Proceedings
IEEE International Workshop on Intelligent Robots and Systems
——Toward the Next Generation Robot and System——

Tokyo, Japan
Oct. 31- Nov. 2, 1988

IEEE Catalog Number 88TH0234-5

COSPONSORED BY

IEEE Industrial Electronics Society
IEEE System Man and Cybernetics Society
Robotic Society of Japan
New Technology Foundation

IN COOPERATION WITH

IEEE Council on Robotics and Automation
International Association for Mathematics and Computers in Simulation (IMACS)
Japan Society of Mechanical Engineers (JSME)
Society of Instrument and Control Engineers (SICE)
Institute of Electric Engineers of Japan (JIEE)
Institute of Electronics, Information and Communication Engineers (IEICE)
Japan Society of Artificial Intelligence (JSAI)
Japan Industrial Robot Association (JIRA)
International Robotics and Factory Automation Center (IROFA)

Cooperative Organizations

Namco Corporation
Pacific Network 21 (PN21)
Yoshida Foundation for Science and Technology
SECOM Science and Technology Foundation
International Information Science Foundation

3-D RECONSTRUCTION OF AN OBJECT FROM A PROJECTED IMAGE BY USING PRIOR KNOWLEDGE

Mutsuhiro Terauchi, Mitsuo Nagamachi, Koji Ito and Toshio Tsuji

Human Factors and Bioengineering Laboratory
Faculty of Engineering
Hiroshima University
Higashi Hiroshima, 724 Japan

Abstract: This paper presents an approach to recover the 3D shape of a rigid body from a projected image by using prior knowledge, which is one of the most important problems to develop the visual sensor system for the intelligent robots. Generally, visual reconstruction from a single view is essentially an indeterministic problem. Therefore the reconstruction is impossible unless some sort of assumption or prior knowledge is given about the image and the object. Our system is based on production rules which represent the knowledge about the projection transform and the constraints on the object in a scene. Furthermore it can infers the structure of invisible part of an object. It is implemented as expert vision system.

1. Introduction

There are several levels of processing in computer vision. The early vision is processed locally, but in high level vision we must consider the structural characteristics of the overall image. In general, we can not obtain the coordinates of an object in 3D space from a single view without the knowledge about the structure or location of the object. It is one of the key points in computer vision to find the assumptions or the constraints which include many objects encountered in various practical applications.

Guzman [1] and Huffman [2] who were frontiers in this area investigated the structure of the object and the mixtured scene. They assumed that the object was polyhedra and used heuristic rules about realizability of combination of edge characters. Waltz took up the images of scenes with shadows [6]. In the practical scene there are many factors which have an effect on the image of the object, e.g. shadows, reflectance and so on. He treated the illumination problem, and utilized the relationship between illumination and shadows. Kanatani used the assumptions that objects were rectangular trihedral polyhedra and were projected orthographically [3]. This assumption is called "block world model" and is often used by other researchers, since a rectangular polyhedron is the most frequently encountered object, e.g. boxes, machine parts, buildings and so on. He proposed the algorithm to compute the spatial orientation of corner edges easily and quickly.

However most of their interpretations were analyzed qualitatively. Moreover they were not represented as solid body but only as connected surfaces. For practical applications, it is necessary to reconstruct quantitatively and to ensure the full shape of the object.

On the other hand, the most successful work in artificial intelligence is expert system. It utilizes the knowledge as the rules. The expert vision system is considered to be valid only when their knowledge is represented explicitly [5].

In this paper, we present the algorithm that can determine the 3D shape of the object from a single line drawing image, and implement as the expert vision system. We assume that the image is originally line drawing. This assumption prevents us from effort to process raw gray level image. Furthermore we assume that the object is a rectangular trihedral polyhedron, and it is projected perspectively on to the image, where the object is at "the general position". Then it is shown that the object can be reconstruct as the full solid body by the inference of its backside shape which is not seen.

2. Geometry of Perspective projection

The objects in three dimensional space is distorted by the perspective projection onto two dimensional space. Its transformation is shown in Fig. 1. There are two types of geometric distortion : (1) As a surface recedes from viewer, the surface edges or areas appear smaller. It is called "railroad track effect". (2) When a surface is inclined off from the frontal plane, the surface edges or shapes appear foreshortened or compressed in the direction of the inclination. For example if the surface is inclined, a circle on the surface will be projected as an ellipse. This distortion makes image analysis and computation complex, but at the same time it tells us relative distances in the image.

2.1 Vanishing Point

If the lines are parallel in 3D scene, the projected lines on 2D image have only one intersection at the same position as follows (see Fig. 2). Here we represent all the coordinates by the vector that originates from the focal point [4]. If $\vec{f}$ is defined as the intersection of the image plane and the optical axis, an arbitrary

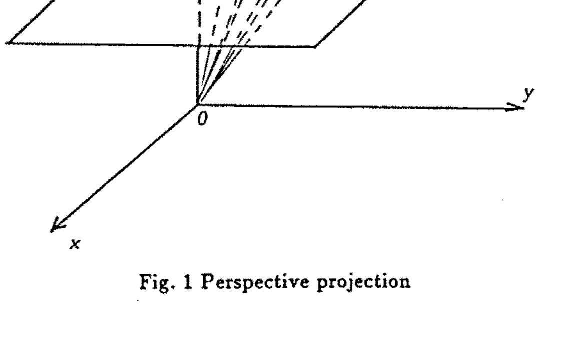

Fig. 1 Perspective projection

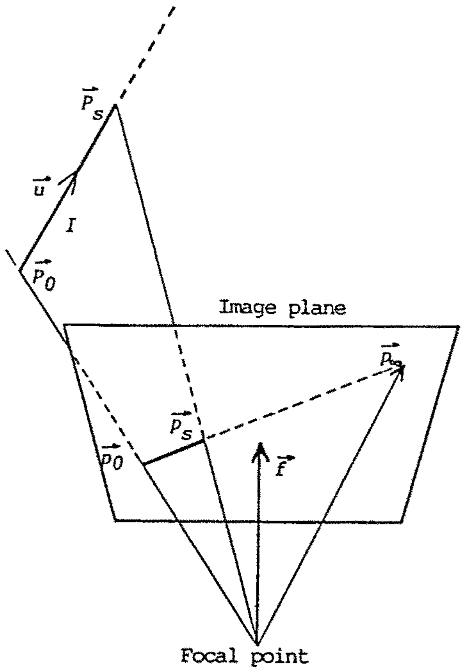

Fig. 2 $\vec{p}_\infty$ depends on only $\vec{u}$.

point $\vec{p}$ is represented as follows,

$$\vec{p}\cdot\vec{f} = \|\vec{f}\|^2 \qquad (1)$$

where "$\cdot$" stands for inner product and $\|*\|$ for Euclidean norm. We assume two points $\vec{P}_0$ and $\vec{P}_S$ on the straight line I in 3D scene, and their projection onto image plane are the points $\vec{p}_0$ and $\vec{p}_S$. Then the relation between the points in 3D space and on an image is described as follows,

$$\vec{P}_0 = a_0\,\vec{p}_0 \quad , \quad a_0 : \text{constant} , \qquad (2)$$

$$\vec{P}_S = a_S\,\vec{p}_S \quad , \quad a_S : \text{constant} , \qquad (3)$$

We define unit vector of the straight line I as $\vec{u}$. Then $\vec{P}_S$ is represented as follows,

$$\vec{P}_S = a_0\,\vec{p}_0 + s\vec{u} \quad , \quad s : \text{constant} . \qquad (4)$$

From eq.(1), (3) and (4) a_S and $\vec{P}_S$ are obtained as

$$a_S = \frac{a_0\|\vec{f}\|^2 + s\vec{u}\cdot\vec{f}}{\|\vec{f}\|^2} , \qquad (5)$$

$$\vec{p}_S = \frac{\|\vec{f}\|^2\,(a_0\vec{p}_0 + s\vec{u}\,)}{a_0\|\vec{f}\|^2 + s\vec{u}\cdot\vec{f}} . \qquad (6)$$

If we assume the straight line I is not parallel to the image plane, the point that is located at infinite distance from point $\vec{P}_0$ is projected to the point $\vec{p}_\infty$, which is computed from eq. (6) as follows

$$\vec{p}_\infty = \lim_{s\to\infty} \vec{p}_S = \frac{\|\vec{f}\|^2\,\vec{u}}{\vec{f}\cdot\vec{u}} . \qquad (7)$$

Therefore p_∞ is determined only by the orientation of $\vec{u}$ of the straight line I. We call the point $\vec{p}_\infty$ as "vanishing point". All the straight line except for parallel ones to an image plane ($\vec{f}\cdot\vec{u} \neq 0$) have vanishing point.

If the coordinates of the point $\vec{p}_0$ and the vector $\vec{u}$ are known from eqs. (3) and (5), we can obtain the 3D coordinates of $\vec{P}_S$ from the point $\vec{p}_S$ on the image. Even when the coordinate of the point $\vec{P}_0$ cannot be found, the relative position $\vec{P}_0$ and $\vec{P}_S$ can be decided. After all, if we know the vector of the straight line in three dimensional space, we can compute the coordinates of the points on the straight line.

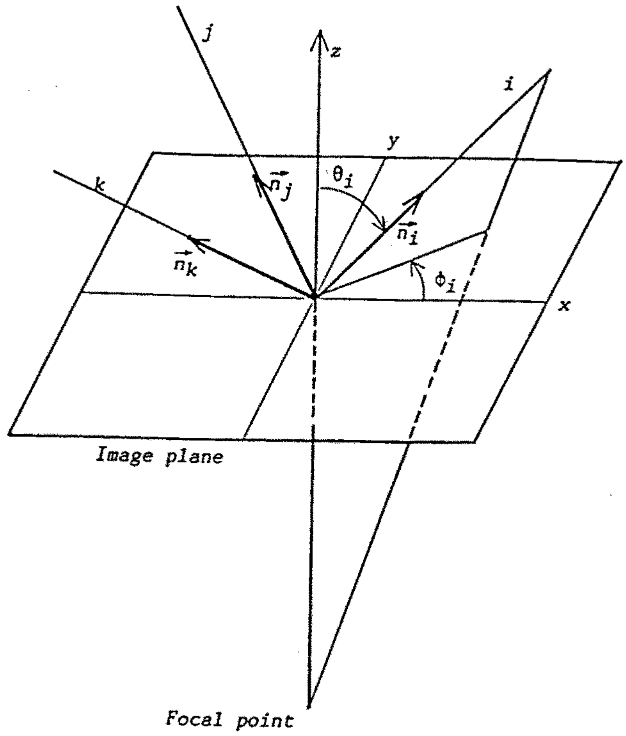

Fig. 3 Definition of coordinates.

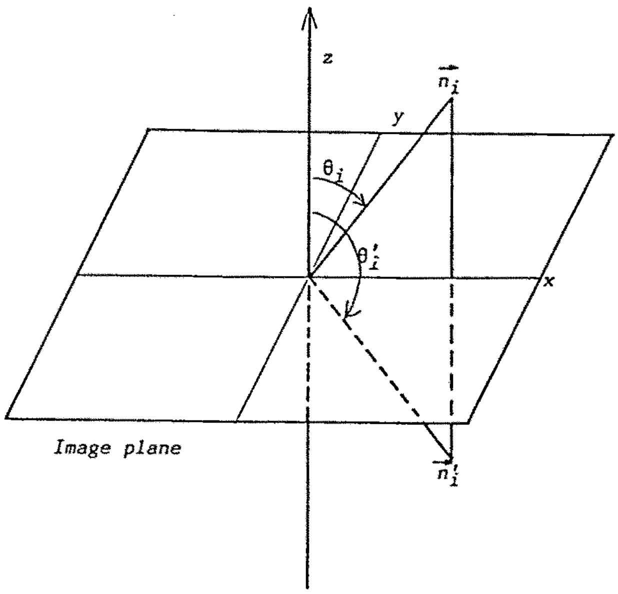

Fig. 4 Two interpretation of a visible edge.

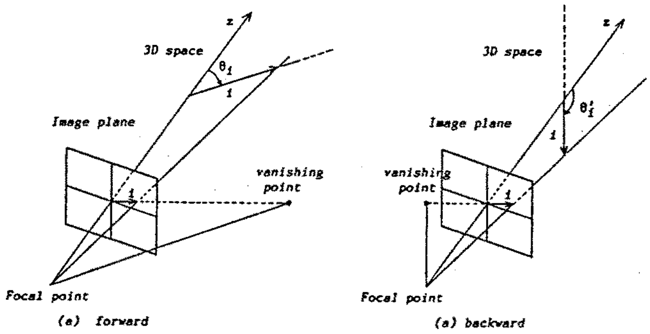

Fig. 5 Relationship between edge orientation and direction toward vanishing point.

2.2 Orientation of Edges

Here we are going to determine of the vector of the edge on a object in a scene from the projected image. We consider the coordinate system shown in Fig.3, where its origin is at the focal point, z axis on the optical axis and $x-y$ plane on the image plane. It is assumed that one of the corners is located at the origin. There are three edges (i=1,2,3) at one vertex. $\vec{n_i}$ denotes the unit vector in the direction of each edge. Then $\vec{n_i}$ is represented by the orthogonal coordinates as follows,

$$\vec{n_i} = (x,y,z)$$

$$x = \sin\theta_i \cos\phi_i$$

$$y = \sin\theta_i \sin\phi_i$$

$$z = \cos\theta_i \qquad (8)$$

The angle between two of three edges is right angle from condition, we obtain

$$\sin\theta_i \sin\theta_j (\cos\phi_i \cos\phi_j + \sin\phi_i \sin\phi_j) + \cos\theta_i \cos\theta_j = 0$$

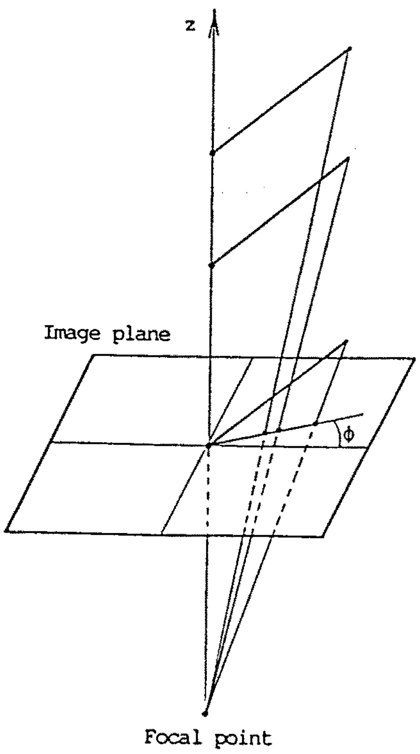

Fig. 6 ϕ does not change when vertex is on z axis.

Furthermore for all combination of two from three edges we obtain θ_i from three equations as follows,

$$\theta_i = \tan^{-1}\sqrt{\frac{-\cos(\phi_j-\phi_k)}{\cos(\phi_i-\phi_j)\cos(\phi_k-\phi_i)}} ,$$

$$\theta'_i = \tan^{-1} - \sqrt{\frac{-\cos(\phi_j-\phi_k)}{\cos(\phi_i-\phi_j)\cos(\phi_k-\phi_i)}} , \qquad (9)$$

$$(i \neq j \neq k) ,$$

where $0 < \theta_i < \pi/2$, $\pi/2 < \theta'_i < \pi$. And the relation between θ_i and θ'_i is represented as

$$\theta'_i = \pi - \theta_i$$

Here we consider unit vector n_i and n'_i for its angle θ_i and θ'_i. The vector n_i is mirror reflection of n'_i. (See Fig.4.) Whether the angle θ_i is greater than $\pi/2$ or not depends on the fact whether the projected edge i orients toward the vanishing point in the image plane or not. This means that the vector of the straight line i is forward or backward. They are shown in Fig.5a and 5b.

3. Computation of corner

As we may see, if one vertex of a object is located at origin, we can obtain directly the angle ϕ from the projected image. The case that the vertex is on z axis, the angle ϕ of the edge of the vertex would not change. (See Fig.6.) However it is not always guaranteed that the vertex lies on z axis. We can settle this problem by considering the virtual vertex shown in Fig.7. The vanishing point of each of the edge i', j', k' of the virtual vertex is same as the original vanishing point. It follows that the corresponding edges are parallel each other in three dimensional space. As we can consider that the virtual vertex is moved parallel to the original vertex, the angle ϕ of the virtual vertex is same as one of the original vertex. In order to do that, we have to obtain the vanishing points at first. Let us consider to the vanishing point.

3.1 Computation of Vanishing point

When we try to obtain the vanishing point, we will extend the edges on the image which is parallel in 3D space, and search their intersection. But we do not know which edges are parallel in 3D space. Therefore we extend all the edges, obtain all the intersections, and select three major groups of the intersections.

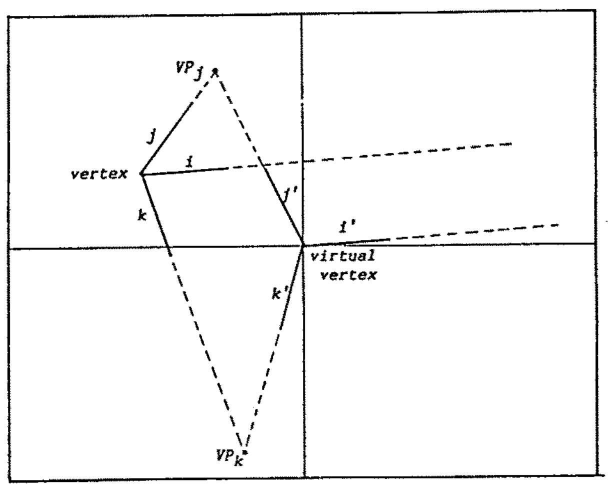

Fig. 7 Virtual vertex : Original vertex (left) is moved to center to remove the projection distortion. The edge of original vertex is re-oriented toward vanishing point (VP_x). Unless VP_x is computed, new edge's direction is taken as same as original one (in this case, i and i').

In this paper, we consider that the object in the scene is the rectangular trihedral polyhedron. Therefore we can classify all edges to three groups. So it is necessary to examine a structure of an object.

In the other hand, a vanishing point must exist except when the edge is parallel to the image plane in 3D space, and the projected edge is located in the center of an image plane. Since it is assumed that the object is in "the general position", it is rare case that the edges that belong to the same group are mutually parallel on the image. However we have to consider the rare case owing to error of digitization, and of distortion of the optical system etc.

When one group of the edges are mutually parallel on the image, we may think the edge of virtual vertex has same direction as corresponding edge which belongs to the same group even if it exist anywhere. As the problem is caused by the fact that the edges are almost parallel to an image plane in 3D space, the this approximation error would be negligible.

3.2 Removement of Perspective Projection Error

In order to remove the distortion error due to the projection, the virtual vertices was assumed. They are moved parallel from the original position in 3D space. On 2D image, they are made in the way that the original vertex is moved to the center of the image and that the edges of the virtual vertex are re-oriented to the vanishing points as shown in Fig. 7. This procedure is the key point in order to make the mathematical problem of projection easier.

3.3 Qualitative method for Corner characteristics

We use Kanatani's algorithms to determine one interpretation from the structure of the object [3]. The rectangular trihedral polyhedron has the distinctive characteristics as follows,

Fact 1: There is only three types of edge if they are regarded as undirected lines.

Fact 2: Given any two corners, their configurations are identical or mirror operationed in the scene.

Fact 3: There are only three types of surface orientation if the distinction between inside and outside is disregarded.

Fact 4: The angle made by two from three edges of a vertex defining a face at a corner is $\pi/2$ or $3\pi/2$.

And he represented three characteristics of the corner by using three triplets, type : $c = (c_1,c_2,c_3)$, state : $t = (t_1,t_2,t_3)$ and orientation $p = (p_1,p_2,p_3)$.

Furthermore the constraint equations between vertices are given as follows,

$$c' = c \oplus t \quad or \quad t = c \oplus c' \quad , \tag{10}$$

$$p' = p \oplus t \quad or \quad t = p \oplus p' \quad , \tag{11}$$

$$p' = p \oplus c \oplus c' \quad , \tag{12}$$

$$s_1s_2 + s_2s_3 + s_3s_1 = 0 \quad , \tag{13}$$

$$s_{[ij]} = s'_{[ij]} \oplus t_j \quad , \quad j \neq i \quad , \tag{14}$$

$$s_{[ij]} = s'_{[ij]} \oplus c_j \oplus c'_j \quad , \quad j \neq i \quad , \tag{15}$$

where $\oplus$ denotes componentwise addition modulo 2. The above implies an additive transformation group whose cycle is order 2, i.e., something like logical function. For the details of this algorithm, see reference [3].

In the other hand, the image corresponding to actual scene are constrained by physical structure. The each vertex of rectangular trihedral polyhedron is represented by some parts of eight divided space as shown in Fig. 8 and 9. From Fig. 9, it is known that the "type" of a visible corner is restricted to only eight ones. These realizable patterns of corner are shown in Fig. 10. At first, the edge orientations are divided into three groups. ,which depends on the angle ϕ of the edge of a corner. After each angle of three edge was decided, it can be divided into eight types of a corner as shown in Fig. 10. Then the consistent solution for all the connected vertices is obtained by using above equations. As a result, we have three triplets for each corner which denote the imply qualitative characters of the corner.

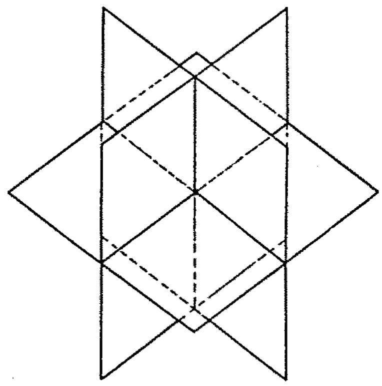

Fig. 8 Eight - divided space.

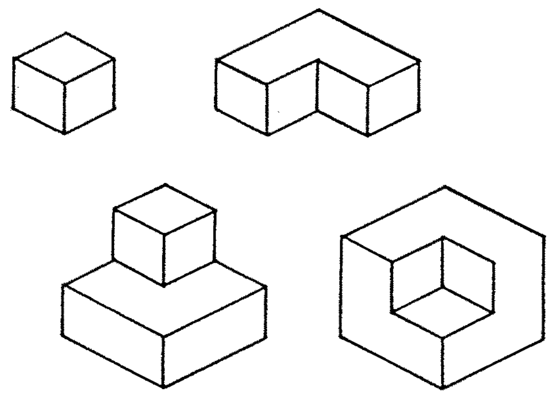

Fig. 9 Four types of rectangular trihedral polyhedra shown in eight- divided space (Fig.8). There are only these types about the types of corner.

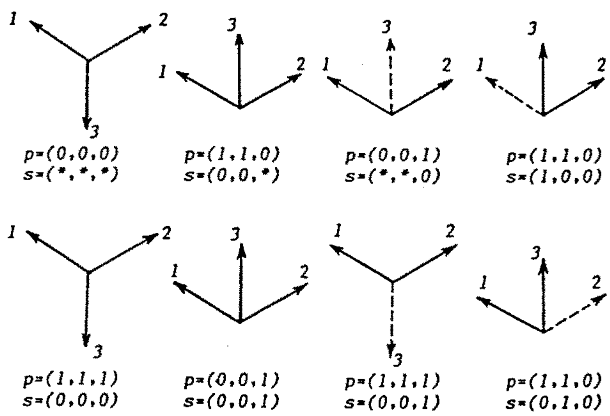

Fig. 10 Realizable patterns of corner [3].

4. Inference of invisible corner

In general, it is impossible to reconstruct hidden edges from only the image data. But we can imagine a solid object even if the perfect image is not given. Furthermore to some extent we can infer the structure which has not been encountered before at all. But it is not ensured at all that the inferred shape agrees with the real object in the scene.

Now it is shown that even if only one image is given, it is possible to reconstruct the 3D full shape of the object by using a prior knowledge.

Noted that all the inference process is all done in 3D space. At first, the process begins by searching an invisible edge at each

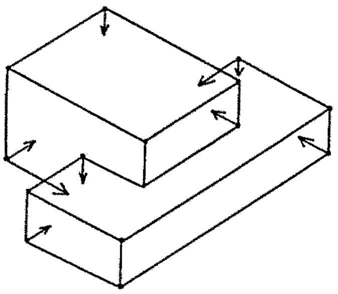

Fig. 11 Invisible edges at visible vertices.

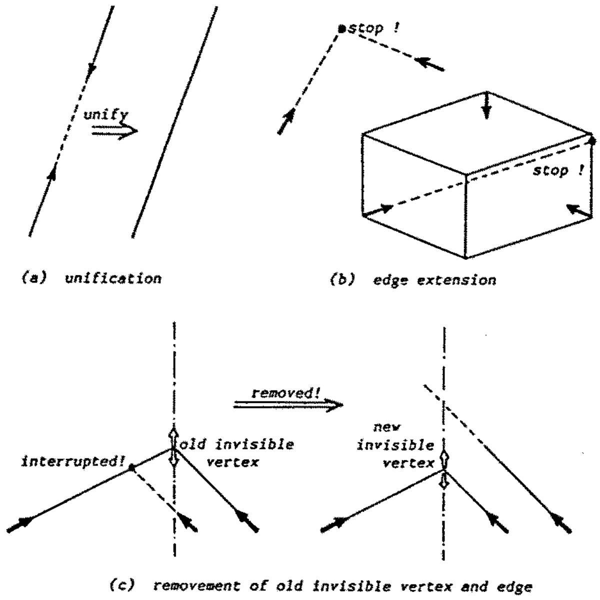

Fig. 12 Inference strategy of backside shape.

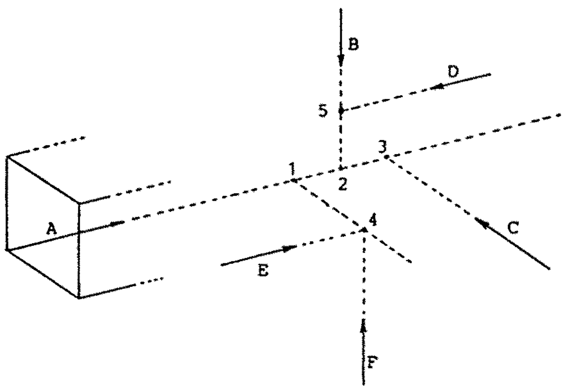

Fig. 13 An example of inference of invible corner.

corner. It can be found easily as shown in Fig. 11. The corner which includes an invisible edge has two visible edge or is "T" type where a boundary line occludes a edge. The latter is due to the condition that the object is located in the "general position" where vertices are not seen as "T" type corner. Then the invisible edge is extended as far as it crosses the other extended edge or the visual line through the contour of the object (see Fig. 12.b). The latter case prevents the edge from being extended infinitely. The two edges facing one another is connected into one line as shown in Fig. 12.a. If the extended edges cross each other, the intersection is considered as an invisible vertex, which must contain one more invisible edge. The invisible edge is extended toward the third direction from the invisible vertex. It is the same case as in the visible vertex. If the invisible edge is intercepted afterward, the new invisible vertex is made at the intersection, and the old vertex and a part of the edge between the new and the old invisible vertices are removed (Fig.12.c). For example, in Fig.13 the invisible vertices which can survive till the end of the iteration are vertex 1, 3 and 5. Above procedures are iterated until all the vertices get three edges. It should be also noted that since the object is described by the object centered coordinates system whose origin is located at one of the vertices of the object, the above procedure are performed easily by using only the two dimensional coordinate.

All the above procedures are implemented in a part of expert vision system as shown in Fig. 14.

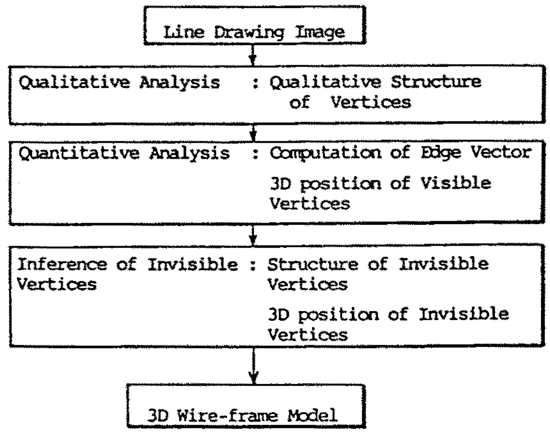

Fig. 14 Process flow of the system.

5. Discussion and Concluding Remarks

When the perspective projected image was given, it was shown that the original block object was reconstructed by using some sort of knowledge.

First the vanishing point was computed, and the edges of the vertex were determined. In order to remove the distortion error of projection, the vertex on the image projected perspectively was transformed to orthographic one by assuming the virtual vertex. Next the qualitative characteristics were analyzed to obtain the structural knowledge. We used Kanatani's algorithm for the analysis of rectangular trihedral polyhedra [3]. It is based on Lie group and Lie algebra.

Furthermore, the back side of the object, i.e. the invisible part, must be inferred for solid reconstruction. Though it has no guarantee for its validity, we can obtain full 3D shape of the object as most imaginable structure. This method reflects the structure of the vertices which have the invisible edge, i.e. utilizes the invisible edges and their intersections. But the object as shown in Fig. 15.a be inferred. Though it is easily seen that it has a square hole in its center axis (Fig. 15.b). Because the method uses only edges and point, not surfaces. When we look at something, we may always imagine its surfaces. Since the surfaces have much information for image understanding, it is necessary to find the method which describe the structure of object surfaces.

Though our reconstruction is performed metrically, we need some information about the length. It is impossible to obtain the distance from only one image without any information about length. But if there are some knowledge for objects' metrics or some function to know the focus, the information about distance can be obtained.

Finally, in the present paper, the objects was restricted to rectangular trihedral polyhedra, which are encountered so often in our lives. But it is artificial one. We can not apply the present system for natural image like a landscape of mountains.

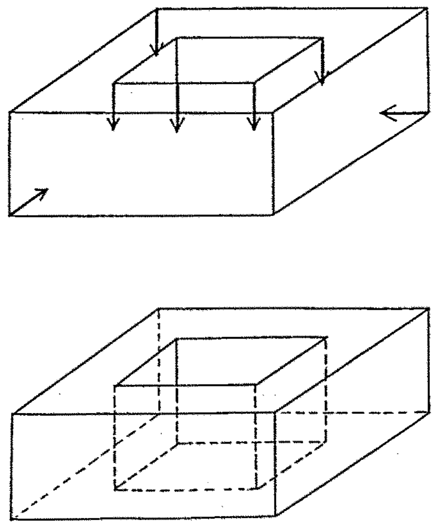

Fig. 15 Impossible shape for inference.

Furthermore we restricted the number of objects to only one. In order to allow existence of many objects, we must to compute vanishing points for each object, because the objects may be located at random.

Acknowledgment

The authors wish to thank Mr. H.Tsukiashi for discussion and implementation of this expert vision system.

References

[1] A.Guzmann, "Computer recognition of three-dimensional objects in a visual scene," Massachusetts Inst. Technol., Cambridge, Project MAC Tech. Rep. MAC-TR-59 (1968)

[2] D.Huffman, "Impossible Objects as Nonsense Sentences," in B.Meltzer and Michie (eds.), Machine Intelligence 6, Edinburgh University Press, Edinburgh, Scotland (1971)

[3] K.Kanatani, "The Constraints on Images of Rectangular Polyhedra," IEEE Trans. Pattern Anal. Machine Intell, Vol.PAMI-8, No.4, pp.456-463 (1986)

[4] H.Nakatani and T.Kitahashi, "Reconstruction of Object in Outdoor Scene Based on Vanishing Point," Trans. IECE, Vol.J68-D, No.8, pp.1481-1488 (1985)

[5] A.Rosenfeld et al., "Expert Vision Systems : Some Issues," Comput. Vision Graphics Image Processing, Vol.34, pp.99-117 (1986)

[6] D.Waltz, "Understanding Line Drawings of Scenes with Shadows," in P.H.Winston (eds.), The Psychology of Computer Vision, McGraw-Hill Inc. New York (1975)