle for robust estimation is effectively used in various fields [8], however, the ratio of outlying data is limited upto fifty percent and actually rather small, and furthermore, the computation cost for sorting is not expensive.

In this study, we address the problem of robustly matching three dimensional contours of rigid bodies with no additive measurement but only depth data. We modify the iterative closest point (hereafter, ICP) algorithm [9] by introducing the M-estimation principle for coefficient determination based on depth residuals in order to perform fitting without any wrong influence from the outlying data [10]. ICP can be applied to straightforwardly obtain rigid transforms which matches two depth data by utilizing provisional correspondence between points and the least square fitting, but it is assumed for the data to have few outlying data. The proposed method improved by M-estimation which has a theoretical background of robust parameter estimation [11] enables ICP to cope with the outlying data, and then it can be used with some other principle such as using attributed depth data.

2. Robust matching

2.1. Iterative closest point algorithm

Fig. 1 shows the schematic procedures of the proposed method. While the original version of ICP consists of the five major procedures from a to e in the improved one, the new iterative procedures f and g are introduced. These procedures are dedicated for optimizing an objective function through iterative estimation of weighting coefficients.

We formalize the original version of ICP [9] in the remainder of this section for the extension that will be provided in the next section to the detail. Let $F = \{\boldsymbol{f}_i\}_{i=1}^{N}$ and $G=\{\boldsymbol{g}_k\}_{k=1}^{K}$ denote two sets of depth data to be matched, the indices and sizes of which can be different. F is then transformed to G through each iterative registration step, in which we have a temporal shifted data $F^{(m)} = \{\boldsymbol{f}_i^{(m)}\}$ subscripted with the step number (m). We assume $m = 0, 1, 2, \ldots,$ and $F^{(0)} = F$. $\mathscr{C}^{(m)} = \{k_i^{(m)}\}_{i=1}^{N}$ is a list of correspondence between both the data defined by the minimum distance rule as

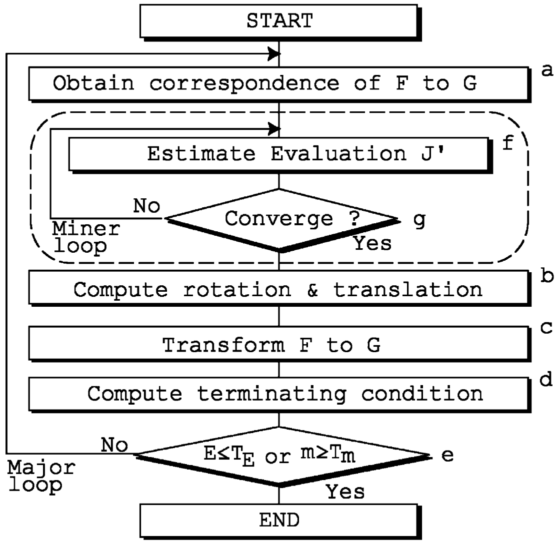

Fig. 1. Improved ICP algorithm.

follows:

$$k_i^{(m)} = \left\{ k \,\Big|\min_k |\boldsymbol{f}_i^{(m)} - \boldsymbol{g}_k| \right\}, \tag{1}$$

where multiple $k_i^{(m)}$ possibly share the same point k in G. Each correspondence depends on relative alignment of the points, therefore the initial condition is much important for good convergence. A rigid transform consists of the rotation matrix $R^{(m)}$ and the translation vector $\boldsymbol{t}^{(m)}$, resulting the iterative transformation as $\boldsymbol{f}_i^{(m)} = R^{(m)}\,\boldsymbol{f}_i^{(m-1)} + \boldsymbol{t}^{(m)}$. We have to determine the six degrees of freedom including the three for rotation and the other three for translation by ICP. While the three dimensional translation vector has simply three parameters as $\boldsymbol{t} = (t_x, t_y, t_z)^{\mathrm{T}}$, the rotation matrix is apparently composed of nine elements which should go along with six conditions for orthonormality. Simple iterative optimization based on the least square principle can not guarantee this orthonormality. Hence, ICP employs *quaternion* (q_1, q_2, q_3, q_4) for representing the rotation parameters in order to reduce this problem, defining the rotation matrix as

$$R = \begin{pmatrix} \alpha_{1234} & \beta_{2314}^{-} & \beta_{2413}^{+} \\ \beta_{2314}^{+} & \alpha_{1324} & \beta_{3412}^{-} \\ \beta_{2413}^{-} & \beta_{3412}^{+} & \alpha_{1423} \end{pmatrix}, \tag{2}$$

where $\alpha_{abcd} = (q_a)^2 + (q_b)^2 - (q_c)^2 - (q_d)^2$, $\beta_{abcd}^{\pm} = 2(q_a q_b \pm q_c q_d)$, $a, b, c, d \in \{1,2,3,4\}$, and $\det(R) = 1$. For simplicity, we omitted the subscript (m) when representing parameters. The parameter vector $\boldsymbol{q}^{(m)} = (q_1^{(m)}, q_2^{(m)}, q_3^{(m)}, q_4^{(m)}, t_x^{(m)}, t_y^{(m)}, t_z^{(m)})^{\mathrm{T}}$ should is obtained as a solution of minimization based on the residual vector $\boldsymbol{e}_i^{(m)} = \boldsymbol{g}_{k_i^{(m)}} - \boldsymbol{f}_i^{(m)}$ evaluated in the objective function of the

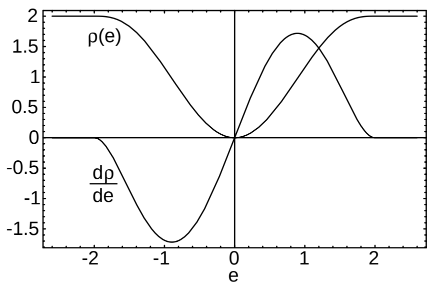

Fig. 2. Biweight function $\rho(e)$. $B=2$.

type of sum of squared residual as

$$J=\sum_{i=1}^{N}\|\boldsymbol{e}_i^{(m)}\|^2. \tag{3}$$

An average shift value $E=1/N\sum_{i=1}^{N}\|\boldsymbol{f}_i^{(m)}-\boldsymbol{f}_i^{(m-1)}\|$ at each pair of successive steps is tested for terminating the iterations when $E\leqslant T_E$ or $m\geqslant T_m$, where T_E and T_m are thresholds given in advance for specifications.

2.2. Introduction of M-estimation

In the proposed ICP modified by M-estimation (hereafter, M-ICP), it is tried to make outlying data invalid in evaluating a new objective function through automatical control of weighting coefficients for residuals [11]. Based on each residual $\boldsymbol{e}_i=(e_{xi},e_{yi},e_{zi})$ for the temporal correspondence (i,k_i) at the mth step, the modified objective function J' is defined as

$$J'=\sum_{i=1}^{N}\{\rho_x(e_{xi})+\rho_y(e_{yi})+\rho_z(e_{zi})\}, \tag{4}$$

where $\rho_x(\cdot)$, $\rho_y(\cdot)$ and $\rho_z(\cdot)$ are the functions, all of which are non-negative even functions having global minima at the origin. They work for making uneven or non-uniform contribution of residuals depending on their magnitudes. For example, $\rho(e)=e^2$ is such an example that yields least square estimates by arithmetic mean, but it is well known that this has a remarkable effect on estimates with large residuals such an outlying noise, degrading precision and stability of the estimates. This is one of the motivation of introducing these functions into the residual evaluation. In (4), we assign the separate functions for individual coordinates so that the sensitivity design is possible for each coordinate, while it is also possible to simply provide functions of distance such as $d=\sqrt{e_x^2+e_y^2+e_z^2}$. The strategy of using these separate functions is much effective in the case of stereoscopic depth reconstruction where errors in the z coordinates along with an optical axis is generally larger than others, and furthermore, more outlying depth may occur in the coordinates. Fig. 2 shows the profile of the bi-weight function [12] defined as

$$\rho(e)=\begin{cases}\frac{B^2}{2}(1-(1-(\frac{e}{B})^2)^3) & (\text{if } |e|\leqslant B),\\ \frac{B^2}{2} & (\text{if } |e|>B).\end{cases} \tag{5}$$

The profile monotonically increases from the origin and then abruptly converges to the constant value $B^2/2$ at B. The figure shows another plot of the profile of $\mathrm{d}\rho/\mathrm{d}e$ which is called *influence function* [11] and represents the sensitivity depending on magnitude of the residuals. It is well represented that there is a band width from the origin to B and larger residuals than B have no influence on estimates.

The modified objective function J' in (4) enables much robust matching or fitting because of the abovementioned characteristics of sensitivity assignment to residuals which are drawn from wrong illegal correspondences in $\mathscr{C}$ at every iteration.

3. Iterative matching algorithm

The proposed matching has two loops as shown in Fig. 1: the miner loop for parameter estimation by M-estimation and the major one for contour data matching. In the miner loop, the objective function J' is minimized based on the temporal correspondence hypothesis $\mathscr{C}$, while the major loop is mainly for shifting and rotating F by the transform $(R,\boldsymbol{t})$ derived from the estimated $\boldsymbol{q}$, and for modifying temporal correspondence.

3.1. Optimization

We perform minimization of the objective function for estimating the parameter vector $\boldsymbol{q}$ in the miner loop by use of the multi-dimensional descending symplex method [13] that has been known to be not so efficient but robust because of no differentiation. In the seven dimensional search space of vectors $\boldsymbol{q}$, a symplex with eight vertices is deformed and moved to a descending direction inferred in terms of distribution of sampled objective function at the vertices. The three motions such as reflection, expansion and contraction are prepared for the symplex to scan over the search space. Provided with an initial parameter vector $\boldsymbol{q}_0^{(0)}$ at the first iteration, an alignment of the symplex vertices is given by $\boldsymbol{q}_0^{(0)}$ and $\{\boldsymbol{q}_i=\boldsymbol{q}_0^{(0)}+\lambda\boldsymbol{e}_i\}_{i=1}^{7}$ $(\lambda>0)$, where $\{\boldsymbol{e}_i\}_{i=1}^{7}$ are the normal basis vectors. Initial parameter vectors $\boldsymbol{q}_0^{(m)}$ $(m=1,2,\ldots)$ at every successive iteration are provided as the solution vectors estimated at the previous step by M-ICP. Every miner loop terminates when the function J' decreases more slightly than specified.

3.2. Initial conditions

We consider how to give the global initial estimates $\boldsymbol{q}_0^{(0)}$, components of which are divided into the translation and rotation parameters. Since either the original ICP algorithm or M-estimation is based on iterative calculation, the initial alignment of F and G affects on their convergence rates and precisions. Registration the geometric centers of both the data [9] solves their initial shift so that $\boldsymbol{t}_0^{(0)} = 1/K \sum \boldsymbol{g} - 1/N \sum \boldsymbol{f}$. In the situation addressed in this paper, the bias in the translation parameters raised by outlying data may degrade their convergence, however, according to the experimental results described afterward, the effect can be assumed to get relatively small.

On the other hand, the bias in the rotation parameters (q_1, q_2, q_3, q_4) due to inconsistent data is likely to mislead their convergence. Such an initialization in numerical analysis has been one of the open problems and there is no general technique to work well [13]. In this study, we are trying to compare the proposed matching with the original ICP matching in order to investigate their robustness for the initialization.

4. Experiments

4.1. Two dimensional matching

Table 1 is the experimental specifications in this study. The two dimensional contour data F and G in Fig. 3(a), which were detected from brightness images by edge extraction, were used for verification of the fundamental performance of M-ICP. We can find large outlying portions in both the data. Fig. 3(b) are the matching results by ICP and M-ICP. For those portions of outlying data, the iterations in ICP converged to an incorrect position with a local minimum after 29 steps, while M-ICP could bring the solution to a satisfactory position even if it took even 95 steps until termination. Although rigorous correspondence could not be expected to obtain because of additive noise that injured most of data positions and further outlying data, the proposed method could reach a good solution in the search space. The outlying data in F were detected as shown in Fig. 3(c). The transitions of the objective functions J and J' and the accumulated movement in pixel for both the methods were obtained as shown in Fig. 4. Even though it was not possible to directly compare the two objective functions since their definitions are different, we observed their

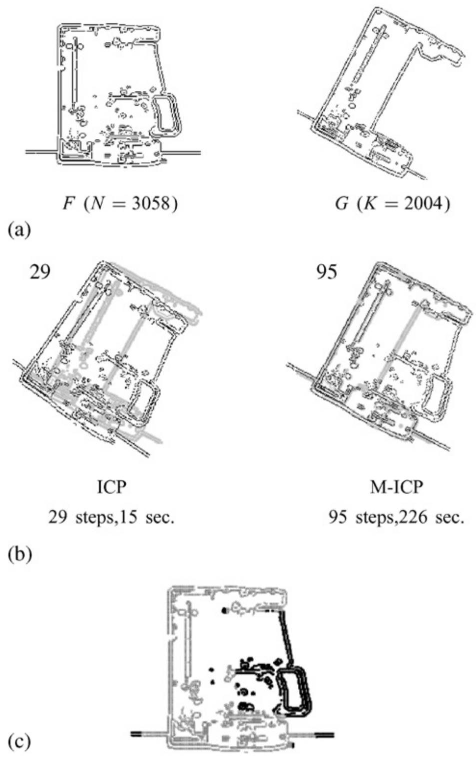

Fig. 3. Matching two dimensional contours (Coffee maker). (a) Two dimensional contour data. (b) Matching solutions (CPU: Pentium III, 0.7 GHz, 256 MB). (c) Outlying points in F (black).

descending trends in the figures. In ICP, J was reduced rapidly at the beginning and then smoothly to the constant level, while J' in M-ICP decreased in a slow pace with small ripples, but in rather abrupt change after around 70 steps, which gave suggestion that it begun to assign small weights to the outlying data. This observation was supported by change in the amount of corresponding data with non-zero contribution as shown in Fig. 4(c).

Table 1
Experimental specifications

	2D	3D
B_x, B_y	4.5	5.0
B_z	—	7.0
T_E	0.05	0.08
T_m	300	300

4.2. Three dimensional matching

The three dimensional contours in Fig. 5(a) were used matched and the results were obtained as shown in Fig. 5(b). M-ICP could reach a successful state of matching, while ICP yielded a clear erroneous solution. For comparison of both the convergence of ICP and M-ICP, their final positions were shown in Fig. 5(b) in different gray levels. A different kind of transition characteristics in the amount of corresponding points in M-ICP was shown in Fig. 5(c), which showed that enclosure of correct matching proceeded gradually to the convergence. Fig. 6 was another example of three dimensional matching. The proposed M-ICP could find a good solution after 300 steps as shown in this figure,

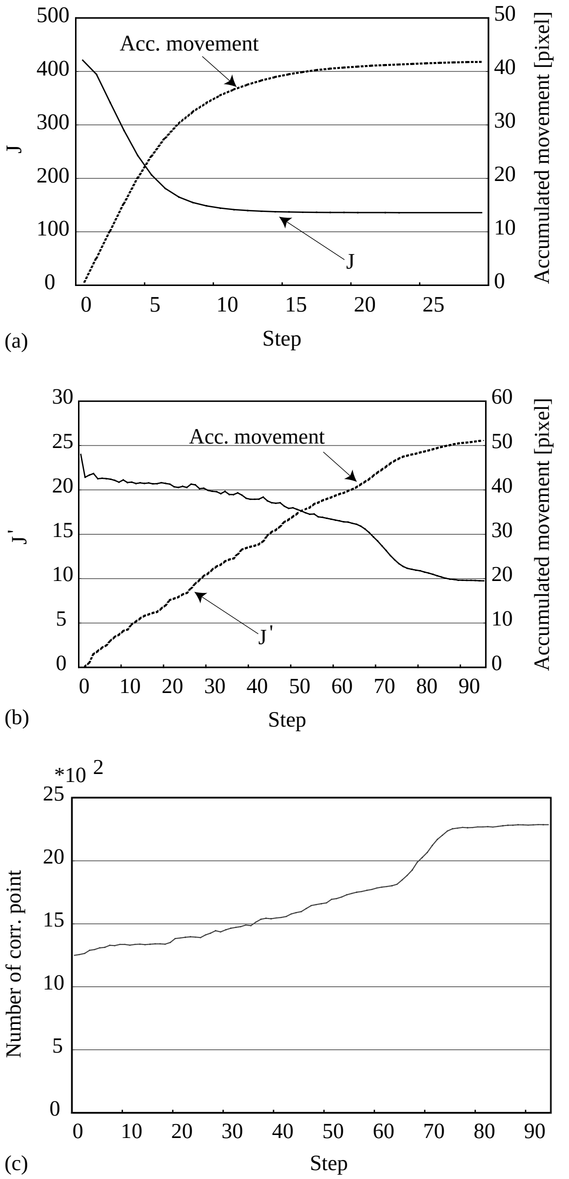

Fig. 4. Comparison of transition (Coffee maker). (a) ICP, (b) M-ICP and (c) M-ICP (corresponding points).

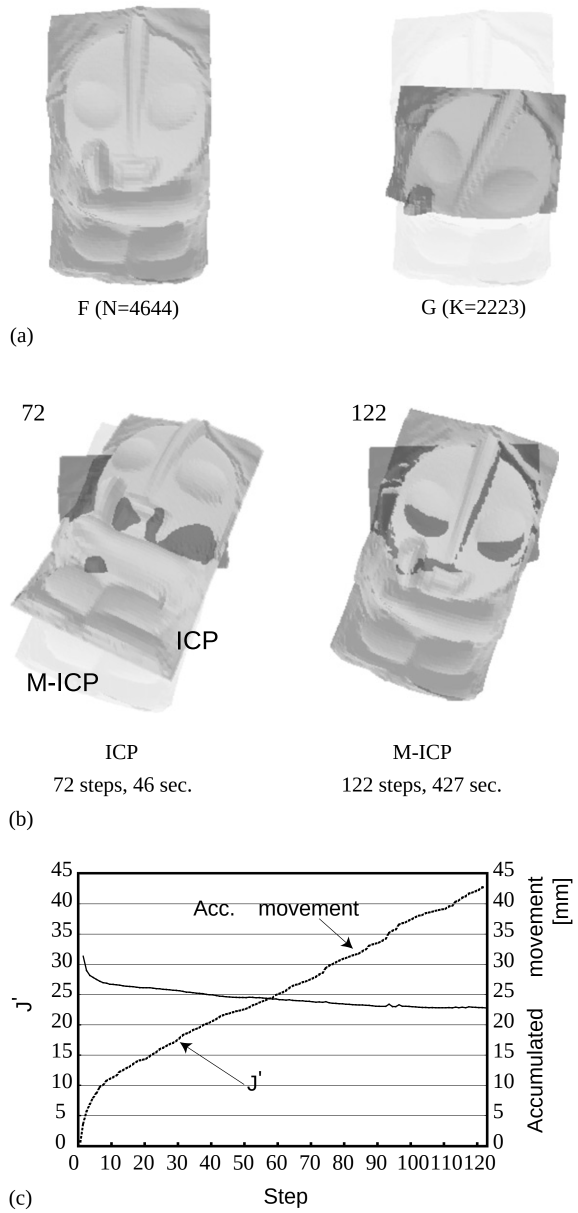

Fig. 5. Matching three dimensional contours (Toy). (a) Data set of 3D contour. (b) Results of matching. (c) Transition of corresponding points (M-ICP).

where in (b) the frontal and side views of the matching were shown.

4.3. Discussion

We have found two different kinds of experimental convergence of M-ICP in the previous experiments, however, it seems difficult to precisely determine when we can terminate by investigating how fast or well the iterations go into convergence.

Initialization of relative alignment of F and G, especially with respect to for rotation parameters, is one of remaining problems in M-ICP also, but we have firmly obtained expectation of robustness in M-ICP to the initialization and the inclusion of outlying data and then verified the effectiveness through the real data. Furthermore, we may have the real world problems where the initialization can be handled rea-

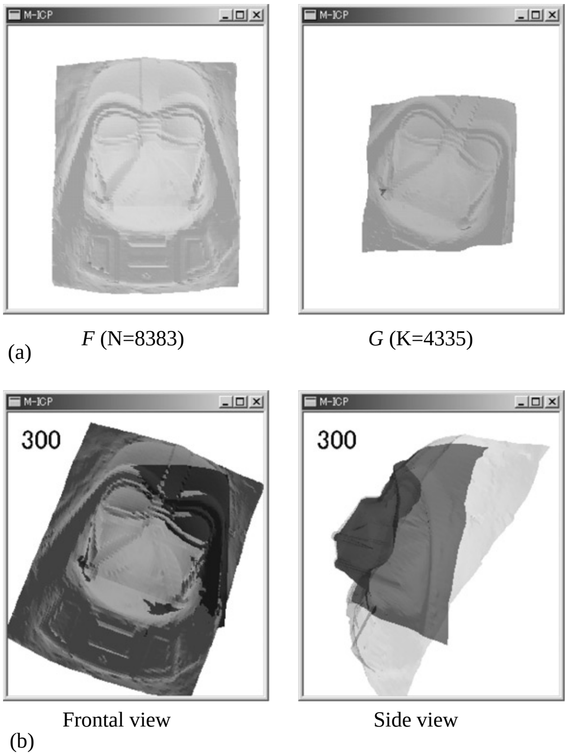

Fig. 6. Matching three dimensional contours by M-ICP (Mask). (a) Data set of 3D contour. (b) Results of matching.

sonably by use of some knowledge and/or know-how about the problems.

On how to set up parameters in M-ICP, essentially for the switching width B in respective coordinate, we may be enable to utilize error distributions, such as additive probabilistic noise and quantization error, in order to determine the concrete values. It is well utilized for this subject to set $B = 2.5\sigma$ for discriminating outlying noise from any standard measurement error [11].

5. Conclusion

As a novel method for matching of three dimensional contours with non-overlapping and/or outlying data, characteristics of which are essentially relative, we proposed the improved Iterative Closest Point algorithm that utilized M-estimation for providing some regular priority to data points through on-line evaluating the modified objective function. The effectiveness of the proposed method were shown by use of the real contour data with ill-conditions in comparison with the original ICP method.

6. Summary

An extension of the Iterative Closest Point matching by M-estimation is proposed for realization of robustness to non-overlapping data or outlying data in two sets of contour data or depth images for rigid bodies. An objective function which includes independent residual components for each of x, y and z coordinates is originally defined and proposed to evaluate the fitness, simultaneously dealing with a distribution of outlying gross noise. The proposed procedure is based on modified M-estimation iterations with bi-weighting coefficients for selecting corresponding points for optimization of estimating the transforms for matching. The transforms can be represented by 'quaternions' in the procedure to eliminate redundancy in representation of rotational degree of freedom by linear matrices. Optimization steps are performed by the simplex method because it does not need computation of differentiation. Some fundamental experiments utilizing real data of 2D and 3D measurement show effectiveness of the proposed method. When reasonable initial positions are given, the unique solution of position could be provided in spite of surplus point data in the objects. And then the outlying data could be filtered out from the normal ones by the proposed method.

References

[1] F. Stein, G. Medioni, Structural indexing: efficient 3-D object recognition, IEEE Trans. PAMI 14 (2) (1992) 125–145.
[2] N. Ikawa, et al., Rapid Product Development, Chapman & Hall, London, 1997.
[3] K. Ikeuchi, Y. Sato, Modeling From Reality, Kluwer Academic Publishers, Norwell, 2001.
[4] N. Kehtarnavaz, S. Mohan, A framework for estimation of motion parameters from range images, CVGIP 45 (1989) 88–105.
[5] A. Johnson, S. Kang, Registration and integration of textured 3-D data, Proceedings of 3DIM '97, 1997, pp. 234–241.
[6] Z. Zhang, Iterative point matching for registration of free-form curves and surfaces, IJCV 13 (2) (1994) 119–152.
[7] T. Masuda, N. Yokoya, A robust method for registration and segmentation of multiple range images, CVIU 61 (3) (1995) 295–307.
[8] P.J. Rousseeuw, A.M. Leroy, Robust Regression and Outlier Detection, Wiley, New York, 1987.
[9] P.J. Besl, N.D. McKay, A Method for Registration of 3-D Shapes, IEEE Trans. PAMI 14 (2) (1992) 239–256.
[10] S. Kaneko, T. Kondo, A. Miyamoto, Robust matching of 3D contours using iterative closest point algorithm improved by M-estimation, Proceedings of the third China-Japan Symposium on Mechatronics, 2002, pp. 37–42.
[11] F.R. Hampel, et al., Robust Statistics—The Approach based on Influence Functions, Wiley, New York, 1986.
[12] P.W. Holland, R.E. Welsch, Robust regression using iteratively reweighted least-squares, Commun. Statist. Theor. Meth. A 6 (9) (1977) 813–827.
[13] R. Fletcher, Practical Methods of Optimization, Wiley, New York, 1997.

About the Author—SHUN'ICHI KANEKO received the B.S. degree in precision engineering and the M.S. degree in information engineering from Hokkaido University, Japan, in 1978 and 1980, respectively, and then the Ph.D. degree in Systems Engineering from the University of Tokyo, Japan, in 1990. He had been a research assistant of the Department of Computer Science since 1980 to 1991, an associate professor of the Department of Electronic Engineering since 1991 to 1995, and an associate professor of the Department of Bio-application and Systems Engineering since 1996 to 1996, in Tokyo University of Agriculture and Technology, Japan. He is an associate professor of the Department of Control and Information Engineering in Hokkaido University from 1996. He received the Best Paper Award in 1990, the Society Award in 1998, respectively, from Japan Society of Precision Engineering. His research interest includes machine vision, image sensing and understanding, robust image registration. He is a member of IEICE, JSPE, IEEJ, SICE and the IEEE computer society.

About the Author—TOMONORI KONDO received the B.S. and M.S. degrees in Systems Engineering from Hokkaido University, Japan, in 1999 and 2001, respectively. His research topics included motion analysis by computer vision and robust matching of depth data. He joined Oki Data Co., Ltd. in 2002.

About the Author—ATSUSHI MIYAMOTO received the B.S., M.S. and Ph.D. degrees in Precision Engineering from Hokkaido University, Japan, in 1996, 1998, and 2001, respectively. He received the Society Award from Japan Society of Precision Engineering in 1998. His research interest includes computer vision and visual measurement and recognition of three dimensional scene. He is a member of JSPE.