2006 IEEE International Conference on Fuzzy Systems
Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada
July 16-21, 2006

3D Shape Modeling Employing Fuzzy Clustering for Stereo Vision

Hideaki Kawano, Atsumori Sasaki, Hiroshi Maeda, and Norikazu Ikoma,
Faculty of Engineering, Kyushu Institute of Technology, Japan

***Abstract*—In this paper, a novel fuzzy-clustering-based method for 3D shape modeling is proposed. This method is intended for scenes involving multiple objects, where each object is replaced by a primitive model. The proposed method is composed of three stages. In the first stage, 3D data is reconstructed using stereo matching technique from a stereo image taking multiple objects. Next, the 3D data is divided into a single object by employing a Fuzzy c-Means augmented with Principal Component Analysis (PCA) and a criterion about the number of clusters. Finally, the shape of each object is extracted by Fuzzy c-Varieties with Noise Clustering.**

I. INTRODUCTION

Techniques for acquisition of three-dimensional (3D) information can be roughly classified into two categories: active and passive methods. The former is usually accomplished by range finders. This method can measure the exact distance. However it has a limit of object and a lack of portability, and the instrument is expensive. The latter is generally based on stereo vision. Compared with active methods, the accuracy of stereo vision is lower. Those methods, however, have a high versatility, because specific objects are not targeted. In addition, 3D information can be obtained without affecting the scene, and the instruments have comparatively lower costs. For these reasons, the passive methods are suited for applications of robot vision.

Most of past studies of shape modeling have focused on processing data obtained from active sensors [1]-[7]. This fact is caused by the following difficulties: 3D information obtained from stereo vision contains a lot of outliers due to mismatching, and usually sparse. To achieve shape modeling using stereo vision [8], those difficulties should be overcome.

In this paper, a novel method of 3D shape modeling is proposed in the framework of fuzzy clustering by considering 3D noisy data obtained from stereo vision as fuzzy data in order to deal with the difficulties mentioned above. This method is intended for scenes involving multiple objects, where each object is represented by a primitive model. The proposed method is composed of three stages. In the first stage, 3D data is reconstructed using stereo matching from a stereo image of a scene composed of multiple objects with primitive shapes (box, cylinder, cone and sphere). Next, the 3D data is separated into objects by using a Fuzzy c-Means algorithm augmented with PCA and a criterion about the number of clusters based on variance ratio. Finally, the shape of each object is extracted by Fuzzy c-Varieties with Noise Clustering. Figure 1 shows the process flow of the proposed shape modeling method. In Fig. 1, the related techniques to each process are also summarized.

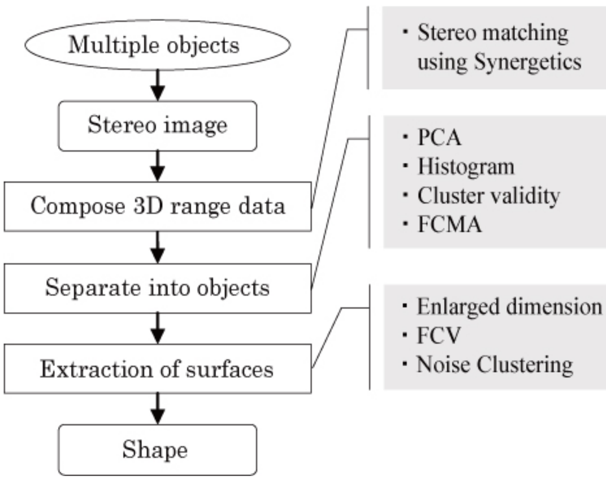


Fig. 1. Overview of shape recognition system

II. ESTIMATION OF THE NUMBER OF OBJECTS AND EXTRACTION OF EACH OBJECT

This section discusses the technique that the 3D data is separated into objects by using a Fuzzy c-Means algorithm augmented with a criterion about the number of clusters.

A. Principal Component Analysis

To separate into each object, 3D range data from stereo images are projected on x - z plain and principal component is obtained using PCA. Then, histogram for the first principal component is made. Finally, the rough sketch of range data can be found, initial cluster centers in FCM is calculated. Therefore, vulnerable clustering result which is caused by random initial cluster center and the convergence to local optimal solution are alleviated. So, cluster validity to decide the number of clusters automatically can be used. Next, PCA is explained briefly.

Let $\boldsymbol{T}_{\mathrm{X}} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_l\}$ be a set of training vectors from the n-dimensional input space $\mathbf{R}^n$. The set of vectors $\boldsymbol{T}_{\mathrm{Z}} = \{z_1, \ldots, z_l\}$ is a lower dimensional representation of the input training vectors $\boldsymbol{T}_{\mathrm{X}}$ in the m-dimensional space $\mathbf{R}^m$. The vectors $\boldsymbol{T}_{\mathrm{Z}}$ are obtained by the linear orthonormal projection.

$$z = \mathrm{W}^T \boldsymbol{x} + \boldsymbol{b} \tag{1}$$

where the matrix **W** [n × m] and the vector $\boldsymbol{b}$ [m × 1] are parameters of the projection. The reconstructed vectors $\boldsymbol{T}_{\widetilde{X}} = \{\widetilde{\boldsymbol{x}}_1, \dots, \widetilde{\boldsymbol{x}}_l\}$ are computed by the linear back projection

$$\widetilde{\boldsymbol{x}} = \mathbf{W}(\boldsymbol{z} - \boldsymbol{b}) \tag{2}$$

obtained by inverting (1). The mean square reconstruction error

$$\varepsilon_{MS}(\mathbf{W}, \boldsymbol{b}) = \frac{1}{l} \sum_{i=1}^{l} \left\| x_i - \widetilde{x}_i \right\|^2 \tag{3}$$

is a function of the parameters of the linear projections (1) and (2). The PCA is the linear orthonormal projection (1) which allows for the minimal mean square reconstruction error (3) of the training data $\boldsymbol{T}_X$. The parameters (**W**, $\boldsymbol{b}$) of the linear projection are the solution of the optimization task

$$(\mathbf{W}, \boldsymbol{b}) = \arg\min_{\mathbf{W}', \boldsymbol{b}'} \varepsilon_{MS}(\mathbf{W}', \boldsymbol{b}') \tag{4}$$

subject to

$$\left\langle \boldsymbol{w}_i \cdot \boldsymbol{w}_j \right\rangle = \delta(i, j), \qquad \forall i, j$$

where $\boldsymbol{w}_i$, $i = 1, \dots, m$ are column vectors of the matrix **W** = [$\boldsymbol{w}_1, \dots, \boldsymbol{w}_m$] and $\delta(i, j)$ is the Kronecker delta function. The solution of the task (4) is the matrix **W** = [$\boldsymbol{w}_1, \dots, \boldsymbol{w}_m$] containing the m eigenvectors of the sample covariance matrix which have the largest eigen values. The vector $\boldsymbol{b}$ equals to $\mathbf{W}^T\boldsymbol{\mu}$, where μ is the sample mean of the training data.

B. Cluster Validity

In Fuzzy *c*-Means algorithm [10], when the number of clusters is automatically decided, another measure is generally needed with the technique of making to the cluster. The validity which quantitatively decides the number of cluster can get if the measure is used. Here, the cluster is seen from the aggregative aspect. When the data is divided into k clusters that are the rebellion and not empty, let S_r: Sum of Squares Total, S_W: Sum of Squares Within, S_B: Sum of Squares Between. It is well-known the following relational expression.

$$S_r = S_W + S_B \tag{5}$$

Eq.(5) is the measure of the aggregative. However, let k variable and the aggregative is calculated. So, by using Eq.(5), the validity based on subgroup both sum of square and determinant are considered.

C. Cluster Validity by Calinski and Harabasz (C-H criterion)

Cluster Validity based on variance ratio proposed by Calinski and Harabasz is called VRC (Variance Ratio Criterion) [9],

$$VRC(k) = \frac{tr(S_B)}{k-1} \Big/ \frac{tr(S_W)}{n-k} \quad , \tag{6}$$

where $tr(A)$ is the trace of matrix A. When the VRC value takes the maximum value or it increases rapidly compared with the change from $(k-1)$ to k with the number of clusters, the value of the k is chosen as measure of the number of clusters. However, if the VRC value includes some maximum values, the minimum value of the value of k at that time is assumed to be the number of clusters.

D. Fuzzy c-means with a Variable for cluster Sizes

In standard fuzzy c-means, all clusters are classified with equal sizes. In this study, 3D range data of each object have each different size. When these kinds of data are classified by FCM, the bigger cluster is absorbed by the smaller one. To solve this problem, fuzzy c-means with a variable for cluster sizes [12] is used. The following is this algorithm.

Let $X = \{\boldsymbol{x}_k \mid k = 1 \cdots N\}$ be a finite subset of a $\boldsymbol{p}$ dimensional vector space over the reals Here, $\boldsymbol{x}_k$ denotes the set of feature vectors. Let c denote the number of clusters. And $c \times n$ matrix $U = \boldsymbol{u}_{ik}$ denotes the grade of membership of $\boldsymbol{x}_k$ in the i th fuzzy subset of X. And let $\boldsymbol{v}_i = (v_i^1, \dots, v_i^p)$ be the vector collecting all cluster centers, and $V = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_c)$. Following the idea that cluster which has the less data has the smaller field, size variables α for clusters are applied. The constrain of α is

$$A = \{\alpha : \sum_{i=1}^{c} \alpha_i = 1, \alpha_i \geq 0, i = 1, \dots, c\} \tag{7}$$

D_{ik} is assumed to be the square of Euclidean distance between the individual $\boldsymbol{x}_k$ and the center $\boldsymbol{v}_i$ of the cluster i. Objective function is:

$$J(U, V, \alpha) = \sum_{i=1}^{c} \alpha_i \sum_{k=1}^{n} \left(\frac{u_{ik}}{\alpha_i} \right)^m D_{ik} \tag{8}$$

Algorithm : Fuzzy c-Means with a Variable for cluster Sizes

Step 1. Initialize $\overline{V}$, $\overline{\alpha}$ properly.

Step 2. Calculate $\min_{U \in M} J(U, \overline{V}, \overline{\alpha})$ using Eq. (9) and update $\overline{U}$.

$$u_{ik} = \left\{ \sum_{j=1}^{c} \left(\frac{\alpha_j}{\alpha_i} \right) \left(\frac{D_{ik}}{D_{jk}} \right)^{\frac{1}{m-1}} \right\}^{-1} \tag{9}$$

Step 3. Calculate $\min_{U \in M} J(\overline{U}, V, \overline{\alpha})$ using Eq. (10) and update $\overline{V}$.

$$\boldsymbol{v}_i = \frac{\sum_{k=1}^{n} (u_{ik})^m \boldsymbol{x}_k}{\sum_{k=1}^{n} (u_{ik})^m} \tag{10}$$

Step 4. Calculate $\min_{U \in M} J(\overline{U}, \overline{V}, \alpha)$ using Eq. (11) and update $\overline{\alpha}$.

$$\alpha_i = \left\{ \sum_{j=1}^{c} \left(\frac{\sum_{k=1}^{n} (u_{jk})^m D_{jk}}{\sum_{k=1}^{n} (u_{ik})^m D_{ik}} \right)^{\frac{1}{m}} \right\}^{-1} \tag{11}$$

Step 5. If $\left\| \hat{U} - \overline{U} \right\| < \varepsilon$, then stop. Otherwise, return to Step 2.

III. Surface Extraction and Shape Recognition

In this section, first, Fuzzy c-Varieties for extraction of surfaces are presented. Next, a principle of 'Noise Clustering' is introduced such that noisy data points may be assigned to the noise class. Finally, FCV algorithm in the enlarged data space is presented.

A. Fuzzy c-Varieties Clustering

Fuzzy *c*-Varieties (FCV) [10, 11, 13] partitions data using linear varieties as the prototypes of the clusters, local principal component vectors as the basis of the prototypical linear varieties can be also extracted. The linear variety (or manifold) of dimension q ($\geq$1) through the point $\boldsymbol{w}_i \in \mathrm{R}^p$ spanned by the linearly independent vectors $\{\boldsymbol{s}_{i1}, \ldots, \boldsymbol{s}_{iq}\} \subset \mathrm{R}^p$ is

$$\boldsymbol{y} = \boldsymbol{w}_i + \sum_{l=1}^{q} \beta_l \boldsymbol{s}_{il} \tag{12}$$

When the $\{\boldsymbol{s}_{il}\}$ are an orthonormal basis for their span, the orthogonal projection theorem yields

$$D_{ik} = \left\| \boldsymbol{x}_k - \boldsymbol{w}_i \right\|^2 - \sum_{l=1}^{q} < \boldsymbol{x}_k - \boldsymbol{w}_i, \boldsymbol{s}_{il} >^2 . \tag{13}$$

Then, objective function is

$$J_{fcv}(U, \boldsymbol{w}, \boldsymbol{s}) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m D_{ik}(\boldsymbol{x}_k, \boldsymbol{w}_i, \boldsymbol{s}_i), \tag{14}$$

where $m \in (1, \infty)$ is weighting exponent. $U = (u_{ik})$, $\boldsymbol{w} = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_c)$, $\boldsymbol{s} = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_c)$ are the parameters of this objective function. So, iterative algorithm needs the following steps:

Step 1. Initialize $\overline{U}$ randomly.

Step 2. Calculate $\min_{w} J_{fcv}(\overline{U}, \boldsymbol{w}, \boldsymbol{s})$ using Eq. (15) and update $\overline{\boldsymbol{w}}(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_c)$.

$$\boldsymbol{w}_i = \frac{\sum_{k=1}^{n} (\overline{u}_{ik})^m \boldsymbol{x}_k}{\sum_{k=1}^{n} (\overline{u}_{ik})^m} \tag{15}$$

Step 3. Calculate $\min_{s} J_{fcv}(\overline{U}, \overline{\boldsymbol{w}}, \boldsymbol{s})$ using Eq. (16) and get q eigen vectors with normalization $\overline{\boldsymbol{s}}(\boldsymbol{s}_{i1}, \ldots, \boldsymbol{s}_{iq})$ from eigen value of matrix $A_{i,}$.

$$A_i = \sum_{l=1}^{n} (\overline{u}_{il})^m (\boldsymbol{x}_l - \overline{\boldsymbol{w}}_i)(\boldsymbol{x}_l - \overline{\boldsymbol{w}}_i)^T \tag{16}$$

Step 4. Calculate $\min_{U \in M_f} J_{fcv}(U, \overline{\boldsymbol{w}}, \overline{\boldsymbol{s}})$ using Eq. (17) and update $\overline{U}$.

$$u_{ik} = \left[\sum_{j=1}^{c} \left(\frac{D_{ik}}{D_{ik}} \right)^{\frac{1}{m-1}} \right]^{-1} \tag{17}$$

Step 5. If $\left\| \hat{U} - \overline{U} \right\| < \varepsilon$, then stop. Otherwise, return to Step 2.

B. Noise Clustering

In [12], Davé independently proposed the idea of a noise cluster to deal with noisy data. In this approach referred to as the NC approach, the noise is considered to be a separate class and is represented by a prototype that has a constant distance, δ, from all data points. The membership

Noise prototype is a universal entity such that it is always at the same distance from every point in the data-set. Let $\boldsymbol{v}_c$ be the noise prototype, and $\boldsymbol{x}_k$ be the point in feature space, $\boldsymbol{v}_c, \boldsymbol{x}_k \in \mathrm{R}^p$. Then the noise prototype is such that the distance $d_{ik} = \delta$, distance of point $\boldsymbol{x}_k$ from $\boldsymbol{v}_c$, is

$$d_{ik} = \delta, \forall k \tag{18}$$

Eq.(9) doesn't define what the distance δ is. It simply says that all the points are at the same distance δ from $\boldsymbol{v}_c$. Physically, this means that all the points have equal *a priori* probability of belonging to a noise cluster. This makes sense since given no prior information all the points should have an equal probability of falling into a noise class. It is hoped, however, that as the algorithm progresses, the *good* points increase their probability of being classified into a *good* cluster.

Let there be $c-1$ good clusters in the data-set, and let the *c*-th cluster be the noise cluster. Then the functional J_{nfcv} including the noise cluster is defined:

$$\begin{aligned} J_{nfc} &= \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m D_{ik} \\ &= \sum_{k=1}^{n} \sum_{i=1}^{c-1} (u_{ik})^m D_{ik} + \sum_{k=1}^{n} \left(1 - \sum_{i=1}^{c} u_{ik} \right)^m \delta^2 \end{aligned} \tag{19}$$

The membership $\boldsymbol{u}_{*k}$ of a point $\boldsymbol{x}_k$ in the noise cluster is defined to be

$$u_{*k} = 1 - \sum_{i=1}^{c} u_{ik} \tag{20}$$

Thus, the membership constraint for the good cluster is effectively relaxed to

$$\sum_{i=1}^{c-1}(u_{ik}) = 1 - u_{ck} \quad or \quad 0 \le \sum_{i=1}^{c-1} u_{ik} \le 1 \tag{21}$$

In this paper, noise clustering is applied to Fuzzy c-Varieties. Algorithm is similar to FCV.

C. *Enlarged dimension*

FCV is the clustering technique that can find linear singularities. FCV is the clustering technique that can find linear singularities. It is available for only a linear space. To extract non-linear structure from 3D information that are targeted in this study, non-linear data have to be regarded as linear data in form. The enlarged data space is the $x^2 - y^2 - z^2 - x - y - z$ coordinate system that adds x^2, y^2, z^2 to the $x - y - z$ coordinate system. The equation of the upright cylinder is given (14). So, the data must be enlarged to 4D.

$$(x-a)^2 + (z-b)^2 = r^2 \tag{22}$$

Similarly, the equation of the upright cone and sphere is given (23). So, the data must be enlarged to 6D.

$$(x-a)^2 + (z-b)^2 = d(y-c)^2 \tag{23}$$

To decide the shape that fits closely, FCV is applied to all dimensions namely 3D, 4D and 6D.In each case, the number of points included in each cluster and the noise cluster are examined. In addition, the dimension that has the most numbers of points included in the cluster is decided. Finally, its prototype is extracted as shape of the object.

IV. EXPERIMENTAL RESULTS

In this section, experimental results are shown. At first, multiple objects are taken by stereo camera system as shown in Fig. 2. Figure 3 shows a couple of images obtained by the stereo camera system. By using stereo matching technique, 3D information about the objects is acquired as shown in Fig. 4.

Next, the 3D information is projected onto x-z plane and PCA is conducted. Figure 5 is the result of PCA. A histogram along with the first principal component axis is constructed from data distribution. According to the shape of the histogram, initial points in FCM stage are determined. The initial points are arranged to centers of some agglomerates discovered in the histogram. Each VRC is calculated by FCM with C-H criterion changing the number of cluster centers from two to five. The result is summarized in Fig. 7. Based on the rule previously mentioned in the section 2, the number of clusters, i.e. the number of objects, is determined as two. And then, FCM with variables for cluster-sizes is applied. The result is shown in Fig. 8.

As shown in Fig. 8, sizes of two clusters are $\alpha_1 = 0.7501$, $\alpha_2 = 0.2499$ respectively. This result shows the applicability for the case with different cluster sizes. Then, Fuzzy c-Varieties with noise clustering is applied to each object in the cases of 3D, 4D, and 6D. From Fig. 9, 10 it can be said that shape modeling is achieved by the proposed method.

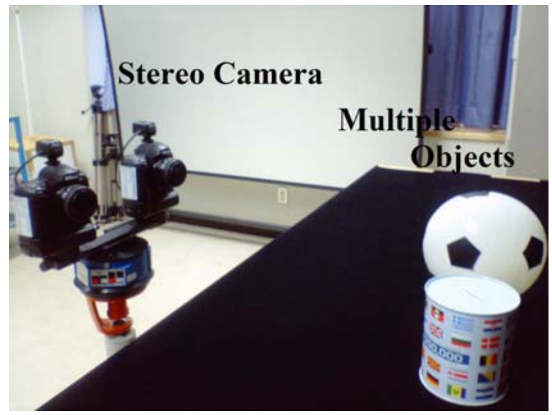


Fig. 2. Experimental setup with stereo camera and multiple objects

(a) Left image (b) Right image

Fig. 3. Binocular stereo images

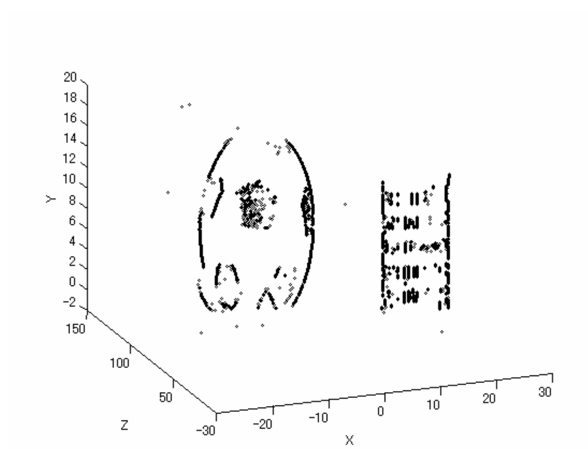


Fig. 4. 3D information obtained by stereo matching

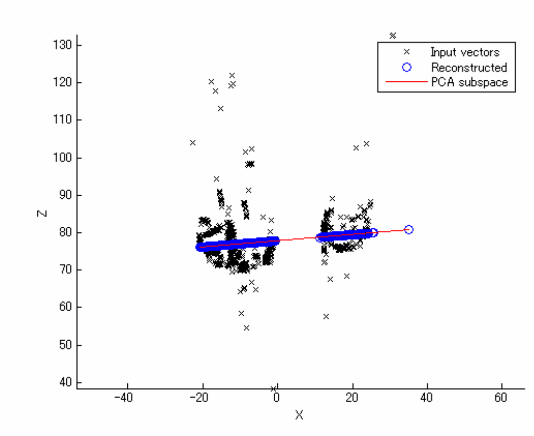


Fig. 5. PCA on *X-Z* plane

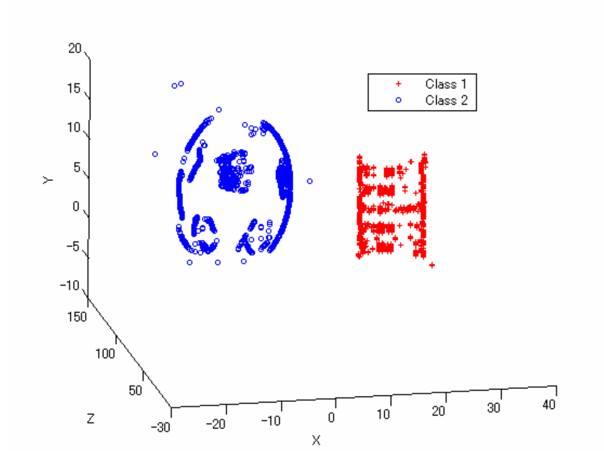


Fig. 8. Result of FCM algorithm on *X-Z* plane
(α_1= 0.7501, α_2= 0.2499)

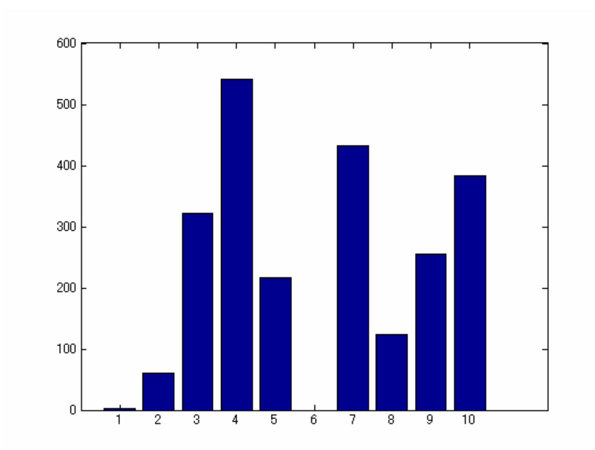
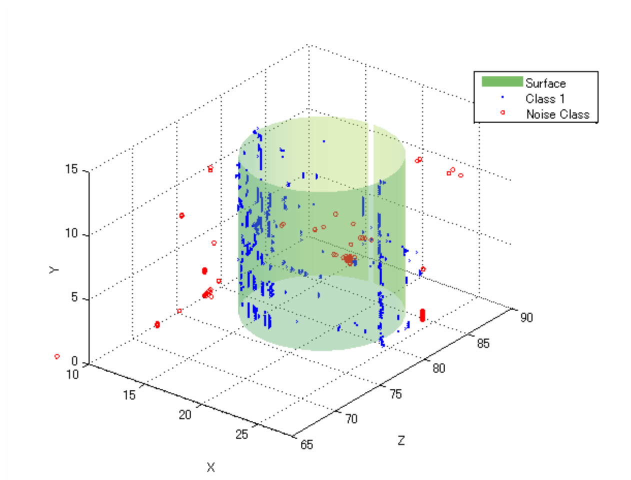


Fig. 6. Histogram on PCA



(a) Cylinder: The number of data 1121/1141, delta=0.5

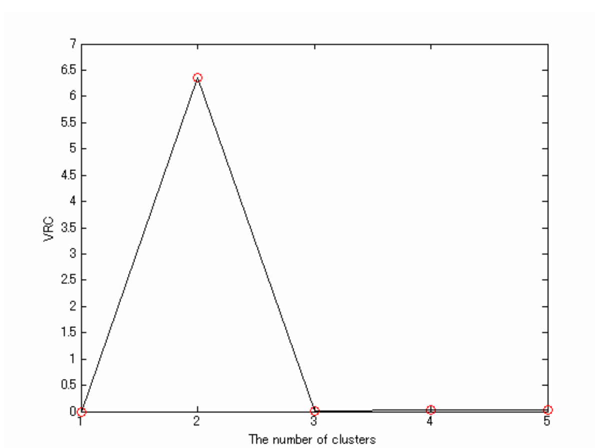
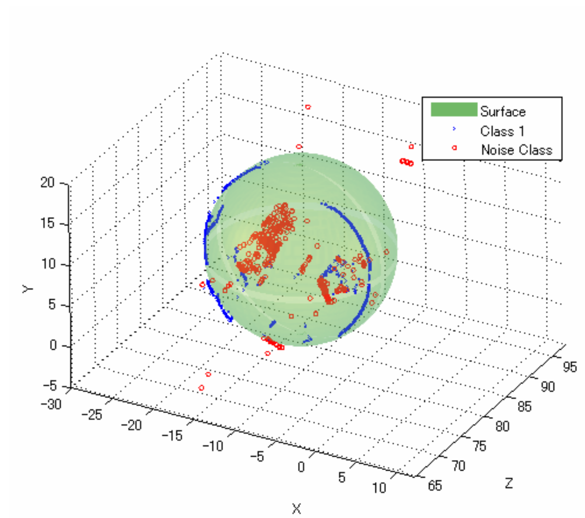


Fig. 7. Variation of VRC



(b) Sphere: The number of data 1093/1195, delta=0.2

Fig. 9. Result of NFCV algorithm

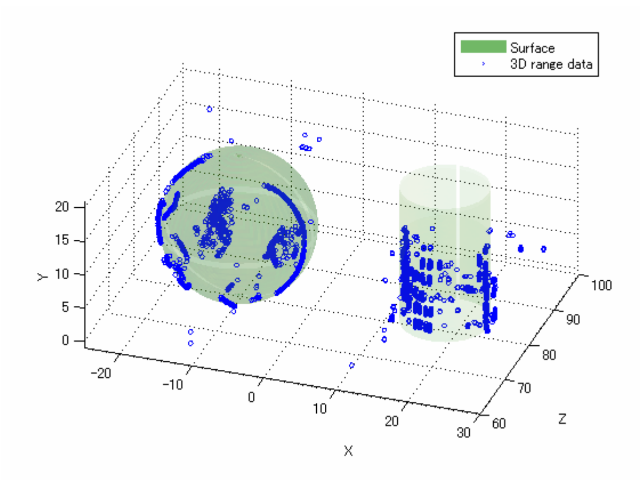


Fig. 10. Result of 3D shape modeling

V. Conclusions

3D range data from a stereo image includes outlier data because of human error by setting camera, technical problem of stereo matching itself, and problems which depends on characteristic of equipments or lighting environment.

In this paper, a new approach based on fuzzy-clustering for object recognition from rough 3D range data is proposed.

Suitable object could be recognized from 3D range data including outlier by fuzzy clustering combined with size function, noise clustering, and enlarged dimension.

It can be said that being able to apply fuzzy clustering that has few utilitarian application to object recognition which is a popular field of study is interesting study.

A further direction of this study will be to enable to recognize more complicated shapes by increasing the kind of the prototype.

Acknowledgment

This work was supported by the Japan Society for the Promotion of Science under the Grant-in-Aid for Scientific Research (C) (No. 16500133).

References

[1] G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.

[2] John WILLIAMS, Mohammed BENNAMOUN: A Multiple View 3D Registration Algorithm with Statistical Error Modeling, IEICE TRANS. INF. & SYST., VOL.E83D, NO.8, pp.1662-1670.

[3] Caihua WANG, Hideki TANAHASHI, Hidekazu HIRAYU, Yoshinori NIWA, Kazuhiko YAMAMOTO: Polyhedral Description of Panoramic Range Data by Stable Plane Extraction, IEICE TRANS. INF. & SYST., Vol.E85-D No.9, pp.1399-1408, 2002.

[4] Gerhard K. Kraetzschmar, Stefan Enderle: Self-localization using sporadic features, Robotics and Autonomous Systems, Volume 40, Issues 2-3, Pages 111-119, 2002.

[5] Dirk Hähnel, Wolfram Burgard, and Sebastian Thrun: Learning compact3D models of indoor and outdoor environments with a mobile robot, Robotics and Autonomous Systems, Volume 44, Pages 15-27, 2003.

[6] Y. Y. Cha, D. G. Gweon: A calibration and range-data extraction algorithm for an omnidirectional laser range finder of a free-ranging mobile robot, Mechatronics, Volume 6, Pages 665-689, 1996.

[7] Adnan A. Y. Mustafa, Linda G. Shapiro and MarkA. Ganter: 3D object identification with color and curvature signatures, Pattern Recognition, Volume 32, Pages 339-355, 1999.

[8] Harrie van Dijck, Ferdinand van der Heijden: Object recognition with stereo vision and geometric hashing, Pattern Recognition Letters, Volume 24, Pages 137-146, 2003.

[9] T. Irie, H. Katou, H. Maeda: Object Recognition from Stereo Images Using Synergetics and Combinatorial Optimization, Proc. of SICE Annual Conference 2003, in CD-ROM, pp. 1221-1226, 2003.

[10] T. Calinski and Harabasz: A Dendrite Method for Cluster Analysis, Communication in Statistics, Vol.3, No.1, pp.1-27, 1974.

[11] J.C.Bezdek: Pattern Recognition with Fuzzy Objective Function Algorithm, Plenum, New York, 1981.

[12] S. Miyamoto, Introduction to Cluster Analysis: Theory and Applications of Fuzzy Clustering, Morikita-Shuppan, Tokyo (1999).

[13] R.N.Davé: Characterization and Detection of Noise in Clustering, Pattern Recognition etters, vol.12, pp.657-664, 1991.