3-D Object Recognition using Adaptive Scale MEGI

Hiroshi Matsuo Junichiro Funabashi Akira Iwata

Dept. of Electrical and Computer Eng.
Nagoya Institute of Technology
Gokiso, Nagoya, 466, JAPAN
matsuo@elcom.nitech.ac.jp

Abstract

We propose a method for recognition of a 3-D object using multi scale description of the object and adaptive matching. MEGI model is a description model to represent arbitrary shapes. However, many MEGI elements are necessary to represent uneven or curved surfaces with accuracy, so it is difficult to use them for recognition. As a solution, we make a tree which corresponds multi scale description of the object. While tracing the tree from the root which corresponds the coarsest representation to leafs, a matching algorithm presented in this paper assigns a different scale to each part of the object adaptively and estimates the matching score effectively.

1 Introduction

In order to recognize an object and to determine its attitude in space, it is necessary to have a method to represent the shape of the object. The extended Gaussian image description model (EGI) makes it easy to determine the attitude of a moving object in space [1]. It is independent of the position of the object (shift invariant). EGI provides a unique description for a convex object, though precise information is limited. That is, no two convex objects have the same EGI.

In recognizing an object, not only shift-invariant features but also scale- and pose-invariant features are needed because we must recognize an object from various angles and distances from the observing position to the object. However, EGI has shift-invariant features, not scale- and pose-invariant features.

To overcome this problem, many algorithms have been proposed. Tanaka et al.[3] proposed a covariant data format for describing 3-D objects which uses spherical harmonic function expansion. By taking the norm of the covariant vector, form information on the object is extracted, and pose information is represented as the transformation equations.

Okada et al. [4] developed a new algorithm. Spherical correlation defined by Fisher and Lee[5], which is a rotational variant, is used for recognition after expressing 3-D data using EGI. However, because only normal vectors of 3-D object surfaces are utilized for description, it is impossible to distinguish between convex and concave shapes. Hence, EGI does not have the capability to describe a 3-D object.

We proposed a new 3-D object description called MEGI (more extended Gaussian image), which has the capability to distinguish not only convex objects, but also concave objects. This description model is a shift-invariant model which consists of a set of position vectors of surfaces originating from the center of a 3-D object and normal vectors of the surfaces[8]. We also proposed a matching scale function called extended spherical correlation, which is rotational and scale invariant, together with a calculation algorithm for this scale function.

But if the object which contains curved surfaces will be described using MEGI, curved part should be described using a set of small plane. In such cases, it is possible to describe it using many small plane, but the matching procedure is difficult and need a lot of computing power by increasing a number of plane.

To solve these problems, we introduced a new multi scale description and adaptive scale matching procedure. Adding these extensions, new 3-D matching procedure named adaptive scale MEGI is proposed in this paper.

In Sect. 2, we will present a definition of the MEGI model. and the extended spherical correlation which is used for 3-D object recognition. Sect. 3 will define the adaptive scale MEGI. In Sect. 4, some experimental results of the proposed method with computer simulated data and human surface data are shown.

2 Object Recognition Using MEGI Model

2.1 MEGI Model

We proposed a new 3-D shape description method called more extended Gaussian image (MEGI), which includes information on surface positions, so it can describe not only convex but also concave objects[8].

The MEGI model consists of a set of position vectors X_i for surfaces originating from an object center and their normal vectors p_i Each length of a normal vector also corresponds with surface area, as in the extended Gaussian image. This model is also shift-invariant since it is expressed by an object-oriented coordinate. Let $\mathbf{M}$ be a vector set which describes vectors of the MEGI element as follows:

$$\begin{aligned} \mathbf{M} &= \{(\mathbf{X}_i, \mathbf{p}_i) | i \in \{0, 1, \cdots, m-1\}\} \quad (1) \\ & \quad \mathbf{X}_i, \mathbf{p}_i \in \mathbf{R}^3. \end{aligned}$$

m indicates number of MEGI elements.

2.2 Extended Spherical Correlation

Let $\mathbf{X}$ and $\mathbf{Y}$ be sets of n-dimensional unit vectors and let the elements be $\mathbf{X} = \{\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{m-1}\}$, $\mathbf{Y} = \{\mathbf{Y}_0, \mathbf{Y}_1, \ldots, \mathbf{Y}_{m-1}\}$ $\quad (\|\mathbf{X}_i\| = \|\mathbf{Y}_i\| = 1)$.

We proposed a new coefficient for recognizing 3-D objects with the MEGI model, which is an extension of the spherical correlation coefficients proposed by Fisher and Lee[5]. The definition for calculating matching scale between object A and object B is as follows.

$$EC = {m_1}^{\alpha} * {m_2}^{\beta} \quad (2)$$

$$m_1 = \frac{1}{2}\left(\frac{\det\left\{\sum_i \mathbf{p}_i \mathbf{q}_i'\right\}}{\sqrt{\det\left\{\sum_i \mathbf{p}_i \mathbf{p}_i'\right\} \det\left\{\sum_i \mathbf{q}_i \mathbf{q}_i'\right\}}} + 1\right) \quad (3)$$

$$m_2 = \frac{1}{(\gamma |\log(S)| + 1) d_q} \sum_i \left(1 - \frac{\left|\sum_{j \in d(i)} \|\mathbf{p}_j\| - \frac{\|\mathbf{q}_i\|}{S}\right|}{\sum_{j \in d(i)} \|\mathbf{p}_j\| + \frac{\|\mathbf{q}_i\|}{S}}\right) \quad (4)$$

$$S = \frac{\|\mathbf{q}_0\|}{\sum_{j \in d(0)} \|\mathbf{p}_j\|}, \quad (5)$$

where $\mathbf{p}_i$ and $\mathbf{q}_i$ are the normal vectors for object A and object B which calculate the correlations, d_q denotes the number of vectors $\mathbf{q}$, α, β and γ are constant values, " $'$ " in the equations denotes the transpose of the matrix, and $d(i)$ denotes the set of the vector number of $\mathbf{p}$ which corresponds to $\mathbf{q}_i$.

Equation (3) is a function having a range of $(0 \leq m_1 \leq 1)$ which is derived from the spherical correlation coefficients by Fisher and Lee. Equation (4) describes the difference in volume between the two objects. α,β and γ are the constant values whose parameters control the behavior of the coefficient. α sets the influence of the original spherical correlation, β controls the influence of the difference between surface areas, and γ controls the influence of the volume difference between objects for the extended spherical correlation.

The value calculated by equation (3) is rotational invariant since Fisher's spherical correlation coefficients are rotational invariant. m_2 derived from equation (4) is a scalar value. Therefore the extended spherical correlation defined by equation (2) is also rotational invariant. Hence using the MEGI model proposed in **2.1** and the extended spherical correction, we can estimate the matching score which is shift and rotational invariant.

2.3 Matching Algorithm Using MEGI

In order to calculate the EC which was defined in **2.2**, the correspondence of each MEGI element (a set of position vectors and normal vectors) must be known. But this correspondence is unknown for general 3-D objects. Therefore, a correspondence procedure for MEGI elements of two objects must be performed in some way. We proposed the correspondence determination function (CF)[8]

3 Adaptive Scale MEGI

3.1 Multi Scale Description

In the proposed matching method using an extended spherical correlation coefficients and MEGI, miss-correspondence between two object have a serious influence on calculation of EC. In the correspondence procedure, miss correspondence may sometimes be occurred by the following reasons.

- If single surface is divided into a few surfaces through the inference of noise, the calculated correlation will be changed drastically.
- In MEGI, a curved surface is expressed as a set of small surfaces in which position vectors and normal vectors are slightly changed respectively. As a result of this description method, number of surfaces and area of each small surface that describes a curved surface will be different between objects A and object B. Same as mentioned above, the calculated correlation will be changed drastically.

To solve these problems, a new 3-D matching procedure named adaptive scale MEGI is proposed. Adaptive scale MEGI consists of two new procedures. One is new 3-D multi scale resolution image description using tree data structure and the other is adaptive scale matching procedure using multi resolution tree.

While making 3-D multi scale resolution images, convolution with Gaussian function is commonly used. But features such as an inflection point do not increase monotonically by increasing the resolution. Therefore multi scale tree structure which has suitable feature for matching is impossible to construct.

On the other hand, the set of MEGI data consists of position vector and normal vector. No surface shape is needed, and the connectivity of each neighboring surface is not required. Using these features, multi resolution elements unification procedure is proposed

3.2 Construct Multi Scale Description

Step0. (definition) $\boldsymbol{M}_l$ is a set of element $m_{l,i} = (\boldsymbol{X}_{l,i}, \boldsymbol{p}_{l,i})$ which describes a 3D object M, description scale is l. if $l = 0$, $\boldsymbol{M}_l(\boldsymbol{M}_l = \{\boldsymbol{m}_{l,0}, \cdots, \boldsymbol{m}_{l,n_l-1}\})$ describes the initial set of element which is calculated by range data or CAD data. $\boldsymbol{M}_{l-1,i}$ is a set of element which contains children elements of $\boldsymbol{m}_{l,i}$. $\boldsymbol{M}_{l-1,i}$ has a feature described as equation (6). l_{max} is maximum depth of multi scale tree (Fig.1).

$$\boldsymbol{M}_l = \bigcup_i \boldsymbol{M}_{l,i}\ ,\ \boldsymbol{M}_{l,i} \bigcap_{i \neq j} \boldsymbol{M}_{l,j} = \emptyset \qquad (6)$$

Step 1. Let l be 1. Let unification threshold of normal vector th_n and unification threshold of position vector th_p be 0 respectively.

Step 2. $th_n \leftarrow th_n + \Delta_n$, $th_p \leftarrow th_p + \Delta_p$, Δ_n and Δ_p are multi scale step of normal and position vectors respectively and let n_l be 0.

Step 3. New set $\boldsymbol{M}_{l-1,n_l} = \{\boldsymbol{m}\}$ is defined over $\forall \boldsymbol{m} \in \boldsymbol{M}_{l-1} \setminus \bigcup_i^{n_l-1} \boldsymbol{M}_{l-1,i}$. Add all element in $\boldsymbol{M}_{l-1} \setminus \bigcup_i^{n_l} \boldsymbol{M}_{l-1,i}$ which satisfies these two conditions simultaneously to $\boldsymbol{M}_{l-1,n_l}$. (1) An angle between normal vector of a target in $\boldsymbol{M}_{l-1} \setminus \bigcup_i^{n_l} \boldsymbol{M}_{l-1,i}$ and normal vector of m is within th_n. (2) An angle between position vector of a target in $\boldsymbol{M}_{l-1} \setminus \bigcup_i^{n_l} \boldsymbol{M}_{l-1,i}$ and position vector of m is within th_p. And using equation (7), $\boldsymbol{m}_{l,n_l}$ which added to $\boldsymbol{M}_{l+1}$ is calculated.

$$m_{\mathrm{l,n_l}} = \left(\sum_{(\mathrm{X,p}) \in \mathrm{M_{l-1,n_l}}} \frac{\|p\|}{\mathrm{t_{l-1,n_l}}} X, \mathrm{t_{l-1,n_l}} \frac{\sum_{(\mathrm{X,m}) \in \mathrm{M_{l-1,n_l}}} p}{\left\| \sum_{(\mathrm{X,m}) \in \mathrm{M_{l-1,n_l}}} p \right\|} \right) \qquad (7)$$

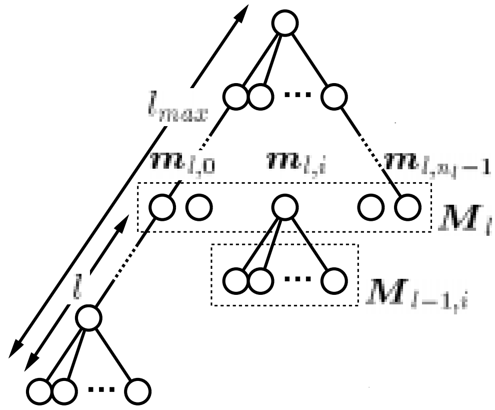
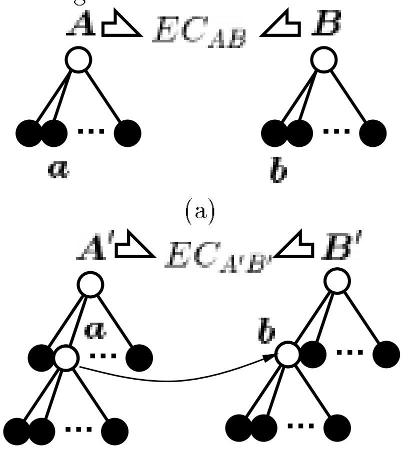


Figure 1: Multi scale tree.



(a)

(b)

Figure 2: Outline of matching algorithm.

where $t_{l,i} = \sum_{(X,p) \in M_{l,i}} \|\boldsymbol{p}\|$ and $n_l \leftarrow n_l + 1$.

Step 4. if $\boldsymbol{M}_{l-1} \setminus \bigcup_i^{n_l-1} \boldsymbol{M}_{l-1,i} \neq \emptyset$ then goto **Step 3.**

Step 5. $l \leftarrow l + 1$. if $l = l_{max}$ then **end** else goto **step 2.**

3.3 Adaptive Scale Matching Algorithm

The second feature of the proposed matching algorithm is that resolution of each regions are selected adaptively on condition that an extension of the spherical correlation coefficients (EC) becomes maximum. To select matching scale adaptively, curved regions that describe a lot of MEGI element can be matched using high resolution, and flat regions describing a few element can be matched using low resolutions. Then computation cost for matching is drastically decreased.

And matching procedure started from low resolutions achieved high robustness matching.

An outline of matching algorithm is as follows. $\boldsymbol{A}, \boldsymbol{B}$ are a set of MEGI elements which describes object A or B respectively. Let an extension of the spherical correlation coefficients between $\boldsymbol{A}$ and $\boldsymbol{B}$ be EC_{AB}.

$\boldsymbol{A}', \boldsymbol{B}'$ are a new set of MEGI element with high resolution element, where one element of $\boldsymbol{A}$ is replaced by a high resolution element $\boldsymbol{a}$ which belongs to the set of$\boldsymbol{A}$ and corresponding element of $\boldsymbol{B}$ is also replaced by $\boldsymbol{b}$(Fig.2). The extension of the spherical correlation coefficients between the two parts is $EC_{A'B'}$.

If $EC_{A'B'} \geq EC_{AB}$, these regions are still pretty much similar when resolution becomes high. If $EC_{A'B'} < EC_{AB}$, matching procedure using high resolution is not performed. These resolution change is described as **resolution translation** hereafter.

Using these resolution control algorithm, adaptive scale matching procedure is described as follows.

step 0. (definition) $\boldsymbol{A}, \boldsymbol{B}$ are set of elements which describe objects A and B. In adaptive scale matching, different resolutions are selected for every regions, then $\boldsymbol{A}, \boldsymbol{B}$ contains various resolutions description. These description divided into two sets, one set denote $\boldsymbol{A}^v$ can perform the next resolution translation, the other set denotes $\boldsymbol{A}^f$ stop the resolution translation. $\boldsymbol{A}$ is described as follows

$$\boldsymbol{A} = \boldsymbol{A}^v \cup \boldsymbol{A}^f \tag{8}$$

$\boldsymbol{A}_l$ denotes a set of elements which shape is A, Description scale l is a length from the initial sets of element $\boldsymbol{A}_0$. The element of $\boldsymbol{A}_l$ are $\{\boldsymbol{a}_{l,0}, \cdots, \boldsymbol{a}_{l,n_l-1}\}$. Children of $\boldsymbol{a}_{l+1,i}$ denotes $\boldsymbol{A}_{l,i}$ which has feature described in equation (9). These above mentioned definition is also applied to $\boldsymbol{B}$.

$$\boldsymbol{A}_l = \bigcup_i \boldsymbol{A}_{l,i} \ , \ \boldsymbol{A}_{l,i} \bigcap_{i \neq j} \boldsymbol{A}_{l,j} = \emptyset \tag{9}$$

step 1. Let l be $l_{max} - 1$. Let $\boldsymbol{A}^f$, $\boldsymbol{B}^f$ be empty sets. $\boldsymbol{A}^v$ is equal to $\boldsymbol{A}_{l,0}$. $\boldsymbol{B}^v$ is equal to $\boldsymbol{B}_{l,0}$ (Fig.3(a)).

step 2. Correspondence procedure is performed between $\boldsymbol{A}^v$ and $\boldsymbol{B}^v$. Using equation (2), EC is calculated, and substitute it to ec. (Fig.3(b)).

step 3. The resolution translation is performed to $\boldsymbol{a}_{l,i}$ which is one element of $\boldsymbol{A}^v$ and the corresponded element $\boldsymbol{b}_{l,i}$. Let $ec_{l,i}$ be calculation result of EC between two elements. This procedure is performed to all element of $\boldsymbol{A}^v$. (Fig.3(c)).

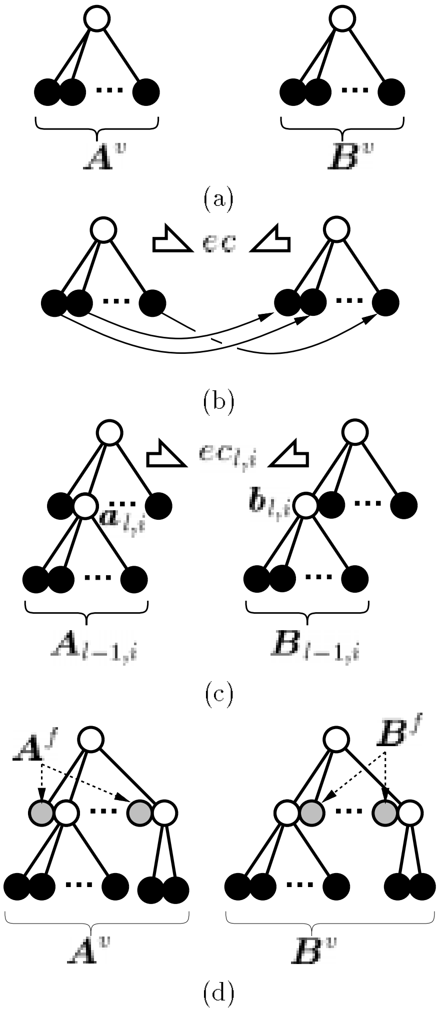


Figure 3: Matching algorithm with adaptive multi scale tree.

step 4. Substitute $a_{l,i}$ to elements in $\boldsymbol{A}_{l-1,i}$ that all i in $ec_{l,i} > ec$. Corresponded elements $\boldsymbol{b}_{l,i}$ are also substitute to elements in $\boldsymbol{B}_{l-1,i}$ (Fig.3(d)). Other element ($ec_{l,i} \leq ec$) are removed from $\boldsymbol{A}^v$ and $\boldsymbol{B}^v$, and add these element to $\boldsymbol{A}^f$ and $\boldsymbol{B}^f$ respectively. After that resolution translation is stopped. Finally EC is calculated and substituted it to ec.

step 5. $l \leftarrow l - 1$. if l is 0 or both $\boldsymbol{A}^v$ and $\boldsymbol{B}^v$ is empty set, then adaptive scale procedure is done. A correlation of coefficient ec is calculated. Else goto **step 3.**

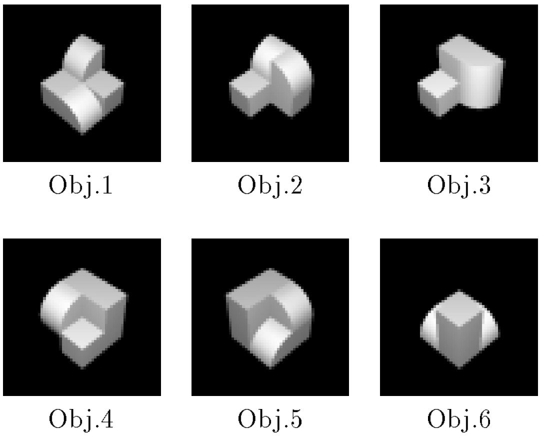


Figure 4: The objects used in experiments.

4 Experimental Results

4.1 Computer Simulation

This matching algorithm using the adaptive scale MEGI has been implemented, and we show examples of 3-D object recognition using some experimental objects constructed on a computer. Fig.4 shows the objects used in this experiment. These objects contains flat and curved regions. We calculated the range data from each object. Performing matching procedures, one matching data is a range data which is a measured object from one view, the other is their 3-D shape models (CAD models) which has whole 3-D shape information.

A view angle of elevation is fixed on 45 degrees, an azimuthal angle is changed from 0 degrees to 90 degrees, range data is measured and construct it to multi scale resolution tree whose height(l_{max}) is 4. In all the experiments, parameters α , β and γ in equations (2),(4) are set at 1.0. Δ_p and Δ_n sets 10 degrees and 5 degrees respectively.

Fig.5 shows the matching correlation between range data Object 2 and CAD models for all six objects including object 2, using the adaptive scale MEGI. Viewing angle (azimuthal) for measuring object 2 changes from 0 degree to 90 degrees. The correlation value of CAD model of object 2 and measured range data amount to a maximum of 6 objects for all aspects, therefore the observed object can be matched to the correct object even while range data which contains only approximately half shape information of original object is used. We also tested using 5 other objects and a recognition rate of 93 % has been achieved.

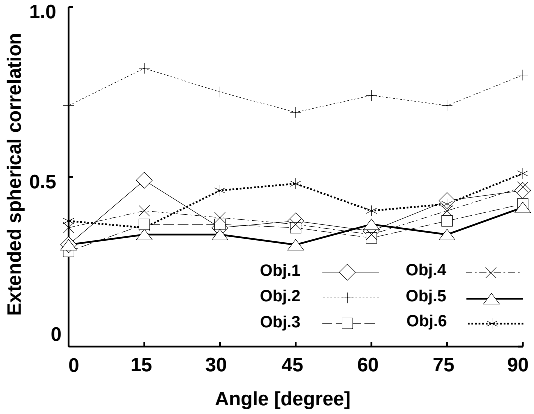


Figure 5: The relationship between direction angle and extended spherical correlation(for Obj.2).

4.2 Human Face Matching using Adaptive Scale MEGI

The experiment using adaptive scale MEGI is performed with the range data of human full face data (25 faces) produced by the National Research Council Canada(NRCC)[7]. Hair part of each full face data is eliminated by hand. In this experiment, 3-D shape model(CAD model) has also generated using range data. Changing view point of range data are also rebuilt using the original range data.

Fig.6 shows one human full face data. In this experiment, a view angle of elevation is fixed at 0 degrees, an azimuthal angle is changed from 0 degree to 10 degrees, range data is measured and it construct multi scale resolution tree whose height(l_{max}) is 4. In all experiments, parameters α , β and γ in equations (2),(4) are set at 1.0. Δ_p and Δ_n are set at 10 degrees and 5 degrees respectively. We did not adjust these parameters to obtain maximum recognition rate. In this research, we want to check the ability of adaptive scale MEGI for recognizing object which has curved surfaces .

Fig.7 shows extended spherical correlation coefficients of adaptive scale MEGI and original MEGI. Recognition rate S_1 is defined the correlation coefficient between rotated range data and original data becomes maximum, and the recognition rate S_3 is defined when correlation coefficient becomes within third orders of magnitude.

Using adaptive scale MEGI, recognition rate becomes high as compared with original MEGI. When in-

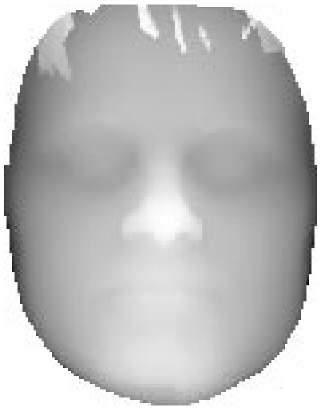


Figure 6: An example of range data images used in experiments.

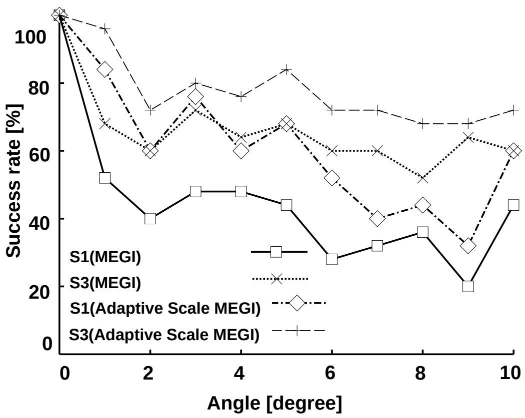


Figure 7: The relationship between direction angle and success rate.

creasing rotation angle, recognition rate becomes low. This is because original face range data is not a complete 3-D image. Therefore rotation image has a lot of occlude part compared with original image. But this result shows high recognition ability of adaptive scale MEGI, even if applying to the images contains curved surface which have very few features like human faces.

5 Conclusions

We proposed the adaptive scale MEGI which has a feature of multi scale description and adaptive scale matching procedure. We also experimented using constructed data and human face range data.

The proposed 3-D description and recognition model "adaptive scale MEGI" has the following features. (1) 3-D description model is a simple model which has no form information and adjacency relationships for surfaces. (2) New 3-D multi scale resolution image description using tree data structure. Using these structure, computation cost for matching is drastically decreased. And matching procedure started from low resolutions achieved high robustness matching. (3) 3-D object which has curved and flat regions can be matched to select matching scale adaptively. In spite of these simplicity, high recognition ability is achieved, even if applying these algorithm to the images contains curved surface which have very few features like human faces.

But a few problem still remains. One is the stability of the constructed adaptive scale tree description in case it contains noise on data or number of measured point are very few; another is is robustness of parameters Δ_n and Δ_t. Changing these two parameters causes the adaptive scale tree to change a lot. Future research will address the development of more powerful schemes for adaptive scale representation, to handle complex objects.

References

[1] Horn, B., " Extended Gaussian Images", Proc. of IEEE, vol.72, no.12, pp.1671-1686, Dec. 1984.

[2] Little, J.J.,"Determining Object Attitude from Extended Gaussian Image", Proc. of 9th IJCAI, pp.960-963, 1985.

[3] Tanaka, K., Sano, M., Meguro, S., Kaneko, H., "3-D Object Pose Estimation Using Spherical Harmonic Function", Information Processing Society of Japan, IPSJ SIG Notes, vol.CV-74-7, pp.1-8, Sep. 1991.

[4] Okada, T., Sano, M., Kaneko, H.,"Three-Dimensional Object Recognition Using Spherical Correlation", Proc. 11th International Conference on Pattern Recognition, pp.250-254, Aug. 1992.

[5] Fisher, N., Lee, A., "Correlation Coefficients for Random Variables on a Unit Sphere or Hypersphere", Biometrica, vol.73, pp.159-164, 1986.

[6] Lyusternik, L.A., "Convex Figures and Polyhedra ", New York, Dover, 1963

[7] Rioux M., and Cournoyer L., "The NRCC Three-dimensional Image Data Files", The Report CNRC 29077, National Research Council Canada, Ottwa, Canada, 1988.

[8] H. Matsuo and A. Iwata, "3-D Object Recognition using MEGI Model from Range Data", Proc. IEEE ICPR, pp.843-846, 1994.