グラフィクスとCAD 60－18
(1992. 12. 18)

ステレオ輪郭線からの立体形状の認識

鄭 在紋　　福島 重広
九州工業大学 情報工学部
〒820 福岡県飯塚市川津 680-4

あらまし　　物体の立体形状をステレオ輪郭線から認識するための新しい幾何学的モデルを提案する．このモデルでは再構成のためにひとつの仮定を設ける．つまり，物体はディスクの集合によって構成されており，それぞれは物体に含まれる球の断面であるとする．各方向からの平面データの解析は Blum による対称軸変換 (SAT) にもとづいている．解析する一対の平面データの間に対応関係を求めるために最大球 Maximal Sphere の概念を導入する．この関係の解析によって導かれる幾何学的ディスク Geometrical Disk はそれぞれ位置，向き，大きさによって十分に記述することができる．全形状は、いろいろな幾何学的ディスクの集合による記述から容易に復元される．この記述は一対のステレオ輪郭線に対して一意的である．このようなディスクは物体の立体形状を認識するための有効なプリミティブであろう．このモデルの有効性を自然な (つまり，滑らかに曲った) 形状と空間曲線を脊椎にもつ薩摩芋について示している．

和文キーワード　　ステレオマッチング，対称軸変換，最大球，幾何学的ディスク，立体形状認識

Recognition of 3-D Shape from Stereo Contours

Jaemoon Chung　and　Shigehiro Fukushima
Faculty of Computer Science and Systems Engineering
Kyushu Institute of Technology
Iizuka 820, Japan

Abstract　　A new geometrical model is proposed to understand and recognize the three-dimentional shape of an object from its stereo contours. One assumption is made for recostruction in this model: the object is composed of a set of disks, each of which is a cross section of a sphere within the object. The analysis of planar data from each direction is based on the Symmetric Axis Transform (SAT) by Blum. The concept of the *Maximal Sphere* is introduced, in order to make the corresponding relationships between a pair of planar data analyzed. Each *Geometrical Disk*, derived from the analysis of above relationships, can be sufficiently described with its position, orientation, and magnitude. Whole shape is easily recoverable from its description with a set of various Geometrical Disks. The description is unique for a pair of stereo contours. Such disks may be effective primitives for recognition of the 3-d shape of an object. The effectiveness of this model is demonstrated with several sweet potatoes, which have natural (that means smoothly curved) shapes with a space-curved spine.

英文 key words　　Stereo Matching, Symmetric Axis Transform, Maximal Sphere, Geometrical Disk, 3-D Shape Recognition

1 Introduction

If a human being sees a planar figure displayed only by its contour, still he may be able to imagine a 3-d shape relating to it. He may be able to describe the 3-d shape with a set of generalized cones as in [4]. In fact, researches have been trying to extract the descriptor of a planar shape from its boundary information [5-9].

In this paper, we propose a new geometrical model for reconstruction, description, and recognition of the 3-d shape of an object. This model is developed on the basis that an object's planar contour may be one of the most important cues, for a human being to understand the shape of an object. It is well known that the binocular approach is necessary to understand any 3-d environment sufficiently. Therefore, the standpoint of this model may be called shape-from-stereo-contours. Determining the correspondence between a pair of images is the most difficult problem in the binocular approach. We use the characteristics of a sphere to solve this problem. The *assumption for reconstruction* of 3-d shape is as follows: an object has certain cross sections, each of which has a shape of a disk as well as is a cross section of a sphere contacting the object from inside. The contact is along the disk edge. In the above assumption, the key concepts of *Maximal Sphere* and *Geometrical Disk* are included. The *maximal sphere* is a 3-d extension of the *maximal disk* in the Symmetric Axis Transform (SAT). The SAT is one of the planar shape descriptors [5]. In this paper, the SAT is used to analyze each planar contour, the silhouette outline of the view of the object from each direction.

A 3-d recognition system has two major parts: the shape description from sensed images through reconstruction, and the matching of the described shape to stored models. As mentioned previously, stereo vision has been difficult due to the problem of finding correspondence between images. Therefore, we think, the 3-d recognition systems have appeared which use a monocular data. One of these systems is a model-based 3-d vision system. This system needs reasoning of a 3-d model from a 2-d silhouette image of an object [1,2]. For the reasoning, view point transformation is applied to the 3-d models stored in a database. The shape of a model is described with generalized cylinders. A generalized cylinder is specified by parameters about its axis curve and cross section shape. A complex object is described by hierarchic representation with a collection of various generalized cylinders [3]. However, the scope of object description is limited within the range of parameters adequately describable. This means that the cut-down of information is inevitable in shape description with the generalized cylinders, in order to retain the model database for efficient matching within acceptable size.

In this paper, a new descriptor of 3-d shape is also proposed. The disk, addressed in the *assumption for reconstruction* above, is related to this. This disk plays roles of primitives not only for 3-d shape description but also for recognition. The disk is called *Geometrical Disk*, in order to emphasize that it is completely characterized by its position, orientation, and magnitude. The whole of a 3-d shape can be described by a set of various these disks. Although the shape is limited to a disk, variety of representation is possible other than a simple disk. This means that all perspective projections of a disk can be included in the parts of the reconstructed shape theoretically. It is possible to design a new type of 3-d vision system by using these disks as primitives for shape representation. These systems incoporate the stereo-vision, and may perform the 3-d recognition function without the burden of the cut-down of information, view point transformation, and hierarchical reasoning. These are inevitable in a model-based 3-d vision system with single vision.

This paper is organized as follows. In Section 2, the concept of the *Maximal Sphere* is described in relation to the SAT. Sections 3 deals with the analysis for recovering. Section 4 includes the representation and the recovering of a 3-d shape with this model. Finally, experimental results are shown in Section 5. In this section, this model was implemented to demonstrate its effectiveness experimentally. Several pairs of images of sweet potatoes were analyzed, whose shapes are relatively complex with space-curved spines.

2 The Maximal Sphere

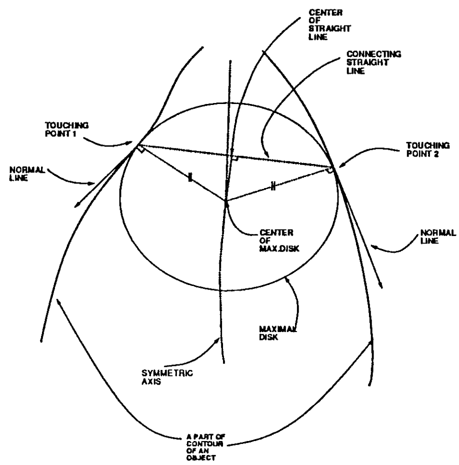


Fig.1 Symmetric axis transform

The *symmetric* (or *medial*) *axis* generated by applying the SAT to a contour can be considered as the locus of the center of a *maximal disk* (MD) as shown in Fig.1 [8]. The MD is a disk which is contained inside the contour and touches at least two points of the boundary contour. Several algorithms have been proposed for computing the SAT for a contour approximated by a polygon [10-15]. In practice to compute the SAT, it is inevitable to sample a train of points from a planar boundary due to the use of digital sensors such as CCD cameras. If the magnitude of an object and the spatial resolution of these digital devices are appropriately adjusted, the problem of inconsistency may be ignored. (The inconsistency may occur due to digitization of the contour.) In this paper, the Voronoi diagram and the Delaunay triangulation were applied to the digital contours [13,14] for implementation of the SAT. The former is used to find symmetric axis and the latter is used to find pairs of symmetric points on the contour. The Delaunay triangulation was implemented using the divide-and-conquer algorithm by Lee and Schachter [11] for efficient computation; the Voronoi diagram was derived from this triangulation.

Proposition 1
For the simply connected contour of a planar figure,

any maximal disk MD is tangent to the contour at its touching point.

Proposition 1 is proved in [16]. This means that the line connecting a touching point and the center of the disk meets perpendicularly with the tangent line at the touching point. For an MD, the perpendicular bisector of the straight line connecting the touching points passes through the center of the MD.

Proposition 2
A disk, contained in a sphere and expressible with a cross section of the sphere, can be uniquely identified by the center position of the sphere, and the center position and the diameter of the disk itself. The orientation of the disk is also needed if the center positions of the sphere and the disk are the same.

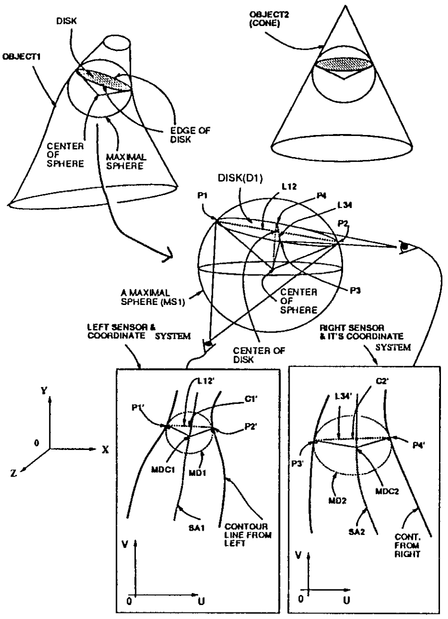


Fig.2 The Maximal Sphere related to SAT

Proposition 2 is trivial. On the bases of Propositions 1 and 2 and the *assumption for reconstruction*, let's consider the shape of Object 1 in Fig.2 as a general case, although it is easier to think about the case of Object 2, a cone. As shown in Fig.2, we assume that a cross section of Object 1 can be expressed as a disk which is a part of a sphere within the object. Object 1 is viewed with a pair of cameras from left and right directions, respectively, to take a pair of silhouette contours. The point pair (P_1, P_2) and (P_3, P_4) are two end-points of the disk within the sphere, viewed from the left and right directions, respectively. $L12$ and $L34$ are connecting line between the point pair (P_1, P_2) and (P_3, P_4), respectively. And $L12$ and $L34$ also pass the center of this disk. The SAT is applied to each of the contours. Then, we assume that $MD1$ and $MD2$ are silhouettes of a sphere $MS1$ viewed from the left and right directions, respectively. Thus, the centers of these MD's correspond to the center of this sphere. The point pairs (P_1',P_2') and (P_3',P_4') correspond to (P_1, P_2) and (P_3, P_4), respectively. The point pairs (P_1',P_2') and (P_3',P_4') can be determined from the fact that they have a common three-dimensional point, the center of $MS1$, which are projected as points $MDC1$ and $MDC2$, respectively. The points $MDC1$ and $MDC2$ are on symmetric axis $SA1$ and $SA2$, respectively. The straight lines $L12'$ and $L34'$ correspond to $L12$ and $L34$. Points $C1'$ and $C2'$ bisect $L12'$ and $L34'$, respectively. These points correspond to the center of the disk $D1$.

The above observation is valid due to the characteristics of a sphere as follows. All normals on the surface of the sphere meet at the center of the sphere. Therefore, all normals on the edge of the disk, a section of the sphere, also meet at the center of the sphere. By Proposition 2, correspondence can be found between a pair of images. This sphere is named *Maximal Sphere* (MS) in relation to MD. This is three dimentional extension of the MD in the SAT.

Using the mutual relations between MS and MD, the analysis for a pair of contour images is implemented through two steps.

At the first step, the SAT is applied to each one of the contours. Then, a set of triples of points are generated for each contour. Each triple is composed of three points: one point is on the symmetric axis as $MDC1$ ($MDC2$); and the other two are corresponding points on the contour as P_1', P_2' (P_3', P_4'). The point $MDC1$ ($MDC2$) is the center of the inscribing circle which touches the contour at the points P_1' and P_2' (P_3' and P_4'). These three points of a triple are related to a specific MD. To find correspondence between a pair of sets of such triples, the following are assumed for the view from a direction.

Assumption :

1. A point on the symmetric axis of a triple is the center point of a certain MS.
2. A pair of corresponding touching points in a triple are two end-points of a disk within the MS.
3. The bisection point of the straight line, connecting the two corresponding touching points in a triple, is the center of the disk.

The correspondence between the two contours can be determined by considering the above relationships with the given sensing condition. The sensing condition comprises the characteristics of the cameras and the coordinate systems of the cameras and of the world.

At the second step, the state of each disk is calculated for a pair of triples, for which the correspondence was determined through the the first step. Consider the straight lines $L12'$ and $L34'$ which connect a pair of the touching points. The center position of the disk is calculated by analyzing the two bisection points $C1'$ and $C2'$ of the straight lines. The orientation of the disk is calculated by analyzing the orientations of the two straight lines. Finally, by analyzing the center positions of both the MS and the disk contained in the MS, the magnitude of the disk can be determined. By applying the second step to all triple pairs, all disks can be obtained. The description with a set of these disks can be used for recovering and recognition of the 3-d shape of the object.

Proposition 3
By applying the SAT to a silhouette contour, the symmetric axis and the set of MD's are uniquely determined.

Proposition 3 is the basis of unique recoverability of the proposed geomerical model. This is because the MS is a 3-d extension of the MD by the SAT. Proposition 3 is proved in [16].

We can know that the SAT embodies a definition of local symmetry and has rich representations in the sense of being information-preserving [8]. The proposed model also has a capability of rich and detailed representation of the 3-d shape of an object, because it completely depends on the SAT. Especially, this model is more effective for smoothly curved objects.

3 Analysis for Recovering

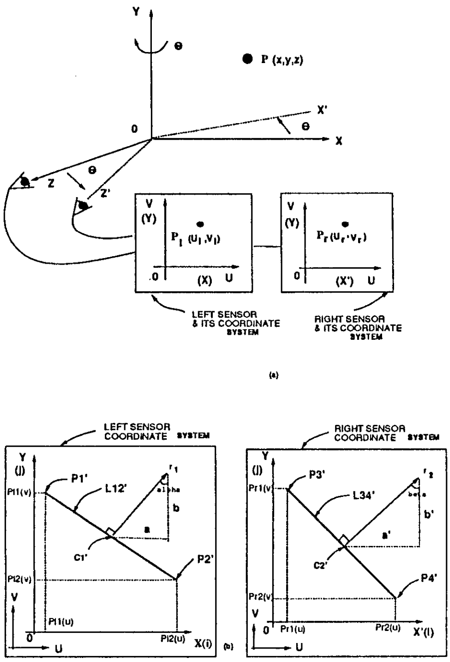


Fig.3 (a) Sensor coordinate systems.
(b) Analysis of orientation

Consider the cartesian world coordinate system arranged based on the left camera system as shown in Fig.3(a). The plane coordinate system of this camera is composed of u- and v-axes which meet perpendicularly. We assume that the x- and y-axes of the world coordinate system are defined parallel to the u- and v-axes of the left camera coordinate system, respectively. The z-axis of the world coordinate system is defined so that the camera is aimed at a point at minus infinity on z-axis. For stereo vision, the right camera is located at the position rotated from the left by θ radian counterclockwise around the y-axis. Let's call z'-axis instead of z-axis and x'-axis instead of x-axis for the right camera. In the right camera system, the u- and v-axes are parallel to the x'- and y-axes, respectively, and the camera itself is aimed at a point at minus infinity on the z'-axis.

The sensing condition is calibrated as follows. As shown in Fig.3(a), the origin of this world coordinate system is on the y-axis. The origin of the coordinate system for each camera may be determined for an adequate corresponding point on the y-axis. The characteristics of the two cameras are assumed to be the same, and the entire camera plane is assumed uniform. For convenience, we assume orthogonal projection for the imaging. This means that the cameras are located sufficiently far away from the imaging planes. Therefore, the camera parameter considered like the focal length[1] is infinity. This means that the magnification ratio between u-, v-axes and x- (or x'-), y-axes is equal to 1.

Consider the necessary conditions to determine the one-to-one correspondence among the triples for the pair of contours. The conditions are derived from the characteristics of the MS and its contained disk.

1. The center of an MS: find a pair of epipolar points on the symmetric axes for the respective contours. This is done by examining the v-values for such points to see if they are equal.
2. The center of the disk: examine the v-values of the bisection points of the straight lines connecting the two corresponding touching points for the respective contours to see if the v-values are equal.
3. Magnitude of the MS: examine if the radii of the MS's are equal for the two contours.
4. If more than two triples pass the above three conditions, the triples with closer u-values of the points on the symmetric axis may be matched.

In examining the equality in terms 1, 2, and 3 of the above conditions, some margins can be permitted within an error bound. This fills the gap between the assumed shape model and the real shape. The further analysis about the proper margins may generate a description closer to the real. Based on the term 4 of the above conditions, multiple objects can be analyzed at same time. For a pair of triples from the above procedure, the analysis is implemented to calculate the state of each GD as follows.

Consider Fig.3(a) for calculation of a point in the world coordinate system. The point $P(x,y,z)$ is imaged as $P_l(u_l,v_l)$ and $P_r(u_r,v_r)$ on the left and right cameras, respectively. When variable w is not zero, and w_l and w_r are dummy terms in z-coordinates, we can represent this relations by Eq.(1) and (2).

$$(u_l, v_l, w_l, 1) = (wx, wy, wz, w) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1/f \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1)$$

$$(u_r, v_r, w_r, 1) = (wx, wy, wz, w)$$
$$\begin{pmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1/f \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2)$$

where f is the focal length; and θ is the angle between the two cameras.

Because f is infinity, the world coordinates of point $P(x,y,z)$ can be expressed by the camera coordinates as Eq.(3).

$$\begin{aligned} x &= u_l \\ y &= v_l \ , \ v_r \\ z &= \frac{u_l \cdot \cos\theta - u_r}{\sin\theta} \end{aligned} \quad (3)$$

[1]The length between the view point and the imaging plane is called focal length here. If the camera (or imaging system) focuses to infinity, this is exactly the same as the focal length of the lens.

Therefore, the center point of the disk can be calculated as follows. A pair of bisection points (for example, $C1'$ and $C2'$), of two straight lines connecting two corresponding touching points in a triple pair, can be applied to this equation. The u- and v-axes values for those points which are imaged on the left and right camera coordinate systems are selected as (u_l, v_l) and (u_r, v_r), respectively. By substituting these values into Eq.(3), the center point of the disk is determined.

The orientation of the disk in the world coordinate system is determined as follows. As shown in Fig.3(b), let the orientations of the straight lines connecting the two corresponding touching points in the left view and the right view be $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ in each respective camera coordinate systems. In Eq.(4) and (5) and Fig.3(b), $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ and $\boldsymbol{l}$ are unit vectors in the directions of x-, y-, z- and x'-axes, respectively. Then, we can describe these orientations by Eq.(4).

$$\begin{aligned} \boldsymbol{r}_1 &= a\boldsymbol{i} + b\boldsymbol{j} \\ \boldsymbol{r}_2 &= a'\boldsymbol{l} + b'\boldsymbol{j} \end{aligned} \qquad (4)$$

where $b = b'$ in the coordinate relation of this camera location. By vector composition of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ with the given coordinate systems, the orientation of the disk $\boldsymbol{r}$ in the world coordinate system is represented by Eq.(5).

$$\boldsymbol{r} = a\boldsymbol{i} + b\boldsymbol{j} + c\boldsymbol{k} \qquad (5)$$

where

$$\begin{aligned} a &= \sin\alpha \\ b &= \cos\alpha \\ c &= \frac{\sin\alpha \cdot \cos\theta - \cos\alpha \cdot \tan\beta}{\sin\theta} \end{aligned} \qquad (6)$$

where

$$\begin{aligned} \alpha &= \tan^{-1}\frac{P_{l1}(v) - P_{l2}(v)}{P_{l2}(u) - P_{l1}(u)} \\ \beta &= \tan^{-1}\frac{P_{r1}(v) - P_{r2}(v)}{P_{r2}(u) - P_{r1}(u)} \end{aligned} \qquad (7)$$

See Fig.3(b) for α, β, $P_{l1}(u)$, $P_{l2}(u)$, $P_{l1}(v)$, $P_{l2}(v)$, $P_{r1}(u)$, $P_{r2}(u)$, $P_{r1}(v)$ and $P_{r2}(v)$.

Finally, the magnitude of the disk can be derived as follows. As mentioned in the previous section, the magnitude of a disk can be obtained by analyzing the lengths of a pair of lines, each of which connects two touching points of it's triples. But, this method requires complex calculation for the perspectives of the slant disk. Therefore, for this calculation, we use the characteristics of a sphere and a circle: locus of all disks with an identical diameter is on a surface of a sphere, which is within the MS and whose center coincides with the center of the MS. Therefore, the diameter of the disk can be calculated from the center position of the MS and the disk itself, and the radius of the MS. The center position of the MS can be found by substituding u-, v-axis values of center positions of a pair of MD's into Eq.(3). From the length of the line connecting the touching point and the center of it's MD, the radius of the MS can be obtained. Let us denote the x-, y-, z-axis values of the center position of the MS and the disk as x_s, y_s, z_s and x_d, y_d, z_d, respectively. The radius of the MS is denoted as r_s. The distance between the above two center points d_{sd} is given by Eq.(8).

$$d_{sd} = \sqrt{(x_d - x_s)^2 + (y_d - y_s)^2 + (z_d - z_s)^2} \qquad (8)$$

Therefore, the diameter d of the disk can be calculated from Eq.(9).

$$d = 2\sqrt{|r_s^2 - d_{sd}^2|} \qquad (9)$$

The state of GD is determined for the world coordinate system through the analysis. By implementing the analysis for all corresponding triple pairs, the set of GD's can be obtained.

4 3-D Shape Representation and Recovering

Through the analysis, a GD can be described with its position, magnitude, and orientation in the world coordinate system. A description can be Expression (10) for the ith GD.

$$x_i \quad y_i \quad z_i \quad d_i \quad a_i \quad b_i \quad c_i \qquad (10)$$

where x_i, y_i and z_i indicate the center position of the GD. This can be determined by Eq.(3). d_i is the diameter of the GD which can be found by Eq.(9). a_i, b_i and c_i are three components in x-, y- and z-axes directions, respectively, of the unit vector representing its orientation. These are determined by Eq.(6). An example of this GD description is shown in Fig.4.

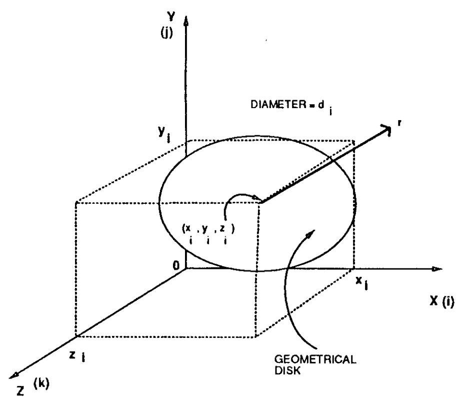


Fig.4 An example of GD

Once the whole part of the object is analyzed, all GD's are known. If the object is described with n-pieces of GD's, the description can be given by Expression (11).

$$\begin{matrix} x_1 & y_1 & z_1 & d_1 & a_1 & b_1 & c_1 \\ x_2 & y_2 & z_2 & d_2 & a_2 & b_2 & c_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_i & y_i & z_i & d_i & a_i & b_i & c_i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & z_n & d_n & a_n & b_n & c_n \end{matrix} \qquad (11)$$

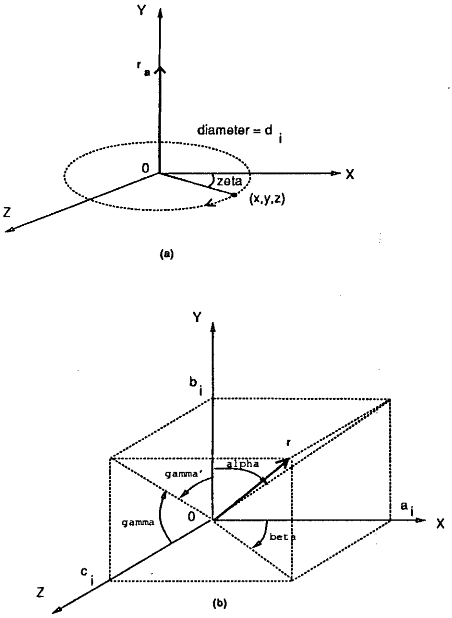


Fig.5 (a) Initial state for recovering
(b) Consideration for orientation

To recover the 3-d shape from Expression (10), the GD's edge must be determined in the world coordinate system. To do this, firstly, we imagine a disk with diameter d_i, which is on the x-z plane. Therefore, its orientation r_a is the same as the direction of the y-axis. The center position of the disk is the origin of the coordinate system as shown in Fig.5. The point of the edge of this disk is given by Eq.(12).

$$\begin{aligned} x &= \frac{d_i}{2}\cdot\cos\zeta \\ y &= 0.0 \\ z &= \frac{d_i}{2}\cdot\sin\zeta \end{aligned} \qquad (12)$$

where ζ varies from 0 *to* 2π radian.

To recover the 3-d shape, the orientation r_a of the disk in Fig.5(a) can be changed to r as shown in Fig.5(b) or Eq.(15).

$$r = a_i i + b_i j + c_i k \qquad (13)$$

where from Fig.5(b),

$$\begin{aligned} \alpha &= \cos^{-1}\frac{b_i}{\sqrt{a_i{}^2 + b_i{}^2}} \\ \beta &= \cos^{-1}\frac{a_i}{\sqrt{a_i{}^2 + c_i{}^2}} \\ \gamma &= \cos^{-1}\frac{c_i}{\sqrt{b_i{}^2 + c_i{}^2}} \end{aligned} \qquad (14)$$

where

if a_i=b_i=0 then $\alpha = 0, \beta = \pi/2$, and γ=0,

if a_i=c_i=0 then $\alpha = 0, \beta = 0$, and $\gamma = \pi/2$,

if b_i=c_i=0 then $\alpha = \pi/2, \beta = 0$, and $\gamma = \pi/2$.

For Expression(10), the point of the disk edge are changed by Eq.(15). Recovered point is obtained by considering the rotation $\boldsymbol{R}$ and translation $\boldsymbol{T}$ of this disk. The rotation is related to Eq.(16). The translation x_i, y_i, z_i is given by Expression (10). Thus,

$$\boldsymbol{O} = \boldsymbol{I}\boldsymbol{R} + \boldsymbol{T} \qquad (15)$$

where $\boldsymbol{O}$ = (x_o, y_o, z_o, 1) denotes the recovered point of GD's edge. $\boldsymbol{I}$ = (x, y, z, 1) is the initial point given by Eq.(10). $\boldsymbol{T}$ = (x_i, y_i, z_i, 1) denotes the translation of the center of the GD and is given by Expression(10). $\boldsymbol{R}$ is the rotation of this GD given by Eq.(16).

$$\boldsymbol{R} = \boldsymbol{R}_\alpha \boldsymbol{R}_\beta \boldsymbol{R}_{\gamma'} \qquad (16)$$

where $\gamma' = \pi/2 - \gamma$,

$$\boldsymbol{R}_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 & 0 \\ \sin\alpha & \cos\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

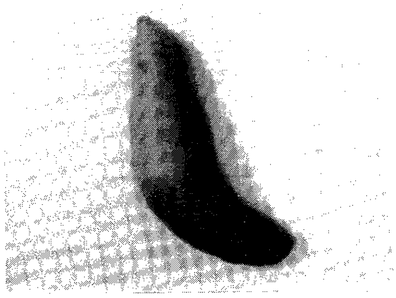
$$\boldsymbol{R}_\beta = \begin{pmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{R}_{\gamma'} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\gamma' & -\sin\gamma' & 0 \\ 0 & \sin\gamma' & \cos\gamma' & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

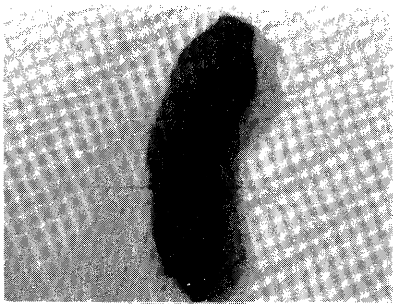
5 Experiments

Instead of using two cameras, only one camera was used in this experiment. The sensing condition was constructed as described previously in Section 3. As shown in Fig.3(a), the object was rotated clockwise by θ radian around the y-axis after taking one view from the left direction. Then, the other view was taken. This procedure guarantees the equality of the characteristics of the assumed two cameras.

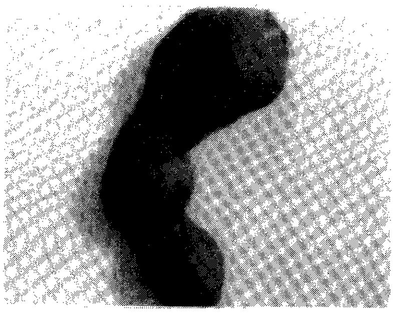
As shown in Fig.6, several sweet potatoes were selected as the objects for the experiment. These are considered to be objects of natural shapes, i.e. smoothly curved, which have relatively complex space-curved spines. Fig.7 shows the overlaid sensor data of the objects in Fig.6, viewed from two directions and processed by the SAT. In order to show drastically the difference between a pair of sensed data, the angle between cameras, θ, was set to about 20 degrees in Fig.7(c). However, smaller angles give the advantage of satisfaction with the fact that this model is more effective especially for smoothly curved objects. The overlaid sensor data is demonstrated in Figs.7(a) and (b) for the experimental analysis actually done. The angle between cameras, θ, is about 5 degrees in this case. Fig.8 shows the recovered results for the three sweet potatoes, viewed from two directions.



(a) Object A



(b) Object B



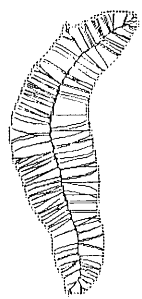
(c) Object C

Fig.6 Sweet potatoes used for experiments

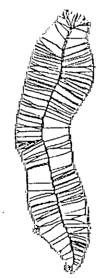
6 Concluding Remarks

A new geometrical model has been proposed for understanding and describing the 3-d shape of an object from its stereo contours. This model is globally composed of the concept of the *Maximal Sphere* and the *Geometrical Disk*. The former is the basis to analyze the 3-d shape for it's recovering. The latter is a primitive for description and recognition of the 3-d shape. The standpoint of this study is shape-from-stereo as well as shape-from-contour. The new model based on geometry named *Maximal Sphere* is basically for resolving the difficulties in finding correspondence between a pair of stereo images. The model has been proved to be sound through experiment. The proposed primitive, *Geometrical Disk*, is very efficient because of the simplicity in description and the richness in representation of a 3-d shape. Therefore, this model may lead to a new type of 3-d vision system, with a pair of contour detecting sensors. By using this system, 3-d recognition is possible without the necessity of reasoning between 3-d models and a 2-d shape, view point transformation, and hierarchical reasoning for complex shapes.

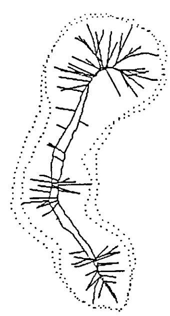
Moreover, the proposed geometrical model is practically available, because correspondance finding based on contour detection is relatively simple, compared with other stereo matching bases. This model produces also uniquely recoverable output for a pair of given inputs as shown in the previous section. Although the theory was implemented for data taken by cameras, this model is applicable to planar contour shapes, taken by any kinds of sensors.



(a) Object A

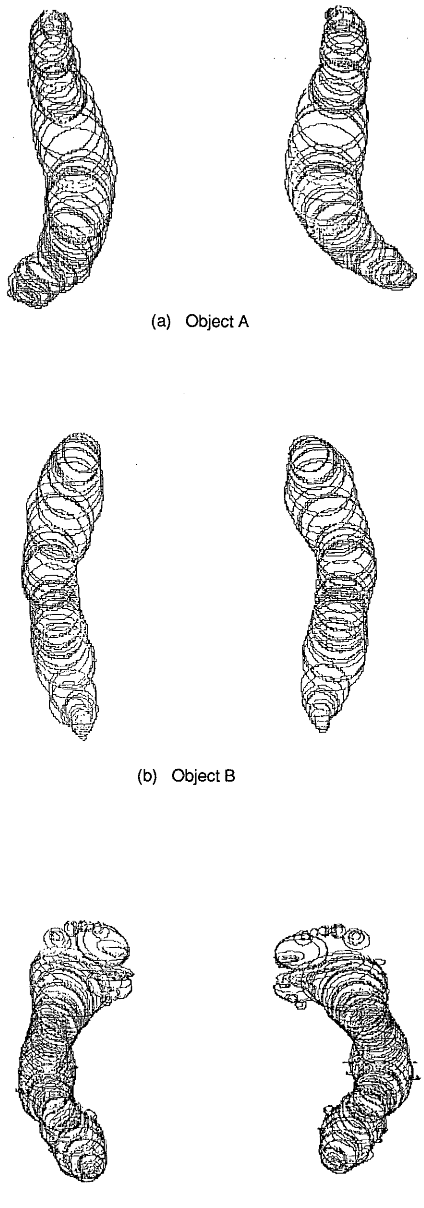


(b) Object B



(c) Object C

Fig.7 Overlaid images taken from stereo sensors



(a) Object A

(b) Object B

(c) Object C

Fig.8 Recovered 3-D shapes

References

[1] R. Brooks, "Symbolic reasoning among 3d models and 2d images," *Art. Intell.*, vol. 17, pp. 285-348, 1981.

[2] R. A. Brooks, R. Greiner, and T. O. Binford, "The ACRONYM: model-based vision system," *IJCAI'79*, pp. 105-113, 1979.

[3] D. Marr and H. K. Nishihara, "Representation and recognition of the spatial organization of three-dimentional shapes," *Proc. Roy. Soc. London B*, vol. 200, pp. 269-294, 1978.

[4] G. J. Agin and T. O. Binford, "Computer description of curved objects," *IEEE Trans. Computers*, vol. C-25, no. 4, pp. 439-449, April 1979.

[5] H. Blum, "A transformation for extracting new descriptors of shape", *Models for perception of speech and visual form*, ed. W. Wathen-Dunn, MIT Press, 1967.

[6] H. Blum and R. N. Nagel, "Shape description using weighted symmetric axis features," *Pattern Recog.*, vol. 10, pp. 167-180, 1978.

[7] M. Brady and H. Asada, "Smoothed local symmetries and their implementation," *Int. J. Robotics Research*, vol. 3, no. 3, pp. 36-60, Fall 1984.

[8] M. Brady, "Criteria for representations of shape," *Human and machine vision*, ed. A. Rosenfeld and J. Beck. New York: Academic Press, 1983.

[9] U. Montanari, "A method for obtaining skeletons using a quasi-Euclidean distance," *J. Assoc. comput. Mach.*, vol. 15, no. 4, pp. 600-624, Oct. 1968.

[10] F. L. Bookstein, "The line-skeleton," *Comput. Gr. Image Proc.*, vol. 11, pp. 123-137, 1979.

[11] D. T. Lee and B. J. Schachter, "Two algorithms for constructing a Delaunay triangulation," *Int. J. Comput. Inform. Sci.*, vol. 9, no. 3, pp. 219-242, 1980.

[12] D. T. Lee, "Medial axis transformation of a planar shape," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, no. 4, pp. 363-369, July 1982.

[13] J. Fairfield, "Segmenting blobs into subregions," *IEEE Trans. Sys. Man, Cybern.*, vol. SMC-13, no. 3, pp. 363-384, May/June 1983.

[14] J. Fairfield, "Segmenting dot patterns by Voronoi diagram concavity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-5, no.1, pp. 104-110, Jan. 1983.

[15] S. Fukushima and T. Okumura, "Estimating the three-dimentional shape from a silhouette by geometrical division of the plane," *Proc. IECON'91*, pp. 1773-1778, Oct. 1991.

[16] A. Rosenfeld, "Axial representations of shape," *Comput. Vision Gr. Image Proc.*, vol. 33, pp. 156-173, 1986.