RECONSTRUCTION OF THREE-DIMENSIONAL SHAPE USING PICTORIAL DEPTH CUES

February 1986

HIROMASA NAKATANI

RECONSTRUCTION OF THREE-DIMENSIONAL SHAPE USING PICTORIAL DEPTH CUES

A Dissertation Presented

By

HIROMASA NAKATANI

Submitted to Osaka University in partial fulfillment
of the requirements for the degree of

DOCTOR OF ENGINEERING

February 1986

ACKNOWLEDGEMENT

I would like to thank my thesis supervisor, Professor Yoshikazu Tezuka whose untiring efforts have been instrumental in the production of this dissertation, and Professors Yoshiroo Nakanishi and Sadao Kurazono who have had the kindness to provide a sound evaluation of this thesis.

Special thanks go to Professor Tadahiro Kitahashi for his generous support and intellectual guidance throughout the years since I was an undergraduate student. This thesis could not have been written without him.

I am deeply grateful to: the late Professor Kokichi Tanaka and Professor Shinichi Tamura for their advice and encouragement during my graduate research at Osaka University; Professors Edward Riseman, Allen Hanson, Richard Weiss and members of the VISIONS group for their generous support of my work at the University of Massachusetts; Professors Kinji Matsumoto, Keiichi Abe and members of Shizuoka University for their valuable comments and friendship; and Professor Robert McGuigan of Westfield State College for helping me with my English in this thesis.

This work was supported in part by The Ministry of Education, Science and Culture under Grand-in-Aid for Encouragement of Young Scientist 479018, 579017, 57780034, 59780030 and 60780034; by The Sakkokai Foundation; and by The Takayanagi Foundation.

ABSTRACT

The primary emphasis of this thesis is on recognition of three-dimensional structure of objects in a scene from a monocular picture by exploiting knowledge about perspective projection. We will introduce vanishing points for spatial information determination in the case of objects which have parallel lines or edges on planar surfaces.

We make clear the role that the location of the vanishing point or vanishing line plays in the three-dimensional measurements, and present two methods for the extraction of vanishing points. We also derive the constraints on the three-dimensional shape of vertices imposed by vanishing points in perspective drawings. Then, as a practical application of the shape reconstruction, we present a process that reconstructs the frontal view of a building from its perspective image.

The other area of research, the display of the three-dimensional information, is also attacked by constructing a system that produces a three-dimensional image of the heart from two-dimensional representations and allows us to view the heart from any orientation at any location.

TABLE OF CONTENTS

ACKNOWLEDGEMENT i

ABSTRACT ii

Chapter 1. INTRODUCTION 1

Chapter 2. DETERMINATION OF THE VANISHING POINT IN OUTDOOR SCENES 4

2.1 Introduction 4
2.2 Spatial Information Determination Based on Perspective Transformation 4
2.2.1 Vanishing Points 5
2.2.2 Calculating the Distance of a Point Based on the Vanishing Point 7
2.2.3 Calculating Surface Orientation Based on the Vanishing Line 8
2.3 Determining a Vanishing Point Using the Hough Transformation 9
2.3.1 The Hough Transformation 9
2.3.2 An Algorithm 10
2.3.3 Improving the Method 13
2.3.4 Deriving the Camera Motion Parameters 13
2.4 Iterative Method for Determining a Vanishing Point . . 16
2.4.1 The Data 16
2.4.2 A Candidate Vanishing Point 16
2.4.3 Iterative Update 18
2.4.4 Test Results 19
2.5 Conclusion 21

Chapter 3. INFERRING THREE–DIMENSIONAL SHAPE FROM LINE DRAWINGS USING VANISHING POINTS . . . 23

3.1 Introduction 23
3.2 Labeling Trihedral Vertices 24
3.2.1 Trihedral World 24
3.2.2 Conventional Approaches 24

	3.3	Constraints by Vanishing Points	26
	3.3.1	Edges and Vanishing Points	27
	3.3.2	Constraints on the Three-Dimensional Shape of a Vertex by Vanishing Points	28
	3.3.3	Constraints on Line Labeling	30
	3.4	Three Perspectives of Trihedral Vertices	32
	3.4.1	Junction Type in One-Point Perspective	33
	3.4.2	Junction Type in Two-Point Perspective	35
	3.5	Comparison with Conventional Approaches	35
	3.6	Conclusion	38
Chapter 4.		SHAPE RECONSTRUCTION OF OBJECTS IN OUTDOOR SCENES	39
	4.1	Introduction	39
	4.2	Reconstruction of the Frontal View of an Object from Its Perspective Image	39
	4.2.1	Assumptions	39
	4.2.2	Determination of the Object Plane Using Constraints at Vertices	40
	4.2.2.1	Constraints on Shape	40
	4.2.2.2	Constraints on Position	41
	4.3	The Reconstruction Process	43
	4.3.1	Vanishing Point Extraction	43
	4.3.2	Determination of the Object Plane	43
	4.3.2.1	Thinning the Edge Image	43
	4.3.2.2	Search for Candidate Corners	44
	4.3.2.3	Search for the Upper and Lower Sides	44
	4.3.2.4	Search for the Object Plane	45
	4.3.3	Reconstruction of the Frontal View	47
	4.3.3.1	Calculation of the Location	47
	4.3.3.2	Interpolation for the Display	47
	4.3.4	Result	47
	4.4	Determination of the Three-Dimensional Shape of Corners	50
	4.4.1	Constraints on the Vertex Shape	50
	4.4.2	Determination of Detailed Structure	51
	4.5	Conclusion	53

Chapter 5. THREE-DIMENSIONAL SHAPE RECONSTRUCTION FROM ULTRASONO-TOMOGRAMS 55

5.1 Introduction 55
5.2 Binocular Stereoscopic Display System for Echocardiography 56
5.2.1 Binocular Stereoscopy 56
5.2.2 Processing of Tomograms 58
5.2.2.1 Input of Tomograms 58
5.2.2.2 Smoothing 58
5.2.2.3 Bilevel Representation 58
5.2.2.4 Boundary Detection and Line Drawing 60
5.2.3 S-Mode Display 62
5.2.4 Results 63
5.3 Tomogram Reconstruction in a Desired Plane 65
5.3.1 C-Mode Display 66
5.3.2 Results 70
5.4 Conclusion 73

Chapter 6. CONCLUSIONS 74

BIBLIOGRAPHY 76

CHAPTER 1

INTRODUCTION

When a scene analysis is done by computer, for example if a robot needs to manipulate an object, the analysis of only the two-dimensional shape, size or location of the object in the image is not sufficient. Information such as shape and location in three-dimensional space is necessary. However, much of the previous work concerning scene analysis concentrated on two-dimensional interpretation of objects on the basis of region segmentations [1,2].

It is not easy for a computer to derive three dimensions from the sensory data directly. Because the picture-taking process is not a one-to-one transformation, the distance of an object cannot be determined uniquely from its image. The shape and disposition of objects are also distorted due to the perspective effects, especially in those pictures which have distant objects. The real size and shape might be quite different from those in the image, though such image features give us important cues when we identify objects and recognize three-dimensional structure of the physical world.

When a person sees a picture, however, he perceives objects in it as three-dimensional objects in three-dimensional space, inferring spatial relationships of the objects from depth cues depicted in the picture.

Investigators interested in monocular scene analysis have been concerned with the use of pictorial depth cues and constraints based on real world knowledge. When some knowledge such as the nature of the scene, lighting and the camera geometry is given in advance, various pictorial cues are used to infer

surface orientation, for example the intensity of a pixel [3], texture gradients [4], etc. In general vision systems, for example see [5], the depth information thus obtained, as well as other visual features, is intended to provide a higher-level process with useful visual knowledge that can facilitate interpretation of the scene.

The primary emphasis of this thesis is on recognition of three-dimensional structure of objects in a scene from a monocular picture by exploiting knowledge about perspective projection. To determine the shape of objects, we will exploit perspective effects, which have been disregarded in most scene analysis. We will introduce vanishing points for spatial information determination in the case of objects which have parallel lines or edges on planar surfaces.

First, in Chapter 2, we make clear the role that the location of the vanishing point or vanishing line plays in the three-dimensional measurements. Then we present two methods for the extraction of vanishing points and show that some spatial information about camera motion can be extracted.

Chapter 3 derives the constraints on the three-dimensional shape of vertices imposed by vanishing points in perspective drawings and describes the feasibility of shape-understanding of scenes and/or objects from partial or imperfect line drawings. It is assumed that objects are made of mutually orthogonal trihedral vertices and that rough positions of vanishing points are *a priori* given in the perspective drawing.

Chapter 4 presents a process that reconstructs the frontal view of a building from its perspective image as a practical application of the three-dimensional shape reconstruction discussed in Chapter 2. We use the location

of the vanishing point as a cue in calculating the distance and shape of the object. Additionally, by using the constraints imposed by vanishing points discussed in Chapter 3, we show how we can recognize the three-dimensional shape of vertices and partially analyze spatial structure of the object in an outdoor scene.

The other area of research, the display of the three-dimensional information, is attacked by constructing a system that produces stereograms and reconstructs cross-sectional images of a human heart. The display of spatial information is as crucial as its analysis. Spatial features should be shown to a human in a three-dimensional fashion. For example, in the field of cardiology the understanding of the shapes of cardiac chambers and the spatial relations of neighboring echoes from a number of tomograms requires much practical experience, and a method which can provide an objective three-dimensional display has long been hoped for.

In Chapter 5, we treat a three-dimensional cardiographic imaging system based on the principle of binocular stereoscopy. The system produces a three-dimensional image of the heart chamber from two-dimensional representations of ultrasono-tomograms. We also present a method that allows us to view the heart from any orientation at any location.

Finally, Chapter 6 concludes the thesis with a discussion of its contributions and possible future improvements.

CHAPTER 2

DETERMINATION OF THE VANISHING POINT IN OUTDOOR SCENES

2.1 Introduction

This chapter shows that spatial information about objects can be extracted from visual observations with only one sensor by exploiting knowledge about perspective projection. First we introduce the concept of vanishing points for the purpose of determining three-dimensional measurements. We derive the equations for the three-dimensional information as a function of vanishing point and vanishing line location.

Orientation of computer-simulated surfaces has been determined by using vanishing points [6,7]. Here, we present two methods for the extraction of vanishing points in natural outdoor scenes [8,9]: one is based on the Hough Transform; the other is an iterative method.

2.2 Spatial Information Determination Based On Perspective Transformation

Vanishing points and vanishing lines can provide three-dimensional information in the case of objects which have parallel lines or edges on planar surfaces. Parallel lines in three-dimensional space are projected onto the image plane to lines which radiate from a single common point. This is called a vanishing point. Once the location of the vanishing point is detected, we can

use it as a cue in calculating the distance and shape of the object to which the parallel lines belong. We will make clear the role that the location of the vanishing point or vanishing line plays in the three-dimensional measurements.

2.2.1 Vanishing Points

We will describe the relation between an object point and its image point (Figure 1). In what follows, each point will be represented as a vector from the view point. We take x-y axis to be parallel to the image plane and z axis in the direction from the view point to the image center. Letting $\vec{n}_f =< 0,0,1 >$ be the normal unit vector and f the distance of the image plane from the view point, then for each image point p,

$$\vec{p} \cdot \vec{n}_f = f \ . \tag{1}$$

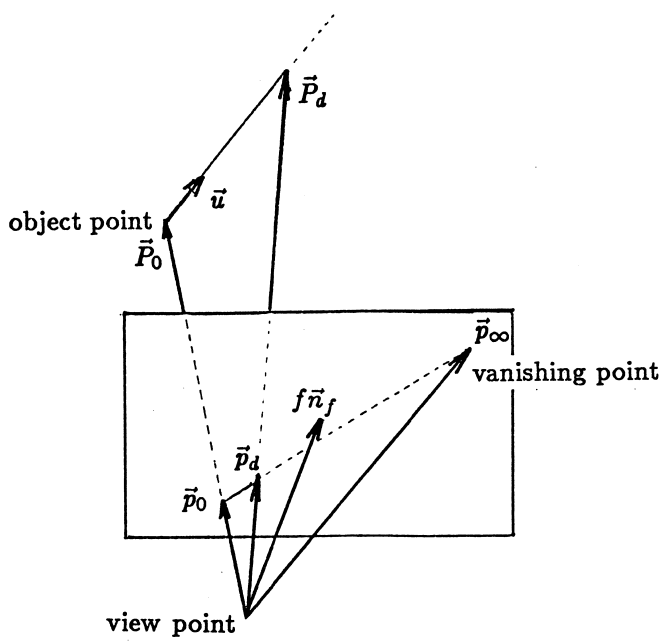


Figure 1. Vanishing point.

When a point P_0 in three-dimensional space is projected to a point p_0 in the image plane, P_0 and p_0 are related as

$$\vec{P}_0 = a_0 \vec{p}_0 \ , \tag{2}$$

where a_0 is a constant depending on the distance of the point. Imagine a line going through the point P_0 with a direction $\vec{u}$ ($\|\vec{u}\| = 1$). A point P_d on this line at a distance d from P_0 is represented as

$$\vec{P}_d = a_0 \vec{p}_0 + d\,\vec{u} \tag{3}$$

$$d = \|\vec{P}_d - \vec{P}_0\| \ . \tag{4}$$

Let p_d be the image point of P_d. Then we can represent it as

$$\vec{P}_d = a_d \vec{p}_d \qquad (a_d : \text{constant}). \tag{5}$$

From (1),(3) and (5), we can calculate the location of the image point p_d:

$$a_d = \frac{a_0 f + d\,\vec{u} \cdot \vec{n}_f}{f} \ , \tag{6}$$

$$\vec{p}_d = \frac{f(a_0 \vec{p}_0 + d\,\vec{u})}{a_0 f + d\,\vec{u} \cdot \vec{n}_f} \ . \tag{7}$$

Now we can determine the location of the image of P_d when the distance d becomes infinite on any such line where $\vec{n}_f \cdot \vec{u} \neq 0$, namely, not parallel to the image plane. From equation (7),

$$\vec{p}_\infty = \lim_{d \to \infty} \vec{p}_d = \frac{f}{\vec{n}_f \cdot \vec{u}} \vec{u} \ . \tag{8}$$

This shows that the location of p_∞ in the image is independent of the distance a_0 and can be calculated from f, $\vec{n}_f$ and $\vec{u}$. That is, all the receding parallel lines in three-dimensional space vanish at a unique image point —

independently of their distances. The point p_∞ is the vanishing point. Every line with a direction $\vec{u}$ which satisfies $\vec{n}_f \cdot \vec{u} \neq 0$, i.e., receding from the view point, has a vanishing point.

Conversely, if we know the location of the vanishing point in the image, we can calculate the line direction $\vec{u}$ in three-dimensional space:

$$\vec{u} = \frac{\vec{p}_\infty}{\|\vec{p}_\infty\|} . \tag{9}$$

2.2.2 Calculating the Distance of a Point Based on the Vanishing Point

Now we can calculate the distance d in three-dimensional space between the two points P_d and P_0 based on the locations of those image points and the vanishing point. From (6),

$$\begin{aligned} d &= \frac{a_0 f \|\vec{p}_d - \vec{p}_0\|}{\|f\vec{u} - (\vec{n}_f \cdot \vec{u})\vec{p}_d\|} \\ &= \frac{a_0 \|\vec{p}_d - \vec{p}_0\| \ \|\vec{p}_\infty\|}{\|\vec{p}_\infty - \vec{p}_d\|} . \end{aligned} \tag{10}$$

That is, when $\vec{p}_\infty$ (the location of the vanishing point), f (the focal length), and a_0 (distance to a standard point) are given, we can calculate the direction $\vec{u}$ and distance d of a point on the line using (9) and (10), respectively.

Though we assumed that $\vec{n}_f \cdot \vec{u} \neq 0$ in equation (8), when $\vec{n}_f \cdot \vec{u} = 0$, i.e. the line $\overline{P_0 P_d}$ is parallel to image, equation (10) becomes

$$d = a_0 \|\vec{p}_d - \vec{p}_0\| . \tag{11}$$

2.2.3 Calculating Surface Orientation Based on the Vanishing Line

We will show that we can determine a surface orientation from the vanishing line in the image. We assume that the vanishing points p_∞^1 and p_∞^2 on a plane are given in the image (Figure 2). The line $\overline{p_\infty^1 p_\infty^2}$ is called the vanishing line of this plane. All the vanishing points of the lines on the plane are located on this line. All the surfaces parallel to the plane, i.e. with the same unit normal, have the same vanishing line in the image. To know the surface orientation is to calculate the unit normal vector perpendicular to the surface. Let us calculate the unit normal vector from the location of the vanishing line in the image. Let $\vec{N}$ denote the unit normal vector of an object surface, $\vec{n}_l$ the unit normal vector of the vanishing line in the image, and l the distance of the vanishing line from the image center $f\vec{n}_f$. Then $\vec{N}$ can be calculated from $\vec{n}_f$, $\vec{n}_l$, f and l:

$$\vec{N} = \frac{l\vec{n}_f - f\vec{n}_l}{\sqrt{l^2 + f^2}} . \tag{12}$$

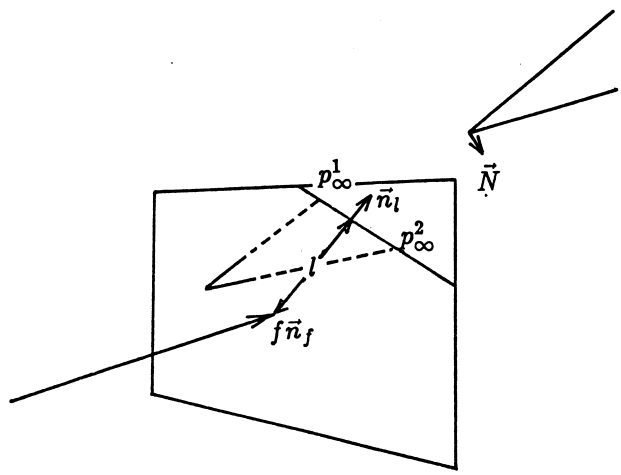


Figure 2. Vanishing line.

2.3 Determining a Vanishing Point Using the Hough Transformation

In the previous section we discussed the role of a vanishing point as a depth cue. Let us now present a technique for extracting a vanishing point from a real outdoor scene. As we discussed above, all the receding parallel lines in the environment are projected onto the image plane to lines radiating from a vanishing point. Many edges of man-made objects such as buildings and roads are mutually parallel. If we can extract edge segments of a straight line from a picture, we can obtain a vanishing point as the intersection of the extensions of those parallel edges.

However, there are two main difficulties in this approach in addition to unsatisfactory performance of line finders: (1) necessity of a line finder which provides the same result on a line composed of plural line segments as on a single contiguous line having the same number of edge elements, and (2) possibility of many intersections due to the combinations of lines, and difficulty in judging which intersection gives the vanishing point.

To overcome those difficulties, we will adopt the Hough Transformation which has usually been used to detect lines[10], and will extract a vanishing point from images of a natural scene.

2.3.1 The Hough Transformation [10]

The Hough transformation maps a point in the x-y plane into a sine curve in the θ-ρ plane, or Hough plane by the following formula:

$$\rho = x\cos\theta + y\sin\theta \ . \tag{13}$$

A straight line is mapped onto a point in the Hough plane; A point

can be thought to be an intersection of lines passing through it and is mapped into a sine curve as a map of those lines. Every edge element in a picture is interpreted to be a line segment with a unit length and is mapped onto a point in the Hough plane, whose information is deposited as a two-dimensional array or histogram $H(\theta, \rho)$.

All the collinear edge elements are mapped onto the same point in the Hough plane and the number of such elements is accumulated in the histogram. Whether a picture contains a single long line or plural short lines, the resulting histogram $H(\theta, \rho)$ will be the same as long as the number of the edge elements is the same. This is an important feature of this transformation. A vanishing point in a picture is represented by a sine curve in the Hough plane on which a large number is accumulated because it is derived from an intersection of several lines.

2.3.2 An Algorithm

We deal with pictures which have a single significant vanishing point. The process is as follows.

Step 1 Edge detection: Using the Sobel edge detector [11], we generate a gradient picture, where the magnitude and the direction at a point (x, y) are given by m_{xy} and θ_{xy}. If the magnitude m_{xy} is above a threshold m_0, an edge element is declared at the point. The threshold should be chosen so that only strong straight-line edge segments, in this experiment 20% of the pixels, are retained.

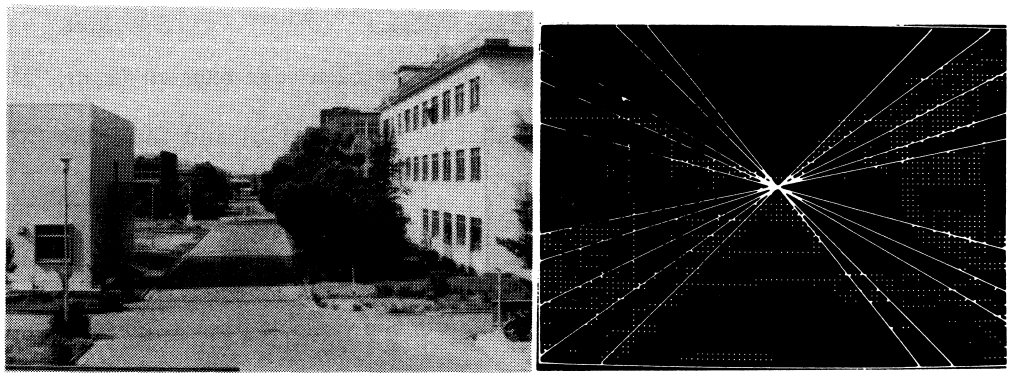
Step 2 Hough Transformation — Mapping the image plane into the Hough plane: We use a two-dimensional array (60 × 60) as a two-dimensional histogram $H(\theta, \rho)$; $0° \leq \theta < 180°$, $|\rho| \leq (x_{\max}^2 + y_{\max}^2)^{1/2}$, where $x_{\max}$ and $y_{\max}$ are the picture width and height, respectively. For each edge element (x, y), corresponding pair of (θ_{xy}, ρ_{xy}) is calculated by equation (13) and the value $H(\theta_{xy}, \rho_{xy})$ is incremented by one.

Step 3 Elimination of vertical and horizontal lines: The vertical and horizontal edges should be neglected in the determination of the vanishing point, for they contribute toward forming the corners of buildings rather than vanishing points. To eliminate them, $H(\theta, \rho)$ is fixed to zero for $\theta \leq \alpha$, $180° - \alpha \leq \theta$, or $90° - \alpha \leq \theta \leq 90° + \alpha$, where the parameter α is 3° on account of noise.

Step 4 Mapping the Hough plane into the image plane: Mutually converging line segments are mapped into a sine curve in the Hough plane. We should search for a sine curve that is the transform of a vanishing point. Numerical analysis offers no effective method other than the following brute force way. A two-dimensional histogram $V(x, y)$ on the image plane is computed by summing the values of $H(\theta, \rho)$ for a set of (θ, ρ) pairs which satisfy (13).

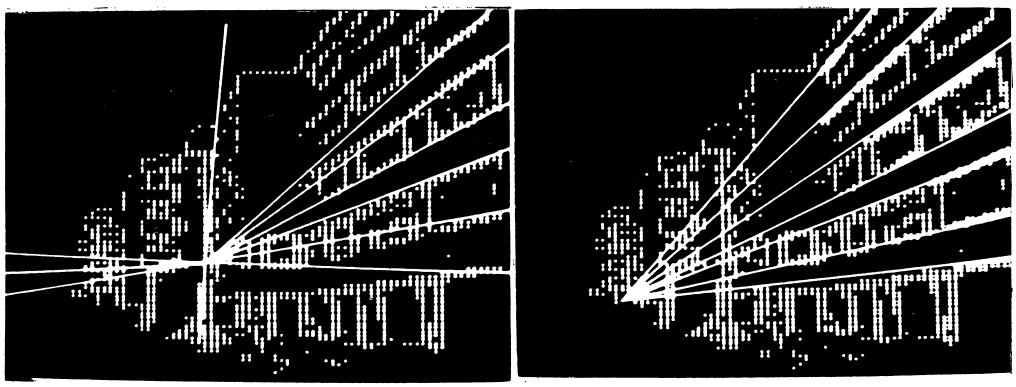
Step 5 Determination of vanishing point location: The vanishing point is defined by the point (x_v, y_v) where the value $V(x_v, y_v)$ is the maximum.

Figure 3 shows the result of the application of the method to an outdoor scene. The significant lines passing through the vanishing point are drawn and edges are represented by dots.



(a) (b)

Figure 3. Extraction of vanishing point.
(a) scene (128×190), and (b)vanishing point.



(a) (b)

Figure 4. Improvement of method.
(a) candidate vanishing point, and (b) final vanishing point.

2.3.3 Improving the Method

The method fails in finding correct vanishing points in some pictures, which often have misleading strong edges such as vivid boundaries between the sky and the scenery. To improve the method we assume the first result (x_v, y_v) to be a candidate for the real vanishing point and take the convergent point in the following procedure as the real vanishing point.

Step 6 Remove from the given picture the circular area which satisfies the following inequality.

$$(x - x_v)^2 + (y - y_v)^2 < \gamma^2, \quad \gamma = 0.3(x_{\max}^2 + y_{\max}^2)^{1/2}. \tag{14}$$

The term γ is chosen so that edges of the scene in the distance may be removed.

Step 7 We calculate m_{xy} and θ_{xy} in the same manner as in step 1. Here edge elements are limited to points (x, y) which satisfy the following conditions:

$$\left.\begin{array}{ll} \text{(i)} & m_{xy} > m_0 \\ \text{(ii)} & (x - x_v)^2 + (y - y_v)^2 > \gamma^2 \\ \text{(iii)} & \left|\tan^{-1}\dfrac{x - x_v}{y - y_v}\right| < \theta_0 \end{array}\right\}, \tag{15}$$

where $\theta_0 = 20°$ in this experiment.

Step 8 We iterate the steps 6,7 and 2-5 until the two consecutive resulting coordinates come to be equal. Figure 4 shows an example.

2.3.4 Deriving the Camera Motion Parameters

We can know the state of the camera from the position of the vanishing point in a picture. The pan and tilt angle of the camera can also be calculated from the shift of the vanishing points in a pair of successive pictures on condition

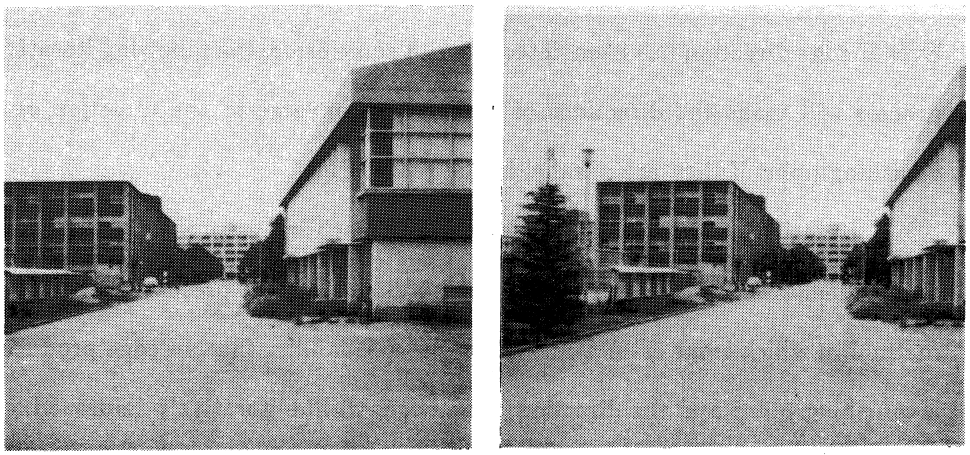
that the camera does not rotate. Suppose we have two pictures taken in different directions and they have a common vanishing point V, whose coordinates in each picture are (x_{v_1}, y_{v_1}) and (x_{v_2}, y_{v_2}). Then the pan and tilt angle are calculated as follows:

$$\left.\begin{aligned} &\text{horizontal direction}: \quad \theta_x = \tan^{-1}\left(\frac{x_{v_2}}{kf}\right) - \tan^{-1}\left(\frac{x_{v_1}}{kf}\right) \\ &\text{vertical direction}: \quad \theta_y = \tan^{-1}\left(\frac{y_{v_2}}{kf}\right) - \tan^{-1}\left(\frac{y_{v_1}}{kf}\right) \end{aligned}\right\}, \tag{16}$$

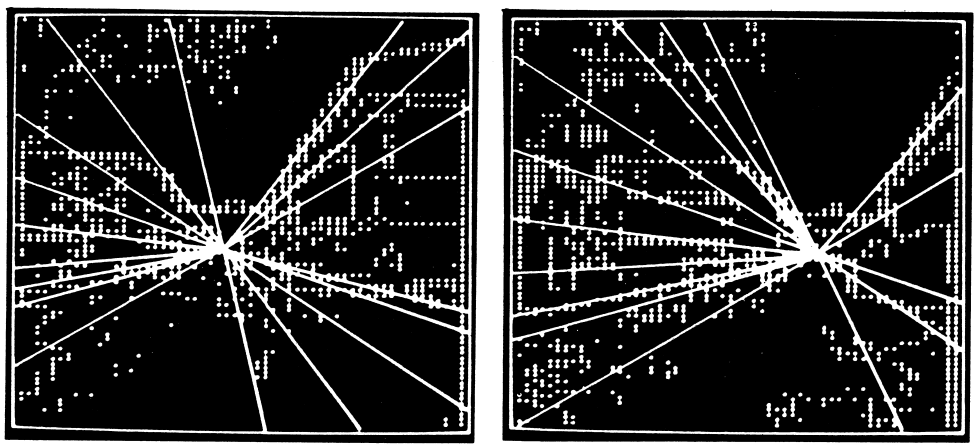
where f: the focal distance,

k: the ratio of the size of a processed picture to that of the film.

An experimental result is shown in Figure 5. The pan angle is calculated to be 8.0 degrees by the computer, while the angle is calculated to be 9.2 degrees by hand-calculation from the original pictures. The error occurs mainly by the incorrect value of the ratio k which is due to the inaccurate framing of the pictures that happens when we store them in the computer. An error in measurement of 1 mm in a picture would lead to an error of 0.5 degrees in the result.



(a)



(b)

Figure 5. Angle of pan.
(a) input image, and (b) vanishing point.

2.4 Iterative Method for Determining a Vanishing Point

This section proposes another method for determining a vanishing point. Even if edge direction has been detected with some error, the following iterative process will make the directions of such edges converge to the direction of a vanishing point if they are parallel in the real world.

Iterative approaches have been used for such image processing as labeling, segmentation, noise compression [12,13,14]. We will present an iterative method in which some estimates, such as the probability that an edge produces a vanishing point and the direction of an edge, will be iteratively improved.

2.4.1 The Data

We will deal with a picture which can be considered to be one-point perspective. We generate a gradient picture by the Sobel operator [11] and detect those edges whose gradient is within the upper 10%. Let θ_e be the direction of an edge $e = (x_e, y_e)$.

2.4.2 A Candidate Vanishing Point

A line associated with an edge e is given by

$$y - y_e = \tan\theta_e(x - x_e) \ . \tag{17}$$

The position of its intersection with a line of another edge $e' = (x_{e'}, y_{e'})$ is calculated, and the corresponding position $L(w)$ in a one-dimensional array L is incremented by one (Figure 6). We compute all the intersections of every edge except for vertical and horizontal lines because they do not produce a vanishing point on the assumption of one-point perspective.

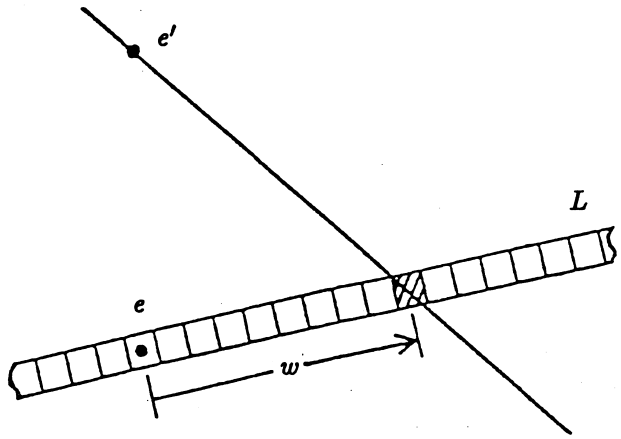


Figure 6. Candidate vanishing point.

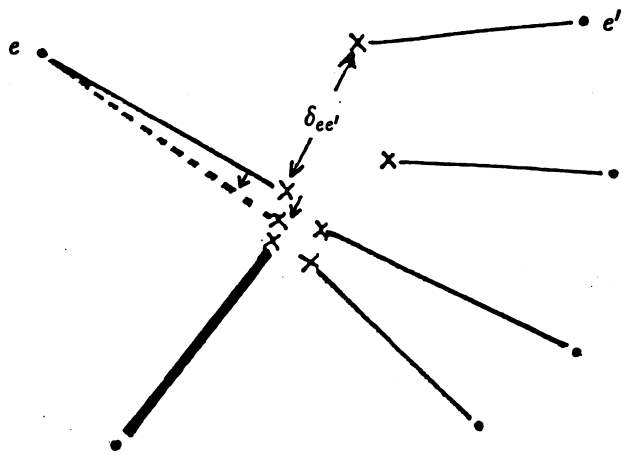


Figure 7. New location of a vanishing point.

The maximum point of the resulting $L(w)$ makes a candidate vanishing point $V_e = (x_{v_e}, y_{v_e})$ for edge e. And the probability that V_e is a vanishing point of e is calculated by

$$W_e = \frac{\max_w L(w)}{\sum_w L(w)} . \tag{18}$$

These values are used as the initial estimates for the following iteration.

2.4.3 Iterative Update

We first define the degree of consistency, $R_{ee'}$, between two edges e and e'. The role of $R_{ee'}$ is to estimate the possibility that those two edges have the same vanishing point. A candidate vanishing point will have a big $R_{ee'}$ when there is another candidate at a short distance and its $W_{e'}$ is big.

$$R_{ee'} = \begin{cases} W_{e'}(1 - \delta_{ee'}/\delta) & (\delta_{ee'} \leq \delta) \\ 0 & (\delta_{ee'} > \delta) , \end{cases} \tag{19}$$

where $\delta_{ee'}$ is the distance between e and e', and δ is a constant.

We now use the following rule to update the location of a candidate vanishing point V_e of edge e using $R_{ee'}$.

$$x_{v_e}^{(k)} = \frac{\sum_{e' \neq e} R_{ee'}\, x_{v_e'}^{(k-1)}}{\sum_{e' \neq e} R_{ee'}} , \quad y_{v_e}^{(k)} = \frac{\sum_{e' \neq e} R_{ee'}\, y_{v_e'}^{(k-1)}}{\sum_{e' \neq e} R_{ee'}} , \tag{20}$$

where $(k-1)$ denotes an old value and (k) denotes a new value. Figure 7 illustrates a scheme of updating the location of a vanishing point, where "$\bullet$" represents an edge, "$\times$" a candidate vanishing point, and a line width is proportional to the value of W. The direction of edge $e = (x_e, y_e)$ is also updated by the new vanishing point location,

$$\theta_e = \tan^{-1}\left(\frac{y_{v_e}^{(k)} - y_e}{x_{v_e}^{(k)} - x_e}\right) . \tag{21}$$

Using these refined values, we repeat the procedure from 2.4.2. After several iterations $R_{ee'}$ becomes 0 or close to 0 and W_e also becomes smaller at the place where a vanishing point hardly exists. We determine the vanishing point V_e at the location that gives the maximum W_e after a certain number of iterations. The direction of those edges, which are parallel in three-dimensional space and produce a vanishing point, will be refined by this iterative process and will all be in the direction of the vanishing point.

2.4.4 Test Results

Figure 8 shows an experimental result of the iterative process, where an edge and its candidate vanishing point are represented by connecting them by a line. In this experiment the size of the outdoor picture is 110×85 points and the parameter δ is set to 5 in equation (19).

Relationships between edges and their vanishing point in the initial state and after 3 iterations are shown in Figure 9. The windows on the second floor of the right side building are not detected because vertical and horizontal edges are eliminated in the process. As this result shows, since edge directions are refined, we can easily detect edges that produce a vanishing point and are parallel in real world.

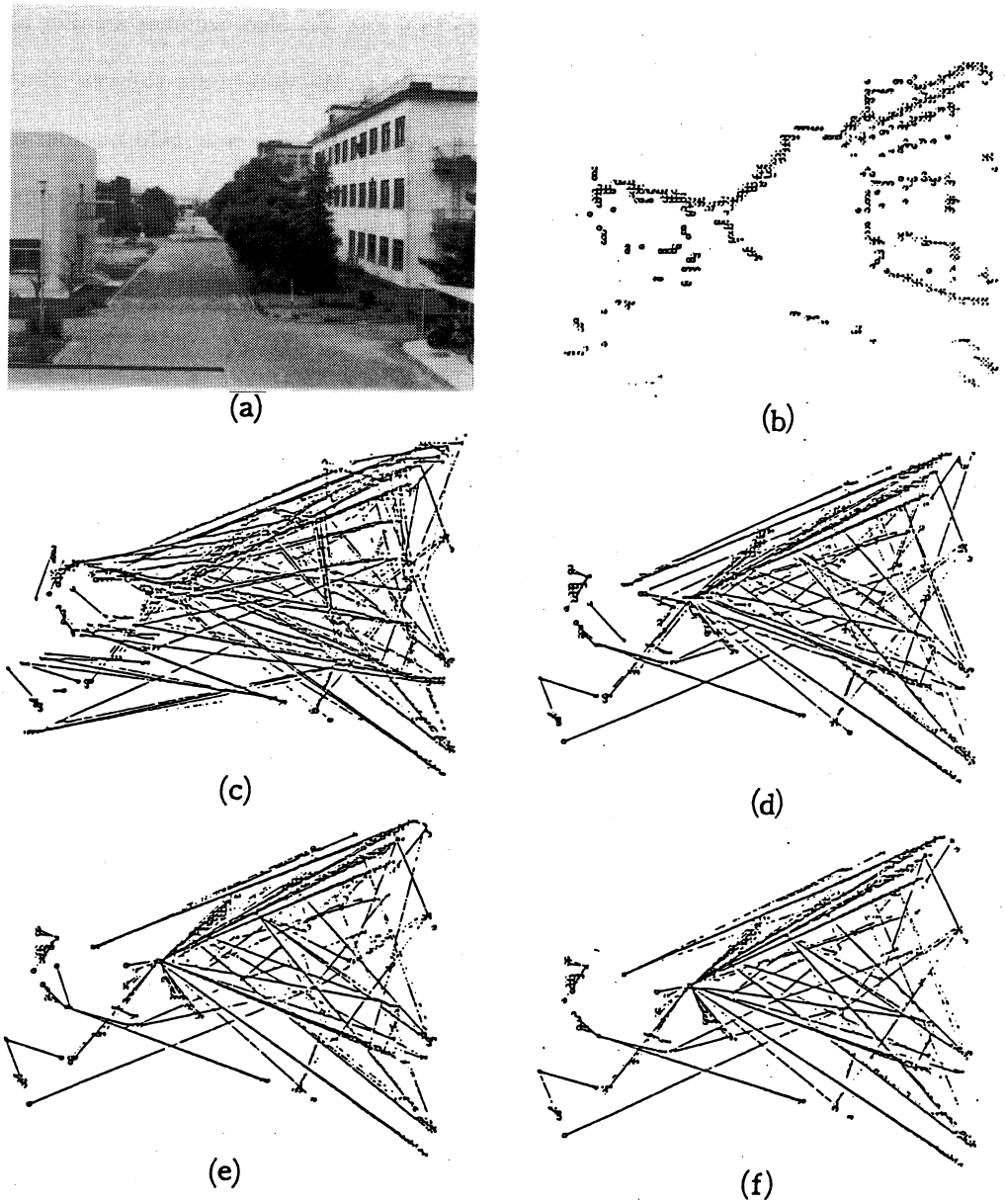


Figure 8. Example. (a)scene, (b)edge, (c)initial, (d)iteration 1, (e)2, (f)3.

Figure 9. Detection of parallel lines.
(a)initial estimate, and (b)on iteration 3.

2.5 Conclusion

Perspective transformation is one of many cues to three-dimensional spatial information about objects in a two-dimensional image. We have presented two algorithms for detecting a vanishing point in an outdoor scene. The role of a vanishing point is as follows:

(1) A cue for reconstructing spatial structure from a two-dimensional image — We have considered the computation of length of line segments and surface orientation, and derived the equations for the three-dimensional information as a function of vanishing point and vanishing line location. The application to the shape reconstruction is discussed later in Chapter 4.

(2) A standard point for registering images — We have shown that we can calculate motion parameters of a camera by using the coordinates of the vanishing point in each picture. Extraction of spatial information from a human observer in a hallway is experimented with in [8].

CHAPTER 3

INFERRING THREE–DIMENSIONAL SHAPE FROM LINE DRAWINGS USING VANISHING POINTS

3.1 Introduction

In this chapter, we discuss how to determine three-dimensional structure of scenes and/or objects in line drawings especially from the view point of the possibility of labeling of partial or imperfect line drawings by the use of vanishing points[15].

A variety of papers have been published on labeling line drawings for the determination of the shape of polyhedra [16,17]. Each line in the drawing is interpreted as a concave edge, convex edge, or a boundary and is identified as such on the drawing by some label. By labeling all the lines consistently, one can understand the physical situation of individual objects and the spatial relations between objects.

We introduce vanishing points for understanding line drawings. We often have to tackle scene analysis within a domain of imperfect and partial line drawings because of the difficulty of extracting line drawings from an image. By previous line labeling techniques we cannot determine a vertex shape if we do not have a perfect line drawing. This chapter shows that the labeling technique using vanishing points is useful in such situations.

Constraints imposed by vanishing points make three-dimensional shape determination of vertices possible from the drawing of the vertex of interest

alone without the necessity of comparing adjacent junctions. For some vertices we can produce a unique labeling and for others we can limit the set of allowable labels, all using vanishing points.

3.2 Labeling Trihedral Vertices

3.2.1 Trihedral World

First we will define terms and give the assumptions used in this chapter. A vertex and an edge of an object in three-dimensional space are referred to as a junction and a line in the two-dimensional line drawings, respectively.

The line drawings of a polyhedron are assumed to meet the following conditions: (1) trihedral vertex - all vertices are the intersection of exactly three plane surfaces which are mutually orthogonal. (2) general position - a small movement of the viewpoints results in an essentially similar picture.

3.2.2 Conventional Approaches [11,16,17]

Huffman [16] and Clowes [17] worked with the trihedral assumption that three planes divide space into eight parts. Thus, the types of trihedral vertex can be characterized by the octants of space around the vertex which are filled by solid material. Out of these, four geometries determine the essential trihedral vertices of the form 1, 3, 5, and 7 as shown in Figure 10, where the names originated from the number of the filled octants.

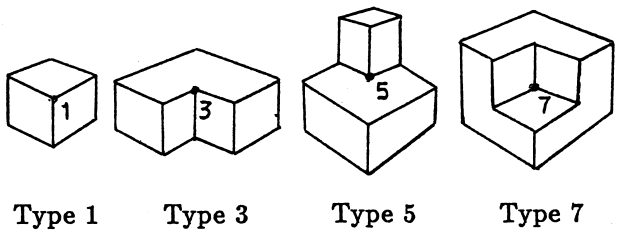


Figure 10. Vertices of trihedral solids.

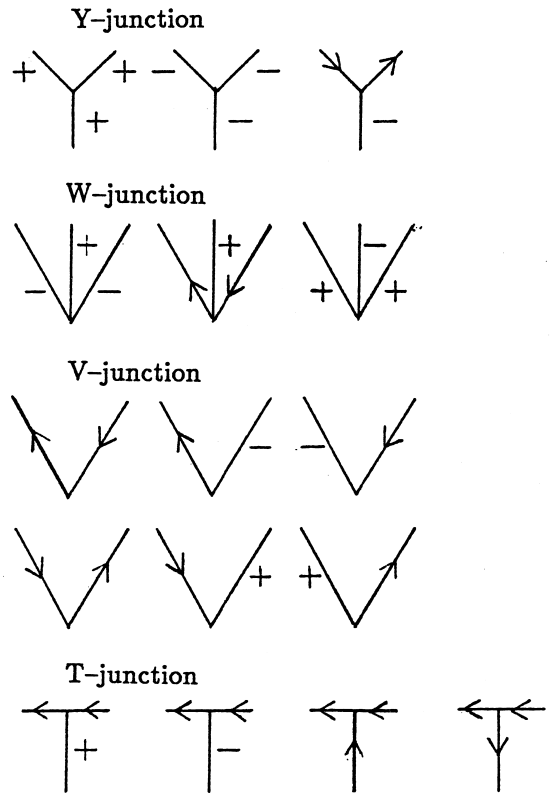


Figure 11. All possible representations of trihedral vertices.

When one looks at essential forms of vertices from each unfilled octant, one can also assign line labels to every form of a drawing junction. The line labels identify the line types: Convex edges are labeled with "+", concave edges with "−", and boundaries with ">" or "<". The direction of the arrow is chosen such that the unoccluded plane lies to the right of it.

In addition to those junctions derived from real trihedral vertices, there are junctions from partial occlusion, i.e., T–junctions as described later. One can divide all junctions into four classes according to the number of lines and the angle between the lines ; Y–, W–, V–, and T–junctions. The resulting catalog of all possible representations of the junctions is shown in Figure 11.

A consistent labeling should satisfy the following two conditions: (1) the labeling assigned to a junction should be one of the physically possible configurations in Figure 11. (2) At the adjacent junctions, the two labels assigned to the common line should be coherent.

Consequently, in order to decide a three-dimensional shape of a vertex as well as the labeling of the junction, it is necessary to find appropriate line labels from all possible labels, "+", "−", ">", or "<", by comparing adjacent junctions.

3.3 Constraints by Vanishing Points

The following terms will be used in describing three-dimensional shapes of vertices, though there may be some arguments about types 3 and 5: "Concave vertices" means vertices of type 1 or 5, and "Convex vertices" means vertices of type 3 or 7.

In general, one cannot determine three-dimensional shapes of vertices solely by inspecting the vertex of interest. Even if one knows the type of a junction, Y–junctions for instance, one cannot determine whether the vertex is concave or convex because its type could be 1, 3, or 7.

Besides the trihedral and the general position assumptions, some additional clues or assumptions influence the three-dimensional shape determination. The analysis of line drawing with shadows [18], with hidden lines [19], and of origami world [20] has been discussed. The mathematical formulation of the problem has also been considered [21,22].

We study the case in which vanishing points are introduced to the analysis of perspective drawings. Since every picture is a perspective projection of a real world view and the line drawings obtained from it also are perspectively deformed, analysis of perspective line-drawings should be considered.

3.3.1 Edges and Vanishing Points

We will show that at most three vanishing points are generated by a trihedral vertex. Let p be the image of an object point P on an edge E (Figure 12). The image point p_∞, which is the limit of the image point P as P recedes to infinite along E, is the vanishing point of the image of the edge E. This is determined uniquely by the intersection of the image plane with the line which goes through the viewpoint and is parallel to E. Therefore, any receding edge in three-dimensional space has a vanishing point in the image plane, and a trihedral vertex has at most three vanishing points.

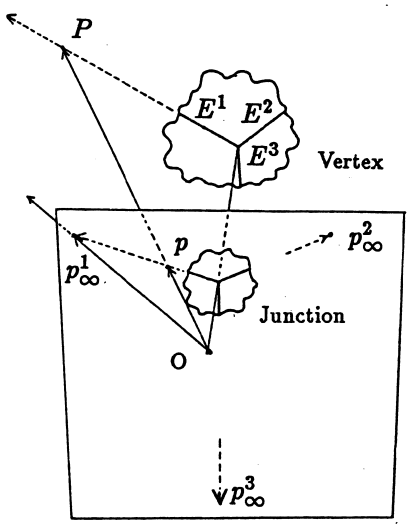


Figure 12. Edges and vanishing points.

3.3.2 Constraints on the Three–Dimensional Shape of a Vertex by Vanishing Points

The constraint added by vanishing points to the three-dimensional shape of vertices can be analyzed by Figure 13. A T–junction is not constrained by vanishing points because it does not correspond to any physical vertex. T–junctions are excluded from consideration hereafter.

Figure 13 shows the catalog of all possible junction labels and vanishing points for every type of vertex. There are generally three vanishing points generated by the three edges. The direction of the vanishing point of an edge is determined, and is marked with the arrow "↑".

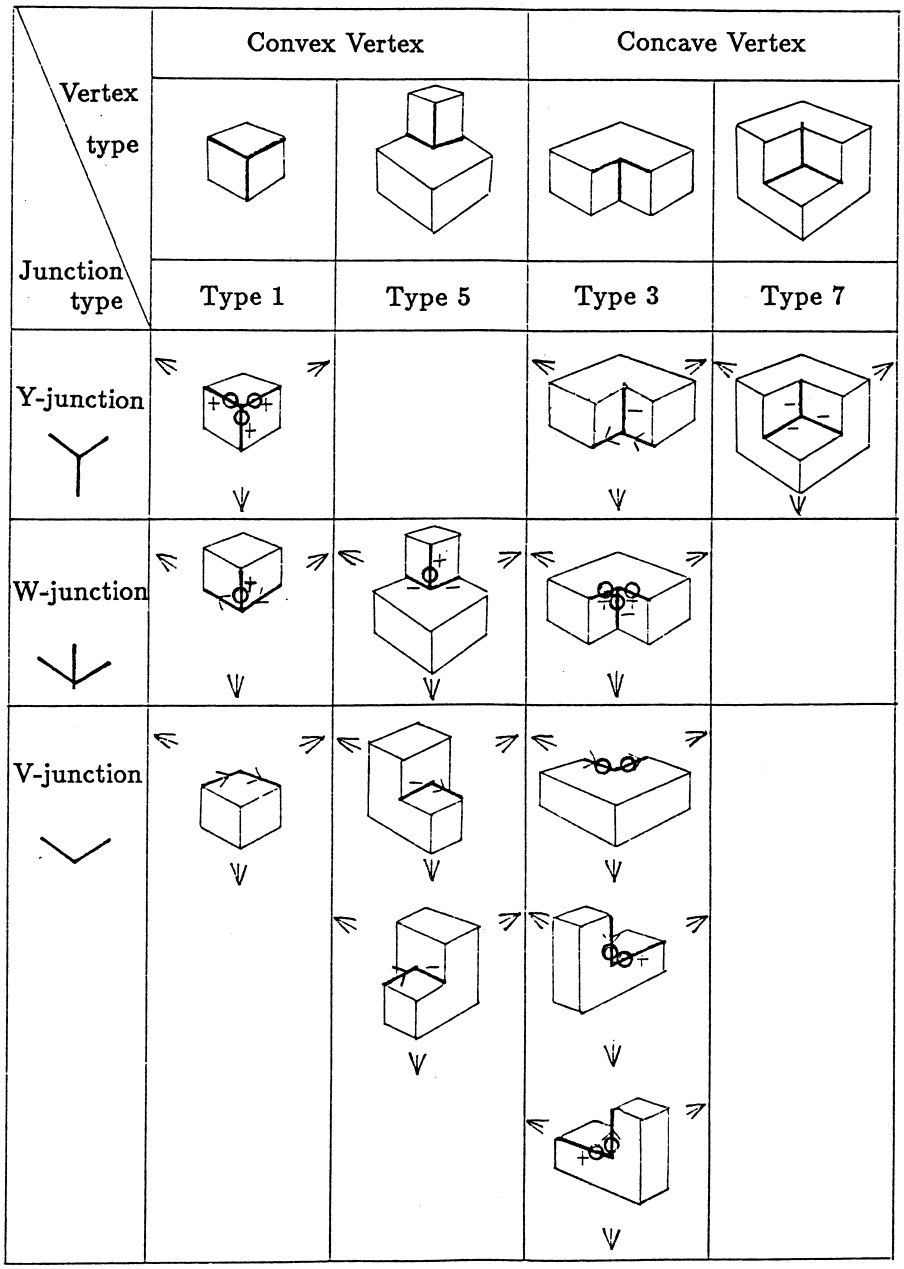


Figure 13. All possible labels and vanishing points.

By inspecting Figure 13, one can see that vanishing points add constraints on the three-dimensional shape determination and on vertex labeling. For example, when the three lines of a Y–junction are in the direction of vanishing points from the junction, the junction can be interpreted as a convex vertex of type 1, and all the lines of the junction are forced to be labeled uniquely with + labels.

Similarly, whether a vertex is concave or convex can be determined according to the number of lines in the direction of vanishing point. Denote the number of those lines by J. The three-dimensional shape of a vertex is determined by J: (1) Y–junctions: concave if J=0; convex if J=3, (2) W–junctions: concave if J=1; convex if J=2, and (3) V–junctions: convex if J=0; concave if J=1 or J=2.

The difference between V–junctions and Y–, W–junctions is explained as follows. A vertex of V–junction type has a hidden edge in addition to two visible edges. Imagine that the hidden edge becomes visible; then such vertex would become a Y– or W–junction. Concave vertices of V–junctions would correspond to convex ones of Y– or W–junctions; convex vertices would correspond to concave ones.

3.3.3 Constraints on Line Labeling

Vanishing points are useful in determining the three-dimensional shape of a vertex as mentioned above. It is important to note that they are also useful in determining three-dimensional properties of an edge, e.g. whether it is concave, convex or boundary, and they limit the set of allowable line labels.

Figure 14 shows possible junction labels based on the type of a junc-

tion and the number of lines in the direction of the vanishing point. One can unambiguously determine the edge geometry and give a unique label to lines marked with circles in Figures 13 and 14 by using the junction type and vanishing points. For labeling V–junctions of type 3 vertices uniquely, the placement of the lines can be used as a constraint. The remaining lines are not marked with circles; even so, it is possible to determine that the edge is either concave or a boundary. Therefore, the number of possible labels can be reduced.

More ambiguity, however, is left in some cases in determining three-dimensional shape of vertices even when we use vanishing points. These are the vertices which cannot be labeled uniquely only by using the junction type and vanishing points. The ambiguity concerning vertex type appears: one cannot determine whether a concave vertex is type 3 or 7, or whether a concave one is type 1 or 5. Concerning an edge, there is also ambiguity in distinguishing boundary lines from interior lines, i.e. a concave edge from a boundary.

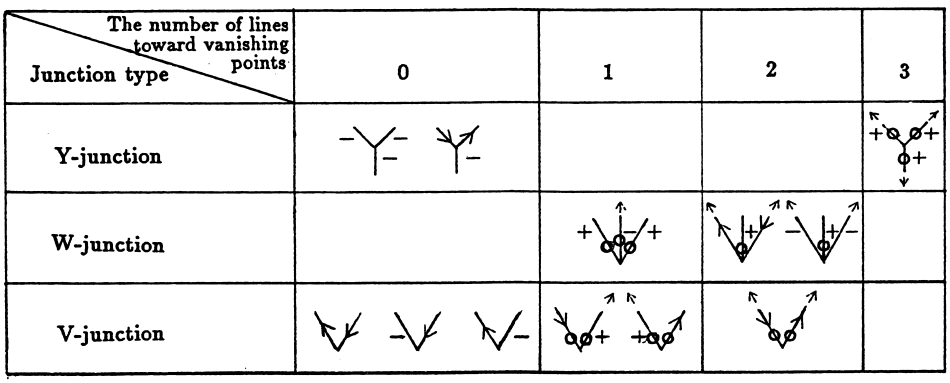


Figure 14. Labeling with vanishing points.

The ambiguity arises from the following two situations: One is that a cube rests on a material of the background and the other is that a cube is suspended in isolation. These situations cannot be distinguished by inspecting only the vicinity of the vertices.

Consider, for example, type 1 and type 5 vertices of W–junction in Figure 13. A vertex is of type 1 if the hidden surfaces behind the remaining two visible surfaces are apart from a background; and it is type 5 if the hidden surface sticks to another surface. Evidently it is impossible to determine if this surface sticks to another or not, solely by inspecting the vertex of interest.

3.4 Three Perspectives of Trihedral Vertices

Introduction of vanishing points to three-dimensional shape determination of trihedral vertices has been discussed in the preceding section, where the drawing has been assumed to be three-point perspective, i.e. each junction has exactly three vanishing points.

Perspective drawing is classified into one-point, two-point, and three-point perspective, where the number of vanishing points is one, two, and three, respectively. No edge generates a vanishing point when it is parallel to the image plane. According to the number of such parallel edges out of three, a drawing can be thus classified.

In this section, we will show that vanishing points add constraints on determining three-dimensional shape of vertices and limit the possible labels in the case of one-point and two-point perspective in the same manner as three-point perspective discussed in Section 3.3.

3.4.1 Junction Type in One–Point Perspective

When two edges of a trihedral vertex are parallel to the image plane, the remaining edge, which is not parallel, generates a vanishing point. In this case the junction is of one-point perspective.

First it is necessary to list all the possible placements of junctions and vanishing points in the case of one-point perspective. A view makes a junction in one-point perspective when two of the three edges are parallel to the image plane. Three junctions in one-point perspective correspond to each junction in three-point perspective according to which vanishing point out of three is chosen to remain. At this time, junctions looking like any other by rotation are made into one.

The correspondence between junctions in one-point perspective and in three-point perspective is shown in Figure 15. Junction types appearing in one-point perspective are classified to Y–, W_1–, W_2–, V_1–, V_2–, and V_3–junctions. A limitation can be placed on labeling Y–, W_1–, W_2–, and V_1–junctions by the number of lines in the direction of the vanishing point, and V_2– and V_3–junctions by the placement of lines and a vanishing point. Moreover, a unique labeling can be provided to lines marked with circles in Figure 15.

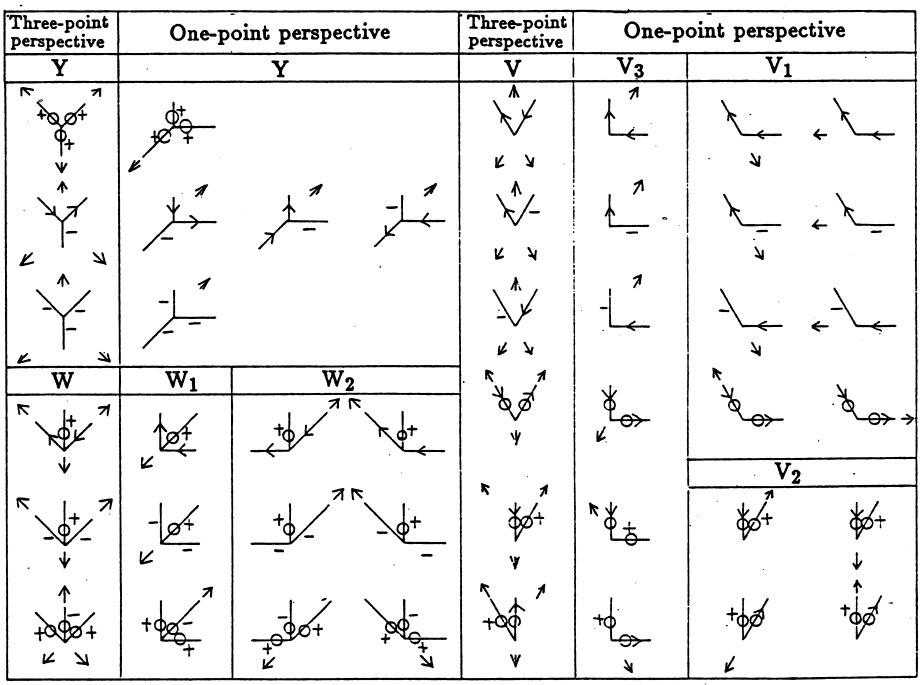


Figure 15. Junctions in one-point perspective.

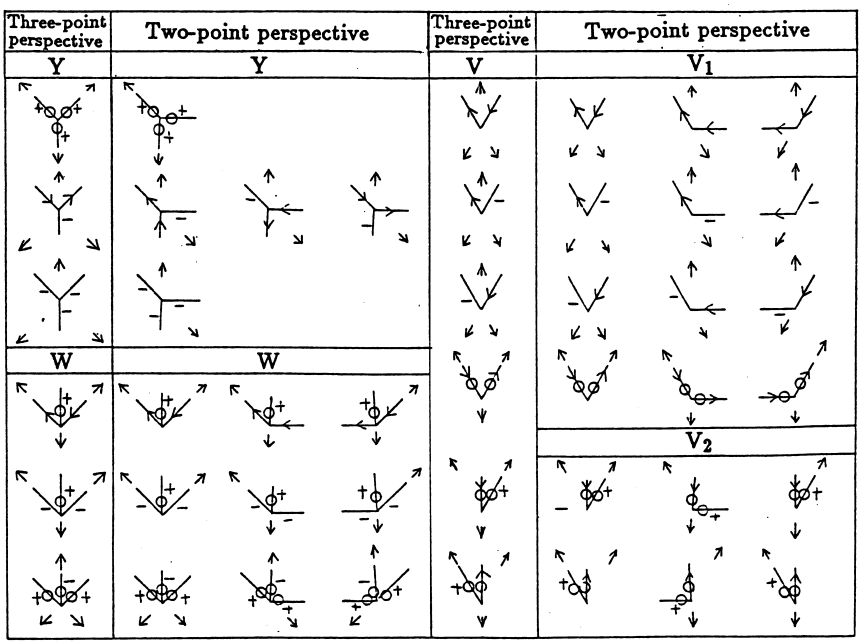


Figure 16. Junctions in tow-point perspective.

3.4.2 Junction Type in Two–Point Perspective

Where one of the three edges is parallel to the image plane, two vanishing points are generated, and the junction is of two-point perspective.

Figure 16 shows a list of all possible junctions and vanishing points configurations. Also in this case, three-dimensional shape of vertices can be determined and the number of possible labels can be reduced using vanishing points in the same manner as 3.4.1.

3.5 Comparison with Conventional Approaches

Figure 17 illustrates an example of labeling of a line drawing with vanishing points, and Figure 18 shows the search tree[11] for appropriate labels of the central part of Figure 17 by the method mentioned in the Section 3.2, i.e. without vanishing points. As can be seen from this example, for the purpose of determining three-dimensional shape of a vertex or for labeling of a junction it is necessary to compare the labels of the surrounding junctions.

On the other hand, vanishing points can provide unique labeling to all the preceding junctions. Concerning a line marked with a circle, one can understand its physical reality, i.e. whether it is concave, convex or boundary, from the partial drawing of the junctions to which the line belongs. For the remaining lines, one can at least determine that the line is either concave or a boundary.

Thus vanishing points place restrictions on junction labeling of perspective drawings, and provide unique labeling for some junctions independently of the surrounding junctions. Consequently, the cost of the search for correct labeling can be reduced.

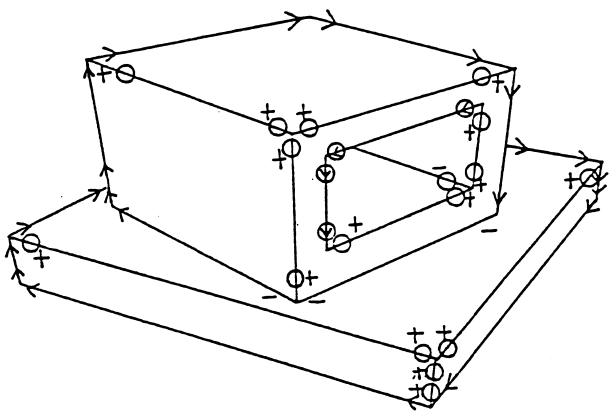


Figure 17. Scene labeling with vanishing points.

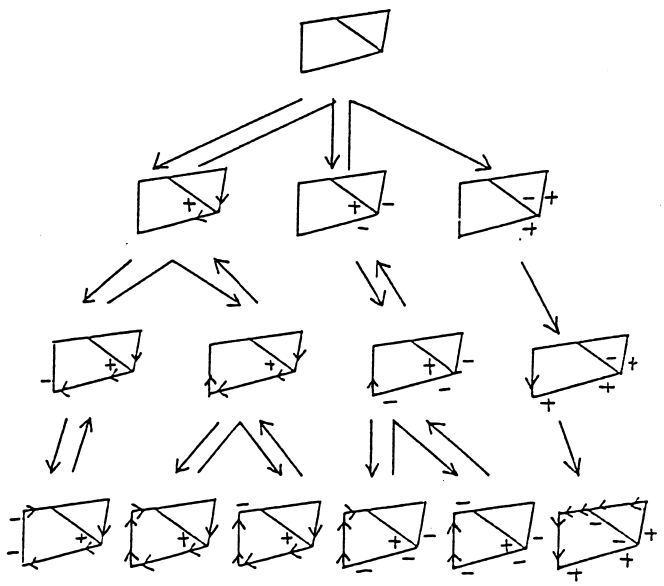


Figure 18. Search tree for labeling without vanishing points.

Another example of three-dimensional shape determination by vanishing point is shown in Figure 19, where only a part of a drawing is given. There is a human visual phenomenon such that each vertex is perceived alternately as convex and as concave. Conventional labeling rules provide these vertices two possible interpretations from the imperfect drawing. Vanishing points permit only a unique interpretation. If there are vanishing points in the direction given in Figure 19(b), one can tell the three-dimensional shape of the vertices; the vertex C_a is convex and C_b is concave.

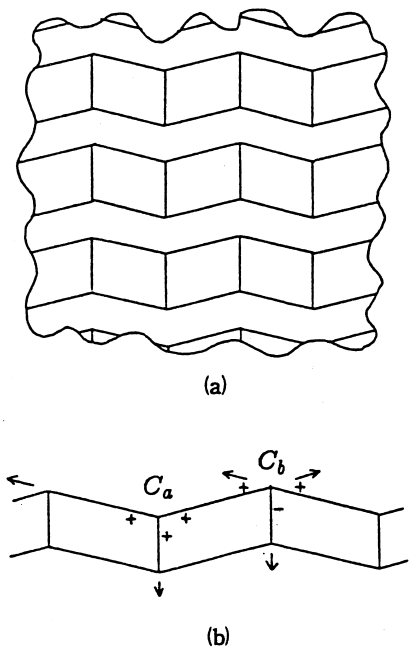


Figure 19. Unique labeling for a partial drawing.

3.6 Conclusion

In this chapter, we have attempted to determine three-dimensional shape of vertices of trihedral solids from perspective drawings. In order to label trihedral vertices without vanishing points, an appropriate label should be assigned taking into account the coherency of the adjacent junctions. If plural labels are allowed for a line, the process involves a search of a tree of possible labels.

This chapter introduced the use of vanishing points in the understanding of line drawings. It was seen that vanishing points add constraints on the three-dimensional shape of vertices, and that the number of possible labels is reduced independently of the surrounding junctions. In particular, for any convex edge one can make a unique labeling, and even for an edge which causes an ambiguity, one can determine that it is either concave or boundary and one can limit the set of labels.

Examples of labeling were given to show that the cost of the search for correct labeling can be reduced by vanishing points.

Though this discussion is applicable only to a mutually orthogonal trihedral vertex, the results of this study will help in reconstructing spatial information of an object in outdoor scenes. This is discussed later in Section 4.4.

CHAPTER 4

SHAPE RECONSTRUCTION OF OBJECTS IN OUTDOOR SCENES

4.1 Introduction

This chapter presents a process that reconstructs the frontal view of a building from its perspective image [23] as a practical application of the three-dimensional shape reconstruction discussed in Chapter 2. It is assumed that the building is a rectangular prism and that the locations of vanishing points of the frontal surface are given. We use the locations of the vanishing points as cues in calculating the distance and shape of the object. Additionally, by using the constraints imposed by vanishing points discussed in Chapter 3, we show how we can recognize the three-dimensional shape of vertices which appear in an outdoor scene.

4.2 Reconstruction of the Frontal View of an Object from Its Perspective Image

4.2.1 Assumptions

In our analysis, we will deal with pictures and buildings under the following assumptions:

(I) The building is a rectangular prism.

(II) The building is depicted in one-point perspective.

(III) The position of the vanishing point for the object plane is *a priori* given.

4.2.2 Determination of the Object Plane Using Constraints at Vertices

We can reconstruct a frontal view of the building using equation (10) or (11) in Chapter 2 with the parameter a_0. When a_0 is not given, although one cannot determine the real size of the object plane, still one can reconstruct its shape using the value d/a_0.

From the one-point perspective assumption, it is seen that one can use equation (10) to reconstruct the shape in the horizontal direction and equation (11) in the vertical direction.

To reconstruct the frontal view, it is necessary to determine the area of the object plane in the picture. That is, one has to specify the four corners of the frontal plane of the building. A feature of our method is to make use of the constraints on the four corners regarding their shape and position in the picture. The vanishing point plays an important role in those constraints.

4.2.2.1 Constraints on Shape. Under the one-point assumption, two of the three edges of a trihedral vertex are parallel to the image plane and one is perpendicular. Here it is seen that the vanishing point adds the following constraint on the shape of the four corners in the picture.

Constraints on shape: The vanishing point lies on an extension of one of the lines at the corner.

The edges which satisfy this condition are limited. This constraint leads us to the introduction of an adaptive template to detect the four corners. The template changes its shape according to the location in the picture (Fig-

ure 20(b)). One should note that one of the segments is always in the direction of the vanishing point.

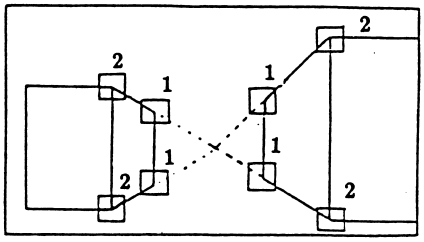
To make the corner detection efficient, we explore a method which we call a minus-image-added template. As shown in Figure 20(c), negative weights are given to those points that are symmetric with respect to the lines of the corner. With this weighting, the output of the corner detection template is emphasized at the specified corners, and de-emphasized at points along lines between the corners.

In practice, several points are extracted from a picture by this template matching. These points are declared as the candidates for the corners. We must then decide which candidates represent the real four corners. Here, we again use constraints imposed by vanishing points.

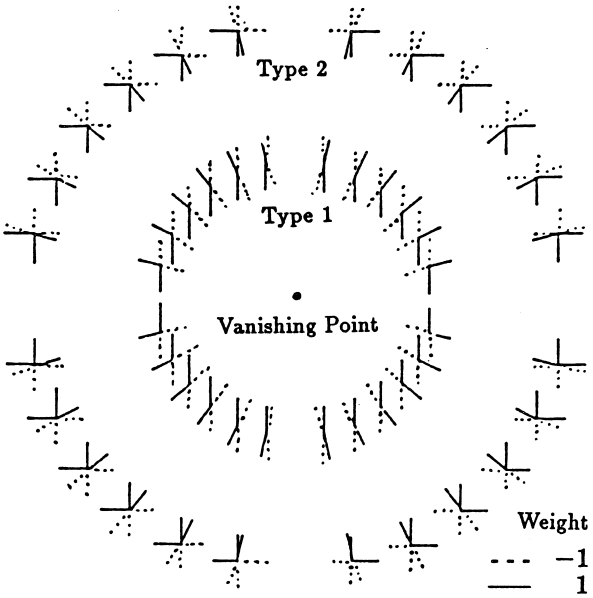
4.2.2.2 Constraints on Position. We divide the four corners of interest into two types depending on their distance; two corners in the far distance are type 1, and two corners in the near distance are type 2 (Figure 20(a)). Two types of corners must satisfy the following conditions on their positions in a picture.

Constraints on position: (I) The vanishing point is on the line which is defined by a type-1 corner and a type-2 corner. (II) A type-1 corner is in the nearer distance from the vanishing point than a type-2.

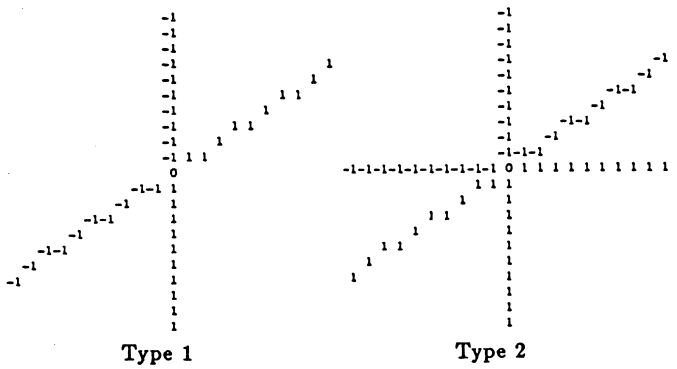
The object plane representing the frontal wall of the building will be obtained based on the constraints described above. Once the plane is determined, the three-dimensional position of a point is calculated using equation (10) and the frontal view will be reconstructed.



(a)



(b)



(c)

Figure 20. Vanishing-point-oriented adaptive template.

4.3 The Reconstruction Process

Based on the discussions in the previous section, we find that the process that reconstructs the frontal view consists of three major phases: (1) the vanishing point detection, (2) the determination of the object plane, and (3) calculation of the location, and interpolation for the frontal view display. This section describes algorithms with the aid of an example which is an application to an outdoor scene. The input picture was taken on our campus, digitized into 256×200 points and quantized into 256 levels.

4.3.1 Vanishing Point Extraction

A single vanishing point is generated on the horizon by the object plane under the one-point perspective presumption described in 4.2.1. We extract the vanishing point from a picture using a method such as that presented in Chapter 2. Let $V = (x_v, y_v)$ be its location.

4.3.2 Determination of the Object Plane

Constraints on shape and position of the four corners imposed by a vanishing point, which we use to determine the object plane, are effective knowledge when we deal with shape patterns, rather than a density distribution. That is, the algorithms are to be applied to line patterns.

4.3.2.1 Thinning the Edge Image. A thinned edge image is generated by spatially differentiating the input image. In our experiment, we use the Sobel operator[11] to obtain an edge image, and retain the upper 10% of the total elements to feed to the Tamura algorithm[21], which provides a

thinned binary image, $B(x,y)$.

4.3.2.2 Search for Candidate Corners. In order to search for candidates for the corners, we use adaptive templates, M_1 for type-1 corners and M_2 for type-2 corners, whose shapes are based on the constraint on corner shape discussed in 4.2.2.1. In our experiment, the size of the template is set to 21×21 as shown in Figure 20(c).

We move the template around the thinned edge image and compute the cross-correlation between the template and the image:

$$\sigma_i(x,y) = \sum_{(a,b)} B(a,b)M_i(a-x, b-y), \quad (i=1,2) \tag{22}$$

where $(a,b) \in \{(a,b)|\ |a-x| \leq 10,\ |b-y| \leq 10\}$. If the cross-correlation at a point is greater than a threshold σ_0^i defined by the corner type, the point is declared as a candidate corner, C_1 or C_2:

$$C_i = (x_i, y_i), \quad \text{if } \sigma_i > \sigma_0^i \quad (i=1,2), \tag{23}$$

where $\sigma_0^1 = 15$ and $\sigma_0^2 = 17$.

4.3.2.3 Search for the Upper and Lower Sides. Using the constraint on positions of the corner given in 4.2.2.2, we determine from the candidates the real corners of the building wall. For each candidate of type-1, $C_1(x_1, y_1)$, we check whether or not any type-2 candidate $C_2(x_2, y_2)$ satisfies the conditions (24) in terms of their positions.

$$\left.\begin{array}{ll} [\mathrm{I}] & (x_2 - x_1)(y_v - y_1) = (y_2 - y_1)(x_v - x_1) \\ [\mathrm{II}] & \|C_1 - V\| < \|C_2 - V\| \end{array}\right\}, \tag{24}$$

These conditions limit the number of possible combinations of type-1

and type-2 corners. A combination $\overline{C_1C_2}$ defines a horizontal side of the object plane.

4.3.2.4 Search for the Object Plane. We need to decide which sides form a pair of the upper and lower side of the plane. The horizontal coordinates of the two sides must be the same because of the one-point perspective presumption. That is, we can determine the pair of the sides by finding two sides, $\overline{C_1(x_1,y_1)C_2(x_2,y_2)}$ and $\overline{C_1'(x_1',y_1')C_2'(x_2',y_2')}$ such that $x_1 = x_1'$ and $x_2 = x_2'$.

If the lower side is occluded in the picture by something like foliage, this technique will fail to find the side. In such a case, one alternative to determining the side might be to specify its location manually. We investigate, however, an automatic way for determining the plane by modifying the technique defined by (22)-(24) as follows.

A counterpart $\overline{C_1'(x_1',y_1')C_2'(x_2',y_2')}$ of a side $\overline{C_1(x_1,y_1)C_2(x_2,y_2)}$ is determined such that two points C_1' and C_2' satisfy

$$\max_{y_1',y_2'} \{\sigma_1(x_1',y_1') + \sigma_2(x_2',y_2')\} , \tag{25}$$

where $(x_2' - x_1')(y_v - y_1') = (y_2' - y_1')(x_v - v_1')$, $x_1' = x_1$, $x_2' = x_2$.

That means C_1' and C_2' do not have to meet the condition (23) but are given such that the sum of their output of the adaptive template is the maximum under the conditions that their horizontal coordinates are the same and that their positions satisfy the constraint (II) in 4.2.2.2 (Figure 21).

One would expect that the output of the template is likely to have a high value at a border between the foliage and the building or the ground. As for

the picture shown as an example, two points at the border between the foliage and the ground satisfy condition (25). Then the object plane includes a side defined by those two points. However, since we can apply this technique only when the building is assumed to be a rectangular prism, it cannot be applied in the case as the border of the foliage is not parallel to the building.

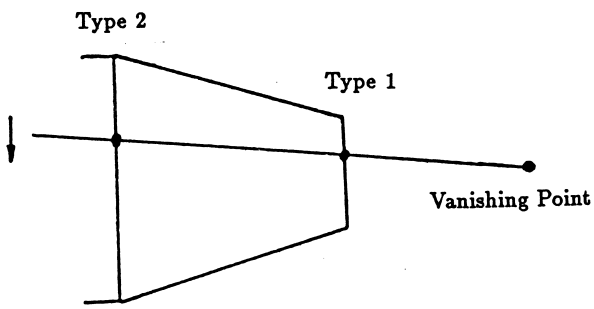


Figure 21. Plane determination.

4.3.3 Reconstruction of the Frontal View

Now we can calculate the three-dimensional location of a point on the object plane and assign gray levels to the points of the reconstructed image by interpolating between the gray levels of the input picture.

4.3.3.1 Calculation of the Location. Given the location of the vanishing point in Section 4.3.1 and the locations of the four corners in Section 4.3.2, we can compute the horizontal and vertical coordinates in three-dimensional space for each point on the object plane by equations (10) and (11).

4.3.3.2 Interpolation for the Display. To make the frontal view, we assign the gray level of the input picture to the point at the calculated coordinates. Only some of the points of the reconstructed picture correspond to coordinates in the input picture, and the number of such points decreases as the vanishing point is approached.

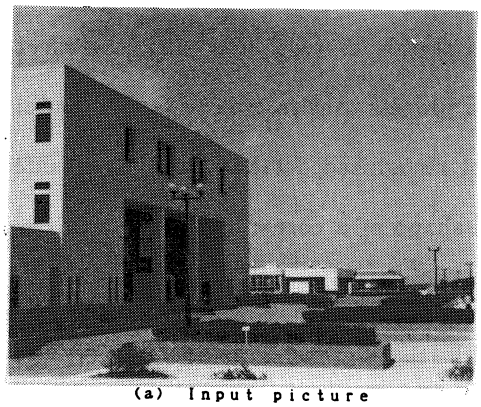
We fit gray levels to the points between given points to make a connected picture. In our experiment, we use a linear interpolation both in the horizontal and vertical directions.

4.3.4 Result

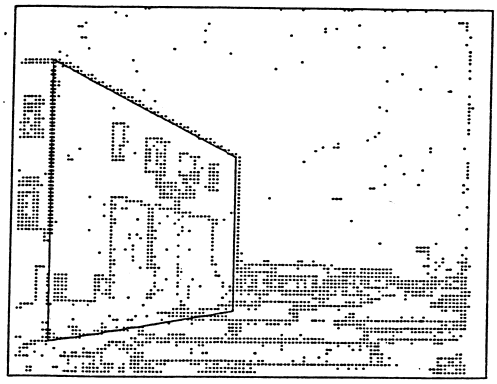
Figure 22 presents an example in which we apply the procedure to an outdoor scene, which is regarded as one-point perspective. The result (e) might give the impression that it is horizontally elongated, judging from the input picture (a). For comparison, we show the real frontal view in (f). The ratio of the height and breadth is 1 : 2.8 in the picture (f) and 1 : 2.6 in the reconstructed

picture (e).

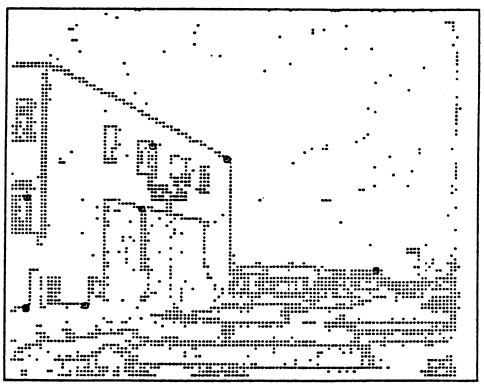
Windows, especially at a distance, are distorted in the resulting picture because only a few points correspond to the original data. In this example, a width of a pixel in the original picture corresponds to two pixel widths in the reconstructed picture at the nearest point, but seven pixel widths at the farthest point and accordingly the accuracy becomes low.



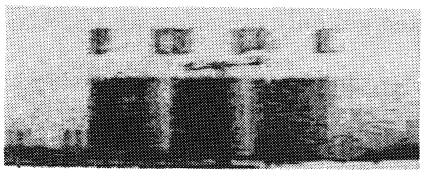
(a) Input picture



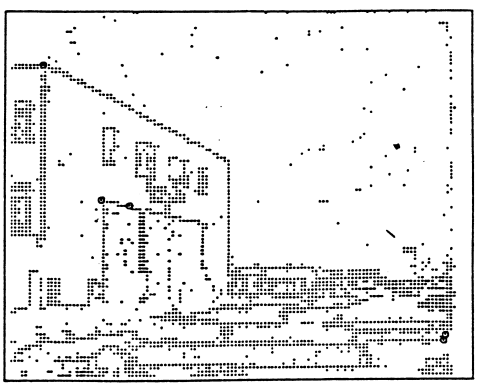
(d) Plane determination



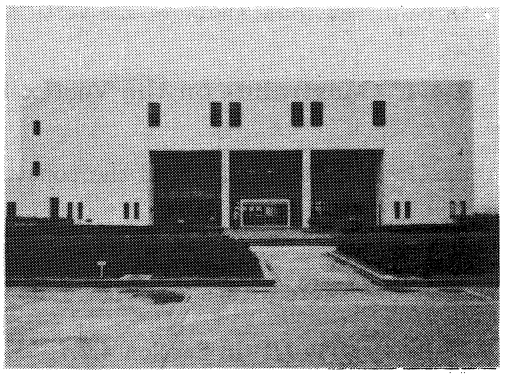
(b) Candidate type 1



(e) Reconstructed frontal view



(c) Candidate type 2



(f) Real frontal view

Figure 22. Reconstruction of frontal view.

4.4 Determination of the Three–Dimensional Shape of Corners

We discussed shape reconstruction of objects which are under the effect of perspective distortion, and presented a process that reconstructs the frontal view of a building. But still we have not recognized detailed structure inside the plane such as the three-dimensional shape of a corner, i.e., whether it is concave or convex.

It might be possible to determine the three-dimensional shape of objects by using line labeling techniques for understanding line drawings. However, it would be necessary to obtain a perfect line drawing of the object and we often get only imperfect and partial line drawings. Even in such a case, we can determine from a partial drawing whether a corner is concave or convex if we use the vanishing point and the adaptive template.

This section shows that the adaptive template, which was used to detect vertices of an object in the last section, facilitates the determination of the three-dimensional shape of the vertices.

4.4.1 Constraints on the Vertex Shape

Chapter 3 discussed the constraints on the three-dimensional shape of objects that are made of mutually orthogonal trihedral vertices in perspective drawings. Here we use the constraints in the case of one-point perspective which is assumption (II) in Section 4.2.1.

From the discussion in Chapter 3, it is seen that the location of the vanishing point and the shape of the adaptive template impose the constraints on the vertex shape. For example, if a V–junction as shown in Figure 23 has a

vanishing point in the direction of "↑", it is found that an edge is convex(+) and another one is boundary(>); and the junction can be interpreted as a concave vertex.

Thus we can determine three-dimensional shapes of corners of the building by classifying the junction shape and the vanishing point location. That means we can do more than reconstruct a frontal view — though only partially at this stage, we can analyze three-dimensional structure.

4.4.2 Determination of Detailed Structure

As shown in Figure 24 we applied the above discussion to the real picture. In the process for reconstructing the frontal view, two types of templates were sufficient for detecting specific convex vertices which formed the four corners at the boundary of the building. Here, in order to detect concave corners inside the frontal wall, we use a different type of adaptive template in a fashion similar to that described in Section 4.3.2.2. The resulting concave corners are marked with "○" in Figure 24(a).

Additionally, we can put the boundary mark "<" if we know a boundary beforehand. And two corners at the top of the frontal wall turn out to be convex vertices; from the list of W–junctions for the top left corner and from the list of V–junctions for the top right corner, both in the catalog of one-point perspective vertices in Figure 15. We illustrate the result of determination of the three-dimensional shape of corners in Figure 24(b).

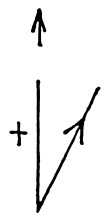
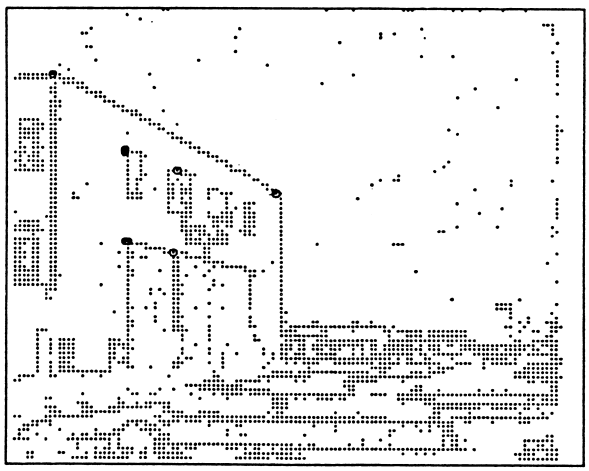
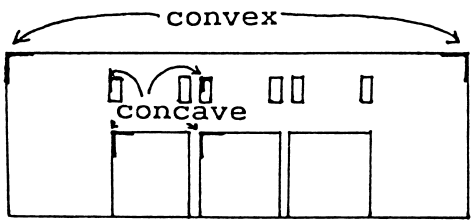


Figure 23. A concave vertex and vanishing point.



(a) Vertex determination



(b) Drawing of result(a)

Figure 24. Determination of 3-D structure.

We succeeded in determining the shape of only six corners in this experiment. However, a feature of this method is that we can determine whether a corner is concave or convex even when a perfect drawing is not available. On the other hand, by using a conventional line labeling techniques, we cannot determine it even partially because line drawings are usually made of discontinuous line pieces.

The result proves that the vanishing point enables reconstruction of the size and shape of objects from their appearance in two-dimensional images. Furthermore, three-dimensional structure of an object is also analyzed from its partial drawing.

4.5 Conclusion

In an attempt to analyze three-dimensional structure in an outdoor scene, we obtained some three-dimensional information such as the distance, size and shape of objects, and we reconstructed an object plane which was distorted perspectively in a picture. Thus we have clarified the analysis of spatial information which has been difficult to tackle with simple techniques such as the Affine transformation.

In the procedure of the reconstruction, we made clear the constraints on the shape and position of corners imposed by vanishing points, and introduced an adaptive template based on the constraints. The template facilitated detection of vertices with the particular shape and determination of the three-dimensional shape from a partial drawing. Though only partially at this stage, we can analyze spatial structure of objects. And we consider that shape deter-

mination by the adaptive templates rationalizes the method for searching the object plane.

We still have to solve some problems. If some obstacles, such as foliage or a lamp post as in our experiment, are in front of the object plane, they will be dealt with as if they were on the object plane though they have no effect on the shape of the remaining part on the object plane. It might be possible to solve this by introducing an interactive method or a higher level processing such as object understanding. In addition to that, making our technique applicable to wider range of objects is a topic that must be studied further.

CHAPTER 5

THREE–DIMENSIONAL SHAPE RECONSTRUCTION FROM ULTRASONO–TOMOGRAMS

5.1 Introduction

So far we have been concerned with the analysis of three-dimensional information depicted in two-dimensional pictures. In this chapter, another area of research, the display of the three-dimensional information, is attacked by constructing a system that produces stereograms and reconstructs cross-section images of a human heart.

By using an ultrasonic imaging system, we can obtain a tomogram of a section of the human body, which is called a B-mode image [25]. The section is visualized as if we were looking at a slice of the body tissue. In the field of cardiology, to understand the shapes of cardiac chambers and the spatial relations of neighboring echoes by a number of tomograms requires much practical experience. Cardiologists take many tomograms and build up a mental three-dimensional picture of a chamber by repeatedly observing many individual tomograms. This procedure is cumbersome and time-consuming even for experienced cardiologists, and its objectivity is poor.

By the B-mode method, we can display a plane which the ultrasonic beam scans; we cannot display a plane which is inclined obliquely to it, however. Furthermore, the greater part of the beam is reflected at the boundary of the tissue and air: we cannot visualize internal body structures beyond the air. The

cardiac B-mode tomograms have therefore been limited by the proximity of the lungs to the heart. But various depth or various angle tomograms are required in bedside diagnosis.

Therefore, a system which can provide an objective three-dimensional display and can synthesize a tomogram at a desire plane has long been hoped for.

5.2 Binocular Stereoscopic Dispay System for Echocardiography

In this section, we propose a system which automatically produces a three-dimensional image of the heart from a two-dimensional representations, i.e., tomograms. The principle of binocular stereoscopy, which is of considerable use in three-dimensional display, see [26] for example, is applied to our system. The boundary of the heart chamber is automatically specified from tomograms by a computer, and its binocular images are displayed on a CRT with plural boundary-lined tomograms.

5.2.1 Binocular Stereoscopy

The perception of depth by humans is achieved by orienting the eyes so that the image of a viewed point falls on corresponding points of the individual retinas [11]. The principle of retinal disparity is sketched in Figure 25. Let $p_{\mathcal{R}}$ with coordinates $(x_{\mathcal{R}}, y)$ and $p_{\mathcal{L}}$ with coordinates $(x_{\mathcal{L}}, y)$ define the right and left stereoscopic point representations of a point P, respectively. The parallax of $p_{\mathcal{R}}$ and $p_{\mathcal{L}}$ produces depth perception. Let us define Z as the horizontal distance from the point P to the eyes. Then the parallax is

$$x_{\mathcal{L}} - x_{\mathcal{R}} = \frac{K}{Z} , \tag{26}$$

where K is a constant.

Conversely, a three-dimensional display can be achieved by presenting to each eye separate two-dimensional images which have binocular parallactic shifts. Our system gives a relief image of the heart by presenting binocular images obtained from conventional ultrasono-tomograms, and is called a binocular stereoscopic (S-mode) display system for echocardiography.

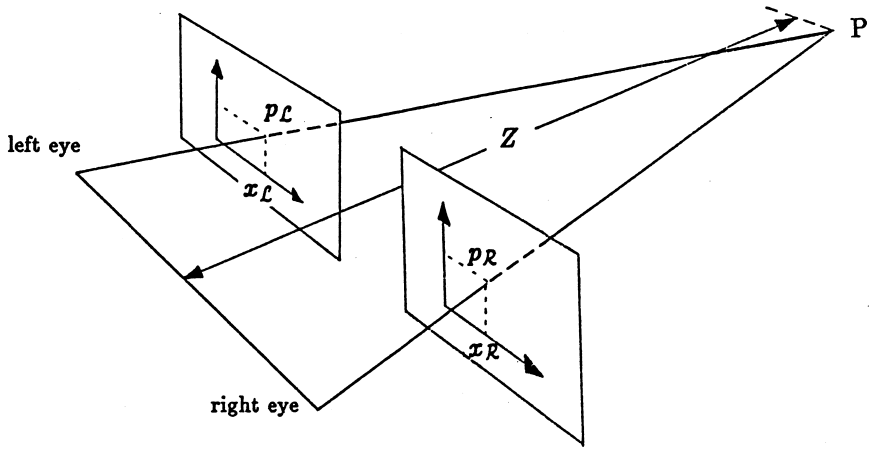


Figure 25. Binocular stereoscopy.

5.2.2 Processing of Tomograms

5.2.2.1 Input of Tomograms. Tomograms are displayed on a bistable storage type CRT which is attached to an ultrasonic diagnosis device (Aloka SSD-10) with frequency 2.25 MHz, 3 mm resolution, 30 mmϕ, and a 200 mm focused concave transducer.

Figure 26 shows an example of the tomograms obtained. These longitudinal tomograms (B-mode image) are obtained at regular intervals of 5 mm. In order to obtain these phase-selected tomograms in a required cardiac phase, the ultrasonic apparatus operates in synchronization with the cardiac cycle.

By means of a flying-spot scanner (KOWA-OS701), the tomograms are digitized to 105×80 picture elements with a linear quantization range of 0-127. Let $g(x,y)$, $1 \leq x \leq 105$; $1 \leq y \leq 80$ be a function that represents a gray-level picture.

5.2.2.2 Smoothing. We use manual scanning that sometimes causes defects in the scan beam (see Figure 26). In order to suppress such artifacts, we compute the spatial average intensity, $g^a(x,y)$, from $g(x,y)$ and its four nearest neighbors by

$$g^a(x,y) = \frac{1}{2}g(x,y) + \frac{1}{8}\sum_{(\mu,\nu)} g(x-\mu, y-\nu) \qquad (27)$$
$$(\mu,\nu) \in \{(0,1),(1,0),(0,-1),(-1,0)\} \ .$$

5.2.2.3 Bilevel Representation. Figure 27 shows a gray level histogram of a tomogram. Because of the binary characteristic of the tomograph's storage type CRT, the histogram shows two peaks, corresponding to the bright areas and dark areas of the picture.

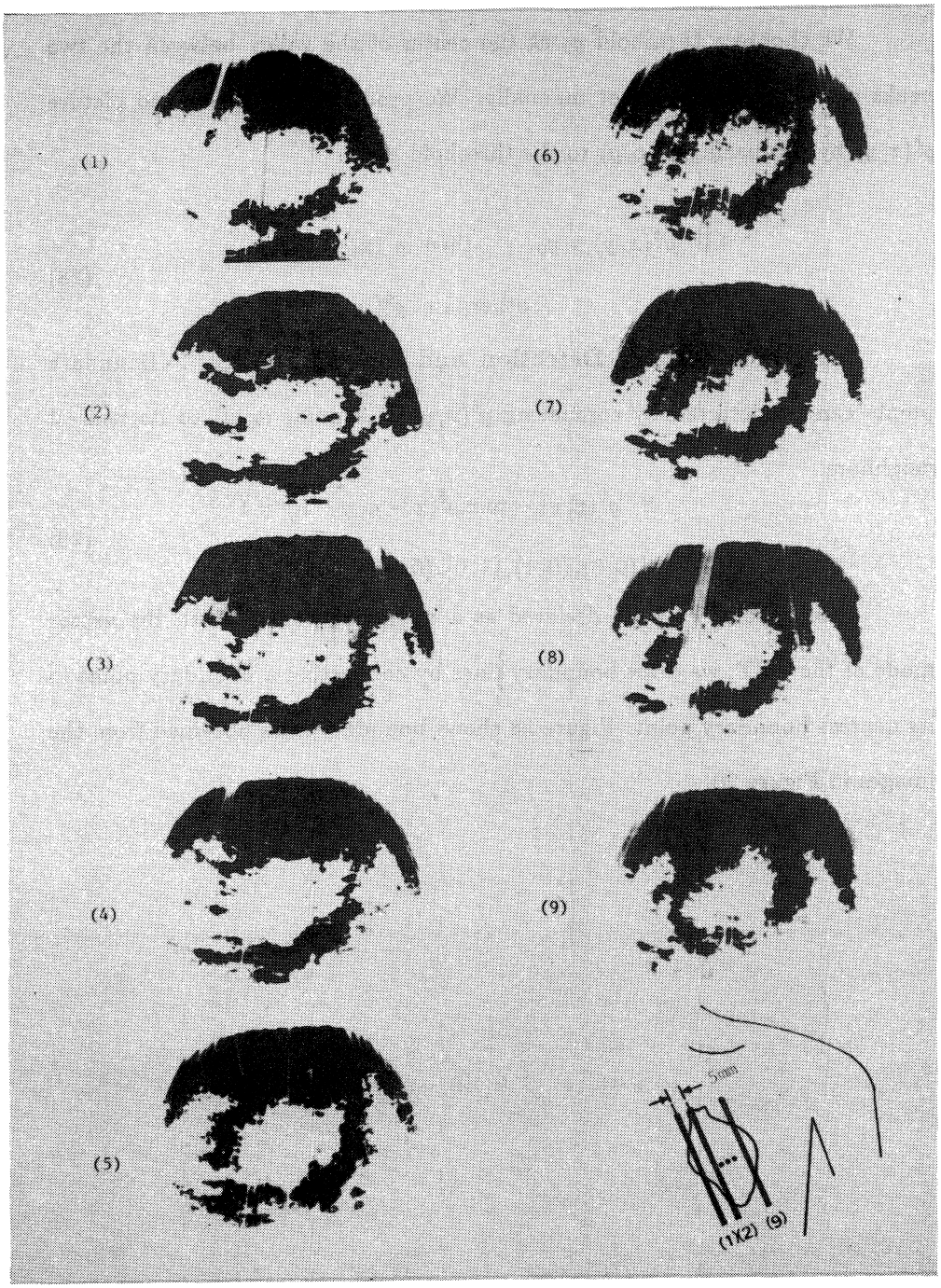


Figure 26. Ultrasono-tomograms (end-diastole; B-mode negative image).

We choose a threshold g_0 at the center of the valley between the two peaks either automatically or manually. We produce a binary-valued picture $g^b(x,y)$ by comparing $g^a(x,y)$ to the threshold g_0:

$$\begin{aligned} \text{If } \; g^a(x,y) > g_0, \qquad & \text{then } \; g^b(x,y) = 1, \\ & \text{otherwise } \; g^b(x,y) = 0. \end{aligned} \tag{28}$$

5.2.2.4 Boundary Detection and Line Drawing. Boundary points can be detected by comparing $g^b(x,y)$ of a point (x,y) to its nearest neighbors:

$$\begin{aligned} &\text{If } \; g^b(x,y) - \min g^b(x-\mu, y-\nu) = 1 \\ &(\mu,\nu) \in \{(0,0),(0,1),(1,0),(0,-1),(-1,0)\} \; , \end{aligned} \tag{29}$$

then a point (x,y) is declared as a boundary point. With the vector mode of the CRT, we draw boundary lines by connecting a boundary point to its nearest boundary point. Figure 28 shows boundary lines obtained from the images in Figure 26.

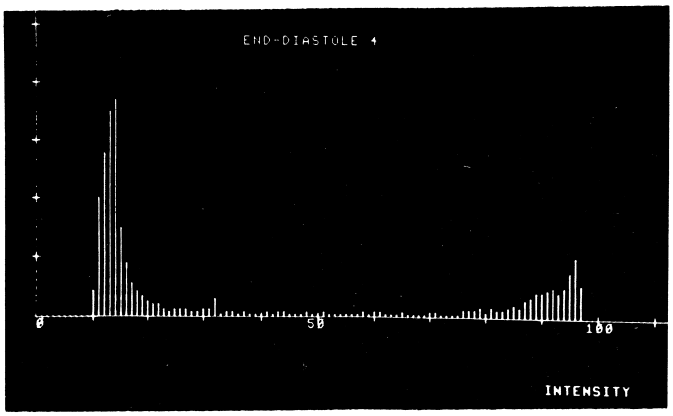


Figure 27. Intensity histogram of Figure 26(4).

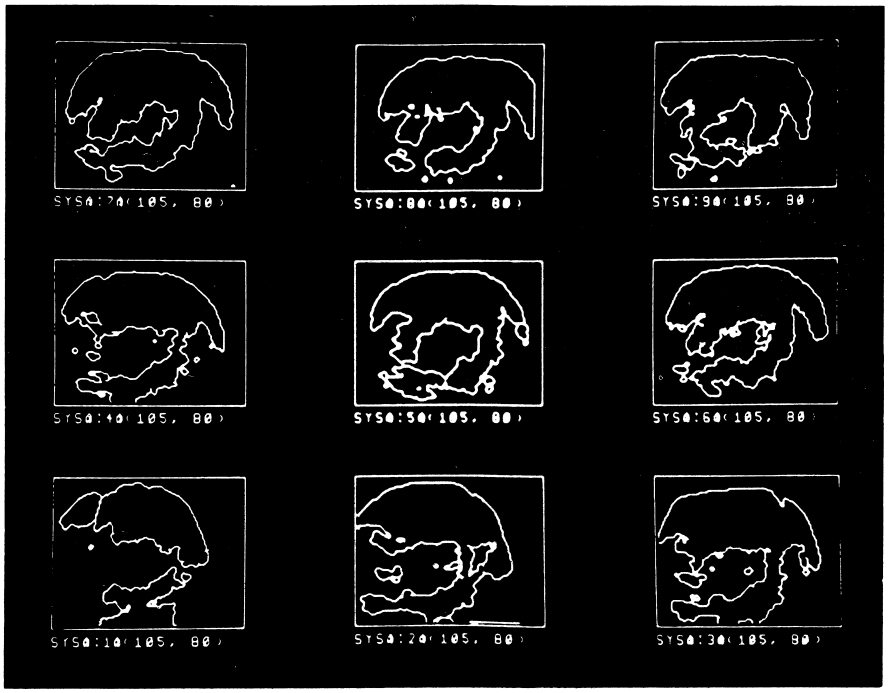


Figure 28. Boundary lines for tomograms in Figure 26.

5.2.3 S–Mode Display

Figure 30 shows an example of binocular images of the living human heart. Boundary lines of some selected sections or the entire collection of a series of the heart sections are reconstructed in overlaid fashion. Each boundary line is drawn with the corresponding parallactic shift given by equation (26).

Stereoscopic viewing of the heart is obtained by converging the right and left views using a simple stereoviewer mounted on the CRT face (Figure 29). If necessary, registration of serial sections is performed by pointing at the corresponding points with a joystick. Trimming is also possible when the area of interest is set by the joystick, only the boundary lines in this area are displayed on the CRT.

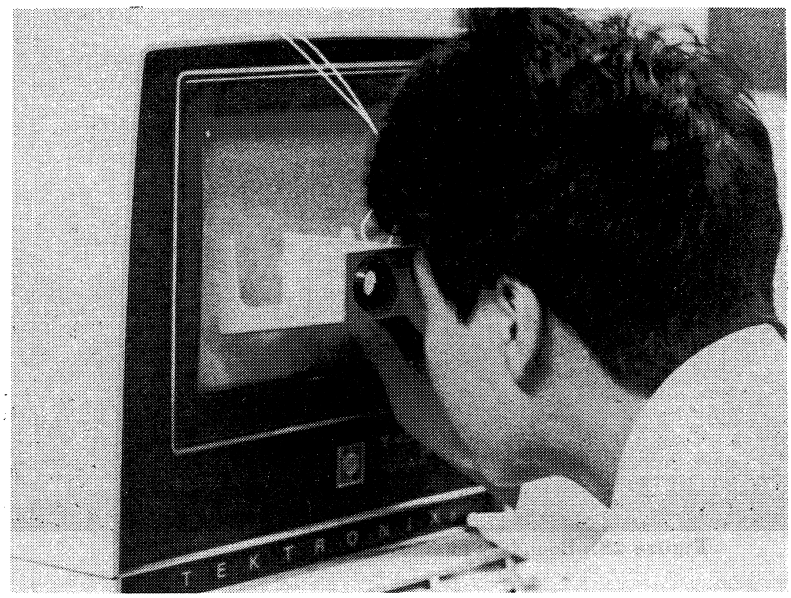
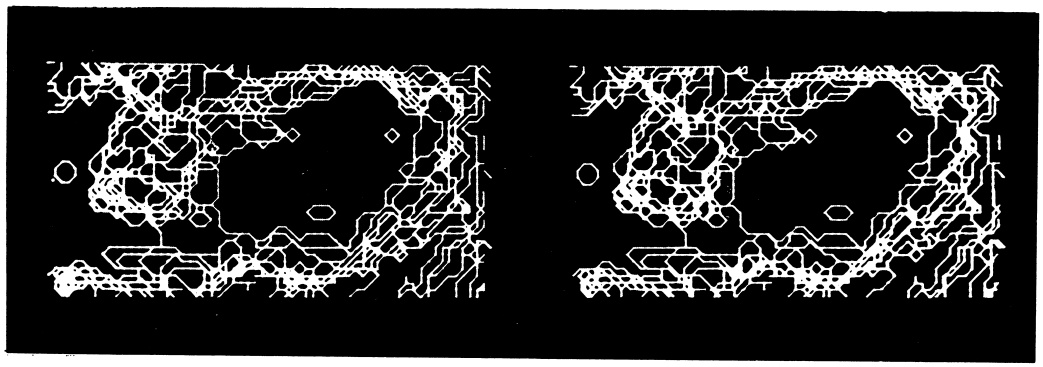


Figure 29. Stereoviewer.

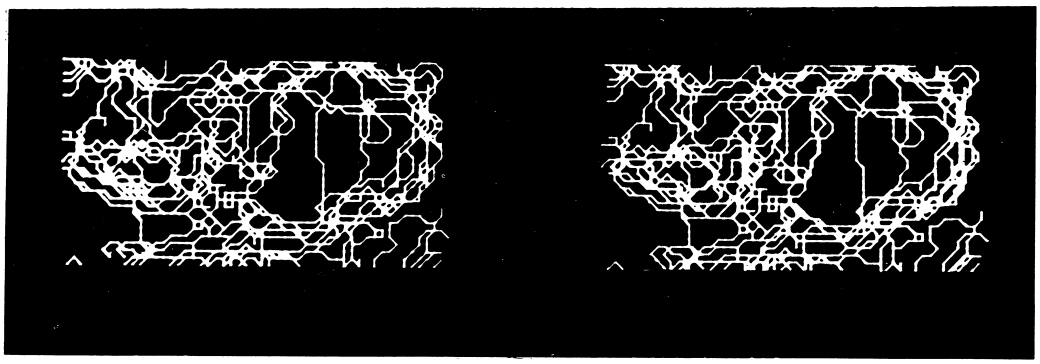
5.2.4 Results

The stereogram (S-mode tomogram) generated from the nine tomograms recorded at end-diastole along the long axis of the heart is shown in Figure 30(a). The nearest boundary line corresponds to the tomogram of Figure 26(1). Therefore, Figure 30(a) shows a view from the center of the body to the left side. The left atrium, the left ventricle and the aorta are visualized in depth.

Another stereogram shown in Figure 30(b) was generated from the tomograms recorded in the same planes as those of Figure 30(a), but in late systole. This presents a more complex configuration than that in late diastole. See [27] for the details of organic structure observed in Figure 30(a) and (b).



(a)



(b)

Figure 30. Binocular stereoscopic display. (a) stereogram composed of nine boundary lines of Figure 26 in end-diastole, and (b) in end-systole.

5.3 Tomogram Reconstruction in a Desired Plane

The last section described a binocular stereoscopic (S-mode) display system. Once the ultrasonic images are digitized and stored in a computer, it is a simple matter to synthesize a new image. In this section, a technique for synthesizing an image from tomograms is presented as an application of the S-mode display system.

There have been some attempts to reconstruct three-dimensional images from their serial sections. Robinson provided a cross-section which is perpendicular to the transverse tomograms of a pregnant uterus [28]. The intersection of each transverse tomogram and the specified section gives line data, and he generated a sectional image by interpolating the line data. Huang and Ledley reconstructed a three-dimensional image of a spinal column in the neck region from its consecutive transverse axial sections [29]. Glenn et al. reconstructed coronal and sagittal planes from computerized axial tomographic (CT) data in conventional transverse planes [30].

They mentioned the practicability of reconstructing an oblique view but did not give any results. Using those methods, considerable experience seems necessary to give the equation of an oblique plane with respect to a fixed coordinate system. Our S-mode display system, however, gives selected border-lined tomograms or an entire collection in an overlaid fashion as mentioned in Section 5.2.3. From this three-dimensional image we can easily determine the spatial relationships of structures and select a proper cross-section for display. In this section an experimental application of our system to the heart and a method for selecting a proper cross-section are described.

5.3.1 C–Mode Display

First, let us consider a general principle for a C-mode tomogram. Assume we have planes (tomograms) $T_1, T_2, \ldots, T_n$ in three-dimensional space and that we want to see a sectional image (C-mode tomogram) on a section plane S (Figure 31). We represent the planes $T_1, T_2, \ldots, T_n$ and S as follows:

$$T_i : \vec{r}\vec{N}_i = t_i \qquad i = 1, 2, \ldots, n \tag{30}$$

$$S : \vec{r}\vec{N} = t, \tag{31}$$

where $\vec{N}_i$ and $\vec{N}$ are unit normal vectors perpendicular to the plane T_i and S, respectively. The t_i and t are the distances between the corresponding plane and the origin. In this case we can choose the origin arbitrarily. The vector $\vec{r}$ is a variable representing each plane. The solution set of (30) and (31) is a line C_i, which is the intersection of the planes T_i and S.

Values of the image function of the C-mode tomogram on S are given only at the intersection C_i of T_i and S. The resolution power at our frequency (2.25 MHz) is approximately 3 mm. Thus, it is necessary to cover S by several mm neighbors of the intersection C_i in order to obtain the C-mode tomogram.

Let us obtain the distribution of the intersections of the plane S. Denote the projection operator onto S and $\vec{N}$ by π_S and $\pi_{\vec{N}}$, respectively. Since $\vec{N} \perp S$,

$$\begin{aligned} \vec{r} &= \pi_S(\vec{r}) + \pi_{\vec{N}}(\vec{r}) \ , \\ \vec{N}_i &= \pi_S(\vec{N}_i) + \pi_{\vec{N}}(\vec{N}_i) \ . \end{aligned} \tag{32}$$

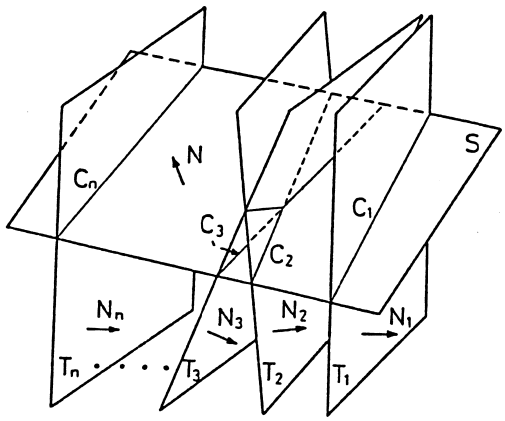


Figure 31. Section plane S.

Then (30) becomes

$$\begin{aligned}
[\pi_S(\vec{r}) + \pi_{\vec{N}}(\vec{r})][\pi_S(\vec{N}_i) + \pi_{\vec{N}}(\vec{N}_i)] &= t_i \\
\pi_S(\vec{r})\pi_S(\vec{N}_i) + \pi_{\vec{N}}(\vec{r})\pi_{\vec{N}}(\vec{N}_i) &= t_i \\
\pi_S(\vec{r})\pi_S(\vec{N}_i) &= t_i - \pi_{\vec{N}}(\vec{r})\pi_{\vec{N}}(\vec{N}_i) \\
&= t_i' \ ,
\end{aligned} \tag{33}$$

where

$$t_i' = t_i - \pi_{\vec{N}}(\vec{r})\pi_{\vec{N}}(\vec{N}_i) \ ,$$

$\pi_S(\vec{r})$ is a vector variable representing the intersection C_i in S; $\pi_S(\vec{N}_i)$, $\pi_{\vec{N}}(\vec{N}_i)$ are constant vectors that are determined by the arrangement of T_i and S; and $\pi_{\vec{N}}(\vec{r})$ is a perpendicular vector from the origin of the three-dimensional space to S. Letting $\pi_{\vec{N}}(\vec{r})$ be the origin of S, $\pi_S(\vec{r})$ in (33) expresses the intersection C_i in S.

For ease of operation and processing, the arrangement of T_i's is usually chosen such that they are all parallel. In this case, $\vec{N}_i$ becomes independent of i. However, there is a case where the parallel set T_i cannot be obtained, namely, when the whole heart cannot be taken only by parallel tomograms because of obstruction by air in the lung. In such a case, some tomograms must be taken obliquely.

In the following, we deal only with the full parallel case. Consider the simplest situation where the T_i's are all parallel and $T_i \perp S$. In this case, since $\vec{N} \perp \vec{N}_i$,

$$\begin{aligned} \pi_{\vec{N}}(\vec{N}_i) &= 0 \\ \pi_S(\vec{N}_i) &= \vec{N}_0 \ , \end{aligned} \tag{34}$$

where

$$\vec{N}_0 = \vec{N}_i \qquad i = 1, 2, \ldots, n \ .$$

Then (33) becomes

$$\pi_S(\vec{r})\vec{N}_0 = t_i, \tag{35}$$

and (35) represents a group of straight lines that are perpendicular to $\vec{N}_0$ with distance t_i from the origin of S which is the projection of the three-dimensional space origin on S.

A plane in three-dimensional space is determined by three points, or a straight line and a point. Therefore, the operator can input the section plane S to the computer by one of the following methods:

(a) from the keyboard by giving the coordinates or gradient of the point, line, or plane.

(b) by looking at the CRT, and indicating the points by a joystick.

In our system, the section plane S is given by the method (b) because of ease of operation. The procedure is as follows:

(i) The case where the section plane S is perpendicular to B-mode tomograms T_i. In this case, S is determined by giving only two points or equivalently a straight line on T_i. Therefore, a selected B-mode tomogram T_i is displayed on the CRT and the operator inputs by the joystick two points which determine the intersection of T_i and S. Each intersection is given by (35). Figure 32(a) shows an example of a straight line specified by the joystick.

(ii) The case where the section plane S is not perpendicular to T_i. In this case, S is determined by indicating another point in some T_j in addition to method (i). Each intersection is given by (33) with $\vec{N}_i = \vec{N}_0$; $i = 1, 2, \ldots, n$.

When the section plane S is given to the computer, the computer collects image data on S. An example of a binary representation of the synthesized C-mode tomogram on the section plane of Figure 32(a) is shown in Figure 32(b). The slice level for displaying C-mode tomogram is automatically or manually chosen.

Usually, we take B-mode tomograms every 5 mm. This is chosen considering the resolution power and the amount of image data. Since the resolution power of right/left direction and that of up/down direction are different, i.e., the length of each pixel is larger that the width, the synthesized C-mode tomogram looks a little bit unnatural especially when the section plane is greatly inclined from the plane parallel to the chest wall.

To get smooth synthesized images under this condition, Robinson had to devise some quite complicated linear interpolation schemes [28] and Glenn

et al. interpolated from thick transverse planes to thin transverse planes by a deconvolution technique [30]. On the other hand, we can get smooth images simply by four-point averaging: first, each pixel in the synthesized C-mode tomogram is divided into some pixels — each of them is nearly square. Then, the C-mode tomogram with divided pixels is smoothed by four-neighbor smoothing with weight 4 at the center and weight one at the four-neighbors.

Our system can also display the C-mode tomogram in density mode. The density mode display of Figure 32(b) is shown in Figure 32(c). The density mode display requires a lot of time because it is an artificial gray-scale created with dots. A better method would use a display controller and a TV monitor.

5.3.2 Results

A synthesized antero-posteriorly oblique C-mode tomogram is shown in Figure 32. The input images consist of nine longitudinal tomograms recorded in end-systole along the long axis of the healthy heart at intervals of 5 mm. The C-mode tomogram is defined by the oblique line on the antero-posterior section (Figure 32(a)). The left ventricle, mitral valve ring, and left atrium during end-systole are presented in the synthesized tomogram.

Figure 33 shows another C-mode tomogram generated from the nine original B-mode tomograms shown in Figure 26. A frontal plane C-mode tomogram is synthesized at a plane specified by a straight line and a point which goes near the bottom of the left ventricle. For a clinical application to a case with atrial septal defect, the reader is referred to [31].

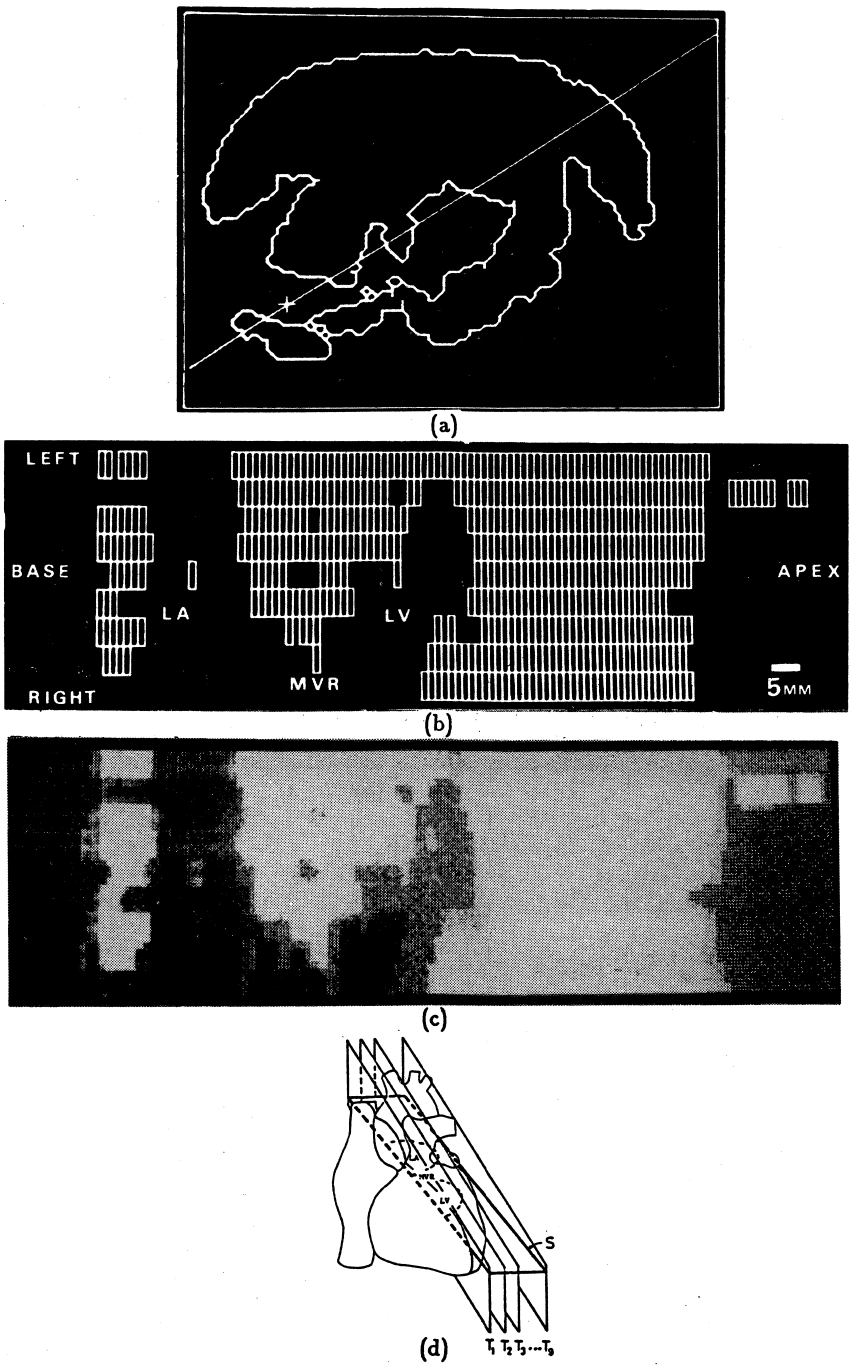


Figure 32. Computer synthesis of cardiac cross-section (end-systole). (a) cross section specified by joystic. (b) synthesized *S*-mode tomogram on section plane of (a). LV = left ventricle, MVR = mitral valve ring, LA = left artium. (c) density mode display. (d) schematic drawing.

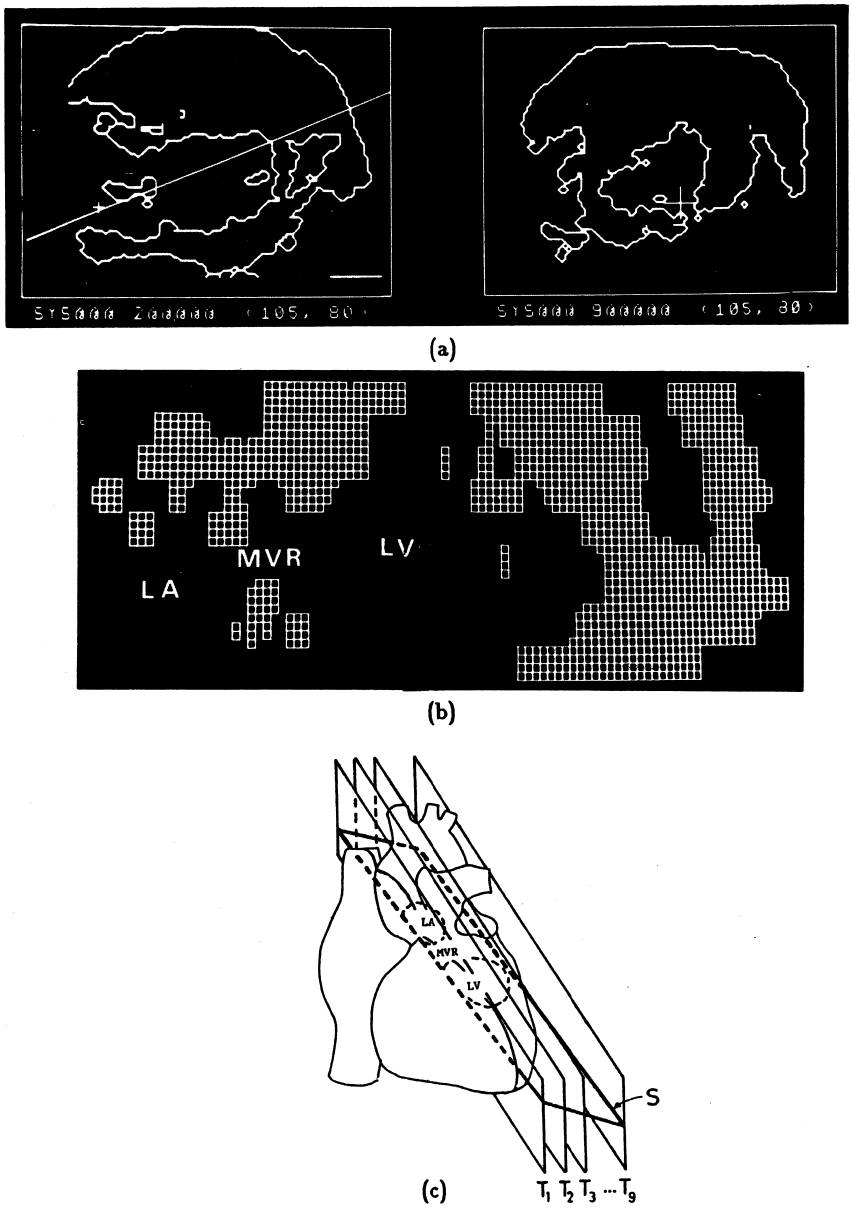


Figure 33. Computer synthesis of cardiac cross-section (end-diastole). (a) longitudinal plane specified by a straight line and a point. (b) *C*-mode tomogram on plane of (a). (c) schematic drawing.

The present reconstruction in desired section is accomplished with original data recorded at intervals of 5 mm. Therefore, these experimental images are gross and may look rather different from conventional echocardiograms, but images can be much improved by reducing the interval between the planes. The focal length of our echocardiograph is 20 cm and the resolution power is 2-3 mm. Therefore, the interval can be reduced to 2-3 mm.

5.4 Conclusion

We have described three-dimensional echocardiographic display based on the principle of binocular stereoscopy, and tomogram reconstruction at an arbitrary angle. Since binocular images are produced on a CRT with plural boundary lines of the heart, cardiologists can see interior structures of the living heart behind the opacities of the tomogram.

Our system aids the user in recognizing the spatial structure of an object and in synthesizing C-mode tomograms at any angle, which are difficult to make without a computer. It will thus become a powerful device for medical diagnosis. So far we have applied the system to the display of internal structures of the human heart. The system is also applicable to the display of three-dimensional structure in other fields, such as industrial problems.

At present, data have been collected off-line between the tomograph and the computer. The system might be improved so as to enable on-line data collection.

CHAPTER 6

CONCLUSIONS

The goal of this thesis is the reconstruction of three-dimensional spatial structure from two-dimensional images by making use of pictorial depth cues. We employ perspective transformation as one of many cues to three-dimensional information about objects in a two-dimensional image. The concept of vanishing points is introduced for the purpose of determining three-dimensional measurements. We can use the location of the vanishing point as a cue in calculating the distance and shape of the object. As an application of the three-dimensional shape reconstruction, a process of reconstructing the frontal view of a building from its perspective image is shown. In addition to the analysis, the display of the three-dimensional shape of an object is explored.

The following is a summary of specific accomplishments:

(1) Development of methods for extracting a vanishing point in natural outdoor scenes.

(2) The use of the constraints on the three-dimensional shape of vertices imposed by vanishing points in perspective drawings.

(3) Development of a means for reconstructing the frontal view of an object from its perspective image.

(4) Development of a system that displays the three-dimensional shape of an object and reconstructs a cross-sectional image at an arbitrary angle.

Possible future improvements are as follows:

(1) Three-dimensional motion analysis. The focus of this thesis has been reconstruction of spatial information from a static image. The vanishing point can be used as a cue to spatial knowledge about a moving object as well [32]. A three-dimensional description of moving objects would be a fundamental function for machine vision systems.

(2) The analysis of the error in the three-dimensional measurements. Since error in locating the vanishing point is inevitable as with any other image process, it is necessary to calculate a range within which the measurement can be considered to be correct [33]. In order to make our technique more reliable in practice, we have to complete the error analysis.

(3) The expansion of the application of the method. The current method based on vanishing points is valid only for objects which have parallel lines or edges on planar surfaces. To make the technique applicable to more general situations is a subject for future research.

BIBLIOGRAPHY

[1] Y.Yakimovsky and J.A.Feldman, "A semantics-based decision theory region analyzer," *Proc. 3rd International Conference on Artificial Intelligence*, Stanford, USA, pp.580-588, 1973.

[2] E.M.Riseman and M.A.Arbib, "Computational techniques in the visual segmentation of static scene", *Computer Graphics and Image Processing*, vol.6, pp.221-276, 1977.

[3] B.K.P.Horn, "Obtaining shape from shading information," in *The Psychology of Computer Vision*, P.Winston, Ed. McGraw-Hill, pp.115-155, 1975.

[4] A.P.Witkin, "Recovering surface shape and orientation from texture," *Artificial Intelligence*, vol.17, pp.17-45, 1981.

[5] A.R.Hanson and E.M.Riseman, "VISIONS: A computer system for interpreting scenes," in *Computer Vision Systems*, A.R.Hanson and E.M.Riseman, Eds. Academic Press, pp.303-333, 1978.

[6] J.R.Kender, "Shape from texture: A computational paradigm," *Proc. DARPA Image Understanding Workshop*, pp.79-84, 1979.

[7] Y.Ohta, K.Maenobu, and T.Sakai, "Obtaining surface orientation from texels under perspective projection," *Proc. 7th International Joint Conference on Artificial Intelligence*, Vancouver, Canada, pp.746-751, 1981.

[8] H.Nakatani, S.Kimura, O.Saito, and T.Kitahashi, "Extraction of vanishing point and its application to scene analysis based on image sequence," *Proc. 5th International Conference on Pattern Recognition*, Miami Beach, USA, pp.370-372, 1980; also in H.Nakatani and T.Kitahashi, "Determination of vanishing point in outdoor scene," *Transactions of IECE Japan*, vol.E64, No.5, pp.357-358, 1981.

[9] H.Nakatani and T.Kitahashi, "An iterative method for locating a vanishing point," *Trans. IECE Japan*, vol.J68-D, No.8, pp.1541-1542, 1985 (in Japanese).

[10] R.O.Duda and P.E.Hart, "Use of the Hough transformation to detect edges and lines," *Communications of the ACM*, vol.15, No.1, 1972.

[11] R.O.Duda and P.E.Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, 1973.

[12] A.Rosenfeld, R.A.Hummel and S.W.Zucker, "Scene labeling by relaxation operations," *IEEE Transactions of System, Man, and Cybernetics*, vol.SMC-6, No.6, pp.420-433, 1976.

[13] S.T.Barnard and W.B.Thompson, "Disparity analysis of images," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, vol.PAMI-2, No.4, pp.333-340, 1980.

[14] K.Sakaue and M.Takagi, "Separation of overlapping particles by iterative method," *Proc. 5th International Conference on Pattern Recognition*, Miami Beach, USA, pp.522-524, 1980.

[15] H.Nakatani and T.Kitahashi, "Effect of vanishing points on scene labeling," *Transactions of IECE Japan*, vol.J65-D, No.10, pp.1273-1279, 1982 (in Japanese); also in H.Nakatani and T.Kitahashi, "Inferring 3-D shape from line drawings using vanishing points," *Proc. 1st International Conference on Computers and Applications*, Beijing, China, pp.683-688, 1984.

[16] D.A.Huffman, "Impossible objects as nonsense sentences," in *Machine Intelligence*, vol.6. B.Meltzer and D.Michie, Eds. Edinburgh Univ. Press, PP.295-323, 1971.

[17] M.B.Clowes, "On seeing things," *Artificial Intelligence*, vol.2, No.1, pp.79-116, 1971.

[18] D.Walts, "Understanding line drawings of scenes with shadows," in *The Psychology of Computer Vision*, P.Winston, Ed. McGraw-Hill, pp.19-91, 1975.

[19] K.Sugihara, "Picture language for skeletal polyhedra," *Computer Graphics and Image Processing*, vol.8, pp.382-402, 1978.

[20] T.Kanade, "A theory of origami world," *Artificial Intelligence*, vol.13, pp.279-311, 1980.

[21] A.K.Machworth, "Interpreting pictures of polyhedral scenes," *Artificial Intelligence*, vol.4, No.2, 1973.

[22] K.Sugihara, "Mathematical structures of line drawings of polyhedrons," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.PAMI-4, No.5, pp.458-469, 1982.

[23] H.Nakatani and T.Kitahashi, "Reconstruction of object in outdoor scene based on vanishing point," *Transactions IECE Japan*, vol.J68-D, No.8, pp.1481-1488, 1985 (in Japanese).

[24] H.Tamura, "A comparison of line thinning algorithms from digital geometry viewpoint," *Proc. 4th International Joint Conference on Pattern Recognition*, Kyoto, Japan, pp.715-719, 1978.

[25] G.W.Stroke, W.E.Kock, Y,Kikuchi and J.Tsujiuchi, Eds., *Ultrasonic Imaging and Holography - Medical, Sonar, and Optical Applications*, Plenum Press, 1974.

[26] H.H.Plott, J.D.Irwin, and L.J.Pinson, "A real-time stereoscopic small-computer graphics display system," *IEEE Transactions on System, Man, and Cybernetics*, vol.SMC-5, pp.527-533, 1975.

[27] H.Nakatani, S.Tamura, K.Tanaka, A.Kitabatake, and M.Inoue, "A binocular stereoscopic display system for echocardiography," *IEEE Transactions on Biomedical Engineering*, vol.BME-26, No.2, pp.65-68, 1979.

[28] D.E.Robinson, "Display of three-dimensional ultrasonic data for medical diagnosis," *The Journal of the Acoustical Society of America*, vol.52, No.2, pp.673-687, 1972.

[29] H.K.Huang and R.S.Ledley, "Three-dimensional image reconstruction from *in vivo* consecutive transverse axial sections," *Computers in Biology and Medicine*, vol.5, pp.165-170, 1975.

[30] W.Glenn, R.Johnston, P.Morton, and S.Dwyer, "Image generation and display techniques for CT scan – thin transverse and reconstructed coronal and sagittal planes," *Investigation Radiography*, vol.10, No.5, pp.403-416, 1975.

[31] S.Tamura, H.Nakatani, K.Tanaka, M.Matsumoto, A.Kitabatake, M.Inoue, "Tomogram reconstruction in desired plane based on serial tomograms," *Computers in Biology and Medicine*, vol.9, No.4, pp.345-355, 1979.

[32] T.Kitahashi and H.Endo, "A new method of 3-D motion analysis using a concept of projective geometry," *Proc. 9th International Joint Conference on Artificial Intelligence*, Los Angeles, USA, pp.902-904, 1985.

[33] H.Nakatani, R.S.Weiss, and E.M.Riseman, "Application of vanishing points to 3-D measurement," *Proc. SPIE*, vol.507, pp.164-169, 1984.