MVA '96 IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

3-D Modeling from a Single Sketch Image Using Genetic Algorithms

Takanori AOI * Hiroshi NAGAHASHI *
Interdisciplinary Graduate School of Science and Engineering,
Tokyo Institute of Technology

Abstract

This article presents a method for constructing a surface model automatically from a single sketch image. The sketch image must be drawn on a paper based on some rules of industrial design. First, the sketch image is digitized by a scanner and a contour image is found by preprocessing. 2-D corners are detected from the contour image and 3-D coordinates of each corner are inferred as vertices of an object. Next, an adequate wire-frame model is constructed using cubic Bézier curves by moving each inferred vertex to an optimum location; the shape of the wire-frame model is optimized using genetic algorithms(GA) so that the projection of the wire-frame model coincides with the contour image. Finally, a surface model is constructed using bicubic Bézier surfaces by interpolating the wire-frame model.

1 Introduction

In the process of designing industrial products many kinds of 3-D CAD(Computer Aided Design) systems have become popular recently. These systems have also been used in the presentation stage of industrial design. In this stage the shapes and colors of products are determined. So far, some mock-ups have been constructed out of a lot of sketch images. Constructing a lot of mock-ups is obviously time-consuming. Therefore, it has been expected to use different type of model which replace a mock-up. Though 3-D CAD systems are generally used for drawings of industrial products, it is difficult to operate them easily. Some methods like acquiring 3-D data from three orthographic views have been presented. However, sketch images are not always drawn based on fixed projective directions. In computer vision many works of analizing 3-D structure of objects from a single image have been done[1], while there are few works where a 3-D object model is constructed from a single picture like a sketch. This article presents a method for constructing a surface model automatically from a 2-D contour image obtained from a sketch. If a surface model is constructed from a sketch image, it is considered that the shapes and colors of the products are determined easily during the design process. Sketch images used are drawn based on some rules of an industrial design. The surface models are represented by bicubic Bézier surfaces. Because a parametric surface such as a Bézier surface, which are widely used in most CAD systems, are capable of describing a variety of shapes.

2 Overview of our method

A sketch image is drawn on a paper using markers and pastels based on the following rules:

- Source of light with parallel rays is placed in the upper left direction, as shown in Fig. 1. Shading and shadowing are limited only on the surfaces of an object.
- An object has emphatic, strong edges. In particular the border of an object is black.
- The projection of an object is based on the three-point perspective projection, which has vanishing points in the upper right, upper left and down direction.

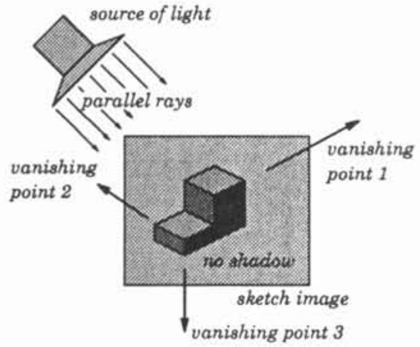


Fig. 1 Principal rules for drawing a sketch

Here we mention an overview of our method. A sketch image digitized by a scanner is differentiated, binarized and thinned to get a contour image. The process of constructing a surface model from the contour image consists of the following three stages:

---

*Address: 4259 Nagatsuta-cho, Midori-ku, Yokohama 226 Japan. E-mail: `takanori@an.ip.titech.ac.jp`, `longb@an.ip.titech.ac.jp`

1. Inferring 3-D coordinates from 2-D corners.

   From the contour 2-D corners are detected. The corners are found based on curvatures obtained by contour tracking. The contours are segmented into short curves at detected corners. The projection of a model coordinate system is defined based on mutual relations of corners. Based on the projected coordinate system, 3-D coordinates of corners are inferred.

2. Constructing an adequate wire-frame model using GA[2].

   Optimum locations of inferred vertices are supposed to exist near the inferred vertices. A wire-frame model constructed by cubic Bézier curves is projected on a 2-D plane. Comparing with the contour image the projected one is optimized by applying GA.

3. Constructing a surface model using bicubic Bézier surfaces.

   A surface model is constructed using bicubic Bézier surfaces by interpolating the wire-frame model.

3 Inferring 3-D coordinates based on a model coordinate system

An algorithm of detecting 2-D corners is described as follows:

1. One assumes that a closed region in a contour image corresponds to a surface of an object. When a first pixel is found by raster scanning, the contour tracking is started counterclockwise.

2. Sampling points $p_i(x_i, y_i)(i = 1, 2, \ldots)$ are placed on an contour at regular intervals. For all sampling points the curvatures given in equation (1) are calculated.

$$\theta_i = \tan^{-1}\frac{y_{i+1} - y_i}{x_{i+1} - x_i} \tag{1}$$

3. When some sampling points exist in the neighboring area, only the point which has the largest curvature is left.

4. A polyhedral scene is defined by tying up two adjacent sampling points and each vertex is labeled by Huffman's labeling method. The vertices which have not been labeled are deleted and new polyhedral scene is defined.

In this article all vanishing points are assumed to exist in the infinite distance. Three-point perspective projection corresponds to an axonometric projection when all vanishing points exists in the infinite distance. In axonometric projection the angles of intersection of three orthogonal coordinate axes are equiangular or unequiangular on the screen. As shown in Fig. 2, we consider that a 3-D object is projected to ZX plane after rotating angles β and α about Z and X coordinate axes, respectively. The polyhedral scene is assumed to be the projected image. Angles between edges of a polyhedron and horizontal line are computed. Angles of X, Y and Z coordinate axes for a horizontal line are considered to be $\phi_X < 90°, \phi_Y > 90°$ and $\phi_Z = 90°$. Each angle is defined by an average value of classified ones, for example, ϕ_X in equation (2) is given by an average value of angles less than 90°,

$$\phi_X = \frac{1}{n}\sum_{i=1}^{n}\phi_i \tag{2}$$

where ϕ_i denotes i^{th} angle less than 90°.

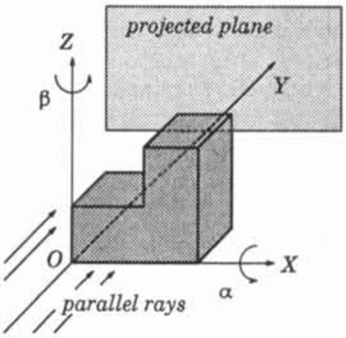


Fig. 2 Construction of axonometric projection

The vertex at the lowest location is assigned at an original point O, and X, Y, Z coordinate axes which have angles ϕ_X, ϕ_Y, ϕ_Z through a point O are defined in a polyhedral scene. Surfaces S_X, S_Y and S_Z denote surfaces perpendicular to X, Y and Z axes of the model coordinate system, respectively. Let us suppose that the polyhedron has many surfaces parallel to S_X, S_Y and S_Z, because many edges are parallel to each coordinate axis. Each surface is judged to correspond to one of S_X, S_Y and S_Z as follows:

1. The surface which corresponds to the darkest region is S_Y and one on the top of the scene is S_Z.

2. The surface which consists of edges parallel only to X and Y coordinate axes is assumed to be S_Z. S_X and S_Y are also assumed.

3. The polyhedron includes all kinds of surfaces, i.e., S_X, S_Y and S_Z.

4. When two surfaces are parallel, two regions corresponding to them have the almost same intensities.

In axonometric projection, 3-D coordinates (X, Y, Z) are transformed into 2-D coordinates (x, y) by equations (3), (4) and (5).

$$\begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha & 0 \\ 0 & \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \cos\beta & -\sin\beta & 0 & 0 \\ \sin\beta & \cos\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3}$$

$$[X_M, Y_M, Z_M] = [x, 0, -y] \tag{4}$$

$$\begin{aligned} x &= X\cos\beta + Y\sin\beta \\ y &= (X\sin\beta + Y\cos\beta)\sin\alpha - Z\cos\alpha \end{aligned} \tag{5}$$

The projection of a model coordinate system is defined in the polyhedral scene, as shown in Fig. 3. Let δ, λ be the angles between X, Y coordinate axes and a horizontal line, respectively. The rotation angles α and β about X and Z coordinate axes are calculated by equation (6).

$$\begin{aligned} \alpha &= \sin^{-1}\sqrt{\tan\delta\tan\lambda} \\ \beta &= \tan^{-1}\sqrt{\frac{\tan\delta}{\tan\lambda}} \end{aligned} \tag{6}$$

Coordinates (x, y) of a vertex on the classified surface are transformed into 3-D coordinates (X, Y, Z) using angles α and β. All 3-D coordinates of vertices for a model coordinate system are inferred using relations among all surfaces.

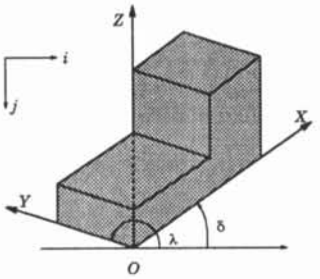


Fig. 3 Model coordinate system and angles δ, λ

4 Optimization of a wire-frame model using GA

Equations (7) and (8) show a set of vertices $P_i (i = 1, \ldots, s)$ and edges $l_j (j = 1, \ldots, t)$, which result in a wire-frame model.

$$\begin{aligned} T &= (P_1(X_1, Y_1, Z_1), \\ &\quad \ldots, P_s(X_s, Y_s, Z_s)) \end{aligned} \tag{7}$$

$$E = (l_1, \ldots, l_t) \tag{8}$$

An edge l_j is a cubic Bézier curve. The shape of the wire-frame model is evaluated in comparison with the contour image. Because each inferred vertex is not always detected at each adequate location, it is necessary to move each inferred vertex to optimum location. A wire-frame model is represented by cubic Bézier curves, which are built by control points chosen out of the finite space near each inferred vertex. Comparing with the contour image the wire-frame model is optimized by GA. The GA consist of three operations which are a selection, a crossover and a mutation. In this article optimization of the model using GA are performed as follows:

1. The genotype G_k is defined as a set of control points of cubic Bézier curves by equation (9),

   $$\begin{aligned} G_k &= (X_1, Y_1, Z_1, \ldots, X_s, Y_s, Z_s, \\ &\quad X_{c11}, Y_{c11}, Z_{c11}, \ldots, X_{c2t}, Y_{c2t}, Z_{c2t}) \end{aligned} \tag{9}$$

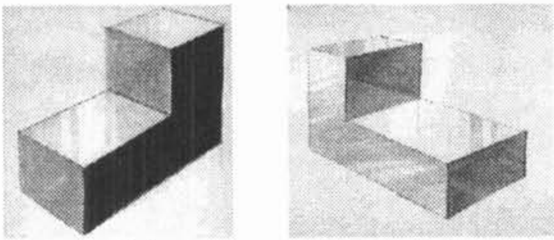
   where (X_i, Y_i, Z_i) denote i^{th} 3-D coordinates of s pieces of vertices given by equation (7). $(X_{c1j}, Y_{c1j}, Z_{c1j})$ and $(X_{c2j}, Y_{c2j}, Z_{c2j})$ denote j^{th} control points between both vertices of edge l_j. The phenotypes of individuals are wire-frame models.

2. The contour image is smoothed in order to cope with a little difference of the shape. The pixel value $f(x_i, y_i)$ of coordinates (x_i, y_i) is ranged from 0 to L. Let m and n be the number of pixels of the contour and the projection of the wire-frame model, respectively. The fitness F of each individual is given by equation (10).

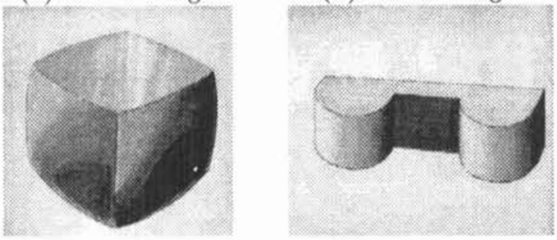
   $$F = \begin{cases} \frac{\sum_{j=1}^{m} f(x_j, y_j)}{L \times m} & (\mathrm{m} > \mathrm{n}) \\ \frac{\sum_{j=1}^{m} f(x_j, y_j)}{L \times n} & (\mathrm{n} \geq \mathrm{m}) \end{cases} \tag{10}$$

3. The procedure of the alternation of generations is performed under the same population size as follows:
   - (a) When the initial individuals are given randomly, each fitness is calculated.
   - (b) Individuals which have low fitness disappear. New individuals are created by those which have high fitness.
   - (c) A mutation is occurred at a constant ratio.

In Fig. 4, we show four examples of sketch images digitized by a scanner with 256 gray levels. These image sizes are 400 × 400 and 470 × 400 pixels. We use two-point crossover operator. The mutation ratio ranges from 10% to 30%. The selection ratio is fixed at 40%. The range for searching vertices is 32 × 32 × 32 pixels that center around each vertex. The searching range of another points is 128 × 128 × 128 pixels. As for initial individuals, vertices (X_i, Y_i, Z_i) are given randomly out of the finite space. Another control points are given at the locations which are divided between two vertices into three equal parts. After alternation of generations the gene which has the largest fitness is regarded as a set of control points of the optimum wire-frame model. In Fig. 5, we show the results of the optimum shapes of wire-frame models.

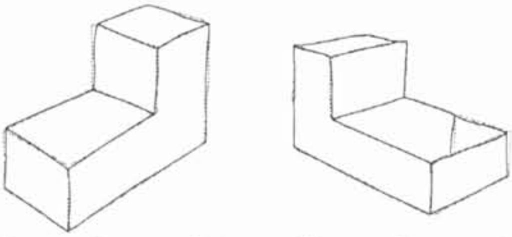


(a) sketch image 1 (b) sketch image 2



(c) sketch image 3 (d) sketch image 4

Fig. 4 Sketch images



(a) matching result 1 (b) matching result 2

Fig. 5 Experimental result 1

5 Constructing a surface model

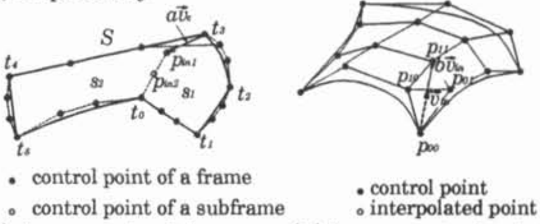
The wire-frame model corresponds to the boundary of a surface model. The boundary is called a frame. Since each surface is represented by a collection of quadrilateral and trilateral patches, a frame is segmented into subframes as follows:

1. A fiducial point for segmenting a frame is assigned to the vertex with a maximum internal angle in each surface of the polyhedron, as defined in section **3**.
2. A frame is segmented into two subframes under the same number of vertices or one difference. In Fig. 6(a), for instance, a frame S is segmented into two subframes s_1, s_2. When a vertex t_0 is assumed to be a fiducial point, one subframe s_1 has four vertices t_0, t_1, t_2, t_3 and the other s_2 has four vertices t_0, t_3, t_4, t_5 as well.
3. Frame segmentation is terminated when all subframes have three or four vertices.

Control points constructed a subframe correspond to boundary points of a bicubic Bézier surface, which consists of twelve boundary points and four inner control points. These sixteen control points are assigned as follows:

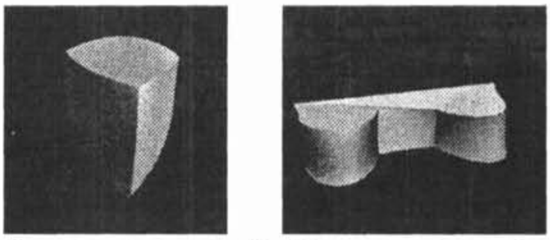
1. When a side of a subframe coincides with one of a frame, boundary points are equivalent to control points of the wire-frame model.
2. When a side of a subframe does not coincide with one of frames, boundary points are interpolated using neighboring points, as shown in Fig. 6(a). Let $\vec{v_c}$ be the vector from t_3 to the center of both adjacent boundary points. Boundary point p_{in1} is placed at the position $a\vec{v_c}$ from t_3, where a is constant. Boundary point p_{in2} is placed at the center between t_0 and p_{in1}.
3. The remaining control points of a bicubic Bézier surface are four ones which are interpolated using boundary points. Let $\vec{v_{in}}$ be the vector from p_{00} to the center of p_{10} and p_{01} in Fig.6(b). Point p_{11} is placed at the position $b\vec{v_i}$ from p_{00}.

After the segmentation of frames and the assignment of control points, a surface model is constructed by trilateral and quadrilateral Bézier patches. Fig. 7 shows the renderings of surface models. The constant value a and b are fixed at 1.0 and 2.0, respectively.



(a) frame and subframe (b) inner control point

Fig. 6 Interpolation of bicubic Bézier surface



(a) constructed model 3 (b) constructed model 4

Fig. 7 Experimental result 2

6 Conclusions

We have presented a method for constructing a surface model using bicubic Bézier surfaces from a single sketch image. The shape of the wire-frame model is optimized by searching control points out of each finite space. Boundaries of all wire-frame models are represented by cubic Bézier curves in order to cope with curved surfaces. The results show that the method makes it possible to find optimum 3-D control points by the use of GA.

References

[1] Ulupinar, F., Nevatia, R. : "Perception of 3-D Surfaces from 2-D Contours", IEEE Trans. Pattern Anal. & Mach. Intell, vol.15, no. 1, pp.3–18, 1993.

[2] Goldberg, D. E. : "Genetic Algorithms in Search, Optimization and Machine Learning", Addison Wesley, 1989.