Volumetric Shape Description of Range Data using "Blobby Model"

Shigeru Muraki

Electrotechnical Laboratory
Tsukuba, Ibaraki, 305 Japan

Abstract

Recently in the field of computer vision, there have been many attempts to obtain a symbolic shape description of an object by fitting simple primitives to the range data of the object. In this paper, we introduce the "*Blobby Model*" for automatically generating a shape description from range data. This model can express a 3D surface as an isosurface of a scalar field which is produced by a number of field generating primitives. The fields from many primitives are blended with each other and can form a very complicated shape. To determine the number and distribution of primitives required to adequately represent a complex 3D surface, an energy function is minimized which measures the shape difference between the range data and the "*Blobby Model*". We start with a single primitive and introduce more primitives by splitting each primitive into two further primitives so as to reduce the energy value. In this manner, the shape of the 3D object is slowly recovered as the isosurface produced by many primitives. We have successfully applied this method to human face range data and typical results are shown. The method herein does not require any prior range segmentation.

Keywords: blobby model, generalized algebraic surface, implicit surface, volumetric shape description, range data analysis, energy minimization, ray tracing.

1 Introduction

In the field of computer vision, one of the most important problems is to obtain scene information from 2D images. The typical method is stereo matching which obtains the depth information from the disparity of the images of two cameras placed parallel. The recovered information from this method is often called a "$2\frac{1}{2}D$ *model*"[1], because it consists of depth data which is measured from a single direction and so it does not form a full 3D description of the object. There have been many attempts to fit 3D volumetric shape description models, such as superquadrics[2] to range data[3~8]. However, the shape primitive of these models is usually very simple and so one must combine many primitives in order to adequately express the shape of a complicated object. Then one needs to divide the range data into segments so each segment can be approximated by a single primitive. However, this *segmentation problem* is also a serious problem in the field of computer vision. Further, the connections between primitives are not smooth and it is difficult to express soft objects with smooth changing shapes. Hence, a 3D shape description model is needed that can express a smooth object with a small number of primitives which preferably avoids the segmentation problems.

In the field of computer graphics, modeling and rendering of 3D objects are both important problems. Regular shapes such as machine parts can be simply described. However, a large amount of numerical data is necessary to describe a smooth and soft object such as a human body. At present, designers obtain such descriptions by tedious manual methods. Recently, a new modeling method, which is called "*Blobby Model*"[9], has been used to describe smooth objects. This method expresses a surface of an object as an isosurface of a scalar field which is generated from field generating primitives. Since the shapes of the primitives are blended with each other, it is possible to express the surface of a complicated object with a small number of primitives. However, because of the fusion of the primitives, it is very difficult to design "*Blobby Models*" manually and so an automatic method of obtaining "*Blobby Models*" of 3D objects is desired.

In this paper we present a method for automatically generating a "*Blobby Model*" of a complex 3D object, given a set of range data. The $2\frac{1}{2}D$ *model* obtained by a computer vision technique can be precisely described with a set of blended shape primitives. If the number of primitives is small, then this method also becomes an efficient $2\frac{1}{2}D$ or 3D data compression method. Since the "*Blobby Model*" obtained by this method changes its form from a simple shape to a complicated shape as the number of primitives is increased, we don't need to segment the range data in advance. Further, the history of the changes of the shape show the hierarchical structure of the object and this may be used to further analyze the structure of the object. By adding different kinds of primitives, such as superquadrics, this method can be used as a general modeling tool for computer graphics. In the following section, the "*Blobby Model*" concept is explained and then our method for automatically fitting the model to a set of range data is presented in section 3. Experimental results for human face range data are shown in section 4.

2 Blobby Model

Blinn(1982) developed a generalized algebraic modeling method which is now called the "*Blobby Model*". This model can express a 3D object in terms of the isosurface of a scalar field which is generated from many field generating primitives[9]. The field value at any point (x, y, z) created by a primitive P_i at a point (x_i, y_i, z_i), is expressed as follows.

$$V_i(x,y,z) = b_i exp\{-a_i f_i(x,y,z)\} \qquad (1)$$

The function $f_i(x, y, z)$ defines the shape of the scalar field. For example, in the case of a spherically symmetric field, $f_i(x, y, z)$ is the square of the distance between (x, y, z) and (x_i, y_i, z_i)

$$f_i(x,y,z) = (x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2, \qquad (2)$$

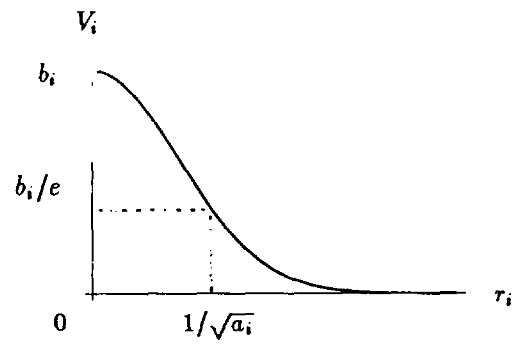


Figure 1: The decay of field value $V_i(x, y, z)$ according to the distance r_i from the point (x_i, y_i, z_i). Eq.(2) was used for the function $f_i(x, y, z)$ of Eq.(1).

and in the case of a superquadric[2] shaped field, it is written as

$$f_i(x,y,z) = \{(x-x_i)^{2/\nu_i} + (y-y_i)^{2/\nu_i}\}^{\nu_i/\mu_i} + (z-z_i)^{2/\mu_i}, \qquad (3)$$

where μ_i and ν_i are the parameters related to the shape of the superquadrics. If Eq.(2) is used, the field value V_i decays exponentially with the distance from (x_i, y_i, z_i) as shown in Fig.1. Parameter $a_i(>0)$ affects the degree of the decay and b_i affects the strength of the field. If several primitives are used at once, then the scalar field from each primitive is summed and the resulting isosurface can show a very complicated shape. From Eq.(1) the field which is produced from N primitives for any point (x, y, z) is expressed as follows.

$$V(x,y,z) = \sum_{i=1}^{N} b_i exp\{-a_i f_i(x,y,z)\} \qquad (4)$$

Consequently, the isosurface of value $T(>0)$ is expressed as an implicit function:

$$V(x,y,z) = T. \qquad (5)$$

If an attribute value is defined, such as the color component C_i, for each primitive, we can calculate the value for a point (x, y, z) as follows:

$$C(x,y,z) = \frac{1}{V(x,y,z)} \sum_{i=1}^{N} C_i V_i(x,y,z). \qquad (6)$$

If there is only one primitive, the primitive makes an isosurface of the function $f_i(x, y, z)$. Fig.2(a) shows

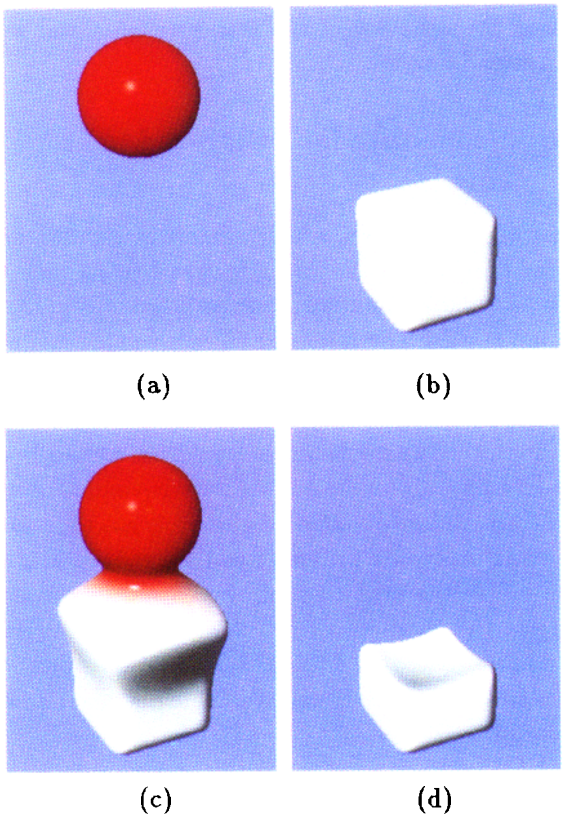


Figure 2: The features of "*Blobby Model*". The parameters for (a) are $a_i = 1.0$ and $b_i = 2.718$. The parameters for (b) are $\mu_i = \nu_i = 0.2$, $a_i = 0.001$ and $b_i = 1.001$.

the isosurface of a spherical field of Eq.(2) and Fig.2(b) shows the isosurface of a superquadric shaped field of Eq.(3). If there are more than two primitives, each primitive is blended and makes a different shape. Fig.2 (c) shows the blended shape of Fig.2(a) and (b). Another interesting feature of the "*Blobby Model*" is the effect of primitives which have a negative b_i value. Since we have assumed that T is positive, we never directly see this kind of primitive. However, such a primitive will form a concavity in any neighboring primitive. Fig.2(d) shows the blended isosurface of Fig.2(a) and (b) when the sign of b_i of the primitive of Fig.2(b) was changed to a negative value. By changing a_is and b_is of Eq.(4), we can control the blending condition of the primitives.

Besides Blinn's "*Blobby Model*", there are similar methods such as "*Metaball*"[10] and "*Soft object*"[11]. These methods change the field function according to the distance from the primitive so that the effect of the field of the primitive is finite. Although this character makes it possible to neglect the effect of the outlying primitives, it creates a problem of necessitating a change in the energy function, which is defined in the next section, according to the distance from the primitives. In this paper, we only consider Blinn's "*Blob by Model*" and we use f_i as defined in Eq.(2) in order to simplify the numerical calculations.

3 Fitting of "Blobby Model" to Range Data

3.1 Formulation of the Optimization Problem

Let us assume that the surface coordinates of M points on the object are measured. If this object is approximated by an isosurface of a field value T, the potential value V of a range data point (x_j, y_j, z_j) should be close to T. Then the function

$$E_{value} = \sum_{j=1}^{M} \{V(x_j, y_j, z_j) - T\}^2 \qquad (7)$$

should be minimized. In other words, the most suitable set of primitives to approximate the range data is obtained by solving the minimization problem of Eq.(7). However, the "*Blobby Model*" has distinctions of both an inside and an outside surface and so the problem arises as to which side should be fitted to the range data. To avoid this problem, we consider the direction of the surface normals of the range data and the model. For the unit surface normal $\mathbf{n}_j$ of a range data point j, the normal vector which is calculated from the depth value of the neighboring pixels in the range data array can be used, as shown in Fig.3. The direction of the normal vector $\mathbf{N}(x, y, z)$ of the "*Blobby Model*" is defined so that it coincides with the negative direction of the gradient of the scalar field, which is represented as:

$$\mathbf{N}(x, y, z) = -\nabla V(x, y, z), \qquad (8)$$

where ∇ is the vector operator,

$$\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z). \qquad (9)$$

Therefore, the most suitable correction of primitives must minimize not only Eq.(7) but also the following function.

$$E_{normal} = \sum_{j=1}^{M} \left|\mathbf{n}_j - \frac{\mathbf{N}(x_j, y_j, z_j)}{|\mathbf{N}(x_j, y_j, z_j)|}\right|^2 \qquad (10)$$

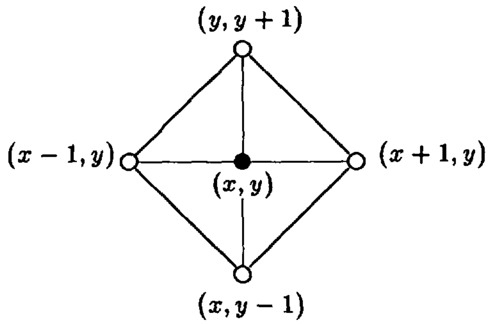


Figure 3: The surface normal vector of each pixel (x, y) is calculated by averaging the normal vectors of the 4 triangles that consist of the pixel and its 4 neighbors.

However, if the range data forms a flat surface, the primitive which is placed infinitely far from the surface and has values $a_i = 0$ and $b_i = T$ exactly satisfies both Eq.(7) and (10). But the computed surface is not limited only to the vicinity of the range data points. Further, the constraints of Eq.(7) and (10) are defined only at range data points, and consequently there is no constraint on the shape forming primitive in the area where there is no range data. Therefore, there is a possibility that the primitive which fits to the range data makes strange shapes away from the vicinity of the range data points. To avoid these problems, a new constraint is added which minimizes the influence of the field of each primitive. From Eq.(1) and (2), the integration of the field of a primitive over 3D space can be expressed as,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V_i(x, y, z)dxdydz = (\frac{\pi}{a_i})^{\frac{3}{2}} b_i. \quad (11)$$

Considering the case where b_i has a negative value, the new constraint is defined as follows:

$$E_{shrink} = (\sum_{i=1}^{N} a_i{}^{-\frac{3}{2}}|b_i|)^2. \quad (12)$$

This constraint has a shrinking effect on the primitives. Consequently, the desired arrangement of primitives can be obtained by minimizing the following energy function, which is the summation of Eq.(7),(10) and (12).

$$E = \frac{1}{M}(E_{value} + \alpha E_{normal}) + \beta E_{shrink} \quad (13)$$

Here α and β are weighting parameters which control the strength of the surface normal constraint and the shrink constraint. By changing these values, we can change the behavior of the fitting to be suitable for the range data.

3.2 Procedure for "Blobby Model" Fitting

A set of N primitives which minimizes Eq.(13) must be found. Since each primitive has five parameters, we must solve the minimization problem of Eq.(13) for $5N$ unknowns. The minimization of Eq.(13) is a nonlinear problem and cannot be solved by an analytical technique. A numerical method such as the Newton method could be used, but it would be extremely difficult to find all of the $5N$ unknowns simultaneously when N is a large number. Our approach is to make an initial fit between a primitive and the range data, and then divide the primitive into two primitives so as to increase the goodness of fit. Continuing this division for all primitives, the detailed surface of the object can be expressed by the isosurface, which is generated by the primitives.

A 5D vector is used to express the parameters of a primitive P_i as follows:

$$P_i = (x_i, y_i, z_i, a_i, b_i). \quad (14)$$

For the initial primitive P_0, the center of the mass of the range data is used for (x_0, y_0, z_0), the reciprocal of the variance of the range data is used for a_0, and b_0 is set at the value $eT(e \simeq 2.718)$. The minimization problem of Eq.(13) is solved using these initial values and then the five parameters of P_0 are determined. A primitive list is created and the primitive P_0 is added to the list. Our method is based on the selection of a primitive P_i from this list and its division into two new primitives. At this stage P_0 is the only primitive in the list and so P_0 is used as P_i. Then P_i is deleted from the primitive list and two new primitives P_i' and P_i'' are appended to the list instead. The initial parameter values of P_i' and P_i'' are calculated as,

$$P_i' = P_i'' = (x_i, y_i, z_i, a_i, b_i/2). \quad (15)$$

We then solve the minimization problem of Eq.(13) for the 10 parameters of P_i' and P_i'' and determine the parameter values. Now there are two primitives in the primitive list. After this the procedure that selects a primitive P_i from the primitive list is repeated and then it is divided into two primitives while holding fixed the parameters of the other primitives in the list. Since the potential field of each primitive is blended,

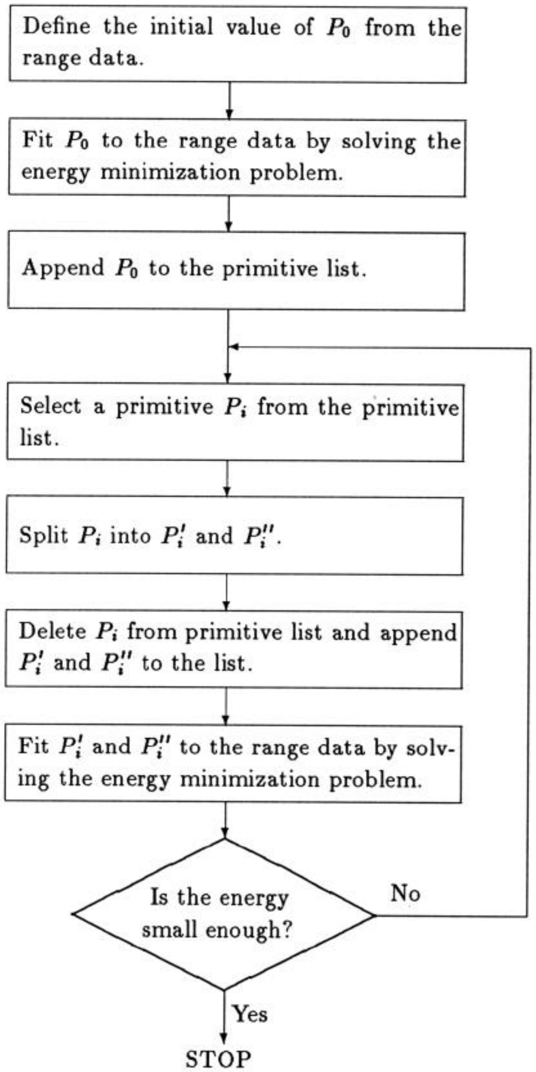


Figure 4: The flow chart of the "*Blobby Model*" fitting procedure.

the selection order of the primitives strongly influences the result of the minimization of Eq.(13). So it is best to choose the effective division in order to obtain a "*Blobby Model*" which preferably approximates the range data with a small number of primitives. To find this effective division, one must examine all of the primitives in the primitive list and determine how much the energy value is reduced by the division of the primitive, and then adopt the division which reduces the energy value the most. However this selection method consumes so much time, so other methods should be used when N becomes a large number. In any case, by continuing this "selection and division" sequence until the energy value becomes sufficiently small, the "*Blobby Model*" gradually comes to approximate the range data as the number of primitives is increased. Fig.4 shows the flowchart of this procedure.

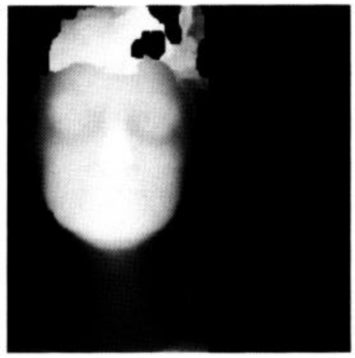
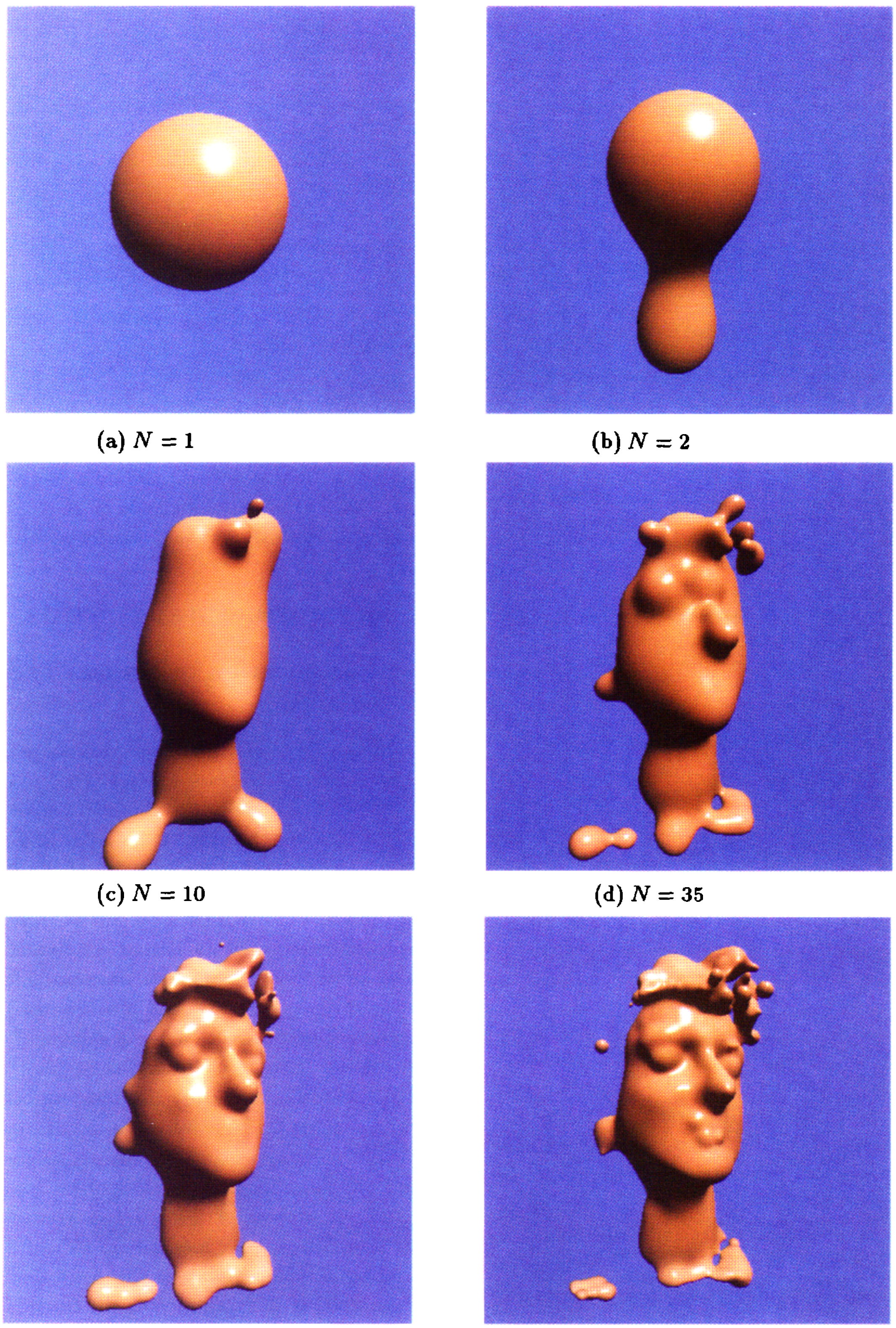


Figure 5: A range image of a human face distributed by NRCC (Face 5).

4 Experimental Results

4.1 Application to Human Face Range Image

We have applied our method to real range data. Fig.5 shows a human face range image (Face 5) distributed by the National Research Council Canada (NRCC) [12]. This image has 256×256 pixels. The depth values are expressed by the intensity of the pixels. The surface normal vector of each pixel is calculated by using the value of neighboring pixels as shown in Fig.3. In order to reduce the amount of calculation, the range image of Fig.5 is blurred by a Gaussian filter of $\sigma = 2$ and one value for every 3×3 pixels is used and 2893 points are obtained with a depth value and a unit normal vector. Then we calculate the parameters of the initial primitive and start the dividing sequence according to Fig.4. To solve the minimization problem of Eq.(13), an approximate solution is determined by using the downhill simplex method[13], which is used as the initial value of the quasi-Newton method[13]. The downhill simplex method is used initially to obtain a reasonable estimate of the unknowns. However, this method is slow. Consequently, the quasi-Newton method is then used, which is much faster, but it does need to have a reasonable estimate of the unknowns. The parameters we used were $\alpha = 1.0$, $\beta = 0.01$ and $T = 1$. Fig.6 shows the change of the "*Blobby Model*"



(a) $N = 1$

(b) $N = 2$

(c) $N = 10$

(d) $N = 35$

(e) $N = 70$

(f) $N = 243$

Figure 6: The transformation of "*Blobby Model*" with the number of primitives N for NRCC range data (Face 5).

Figure 7: A panoramic range image of an actor's face.

with the number of primitives by using a ray tracing technique[14]. It is clearly seen that the detailed features of the face become more apparent as the number of primitives increases. By the image of Fig.6(e), the selection method of a primitive described in section 3.2 was used, however this method was too slow to continue, so we changed the selection method as to choose a primitive merely successively from the primitive list until we obtained Fig.6(f).

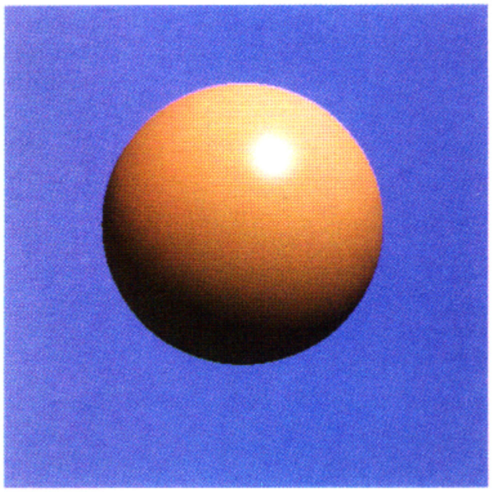
4.2 Application to Panoramic Human Face Range Image

Since the range image of Fig.5 has been taken from a single direction, it does not have any information from behind the head. Hence, from the effect of the shrinking constraint of Eq.(12), the resultant "*Blobby Model*" only represents the facial surface as shown in Fig.6. To obtain the whole shape of a 3D object, one must use multiple range images taken from many directions. Fig.7 is a panoramic range image of a movie actor which has been taken by a special range finder (Cyberware 4020/PS 3-D Digitizer). This image has 512×256 pixels. After we blurred this image by the same Gaussian filter as described in section 4.1, we used one value for every 4×4 pixels from the face area of Fig.7 and obtained 5334 points with a depth value and a normal vector. Fig.8 is the resultant "*Blobby Model*" from this data. In comparison to Fig.6, one can see that the entire shape of the head is correctly reconstructed. The parameters used were $\alpha = 0.1$, $\beta = 0.1$ and $T = 1$. As in Fig.6, we also changed the selection method of a primitive between Fig.8(e) and Fig.8(f).

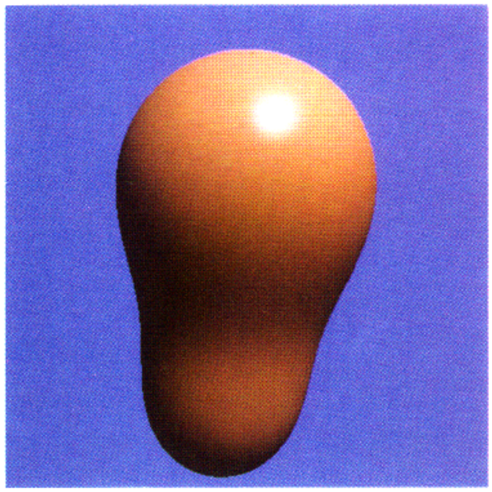
5 Conclusions

We have proposed a method to obtain a volumetric shape description of range data by a "*Blobby Model*" and have successfully applied this method to human face range data. Sufficiently fine features of the faces were restored by using several hundreds of primitives. The history of the primitive division shows a quad tree structure and can be used for hierarchical analysis of an object. For example, one can use a color to represent a branch of the tree and see how the primitives in the branch work. Fig.9(a) shows a "*Blobby Model*" obtained in the experiment of section 4.1. This model consist of 11 primitives and we used red color for the primitive which formed the nose area of the face. Fig.9(b) is a "*Blobby Model*" which consists of 243 primitives. We used a red color for the 13 primitives generated by the division of the initial red primitive of Fig.9(a). Since the primitives are stored in a list structure, this kind of procedure is very simple. Fig.9(c) is the isosurface when these red primitives are removed from Fig.9(b). These pictures show that we can deal with the object structure *by parts* by using the list structure of the primitives of the "*Blobby Model*". In future work, we are going to apply this method to object recognition problems. We also intend to experiment using superquadrics as the primitive.

This method is computationally expensive. For the data of Fig.6(f), it took a few days on a UNIX workstation (Stardent TITAN3000 2CPU). To deal with larger range data sets, further improvement of the algorithm is necessary. The range data we used was obtained by using special devices. To apply this method to the problems of computer vision, one needs to use depth values obtained by stereo matching.



(a) $N = 1$



(b) $N = 2$

(c) $N = 20$

(d) $N = 60$

(e) $N = 120$

(f) $N = 451$

Figure 8: The transformation of "*Blobby Model*" with the number of primitives N for Cyberware 4020/PS image.

(a) (b) (c)

Figure 9: The structure analysis of a "*Blobby Model*".

Acknowledgements

The author thanks Dr. Naokazu Yokoya for helpful discussions about optimization techniques. Special thanks are due to Gaile Gordon of Harvard University and Cyberware Laboratory in making arrangements to use their data. I also thank Richard Baldwin of the Polytechnic of East London and Lisa Bond of Tsukuba University for kindly reading the draft of this paper.

References

[1] Marr, D.: "Vision", Freeman, San Francisco, 1982.

[2] Barr, A.H.: "Superquadrics and Angle-Preserving Transformations", *IEEE Computer Graphics and Applications*, Vol.1, No.1, pp.11-23, 1981.

[3] Kaneta, M., Yokoya, N. and Yamamoto, K.: "Recovery of Superquadric Primitives from Range images by Simulated Annealing",*SIG Notes of the Information Processing Society of Japan*, SIGCV 65-6, 1990. (in Japanese).

[4] Pentland, A.P.: "Recognition by Parts", *SRI International Technical Note*, No.406, 1986.

[5] Boult, T.E. and Gross, A.D.: "Recovery of Superquadrics from Depth Information", *Proc. of Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp.128-137, 1987.

[6] Horikosi, T. and Kasahara, H.: "A 3D Indexing Method for an Image Database", *Technical Report of the Institute of Electronics, Information and Communication Engineers of Japan*, IE88-111, pp.33-40, 1988. (in Japanese).

[7] Solina, F. and Bajcsy, R.: "Recovery of Parametric Models from Range Images: The Case for Superquadrics with Global Deformations", *IEEE Trans. PAMI*, Vol. 12, No.2, pp.131-147, 1990.

[8] Ferrie, F.P., Lagarde,J. and Whaite, P.: "Recovery of Volumetric Object Descriptions From Laser Rangefinder Images", *Proc. of ECCV 90*, pp.387-396, 1990.

[9] Blinn, J.F.: "A Generalization of Algebraic Surface Drawing", *ACM Trans. on Graphics*, Vol.1, No.3, pp.235-256, 1982.

[10] Nishimura, H., Hirai, M., Kawai, T., Kawata, T., Shirakawa, I. and Omura, K."Object Modeling by Distribution Function and a Method of Image Generation", *Trans. IEICE Japan*, Vol.J68-D, No.4, pp.718-725, 1985. (in Japanese).

[11] Wyvill, G., McPheeters, C. and Wyvill, B.:"Data Structure for Soft Objects", *The Visual Computer*, Vol.2, pp.227-234, 1986.

[12] Rioux, M. and Cournoyer, L.: "The NRCC Three-dimensional Image Data Files", *Thec. Report, CNRC* 29077, National Research Council Canada, Ottawa, Canada, 1988.

[13] Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T.: "Numerical Recipes in C", *Cambridge*, 1988.

[14] Kalra, D. and Barr, A.H.: "Guaranteed Ray Intersections with Implicit Surfaces", *Computer Graphics (SIGGRAPH '89 Proceedings)*, Vol.23, No.3, pp. 297-306, 1989.