IAPR Workshop on CV - Special Hardware and Industrial Applications OCT.12-14, 1988, Tokyo

3-D INDUSTRIAL PARTS IDENTIFICATION USING THE O. P. VIEWS GENERATION METHOD

Chung Lin Huang

Electrical Engineering Department, National Tsing Hua University
Hsin Chu, Taiwan, R.O.C.

ABSTRACT

This paper presents a new method to automatic recognition of 3-D industrial parts. Given images taken from different viewpoints of designated 3-D parts, algorithms are developed to interpret them as the same object. This study mainly involves the generation of the footprints of the viewed object, which consists of the top view, the front view, and the side view, from its input images.

I INTRODUCTION

During the past three decades, research and development in computer vision has increased dramatically. Recently, machine vision applied to manufacturing automation has received considerable attention [1,2]. Most of the significant progress has been made in two-dimensional pattern recognition. There are still many unsolved problems in the field of 3-D industrial parts identification.

A large body of research has been devoted to the problem of interpreting three-dimensional objects from digitized images [3,4], but the problem of recognizing a rigid 3-D industrial parts taken from any unknown position still remains mostly unsolved. The major problem is that the 2-D projection image is generally under-constrained, and complicated by the fact that metric properties such as orientation and length are not invariant under projection. Features and primitives, which can be easily fetched in 2-D pattern analysis, can hardly be extracted in 3-D scene analysis. New approaches in image understanding are essential to compute the descriptions of the three-dimensional layout from the two-dimensional projection image.

The main propose of this paper is to develop a computer-based system to identify 3-D objects from different viewpoints. These objects will be restricted to man-made industrial parts which are geometrically simpler than a natural scene (see Figure 1). A vision system developed to recognize 3-D objects with a camera is shown in Figure 2. A gray-level picture of a 3-D object through the image preprocessing technique is converted into a pictorial drawing. With the knowledge of shape from contour, we can develop the orientation of each surface of the 3-D object. The vision system is designated to generate the footprints of the 3-D object: the top view, the front view and the side view. The three views form a unique representation of the 3-D object of which the features are matched with those in the data base for positive identification of the object.

II PICTORIAL DRAWING GENERATION

The image processing techniques, which consist of edge detection, thinning, and noise removal, identify the objects only in terms of the edge pixels. The gap between the output of edge detection techniques and the smoothed, segmented curve or line segment descriptions of the 2-D image will be discussed. The 2-D description is needed for further spatial interpretation. Two of the basic problems for the curve segment description are how to choose the best scale of curve and how to decide the locations of best tangential discontinuities.

Since most industrial parts are composed of either planar or cylindrical surfaces, the boundaries of these surfaces can be described by three features: (1) the line segments, (2) the ellipses or circles, and (3) the elliptical arcs. The line finder is implemented basically on the transformation from the image space to the parameter space. By using the Hough Transform technique each figure point in image space can be transferred to several parametric points in the parameter space. Each point in the parameter space represents a straight line of the figure in the image space. The method is generalized to detect circles, ellipses, and elliptical arcs, which appear in a digitized image in the form of a cluster of edge pixels [5].

However, the small intensity differences in the 2-D image, which are caused by the rounded edge of the object, may produce some dangling lines and missing lines. A method had been presented to identify these defects by using the properties of labelling techniques, surface gradient and heuristic rules. How to connect the dangling lines correctly, to restore the missing lines, and to generate a pictorial drawing is not a trivial problem [6].

A gray level picture of 3-D objects through the image preprocessing techniques is first converted into a pictorial drawing. It is known that pictorial drawings can narrate 3-D shape very effectively. However, pictorial drawings are viewpoint-dependent and a 3-D object may be represented by numerous pictorial drawings. The pictorial drawing interpretation problem is in principle greatly underconstrained.

III ORTHOGRAPHIC PROJECTION VIEW GENERATION

This section presents the spatial understanding of the 3-D object pictorial drawing. By making use of the knowledge of regularity constraints and the property of the gradient space, we can develop the orientation of each surface of the 3-D object. Once the surface orientations have been obtained, the orientation of the edges and the recovery of the depth information for each vertex can be derived. The footprints of the 3-D object, which are the top view, the front view, and the side view, are generated by integrating the surfaces projected on each view.

In this paper, we assume (1) all 3-D parts having planar surfaces; (2) the distance of the viewed object being much greater than the camera focal length which makes the parallel projection a reasonable approximation; (3) all the pictures being taken from a general position. A general position means that a slight change of the position from which the picture is taken would not change the number of lines in the picture or the configuration in which they come together. In much of the following discussion, we also assume that the shapes can be described adequately by straight lines and planes. The surface, which is a plane enclosed inside the intersecting lines, is the basic element for the boundary representation of the solid object. A solid object can be represented by segmenting its boundary into a finite number of bounded subsets called surfaces and then describing each surface by its bounding edges or lines and vertex.

Let the 3-D surface in the viewer-centered coordinate system be described by the following equation,

$$f(x,y) - z = 0 \tag{1}$$

Taking the differential of both terms, we can rewrite it as

$$[\ \partial f/\partial x,\ \partial f/\partial y\ ,\ -1\ ]\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = 0 \tag{2}$$

The vector $[\partial f/\partial x\ ,\partial f/\partial y\ ,\ -1\ ]$ represents the 3-D gradient and is referred to as the surface normal which is perpendicular to any unit vector [dx, dy, dz] on the surface. We define the gradient $G=(p,q)$, where $p = \partial f/\partial x$, $q = \partial f/\partial y$, to represent how the plane is slanting relative to the viewing direction which is in the direction of the Z axis in the viewer-centered coordinate.

Let $G = (p,q)$ be the gradient of a surface and the 3-D line vector $\underline{A}$ on the surface projected on the 2-D picture plane as another 2-D line vector $\underline{a}$ with the orientation (i.e. $\underline{a} = [\ \cos\alpha,\ \sin\alpha\ ]$). The 3-D line vector $\underline{A}$ can be illustrated as,

$$\underline{A} = (\ \cos\alpha,\ \sin\alpha,\ \ -\underline{G}\cdot\underline{a}\ ). \tag{3}$$

During image formation projection, any two perpendicular 3-D vectors will distort as two skewed 2-D vectors on the picture, and the vectors then being no longer orthogonal. However, the heuristic inference determines these two vectors to be orthogonal in the 3-D scene[7]. If there are two 2-D line vectors, $\underline{a}$ and $\underline{b}$ (with orientations and), of which the two corresponding orthogonal 3-D line vectors, $\underline{A}$ and $\underline{B}$, are on the surface with gradient $\underline{G}$ (how to find these two 2-D line vectors will be discussed in next section), then from Equation 3 we can obtain the following equations, $\underline{A}\cdot\underline{B} = 0$. It can be demonstrated as

$$\cos\ (\ \alpha - \beta\ ) + (\underline{G}\cdot\underline{a})\ (\underline{G}\cdot\underline{b}) = 0. \tag{4}$$

Suppose there are three surfaces sharing one vertex which have the unknown gradient $\underline{G1}$, $\underline{G2}$ and $\underline{G3}$. By applying the properties already mentioned, we can derive the following constraints:

$$\cos\ (\alpha_1 - \beta_1) + (\underline{G1}\cdot\underline{a1})\ (\underline{G1}\cdot\underline{b1}) = 0 \tag{5}$$
$$\cos\ (\alpha_2 - \beta_2) + (\underline{G2}\cdot\underline{a2})\ (\underline{G2}\cdot\underline{b2}) = 0 \tag{6}$$
$$\cos\ (\alpha_3 - \beta_3) + (\underline{G3}\cdot\underline{a3})\ (\underline{G3}\cdot\underline{b3}) = 0 \tag{7}$$

where

$\underline{a1}=(\cos\alpha_1,\ \sin\alpha_1)$, $\underline{b1}=(\cos\beta_1,\ \sin\beta_1)$
$\underline{a2}=(\cos\alpha_2,\ \sin\alpha_2)$, $\underline{b2}=(\cos\beta_2,\ \sin\beta_2)$
$\underline{a3}=(\cos\alpha_3,\ \sin\alpha_3)$, $\underline{b3}=(\cos\beta_3,\ \sin\beta_3)$

are three pairs of the orthogonal line vectors on the three adjacent surfaces.

One of the most important properties of gradient space is the duality between the picture plane and the gradient space. For instance, if two surfaces meet and the intersection line is projected on the picture plane as line L, then the gradients of the two surfaces are on a gradient space line which is perpendicular to L. The duality property between the gradient space and the picture plane can be described in the following equations:

$$\underline{G1}\cdot\underline{c2} = \underline{G2}\cdot\underline{c2} \tag{8}$$
$$\underline{G2}\cdot\underline{c3} = \underline{G3}\cdot\underline{c3} \tag{9}$$
$$\underline{G3}\cdot\underline{c1} = \underline{G1}\cdot\underline{c1} \tag{10}$$

where $\underline{c1} = (\cos\phi_1,\ \sin\phi_1)$, $\underline{c2} = (\cos\phi_2,\ \sin\phi_2)$ and $\underline{c3} = (\cos\phi_3,\ \sin\phi_3)$ are the intersection line vectors on every two surfaces. For every three adjacent surfaces, there are six constraints for the three gradients $\underline{G1}$, $\underline{G2}$, and $\underline{G3}$.

To solve the above six nonlinear equations effectively, we introduce the hill climbing technique illustrated in Figure 3. Beginning with the lowest point (i.e. point 1) in the valley of the hyperbola G3 and climbing in two directions (i.e. from point 1 to point 2 or 3), the corresponding points (i.e. point 1' and 1") on the other two hyperbola G1 and G2 can be obtained by the two constraints (i.e. Eqs.9 and 10). If the third constraint (i.e. from Eq. 8 , $|\underline{G1} \cdot \underline{c2} - \underline{G2} \cdot \underline{c2}| < \theta$, where θ is the tolerance) is satisfied, then the solutions for the three gradients are found. As illustrated in Figure 3, the first constraint implies that the line connecting point P and point P' has to be perpendicular to vector $\underline{c1}$, the second constraint implies that the line connecting point P and P" has to be perpendicular to vector $\underline{c2}$, and the final constraint requires the line connecting points P' and P" to be perpendicular to vector $\underline{c3}$. As illustrated in Figure 3 the final solution is found when the line connecting points 4' and 4" and the vector $\underline{c3}$ meet the requirement.

The gradient calculation approach is carried on the viewer-centered coordinate system which is not sufficient for generating the three orthographic projection views of the object. Another coordinate system called the O.P. coordinate system which is object-oriented is introduced. It is determined by the base and the primary surface. The primary surface is selected from one of the surfaces which have their normal perpendicular to that of the base. The normal of the primary surface is selected as the Z axis of the O.P. coordinate system. The base of the object, by which the object may rest stably on the ground, is supposed to be flat and has its orientation determined as the Y axis of the O.P. coordinate. In the O.P. coordinate system, the viewer is no longer in the direction of the Z axis, but in a certain direction determined by coordinate transformation between the viewer-centered coordinate and the O.P. coordinate systems. The entire procedure of the O.P. views generation is illustrated in Figure 4.

An arbitrary polyhedron can be described by the surfaces. Each surface is a planar polygon that can be described by an ordered list of the vertices or the line segment boundaries. The locations of the vertices have a significant influence on the faces interpretation which need to be defined relative to a certain reference point. By choosing the lowest vertex in the picture, which seems to be the closest point to the viewer, as a reference point (i.e. [0,0,0] in O.P. coordinate system), and tracing from the reference vertex to the other vertex following the line segment boundary joining them, the positions of all the other vertices can be developed.

The O.P. views of the slant surface can be developed by projecting the slant surface from three projection planes whose normals are the same as the directions of the three major axes of the O.P. coordinate system. The slant surface is decomposed into two or three subsurfaces in the three projection planes. By integrating all the subsurfaces on each projection plane, the three O.P. views of the 3-D object can be developed. However, the subsurfaces may overlap one another because of their different depths. The subsurface which has the smaller depth, sometimes, will overlap the subsurface with larger depth.

The experimental results of the 3-D object image taken at different viewing angles are illustrated in Figure 5. These two pictures show one surface being occluded. The generated O.P. views demonstrate that one edge of the top view and front view from the two pictures are not complete due to the edge being occluded by the other two surfaces.

IV INEXACT MATCHING TO IDENTIFY THE 3-D PARTS

The 3-D industrial parts identification system has two goals: one is model generation and the other is recognition. Model generation deserves great care because it helps to increase the recognition reliability. The previous section illustrates the model generation which generates the O.P. views of the line drawings of objects from their 2-D intensity images. The object recognition is based on the matching between the generated model and the stored model. Due to the unpreciseness of the image processing techniques, the generated model can never be exact. However, the best match is required to interpret the object.

The generated and the stored models are described in terms of an undirect graph where the nodes of the graph correspond to the primitives and the branches connecting nodes correspond to the relationship among the primitives. The model described in this kind of graph is called a relational model. Relational models, which describe a scene in terms of its primitives and their relationships, have proved to be useful in the recognition process [7,8,9].

V SUMMARY

We have proposed a new method to identify 3-D industrial parts from 2-D input intensity images. In comparison with previous 3-D vision research, this study used heuristic constraints and surface gradient analysis to recover the 3-D information quantitatively. This method interprets the scene of objects much more robustly in that it can differentiate the cube from the truncated pyramid. It is based on the following procedures: firstly convert the images to pictorial drawings, secondly use some heuristic inference of the spatial information and the constraints of the orientations of intersecting surfaces to restore 3-D information of the objects from 2-D images, and thirdly compare the generated model with the prestored model for the best match to identify the object in the image.

VI REFERENCES

[1] Perkins, W.A.,IEEE Trans. on Computers, Vol.C-27, No.2, pp.126-143, Feb. 1978.

[2] Kelley, R.B., Martins, H.A.S., Birk, J.R., and Dessimoz, J.D., Proceedings, Vol.71, No.7, pp.803-820, July 1983.

[3] Barrow, H.G., and Tenenbaum, J.M., IEEE Proc., Vol.69, No.5, pp.572-595, May 1981.

[4] Brandy, M., ACM Computing Surveys, Vol. 14, No.1, pp.3-71, March 1982.

[5] Huang C.L. and Tou J., SPIE Conf. on Appl. of Art. Int. V, May 1987.
[6] Huang C.L. and Tou J., SPIE Conf. on Visual Comm. and Image Proc. II, Oct. 1987.
[7] Huang C.L., Ph.D. Dissertation, Univ. of Florida, Aug., 1987.
[8] Shapiro, L.G. and Haralick, R.M., IEEE Trans. on PAMI, Vol. PAMI-3, Sept. 1981.
[9] Sanfelieu, A. and Fu, K.S., IEEE Trans. Syst., Man, Cybern., Vol.SMC-13, May-June 1983.

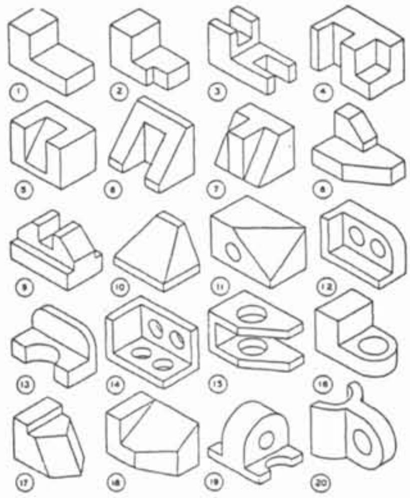


Figure 1. Mechanical Parts.

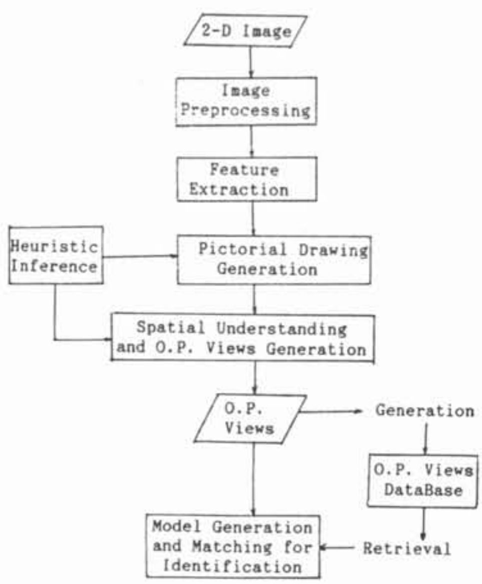
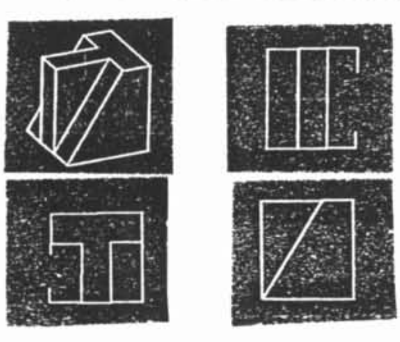


Figure 2. System Configuration.



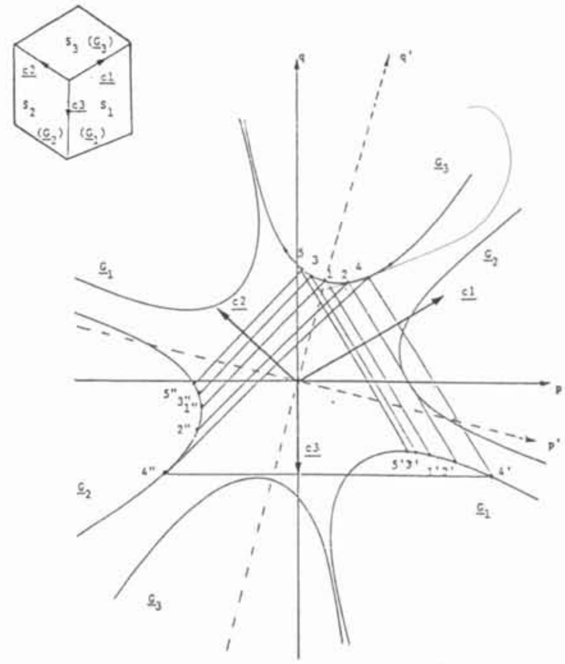


Figure 3. Hill Climbing Technique.

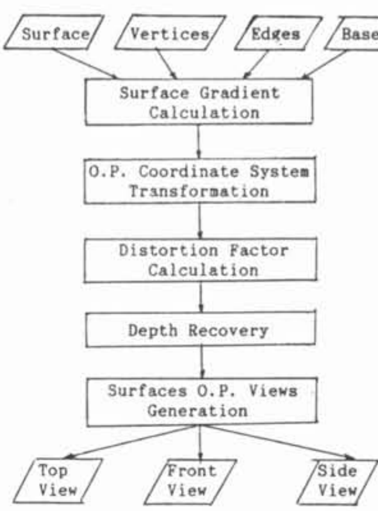


Figure 4. O.P. Views Generation Flow Diagram.

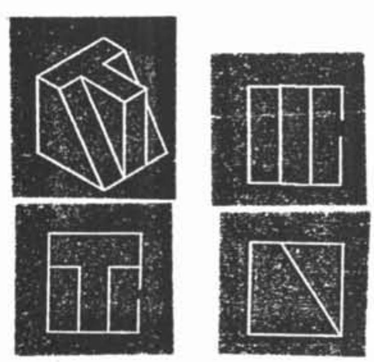


Figure 5. O.P. Views for two images of the same object.