The Constraints on Images of Rectangular Polyhedra

KEN-ICHI KANATANI

Abstract—**This paper discusses how polyhedron interpretation techniques are simplified if the objects are rectangular trihedral polyhedra. This restriction enables one to compute the spatial orientation of a given corner and its motion from its image in terms of polar coordinates, Eulerian angles, and quaternions. One can also interpret the shape and the face adjacency from local information only. The necessary constraints are listed, and some examples are given to compare the presented scheme to existing ones. The possible nonuniqueness of the interpretation is also discussed.**

Index Terms—**Computer vision, constraint, Eulerian angle, machine intelligence, quaternion, rectangular polyhedron, shape analysis, spatial orientation.**

I. Introduction

DETERMINATION of the shape, position, and orientation of an object from its projected image is one of the most fundamental problems of computer vision and machine intelligence. In general, reconstruction from only one projection is impossible unless some sort of prior knowledge or assumptions are given about the true shape of the object. Recognition becomes easier as the prior information increases. However, these assumptions, or "constraints," should preferably be such that they include many objects encountered in various practical applications. One frequent assumption is that the objects are polyhedra. This assumption, called the "blocks world model," was initiated by Guzman [1], Huffman [2], and Clowes, [3] and further developed by Waltz [4], Mackworth [5], Sugihara [6], Kanade [7], and others using various types of additional information such as shading and shadows, hidden lines, or paper-folding properties. Another approach is to confine our attention to a portion of the object which is regarded as a plane surface. In other words, the object under consideration is idealized as a plane surface equipped with various kinds of additional information such as light reflectance (Horn [8], Ikeuchi and Horn [9], etc.), texture (Witkin [10], Stevens [11], Kanatani [12], etc.), bounding curves (Kanatani [13], [14]), and so on.

In this paper, the objects are assumed to be rectangular trihedral polyhedra, i.e., polyhedra whose corners consist of three mutually perpendicular edges. Shape interpretation of such objects can be done, say, by Huffman's scheme [2], which is applicable to general polyhedra. However, the rectangularity constraint simplifies the interpretation process to a large extent, as first pointed out by Mackworth [15] and later demonstrated by Kanade [16]. Here, we describe this rectangularity constraint in rigorous but concise mathematical terms. This problem well deserves special attention and careful study, because rectangular polyhedra are the most frequently encountered manmade objects—buildings, furniture, machine parts, and so on. Our conclusions are as follows. First, as an obvious consequence, "quantitative" information is extracted as to the three-dimensional orientation and motion of the object from its two-dimensional image alone. All relevant mathematical relationships are listed in terms of polar coordinates, Eulerian angles, and quaternions. As for shape interpretation, the rectangularity constraint makes it possible to give "local interpretations" of the portions under consideration. This is in contrast to Huffman's scheme, which makes use of the "global consistency" of a polyhedron as a whole. Thus, the present study is also applicable to objects that are not polyhedra but "partly rectangular" objects (cf. Section V and Fig. 12). All the necessary constraints are listed and the interpretation procedure based on them is described. Some examples are given to illustrate of our scheme and compare it to existing ones. Some discussion about the possible nonuniqueness of the interpretation is also presented.

Manuscript received June 11, 1984; revised December 3, 1985. Recommended for acceptance by S. L. Tanimoto. This work was supported in part by the Defense Advanced Research Projects Agency and the U.S. Army Night Vision and Electro-Optics Laboratory under Contract DAAK70-83-K-0018 (DARPA Order 3206).

The author is with the Center for Automation Research, University of Maryland, College Park, MD 20742, on leave from the Department of Computer Science, Gunma University, Kiryu, Gunma 376, Japan.

IEEE Log Number 8608014.

II. Determination of Spatial Orientation of a Corner

Let us fix a Cartesian xy-coordinate system on the image plane and take the z-axis perpendicular to it so that the xyz-axes form a right-hand system. An object is projected orthographically onto the xy-plane along the z-axis. Given an image of a corner, we want to know its spatial orientation. If the true length of an edge is known, its spatial orientation is immediately given. Suppose the true size of the polyhedron is not known. Yet, the spatial orientation can be determined from the image alone if the corner is rectangular and in general position, and this is a well-known fact. For example, Kanade [16] made use of "gradient space." The relevant principles are found in classical 3-D geometry and are well known to mathematicians. However, they are usually scattered in books on mathematics intermingled with highly sophisticated math-

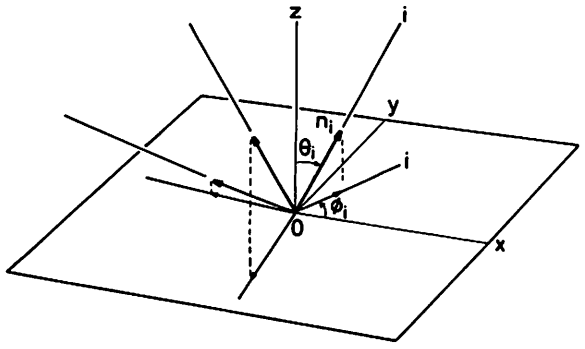

Fig. 1. An xyz-coordinate system is placed in such a way that the xy-plane is the image plane. An object is projected along the z-axis. The spatial orientations of the three edges at a corner are described by unit vectors emanating from the vertex along them or their spherical coordinates.

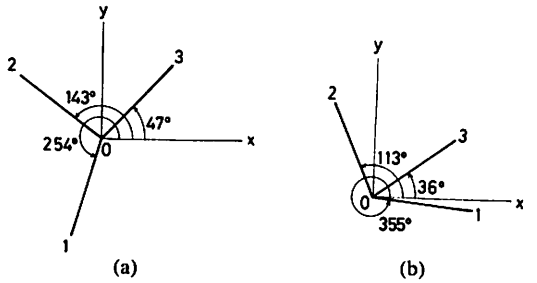

Fig. 2. Two images of a corner: (a) a "fork" and (b) an "arrow."

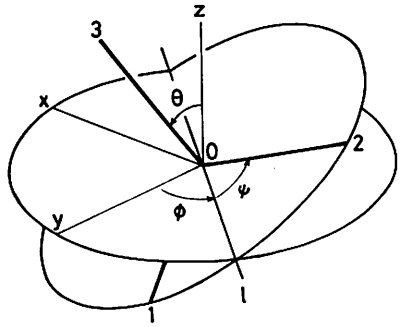

Fig. 3. Definition of the Eulerian angles θ, ϕ, ψ. Angles θ and ϕ are the spherical coordinates of the 3-edge. Let l be the intersection of the xy-plane and the 12-plane. Then, ψ is the angle of the 2-edge from l measured screwwise about the 3-edge.

ematical concepts—the representation of a Lie group and its Lie algebra, manifolds, topology, homotopy, spinors, etc. Hence, we summarize those facts that are relevant in the present context for the convenience of those who are not familiar with these kinds of mathematics. In the following, the term "corner" means both the corner of a 3-D polyhedron and its 2-D image consisting of a vertex and three edges emanating from it. We do not make a great distinction between a 3-D object and its 2-D image because we consider the case where the image faithfully represents the object.

Number the edges at the corner in question 1, 2, and 3 arbitrarily. Since the location of the corner is irrelevant for computation of orientation, the vertex is assumed to be at the origin. Let $\boldsymbol{n}_i$ be the unit vector emanating from the vertex along the i-edge (Fig. 1), and let θ_i and ϕ_i be its spherical coordinates, i.e.,

$$\boldsymbol{n}_i = (\sin\theta_i \cos\phi_i, \sin\theta_i \sin\phi_i, \cos\theta_i). \tag{1}$$

The condition that $\boldsymbol{n}_i$ is perpendicular to $\boldsymbol{n}_j$ is

$$\sin\theta_i \sin\theta_j (\cos\phi_i \cos\phi_j + \sin\phi_i \sin\phi_j) + \cos\theta_i \cos\theta_j = 0, \tag{2}$$

or

$$\tan\theta_i \tan\theta_j = -1/\cos(\phi_i - \phi_j). \tag{3}$$

This holds for $(i, j) = (1, 2), (2, 3), (3, 1)$. Solving the three equations for $\tan\theta_1$, $\tan\theta_2$, $\tan\theta_3$, we can express θ_1, θ_2, θ_3 in terms of ϕ_1, ϕ_2, ϕ_3. If all the edges go "upward," i.e., $0 < \theta_i < \pi/2$, $i = 1, 2, 3$, we have

$$\theta_1 = \tan^{-1}\sqrt{-\cos(\phi_2 - \phi_3)/\cos(\phi_1 - \phi_2)\cos(\phi_3 - \phi_1)},$$

$$\theta_2 = \tan^{-1}\sqrt{-\cos(\phi_3 - \phi_1)/\cos(\phi_2 - \phi_3)\cos(\phi_1 - \phi_2)},$$

$$\theta_3 = \tan^{-1}\sqrt{-\cos(\phi_1 - \phi_2)/\cos(\phi_3 - \phi_1)\cos(\phi_2 - \phi_3)}. \tag{4}$$

If the i-edge goes "downward," i.e., $\pi/2 < \theta_i < \pi$, then θ_i calculated above is replaced by $\pi - \theta_i$, i.e., by the "mirror image" with respect to the xy-plane. In deciding which one goes upward or downward, we must distinguish two configurations. One is the "fork" (or "Y") where $|\phi_i - \phi_j|$ is larger than $\pi/2$ for all (i, j) pairs, $i \neq j$ [e.g., Fig. 2(a)]. In this case, the three edges go either all upward or all downward, being the mirror images of each other. The other configuration is the "arrow" where $|\phi_i - \phi_j|$ is larger than $\pi/2$ for one (i, j) pair and less than $\pi/2$ for the other two pairs [e.g., Fig. 2(b)]. Then, either the side edges go downward and the central edge upward or the side edges go upward and the central edge downward, being the mirror images of each other. (Here, we do not consider the degenerate case where two edges are projected on the same line ("L" or "T"), i.e., we assume that the object is in "general position.")

Now, suppose the 1-, 2-, 3-edges form a right-hand system in that order. Consider the Eulerian angles θ, ϕ, ψ of the rotation that transforms the x-, y-, z-axes to the 1-, 2-, 3-edges (Fig. 3). The unit vectors $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{n}_3$ of (1) are expressed in terms of the Eulerian angles θ, ϕ, ψ by

$$\begin{aligned}
\boldsymbol{n}_1 &= (\cos\theta \cos\phi \cos\psi - \sin\phi \sin\psi, \\
&\quad -\sin\theta \cos\psi), \\
&\quad \cos\theta \sin\phi \cos\psi + \cos\phi \sin\psi, \\
\boldsymbol{n}_2 &= (-\cos\theta \cos\phi \sin\psi - \sin\phi \cos\psi, \\
&\quad -\cos\theta \sin\phi \sin\psi + \cos\phi \cos\psi, \\
&\quad \sin\theta \sin\phi), \\
\boldsymbol{n}_3 &= (\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta).
\end{aligned} \tag{5}$$

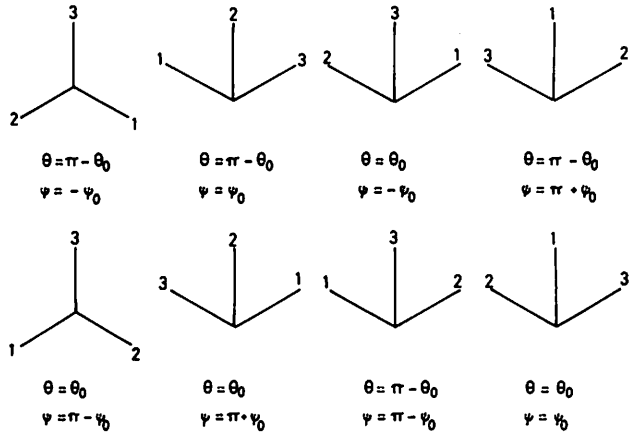

Fig. 4. Determination of Eulerian angles from the image configuration.

If ϕ_1, ϕ_2, ϕ_3 are the polar coordinates of the images of the 1-, 2-, 3-edges, respectively, as before, we get $\phi = \phi_3$ by definition. Comparing (1) and (5), we obtain

$$\tan \phi_1 = \frac{\cos\theta \sin\phi \cos\psi + \cos\phi \sin\psi}{\cos\theta \cos\phi \cos\psi - \sin\phi \sin\psi},$$
$$\tan \phi_2 = \frac{\cos\theta \sin\phi \sin\psi - \cos\phi \cos\psi}{\cos\theta \cos\phi \sin\psi + \sin\phi \sin\psi}. \quad (6)$$

Hence, θ and ψ are given as follows[1]:

$$\theta = \theta_0 \quad \text{or} \quad \pi - \theta_0,$$
$$\psi = \psi_0, -\psi_0, \pi - \psi_0 \quad \text{or} \quad \pi + \psi_0, \quad (7)$$

where

$$\theta_0 = \cos^{-1} \sqrt{\cot(\phi_3 - \phi_1) \cot(\phi_2 - \phi_3)},$$
$$\psi_0 = \tan^{-1} \sqrt{\tan(\phi_3 - \phi_1)/\tan(\phi_2 - \phi_3)}. \quad (8)$$

The choice of the value in (7) is made according to the configuration of the corner image as shown in Fig. 4. Since the unit vectors n_1, n_2, n_3 of (1) are given in terms of θ, ϕ, ψ by (5), the validity of Fig. 4 is seen by checking the sign of the z-component.

As is well known, a three-dimensional rotation is represented by a "quaternion" of the form $q = q_0 + q_1 i + q_2 j + q_3 k$, $(q_0)^2 + (q_1)^2 + (q_2)^2 + (q_3)^2 = 1$. If two rotations represented by the quaternions q, q' are composed in that order, the resulting rotation is represented by the quaternion

$$q'' = qq', \quad (9)$$

where multiplication of quaternions is the same as for real numbers except for the following rules:

$$ii = -1,$$
$$jj = -1,$$
$$kk = -1,$$
$$ij = -ji = k,$$
$$jk = -kj = i$$
$$ki = -ik = j. \quad (10)$$

If $q = q_0 + q_1 i + q_2 j + q_3 k$ represents a rotation, its inverse rotation is represented by its "conjugate" $q^* = q_0 - q_1 i - q_2 j - q_3 k$. Also, given a quaternion, the rotation axis and rotation angle which realize that rotation are easily obtained. The quaternion of the rotation about unit vector $n = (n_1, n_2, n_3)$ by rotation angle Ω (screwwise) is given by

$$q = \cos\frac{\Omega}{2} + n \sin\frac{\Omega}{2}, \quad (11)$$

where n is identified with quaternion $n_1 i + n_2 j + n_3 k$. In other words, the "scalar part" gives the rotation angle, and the "vector part" gives the orientation of the rotation axis, i.e.,

$$\Omega = 2\cos^{-1} q_0, \qquad n = (q_1, q_2, q_3)/\sin\frac{\Omega}{2}. \quad (12)$$

Thus, the use of quaternions makes it very easy to obtain compositions, inverses, rotation axes, and rotation angles, giving them in terms of explicit analytical expressions.

Now, the rotation that has Eulerian angles θ, ϕ, ψ is represented by

$$q = \cos\frac{\theta}{2}\left(\cos\frac{\psi+\phi}{2} + k \sin\frac{\psi+\phi}{2}\right) + \sin\frac{\theta}{2}\left(i \sin\frac{\psi-\phi}{2} + j\cos\frac{\psi-\phi}{2}\right). \quad (13)$$

Hence, we can describe the corner orientation by a quaternion which represents the rotation transforming the reference xyz-axes to the 1-, 2-, 3-edges of the corner. If its quaternion is $q = q_0 + q_1 i + q_2 j + q_3 k$, the unit vectors of the edges are given by

$$n_1 = ((q_0)^2 + (q_1)^2 - (q_2)^2 - (q_3)^2, 2(q_0 q_3 + q_1 q_2), 2(-q_0 q_2 + q_1 q_3)),$$
$$n_2 = (2(-q_0 q_3 + q_1 q_2), (q_0)^2 - (q_1)^2 + (q_2)^2 - (q_3)^2, 2(q_0 q_1 + q_2 q_3)),$$
$$n_3 = (2(q_0 q_2 + q_1 q_3), 2(-q_0 q_1 + q_2 q_3), (q_0)^2 - (q_1)^2 - (q_2)^2 + (q_3)^2). \quad (14)$$

If two corner images have quaternions q and q', the relative rotation is given by

$$q'' = q'q^*. \quad (15)$$

Hence, by (12) we can immediately determine the axis and angle of the rotation which transforms one corner to another. This fact can be used most effectively when we are tracing the motion of an object from its projected image alone.

[1]Since $\theta = \theta_3$ by definition, the first equation of (8) is actually equivalent to the last equation of (4).

Example 1: Consider the images in Fig. 2. For Fig. 2(a), which is a "fork," we have $\phi_1 = 254°$, $\phi_2 = 143°$, $\phi_3 = 47°$. Assume that all the edges go downward. From (4), we obtain $\theta_1 = 150.2°$, $\theta_2 = 101.6°$, $\theta_3 = 117.0°$. Hence, the unit vectors along the edges are given from (1) by

$$\boldsymbol{n}_1 = (-0.1369, -0.4774, -0.8680),$$
$$\boldsymbol{n}_2 = (-0.7824, 0.5896, -0.2009),$$
$$\boldsymbol{n}_3 = (0.6076, 0.6516, -0.4542). \quad (16)$$

Similarly, we have $\phi_1 = 355°$, $\phi_2 = 113°$, $\phi_3 = 36°$ for Fig. 2(b), which is an "arrow." Assuming that the 3-edge goes upward and the others downward, we obtain $\theta_1 = 141.5°$, $\theta_2 = 110.5°$, $\theta_3 = 59.0°$, so that

$$\boldsymbol{n}_1 = (0.6208, -0.0543, -0.7821),$$
$$\boldsymbol{n}_2 = (-0.3660, 0.8622, -0.3504),$$
$$\boldsymbol{n}_3 = (0.6933, 0.5037, 0.5153). \quad (17)$$

The Eulerian angles are given by (7) and (8), and we obtain $\theta = 117.0°$, $\phi = 47.0°$, $\psi = -13.0°$ for Fig. 2(a) and $\theta = 59.0°$, $\phi = 36.0°$, $\psi = -24.1°$ for Fig. 2(b). Substitution of these values in (5) yields (16) and (17) again. From (13), the quaternions associated with Fig. 2(a) and (b) are given by

$$q = 0.4996 - 0.4265i + 0.7383j + 0.1526k,$$
$$q' = 0.8658 - 0.2466i + 0.4260j + 0.0900k, \quad (18)$$

respectively. Substitution of these in (14) again yields (16) and (17). From (15), the relative rotation is expressed by

$$q'' = 0.8660 + 0.2475i - 0.4256j - 0.0868k, \quad (19)$$

which represents, according to (11), a rotation having axis

$$\boldsymbol{n} = (0.4950, -0.8514, -0.1736) \quad (20)$$

and angle $\Omega = 60.0°$ screwwise. In other words, the corner of Fig. 2(b) is obtained by rotating the corner of Fig. 2(a) around this vector by this angle.

III. Characterization of a Rectangular Polyhedron

Now, we consider shape interpretation, exploiting the fact that the object is a rectangular polyhedron. In contrast to a general polyhedron, a rectangular trihedral polyhedron has the following distinctive characteristics.

Fact 1: There are only three different edge orientations if they are regarded as undirected lines. This holds both for the scene and for the image. (As before, we do not consider the degenerate case where two of them coincide by projection.)

Fact 2: Given any two corners, either their configurations are identical or one is obtained from the other by reversing one or more edges. This again holds both for the scene and for the image.

Fact 3: There are only three face orientations in the 3-D space if the distinction between inside and outside is disregarded.

Fact 4: The angle made by two edges defining a face at a corner is either $\pi/2$ or $3\pi/2$ in the scene.

Fact 1 enables us to fix reference edge orientations and describe the configuration of a corner in reference to these orientations on the image plane. Number the three possible orientations arbitrarily as 1, 2, and 3, and give them directions arbitrarily. These three directed orientations play the role of a "coordinate system," and we refer to each directed orientation as the 1-, 2-, 3-axis accordingly. At this stage, we do not consider the "visibility" and no distinction is made between "visible" lines and "hidden" lines, since we are not considering any face or "substance" inside the polyhedron. Consider a corner image and regard its edges as directed vectors emanating from the vertex pont. We call the edge parallel to the i-axis the i-edge of the corner. The "type" of a corner is defined by a triplet $\boldsymbol{c} = (c_1, c_2, c_3)$, where $c_i = 0$ if the i-edge has the same direction as the i-axis and $c_i = 1$ if it has the opposite direction.

Fact 2 implies a transformation group T generated by the identity and the reversals of each edge direction. There are eight elements in T. An element of T is expressed by a triplet $\boldsymbol{t} = (t_1, t_2, t_3)$, where $t_i = 0$ if it preserves the direction of the i-edge and $t_i = 1$ if it reverses it. The identity is given by $\boldsymbol{0} = (0, 0, 0)$. The group operation, i.e., the rule of composition, is given by $\boldsymbol{t} \oplus \boldsymbol{t}'$, where $\oplus$ denotes componentwise addition modulo 2 ("exclusive OR"). It is immediately seen that T forms a group, because each component follows the group operation of addition modulo 2 (i.e., of Z_2, the cyclic group of order 2). Then, T is isomorphic to their direct product[2]:

$$T \simeq Z_2 \times Z_2 \times Z_2. \quad (21)$$

Hence, T is Abelian, and each element is the inverse (or negative) of itself:

$$(\boldsymbol{t} \oplus \boldsymbol{t}') \oplus \boldsymbol{t}'' = \boldsymbol{t} \oplus (\boldsymbol{t}' \oplus \boldsymbol{t}'')$$
$$\boldsymbol{t} \oplus \boldsymbol{t}' = \boldsymbol{t}' \oplus \boldsymbol{t}, \boldsymbol{t} \oplus \boldsymbol{t} = \boldsymbol{0}. \quad (22)$$

Proposition 1: If transformation $\boldsymbol{t}$ transforms a corner of type $\boldsymbol{c}$ to a corner of type $\boldsymbol{c}'$, then

$$\boldsymbol{c}' = \boldsymbol{c} \oplus \boldsymbol{t} \quad \text{or} \quad \boldsymbol{t} = \boldsymbol{c} \oplus \boldsymbol{c}'. \quad (23)$$

The above statement is a characterization of the projected image. Next, consider the spatial configuration of a corner and define its "orientation" by a triplet $\boldsymbol{p} = (p_1, p_2, p_3)$, where $p_i = 0$ if its i-edge goes upward, i.e., $0 < \theta_i < \pi/2$, θ_i being its polar angle, and $p_i = 1$ if its i-edge goes downward, i.e., $\pi/2 < \theta_i < \pi$. In the previous section, we showed that for a given corner image, two spatial configurations are possible, one being the mirror image of the other. We repeat this fact as follows.

Proposition 2: Two orientations can be assigned to a

[2]The three groups are generated by different generators, but they are all isomorphic, and consequently T can be represented as a direct product of the same group Z_2 as in (21).

corner image, one being the complement of the other (i.e., 0 and 1 are interchanged).

The assignment cannot be done independently for each corner. As is easily seen, we have the following.

Proposition 3: If transformation t transforms a corner of orientation p to a corner of orientation p', then

$$p' = p \oplus t \quad \text{or} \quad t = p \oplus p'. \tag{24}$$

Combining this with Proposition 1, we obtain the following.

Corollary: If one corner is of type c and orientation p and another is of type c' and orientation p', then

$$p' = p \oplus c \oplus c'. \tag{25}$$

From this we can see that, once we assign an orientation to an arbitrary corner, the orientations of all the other corners are uniquely determined by (25). Combining this with Proposition 2, we obtain the following.

Proposition 4: Two spatial configurations are possible for the skeleton (note that we are considering edges only) corresponding to a given image, one being the mirror image of the other. The p's of one are obtained from those of the other by taking complements.

This is a well known familiar fact sometimes referred to as the "Necker cube phenomenon."

Up to now, we have considered only skeletons. Now, consider face adjacency. As we pointed out as Fact 3, there are three types of faces. Call the face defined by the i- and j-edges the ij-face. Fact 4 enables us to characterize the face adjacency at a corner by defining the "state" of the corner as a triplet $s = (s_1, s_2, s_3)$, where $s_{[ij]} = 0$ if the ij-face is at angle $\pi/2$ and $s_{[ij]} = 0$ if it is at angle $3\pi/2$. Here, we adopt a notational convention that $[12] = [21] = 3$, $[23] = [32] = 1$ and $[31] = [13] = 2$. (Note that at this stage we are not considering on which side the "substance" exists.) It is easily seen that there are two constraints concerning the state. First, no two faces at a corner can be at angle $3\pi/2$. Next, consider two adjacent corners linked by an i-edge. The ij-face has the same angle at the two corners if and only if the two j-edges have the same orientation, and it has different angles otherwise. Our observation is summarized as follows.

Proposition 5: No two components of any state s are 1, so that

$$s_1 s_2 + s_2 s_3 + s_3 s_1 = 0. \tag{26}$$

Proposition 6: If two corners of state s and s' are linked by an i-edge, and t is the transformation associated with the two corners, then

$$s_{[ij]} = s'_{[ij]} \oplus t_j, \; j \neq i. \tag{27}$$

Corollary: If two adjacent corners linked by an i-edge have types c and c' and states s and s', respectively, then

$$s_{[ij]} = s'_{[ij]} \oplus c_j \oplus c'_j, \; j \neq i. \tag{28}$$

IV. Visibility Conditions and Hidden Line Detection

Up to now, we have treated the faces of a polyhedron as "membranes." Now, we consider "visibility conditions," viewing the object as a polyhedron filled with an opaque "substance." If a corner image has three visible edges, they form either a "fork" or an "arrow." If a corner image has only two visible edges, they make either an obtuse angle or an acute angle and for each case there are two possibilities for the direction of the remaining missing edge. Hence, we must first get algebraic expressions telling whether a corner is a "fork" or an "arrow" and whether two edges make an obtuse angle or an acute angle. This can be done by using a special "coordinate system." So far, we have arbitrarily numbered the three reference orientations and assigned directions to them arbitrarily. Now, assign directions to them in such a way that these three directed lines form a "fork" on the image plane emanating from their intersection and call such a set of three directed axes a "fork coordinate system." Then, we obtain the following.

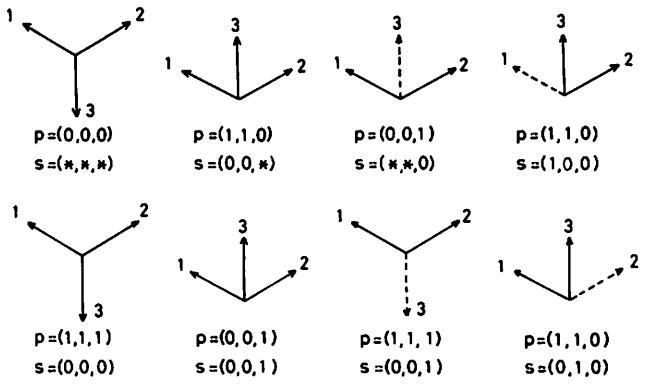

Fig. 5. The type, the orientation, and the state of a visible corner are those shown here and those obtained from them by permutation of 1, 2, and 3 components. Dashed lines indicate hidden edges, and * denotes an indeterminate bit.

Observation 1: The i- and j-edges of a corner of type c make an obtuse angle if $c_i \oplus c_j = 0$ and an acute angle if $c_i \oplus c_j = 1$.

Observation 2: A corner is a "fork" if its type is (0, 0, 0) or (1, 1, 1) and is an "arrow" otherwise.

Now, we consider the visibility conditions by checking possible corner configurations exhaustively. Here, the following obvious facts play a fundamental role.

Observation 3: If a corner is a "fork," its orientation is either (0, 0, 0) or (1, 1, 1). If it is an "arrow," the i-edge being the central edge and the j- and k-edges being the side edges, then either $p_i = 0$, $p_j = p_k = 1$ or $p_i = 1$, $p_j = p_k = 0$.

Observation 4: For a corner with two visible edges, the missing edge always goes downward.

Observation 4 is obtained by exhaustively considering all the cases in which an edge is hidden by the face defined by the other two edges. Then, we obtain the following.

Proposition 7 (visibility conditions): For a corner of a given type c, the orientation p and the state s must satisfy the constraints in Fig. 5 and those obtained from them by permutations of the 1, 2, 3 components.

An actual procedure goes as follows. First, we choose a "fork coordinate system" and define types for all visi-

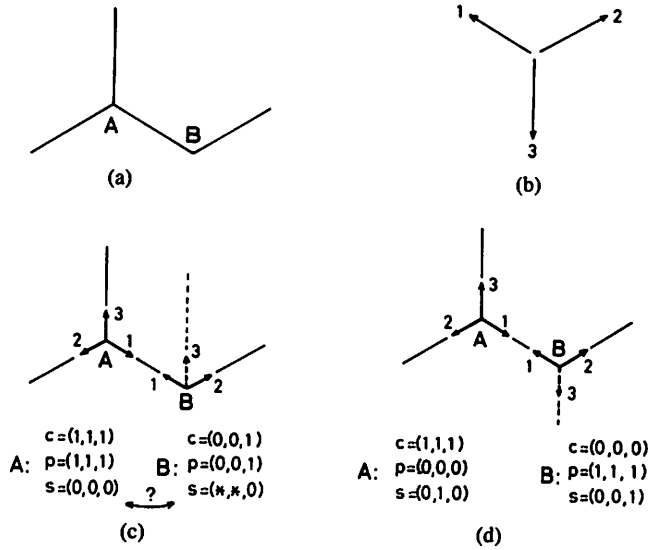

Fig. 6. (a) A given image. (b) A "fork coordinate system." (c) An assignment of orientations and states which yields inconsistency. (d) A uniquely determined consistent interpretation.

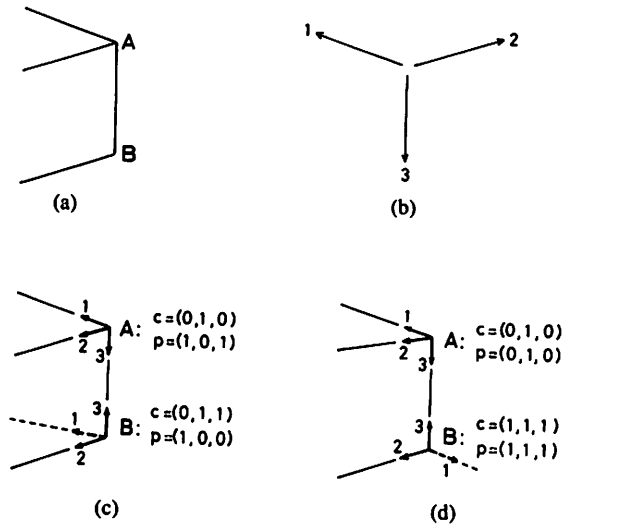

Fig. 7. (a) A given image. (b) A "fork coordinate system." (c) The states are not uniquely determined—part of an isolated box, part of a box sticking out of a vertical wall, or part of a box on a horizontal plane. (d) An assignment of orientations and states which yields inconsistency.

ble corners. For a corner having two visible edges, only those corresponding components of c are determined. Then, according to Observation 1, we can freely talk of obtuse angles, acute angles, "forks," and "arrows." Next, find a corner with a missing edge and check the two possibilities for the direction of the missing edge. Once one of them is assumed, the orientation of that corner is determined due to Observations 3 and 4, which means that the spatial orientations of all the visible corners are uniquely determined due to Proposition 4. Next, consider a neighboring corner and test the validity of the assumption by checking if the visibility conditions and Propositions 5 and 6 are violated or not. If not, go to another incomplete corner and repeat the same process. If we proceed in this way and obtain an assignment of types, orientations, and states without any inconsistency over a portion of the image under consideration (not necessarily the entire image), we say that an "interpretation" of that portion has been obtained. The interpretation may not be unique.

Example 2: Consider the image of Fig. 6(a). We use the "coordinate system" of Fig. 6(b), and hence corner A is of type (1, 1, 1) and corner B is of type (0, 0, *) with the 3-edge missing. If we assume its types to be (0, 0, 1) as shown in Fig. 6(c), its orientation and state must be (0, 0, 1) and (*, *, 0), respectively, from the visibility conditions. Then, according to Proposition 3, the orientation of A is (1, 1, 1), which implies from the visibility conditions that its state is (0, 0, 0). However, this contradicts Proposition 6, because $s_{[12]} = 0$ at B would imply $s_{[12]} = 1$ at A. Hence, the hidden line is drawn as in Fig. 6(d) and corner B must be of type (0, 0, 0). Its orientation and state are (1, 1, 1) and (0, 0, 1), respectively, from the visibility conditions. Then, corner A has orientation (0, 0, 0) from Proposition 3. Its state is (0, 1, 0) from Propositions 5 and 6.

Example 3: Consider the image of Fig. 7(a). We use the "coordinate system" of Fig. 7(b), and hence corner A is of type (0, 1, 0) and corner B is of type (*, 1, 1) with the 1-edge missing. If we assume its type to be (1, 1, 1) as shown in Fig. 7(d), its orientation and state must be (1, 1, 1) and (1, 0, 0), respectively, from the visibility conditions. Then, according to Proposition 3, the orientation of A is (0, 1, 0), which implies from the visibility conditions that its state is (0, 1, 0). However, this contradicts Proposition 6, because $s_{[23]} = 1$ at B would imply $s_{[23]} = 1$ at A. Hence, the hidden line is drawn as in Fig. 7(c) and corner B must of of type (0, 1, 1). Its orientation is (1, 0, 0) from the visibility conditions. Then, corner A has orientation (1, 0, 1) from Proposition 3. However, the states of A and B are not uniquely determined. Possible states of A and B are $s(A) = s(B) = (0, 0, 0)$ (part of an isolated box), $s(A) = s(B) = (0, 1, 0)$ (part of a box sticking out of a vertical wall) or $s(A) = (0, 0, 0)$ and $s(B) = (0, 0, 1)$ (part of a box on a horizontal plane).

V. Discussion and Concluding Remarks

We have studied the consequences of the fact that the object is a rectangular polyhedron. They are summarized as follows. First, we can make quantitative deductions of the spatial edge orientations and their relative motions in terms of polar angles, Eulerian angles, and quaternions. Second, we can obtain an interpretation of a "specified portion" of a given image. This is also possible by existing schemes, but they are less powerful because they are essentially based on the global consistency of a polyhedron as a whole. For example, if we apply Huffman's scheme [2] of labeling edges to the image of Fig. 6(a), we end up with the seven different interpretations shown in Fig. 8. However, the knowledge of rectangularity reduces them to only one, namely Fig. 8(e), as we have shown. A third point is that the characterization of a polyhedron can be constructed "hierarchically." We first regard the image as a two-dimensional "graph" and characterize its corners by assigning types to them. The spatial configuration as a "skeleton" is then described by as-

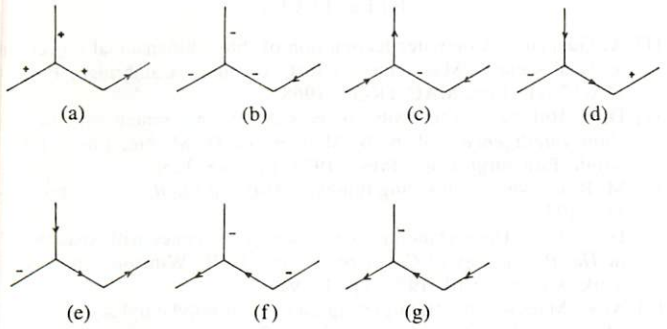

Fig. 8. Seven possible interpretations of Fig. 6(a) according to Huffman's labeling scheme. Here, + and − designate convex and concave edges, respectively, and an arrow designates a boundary edge to which a face is adjacent on the right side. If the object is known to be rectangular, only (e) is possible.

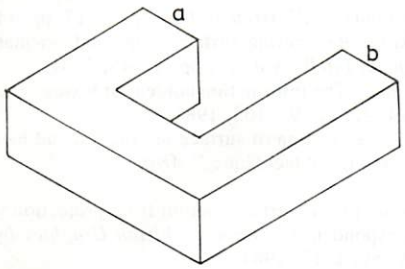

Fig. 9. A false interpretation can arise that both edge *a* and edge *b* are adjacent to an outside "wall."

signing orientations to the corners. Next, the face adjacency is specified by assigning states to the corners, regarding the faces as "membranes." Finally, we regard the object as a polyhedron filled with an opaque "substance," obtaining the visibility conditions. Each time we ascend the hierarchy, we obtain new constraints.

Now, let us consider the possible nonuniqueness of the interpretation. An important question is whether all the resulting interpretations are physically admissible or not. First, assume that the object is a "real" rectangular polyhedron. Then, interpretations which are physically impossible arise only from what can be termed "wall adjacency." Usually, the outermost edges encircling a domain are assumed to be the outer boundary of the polyhedron, while here no such assumptions are made. (Of course, we could assume so, if we wished.) Hence, we have two interpretations for such an edge—one as an outer boundary and the other as the adjacency to an outside "wall," as we have seen in Example 3. Note that no metric properties like the *xy*-coordinates of the corner vertices on the image plane are incorporated in our scheme of interpretation. Hence, we do not know the distance between faces. Thus, according to our scheme, the possible interpretations of Fig. 9 include one telling us that both edge *a* and edge *b* are adjacent to an outside wall. The "misinterpretations" of our scheme are only of this type and are unavoidable unless metric properties based on the *xy*-coordinates of vertices are incorporated. They cannot be removed by Huffman's scheme, either.

On the other hand, suppose we have a false drawing

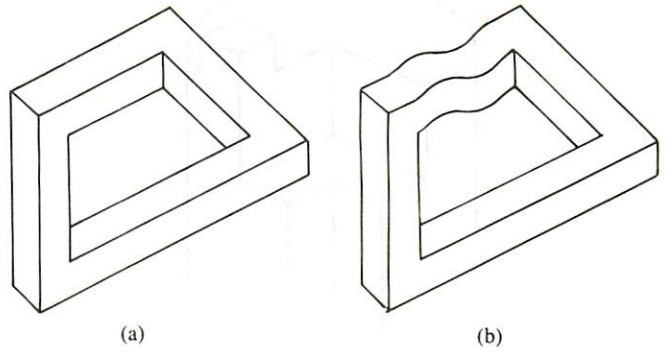

Fig. 10. The drawing of (a) is impossible as a rectangular polyhedron but possible if the real edges and surfaces are curved, all the corners being rectangular. For example, it may look like (b) from a slightly different viewpoint.

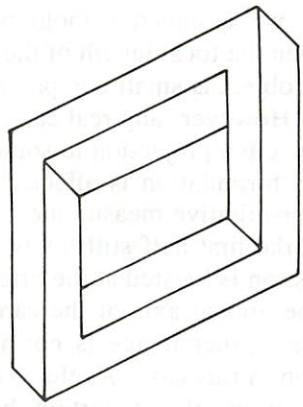

Fig. 11. An impossible rectangular polyhedron rejected by the present process of interpretation.

purporting to be a projection of a real rectangular polyhedron like those in Fig. 10. These impossible drawings can be classified into two categories. One consists of those which can be projections of real objects if curved edges and surfaces are allowed, yet satisfying Facts 1–4 of Section III near corners, like the one in Fig. 10(a), which might look like Fig. 10(b) from a slightly different viewpoint. Namely, an edge has the same orientation at corners but can be curved elsewhere, a surface has the same orientation at corners but can be curved elsewhere, and they make right angles at corners with only three distinct edge and surface orientations. The other category consists of those which cannot be interpreted that way like the one in Fig. 11. No interpretation results for the latter type and our scheme rejects those drawings as impossible. In other words, only drawings of the former type cannot be rejected by our scheme. (Huffman's scheme is unable to reject both Fig. 10(a) and Fig. 11.) Of course, the above statement is a rough description, and the real criterion is whether or not the drawing satisfies Facts 1–4. In order to remove "misinterpretations" completely, we must check the "gradient space," or preferably, measure the *xy*-coordinates of the vertices and algebraically check the "planarity," i.e., test if there exists a solution composed of the desired lines and planes in three dimensions. This procedure as well as a complete classification of "impos-

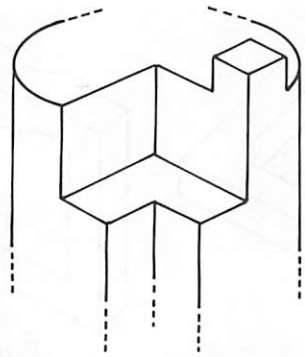

Fig. 12. A partly rectangular object. The present interpretation process can be applied to the part known or assumed to be rectangular.

sible objects'' is described by Sugihara [17], [18] for a general polyhedron, so that we do not discuss them here.

In this paper, we assumed orthographic projection, which applies when the focal length of the camera is large or the size of the object is small compared to its distance from the camera. However, any real camera images have the effect of perspective projection to some extent. Let us consider how our formulation is affected by perspective projection. The quantitative measurement of 3-D orientation described in the first half still holds if the vertex of the corner in question is located at the origin of the image plane (i.e., on the optical axis of the camera), since, as is easily seen, the corner image is not affected by perspective projection in this case. As the vertex lies further away from the origin, the distortion becomes larger. Hence, our formulation is applicable if the camera position is chosen so that the object of interest can lie near the center of the image plane. On the other hand, the interpretation process described in the latter half is not affected by perspective projection at all. This is because we do not use metric properties such as length and angle. All we need is ''identification'' of the edge orientation, i.e., to know which edges are parallel. This should not be so difficult if the projective distortion is not too extreme.

Finally, we should note that the present study is useful in practice even if the objects are not necessarily rectangular polyhedra. As we mentioned before, rectangular polyhedra are the most familiar objects in our daily life. Besides, many objects we encounter which are not rectangular polyhedra are ''partly'' rectangular like the one in Fig. 12. Since the study in this paper can be applied ''locally,'' our results can be applied to that portion

Acknowledgment

The author thanks Prof. K. Sugihara of the University of Tokyo for helpful communications. He also thanks Prof. A. Rosenfeld and Prof. L. S. Davis of the University of Maryland for discussions and comments.

References

[1] A. Guzman, ''Computer recognition of three-dimensional objects in a visual scene,'' Massachusetts Inst. Technol., Cambridge, Project MAC Tech. Rep. MAC-TR-59, 1968.
[2] D. A. Huffman, ''Impossible objects as nonsense sentences,'' in *Machine Intelligence*, vol. 6, B. Meltzer and D. Michie, Eds. Edinburgh: Edinburgh Univ. Press, 1971, pp. 295–323.
[3] M. B. Clowes, ''On seeing things,'' *Artificial Intell.*, vol. 2, pp. 79–116, 1971.
[4] D. Waltz, ''Understanding line drawings of scenes with shadows,'' in *The Psychology of Computer Vision*, P. H. Winston, Ed. New York: McGraw-Hill, 1975, pp. 19–91.
[5] A. K. Mackworth, ''Interpreting pictures of polyhedral scenes,'' *Artificial Intell.*, vol. 4, pp. 121–137, 1973.
[6] K. Sugihara, ''Picture language for skeletal polyhedra,'' *Comput. Graphics Image Processing*, vol. 8, pp. 382–405, 1978.
[7] T. Kanade, ''A theory of origami world,'' *Artificial Intell.*, vol. 13, pp. 279–311, 1980.
[8] B. K. P. Horn, ''Understanding image intensities,'' *Artificial Intell.*, vol. 8, pp. 201–231, 1977.
[9] K. Ikeuchi and B. K. P. Horn, ''Numerical shape from shading and occluding boundaries,'' *Artificial Intell.*, vol. 17, pp. 141–184, 1981.
[10] A. P. Witkin, ''Recovering surface shape and orientation from texture,'' *Artificial Intell.*, vol. 17, pp. 17–45, 1981.
[11] K. A. Stevens, ''The information content of texture gradients,'' *Biol. Cybern.*, vol. 42, pp. 95–105, 1981.
[12] K. Kanatani, ''Detection of surface orientation and motion from texture by a stereological technique,'' *Artificial Intell.*, vol. 23, pp. 213–237, 1984.
[13] —, ''Tracing planar surface motion from projection without knowing the correspondence,'' *Comput. Vision Graphics Image Processing*, vol. 29, pp. 1–12, 1984.
[14] —, ''Detecting the motion of a planar surface by line and surface integrals,'' *Comput. Vision Graphics Image Processing*, vol. 29, pp. 13–22, 1985.
[15] A. K. Mackworth, ''Model-driven interpretation in intelligent vision systems,'' *Perception*, vol. 5, pp. 349–370, 1976.
[16] T. Kanade, ''Recovery of the three-dimensional shape of an object from a single view,'' *Artificial Intell.*, vol. 17, pp. 409–460, 1981.
[17] K. Sugihara, ''Classification of impossible objects,'' *Perception*, vol. 11, pp. 65–74, 1982.
[18] —, ''A necessary and sufficient condition for a picture to represent a polyhedral scene,'' *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 578–586, Sept. 1984.

Ken-ichi Kanatani was born in Japan in 1947. He received the B.Eng., M.Eng., and Ph.D. degrees from the University of Tokyo, Tokyo, Japan, in 1972, 1974, and 1979, respectively. From 1969 to 1970, he studied physics at Case Western Reserve University, Cleveland, OH.

In 1979 he became Assistant Professor at the Department of Computer Science, Gunma University, Kiryu, Gunma, Japan, and in 1983 Associate Professor there. He teaches numerical analysis and computer vision. He is currently visiting the Center for Automation Research, University of Maryland, College Park. His interest is applications of advanced mathematics to various physics and engineering problems. His research areas have included mechanics of microstructured materials such as polycrystalline metals and granular materials and its applications to soil mechanics and powder technology, especially plastic flow properties; mathematical study of nervous networks; numerical analysis of physics and computer graphics related problems; stereology and its applications to measurement by computer image processing; image understanding; and 3-D reconstruction from optical flow.