COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 42, 345–370 (1988)

Shape from Random Planar Features

SEIICHIRO NAITO* AND AZRIEL ROSENFELD

Center for Automation Research, University of Maryland, College Park, Maryland 20742

Received November 1, 1984; accepted December 29, 1987

This paper describes an approach to interpreting line drawings under assumptions which are ubiquitous in natural scenes. The assumptions are that many identical, essentially two-dimensional features are depicted and they are arranged in random orientations. We assume that at least one of the many features is parallel to the image plane, and thus gives the real dimensions of a feature. From this, the orientations of the other features can be easily recovered. Four examples of this approach are shown to give quite natural results. © 1988 Academic Press, Inc.

1. INTRODUCTION

One of the most marvelous image understanding abilities of human beings is the ability to perceive three-dimensional shapes in line drawings. Even though line drawings contain no information such as stereo, shading, or texture, generally we have no difficulty in visualizing 3D shapes from such drawings. In this paper, we propose an approach to interpreting some classes of line drawings under probable assumptions which are frequently valid in natural scenes.

There have been a number of other approaches to 3D line drawing interpretation. In one approach, known as "shape from shape" [1], the basic idea is that of regularity—for example, parallelism or symmetry. Combining regularity assumptions with gradient space techniques, sufficient constraints on surface normal orientations can be obtained. If some type of uniformity is assumed for a spatial curve, its 2D image can be interpreted as a 3D shape [2, 3]. The ACRONYM system [4, 5] is a model-based system whose representation of shape is very general, though domain specific data is necessary.

This paper deals with a class of line drawing images in which many identical objects are depicted in random orientations. Conventional line drawing interpretation is carried out on single objects. If we are given many different representations of a feature, that is, many identical objects in random orientations are given, they present a useful cue for 3D interpretation. This situation is very common in natural scenes encountered in our daily life. Animals, plants, and many types of artificial objects have almost identical shapes, and they often occur in large groups in a scene. The purpose of the interpretation scheme proposed here is to indicate how 3D shape can be estimated from a line drawing containing a large group of such objects. An advantage of our scheme is that it is not model based, and so does not require knowledge of domain specific parameters.

Since our scheme uses many features, it is related to the techniques for deriving 3D shape from texture [6]. In a sense, the technique proposed here can be considered to be somewhere between shape from shape and shape from texture.

*Permanent address: Musashino Electrical Communication Lab., N.T.T., Musashino-shi Tokyo, 180 Japan.

0734-189X/88 $3.00

In the next section, the general assumptions on which our interpretation scheme is based, and the basic steps of the scheme, will be explained. In Section 3, four classes of models based on our general scheme will be introduced, and typical examples and interpretation results using these models will be presented. It will be shown that the line drawing interpretation scheme proposed here gives natural results under a reasonable set of assumptions. In Section 4, some of the advantages of and problems with our scheme will be discussed and brief comments on possible future work will be made.

2. THE GENERAL APPROACH

2.1 Assumptions

The general idea of our approach, and the assumptions on which it is based, are presented in this section.

Our purpose is to propose an approach to 3D interpretation and to demonstrate its ability. Like much recent research, our interpretation scheme is module or function specific rather than domain specific. The scheme assumes the results of lower level feature extraction. The examples in this paper show natural objects represented by line drawings. The drawings are simplified in the sense that only information essential to the interpretation scheme remains. The real scene, of course, is not a line drawing, and the process of extracting the important lines from among the edges present in the image would be a nontrivial problem.

Our interpretation scheme assumes that the following two conditions hold:

1. Many identical objects are depicted.
2. They are arranged in random orientations.

Under these assumptions, the image contains many instances of the same object in various directions. Based on our prior process of feature extraction, we also assume that we have complete information about the correspondences of lines, endpoints, or corners between the various instances.

2.2. Underlying Concept

To clarify the significance of our assumptions and our interpretation strategy, some comparative remarks may be helpful.

In conventional "shape from shape" line drawing interpretation, the assumption of uniformity is one of the basic ideas. For example, under the assumption of zero or minimal torsion, some spatial curve images can be interpreted [2, 3]. In this paper, we also utilize a certain type of uniformity. This is not uniformity of the features in one object, but rather it is uniformity in the sense that many identical objects are present in the image. Non-uniformity, complexity, or heterogeneity within one object is acceptable in our interpretation scheme.

From a different viewpoint, interpretation techniques using many objects are relevant to shape from texture techniques [6]. They are based on statistical information in some region where the texture composed of many primitive objects can be considered to be approximately uniform. On this basis, the orientation of the region is estimated from the observed texture.

The assumption in this paper is that the objects have random orientations, so there is no uniform region. Three-dimensional interpretation is carried out for every object, not for a surface containing a group of them.

We have so far used the term "object" very vaguely. In general, it is an object as it appears in an image, but more specifically we shall assume that it is a geometrical feature such as a triangle formed by a given triple of points on the object in the scene, a line segment, an angle between a pair of lines, or a combination of these. In this paper they are called unit features. An important property of all the unit features just mentioned is that they are all planar features.

A general interpretation scheme based on the properties of planar features can be defined as explained in the next subsection. Depending on which kind of features are used, many different specific models can be designed. In the next section four such models will be introduced and examples of their application will be given.

2.3. Interpretation Scheme

Although each model defined in the next section has its own specific scheme, all of these schemes involve two common steps as described below.

Step 1. *Estimating the Actual Features*

An actual feature is a real length or angle in 3D space, not in its projection on the image. Such features can be estimated using cues such as the following:

* Longest line. Assume that a line segment of length L is given in 3D space, and that it is projected orthographically onto the image. Then its length in the image is $L \cos \theta$, where θ is the angle between the image plane and the direction of the line.

By our basic assumption, many identical line segments are present in random orientations. Thus there is a good possibility that some of them are oriented at angles θ that are nearly equal to zero. It is therefore a reasonable assumption that the longest length among the line segments in the image is the real length L.

* Features in special positions. In the case of line segments, the longest one is assumed to be oriented parallel to the image plane. Imagine a plane which contains that longest line segment, and a feature on it. The image of the feature is a projection of the feature rotated around the line as an axis. Then, relative to the direction of the line, we have some mathematical constraints on estimating the real feature from its image. This is one of the possible cues for interpretation, which depends on special arrangements or positions of the features.

The angle between two branching line segments is another important geometrical feature. Assume that an angle between two branching line segments in 3D space is projected onto the image and lengths of line segments are measurable. Again, by our basic assumption, many identical angles are presented in random orientations. Then some angles on the image are near zero. This means that the plane containing the two branching line segments is perpendicular to the image plane. Thus, the lengths of these line segments in the image provide a good basis for estimating the real angle or the real lengths.

In the same way, the maximum length of the perpendicular from the end of one line to the other line is another good cue. The perpendicular should be parallel to the image plane. Then, geometrical interpretation can be carried out using combinations of this and other cues.

* Most frequent angle. In the preceding paragraphs, we discussed interpretation cues in the case where both angles and lengths of line segments are measurable.

Even in the situation where only angles are measurable, we can obtain some information and estimate the real angles. This interpretation scheme too depends on the randomness of the distribution of 3D orientations of the angles. It will be shown in Section 3 that the most frequent angle is likely to be the real angle.

Step 2. Estimating the 3D Angle

Once a real feature in 3D space has been estimated, it is not difficult to determine the value of the angle in 3D space that projects the given feature onto the image plane. Generally, there will be two possible values that are mirror images. To eliminate this ambiguity, other heuristic information would be required. Although we discuss some ways to resolve the ambiguity in this paper, this is essentially a domain specific problem. We will therefore sometimes use random selection to chose one of the two possible answers.

3. SPECIFIC MODELS

In this section, four specific models are introduced. They are based on the general scheme described in Section 2. Typical examples of these models and the results of applying them are presented.

3.1. Base Line and Height

As indicated in section 2, we assume that the image is composed of many identical planar features in random orientations.

A simple example of a feature is a planar curve. In this section we present a model that can estimate 3D information from an image in which many planar curves are depicted. These curves will have the same shape and are oriented in random direction in 3D space.

Let C be a planar curve from point P to Q as shown in Fig. 1. The actual shape of C is arbitrary. The only information required in our interpretation scheme is the length L of the base line between P and Q, and the height H of C from the base line.

A coordinate system is chosen as shown in Fig. 2, where orthographical projection is assumed. The planar curve C is located at point P in 3D space and in a direction designated by the two angle parameters θ and ψ.

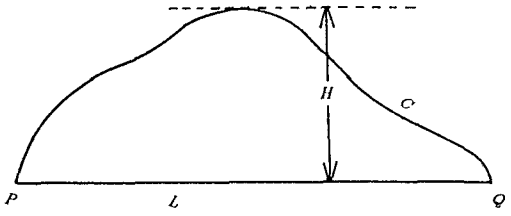

FIG. 1. Planar curve notation.

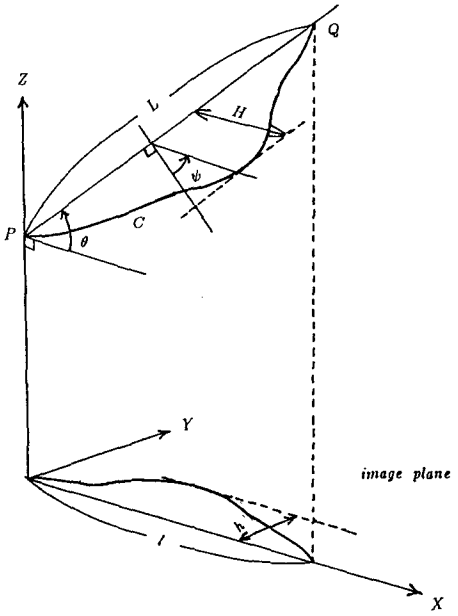

FIG. 2. Projection of a planar curve C.

The following relations on the base line length and the height are easily proved:

$$l = L\cos\theta \tag{1.a}$$

$$h = H\sin\psi \tag{1.b}$$

As described in Section 2, let us assume that there are many identical curves in the scene, oriented in various random directions (θ's and ψ's). If they are really many and randomly oriented, there is a good chance that the θ of some curve is close to zero, and also that some curve has ψ close to $\pi/2$. Hence, the unknown real 3D length L and height H can be estimated using the maximum lengths among the l's and h's. This estimations scheme is represented as Eq. (2):

$$L \geq \max_i \{l_i\} \tag{2.a}$$

$$H \geq \max_j \{h_j\}, \tag{2.b}$$

where i or j is the suffix of each curve.

This scheme does not require that the i and j that give the maxima are the same. The necessary condition for the scheme is that at least one base line among all the curves is parallel to the image plane, and that at least one direction of the height of a curve is also parallel to the image plane.

The assumptions of a large number of curves and of randomness are sufficient conditions for the estimation scheme being accurate. Although the example given later will be interpreted in the way described above, that is, by simply using the

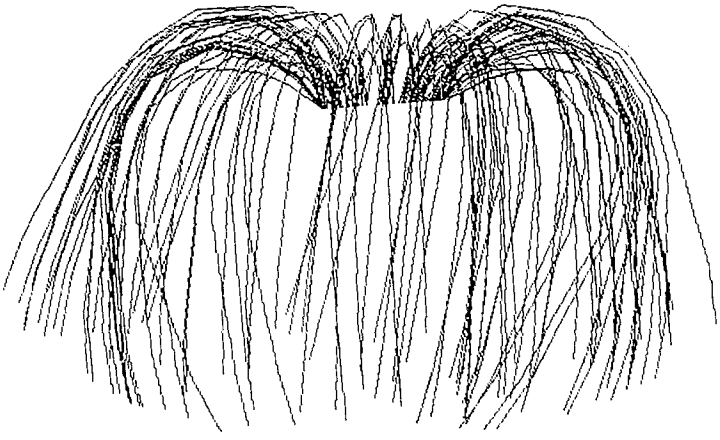

FIG. 3. House plant.

maxima, a more mathematically precise estimation is available. By Eq. (1), the distributions of l's and h's should have particular shapes related to the cos and sin curves. The accurate maximum value for l or h can be estimated by fitting the theoretical distribution curve to the data. This scheme would be effective in the case where only a small number of curves are given.

Once the real length L and the height H have been obtained, interpreting the 3D information is easy. First, every θ_i and ψ_i is determined by l_i and h_i as in Eq. (3):

$$\theta_i = \cos^{-1}(l_i/L) \tag{3.a}$$

$$\psi_i = \sin^{-1}(h_i/H). \tag{3.b}$$

Mathematically θ and ψ each have two possibilities. Combining these, there are four answers that satisfy Eq. (3). In practical cases, this ambiguity would not necessarily be serious. Some natural heuristics or some other information would be used to resolve the ambiguity.

Suppose that the θ and ψ for every planar curve have been determined. Because θ and ψ designate the plane which contains the curve, a z axis value relative to the point P for every point on the curve can then be easily obtained.

EXAMPLE. A house plant. Figure 3 is a simplified house plant. Many thin leaves are hanging down from a pot. Assume that every root of a leaf comes out from the circular edge of the pot, and every leaf has the same shape. Let p be the root point of a leaf, and q the end point of the leaf. Then the base line l is the line pq and the height h is the maximum distance between the leaf and the base line. Choose the X axis as the horizontal direction, the Y axis as the vertical direction, and the Z axis as coming toward the viewer perpendicular to the XY plane. The (x, y) coordinate values of every leaf are given. The 3D information can then be recovered by the scheme described above.

Now, heuristics specific to this example may be introduced. If θ is positive, that is, the leaf is coming toward the viewer, the root of the leaf may be on the front edge of the pot and ψ is between $\pi/2$ and $3\pi/2$. On the other hand, if θ is negative, that is, the leaf is going away from the viewer, the root of the leaf may be on the

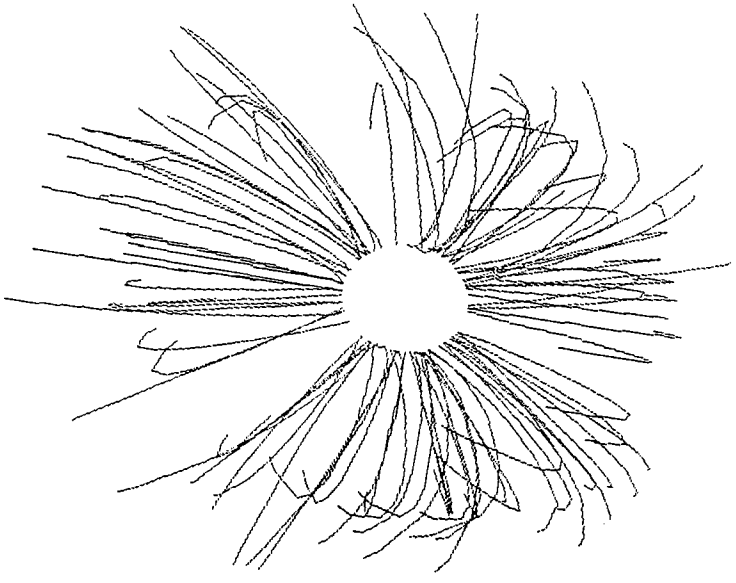

FIG. 4. Top view of interpreted 3D house plant.

back edge of the pot and ψ is between $-\pi/2$ and $\pi/2$. We may also assume that we are given the z position of the pot.

These are natural heuristic interpretations. However, there is still an ambiguity; namely, there is no information about the sign of θ in a silhouette image like Fig. 3. Here, we choose the sign of θ randomly.

The complete 3D shape can then be recovered from the image in Fig. 3. Figure 4 is the top view of the recovered 3D house plant, that is, the silhouette in the XZ plane. Figure 5 is another side view, the silhouette in the YZ plane. The results look very natural.

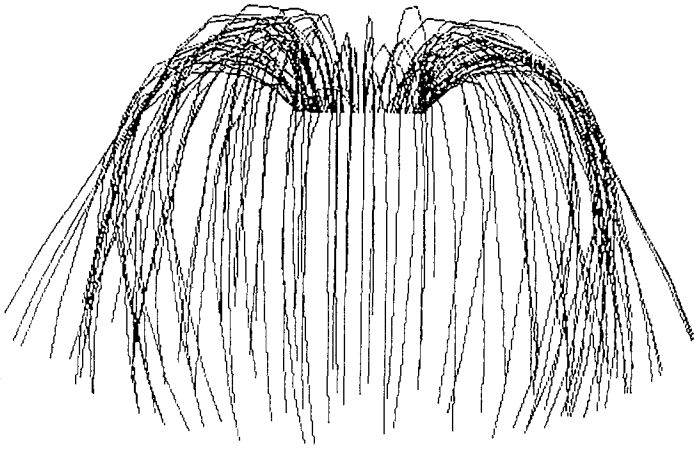

FIG. 5. Side view of interpreted 3D house plant.

Generally, the interpreted result in the XY plane is exactly the same as the original image Fig. 3. Thus, we will skip all the XY plane results in the examples.

3.2. *Length and Angle*

Many natural or artificial objects have spines, and around the spine many needles of the same length branch out at a fixed angle in various directions. If there really are many needles in random directions, it is possible to interpret the 3D shape of such an object from its image.

Let the spine be at an unknown angle θ to the image plane, and let many needles whose lengths all have the unknown value L be distributed in random directions around the spine, at an unknown fixed angle A to the spine.

Generally, the needles are scattered over many points on the spine. For simplicity, however, a preprocessing of the image is useful in order to make it easy to understand our interpretation scheme. The preprocessing causes all the needles to be temporarily gathered together at a single point p on the spine; that is, every needle is translated along the spine to the point p without changing its length and angle. The geometrical relations at this stage are illustrated in Fig. 6. Passing through the end points of the needles on the image, a hypothetical ellipse can be constructed. Let the center of the ellipse be q. The three lengths a, b, and c shown in Fig. 6 are important parameters for our interpretation scheme.

Rough estimates of these parameters are available in a very simple way without the need for preprocessing. The parameters can be estimated using the longest and

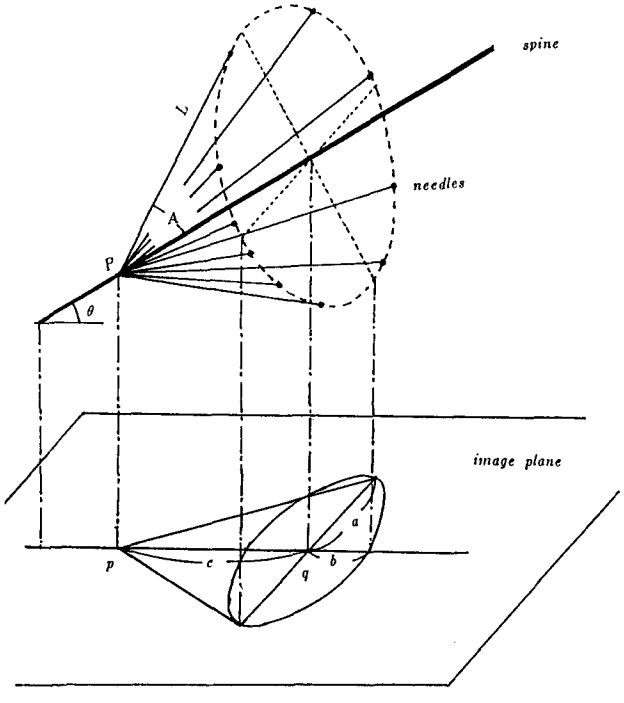

FIG. 6. Geometrical relations on spine and needles.

shortest lengths of the needles which are in almost the same direction relative to the spine, and finding the maximum distrance of the end points of these needles from the spine. Then the following geometrical relations are easily proved:

$$a = L \sin A \tag{4.a}$$

$$b = a \sin \theta \tag{4.b}$$

$$c = L \cos A \cos \theta. \tag{4.c}$$

Solving these equations, we obtain

$$\theta = \sin^{-1} \frac{b}{a} \tag{5.a}$$

$$A = \tan^{-1} \frac{\sqrt{a^2 - b^2}}{c} \tag{5.b}$$

$$L = \frac{a\sqrt{a^2 - b^2 + c^2}}{\sqrt{a^2 - b^2}}. \tag{5.c}$$

These solutions provide complete 3D information, but there is still an ambiguity between two possible solutions that are mirror images. This ambiguity depends on the sign of b. In real examples, it is usually not an important problem and may be determined using heuristics or domain specific information.

EXAMPLE. A pine branch. A pine branch is a typical example that has a spine and needles. In Fig. 7 a simplified pine branch is shown. It has a spine which is divided into four sections. Let us assume that the spine is coming toward the viewer, that is, the θ's are positive. In each spine section, there are many needles which have the same lengths and make the same angles relative to the spine in 3D space.

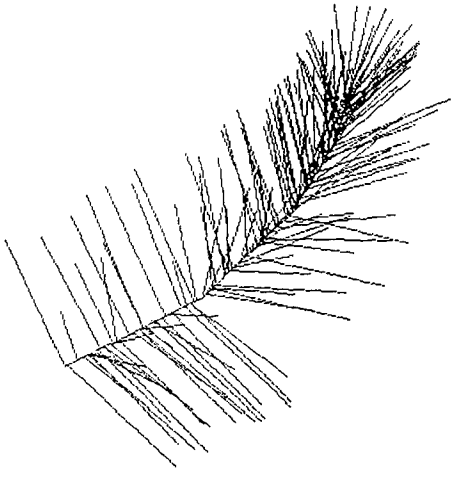

FIG. 7. Pine branch.

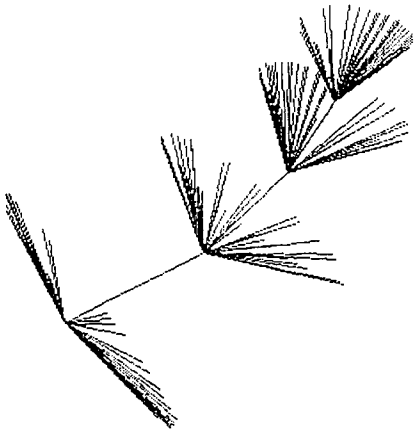

FIG. 8. Preprocessed pine branch.

As in the previous example let the X axis be in the horizontal direction, the Y axis in the vertical direction, and the Z axis perpendicular to the XY plane.

In Fig. 8, an image of the preprocessed pine branch is shown. The preprocessing is done for convenience in order to provide a better intuitive understanding of the parameters. It is not difficult to fit an ellipse to the end points of the needles. However, the actual interpretation was carried out in the simple way described above, using the lengths in the special positions.

For each section, θ, L, and A can be recovered. These provide relative information. Let us finally assume an absolute Z axis position for the root point of the branch, and interpret each section as connecting to the top of the previous section.

The top view, that is, the silhouette projected onto the XZ plane, of the interpreted 3D shape is shown in Fig. 9. The side view, the YZ plane silhouette, is shown in Fig. 10.

The accuracy depends on the image. In this case it depends on how close to the special position the needles can be found.

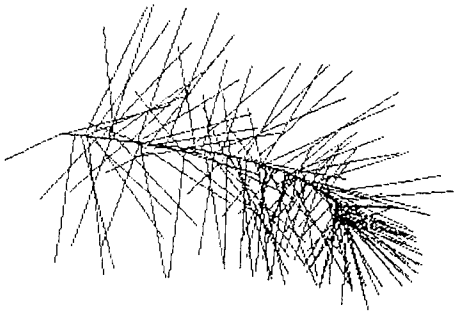

FIG. 9. Top view of interpreted 3D pine branch.

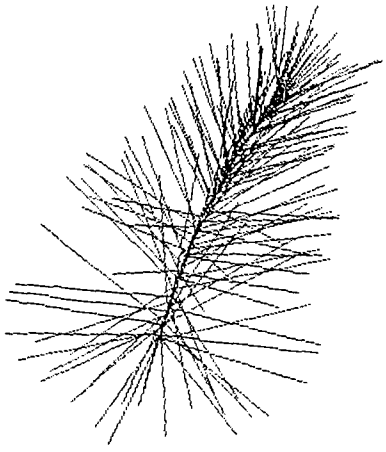

FIG. 10. Side view of interpreted 3D pine branch.

3.3. Triangle

Many objects can be represented by fitting triangles. A triangle is a natural and simple way to specify a plane in the space. Assume that there are many identical objects in the scene represented by triangles and that they are oriented in various directions. Then 3D interpretation of the objects from their images is possible. An advantage of using triangles is that since some triangles may be smaller than others, a scale factor can be estimated.

The situation assumed above is very common in natural objects, such as animals or plants. They are similar to each other, tend to occur grouped together, and some of them are immature, having similar shapes but smaller than the adults.

We assume that preprocessing provides reliable information about the correspondences among the triangles, that is, every vertex or edge can be recognized and distinguished from the others.

For a triangle which is not small, notation is defined as shown in Fig. 11. The vertex A is chosen as a root point, the edge AB is chosen as a base line, and the perpendicular line CD is drawn. The length of the perpendicular line is H, the length of the line AD is L, and the length of the line DB is R. The quantities H, L, and R determine the triangle.

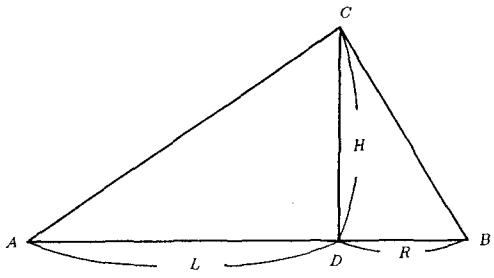

FIG. 11. Notation for a triangle.

Next, a small triangle is represented using a scale factor λ ($0 < \lambda < 1$):

$$H' = \lambda H \tag{6.a}$$

$$L' = \lambda L \tag{6.b}$$

$$R' = \lambda R. \tag{6.c}$$

H', L', and R' are the lengths of the lines corresponding to the sides of a normal size triangle. H, L, R, and λ are unknown values.

Now, assume that many triangles T_i lie in various orientations in the scene. Each of them is defined by H, L, R, and λ_i. All of these triangles are projected onto the image plane XY.

Let the triangle T_i have root vertex A at z_0 on the Z axis; let the angle between the base line AB and the XY image plane be θ_i, and let the angle between the perpendicular line DC and the XZ plane be ψ_i.

The geometrical relations between the triangle and the X, Y, and Z axes are illustrated in Fig. 12. For simplicity, in Fig. 12 and from now on, the suffixes i are omitted. The triangle T is projected onto the image plane and forms a triangle t. The triangle t can be represented in the same way as T: the length of the perpendicular line ce is h, the length of the line ae is l, and the length of the line eb is r.

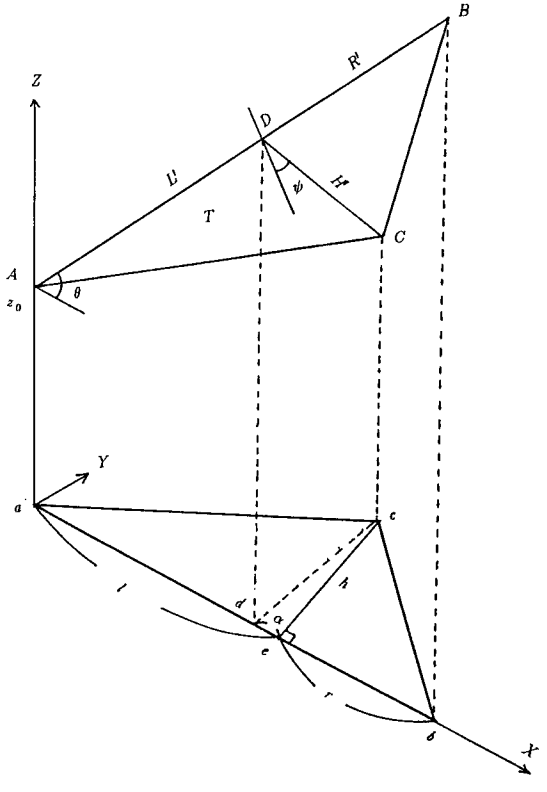

FIG. 12. Geometrical relations for a triangle and its image.

In the image, the point d is the projection of the point D, the root of the perpendicular line in space. Generally the points d and e are not the same, having disparity that depends on the angles θ and ψ. The position of the point d is unknown. Let the unknown distance between d and e be α. H, L, R, λ, θ, and ψ are unknown, and h, l, and r can be measured on the image. The relations between them are as follows:

$$\alpha = H'\cos\psi\sin\theta \tag{7.a}$$

$$h = H'\sin\psi \tag{7.b}$$

$$l = L'\cos\theta + \alpha \tag{7.c}$$

$$r = R'\cos\theta - \alpha. \tag{7.d}$$

The first step of interpretation is the estimation of the real lengths H, L, and R. The basic scheme is the same as that of Section 3.1. If there really are many triangles in random orientations, there is a good possibility that at least one of the edges of a triangle is close to flat. Hence, the estimation can be carried out by using the maximum length. Once the lengths of the edges are obtained, they are easily converted to the lengths H, L, and R.

This is the general, fundamental scheme used above. However, by taking advantage of a triangle's special properties, a more sophisticated method of estimation is available.

The general idea is sketched in Fig. 13. Let the real lengths of the edges be L_1, L_2, L_3, and the lengths in the image be l_1, l_2, and l_3. Now, among the many triangles in the image, some triangle has maximum length of l_1; let the maximum value be $l_{1\,\max}$, and in that triangle let the other two lengths be l_{12} for l_2, and l_{13} for l_3. At the same time, there are two other triangles which have the maximum lengths of l_2 and l_3. The notation for these triangles is shown in Fig. 13.

Consider the trinagle abc where the length ab in the image is the maximum, $l_{1\,\max}$. The line ab in the space can be assumed to be parallel to the image plane; hence the only distortion in the shape of the image triangle is caused by a rotation around the axis ab in space. Depending on the angle of rotation, the point c moves along the dotted line perpendicular to the line ab in Fig. 13.

On the other hand, from another triangle, $l_{3\,\max}$ has been obtained, which is usually longer than l_{13}. It is then a reasonable condition that triangle abp is one of the largest triangles, even though we could not find the length bp for line l_2 in the actual image.

The estimated length bp may be longer than the $l_{2\,\max}$, the maximum value found in the image. Thus, this is a superior scheme as compared with the elementary scheme described above.

Using l_{13}, l_{12}, and $L_{2\,\max}$, l_2^*, the length of bp is obtained by

$$\{l_{3\,\max}\}^2 - \{l_{13}\}^2 = \{l_2^*\}^2 - \{l_{12}\}^2. \tag{8}$$

Then, the estimation of L_2 is carried out by choosing the maximum value among $l_{2\,\max}$, l_2^* obtained above, and another l_2^* obtained by the same scheme from the

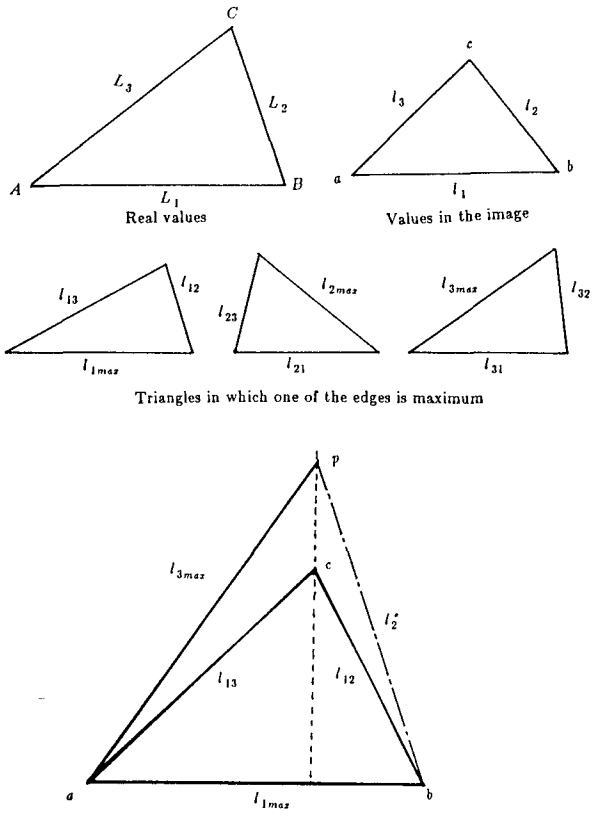

FIG. 13. Estimating the lengths of edges.

triangle which has $l_{3\,\max}$. The complete results are represented in Eq. (9):

$$L_1 \geq \max\left\{ l_{1\,\max}, \sqrt{l_{21}^2 + l_{3\,\max}^2 - l_{23}^2}, \sqrt{l_{31}^2 + l_{2\,\max}^2 - l_{32}^2} \right\} \tag{9.a}$$

$$L_2 \geq \max\left\{ l_{2\,\max}, \sqrt{l_{12}^2 + l_{3\,\max}^2 - l_{13}^2}, \sqrt{l_{32}^2 + l_{1\,\max}^2 - l_{31}^2} \right\} \tag{9.b}$$

$$L_3 \geq \max\left\{ l_{3\,\max}, \sqrt{l_{13}^2 + l_{2\,\max}^2 - l_{12}^2}, \sqrt{l_{23}^2 + l_{1\,\max}^2 - l_{21}^2} \right\}. \tag{9.c}$$

Once L_1, L_2, and L_3 are obtained, they are easily converted to H, L, and R. Then, Eq. (6) and Eq. (7) are combined together to yield

$$h = \lambda H \sin \psi \tag{10.a}$$

$$l = \lambda L \cos \theta + \lambda H \cos \psi \sin \theta \tag{10.b}$$

$$r = \lambda R \cos \theta - \lambda H \cos \psi \sin \theta. \tag{10.c}$$

H, L, and R are constant parameters over all triangles. The lengths h, l, and r can be measured on each triangle in the image. Equation (10) can then be solved for

λ, θ, and ψ for each triangle. The solutions are

$$\lambda = \sqrt{\frac{\eta \pm \sqrt{\gamma}}{2\xi}} \tag{11.a}$$

$$\psi = \sin^{-1}\left\{\frac{h}{\lambda H}\right\} \tag{11.b}$$

$$\theta = \cos^{-1}\left\{\frac{(r+l)}{\lambda(R+L)}\right\}, \tag{11.c}$$

where

$$\xi = H^2(R+L)^2 \tag{11.d}$$

$$\eta = h^2(R+L)^2 + H^2(r+l)^2 + (lR - Lr)^2 \tag{11.e}$$

$$\gamma = \left\{h^2(R+L)^2 - H^2(r+l)^2\right\} + (lR - Lr)^2\left[(lR - Lr)^2 + 2\left\{h^2(R+L)^2 + H^2(r+l)^2\right\}\right]. \tag{11.f}$$

Formally, there are two answers for λ; let the smaller one be λ^- and the bigger one λ^+. It is easily proved that λ^- is not a solution of Eq. (10), because λ^- satisfies

$$\sin\psi = \frac{h}{\lambda^- H} > 1 \tag{12.a}$$

$$\cos\theta = \frac{(r+l)}{\lambda^-(R+L)} > 1. \tag{12.b}$$

Thus λ^+ is the only answer for λ; however, there are still two possible answers for θ and ψ. They are mirror images, just as in the previous examples. In this case, they are summarized by

$$\begin{Bmatrix} \lambda^+ \\ \psi \\ \theta \end{Bmatrix}, \quad \begin{Bmatrix} \lambda^+ \\ \pi - \psi \\ -\theta \end{Bmatrix}. \tag{13}$$

The other two answers using other combinations of ψ, θ, $\pi - \psi$, and $-\theta$ do not satisfy Eq. (10). Thus (13) gives complete information about the 3D shape of each triangle.

EXAMPLE. Flowers. Many natural objects satisfy the assumptions required in the triangle interpretation scheme. Leaves or flowers of plants are good examples. They have almost the same shapes and are easily represented by triangles. Moreover, some of them are small in size.

In this section, an interpretation for a collection of flowers is carried out. Figure 14 is the input image in which there are many flowers, some of which are small. Now, assume that all petals are the same shape and planar. Then they are easily

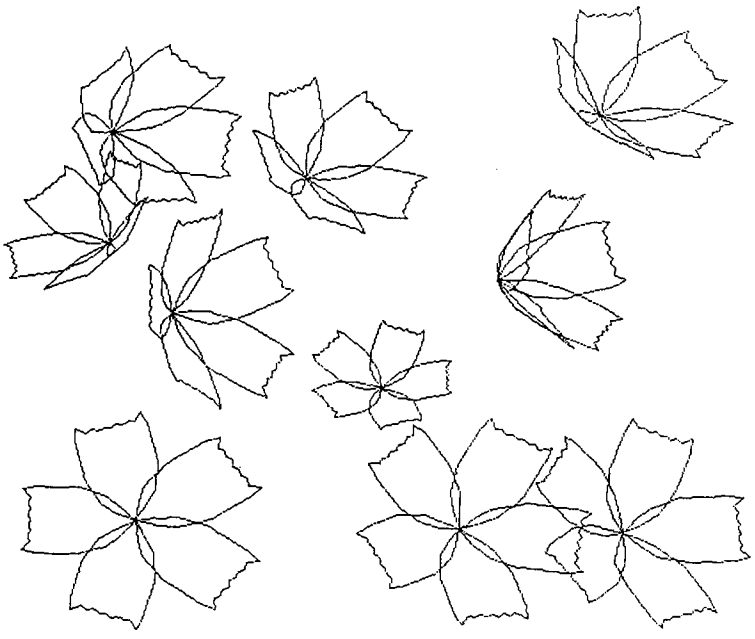

FIG. 14. Flowers.

represented by triangles, where the three vertices of the triangle are a root point and the two extreme points of a petal.

As in the previous examples, on the image plane the X axis is in the horizontal direction, the Y axis in the vertical direction, and the Z axis perpendicular to the XY image plane. The center cores of the flowers are the common points; that is, the root points of the petals in a flower have the same Z axis values. In a more general scene, other information such as location on stalks or branches will give absolute Z values for the cores of flowers. Here, we assume that Z axis values are given a priori for every flower.

In Fig. 15 flowers represented by triangles are illustrated. The first step is the estimation of the real lengths of the triangle edges. The values are converted to H, L, and R, and Eq. (10) is solved.

Combined with the given Z position of the flower core, the solution for λ, θ, and ψ gives 3D information for each petal. We choose one of the two answers for which the flower shape looks natural, that is, the petals have edges that are close together.

The interpretation results for the triangles are shown in Figs. 16 and 17. Figure 16 is the top view, that is, the silhouette image of the interpreted flowers projected onto the XZ plane. Figure 17 is the other side view on the YZ plane. Since each triangle determines a plane, the original line drawings in Fig. 14 are easily interpreted, because they are contained in that plane. The final results are shown in Figs. 18 and 19, which are analogous to Figs. 16 and 17. Although some petals are interpreted as lying in unnatural orientations, generally speaking, the interpretation was successful for these flowers.

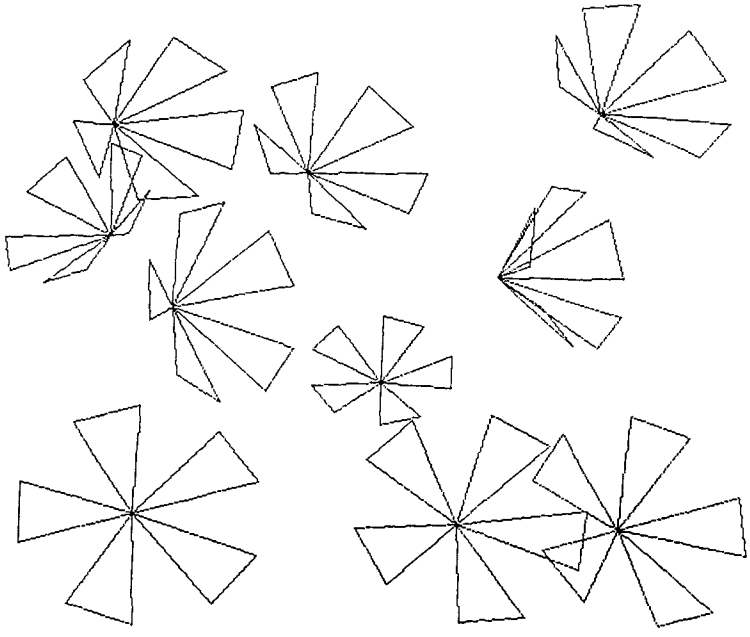

FIG. 15. Flowers represented by triangles.

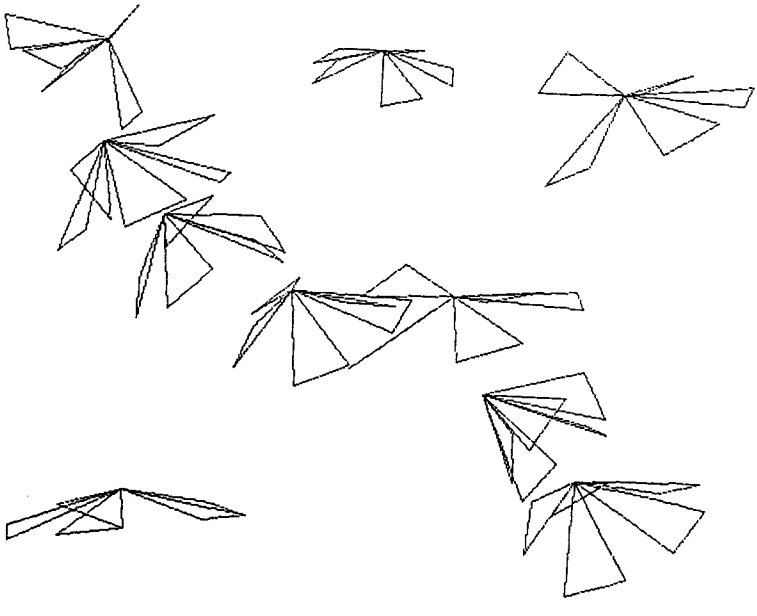

FIG. 16. Top view of interpreted triangle flowers.

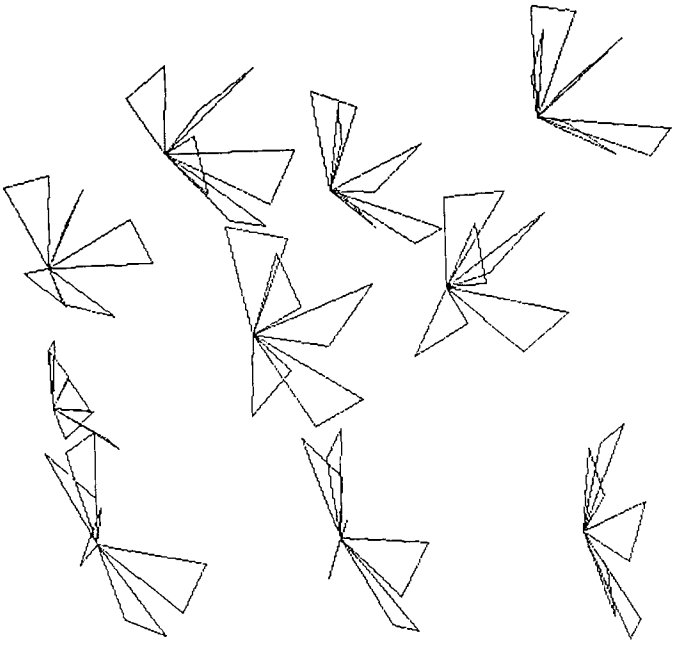

FIG. 17. Side view of interpreted triangle flowers.

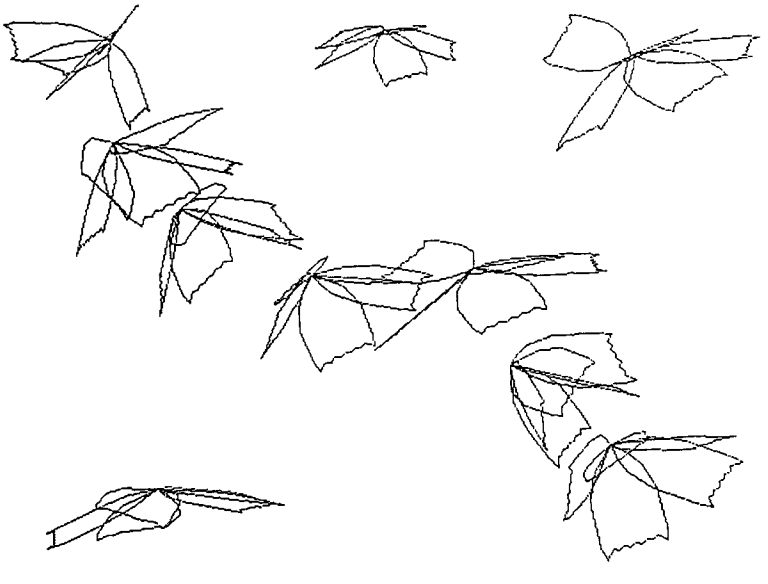

FIG. 18. Top view of interpreted flowers.

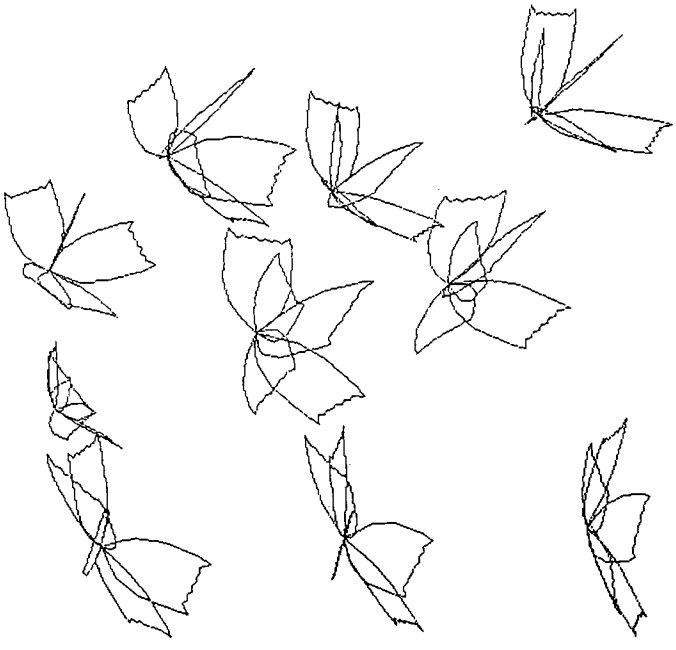

FIG. 19. Side view of interpreted flowers.

3.4. *Angle*

In this section, a model which is based only on angles in introduced. The assumption for the interpretation is that many identical angles are distributed in space in random orientations. An angle is a more ambiguous geometrical feature than the features in the previous sections, because there is no information on the lengths of the line segments. Even though it provides less information, it is possible to estimate the real angle, and from it the 3D shape, as shown in the following example.

The geometrical relations are illustrated in Fig. 20. In 3D space the angle POQ is oriented in the direction determined by two angles θ and ψ, where θ is the angle between the line OP and the image plane, and ψ is the angle between the plane containing the angle POQ and the plane perpendicular to the image plane and containing the line OP.

The relation between the spatial angle A and the angle a on the image plane is

$$\tan a = \frac{\sin \psi \tan A}{\cos \theta + \cos \psi \sin \theta \tan A}. \tag{14}$$

An overview of the appearance of $\tan a$ as represented by Eq. (14) is illustrated in Fig. 21, as a function of θ and ψ, where $0 \leq \theta \leq \pi/2$ and $0 \leq \psi \leq \pi/2$. In the area outside of Fig. 21, the $\tan a$ function has many singular points.

Assume that many angles are distributed in random orientations. In other words, over the θ, ψ plane, where $-\pi \leq \theta \leq \pi$ and $-\pi \leq \psi \leq \pi$, many angle samples are

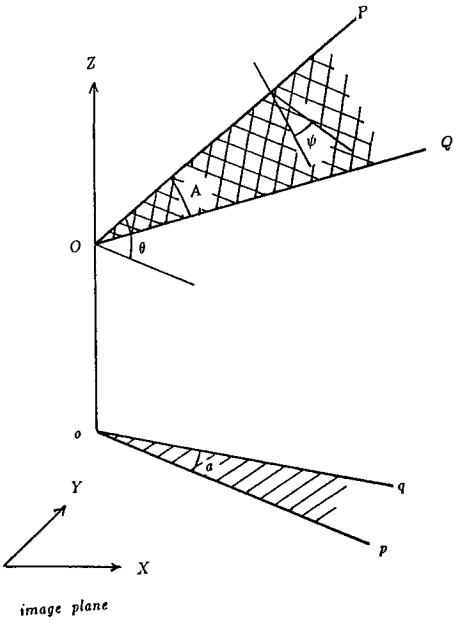

FIG. 20. Geometrical relations for an angle and its image.

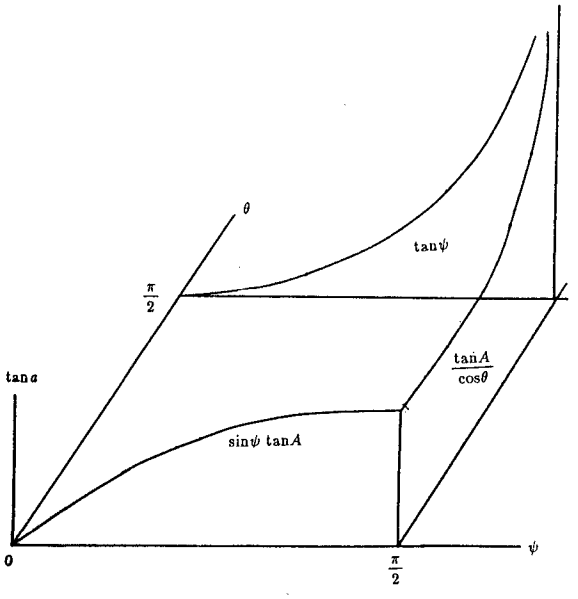

FIG. 21. Overview of the appearance of tan a.

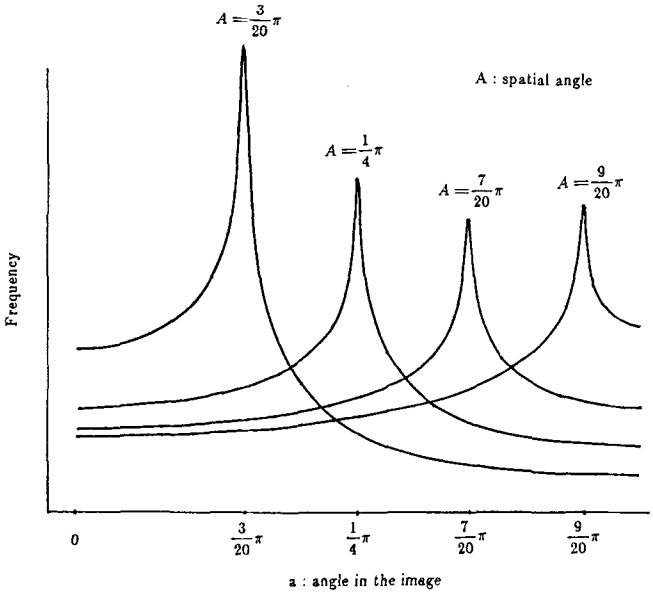

FIG. 22. Histogram of angles in the image.

distributed with uniform density. Then, each sample is projected onto the image plane using Eq. (14).

Let us consider the statistical frequency distribution of $\tan a$, that is, how many samples are in a unit interval of $\tan a$ values. The frequency at a given value of $\tan a$ depends on the gradient of the $\tan a$ function. For example, in a small unit area around the point, $\theta = 0$ and $\psi = \pi/2$, the gradient is nearly equal to zero, and we have many samples which satisfy $\tan a = \tan A$. On the other hand, in a small unit area around the point $\theta = \pi/2$ and $\psi = \pi/2$ the gradient is nearly infinite, and samples in the unit area are scattered over a wide range of $\tan a$.

The discussion above suggests that the density distribution of $\tan a$ is not uniform over its range, even though the samples were originally distributed on the θ, ψ plane uniformly. Moreover, it is likely that there are more samples that satisfy $\tan a = \tan A$. This can be proved by analysis of Eq. (14). For example, let the real spatial angle A be $\pi/4$ and let many of the $\pi/4$ angles be distributed uniformly on the θ, ψ plane; then the histogram of the image angle a has a peak that is exactly at $\pi/4$ as shown in Fig. 22. Some other cases of A's equal to $\frac{3}{20}\pi$, $\frac{7}{20}\pi$, and $\frac{9}{20}\pi$ are also shown in the same figure. These examples indicate that it is possible to guess the real spatial angle by looking for a significant peak on the density distribution.

Actual 3D information for the spatial z position is easily obtained as follows. For convenience in explanation, let us translate the vertex to the origin and assume a 3D direction for one of the edges, for example OP in Fig. 20, where the 3D direction is given by a point on $OP = (\alpha, \beta, \gamma)$. Then the direction of the other edge OQ is represented in the same way by a point on $OQ = (x, y, z)$, where z is unknown. We

can obtain z from

$$z = \frac{b \pm \sqrt{b^2 - ac}}{a} \tag{15.a}$$

where

$$a = (\alpha^2 + \beta^2 + \gamma^2)\cos^2 A - \gamma^2 \tag{15.b}$$

$$b = \gamma(\alpha x + \beta y) \tag{15.c}$$

$$c = (\alpha^2 + \beta^2 + \gamma^2)\cos^2 A(x^2 + y^2) - (\alpha x + \beta y)^2. \tag{15.d}$$

Generally there are two solutions for z, but in some cases only one of them is available. This can be determined by verifying the conditions that

$$-\frac{\pi}{2} \leq A \leq \frac{\pi}{2} \tag{16.a}$$

and

$$\cos A = \frac{\alpha x + \beta y + \gamma z}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}\sqrt{x^2 + y^2 + z^2}} \geq 0. \tag{16.b}$$

In the special case where $b^2 - ac < 0$, there is no solution.

EXAMPLE. A tree. A good example of interpretation based on angles is a tree, where the spatial angles between the trunk and the branches have the same value. Recent research on graphics has created beautiful tree images in which the angles are basically fixed [7]. In our example, a tree image was generated in a rather primitive way: the angles between the branches and the trunk are fixed at 30°; the lengths of the branches are defined using a random function; and the directions of the branches are random. The image is shown in Fig. 23. The image has about 3500 angles. The histogram of the image angles is shown in Fig. 24, which is the result obtained by a simple filtering operation on the original histogram. This filtering eliminates small jagged peaks.

The peak is located at 29°. Thus we can guess that the real spatial angles in the tree image shown in Fig. 23 are equal to 29°. The correct exact answer is 30°. Since the histogram was constructed for angles spaced 1° apart, this is the minimum error.

Let assume that the X axis is horizontal and the Y axis is vertical in Fig. 23, and that the root trunk of the tree is parallel to the XY plane. Then we can recover the 3D shape of the tree using Eq. (15).

For many branches we have two possible solutions. In practice it would not be difficult to choose one of the two answers, because we may have much more information on the branch directions, such as detailed image data at the branching points, or occlusion information between branches.

It is possible, however, to interpret the example by using the following "pure" algorithm: We first choose one of the two solutions at random, and continue to reconstruct the 3D shape. If a sub-branch yields a contradiction, that is, it has no

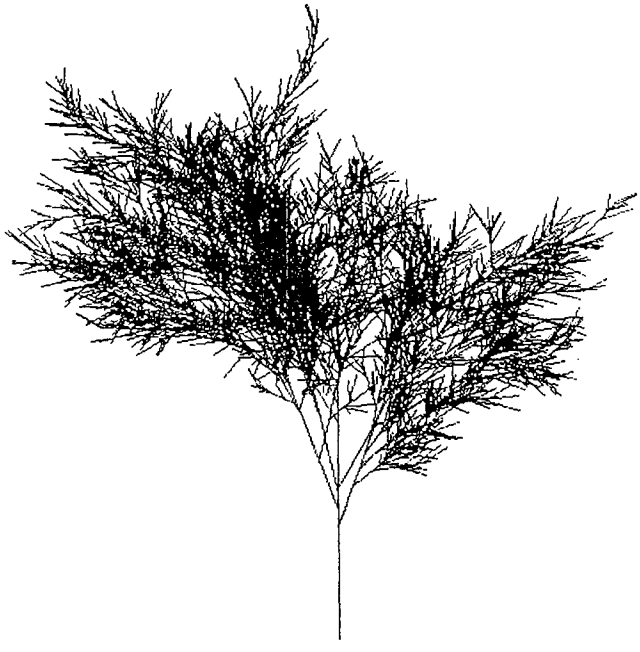

FIG. 23. Tree image.

possible solution, we trace back to the nearest parent trunk which has another possible solution, discard the old interpretation, and try to reconstruct the 3D shape using the other solution. Because of the estimation error for the angles, the possibility remains that some branches have no solution, even after trying all possibilities. In our example, 52 branches had no solutions. For simplicity, they were interpreted as parallel to the XY plane.

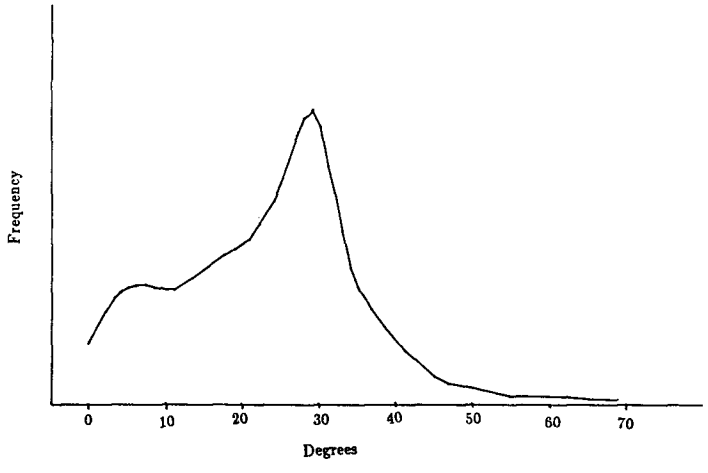

FIG. 24. Histogram of angles in the tree image.

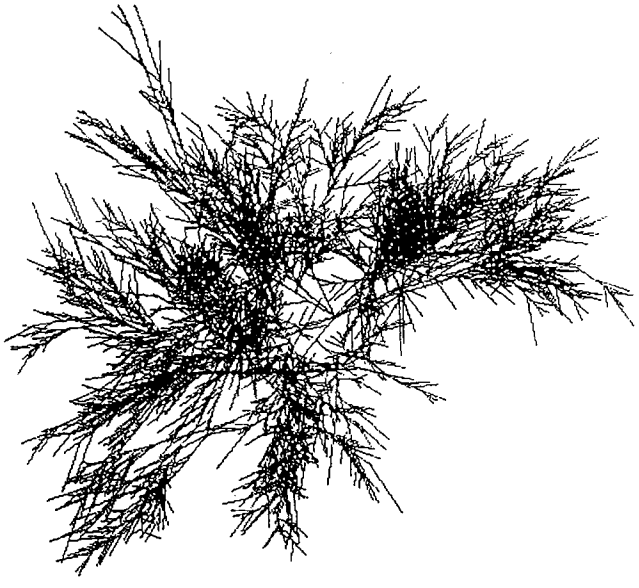

FIG. 25. Top view of interpreted 3D tree.

FIG. 26. Side view of interpreted 3D tree.

The top view, the silhouette of the recovered 3D tree on the XZ plane, is shown in Fig. 25. The other side view, the silhouette on the YZ plane, is shown in Fig. 26. For this 3D interpretation, the input information was very vague, because there was no information about lengths. However, the result is quite natural.

4. DISCUSSION

4.1. Advantages

Some advantages of our interpretation scheme are summarized here. First, the scheme is a quite general approach for interpreting various kinds of natural or artificial objects. Its assumptions of many identical objects in the scene and of randomness in their orientations are characteristic of many natural scenes that we observe in daily life. Not only natural objects, but also artificial objects such as industrial parts in a pile, or architectural features of a large building, may satisfy the assumption. The scheme does not require object models. Domain specific knowledge is not important, though it can help in the interpretation by reducing the ambiguity of the results obtained by the basic scheme.

Second, the mathematical cncepts on which the scheme is based are quite simple. Generally, the equations can be solved by elementary algebra. They do not involve complicated differential components, which are usually unstable. The data that are required for the equations are basically a length between two points or an angle between two lines. Such quantities can be easily measured with high accuracy on the given image.

From the standpoint of the complexity of the algorithm, the scheme is very simple, since it does not require any iterative calculations. The interpretation of every feature is carried out simply by calculating the solution of the equations, whose parameters are measured in the image. In fact, the implementation of the scheme required only a small amount of work. For each example in the last section, the programs contained about 30–70 Lisp functions, half of which were display controls.

Finally, from a psychological standpoint, it is probably not difficult for a human to imagine the 3D shape of the house plant in Section 3.1. As regards quantitative accuracy, the program works almost perfectly on the data. The numerical calculations in the interpretations are apparently very accurate. On the other hand, the accuracy of human beings would be very poor. In the case of the tree in Section 3.4, it may be possible for humans to vaguely infer the most frequent angle, but it would be difficult for most humans to imagine the consistent 3D shape shown in Figs. 25 and 26.

4.2. Problems

The low level feature extraction stage, including feature selection and defining the correspondences between features, was skipped in this paper. If the scheme is to be applied to the interpretation of a real natual scene, those low level processes become serious problems.

The examples given here were all wire-frame-like objects. No occlusion between objects was considered in this paper. In the case of solid objects, the interpretation would become somewhat more complicated, although essentially the scheme would be the same. In some situations, occlusion helps the interpretation, because it gives

relative 3D position. For example, in case of the tree example, occlusions between branches would be of great help in obtaining 3D shape.

A more fundamental problem is that it is not realistic to expect natural objects to be exactly identical to each other in their lengths or angles. If the lengths or angles vary, the 3D shape obtained by the scheme would be deformed. Intrinsically, it is impossible to infer the exact answer, when the measurements are variable. The scheme proposed here gives the most credible answer from a statistical standpoint by assuming that the features are exactly the same.

4.3. Future Work

In case of the house plant in Section 3.1, it may be possible to derive a more sophisticated interpretation, if the continuous curves of the leaves can be compared and the disparity between them can be evaluated and interpreted. Generally speaking, continuous curves are difficult objects in the sense of the scheme proposed here, because it is difficult to pick up exact feature points and measure lengths or angles.

For more complicated scenes, a combination of the techniques proposed in Section 3 may be required. If the initial feature extraction is sufficiently complete, such a combination should not be difficult.

Another interesting extension of the scheme is possible. In this paper, the condition that many objects are present in the scene is required. Suppose that only one object is shown and that it is moving around and rotating in a random way. Then a time sequence of images gives the same information that is assumed in this paper. For example, suppose the skeleton of a person moves around in various ways for a while; then gradually the real lengths of its bones can be obtained, and eventually we can interpret its 3D features.

Another more practical example is a shape measurement method for industrial parts. Let us put a mark on the object at every point which is important for representing its 3D shape. These marks form many triangles. Let the object turn in various directions so that we can measure every distance between the marks. This information allows us to interpret the actual shapes of the triangles using the scheme in this paper, so that the 3D shape of the object can be determined. This approach may constitute a kind of learning scheme based on image sequences for 3D object recognition.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Larry S. Davis and Dr. Allen M. Waxman for their helpful comments.

REFERENCES

1. T. Kanade, Recovery of the three-dimensional shape of an object from a single view, *Artif. Intell.* **17**, 1981, 75–116.
2. H. G. Barrow and J. M. Tenenbaum, Interpreting line drawings as three dimensional surfaces, *Artif. Intell.* **17**, 1981, 75–116.
3. S. T. Barnard and A. P. Pentland, Three dimensional shape from line drawings, in *Proceedings, Image Understanding Wrokshop, 1983*, pp. 282–284.
4. T. O. Binford, Inferring surfaces from images, *Artif. Intell.* **17**, 1981, 205–244.
5. R. A. Brooks, Symbolic reasoning among 3-D models and 2-D images, *Artif. Intell.* **17**, 1981, 285–348.
6. A. P. Witkin, Recovering surface shape and orientation form texture, *Artif. Intell.* **17**, 1981, 17–45.
7. M. Aono and T. Kunii, Botanical tree image generation, *IEEE Comput. Graphics Appl.* **4**, 1984, 16–34.