Polyhedral Description of Panoramic Range Data by Stable Plane Extraction

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

†Office of Regional Intensive Research Project, Softopia Japan, JST
{c-wang,tana,hirayu,niwa}@softopia.pref.gifu.jp
‡ Department of Information Science, Faculty of Engineering, Gifu University
yamamoto@info.gifu-u.ac.jp

Abstract

In this paper, we describe a novel technique to extract a polyhedral description from panoramic range data of a scene taken by a SceneModeler laser range finder. First, we introduce a reasonable noise model of the range data acquired with a laser radar range finder, and we derive a simple and efficient approximate solution of the optimal fitting of local planes in the range data under the assumed noise model. Then, we extract stable planar regions from the range data by using both the distribution information of local surface normals and their spatial information in the range image. Finally, we describe a method which builds a polyhedral description of the scene using the extracted stable planar regions of the panoramic range data with 360° *field of view in a polar coordinate system. Experimental results on complex real range data show the effectiveness of the proposed method.*

1 Introduction

In recent years, in connection with the appearance of fast and safe laser range finders with high ability of wide range sensing, such as Cyrax and SceneModeler, range data has become a convenient mean for acquiring 3-dimensional information from a scene. Scene modeling and understanding using range data became attractive because 3D information about the scene, which is very difficult to be recovered accurately and stably from 2D image or 2D stereo images, is directly available from the range data. In this research, we consider the problem of extracting a polyhedral description of a scene from panoramic range data taken with the SceneModeler laser range finder. This problem is important because the polyhedral description is not only useful and efficient for modeling from huge amount of range data (for example, there are 8000×1400 points in the range data of one scene taken by SceneModeler), but also plays a basic role in a hierarchical scheme of model description. In such a modeling scheme, the polyhedral description can be used as a base for generating a mesh model for more accurate representation of details of the range data, as well as for constructing a high-level description by deriving higher order surfaces from the first-order planar surfaces.

Many techniques for range data modeling have been proposed in the past, but here we only review some methods which are related to the construction of polyhedral models from range data[1-10]. Faugeras et al.[1,2] proposed a method which segments the range data into planar and quadric patches by region growing, using local surface normals and quadric features. Taylor[7] segmented range images into planar regions by a spit and merge approach using the local surface normals. There are also other methods which used a similar approach of using the local surface normals with a region growing or clustering techniques[3,4]. On the other hand, Jiang et al.[5] and Haindl et al.[6] chose a special approach in which each scan line of the range data is segmented into straight line segments and then the straight line segments are merged into planar regions. Bock et al.[8] also proposed a method to extract planar regions with the Hough transform using the edges in range data.

While most of the proposed methods consist in the segmentation of range images, Hoover et al.[9] tried to build a polyhedral model of the scene using the planar regions segmented from range images. They used the topological relations of planar surfaces to describe the visible space of the scene and introduced virtual planes to represent the invisible space, that is, the occluded space of the scene. Fitzgibbon et al.[10] built CAD models from range data using the planar and quadric regions obtained from the segmentation of range data.

In this paper we propose a novel technique to extract a polyhedral description from panoramic range data obtained with a laser range finder. First, we derive an efficient and effective method to approximately estimate the optimal local plane fitting under the noise model. Second, we use both the distribution and spatial information of the local normals to extract the stable planar regions from range data and use them to build a polyhedral description of the scene. Third, we deal with panoramic range data with 360° field of

view, and consider how to derive a polyhedral description of the panoramic scene in the polar coordinate system, using the stable planar regions extracted in the previous steps. Experimental results on complex real range data show the effectiveness of the proposed method.

2 Local Plane Fitting

2.1 Noise model

The range data, which is acquired by a SceneModeler laser radar range finder, is given by a set of 3-dimensional points $\{(X_{ij}, Y_{ij}, Z_{ij})^\top\}$. Here, ij is the lattice point in a range image, and $^\top$ stands for transposition of a vector. In the panoramic range image, i and j stand for the indices of pitch and yaw angles, respectively. The spatial sampling error and the error in distance measurement are included in the range data, and this can be modeled as noise. Compared to the latter, the former is usually much smaller. Moreover, it can be further reduced by calibration. Therefore, we can omit the error in spatial sampling and consider only the error in distance measurement. Fig. 1 shows this noise model.

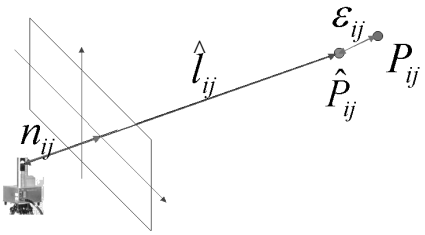


Figure 1. Noise model of range data

Let $\boldsymbol{n}_{ij} = (x_{ij}, y_{ij}, z_{ij})^\top, \|\boldsymbol{n}_{ij}\| = 1$ be the radial direction of a 3-dimensional point $\boldsymbol{P}_{ij} = (X_{ij}, Y_{ij}, Z_{ij})$ and l_{ij} be the distance from a sensor to $\boldsymbol{P}_{ij}$. Let $\hat{\boldsymbol{P}}_{ij}$ be the real position of $\boldsymbol{P}_{ij}$, and $\hat{l}_{ij}$ be the distance from $\hat{\boldsymbol{P}}_{ij}$ to the sensor (that is, the real distance of $\boldsymbol{P}_{ij}$). Assume that the error between l_{ij} and $\hat{l}_{ij}$ obeys a Gaussian distribution $N(0, \sigma)$. As we can assume that a spatial sampling error is close to 0, $\boldsymbol{n}_{ij}$ can be regarded as the real radial direction of point $\boldsymbol{P}_{ij}$. Therefore, the relation of the observed point $\boldsymbol{P_{ij}} = (X_{ij}, Y_{ij}, Z_{ij})^\top$ and its real position $\hat{\boldsymbol{P}}_{ij}$ can be written as the following equation.

$$\boldsymbol{P}_{ij} = \boldsymbol{n}_{ij}(\hat{l}_{ij} + \epsilon_{ij}) = \hat{\boldsymbol{P}}_{ij} + \boldsymbol{n}_{ij}\epsilon_{ij} \tag{1}$$

where ϵ_{ij} obeys the Gaussian distribution $N(0, \sigma)$.

When the real position $\hat{\boldsymbol{P}}_{ij}$ of the observed point $\boldsymbol{P}_{ij}$ is on a plane, the following formula is obtained.

$$\boldsymbol{A}\boldsymbol{P}_{ij} + d = \boldsymbol{A}\boldsymbol{n}_{ij}\epsilon_{ij} \tag{2}$$

where $\boldsymbol{A} = (a, b, c)$.

Given some sample points of a plane, an optimal fitting of the plane can be acquired by solving the following weighted least-squares problem.

$$\begin{aligned} J[\boldsymbol{v}] &= \frac{1}{N}\sum_{i,j} W_{ij}\boldsymbol{v}^\top \begin{pmatrix} \boldsymbol{P}_{ij}\boldsymbol{P}_{ij}^\top & \boldsymbol{P}_{ij} \\ \boldsymbol{P}_{ij}^\top & 1 \end{pmatrix} \boldsymbol{v} \\ &\longrightarrow \min \end{aligned} \tag{3}$$

$$W_{ij} = \frac{1}{(\boldsymbol{A}\boldsymbol{n}_{ij})^2} \tag{4}$$

$$\text{Subjected to:} \quad \|\boldsymbol{v}\| = 1$$

where $\boldsymbol{v} = (\boldsymbol{A}, d)^\top$ is the parameter of a plane to be estimated, and N is the number of sample points.

Since the weights contain the parameters to be estimated, it is difficult to obtain the analytical solution of the above minimization problem directly. An iterative method called the re-normalization method[11] was proposed to compute the optimal estimation. However, since it adopts an iterative strategy, it is time consuming. In this section, we give an approximate solution for the minimization problem in Equation (3).

2.2 Least-squares approximation

Using Eq. (2) and the relation $\boldsymbol{P}_{ij} = \boldsymbol{n}_{ij}l_{ij}$, we obtain the following formula.

$$\begin{aligned} \boldsymbol{A}\boldsymbol{P}_{ij} + d &= \frac{\boldsymbol{A}\boldsymbol{P}_{ij}}{l_{ij}}\epsilon_{ij} = \frac{\boldsymbol{A}\boldsymbol{n}_{ij}\epsilon_{ij} - d}{l_{ij}}\epsilon_{ij} \\ &= \boldsymbol{A}\boldsymbol{n}_{ij}\frac{\epsilon_{ij}^2}{l_{ij}} - \frac{d}{l_{ij}}\epsilon_{ij} \end{aligned} \tag{5}$$

Let $\bar{l}$ be the average value of l_{ij} in a local region, and $\alpha_{ij} = \frac{l_{ij}}{\bar{l}}$. If we multiply both sides of Eq. (5) by α_{ij}/d, we obtain the following equation.

$$\frac{\boldsymbol{A}}{d}\alpha_{ij}\boldsymbol{P}_{ij} + \alpha_{ij} = \boldsymbol{A}\boldsymbol{n}_{ij}\left(\frac{\epsilon_{ij}}{\sqrt{\bar{l}d}}\right)^2 - \frac{\epsilon_{ij}}{\bar{l}} \tag{6}$$

The right-hand side of the above equation consists of two items of the first and second order of the noise ratio. Generally, since the ratio of the noise is far smaller than 1, the second order item can be omitted. Therefore, the optimal plane fitting problem can be approximated by a weighted least-squares problem, and the direction of the normal of the plane is given by the following formula.

$$\boldsymbol{d} = [\sum_{i,j}\alpha_{ij}^2\boldsymbol{P}_{ij}\boldsymbol{P}_{ij}^\top]^{-1}\sum_{i,j}\alpha_{ij}^2\boldsymbol{P}_{ij} \tag{7}$$

where $[M]^{-1}$ stands for the inversion of matrix M.

The comparison of experimented results[12] on synthetic noise range data and real range data with ground truth shows that the proposed method has similar performance as the renormalization method, which was seen as the optimal method under the noise model hypothesis.

3 Stable Planar Region Extraction

The local surface normal at each pixel is estimated by the above method, using a 7×7 window centered at the concerned pixel. Note that $j_{max} + 1 \rightarrow 0$ in the polar coordinate system. The local surface normals of a plane in the range data have similar values which form a distribution in the space of the local surface normals. On the other hand, the plane appears as a region of pixels connected together in the range image. Therefore, in this research, we use both the distribution of local surface normals in the normal space and the spatial information of the corresponding pixels to extract planar regions from the range data.

3.1 Histogram of local surface normals

The local surface normals $\boldsymbol{n} = (n_x, n_y, n_z)$ can be obtained by normalizing (a, b, c), which were estimated in the above section, with the sign undetermined. If we constrain the direction of the local surface normals to point to the outside from the sensor, the signs of the local surface normals can be determined uniquely. The obtained local surface normals are distributed on a sphere, and we represented the spherical space by two ring-like regions on the sphere as shown in Figure 2.

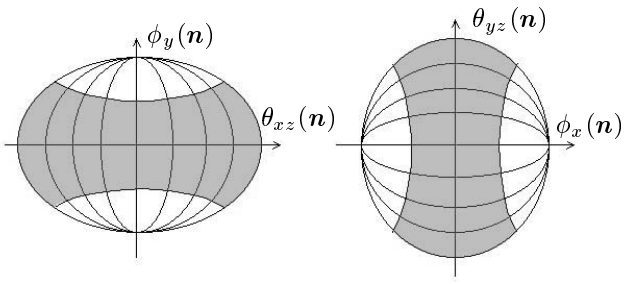


Figure 2. Histogram space of the local normals

Here $\theta_{xz}(\boldsymbol{n})$ is the yaw angle of the vector in xz plane which is projected from local surface normal $\boldsymbol{n}$. $\phi_y(\boldsymbol{n})$ is the pitch angle in the y direction of $\boldsymbol{n}$. $\theta_{xz}(\boldsymbol{n})$ and $\phi_y(\boldsymbol{n})$ are computed as follows.

$$\theta_{xz}(\boldsymbol{n}) = \begin{cases} \arccos(\frac{n_x}{\sqrt{n_x^2+n_z^2}}) & \text{if } n_z \leq 0 \\ \arccos(\frac{n_x}{\sqrt{n_x^2+n_z^2}}) + \pi & \text{otherwise} \end{cases} \quad (8)$$

$$\phi_y(\boldsymbol{n}) = \arcsin(n_y) \quad (9)$$

Similarly, $\theta_{yz}(\boldsymbol{n})$ and $\phi_x(\boldsymbol{n})$ are computed as follows.

$$\theta_{yz}(\boldsymbol{n}) = \begin{cases} \arccos(\frac{n_y}{\sqrt{n_y^2+n_z^2}}) & \text{if } n_z \leq 0 \\ \arccos(\frac{n_y}{\sqrt{n_y^2+n_z^2}}) + \pi & \text{otherwise} \end{cases} \quad (10)$$

$$\phi_x(\boldsymbol{n}) = \arcsin(n_x) \quad (11)$$

In the overlap region in the two rectangular coordinate systems, there will exist two equivalent distributions which correspond to the same distribution in the original spherical space. In this case, we use any of them for region extraction, and the equivalent distribution in the other rectangular coordinate system will be extracted simultaneously.

Because the latticed squares in the rectangular histogram spaces are projected from trapezoid-like patches with different areas in the original spherical space, we scaled the histogram along the ϕ axis in the rectangular coordinate systems to preserve the homogeneity of the voting density of the histogram, by the following formula:

$$h'(\theta_*, \phi_*) = \frac{1}{sin(\phi_*)} h(\theta_*, \phi_*) \quad (12)$$

where $h(\theta, \phi)$ stands for the histogram, and (θ_*, ϕ_*) should be substituted by (θ_{xz}, ϕ_y) or (θ_{yz}, ϕ_x) in the actual computation.

3.2 Estimation of distribution of local normals

First, the highest peak $Peak_{max}$ is detected from the two histogram spaces of Fig. 2, and a plane is assumed to exist in the range data, whose local surface normals formed the detected peak. Then, the range data points $\{P_{i_m j_m}\}$ whose local surface normals are equal to $Peak_{max}$ are detected from the range image. When $P_{i_m j_m}$ belongs to a plane, its neighborhood $N(P_{i_m j_m})$, including $P_{i_m j_m}$, is much likely to belong to the same plane. Therefore, we can use $N(P_{i_m j_m})$ as sample points to estimate mean vector M and the covariance C of the distribution of the local surface normals of the plane.

Using the estimated distribution of the local surface normals of the plane with respect to $Peak_{max}$, we can extract its corresponding region(s) in the range image. First, we extract the points whose local surface normals are located within σ interval of the distribution. That is, the extracted points $\{P_r\}$ satisfy the following condition:

$$d_r = (\boldsymbol{n}(P_r) - M)^\top C^{-1} (\boldsymbol{n}(P_r) - M) < 1.0 \quad (13)$$

where $\boldsymbol{n}(P_r)$ is the local surface normal at P_r.

The extracted points include nearly $2/3$ of the population of the distribution, and provide a dense sampling of the plane. Based on these points, we perform a dilation operation with a constraint that the local surface normal of the extended pixel P_e must be within 4σ range from the center of the estimated distribution, that is, $d_e < 4.0$, which corresponds to the interval containing 99.9% of the population of the distribution. After the constrained dilation, the significant regions whose area is bigger than a threshold T are extracted. In our experiments, we adopted a hierarchical plane extraction strategy which set T to different values at each stage.

3.3 Hierarchical strategy for plane extraction

When there are both planar and curved surfaces in the range data, as a curved surface has variable local surface normals, a part of its local surface normals may be consistent with that of a given plane. In this case, a part of the curved surface will be extracted along with the plane. Therefore, the approximation of curved surfaces by planar patches will be influenced by the planes which exist in the range data. The similar thing may also occur on planes which have similar normals.

To reduce the undesired extraction of small regions described above, we adopted a hierarchical strategy for plane extraction. It consists of three levels. At the top level, only the planar regions with area greater than 1600 pixels are extracted and the smaller regions are left to the next stage. The resolution of the histogram space of local surface normals is set to 1°. In the middle and bottom levels, planar regions greater than 800 and 400 pixels, respectively, are extracted. The resolution of the histogram space is set to 2° to deal with a larger variance of the local surface normals of the patches of some curved surfaces.

3.4 Region merging

Although our method can extract regions properly under a quite hard condition of overlapping of the distribution, there may be some planes which are divided to some parts. In this research, we use the Wilcoxon test to check if two neighboring regions can be merged together.

Let $R_1 = \{P_i^{(1)}\}$ and $R_2 = \{P_j^{(2)}\}$ be two neighboring regions. Let $\{Es_i^{(1)}\}$ and $\{Es_j^{(2)}\}$ be the fitting errors of regions R_1 and R_2, respectively, using their own plane equations, and $\{Eo_i^{(1)}\}$ and $\{Eo_j^{(2)}\}$ be that of R_1 and R_2, using the plane equation of the other region. Then we apply Wilcoxon test to test whether $\{Es_i^{(1)}\}$ and $\{Eo_j^{(2)}\}$ come from the same distribution. The same is done for $\{Es_i^{(2)}\}$ and $\{Eo_j^{(1)}\}$. If one of the tests succeeds, then we regard (R_1, R_2) as a candidate pair of regions which can be possibly merged. The details of the Wilcoxon test are omitted as they can be found easily in some statistics textbooks.

Suppose that $\{Es_i^{(1)}\}$ and $\{Eo_j^{(2)}\}$ passed Wilcoxon test, we compute the variance σ_1 of $Es_i^{(1)}$ and the mean $m2$ of $Eo_j^{(2)}$. If $m_2 \leq \sigma_1$, we merge R_1 and R_2.

4 Polyhedral Description

The planar regions extracted from the range data are surrounded by edges of three kinds: boundary edges indicate the boundary of the data, jump edges indicate a significant variation of depth in neighboring pixels, and crease (or roof) edges are formed by two neighboring planar regions. While the crease edge lines can be computed directly from the plane equations of two neighboring planar regions, the edges of the other two kinds, which are called discontinuity edges in the rest of this paper, must be transformed into line segments in order to represent each planar region with a polygon. In a rectangular image coordinate system, the Hough transform can be applied for line extraction, but in the panoramic range data of a 360° field of view, the range data is represented in a polar coordinate system, in which a 3D straight line is projected as a curve. In this section, we first create an ordinary rectangular image coordinate system for each region and establish a one-to-one mapping relation between the two coordinate systems. Then we propose a modified Hough transform which uses the rectangular image coordinate system for voting and peak detection, and the polar coordinate system for line segment extraction. Finally, a polyhedral description of the scene is built by using the extracted line segments of discontinuity edges and the crease edges.

4.1 Correspondence between polar and rectangular coordinate systems

Let the world coordinate system be denoted as (X_w, Y_w, Z_w). The local coordinate system for a planar region R is set to an orthogonal coordinate system whose Z-axis is equal to the normal of R, and it is denoted as (X_r, Y_r, Z_r). The two coordinate systems have the following relation:

$$(X_r, Y_r, Z_r)^\top = R_r (X_w, Y_w, Z_w)^\top \tag{14}$$

where R_r is a 3×3 rotation matrix which transforms the normal of R to the Z-axis in the world coordinate system.

The problem here is to establish a one-to-one correspondence between the lattice points (i_p, j_p) of the rectangular image plane in (X_r, Y_r, Z_r) coordinate system and the spherical lattice points (i_s, j_s) of the polar coordinate system in (X_w, Y_w, Z_w). If we map (i_p, j_p) and (i_s, j_s) to the image plane of $Z = 1$ in (X_r, Y_r, Z_r) and the image sphere with radius of 1 in (X_w, Y_w, Z_w), respectively, as shown in Eqs. (15) and (16), then their correspondence relation can be obtained easily using Equation (14):

$$\begin{pmatrix} x_p \\ y_p \\ z_p \end{pmatrix} = \begin{pmatrix} (j_p - j_0)/s_x \\ (i_p - i_0)/s_y \\ 1 \end{pmatrix} \tag{15}$$

where (i_0, j_0) is the image center, and s_x and s_y are the horizontal and vertical scaling factors, respectively, which map the coordinates in the standard image plane to the lattice points in an image.

$$\begin{pmatrix} x_s \\ y_s \\ z_s \end{pmatrix} = \begin{pmatrix} cos(j_s/s_a)cos(i_s/s_a + \phi_{min}) \\ sin(i_s/s_a + \phi_{min}) \\ sin(j_s/s_a)cos(i_s/s_a + \phi_{min}) \end{pmatrix} \tag{16}$$

where s_a is the angle resolution and ϕ_{min} is the minimum of the pitch angle in range data.

Using Eqs. (16) to (18), we can map (i_s, j_s) to (i_p, j_p) as follow.

$$\begin{pmatrix} j_p \\ i_p \end{pmatrix} = \begin{pmatrix} s_x \frac{r_{11}x_s + r_{12}y_s + r_{13}z_s}{r_{31}x_s + r_{32}y_s + r_{33}z_s} + j_0 \\ s_y \frac{r_{21}x_s + r_{22}y_s + r_{23}z_s}{r_{31}x_s + r_{32}y_s + r_{33}z_s} + i_0 \end{pmatrix} \quad (17)$$

where r_{ij} is the ijth element in the rotation matrix R_r, and x_s, y_s and z_s are represented as functions of i_s and j_s, which are defined in Eq (18).

A straight line $aj_p + bi_p + c = 0$ in the rectangular image (i_p, j_p) will be mapped to a curve in the sphere image (i_s, j_s) of the polar coordinate system as follows.

$$as_x(r_{11}x_s + r_{12}y_s + r_{13}z_s) + bs_y(r_{21}x_s + r_{22}y_s + r_{23}z_s) + (c + aj_0 + bi_0)(r_{31}x_s + r_{32}y_s + r_{33}z_s) = 0 \quad (18)$$

where x_s, y_s and z_s should be substituted by functions defined on i_s and j_s in Eq (16).

4.2 Hough transform of discontinuity edges

We use Eqs. (17) and (18) to develop a modified Hough transform algorithm which extracts curve segments corresponding to straight lines in the orthogonal coordinate system from the spherical image in the polar coordinate system.

Let us consider the discontinuity edges of region R. The discontinuity edge pixels $\{(i_s, j_s)\}$ in the spherical image in the polar coordinate system are transformed to $\{(i_p, j_p)\}$ in the rectangular image, using Eq. (17). Then, we use $\{(i_p, j_p)\}$ to vote in the Hough space (θ_h, ρ_h), and detect the highest peak $(\theta_{peak}, \rho_{peak})$ in the Hough space. The peak $(\theta_{peak}, \rho_{peak})$ corresponds to a straight line of $cos(\theta_{peak})j_p + sin(\theta_{peak})i_p - \rho_{peak} = 0$.

Using Eq. (18), we can obtain the curve $C(i_s, j_s)$ in the spherical image in the polar coordinate system. We extract the longest curve segment S_1 from $\{(i_s, j_s)\}$, which lays within a range of width W of $C(i_s, j_s)$. In the experiments, W is set to 3 pixels.

The curve extraction processing is repeated on the remaining edge pixels until no curve longer than L pixels can be extracted. Here L is a threshold and is set to 3 pixels in our experiments.

While the curves are extracted from the spherical image in the polar coordinate system, their connective relations are also saved for the computation of their crossing points.

4.3 Crease edges

An iterative expanding operation, bounded by jump and boundary edges, is carried out on the extracted regions $\{R_i, i = 1, \cdot, r\}$ so that the neighboring regions touch each other, and the adjacent relationship $\{(R_i, R_j)\}$ between regions is found.

Let the plane equations of R_i and R_j be $a_iX + b_iY + c_iZ + d_i = 0$ and $a_jX + b_jY + c_jZ + d_j = 0$. The equation of the crease edge of the two neighboring regions (R_i, R_j) in the image sphere of the polar coordinate system can be computed directly.

By eliminating the constant items of d_i and d_j of the two plane equations of R_i and R_j, we obtain the following equation.

$$(a_id_j - a_jd_i)X + (b_id_j - b_jd_i)Y + (c_id_j - c_jd_i)Z = 0 \quad (19)$$

Dividing two sides of Eq. (19) by $\sqrt{X^2 + Y^2 + Z^2}$, we get the curve equation of the crease edge in the sphere image of the polar coordinate system as follow.

$$(a_id_j - a_jd_i)x_s + (b_id_j - b_jd_i)y_s + (c_id_j - c_jd_i)z_s = 0 \quad (20)$$

where x_s, y_s and z_s are represented as functions defined on sphere image coordinates (i_s, j_s) in Eq (16).

For any pair of neighboring regions, if the crease edge computed above does not pass through near the adjacent area of the two regions, the crease edge is judged as a false crease edge and the adjacent boundary is regarded as a discontinuity edge, which is passed to the Hough transform algorithm described above.

4.4 Building polyhedral description

A polyhedral description of the scene can be easily built by using the extracted line segments of discontinuity edges and the crease edges. As the connective relations can be detected from the edge pixels in the sphere image in the polar coordinate system, the cross points of the edges, i.e., the vertices of the polygons of planar surfaces, can be computed easily. Using the plane equation of each planar region, the 3D coordinates of the vertices can be obtained.

The extracted polygon planar regions and their neighboring relations are described by an attributed graph denoted as:

$$G = \{\mathcal{R}, F, \mathcal{E}, H\}$$

where $\mathcal{R} = \{R_i\}$ is the set of nodes which corresponds to polygon regions in the range image, $F = \{(f_1(R_i), ..., f_m(R_i))\}$ denotes the set of feature vectors which characterizes the regions. In our work, F denotes the coefficients of the plane equations of the regions. $\mathcal{E} = \{E_i\}$ is the set of edges of the polygon regions, and $H = \{(h_1(E_j), ..., h_n(E_j)\}$ denotes the set of feature vectors which characterizes the edges, including crease edges or discontinuity edges, edge directions and the two endpoints of the edges.

5 Experimental results

We applied the proposed method to the range data of some scenes taken by SceneModeler. In this section, we present the experimental results for one of these scenes.

Figure 3 shows the panoramic range data acquired in our laboratory with the SceneModeler. The range data is shown

Figure 3. Panoramic range data acquired with SceneModeler.

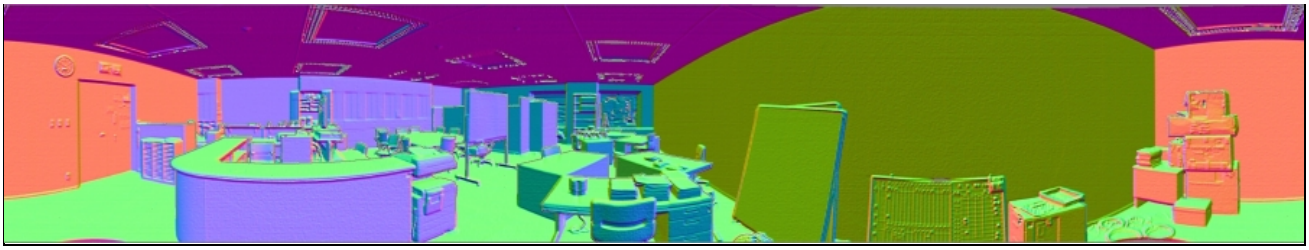


Figure 4. Local surface normals.

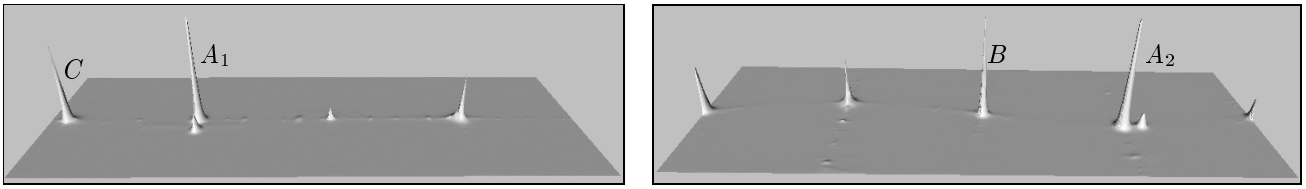


Figure 5. Histogram of local surface normals (left: histogram space of (θ_{xz}, ϕ_y); right: that of (θ_{yz}, ϕ_x)).

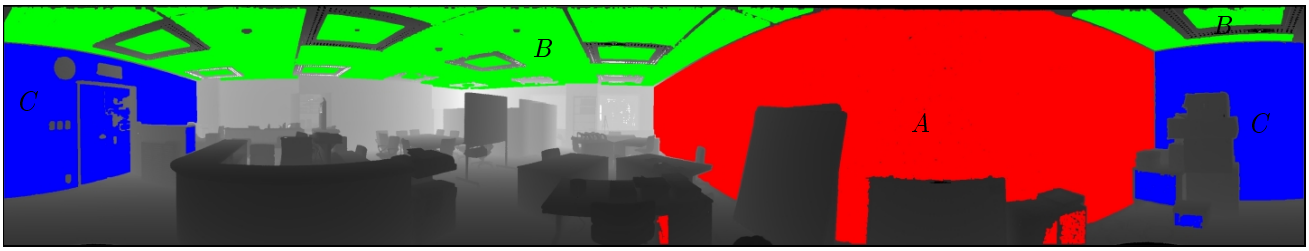


Figure 6. Regions corresponding to the three highest peaks.

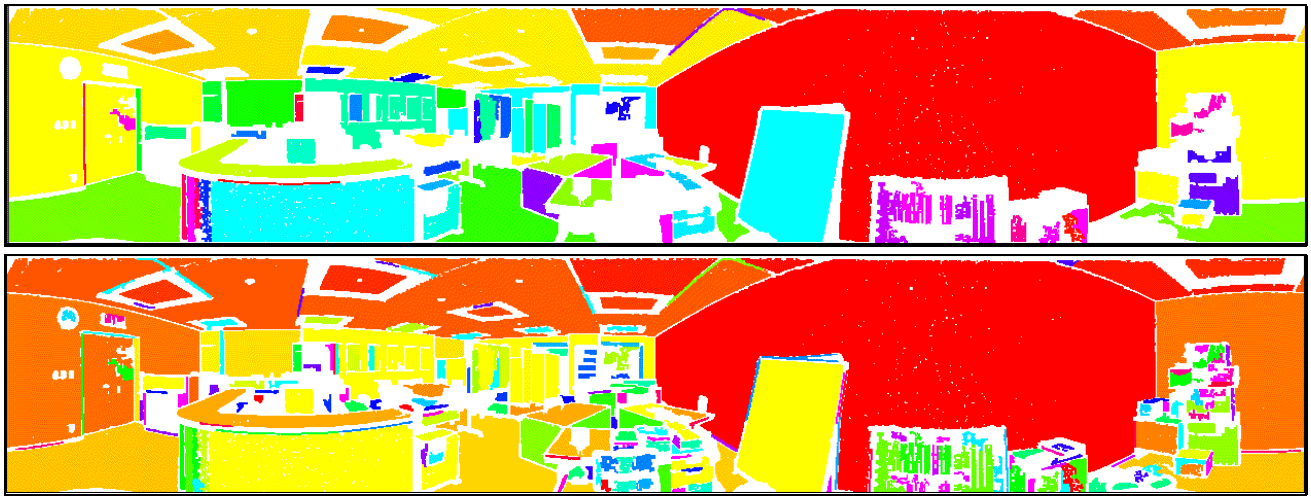
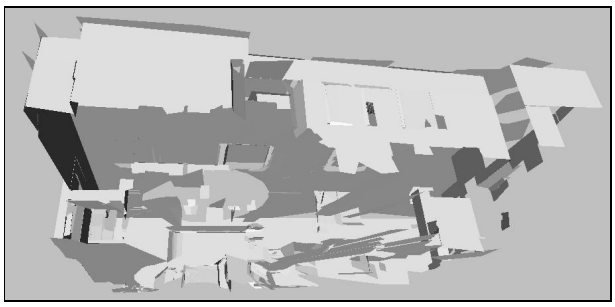


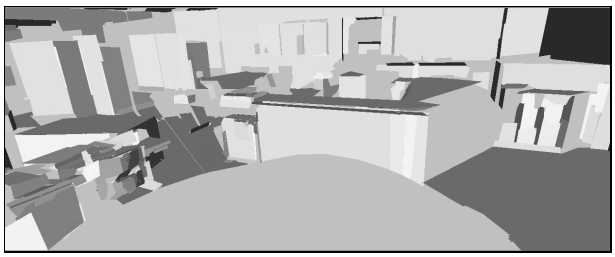
Figure 7. Regions extracted (top: at the first level; bottom: the final results).

by the grey levels of the distance of the scene from the sensor. Figure 4 shows the local surface normals computed by the method described in Section 2. The local surface normals are shown in pseudo color. Figure 7 shows the histograms, which were described in Section 3. We can see that there are two equivalent distributions A_1 and A_2 of the local surface normals in the two spaces, due to the overlap between them. Figure 6 shows the planar regions of red, green and blue colors which correspond to the first, second and third highest peaks, respectively.

The top image in Figure 7 shows the planar regions which are extracted at the first level, where the labels of the regions are shown in pseudo-color. The bottom image in Figure 7 shows the final results of planar region extraction, which is an integration of the regions extracted at different levels. Figure 8(a) and (b) shows the 3D views from outside and inside, respectively, of the scene displayed by COSMO player using the VRML model transformed from the polyhedral description. From the experimental results, we can see that a good approximation of the polyhedral representation of range data has been acquired.



a) View from outside.



b) View from inside.

Figure 8. VRML model of polyhedral description.

6 Conclusion and future work

In this paper, we proposed a novel method to extract a polyhedral description from the panoramic range data of a scene taken by the SceneModeler laser range finder. First, we derived a simple and efficient method to approximately estimate the optimal local plane fitting under a noise model of laser radar range finder. Then, we extracted stable planar regions from the range data by using both the distribution information of local surface normals, which were computed by the proposed approximation method, and their spatial information in the range image. Finally, we proposed a modified Hough transform algorithm, which uses the rectangular image coordinate system for voting and peak detection, and the polar coordinate system for line segments extraction, to transform jump and boundary edges of the planar regions into line segments. Using those line segments together with the crease edges of neighboring regions, we built a polyhedral description of the scene. Experimental results on complex real range data show the effectiveness of the proposed method.

We are now planning to integrate the polyhedral descriptions obtained from the range data at multiple viewpoints. Texture mapping using images of omni-directional cameras also arises as a future work.

References

[1] O.D. Faugeras M. Herbert and E. Pauchon, "Segmentation of Range Data into Planar and Quadratic Patches", *Proc. of CVPR-83*, pp. 8-13, 1983.

[2] O.D. Faugeras and M. Hebert, "The Representation, Recognition, and Locating of 3-D Objects", *The International Journal of Robotics Research*, Vol. 5, No. 3, pp.27-52, 1986.

[3] R. L. Hoffman and A. K. Jain, "Segmentation and Classification of Range Data", *Trans. on PAMI*, Vol. 9, No. 5, pp.608-620, 1987.

[4] A. Hoover, G. J. Baptiste, et al., "An Experimental Comparison of Range Image Segmentation Algorithms," *IEEE Trans. on PAMI*, Vol. 18, No. 7, pp.673-689, 1996.

[5] X. Jiang and H. Bunke, "Fast Segmentation of Range Data into Planar Regions by Scan Line Grouping," *Machine Vision and Application*, Vol. 7, No. 2, pp.115-122, 1994.

[6] M. Haindl and P. Zid, "Fast Segmentation of Planar Surfaces in Range Images", *Proc. of 14th ICPR*, pp.985-987, 1998.

[7] R.W. Taylor M. Savini and A. P. Reeves, "Fast Segmentation of Range Imagery into Planar Regions", *Computer Vision, Graphics, and Image Processing*, Vol. 45, No.1, pp. 42-60, 1989

[8] M.E. Bock and C. Guerra, "A Geometric Approach to the Segmentation of Range Images", *Proc. of 3DIM*, pp.261-169, 2000.

[9] A. Hoover, D. Goldgof and K. W. Bowyer, "The space envelope: A representation for 3D scenes," *Computer Vision and Image Understanding*, Vol. 69, No. 3, pp. 310-329, 1998.

[10] A. Fitzgibbon, D. Eggert and R. Fisher, "High level CAD model acquisition from range images," *Computer Aided Design*, Vol. 29, No. 4, pp. 321-330, 1997.

[11] Y. Kanazawa and K. Kanatani, "Reliability of Fitting a Plane to Range Data," *IEICE Trans. on Inf. Syst.*, Vol. E78-D, No.12, pp. 1630-1635, 1995.

[12] C. Wang, H. Tanahashi, H. Hirayu, Y. Niwa and K. Yamamoto, "Comparison of Local Plane Fitting Methods for Range Data", *Proc. of CVPR-01*, to appear.