A Hybrid Approach to Range Image Segmentation Based on Differential Geometry[1]

NAOKAZU YOKOYA* and MARTIN D. LEVINE**

One of the most significant problems arising out of understanding range images of 3D objects is *image segmentation*. This paper describes a hybrid approach to the problem, where "*hybrid*" refers to a combination of both region- and edge-based considerations. It is assumed that the range image of objects, which may be constructed of both curved and planar surfaces, is divided into *regions*. These are meant to correspond to *surface primitives* that are homogeneous in their intrinsic differential geometric properties and do not contain discontinuities in either depth or surface orientation. The method is based on the computation of first and second partial derivatives, which are obtained by locally approximating object surfaces, using biquadratic polynomials. By computing the Gaussian and mean curvatures and examining their signs, an initial region-based segmentation is then obtained in the form of a curvature sign map. Two initial edge-based segmentations are also computed from the partial derivatives and depth values. One detects *jump edges* while the other highlights *roof edges*. The three image maps are then combined to produce the final segmentation. Experimental results are presented for both synthetic and real range data. These indicate that the proposed segmentation method is useful for describing both polyhedral and curved objects.

1. Introduction

One of the main objectives in computer vision research is to enable a machine (computer) to understand its three-dimensional (3D) environment. In recent years, sets of digital range data, which are referred to as *range images* or *depth maps*, have become widely available for the analysis of 3D objects [1, 2], thanks to the development of various active and passive range-finding techniques [3–7]. Range data directly provide geometrical information about the shape of visible 3D object surfaces, and some problems in 3D object description and recognition can therefore be solved more easily in range images than in conventional intensity images. In particular, a dense depth map, which is usually produced by an active rangefinder, is useful for analyzing curved as well as polyhedral objects. Recently, considerable attention has been paid to the problem of analyzing 3D objects by using dense depth maps.

The most significant problem in the early stages of image understanding is *image segmentation*, a process of partitioning an image into meaningful parts or extracting important image features from it. In general, image segmentation techniques are based on the common assumption that meaningful object components or surface primitives are homogeneous in image properties and that there exist abrupt changes in properties between the components. Segmentation techniques for range images can be classified into two categories: region- and edge-based approaches. A region-based approach, which is called *region segmentation*, attempts to group pixels into surface regions based on the homogeneity or similarity of image and physical properties. On the other hand, an edge-based approach, which is also referred to as *edge detection*, attempts to extract discontinuities in properties.

Region-based methods for range images usually partition the image into surface regions which can be approximated by using analytical functions. The simplest such function is a first-order polynomial [8]. In this case, object surfaces are approximated by planer patches, with the result that objects are modeled as polyhedra. However, higher-order functions, such as second-order polynomials [8–10], are needed for working with more complex curved objects. These approaches require the assumption that parts of object surfaces can be globally approximated by a particular function. More recently, an attractive idea has been proposed for locally describing surfaces by using curvatures from the point of view of *differential geometry* [11–15]. Smooth surfaces can be characterized by Gaussian and mean curvatures, and

This is a translation of the IPSJ Best Paper Award paper that appeared originally in Japanese in Transactions of IPSJ, Vol. 30, No. 8 (1989), pp. 944–953.

[1]This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grant G1840.

*Image Understanding Section, Machine Understanding Division, Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba, Ibaraki 305, Japan.

**McGill Research Centre for Intelligent Machines, McGill University, 3480 University Street, Montréal, Québec, Canada H3A 2A7.

in such an approach, surface primitives are defined by a combination of the signs of these surface curvatures [12]. Each object surface is locally classified into one of eight descriptive surface types. However, there still remains a problem. Differential geometry is a theory for smooth differentiable surfaces, whereas most object surfaces are not entirely smooth but have discontinuities. The problem is how the surface curvatures can be accurately calculated for surfaces with discontinuities.

On the other hand, edge detection techniques in range images are intended to isolate discontinuities in both depth and surface orientation. A depth discontinuity, which is called a *jump edge*, is meant to correspond to an object contour or occluding boundary. A surface orientation discontinuity, referred to as a *roof edge*, is an internal edge of an object. Whereas differentiation operators developed for intensity images are also applicable for detecting jump edges in range images, roof edges require different techniques. Typical approaches to the detection of roof edges have been based primarily on surface normal analysis [16, 17], second-order differentiation of the depth function [16], or computation of the principal curvatures of surfaces [13, 18]. The difficulties and drawbacks of edge-based approaches are that (1) it is difficult to form closed contours of objects; (2) the resulting line drawing is generally not complete and is often difficult to interpret uniquely; and (3) it is difficult to extract smooth surface changes, which are usually only available from curvature-based approaches. The shortcomings of (3) are particularly severe for the analysis of curved objects.

We propose a hybrid approach to the problem of range image segmentation, where "*hybrid*" refers to a combination of region- and edge-based considerations. The range image of a set of objects, which may have both curved and planar surfaces, is divided into regions. These regions are meant to correspond to surface primitives that are homogeneous in their intrinsic differential geometric properties and do not contain discontinuities in either depth or surface orientation. The method employs a selective local surface fit and is based on the computation of first and second partial derivatives for piecewise smooth surfaces. By computing the Gaussian and mean curvatures and examining their signs, an initial region-based segmentation is then obtained in the form of a curvature sign map. Two initial edge-based segmentations are also computed from the partial derivatives and depth values. One detects jump edges by computing differences in depth, while the other highlights roof edges by differences in surface normals. The three initial image maps are then combined to produce the final range image segmentation, which can be represented in the form of a region adjacency graph containing information about surface and edge types. The main advantages of the proposed method are (1) the accurate local estimation of view-independent surface properties for piecewise smooth surfaces and (2) the isolation of adjacent distinct surface regions of the same surface type. The first advantage is accomplished by the use of a new local surface fit technique, and the second by the integration of edge information into a region-based segmentation scheme. The method has been experimentally proven to be useful for describing both polyhedral and curved objects.

2. Computing Differential Geometric Properties of Piecewise Smooth Surfaces

In order to recognize free-formed objects by using a range image, it is important to estimate the view-independent local properties of the object surfaces. For this purpose, we employ differential geometry [19, 20] as a mathematical basis of our approach. In the following, we first give a brief introduction to the differential geometry of surfaces, especially the definition of surface normals and surface curvatures. A new technique for computing these quantities from range data is then described.

2.1 Differential Geometry of Surfaces

2.1.1 Surface Normal, Gaussian Curvature, and Mean Curvature for a Graph Surface

In this paper, we assume that a range image from a single view is provided in the form of a digital graph surface, which is sometimes referred to as the orthogonal-range type [2]. Thus the definition of a surface S in the range image is explicitly defined with respect to an image coordinate system by

$$S=\{(x, y, z(x, y)): (x, y)\in D\subseteq E^2\}, \tag{1}$$

where $z(x, y)$ is meant to be a depth value from a reference plane at a point (x, y) in the range image.

A unit surface normal $\boldsymbol{n}$ at a point (x, y) is defined to be

$$\boldsymbol{n}=\frac{1}{\sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}}\left(-\frac{\partial z}{\partial x}, -\frac{\partial z}{\partial y}, 1\right). \tag{2}$$

The Gaussian curvature K and mean curvature H at a point (x, y) of the surface S are defined as follows:

$$K=\frac{\frac{\partial^2 z}{\partial x^2}\cdot\frac{\partial^2 z}{\partial y^2}-\left(\frac{\partial^2 z}{\partial x\partial y}\right)^2}{\left\{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2\right\}^2}, \tag{3}$$

$$H=\frac{\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}+\frac{\partial^2 z}{\partial x^2}\cdot\left(\frac{\partial z}{\partial y}\right)^2+\frac{\partial^2 z}{\partial y^2}\cdot\left(\frac{\partial z}{\partial x}\right)^2-2\frac{\partial z}{\partial x}\cdot\frac{\partial z}{\partial y}\cdot\frac{\partial^2 z}{\partial x\partial y}}{2\left\{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2\right\}^{3/2}}. \tag{4}$$

See Refs. 12 and 19–21 for detailed discussion of the geometrical meaning of these quantities.

2.1.2 Classification of Surfaces According to Curvature Signs

Both Gaussian and mean curvatures are local properties of a surface that have the attractive characteristics of translational and rotational invariance; that is, at a point on a surface, these curvatures are invariant to translation and rotation of the object as long as the surface is visible [12, 20]. Besl and Jain [12] have pointed out that the signs of Gaussian and mean curvatures yield a set of eight surface primitives with desirable invariance properties including view-independency, and with sufficient power to describe visible surfaces qualitatively. Given a coordinate system in which the z axis is directed toward the viewer, eight surface primitives are defined as follows [12]:

(1) $K>0$ and $H<0\cdots$ *peak surface*;
(2) $K>0$ and $H>0\cdots$ *pit surface*;
(3) $K=0$ and $H<0\cdots$ *ridge surface*;
(4) $K=0$ and $H>0\cdots$ *valley surface*;
(5) $K=0$ and $H=0\cdots$ *flat surface*;
(6) $K<0$ and $H=0\cdots$ *minimal surface*;
(7) $K<0$ and $H<0\cdots$ *saddle ridge surface*;
(8) $K<0$ and $H>0\cdots$ *saddle valley surface*.

Such a classification of surfaces is useful for describing 3D objects from the viewpoint of object recognition. In a later section, these surface types will be employed as surface primitives into which a range image is segmented. Incidentally, it is obvious that the case of $K>0$ and $H=0$ does not occur, since K and H must satisfy the inequality $K\leq H^2$ from the definition of surface curvatures. (The area defined by $K>H^2$ in (K, H) space is called the prohibitive zone [21].)

2.2 Computing Differential Geometric Properties by Using a Selective Local Surface Fit

This section proposes a new technique for locally estimating the differential geometric properties of a surface, especially for computing Gaussian and mean curvatures. It should be noted that differential geometry is a theory of smooth differentiable surfaces. However, objects are usually not entirely smooth over all their surfaces, but are piecewise smooth. The problem is how to estimate curvature properties accurately for piecewise smooth surfaces. The proposed method has three steps: (1) fitting of an analytical function to a local window centered at an individual point; (2) determination of the best window orientation for each point; (3) computation of surface curvatures by using the selected best window.

2.2.1 Polynomial Approximation of Local Windows

At each pixel in the range image, we locally determine a continuous differentiable function that fits the underlying depth surface. Note that what is considered here is not a global fit but a local one. We can use a single valued function $z=z(x, y)$ as an approximating function, since the range data are assumed to be given in the form of a graph surface as defined by Eq. (1). At least second-order differentiability is needed for the function, as is obvious from the definition of curvatures, and it should be analytically simple. The surface fitting must not be sensitive to noise and quantization effects in the range images. Generally, higher-order functions fit the given data well; however, this implies that such functions will fit not only an object shape but also the undesirable noise effects. These considerations lead us to choose the following biquadratic polynomial, which includes six independent parameters:

$$z(x, y)=ax^2+by^2+cxy+dx+ey+f. \tag{5}$$

We employ a local surface fit within a $(2m+1)\times(2m+1)$ window centered at each point (x, y) and denoted by $W(x, y)$. The coefficients a–f of the polynomial are determined for each window $W(x, y)$ by using a standard least-squares method in which the sum of the squared fitting error is minimized within the window. The resulting convolution operators for determining the coefficients are shown in Fig. 1, where the window size is assumed to be $5\times5(m=2)$. The computed coefficients of the fitted polynomial are associated with each window $W(x, y)$ and therefore represented as $a(x, y)$, $b(x, y)$, $c(x, y)$, $d(x, y)$, $e(x, y)$, and $f(x, y)$, respectively. See Ref. 21 for other sets of operators for different values of m.

$$a:\ \frac{1}{70}\begin{bmatrix}2&-1&-2&-1&2\\2&-1&-2&-1&2\\2&-1&-2&-1&2\\2&-1&-2&-1&2\\2&-1&-2&-1&2\end{bmatrix}\qquad b:\ \frac{1}{70}\begin{bmatrix}2&2&2&2&2\\-1&-1&-1&-1&-1\\-2&-2&-2&-2&-2\\-1&-1&-1&-1&-1\\2&2&2&2&2\end{bmatrix}$$

$$c:\ \frac{1}{100}\begin{bmatrix}-4&-2&0&2&4\\-2&-1&0&1&2\\0&0&0&0&0\\2&1&0&-1&-2\\4&2&0&-2&-4\end{bmatrix}\qquad d:\ \frac{1}{50}\begin{bmatrix}-2&-1&0&1&2\\-2&-1&0&1&2\\-2&-1&0&1&2\\-2&-1&0&1&2\\-2&-1&0&1&2\end{bmatrix}$$

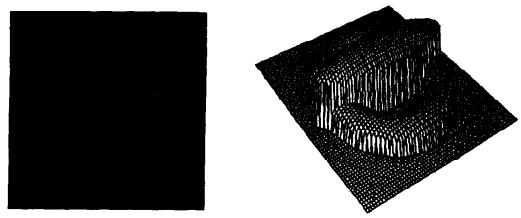
$$e:\ \frac{1}{50}\begin{bmatrix}2&2&2&2&2\\1&1&1&1&1\\0&0&0&0&0\\-1&-1&-1&-1&-1\\-2&-2&-2&-2&-2\end{bmatrix}\qquad f:\ \frac{1}{175}\begin{bmatrix}-13&2&7&2&-13\\2&17&22&17&2\\7&22&27&22&7\\2&17&22&17&2\\-13&2&7&2&-13\end{bmatrix}$$

Fig. 1 A set of operators for local biquadratic surface fit using a 5×5 window.

2.2.2 Selecting the Best Window

Previous researchers have adopted similar ideas on local surface fitting in order to obtain differentiation operators for range data [12, 14, 22]. However, such operators tend to blur or deform the original shape of an object, because they are uniformly applied even in the vicinity of discontinuities. In particular, the blurring of a depth discontinuity creates a distorted surface shape and therefore generates spurious shape descriptions [23].

The above considerations make it clear that the surface fit window must not overlap a discontinuity in the range data. We propose a method for selecting the best window that does not overlap a discontinuity. A point in a range image is covered by $(2m+1)\times(2m+1)$ win-



(a) Synthetic range image (b) Wireframe representation

(c) Squared errors of local surface fit

Fig. 2 Distribution of the least squared errors of local biquadratic surface fit.

dows, since the window size is assumed to be $(2m+1) \times (2m+1)$. For each point (x, y), we select the best window that provides a minimum fitting error among those windows covering the point (x, y). This method is based on the observation that the surface fitting error increases in the neighborhood of a discontinuity, as is shown in Fig. 2. Figure 2(a) is a synthetic range image that contains an ellipsoid and a partially occluded torus on a flat background. The magnitude of the fitting error in Fig. 2(c) is proportional to the intensity. Note that the synthetic range image in Fig. 2(a) is frequently used throughout this paper to illustrate the performance of algorithms at each stage.

Once the best window $W(x-u, y-v)$ has been determined, the surface fit at a point (x, y) is represented by a set of coefficients $\{a(x-u, y-v), b(x-u, y-v), c(x-u, y-v), d(x-u, y-v), e(x-u, y-v), f(x-u, y-v)\}$, where $(-u, -v)$ is a displacement (offset) from the point (x, y) to the center of the selected best window. This surface fit may be referred to as the *selective local surface fit*. While the basic idea behind the method comes from slope facet model smoothing of intensity images [24], this is the first application of the idea in range images.

2.2.3 Computing Derivatives and Curvatures

We discuss the computation of surface properties at an individual point in a range image. At the outset, the first and second partial derivatives are estimated in accordance with the selective local surface fit described above. The surface curvatures are then calculated from the partial derivative estimates.

By using the selective local surface fit and symbolically differentiating Eq. (5), the first and second partial

$$\frac{\partial z}{\partial x}: \frac{1}{700}\begin{bmatrix} 40u-28v-28 & -20u-14v-14 & -40u & -20u+14v+14 & 40u+28v+28 \\ 40u-14v-28 & -20u-7v-14 & -40u & -20u+7v+14 & 40u+14v+28 \\ 40u-28 & -20u-14 & -40u & -20u+14 & 40u+28 \\ 40u+14v-28 & -20u+7v-14 & -40u & -20u-7v+14 & 40u-14v+28 \\ 40u+28v-28 & -20u+14v-14 & -40u & -20u-14v+14 & 40u-28v+28 \end{bmatrix}$$

$$\frac{\partial z}{\partial y}: \frac{1}{700}\begin{bmatrix} -28u+40v+28 & -14u+40v+28 & 40v+28 & 14u+40v+28 & 28u+40v+28 \\ -14u-20v+14 & -7u-20v+14 & -20v+14 & 7u-20v+14 & 14u-20v+14 \\ -40v & -40v & -40v & -40v & -40v \\ 14u-20v-14 & 7u-20v-14 & -20v-14 & -7u-20v-14 & -14u-20v-14 \\ 28u+40v-28 & 14u+40v-28 & 40v-28 & -14u+40v-28 & -28u+40v-28 \end{bmatrix}$$

$$\frac{\partial^2 z}{\partial x^2}: \frac{1}{35}\begin{bmatrix} 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \end{bmatrix} \qquad \frac{\partial^2 z}{\partial y^2}: \frac{1}{35}\begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ -1 & -1 & -1 & -1 & -1 \\ -2 & -2 & -2 & -2 & -2 \\ -1 & -1 & -1 & -1 & -1 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix}$$

$$\frac{\partial^2 z}{\partial x \partial y}, \frac{\partial^2 z}{\partial y \partial x}: \frac{1}{100}\begin{bmatrix} -4 & -2 & 0 & 2 & 4 \\ -2 & -1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & -1 & -2 \\ 4 & 2 & 0 & -2 & -4 \end{bmatrix}$$

Fig. 3 A set of 5×5 operators for estimating the first and second partial derivatives by using the selective local surface fit.

derivative estimates are obtained at a point (x, y) as follows:

$$\frac{\partial z}{\partial x}(x, y) = 2a(x-u, y-v)u + c(x-u, y-v)v + d(x-u, y-v) \tag{6}$$

$$\frac{\partial z}{\partial y}(x, y) = 2b(x-u, y-v)v + c(x-u, y-v)u + e(x-u, y-v) \tag{7}$$

$$\frac{\partial^2 z}{\partial x^2}(x, y) = 2a(x-u, y-v) \tag{8}$$

$$\frac{\partial^2 z}{\partial y^2}(x, y) = 2b(x-u, y-v) \tag{9}$$

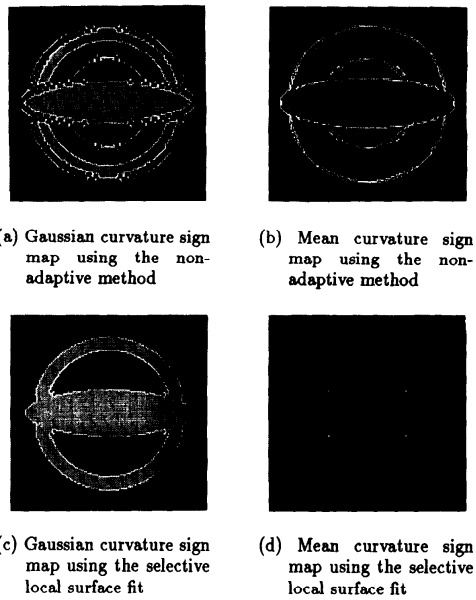
$$\frac{\partial^2 z}{\partial x \partial y}(x, y) = \frac{\partial^2 z}{\partial y \partial x}(x, y) = c(x-u, y-v), \tag{10}$$

where the offset from the point (x, y) to the center of its best window is assumed to be $(-u, -v)$. Substituting the oprators in Fig. 1 into Eqs. (6)-(10), we obtain a set of the derivative estimators that include the offset $(-u, -v)$. A set of 5×5 finite difference operators is shown in Fig. 3. Note that the operators with no offsets, $(-u, -v) = (0, 0)$, are equivalent to the second-order operators proposed by Beaudet [25].

The direct execution of differentiation operators in Fig. 3 with a range image is computationally expensive because it requires the determination of the offset $(-u, -v)$ at every pixel. A practical method is as follows: (1) the operators in Fig. 1 are convolved with the image,

and the computed coefficients a-f and the local surface fit error are stored for each pixel; (2) the offset $(-u, -v)$ is determined at each pixel by searching for the minimum number of surface fit errors within a window; and (3) Eqs. (6)-(10) are evaluated in order to determine the partial derivatives. This method is composed of the convolution, min-filter, and evaluation of at most first-order polynomials, and can be efficiently carried out by using special hardware.

Using the first and second partial derivatives estimated above, the Gaussian and mean curvatures are computed from Eqs. (3) and (4), respectively. Unit surface normals are also computed from Eq. (2). Figure 4 shows that the selective local surface fit method improves the curvature estimates when compared to the fixed non-adaptive window method. The Gaussian and mean curvature sign maps in Fig. 4 were obtained by assigning $K=0$ when $|K| \leq 0.00001$ and $H=0$ when $|H| \leq 0.003$. The curvature sign maps represent positive curvature in white, zero curvature in gray, and negative curvature in black. It can be observed in Fig. 4 that surface points are correctly characterized by using the proposed surface fit technique, even in the neighborhood of discontinuities.



(a) Gaussian curvature sign map using the non-adaptive method

(b) Mean curvature sign map using the non-adaptive method

(c) Gaussian curvature sign map using the selective local surface fit

(d) Mean curvature sign map using the selective local surface fit

Fig. 4 Comparison of the selective local surface fit with the conventional non-adaptive window method for computing curvature sign maps.

3. Segmentation and Description of Range Images

3.1 Overview of the Approach

We wish to partition a range image into meaningful surface regions without using any domain-specific knowledge. Special types of objects and their shapes are not assumed here. A scene to be analyzed may be composed of several objects with occlusions, and each object may be constructed of both curved and planar surfaces. The only assumption used here is that an object surface is partially composed of smooth differentiable surfaces (piecewise second-order differentiability). The problem to be discussed here is how to define a domain-independent segmentation of range data; in other words, what is a meaningful surface region in the range data? The surface region should be a geometrical primitive and must provide important cues for recognizing 3D objects. Considering the above, we define the surface region to be the largest four-connected component that satisfies the following two conditions:

(1) The signs of the Gaussian and mean curvatures are constant within the surface region.

(2) The region does not contain any significant discontinuities in either depth or surface orientation.

In order to accomplish the segmentation defined above, we propose a new segmentation method that consists of the following three major steps:

(I) *Local Surface Characterization*: The first and second partial derivatives are first estimated at each surface point. These are accurately computed not only on a smooth surface but also in the neighborhood of a discontinuity, using the selective local surface fit described in Section 2.2. By using these quantities, the surface normal and surface curvatures are then calculated.

(II) *Initial Segmentation*: Three kinds of initial segmentation maps are computed; a region-based segmentation in the form of a curvature sign map (*KH-sign map*) suggested in Section 2.1.2 and two edge-based segmentations embodying jump and roof edge maps. Essentially, these maps can be computed in parallel.

(III) *Final Segmentation*: The three initial segmentation maps are combined to produce the final segmentation, in which each region must satisfy the surface region conditions mentioned earlier.

Figure 5 illustrates the overall control flow of the segmentation algorithm. Since the first step has been described in detail earlier, the following discussion focusses on steps (II) and (III).

3.2 Region-Based Initial Segmentation

We recall that each point on a surface can be classified into one of eight possible surface primitives according to the signs of the Gaussian (K) and mean (H) curvatures. These surface primitives are view-independent as long as the surface is visible. The theoretical aspects and practical computational schemes for this approach have been fully discussed in Section 2.

However, in a practical sense, thresholding about zero is required in order to obtain the curvature sign map. For example, the real range image of a flat surface does not yield Gaussian and mean curvature values that are exactly equal to zero. This is usually caused by noisy variations in depth value and quantization effects. The

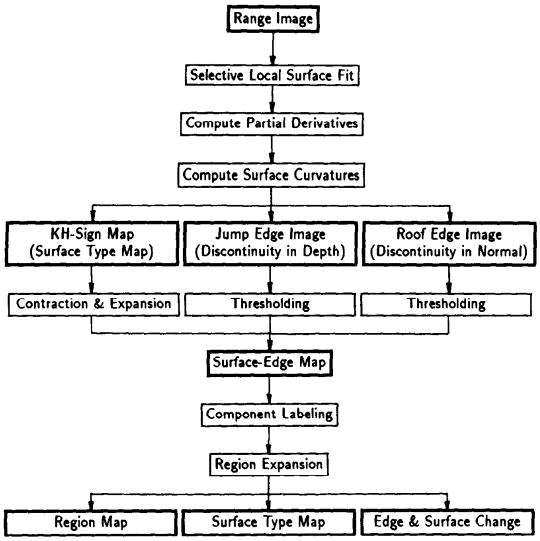
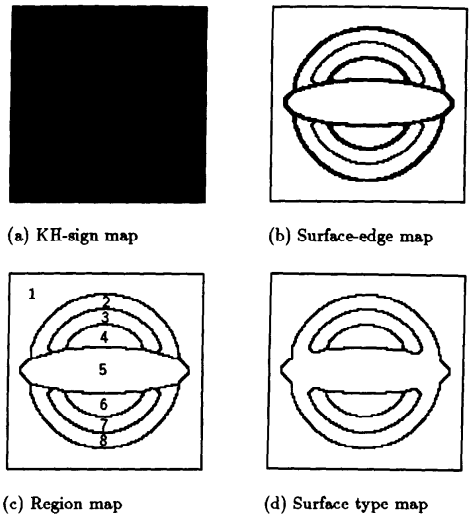


Fig. 5 The overall control flow for the range image segmentation algorithm.



(a) KH-sign map (b) Surface-edge map

(c) Region map (d) Surface type map

Fig. 6 Results at each stage of region segmentation.

thresholds about zero for the Gaussian and mean curvatures are selected symmetrically, and are referred to as K_0 and H_0, respectively; that is, we assign $K=0$ when $-K_0 \le K \le K_0$ and $H=0$ when $-H_0 \le H \le H_0$. From the definition of K and H, we use a constraint on the selection of K_0 and H_0 as follows: (1) $K_0 \approx H_0^2$; (2) $K_0 \ge H_0^2$. The second condition comes from the fact that if the inequality is not satisfied, some surface points will fall into the prohibitive zone in (K, H) space.

In order to erase small surface regions regarded as noise, the KH-sign map is contracted and then expanded. First, each surface region in the map is contracted by one pixel, considering four-connectedness for pixel connectivity. The surface region is then iteratively expanded until all unlabeled pixels are eliminated. The refined KH-sign map at this stage is shown in Fig. 6(a). It should be noted that at this stage the surface types are correctly determined but the adjacent distinct surfaces of the same type are not discriminated.

3.3 Edge-Based Initial Segmentation

3.3.1 Detecting Jump Edges

For each point in the range image, the jump edge magnitude, which is denoted by M_{jump}, is computed as the maximum difference in depth between the point and its eight neighbors. This is formulated for a point (x, y) as follows:

$$M_{jump}(x, y) = \max\{|z(x, y) - z(x+k, y+l)| : -1 \le k, l \le 1\}, \quad (11)$$

where z represents the value of the fitted polynomial computed at the point (x, y). The edge image is thresholded to produce the *jump edge map*. At present, the threshold is determined according to the mean and standard deviation of the edge magnitude.

3.3.2 Detecting Roof Edges

A roof edge can be located by surface normal analysis. At each point in the range image, the roof edge magnitude, denoted by M_{roof}, is computed as the maximum angular difference between adjacent unit surface normals. This is formally given as

$$M_{roof}(x, y) = \max\left\{ Cos^{-1}\left(\frac{\boldsymbol{n}(x, y) \cdot \boldsymbol{n}(x+k, y+l)}{|\boldsymbol{n}(x, y)| \cdot |\boldsymbol{n}(x+k, y+l)|} \right) : -1 \le k, l \le 1 \right\}$$
$$= \max\{Cos^{-1}(\boldsymbol{n}(x, y) \cdot \boldsymbol{n}(x+k, y+l)) : -1 \le k, l \le 1\} \quad (12)$$

The unit surface normal $\boldsymbol{n}$ is computed from Eq. (2), using the first partial derivative estimates. This should also be thresholded to produce the *roof edge map* and combined with the KH-sign and jump edge maps at the final segmentation stage.

3.4 Final Segmentation

In this section, the final segmentation stage is described in detail. We present a method for integrating the three initial segmentation maps into the final segmentation. The process of generating the final region-based segmentation is composed of the following three steps: (1) superimposition of the two edge maps onto the KH-sign map; (2) component analysis of surface regions; and (3) expansion of surface regions.

3.4.1 Superimposing Edge Maps onto the KH-Sign Map

After the initial region-based segmentation, the KH-sign map has been generated in the form of a labeled image in which each point has a value of 1 to 8 corre-

sponding to the determined surface type. At this stage, the thresholded edge maps are superimposed onto the KH-sign map to produce a *surface-edge map*, in which each edge point has a value of -2 or -1 if it is regarded as a jump or roof edge, respectively. Figure 6(b) depicts the surface-edge map, in which boundaries of the KH-sign map are indicated by black lines, while jump edges are shown as deep gray and roof edges as light gray. The thresholds for the edge maps were selected as the mean plus the standard deviation of their magnitudes.

3.4.2 Component Analysis

To generate a region map, an existing four-connected component labeling algorithm is independently employed for each surface type in the surface-edge map. The edge points are not labeled with positive number and are kept unchanged. Thus, the resulting region map is not a complete segmentation; that is, it is composed of surface points (non-edge points) that have region numbers as their labels and edge points with negative labels.

3.4.3 Expansion of Surface Regions

Each surface region in the previously obtained partial region map is expanded in parallel to generate a final region map, using the following two steps. First, each surface region having a positive label is conditionally expanded so that the expansion does not cross the boundaries of surface regions in the KH-sign map. However, there still remain edge points with negative labels that correspond to the surface points entirely overlapped by edge points. Each surface region is then unconditionally expanded in parallel until all edge points are erased.

Simultaneously, to produce a final surface type map corresponding to the above region map, each surface region in the surface-edge map is expanded in the manner described above. As a consequence, the surface regions entirely overlapped by edge points are not recovered; that is, small surface regions on discontinuities are eliminated.

Figures 6(c) and (d) show boundary representations of the final region and surface type maps. It can be seen that the horizontal ellipsoid and the outside surface of the torus are perfectly discriminated in the final region map.

3.5 Description of a Segmented Range Image

We now suggest a method for describing a 3D scene by using the segmented regions and related information. The final segmentation maps contain the following rich information about the scene: (1) surface region discrimination (region number); (2) view-independent surface type (KH-sign); and (3) edge types between adjacent regions (jump edge, roof edge, or smooth surface change). From the segmentation in Fig. 6(c), these region and boundary properties and their relationships can be represented as a *region adjacency graph*, as shown in Fig. 7, in which each surface region is indicated by a node and a boundary type label is associated with an edge between the nodes.

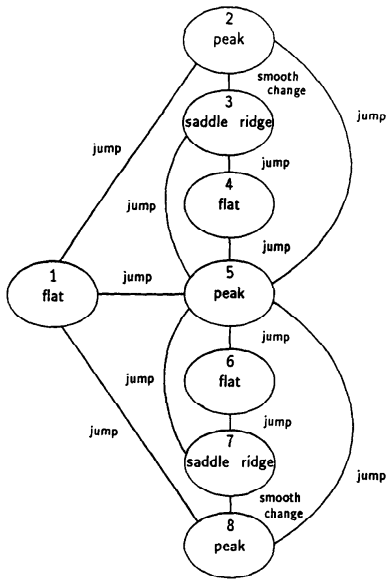


Fig. 7 A region adjacency graph for the segmented image in Fig. 6(c).

A jump edge implies an occluding contour of an object in the range image, while both roof edges and smooth surface changes correspond to internal edges of an object. If we suppose that a jump edge necessarily isolates an object from others, the segmented image in Fig. 6(c), which consists of *Regions 1–8*, can be interpreted as being composed of the following six objects: *Object 1* (*Region 1*); *Object 2* (*Regions 2* and *3*); *Object 3* (*Region 4*); *Object 4* (*Region 5*); *Object 5* (*Region 6*); *Object 6* (*Regions 7* and *8*). However, in order to conclude that *Objects 1 3*, and *5* correspond to a single background and that *Objects 2* and *6* construct a single torus, further analysis such as a hypothesis-and-test process is needed.

Only the computed surface type is associated with each region (node) in Fig. 7. The list of calculated attributes for each region, such as the surface area, may be attached to the corresponding node in the graph. Such an improvement is expected to be useful for a model-based 3D object recognition system that uses the region adjacency graph suggested here.

4. Experimental Results

We present an example of experimental results for the proposed segmentation method, using real range data. The real range images used in the experiments were obtained by using a laser rangefinder developed at the National Research Council of Canada [5]. The im-

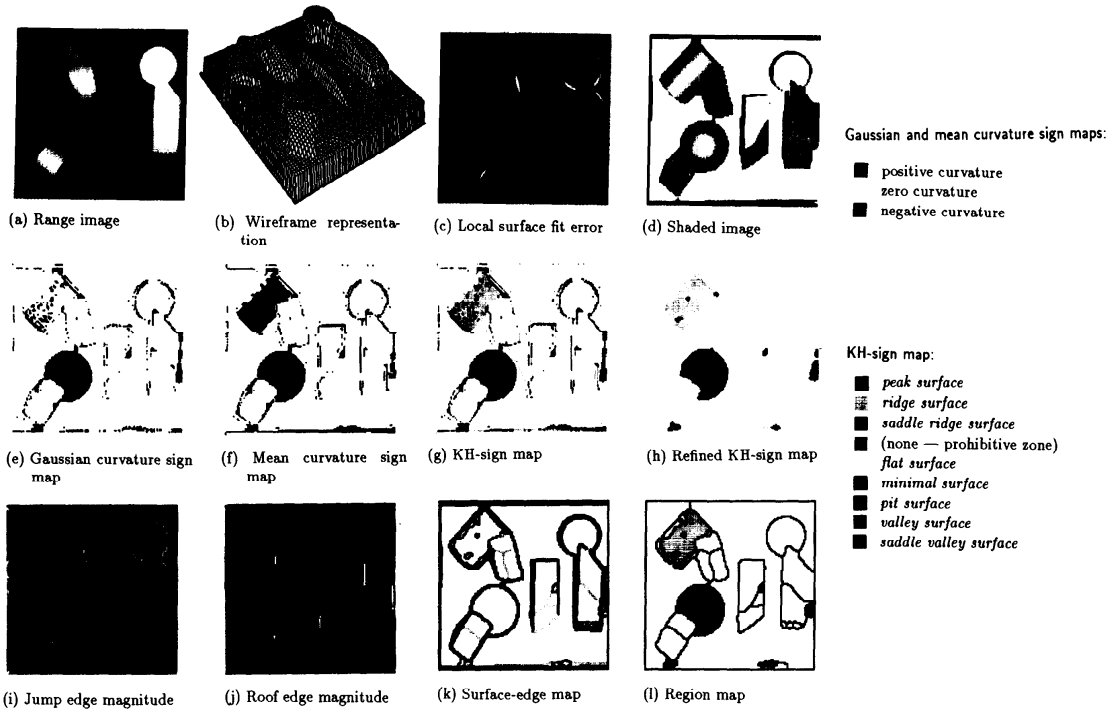


Fig. 8 Experimental results for a real range image (DYPRO1) consisting of both polyhedral and curved objects.

ages were originally obtained as a 2D array of pixels with values 12 bits long and proportional to the z values. These were then calibrated, resulting in floating-point data. The original range image contains shadow and quantization effects that cannot be completely eliminated but are rather reduced in the calibrated image. Although the typical image size is 256×256, experiments were carried out with 128×128 images obtained by resampling. In the results shown below, the first and second partial derivatives were computed by using the 5×5 operators.

Figure 8 shows experimental results for real range data (DYPRO1) of a scene consisting of both polyhedral and curved objects on a table. The differences in curvature signs and surface types are shown by the colors. Note that region boundaries are superimposed on the surface type map in Fig. 8(l). The thresholds K_0 and H_0 are set to 0.0004 and 0.02, respectively. It can be seen that the spherical and cylindrical surfaces are correctly extracted as the *peak* and *ridge surfaces*, respectively. Note that the polyhedral objects are excellently partitioned into their planar surfaces owing to the integration of the edges with the KH-sign map. However, a few small surface regions appear at the corner of planar surfaces and in the vicinity of image borders. These misclassifications are created by the impossibility of selecting windows that do not overlap any discontinuities.

5. Conclusions

In this paper, we have proposed a hybrid approach to range image segmentation. We first locally approximate a depth surface by using biquadratic polynomials, and analytically compute the first and second partial derivatives. The Gaussian and mean curvatures and the surface normal are then computed at each point, using the partial derivative estimates. By using these quantities, three kinds of initial segmentation maps are generated in parallel: (1) a KH-sign map, (2) a jump edge map, and (3) a roof edge map. These image maps are then combined into the final segmentation.

The usefulness of this approach has been proven experimentally for both synthetic and real range data. The advantages of the proposed method are summarized as follows:

(1) Each surface point is accurately characterized as one of eight possible view-independent surface types, even in the vicinity of discontinuities in depth and surface orientation. This is accomplished by the use of the proposed local surface fit technique.

(2) The proposed method is useful for segmenting range data of free-formed objects, since it does not rely on the global approximation of data using special types of functions.

(3) Adjacent distinct surface regions of the same surface type can be discriminated. This advantage is due to

the integration of edge information into a region-based segmentation scheme; in other words, to the hybrid approach.

(4) In addition to segmented region maps, rich descriptions of the surface and boundary of a region are obtained. These are expected to be useful for further object recognition.

The major problems remaining can be stated as follows:

(a) At present, the thresholds K_0 and H_0 for the Gaussian and mean curvature sign maps are selected manually. Different values of the thresholds yield different KH-sign maps, and subsequently different region maps. The threshold selection technique should be investigated further.

(b) The proposed selective local surface fit usually allows accurate estimation of surface curvatures in the neighborhood of discontinuities. However, as can be seen in the experimental results, sometimes all possible windows overlap the discontinuities and wrong curvature estimates are subsequently obtained. A possible solution to this problem is to use a variable window size (the multiscale approach) or a variable window shape.

While our segmentation method will be computationally expensive on ordinary sequential machines, its efficiency could be improved by using a special-purpose image processor or a highly parallel machine, since the main part of the computational cost is for the selective local surface fit and curvature estimation.

In addition to improving the method of handling the problems listed above, we suggest two future research directions: (1) development of a model-based object recognition system using scene descriptions obtained by our method as discussed in Section 3.5, and (2) automatic object-model acquisition using range data from multiple views.

Acknowledgements

The authors wish to acknowledge the assistance of the National Research Council of Canada, who provided us with laser range images. They also would like to thank Baris I. Demir, David Gauthier, and Andre Sokalski for their helpful discussions and assistance in developing programs for this work. Naokazu Yokoya would like to thank the members of the Image Understanding Section, Electrotechnical Laboratory, for their support. Martin D. Levine would like to thank the Canadian Institute for Advanced Research for its support.

References

1. BESL, O. J. and JAIN, R. C. Three-dimensional object recognition, *ACM Computing Surveys*, **17**, 1 (March 1985), 75-145.
2. YAMAMOTO, H., TAMUNE, M. and TAMURA, H. Range finding and range data processing: A survey, *Technical Report of the Institute of Electronics, Information and Communication Engineers of Japan*, **IE 86-128** (March 1987) (in Japanese).
3. JARVIS, R. A. A perspective on range finding techniques for computer vision, *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-5**, 2 (March 1983), 122-139.
4. SUGIHARA, K. Survey: Extraction of surface structures from visual information, *SIG Notes of the IPS Japan*, **SIGCV 33-4** (Nov. 1984) (in Japanese).
5. RIOUX, M. Laser range finder based on synchronized scanners, *Applied Optics*, **23**, 21 (Nov. 1984), 3837-3844.
6. INOKUCHI, S., SATO, K. and MATSUDA, F. Range imaging system for 3-D object recognition, *Proc. 7th Int. Conf. on Pattern Recognition* (August 1984), 806-808.
7. OHTA, Y. and KANADE, T. Stereo by intra- and inter-scanline search using dynamic programming, *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-7**, 2 (March 1985), 139-154.
8. OSHIMA, M. and SHIRAI, Y. Object recognition using three-dimensional information, *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-5**, 4 (July 1983), 353-361.
9. FAUGERAS, O. D., HEBERT, M. and PAUCHON, E. Segmentation of range data into planar and quadratic patches, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (June 1983), 8-13.
10. MULLER, Y. and MOHR, P. Planes and quadrics detection using Hough transform, *Proc. 7th Int. Conf. on Pattern Recognition* (Aug. 1984), 1101-1103.
11. WAKAYAMA, T. On deriving and drawing the structural lines of curved objects, *Technical Report of the Institute of Electronics and Communication Engineers of Japan*, **IE 76-94** (March 1977) (in Japanese).
12. BESL, P. J. and JAIN, R. C. Invariant surface characteristics for 3D object recognition in range images, *Computer Vision, Graphics, and Image Processing*, **33**, 1 (Jan. 1986), 33-80.
13. VEMURI, B. C., MITICHE, A. and AGGARWAL, J. K. Curvature-based representation of objects from range data, *Image and Vision Computing*, **4**, 2 (May 1986), 107-114.
14. YANG, H. S. and KAK, A. C. Determination of the identity, position and orientation of the topmost object in a pile, *Computer Vision, Graphics, and Image Processing*, **36**, 2/3 (Nov./Dec. 1986), 229-255.
15. SHIRASAWA, H., SATO, K. and INOKUCHI, S. Recognition of continuous quadratic curved surface, *Technical Report of the Institute of Electronics, Information and Communication Engineers of Japan*, **PRU 87-15** (May 1987) (in Japanese).
16. BHANU, B., LEE, S., HO, C. C. and HENDERSON, T. Range data processing: Representation of surfaces by edges, *Proc. 8th Int. Conf. on Pattern Recognition* (Oct. 1986), 236-238.
17. TOMITA, F. and KANADE, T. A 3D vision system: Generating and matching shape descriptions in range images, *Proc. 1st Conf. on Artificial Intelligence Applications* (Dec. 1984), 186-191.
18. HERMAN, M. Generating detailed scene descriptions from range images, *Proc. 1985 IEEE Int. Conf. on Robotics and Automation* (March 1985), 426-431.
19. LIPSCHUTZ, M. M. *Theory and Problems of Differential Geometry*, McGraw-Hill, New York, 1969.
20. DO CARMO, M. P. *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, 1976.
21. YOKOYA, N. and LEVINE, M. D. Range image segmentation based on differential geometry: A hybrid approach, *McGill Research Centre for Intelligent Machines Technical Report*, **McRCIM-TR-CIM 87-16**, McGill University (Sep. 1987).
22. BRADY, M., PONCE, J., YUILLE, A. and ASADA, H. Describing surfaces, *Computer Vision, Graphics, and Image Processing*, **32**, 1 (Oct. 1985), 1-28.
23. BESL, P. J. and JAIN, R. C. Segmentation through symbolic surface descriptions, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (June 1986), 77-85.
24. PONG, T. C., SHAPIRO, L. G. and HARALICK, R. M. A facet model region growing algorithm, *Proc. IEEE Conf. on Pattern Recognition and Image Processing* (Aug. 1981), 279-284.
25. BEAUDET, P. R. Rotationally invariant image operators, *Proc. 4th Int. Joint Conf. on Pattern Recognition* (Nov. 1978), 579-583.