A Probabilistic Approach to Surface Extraction from Range Data

Hidekazu Hirayu† Caihua Wang† Hideki Tanahashi†
Yoshinori Niwa† Kazuhiko Yamamoto‡

†Office of Regional Intensive Research Project, Softopia Japan, JST
Kagano 4–1–7, Ogaki City, Gifu 503-8569, JAPAN
{hirayu, c-wang, tana, niwa}@softopia.pref.gifu.jp
‡ Department of Information Science, Faculty of Engineering, Gifu University
Yanagido 1–1, Gifu City, Gifu 501-1193, JAPAN
yamamoto@info.gifu-u.ac.jp

Abstract

In this research, we proposed a new method to model complex range data using a parametric model. We first compute local surface curvatures by quadratic fitting of local surfaces, and then extract the curved surfaces and plane regions from the range data by examining the probability distribution of local surface curvatures. After extracting these regions, one shape may be divided into multiple regions by the characteristics of the feature space. Then, after the parameters of each region are estimated, the fitting error between regions of the integration candidate is calculated. Finally, by applying statistical testing techniques on the fitting error, we establish integration criteria and integrate the regions that satisfy this criteria. From this process, it is possible to perform a robust surface extraction. Experimental results on real range data show the effectiveness of the proposed method.

1. Introduction

In recent years, the acquisition of three-dimensional information of a scene or an object has become vital in many fields such as virtual reality and object recognition. In addition, with a rise in the availability of fast, safe, and economical range scanners' capability of wide range sensing, range data has become an adequate means of acquiring 3D information of a scene. However, a range image consists of numerous three-dimensional points, and it is an important task to derive a model description from the range data. In the past, many people have proposed methods concerning the description of three-dimensional objects from the range data. The approaches for range data modeling can be roughly divided into methods using volume modeling, and those using surface modeling. Volume modeling describes an object using function primitives [1]. In order to describe the range data of multiple objects or an object with a complicated shape, it is necessary to segment the range data into primitive parts and then apply a primitive function to them. On the other hand, surface modeling, such as B-rep, is expressed by the relationship between the surfaces and their equations [2].

Using surface methods, range data is first segmented into regions corresponding to surface patches. Then a description of the scene is constructed from region geometries and topologies. In this case, the accuracy of the object description is dependent on the results of the region segmentation, which is the most difficult task in range data processing. At present, there are no established methods to segment noisy range data.

In this present research, we generate a feature space histogram using local surface curvatures. The stable planar and curved surface regions are extracted by examining the probability distribution of the histogram. For the range data containing only a polyhedron, a feature space histogram may be generated using a normal vector. However, since our aim is to process range data containing both planes and curved surfaces with divided regions, the feature space histogram may be generated using local surface curvature. After some regions are obtained, on the basis of the parameter of the quadric surface equation of each region and geometric features, the integration between multiple segmented regions has been carried out. Figure 1 shows an example of the range data, which consists of some planes and quadric surfaces used in this research.

In our approach, we first compute local surface curvatures by quadratic fitting of local surfaces, and then extract curved surfaces and plane regions from the range data by examining the probability distribution of local surface curvatures. As a result, one shape may be divided into the multiple regions according to the characteristics of the feature space. Then, after the parameters of each region are esti-

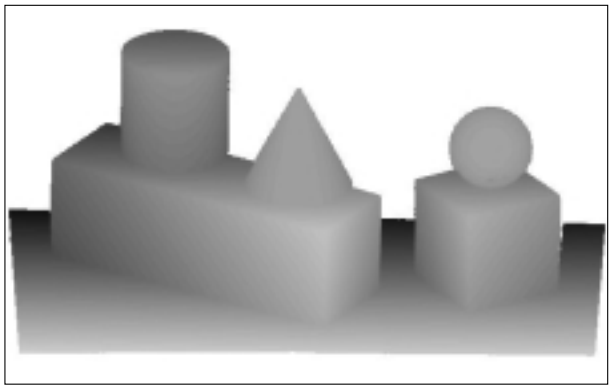


Figure 1. Test range image.

mated, the fitting error between regions of the integration candidate is calculated. Subsequently, by applying a statistical testing technique on the fitting error, we establish integration criteria and integrate the regions that satisfy the criteria. Finally, it becomes possible to accomplish a robust surface extraction.

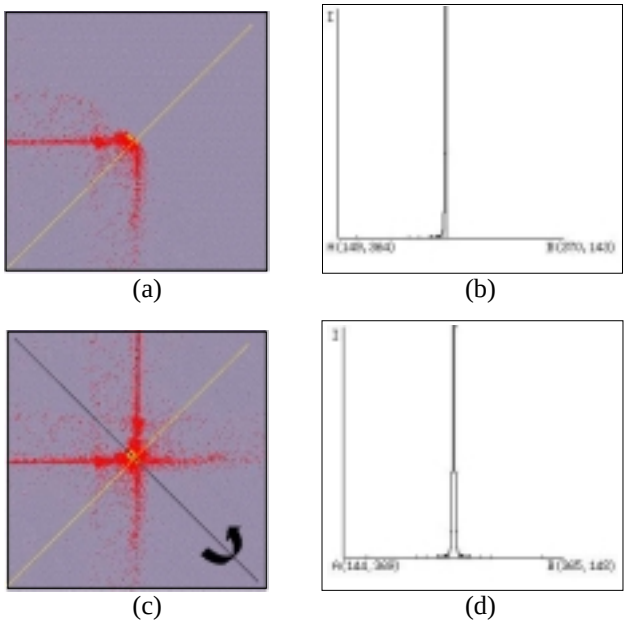


Figure 2. (κ_1, κ_2) **Feature space histogram.**

2. Segmentation of range data

As for segmentation using curvature, there are techniques based on the signs of both Gaussian Curvature (K) and Mean Curvature (H) [3, 4, 5]. In these methods, H and K are used to classify the local surfaces to specified types of surfaces, due to the combination of the signs of H and K. The signs of H and K alone cannot help to distinguish surfaces of the same type with a significant difference in radius length. In this paper, we propose a method that can solve this problem using the curvature feature space. First, we compute local surface curvatures by quadratic fitting of local surfaces, and then the regions with different radius lengths of curvature are extracted by examining the probability distribution of local surface curvatures.

2.1. A feature space histogram using curvatures

First, we fit quadratic equations to each local surface region of the range data by the least-squares method using the following quadratic equation (1). A 15×15 window centered on an observation point is used as the local surface region.

$$Z = aX^2 + bY^2 + cXY + dX + eY + f \qquad (1)$$

(Where a, b, c, d, e and f are parameters)

The local surface curvatures are computed from the local surface equation, projected into the feature space, and then transformed into a histogram. If the local surface curvatures of a curved surface in the range data obey a Gaussian distribution in the feature space, we can segment the range data into curved surfaces by extracting Gaussian distributions in the feature space. A feature space histogram generated from Gaussian and Mean Curvature is used in some methods, but the usage of $K > 0$ and $H = 0$ does not exist in this space. Moreover, due to the non-linear relation between H and K, the distributions of the curvatures are deformed Gaussian. In this research, we use the feature space composed of the principal curvatures of κ_1 and κ_2.

In this case, as shown in Figure 2(a), since κ_1 and κ_2 are projected only into half of the histogram space, the region with curvatures near $\kappa_1 = \kappa_2$, like a plane or spherical surface, does not become a Gaussian distribution, as shown in Figure 2(b). Then, by folding back the feature space around the axis of $\kappa_1 = \kappa_2$, the range data can be projected into all of the histogram space, as shown in Figure 2(c). By projecting the feature space around the axis of $\kappa_1 = \kappa_2$, the distribution on the axis of $\kappa_1 = \kappa_2$ becomes a Gaussian distribution with symmetry, as shown in Figure 2(d).

Until now, we have used a rectangular coordinate system that sets an axis of coordinates as (κ_1, κ_2) for the feature space histogram. However, the following problems have been detected:

1. Since the change of curvature is very large compared with the change of radius, a surface with a small radius has a variance that is larger than a surface with a larger radius, causing the radius and variance to become nonlinear.

2. In the case of scaling (κ_1, κ_2), since all the curvatures in the range data are projected into feature space, a very large space is needed.

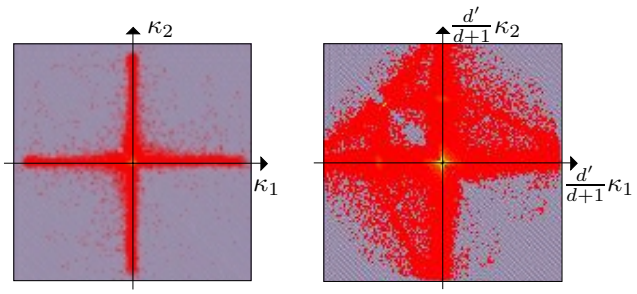


Figure 3. Feature space histogram of test image.

By using the logarithm of (κ_1, κ_2), the aforementioned problems have been solved. When the logarithm is calculated for κ_1 and κ_2 independently, the distribution of the histogram collapses near the κ_1 and κ_2 axis. Therefore, we use space $(\frac{d'}{d+1}\kappa_1, \frac{d'}{d+1}\kappa_2)$, which was calculated from the logarithm distance $d = \sqrt{{\kappa_1}^2 + {\kappa_2}^2}$ of (κ_1, κ_2).

$$d = \sqrt{{\kappa_1}^2 + {\kappa_2}^2} \tag{2}$$

$$d' = \frac{A}{log(1 + M_d)} log(1 + d) \tag{3}$$

where A is the resolution of the histogram, and M_d is the maximum of $d + 1$. In order to use histogram space efficiently, a 1% of the range data which has the largest curvature (i.e., the edge portion in the range data) is removed, and the rest of the data is projected into the histogram. Figure 3(a) shows the feature space histogram that only scales κ_1, κ_2 in Figure 1. Figure 3(b) shows the feature space histogram, which was calculated from the logarithm of distance $d = \sqrt{{\kappa_1}^2 + {\kappa_2}^2}$ of (κ_1, κ_2).

2.2. Extraction of planes and curved surfaces

There are many planar and curved surfaces in range data, and their local curvature distributions may overlap in the histogram space creating the possibility of a multi-distribution. It is very difficult to separate each distribution from the multi-distribution. In this work, we use the histogram of local surface curvatures together with space information in the range data to extract the distribution of each plane and curved surface.

From this histogram, we identified the highest peak value and projected it back into the range data. This process of re-projection produced some discrete points $P_i (i = 1, \cdots, m)$ in the curved surface S_p (or a set of surfaces S_{pi}), whose values of $\frac{d'}{d+1}\kappa_1$ and $\frac{d'}{d+1}\kappa_2$ correspond to the detected peak. We assumed the points to be approximately equal to the true curvatures of the surface(s). Because curved surfaces appear in the range data as regions, we assumed that for the majority of P_i that their neighbors $N_{ij} (j = 1, \cdots, n)$ also belong to the same surface as P_i. Therefore, the neighbors N_{ij}, which can be obtained by a dilation operation centered about P_i, can be regarded as a set of sample points on the surface S_p. The neighbors N_{ij} are then used to estimate the mean M and the covariance matrix C of the distribution of the curvatures of the surface S_p in $(\frac{d'}{d+1}\kappa_1, \frac{d'}{d+1}\kappa_2)$ space, as shown in the following equations:

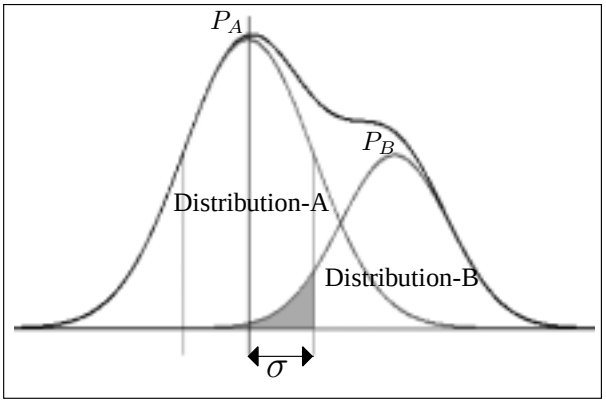


Figure 4. Region extraction from mixed distribution.

$$M = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} k_{ij} \tag{4}$$

$$C = \frac{1}{mn - 1} \sum_{i=1}^{m} \sum_{j=1}^{n} (k_{ij} k_{ij}^{\top} - M M^{\top}) \tag{5}$$

where k_{ij} is the projected point of the local surface curvature of the sample points in the $(\frac{d'}{d+1}\kappa_1, \frac{d'}{d+1}\kappa_2)$ space histogram.

The planes and curved surfaces are then extracted based on the distribution estimated above. We first extract the points whose curvatures are located within the confidence interval containing $2/3$ of the population of the distributions. The extracted points may represent multiple regions on range data. Then, when we expand these points with a dilation operation constrained to the confidence interval of 99.9% of the population, some large collected regions are extracted.

As shown in Figure 4, even if there is a distribution B with peak P_B near another distribution A with peak P_A, two regions were not extracted as one collected region without entering the majority of distribution B in the range of σ of distribution A. Moreover, since the small region is not considered to be the object of interest, only the large region is extracted.

The above process was applied to the remaining range data until no significant region was extracted.

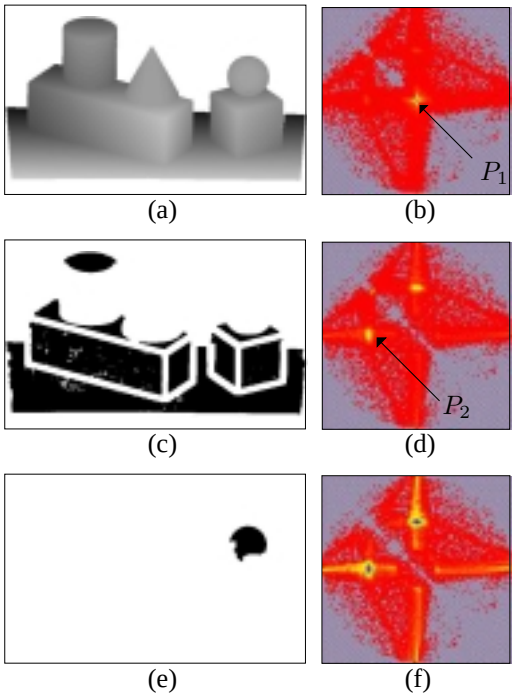


Figure 5. Results of segmentation after two iterations.

Figure 5 shows the results of plane and curved surface extraction from the range data (Figure 1), where the results of the first two iterations of surface extraction are illustrated.

Figure 5(a) shows range data and Figure 5(b) shows the feature space histogram of principle curvature computed from Figure 5(a). The planar regions in Figure 5(c) are first extracted from the distribution making point P_1 (Figure 5(b)) the highest peak. Next, the feature space histogram in Figure 5(d) is generated with the exception of the distribution of the planar regions in Figure 5(c). The spherical region in Figure 5(e) is extracted from the distribution, which makes point P_2 (Figure 5(d)) the highest peak. Then the feature space histogram (Figure 5(f)) is generated with the remaining data excluding the distributions in Figure 5(c) and (e).

Figure 6 shows planar and curved surface regions that have finally been extracted. A total 35 regions were extracted, as shown in Figure 6 from the range data.

3. Integration of surface regions

Since the local curvatures in the hyperboloid and ellipsoid surfaces change, they may be divided into some piecewise surfaces. However, these piecewise surfaces will have similar parameters. Therefore, integration of surface regions is needed. After we estimated the parameters of the equation of each region, we tested the possibility of integration of neighboring regions.

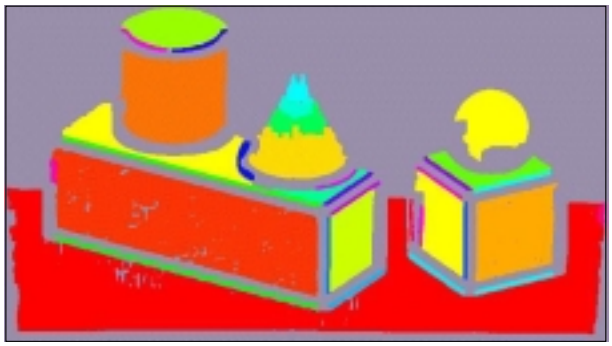


Figure 6. Extracted plane and curved surface regions from range data.

In addition, in the segmentation process the regions are detected without overlapping each other. Dilation operation bounded by the discontinuity edges was carried out for each region. Then the regions of integrating candidates are obtained from the overlapping dilated regions.

The fitting errors $D_{Bi}(i = 1, \cdots, m)$ of point P_B of region B in the equation of region A are calculated from regions A and B as integrating candidates. Finally, examination of whether it can be integrated by the statistical method for fitting error D_{Bi} is carried out.

3.1. Estimation of surface equations

For the extracted regions, we estimated their parameters by fitting the three-dimensional points in corresponding regions by the following equation [6, 7, 8, 9].

$$(X, Y, Z)A \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + B \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + C = 0 \qquad (6)$$

$$A = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix} \quad B = (g, h, i) \quad C = j$$

3.2. Calculating fitting error of regions for integrating candidates

There are some methods, which calculate the fitting errors $D_{Pi}(i = 1, \cdots, m)$ of point P_B in region R_B in the equation of region R_A by assuming regions R_A and R_B are integrating candidates.

3.2.1 Fitting error according to quadric surface equation

The residual error, which is obtained by substituting three-dimensional points for the quadric surface equation can be

used as the fitting error between the point and the curved surface. However, the error is not the Euclidean distance from the three-dimensional point to the curved surface, but is involved in the scale of the parameter of the quadric surface equation or the value of the parameter influencing the residual. For example, for the point which protruded from the spherical surface in distance ε, the residual error becomes equation (7). The residual error is clearly proportional to the size of the radius of the curved surface.

$$X^2 + Y^2 + Z^2 = (r + \varepsilon)^2 = r^2 + \underline{2r\varepsilon} + \varepsilon^2 \qquad (7)$$

On the other hand, if a quadric surface equation is fitted using both region R_A and region R_B, there is a problem when the residual error D_{Pi} is calculated for the points in region R_A and R_B. For example, when a quadric surface equation is fitted to the cluster of points from two planer regions, it becomes equation (8).

$$(a_1X + b_1Y + c_1Z + d_1)(a_2X + b_2Y + c_2Z + d_2) = 0 \qquad (8)$$

Thus, small fitting error D_{Pi} will be obtained for all the points as the associated values with small errors are multiplied together.

3.2.2 Error by shortest distance from point to curved surface

When a segment $\overline{P_AP_B}$ in Figure 7, which connects point P_A in region R_A and point P_B in region R_B, becomes perpendicular to region R_A at the point P_A, the segment $\overline{P_AP_B}$ becomes its shortest. However, it is very difficult to find the normal vector that passes through both the point P_A and P_B in the region R_A.

3.2.3 Error by approximate shortest distance from point to curved surface

The approximate shortest distance from a point to a curved surface is calculated based on the normal vector of the curved surface. As described above (section 3.2.2), it is very difficult to find the point in the region that becomes perpendicular to a region and passes through a point P of another region. In this paper, for two regions of integrating candidates, the approximate shortest distance D_{Pi} from each point P_i in one region to the surface of the other region has been calculated.

As shown in Figure 7, for region R_A and R_B the number of points in region R_A is assumed to be greater than the number of points in region R_B. The following steps have been adopted to obtain D_{Pi}:

1. Assume that the point P_B in region R_B is on the surface of region R_A, and compute the normal vector N_{PB} at point P_B.

2. Let the direction vector be N_{PB} intersecting point P_A by a straight line L with region R_A being obtained from straight line L, which passes through point P_B.

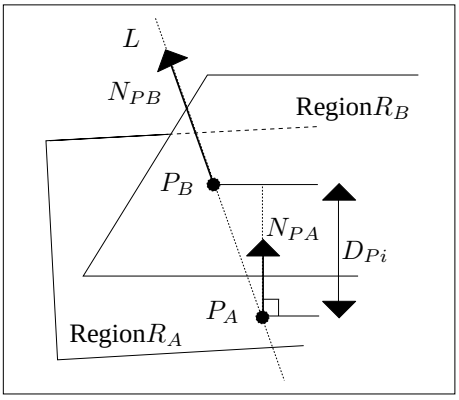


Figure 7. Estimating distance D_{Pi} from normal vector.

3. The normal vector N_{PA} at point P_A is calculated, and the approximate shortest distance D_{Pi} between region R_A and region R_B is obtained using the equation (9).

$$\frac{|\overrightarrow{P_{Ai}P_{Bi}} \cdot \overrightarrow{N_{PAi}}|}{|\overrightarrow{N_{PAi}}|} = D_{Pi} \qquad (9)$$

3.3. Integration of regions

For error D_{Pi} of an approximate shortest distance between a point and a curved surface obtained from the equation (9), the method which judges whether two regions can be integrated together is examined in this section.

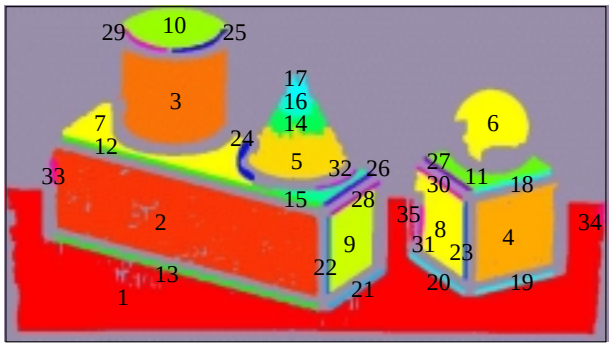
The largest error permitted in the model is set as a tolerance. The method is that if all the fitting errors from the points $P_{Bi}(i = 1, \cdots, m)$ in region R_B to region R_A are within the largest error. Then these points P_{Bi} are integrated into region R_A.

However, in the case of actual data, all points do not always satisfy the tolerance, since there is a dispersion of points at the shortest distance. Therefore, a permissible error does not need to be one interval. Instead it is made to be a tolerance using a probability distribution (Gaussian distribution). Hence it is possible to judge whether the regions can be integrated together using a statistical method.

For example, first let us consider the distribution modeling the accuracy necessary for integration when the *t*-test is used. This is decided by the measurement accuracy of the range finder. In the experiment, the mean value is assumed to be 0 with a variance of 0.5% of the maximum width of the largest object.

Thus in the distribution modeling, the *t*-test is carried out for the necessary precision and the distribution that is actually composed of the approximate shortest distance D_{Pi} between regions.

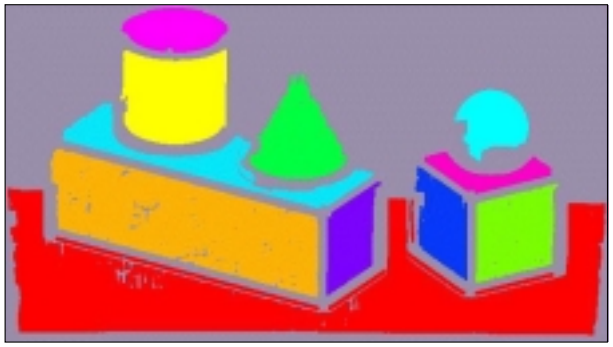
Furthermore, there is a test of the absolute distance using a simpler method. At first, the upper and lower values



(a) Regions and region labels

Region:1-13-19-20-21-34-35	Region: 8-23-30-31
Region:2-23	Region: 9-22-28
Region:5-14-16-17	Region:10-25-29
Region:7-12-15-24-26-32	Region:11-18-27

(b) Region labels which should be integrated



(c) Results of integration

Figure 8. Results of integration.

falling within the 95% of this distribution are obtained from the mean and variance of the approximate shortest distance D_{Pi} between regions. Then, the value with a greater distance from the origin to upper value or to the lower value is compared with the threshold. If the threshold is smaller than this value, the integration is carried out. In the experiment, the threshold is the same value as 0.5% of the maximum width of the largest object as mentioned above.

3.4. The experiment

In this paper, for the segmented regions in Figure 6, the approximate shortest distance between the points and curved surfaces were calculated. We performed the experiment by calculating the absolute distance from each point to each curved surface with the following results. Furthermore, very similar results of the T-test were obtained.

Figure 8(a) shows the results of segmentation with the regions labelled. Figure 8(b) shows the regions that should be integrated from range data. Figure 8(c) shows the results of integration. In the experiment, there were 77 groups of regions as integrating candidates, resulting in 31 groups. The segmented 35 regions were integrated to 11 regions.

4. Conclusion

In this paper, we proposed a robust method to extract planer and quadric surface regions from range data. This is a method in which we compute local surface curvatures by quadratic fitting of local surfaces, and then extract stable planer regions and curved surface regions from the range data by examining the probability distribution of local surface curvatures. For the extracted regions, we estimated their parameters by geometric fitting using the three-dimensional points in these regions, including some regions that were integrated based on the approximate distance between regions.

In this paper, we used the space with calculated the logarithm distance $d = \sqrt{{\kappa_1}^2 + {\kappa_2}^2}$ of (κ_1, κ_2) as feature space. In future work, in order to extract a curved surface region with greater accuracy, we are planning to generate the feature space that expedites the extraction accuracy of the region with the curvature within a certain range, rather than the extraction accuracy of the region which curvature is near zero or has a very high value, and carry out a verification method for noise interference. We are planning to integrate multi-view range data as well.

References

[1] H. Tanahashi, K. Yamamoto and K. Kato, "Shape description of two super quadrics from range data using GA", *Trans. IEEJ*, 119-D(1):44–49, January 1999.

[2] O.D. Faugeras, M. Herbert and E. Pauchon, "Segmentation of range data into planar and quadratic patches", *Proc. CVPR*, pages 8–13, 1983.

[3] P.J. Besl and R.C. Jain, "Segmentation through variable order surface fitting", *IEEE Trans. on PAMI*, 10(2):167–192, March 1988.

[4] N. Yokoya and M.D. Levine, "Range image segmentation based on differential geometry: A hybrid approach", *McGill Univ. Tech. Rep. McRCIM-TR-CIM*, pages 87–16, September 1987.

[5] M. Yang and E. Lee, "Segmentation of measured point data using a parametric quadric surface approximation", *Computer-Aided Design*, 31:449–457, 1999.

[6] S. Kaveti, E.K. Teoh and H. Wang, "Second-order implicit polynomials for segmentation of range images", *Pattern Recognition*, 29(6):937–949, 1996.

[7] D. Keren and C. Gotsman, "Fitting curves and surfaces with constrained implicit polynomials", *IEEE Trans. PAMI*, 21(1):31–41, January 1999.

[8] K. Kanatani, "Statistical optimization for geometric computation: Theory and practice", *Elsevier Science*, 1996.

[9] X. Cau, N. Shrikhande and G. Hu, "Approximate orthogonal distance regression method for fitting quadric surfaces to range data", *Pattern Recognition Letters 15*, pages 781–796, 1994.