Locating partially occluded parts from the poses of shape primitives

D. M. Tsai and J. K. Liao Machine Vision Lab. Department of Industrial Engineering and Management Yuan-Ze University, Chung-Li, Taiwan, R.O.C. E-mail: <u>iedmtsai@saturn.yzu.edu.tw</u> (Figures are not available in this report.)

1.INTRODUCTION

The problen of recognizing partiall occluded parts is of considerable interest in industrial automation. The occlusion usually takes place when an object is either overlapped or touched by other objects. Lighting, reflection of physical obstructions may also render the observation of a complete object impossible. In the robotic assembly applications, we can increase the system flexibility by using machine intelligence that allows for the appearance of noisy objects, partially occluded objects, and objects in random positions and orientations. Global features such as areas, perimeters and moments are very robust for the recognition of non-overlapping objects. However, nes features are created when objects are partially occluded and they will not correspond to any features known a priori. This makes the recognition task for occluded objects more difficult and complicated than that for a single isolated object.

Object recognition systems generally use precompiled descriptions of the model objects that can possibly occur in a scene. A model image refers to the image of an object under ideal viewing conditions and with the entire object visible, whereas the scene image refers to an image of multiple objects with occlusions. Object recognition basically involver two tasks:shape representation(feature extraction)followed by a matching process. Polygonal approximations [1,2,3,4]are the most popular representation to objects in the literature of partially occluded object recognition. They represent a digital curve by a sequence of line segments, under the condition that the obtained polygon preserves the shape and size of the original one, within a giver tolerance. In the matching process, local features such as vertex position, line segment slope, interior angle at the vertex and line segment length are employed for

measuring the similarity between the model object and the scene object. The main problem with the polygonal approximation is that under different viewing conditions the approximation is not stable and, therefore, the polygon representation of a curved object may not be unique. Han and Jang [5]extracted the vertices of local maximum curvature of digital boundaries. At each vertex, the interior angle and the length of two adjacent vertices are used to construct a graph of compatible nodes. A zero-one integer programming and a heuristic method are proposed to approach the graph matching problem. Koch and Kashyap [6] used the polygon representation of object boundaries to guide a hypothesis generation algorithm. The algorithm iteratively generates and tests hypotheses for compatibility between models and the scene until it identifies all occluded objects. Bouyakhf [7] used the segments represented by the polygonal approximation as the feature primitives. The matching process is based on the hypothesis propagation which usesstructural relations and hash-key to find the next segment of a polygon. Ayache and Faugeras [8] matched simple descriptions of the occluded objects and the models by a technique called HYPER of hypotheses generation and verification coupled with a recursive estimation of the model to scene transformation. They used a polygonal approximation method to extract the feature primitives form the object boundaries, although the algorithm is basically independent of the types of primitives used to represent the 2-D shapes. Schwarts and Sharir [9] proposed an algorithm that identifies partially occluded objects in two and three dimensions by finding the shortest paths near a given polygonal curve. In [10], the segments obtained from the polygonal approximation of the object boundaries are used as the feature primitives. The matching problem is tackled using the heuristic search with a state space formulation. Grimson and Lozano-Perez [11] modeled the objects as ploygons or polyhedra. They structure the search for consistent matches as the generation and exploration of an interpretation tree.

Lin and Chellappa [12] presented a method for the classification of 2-D partial shapes using the Fourier descriptors [13,14]. The matching problem is formulated as one of estimating the Fourier descriptors of the unknown complete shape from the observations derived from an arbitrarily rotated and scaled shape with missing segments. Mitchell and Grogan [15] proposed the Fourier-Mellin transform for partial 2-D boundary classification. Gorman et al. [16] obtained the local features by splitting a boundary into segments which are described by the Fourier descriptors. A dynamic

programming formulation is developed for matching the segments extracted from the partial object and the model in the database. Shape representation with the Fourier descriptors is time-consuming and may fail to recognize severally occluded objects.

Perkins [17] presented a matching algorithm based upon cross-correlating the tangent angle q at a boundary point as a function of the arc length s between the scene object and the model in the database. Turney et al. [18] used salient subtemplates which most differentiate the objects as features. The model subtemplates are matched with the segments of the scene object using the least-squares fit in q-s space.

Davies [19] examined the use of the generalized Hough transform (GHT) [20] for object recognition under the effects of occlusions. In the GHT process for the recognition of arbitrary shapes, the model of the shape is built by choosing a reference point R and a set of points Pi lying on the shape's boundary. It consists of a table, the so-called reference table, which stores the displacement vector PiR as a function of the directin of the gradient. A voting process then transforms the set of boundary points of a scene object in the image space into a set of accumulated votes in the parameter space defined by the coordinates of the reference point. The gradient direction is the only signature information used in the reference table of the GHT, and is rotation-dependent. David concluded that the sensitivity and accuracy of the GHT are simply related to the length of the visible portion of the boundary. Grimson and Huttenlocher [21] investigated the sensitivity of the Hough transform for object recognition. They found that the GHT methods work well as long as the correct match account for both much of the model and much of the scene image data. For moderate levels of sensor noise, occlusion and image clutter, however, the methods can hypothesize many false solution, and their effectiveness is dramatically reduced.

In this paper, we address the problem of recognizing and locating partially occluded industrial parts lying on a flat surface. The shapes of parts under study consist of linear and curved segments, where the curved segments can be approximated by circular arcs. The contraints on the objects are widely applicable in the industrial environment since most of man-made objects have such geometric features and many industrial parts can be considered planar due to their small thickness.

For an object shape involving linear and circular segments, a gradient direction will be associated to many boundary points and, therefore, many false alarms may be generated if the GHT method is employed. Figures 1(a) and 1(b) show overlapping objects that comprises linear and circular boundaries. The shapes of black dash lines in the figures represent the object poses (translation and rotation) evaluated by the GHT. Unlike the GHT method that describes an object shape by means of the points lying on its boundary, a Hough-clustering object recognition method is introduced, where the linear and circular segments are used as the feature primitives to the construction of reference table and the generation of votes. The entries in the reference table are the binary relation of each primitive pair and unary properties of the primitives, which are used as additional constraints to eliminate false matching of a model to the scene image.

Since segments of object shapes instead of boundary points are used for shape representation, the proposed recognition method is relatively insensitive to the length of the vible protion of the object boundary as long as the primitives can be detected in the feature extraction phase.

Indetification of the top-bottom relation of overlapping objects is important for robotic assembly applications. This problem is not addressed in the literature described above. Range images [22,23] are the most common methods used to detect the depth of objects. However, the equipment of range sensing is more expensive than the regular CCD sensing devices, and the algorithms for detecting object depth is relatively more complex and computationally intensive. In this paper,we also investigate the top-bottom relation of overlapping objects in gray-scale images by tracing the shadowed points along the estimated object boundaries.

This paper is organized as follows: In section 2, a shape representation method that extracts line and circle primitives of object shapes is discussed. In section 3, the definitions of unary properties of primitives and binary relations of primitive pairs are introduced first. The construction of reference table for a model based on the binary relations of primitive pairs is addressed. A vote generation procedure is then carred

out to indentify the location of the object in the scene image. Section 6 summarize this work.

(a)

(b)

figure 1. locating overlapping objects using the GHT

2. SHAPE REPRESENTATION

Images are pre-processed in order to extract all feature primitives that compose the object shapes. Features employed in this paper are the linear and circular segments of object boundaries. Since the shpae of a curve is concentrated at the dominant points, the feature extraction procedure begins with the detection of dominant points by measuring the second-order curvatures along the boundaries. the resulting dominant points partition the boundary into fragments. Then we determine the parameters for the equation of a line of a circle that make the equation a best fit, in the least-squares scene, to the points on each fragment. If a fragment can not fit a line or circle equation subject of some bounded error, it is concluded that additional dominant points exist on the fragment and this fragment should be partitioned into small subfragments. The partition procedure is carried out by detecting the maximal curvature point on the fragment and splitting the fragment into two subfragments that have the maximal curvature point as one end. The least-squares line/circle fitting is then performed to estimate the equation of each splitled subfragments. The partition procedure above is repeated recursively until the least-squares line/circle fitting succeeds for all fragments.

The feature extraction procedure is described in detail as follows.

Let the sequence of n degital points describe a closed boundary P,

$$P = \{P_i = (x_i, y_i), i = 1, 2, \dots, n\}$$

where p_{i+1} is a neighbor of p_i (modulo n), and (x_i, y_i) is the Cartesian coordinates of $p_i, i=1,2,...,n$.

Dominant points on a boundary correspond to the points having high curvature values.Curvature is defined as the instantaneous rate of change of tangents with respect to the arc length at a given boundary point. The discrete version of curvature is computed by the following formulas:

$$\boldsymbol{t}(i) = \frac{1}{h} \sum_{j=1}^{h} \tan^{-1}(\frac{y_{i+j} - y_{i-j}}{x_{i+j} - x_{i-j}}), i = 1, 2, \dots, n$$
(1.a)

$$\boldsymbol{k}(i) = \frac{1}{h} \sum_{j=1}^{h} [\boldsymbol{t}(i+j) - \boldsymbol{t}(i-j)], i = 1, 2, \dots, n$$
 (1.b)

Here, h specifies the region of support. It is used as a smoothing factor to calculate the mean tangent angle t(i) and mean curvature k(i) k(i) at point p_i . The size of region of support h can be a predefined constant or be adaptively determined as proposed in [24]. A point p_i is said to be a dominant point if k(i) is above a given threshold and individual dominant point are separated by a spacing of at least h points.

For each fragment determined by two adjacent dominant points, we perform the leastsquares fitting for a line or a circle to the set of boundary points on the fragment. Let $p_t = (x_t, y_t)$ and $p_{t+k} = (x_{t+k}, y_{t+k})$ correspond to the coordinates of two adjacent dominant points. The optimal parameters of a fitted line y=mx+c, in the least-squares scene, are given by

$$m = \frac{(\sum_{j=t}^{t+k} x_j y_j) - (k+1)\overline{x}\overline{y}}{(\sum_{j=t}^{t+k} x_j^2) - (k+1)\overline{x}^2}$$
$$c = \overline{y} - m\overline{x}$$
where $\overline{x} = \frac{1}{k+1} \sum_{j=t}^{t+k} x_j$
$$\overline{y} = \frac{1}{k+1} \sum_{j=t}^{t+k} y_j$$

The neam distance from a boundary point on the fragment to the estimated line is evaluated by

$$\mathbf{e}_{\rm L} = \frac{1}{{\rm k}+1} \sum_{\rm j=t}^{\rm t+k} \frac{\left|{\rm m} \cdot {\rm x}_{\rm j} - {\rm y}_{\rm j} + {\rm c}\right|}{\sqrt{{\rm m}^2+1}}$$

If the mean distance error e_{L} is sufficiently small against a predefined threshold T_{d} ,

this fragment is categorized as a lline primitive and is represented by L. The unary property of a line primitive is described by the length of the line segment.

If $e_L > T_d$, the fragment is assumed to belong to a circular arc. The arc center (x_0, y_0) and the radius r are estimated by minimizing the sum of the squared errors between the radius and the distances from the boundary points to the center, i.e.,

min f
$$(x_0, y_0, r) = \sum_{j=t}^{t+k} [r^2 - \{(x_j - x_0)^2 + (y_j - y_0)^2\}]$$

Differentiating the function $f(x_0, y_0, r)$ with respect to x_0, y_0 and r and setting them to zero, we have [25]:

$$\mathbf{x}_{0} = \frac{\mathbf{c}_{1}\mathbf{b}_{1} - \mathbf{c}_{2}\mathbf{b}_{1}}{\mathbf{a}_{1}\mathbf{b}_{2} - \mathbf{a}_{2}\mathbf{b}_{1}}$$
(4.a)

$$\mathbf{y}_{0} = \frac{\mathbf{a}_{1}\mathbf{c}_{2} - \mathbf{a}_{2}\mathbf{c}_{1}}{\mathbf{a}_{1}\mathbf{b}_{2} - \mathbf{a}_{2}\mathbf{b}_{1}}$$
(4. b)

$$r = \frac{1}{k+1} \bigvee_{j=t}^{t+k} x_j^2 - 2x_0 \sum_{j=t}^{t+k} x_j + (k+1)x_0^2 + \sum_{j=t}^{t+k} y_j^2 - 2y_0 \sum_{j=t}^{t+k} y_j + (k+1)y_0^2 \bigvee_{j=t}^{t+k} (4.c)$$
where $a_1 = 2 \bigvee_{j=t}^{t+k} x_j \bigvee_{j=t}^{t+k} 2(k+1) \sum_{j=t}^{t+k} x_j^2$

$$a_2 = 2 \bigvee_{j=t}^{t+k} x_j \sum_{j=t}^{t+k} y_j - (k-1) \sum_{j=t}^{t+k} x_j y_j \bigvee_{j=t}^{t+k} b_1$$

$$b_2 = 2 \bigvee_{j=t}^{t+k} y_j \bigvee_{j=t}^{t+k} 2(k+1) \sum_{j=t}^{t+k} y_j^2$$

$$c_1 = \sum_{j=t}^{t+k} x_j^2 \sum_{j=t}^{t+k} x_j - (k+1) \sum_{j=t}^{t+k} x_j^3 + \sum_{j=t}^{t+k} x_j \sum_{j=t}^{t+k} y_j^2 - (k+1) \sum_{j=t}^{t+k} x_j y_j^2$$

$$c_2 = \sum_{j=t}^{t+k} x_j^2 \sum_{j=t}^{t+k} y_j - (k+1) \sum_{j=t}^{t+k} y_j^3 + \sum_{j=t}^{t+k} y_j \sum_{j=t}^{t+k} y_j^2 - (k+1) \sum_{j=t}^{t+k} x_j^2 y_j$$

The mean distance from a boundary point on the fragment to the estimated circle is computed by

$$\boldsymbol{e}_{c} = \frac{1}{k+1} \sum_{j=t}^{t+k} \left| \mathbf{r} - [(\mathbf{x}_{j} - \mathbf{x}_{0})^{2} + (\mathbf{y}_{j} - \mathbf{y}_{0})^{2}]^{\frac{1}{2}} \right|$$

(5) e

If e_c is sufficiently small against the given threshold T_d , the fragment is categorized as a circle primitive, and si represented by C.

The unary properties of a circle primitive include the arc length, the radius and the concavity of the arc. Figures 2(a) and 2(b) show a convex arc and a concave arc, repectively. They arc determined as follows:

$$\mathbf{P}_{i} = (\mathbf{x}_{i}, \mathbf{y}_{i})$$

Select an arbitrary point on the boundary of the circular arc, and form a line segment \overline{L} connecting point P_i and the center of the circular arc (x_0, y_0) . Let $V = (v_x, v_y)$ be the immediate neighboring point of P_i and lie on \overline{L} . If V is an object point, the circular segment is defined as a convex arc (see Figure 2(a)). If V is a background point, the circular segment is defined as a concave arc (see Figure 2(b)). The coordinates of point $V = (v_x, v_y)$ are determined by the formula below.

Let
$$x_i \neq x_0$$

 $v_x = x_i + SGN(x_0 - x_i)$
 $v_y = y_i + SGN(x_0 - x_i) \cdot [(y_0 - y_i)/(x_0 - x_i)]$
where $SGN(x_0 - x_i) = 1$ if $x_0 - x_i > 0$
 $= -1$ otherwise

If the fragment under test results in a large fitting error \mathbf{e}_{c} , the pratition procedure described previously is repeated until the mean distance error \mathbf{e}_{L} or \mathbf{e}_{c} of every splitted subfragment meets the requirement.

Through the use of dominant point detection, least-squares line/curve fitting and partition procedure, all feature primitives of an object shape can be extracted. Since the parameters of each primitive are estimated from a set of boundary points on the fragment rather than a few salient points sampled from the boundary, they are less sensitive to noise and, therefore, a reliable representation scheme for object shapes.

(a) A convex arc

(b) A concave arc

Figure 2. Convexity and concavity of circular arcs

3. CLUSTERING OF POSE PARAMETERS

In this section, we introduce a Hough-clustering technique to construct the reference table of the model based on the binary relations of primitive paris. In the vote generation phase, the binary relations of primitive paris of a scene object is used as the signature to find the counterparts in the reference table, and the coordinates of the reference point in the scene image and the rotated angle between the model object and scene object are computed accordingly. The coordinates of the reference oint and the rotated angle make up the parameter space, and are accumulated for the evidence of a match.

3.1 The construction of reference table

In this paper, the model of the shape is built by choosing a reference point at the centroid of the shape and a set of primitive pairs G. G is a 3-tuple and is defined by

 $G = \{B(a,b),U(a),U(b)\}$

where B(a,b) represents the binary relation of two primitives a and b, $a \neq b$

U(a) represents the unary properties of primitive a

U(b) represents the unary properties of primitive b

Let L_i denote the *i*-th primitive of a line segment and C_i the *j*-th primitive of

a circular arc. The binary relations of different combinations of primitive types are defined by

 $B(L_i, L_j)$ = the angle between linws L_i and L_j , $i \neq j$ $B(C_i, C_j)$ = the distance between the centers of circles C_i and C_j , $i \neq j$ $B(L_i, C_j)$ = the shortest distance from the center of circle C_j to the line L_i .

The unary property of a line primitive $L_i, U(L_i)$, is defined by the length of the line segment $L_i, U(C_i)$ is a three-dimensional vector which defines three properties of the circle primitive C_i , i.e.,

$$U(C_i) = (C_{i1}, C_{i2}, C_{i3})$$

where C_{i1} = the arc length of circle C_i

 C_{i2} = the radius of circle C_i

 C_{i3} = convexity or concavity of circle C_i .

Note that the unary properties and binary relations of primitives defined above are translation- and rotation-invariant. The shape representation by binary relations of primitive pairs is robust for severely occluded objects, where salient features may not exist on the visible boundaries.

Now we can construct the reference table of a model shape by examining the binary relations of primitive pairs. Unlike the GHT that only uses the rotation-dependent gradients of boundary points to construct the reference table, the proposed method uses the binary relations of primitive pairs as the signature and utilizes the unary properties of primitives as additional constraints to eliminate false matches. For each correct match of model and scene primitive pairs in the vote generation phase, the reference table must provide the required geometrical information so that the reference table and rotated angle of the scene object can be computed and votes of the corresponding parameter values can be accumulated. The construction of the table for a model object is accomplished as follows.

Choose the centroid of the model shape as the reference point at coordinates (x^*,y^*) . The geometrical relations between the reference point and the equations of primitives are determined according to the types of primitive pairs below.

1). (L_i, L_j) : a pair consisting of two line primitives L_i and L_j .

The geometrical entries are defined by the distances, d_1 and d_2 , from the reference point to individual lines L_i and L_j (see Figure 3). Let the line equations of L_i and L_j be given by

$$L_i: a_1x + b_1y + c_1 = 0$$

 $L_i: a_2x + b_2y + c_2 = 0$

hence,

$$d_{1} = |a_{1}x^{*} + b_{1}y^{*} + c_{1}| / (a_{1}^{2} + b_{1}^{2})^{\frac{1}{2}}$$
(6.a)

$$d_{2} = \left| a_{2}x^{*} + b_{2}y^{*} + c_{2} \right| / (a_{2}^{2} + b_{2}^{2})^{\frac{1}{2}}$$
(6.b)

2). (C_i, C_j) : a pair consisting of two circle primitives C_i and C_j

The geometrical entries are defined by the distances, d_1 and₂, from the reference point to the individual centers of circles, and the angles q_1 and q_2 are the angles between the individual line segments d_1 and₂ and the line connecting the two centers of circles C_i and C_j . Let the circle equations C_i and C_j be given by

$$C_{i}:(x - x_{1})^{2} + (y - y_{1})^{2} = r_{1}^{2}$$

$$C_{j}:(x - x_{2})^{2} + (y - y_{2})^{2} = r_{2}^{2}$$

Hence,

$$\mathbf{d}_{1} = [(\mathbf{x}^{*} - \mathbf{x}_{1})^{2} + (\mathbf{y}^{*} - \mathbf{y}_{1})^{2}]^{\frac{1}{2}}$$

(7.a)

$$d_2 = [(x^* - x_2)^2 + (y^* - y_2)^2]^{\frac{1}{2}}$$

(7.b)

$$\boldsymbol{q}_1 = \cos^{-1} \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1| \cdot |\mathbf{v}_2|}$$

(7.c)

$$\boldsymbol{q}_2 = \cos^{-1} \frac{\mathbf{v}_3 \cdot \mathbf{v}_4}{\left|\mathbf{v}_3\right| \cdot \left|\mathbf{v}_4\right|}$$

(7.d)

where
$$V_1 = (x^* - x_1, y^* - y_1)$$
, $V_2 = (x_2 - x_1, y_2 - y_1)$
 $V_3 = (x^* - x_2, y^* - y_2)$, $V_4 = (x_1 - x_2, y_1 - y_2)$

3). (L_i, C_j) : a pair consisting of a line primitive L_i and a circle primitive C_j .

The geometrical entries are defined by the distances, d_1 and d_2 , from the reference point to the line L_i and the center of circle C_j , respectively(see Figure 5). Let the line equation of L_i and circle equation of C_j be given by

L_i:
$$ax + by + c = 0$$

C_j: $(x - x_0)^2 + (y - y_0)^2 = r^2$

Hence,

$$d_1 = |ax^* + by^* + c| / (a^2 + b^2)^{\frac{1}{2}}$$

(8.a)

$$\mathbf{d}_2 = \begin{bmatrix} \mathbf{y}_1^* \mathbf{y}_1 & \mathbf{y}_2 \end{bmatrix} \mathbf{y}_1^* \mathbf{y}_2^* \mathbf{y}_2^* \mathbf{y}_1^* \mathbf{y}_2^* \mathbf{y}_1^* \mathbf{y}_2^* \mathbf{y}_2^*$$

(8.b)

Note that the geometrical entries of primitive pairs with vespect to the reference point, d_1 , d_2 , q_1 and q_2 , are invariant to translations and rotations of objects. For each primitive pair of the model object, the geometrical entries are stored in the reference table. Table 1 shows the schematic format of the reference table.

Figure 3. Geometrical entries d_1 and d_2 for primitive pairs (L_i, L_j)

Figure 4. Geometrical entries d_1 , d_2 , \boldsymbol{q}_1 and \boldsymbol{q}_2 for primitive pairs (C_i , C_j)

Figure 5. Geometrical entries d_1 and d_2 for primitive pairs (L_i, C_j)

Table 1. The reference table format

Binary relation	Unary property	Unary property	Geometrical	
B(a,b)	U(a)	U(b)	entries	
$B(L_i, L_j)$	$U(L_i)$	$J(L_j)$	$d_1, d_2(eq.6)$	

$B(C_i, C_j)$	$U(C_i) = (C_{i1}, C_{i2}, C_{i3})$	$U(C_{j}) = (C_{j1}, C_{j2}, C_{j3})$	$\mathbf{d}_1, \mathbf{d}_2, \boldsymbol{q}_1, \boldsymbol{q}_2$
			eq.7

$B(C_i, C_j)$	$U(L_i)$	$U(C_{j}) = (C_{j1}, C_{j2}, C_{j3})$	$l_1, d_2(eq.8)$

The reference table is divided into three groups according to the types of primitive pairs $(L_i, L_j), (C_i, C_j)$ and (L_i, C_j) . The entries in each group are sorted in non-decreasing order by the values of the binar relations. Column 1 of the table lists the binary relations of primitive pairs, which are used as the signature in the vote generation phase. Columns 2 and 3 are the unary properties of primitives, which give additional constraints to eliminate false matches. Column 4 lists the geometrical entries with respect to the reference point, which are stored as a function of binary relations of primitive pairs. This concludes the construction of the reference table.

3.2 The generation of votes

The reference table is used to detect instances of the scene object in an image in the following manner.

Given the overlapping objects in a scene image, we obtain a set of primitive pairs $G = \{B(a',b'), U(a'), U(b')\}$ using the feature extraction method described previously. Divide G' into three groups according to the types of primitive pairs in the image space into a set of accumulated votes in the parameter space. The binary relation B(a',b') is used as the signature to the reference table from which the geometrical entries d_1 , d_2 and/or q_1 and q_2 are obtained, and the coordinates of reference point $(\overline{x}, \overline{y})$ of a scene object and rotated angle q_r are computed. Rondom pairs of model and scene primitives will result in randomly distributed transformations of parameters. Those pairs of correct matching will generate approximately the same coordinates of reference point and rotated angle. Acluster of similar transformation then indicates the evidence of a correct match and shows the pose of the scene object with respect to the model object.

Let L_s and C_t represent the line and circle primitives of a model object,

respectively; \mathbf{L}_{i} and \mathbf{C}_{j} denote the line and circle primitives of a scene object, respectively. A primitive pair $(\mathbf{L}_{i}, \mathbf{L}_{j})$ of the scene object is said to be assignable to the primitive pair $(\mathbf{L}_{s}, \mathbf{L}_{t})$ of the model object, denoted by $(\mathbf{L}_{i}, \mathbf{L}_{j}) \cong (\mathbf{L}_{s}, \mathbf{L}_{t})$, if

$$B(L'_{i}, L'_{j}) = (L_{s}, L_{t})$$

and

$$\min\{u(L_i), u(L_i)\} \le \min\{u(L_s), u(L_t)\}$$

and

$$\max \{ u(L_i), u(L_i) \} \le \max \{ u(L_s), u(L_t) \}$$

If the primitive pairs (L_i, L_j) and (L_s, L_t) have a correct match, the necessary conditions must be that both have the same binary relation (the angle between two lines), and the line segment lengths of the scene object are shorter than or equal to those of the model object since the line segments of the scene object could be occluded. Note that the necessary conditions listed above and a concise expression for simplicity. In real applications, they should be subject to some bounded error due to quantization.

Similarly. a primitive pair (C_i, C_j) is said to be assignable to the primitive pair (C_s, C_t) , denoted by $(C_i, C_j) \cong (C_s, C_t)$, if

$$B(C'_{i}, C'_{j}) = (C_{s}, C_{t})$$

and

$$C'_{i1} \le C_{s1}, C'_{i2} = C_{s2}, C'_{i3} = C_{s3}$$

and

$$C'_{j1} \le C_{t1}, C'_{j2} = C_{t2}, C'_{j3} = C_{t3}$$

given that $C_{i2} \leq C_{j2}$ and $C_{s2} \leq C_{t2}$, Recall that $B(C_i, C_j)$ is the binary relation that measures the distance between two centers of circles C_i and C_j . C_{i1} , C_{i2} and C_{i3} are the arc length, the radius and the concavity of circle C_i , respectively.

Finally, a primitive pair (L_i, C_j) is said to be assignable to the primitive pair

 (L_s, C_t) , denoted by $(L_i, C_j) \cong (L_s, C_t)$, if

$$B(L'_{i}, C'_{i}) = B(L_{s}, C_{t})$$

and

$$U(L_i) \leq U(L_s)$$

and

$$C'_{j1} \le C_{t1}, C'_{j2} = C_{t2}, C'_{j3} = C_{t3}$$

For each compatible assignment of model and scene primitive pairs, the reference point and rotated angle of the scene object with respect to the model can be computed from the geometrical entries given in the reference tabel. The transformation procedure to estimate the pose of the scene object is described in detail for individual types of primitive pairs as follows.

1).(L_i , L_j): a pair consisting of two line primitives L_i and L_j .

Let the equations of lines L_i and L_j be given by

$$\mathbf{L}_{i}: \mathbf{y} = \mathbf{m}_{1}\mathbf{x} + \mathbf{C}_{1}$$
$$\mathbf{L}_{j}: \mathbf{y} = \mathbf{m}_{2}\mathbf{x} + \mathbf{C}_{2}$$

Let $(L_i, L_j) \cong (L_s, L_t)$, and d_1, d_2 be the corresponding geometrical entries of (L_s, L_t) given in the fourth colmn of the reference table (see Table 1). The following two equations nust hold to solue for the coordinates of the scene reference point $(\overline{x}, \overline{y})$:

$$\left| m_{1} \overline{x} - \overline{y} + C_{1} \right| / (m_{1}^{2} + 1)^{\frac{1}{2}} = d_{1}$$
(9.a)

$$|\mathbf{m}_2 \overline{\mathbf{x}} - \overline{\mathbf{y}} + \mathbf{C}_2| / (\mathbf{m}_2^2 + 1)^{\frac{1}{2}} = \mathbf{d}_2$$

(9.b)

Solving egs 9.a ad 9.b, we obtain four possible reference points $(\overline{x}_k, \overline{y}_k)$, k=1,2,3,4 (see figure 6.a):

$$\overline{\mathbf{x}}_1 = (\mathbf{a} - \mathbf{b} - \mathbf{c}_1 + \mathbf{c}_2) / \mathbf{c}$$
, $\overline{\mathbf{y}}_1 = (\mathbf{m}_1 \mathbf{c}_2 - \mathbf{m}_2 \mathbf{c}_1 - \mathbf{m}_1 \mathbf{b} + \mathbf{m}_2 \mathbf{a}) / \mathbf{c}$

or

$$\overline{x}_2 = (a+b-c_1+c_2)/c$$
, $\overline{y}_2 = (m_1c_2-m_2c_1+m_1b+m_2a)/c$

or

$$\overline{\mathbf{x}}_3 = (-\mathbf{a} - \mathbf{b} - \mathbf{c}_1 + \mathbf{c}_2) / \mathbf{c}$$
, $\overline{\mathbf{y}}_3 = (\mathbf{m}_1 \mathbf{c}_2 - \mathbf{m}_2 \mathbf{c}_1 - \mathbf{m}_1 \mathbf{b} + \mathbf{m}_2 \mathbf{a}) / \mathbf{c}$

or

$$\overline{x}_4 = (-a + b - c_1 + c_2) / c$$
, $\overline{y}_4 = (m_1 c_2 - m_2 c_1 + m_1 b + m_2 a) / c$

where
$$a=d_1 \cdot (m_1^2 + 1)^{\frac{1}{2}}$$
,
 $b=d_2 \cdot (m_2^2 + 1)^{\frac{1}{2}}$
 $c=m_1 - m_2$

The retated angle q_r between the model nad the scene object is determined by two directional vector, one connecting the model reference point (x^*, y^*) and the intersection point $(\mathbf{r}_{x} \mathbf{r}_{y})$ of lines L_{s} and L_{t} , and the other one connecting the scene reference point (\bar{x}, \bar{y}) and the intersection point (o_x, o_y) of lines L_i and L_j (see Figure 6.b). Therefere,

$$\boldsymbol{q}_r = \cos^{-1} \frac{V_m \cdot V_o}{|V_m| \cdot |V_o|}$$

,

where $V_m = (r_x - x^*, r_y - y^*)$

$$V_o = (o_x - \overline{x}, o_y - \overline{y})$$

2). (C_i, C_j) : a pair consisting of two circle primitives C_i and C_j . Let the equations of circles C_i and C_j be given by

$$C'_{i}:(x - x_{1})^{2} + (y - y_{1})^{2} = r_{1}^{2}$$

$$C'_{j}:(x - x_{2})^{2} + (y - y_{2})^{2} = r_{2}^{2}$$

Let $(C_i, C_j) \cong (L_s, C_t)$, and d_1, d_2, q_1 and q_2 be the corresponding geometrical entries of (C_s, C_t) in the reference table. The following equations must hold to solve for the scene reference point $(\overline{x}, \overline{y})$:

$$(\overline{x} - x_1)^2 + (\overline{y} - y_1)^2 = d_1^2$$
(10.a)

$$(\overline{x} - x_2)^2 + (\overline{y} - y_2)^2 = d_2^2$$
(10.b)

$$(\overline{x} - x_1) \cdot (\overline{x} - x_2) + (\overline{y} - y_1)(\overline{y} - y_2) = d_1 d_2 \cos(\mathbf{p} - \mathbf{q}_1 - \mathbf{q}_2)$$
(10.c)

Solving eqs. 10.a, 10.b and 10.c, we gain two possible reference points $(\overline{x}_1, \overline{y}_1)$ and $(\overline{x}_2, \overline{y}_2)$ (see Figure 7.a):

$$\overline{x}_1 = a - m\overline{y}_1$$
, $\overline{y}_1 = (-b + [b^2 - 4(m^2 + 1) \cdot c]^{1/2}/2(m^2 + 1))$

or

$$\overline{x}_{2} = a - m\overline{y}_{2} , \ \overline{y}_{2} = (-b + [b^{2} - 4(m^{2} + 1) \cdot c]^{\frac{1}{2}}/2(m^{2} + 1)$$
$$m = (y_{1} - y_{2})/(x_{1} - x_{2})$$

where

$$a = \frac{(d_1^2 - d_2^2) + (y_2^2 - y_1^2)}{2(x_2 - x_1)} + \frac{x_1 + x_2}{2}$$

$$b = m(x_1 + x_2 - 2a) - (y_1 + y_2)$$

$$c = a(a - x_1 - x_2) + x_1x_2 + y_1y_2 - d_1d_2\cos(\mathbf{p} - \mathbf{q}_1 - \mathbf{q}_2)$$

The rotated angle q_r between the model and scene object is determined by two directional vectors, one connecting the model reference point (x^*, y^*) and the midpoint (r_x, r_y) between the centers of circles c_s and c_t , and the other one connecting the scene reference point $(\overline{x}, \overline{y})$, and the mid-point (o_x, o_y) between the centers of circles c_i and c_j (see Figure 7.b). Hence,

$$\boldsymbol{q}_r = \cos^{-1} \frac{\boldsymbol{V}_m \cdot \boldsymbol{V}_o}{|\boldsymbol{V}_m| |\boldsymbol{V}_o|}$$

where $V_m = (r_x - x^*, r_y - y^*)$

$$V_o = (o_x - \overline{x}, o_y - \overline{y})$$

2). $(\vec{L_i}, \vec{C_j})$: a primitive pair consisting of a line $\vec{L_j}$ and a circle $\vec{C_j}$

Let the equations of line L_i and circle C_j be given by

$$L_i$$
: $y = mx + c$
 C_j : $(x - x_0)^2 + (y - y_0)^2 = r^2$

Let $(L_i, C_j) \cong (L_s, C_j)$, and d_1 , d_2 be the correspinding geometrical entries of (\mathbf{q}_1, c_t) in the reference tabel. The following equations must hold to solve for the scene reference point $(\overline{x}, \overline{y})$

$$|m\overline{x} - \overline{y} + c|/(m^{2} + 1)^{\frac{1}{2}} = d_{1}$$
(11.a)
$$(\overline{x} - x_{0})^{2} + (\overline{y} - y_{0})^{2} = d_{2}^{2}$$
(11.b)

Solving eqs. 11.a and 11.b, we obtain four possible reference points $(\overline{x}_k, \overline{y}_k)$, k=1,2,3,4 (see Figure 8.a):

$$\overline{x}_{1} = \frac{-e + [e^{2} - 4f(m^{2} + 1)]^{\frac{1}{2}}}{2(m^{2} + 1)} , \quad \overline{y}_{1} = m\overline{x}_{1} + c + q$$

or

$$\overline{x}_{2} = \frac{-e - [e^{2} - 4f(m^{2} + 1)]^{\frac{1}{2}}}{2(m^{2} + 1)} , \quad \overline{y}_{2} = m\overline{x}_{2} + c + q$$

or

$$\overline{x}_3 = \frac{-g + [g^2 - 4h(m^2 + 1)]^{\frac{1}{2}}}{2(m^2 + 1)} , \ \overline{y}_3 = m\overline{x}_3 + c - q$$

or

$$\overline{x}_4 = \frac{-g - [g^2 - 4h(m^2 + 1)]^{\frac{1}{2}}}{2(m^2 + 1)} , \ \overline{y}_4 = m\overline{x}_4 + c - q$$

where

$$e = 2(mc + mq - x_0 - my_0)$$

$$f = (c + q - y_0)^2 + x_0^2 - d_2^2$$

$$g = 2(mc - mq - x_0 - my_0)$$

$$h = (c - q - y_0)^2 + x_0^2 - d_2^2$$

 $q = d_1(m^2 + 1)^{1/2}$

The rotated angle q_r of the scene object is determined by two directional vectors, one connecting the model rederence point (x^*, y^*) and the center (L_s, r_y) of circle C_t , and the other one connecting the scene reference point $(\overline{x}, \overline{y})$ and the center (x_0, y_0) circle C_j (see Figure 8.b). Therefore,

$$\boldsymbol{q}_r = \cos^{-1} \frac{V_m \cdot V_o}{|V_m| |V_0|}$$

where

 $V_m = (r_x - x^*, r_y - y^*)$ $V_0 = (x_0 - \overline{x}, y_0 - \overline{y})$

Define a three dimensional accumulator array $A(\overline{x}, \overline{y}, q_r)$, where the parameters $(\overline{x}, \overline{y})$ and q_r are the possible reference point and rotated angle for a scene object, respectively. Each assignable primitive pair in the scene image generates possible locations of the parameter space $(\overline{x}, \overline{y}, q_r)$ and contributes as many votes as the entries in the reference table. A peak of vote counts indicates the evidence of a match and specifies the pose (a translation and rotation) of the scene object.

(a) possible reference points

(b) Directional vectors that determine the rotated angle

Figure 6. Evaluate the reference point and rotated angle for primitive pair $(\vec{L_i}, \vec{L_j})$

(a) possible reference points

(b) Directional vectors that determine the rotated angle

Figure 7. Evaluate the reference point and rotated angle for primitive pair (C_i, C_j)

(a) possible reference points

(b) Directional vectors that determine the rotated angle

Figure 8. Evaluate the reference point and rotated angle for primitive pair (L_i, C_j)

Recall that the GHT method uses single gradient information of edge points as the signature, and the gradients and displacement vectors in the reference table are rotation-dependent. However, the vote generation process described above uses the binary relation of a primitive pair as the signature to the reference table and utilizes unary properties of primitives as additional constraints to eliminate non-assignable primitive pairs in early voting. This will result in as few false alarms as possible. Furthermore, all entries in the proposed reference table are translation- and rotation invariant. The reference point and rotated angle of a scene object can be evaluated in

4. TOP-BOTTOM RELATION

Identifying the position, orientation as well as the top-bottom relation of overlapping objects is important for robotic assembly. In this paper, we propose a simple procedure to detect the object on the top (or on the bottom) from a gray-scale image by taking the fact that objects cast shadows on one another. Given fair intensity (not as dark as the shadow) of object surfaces, some points on the shadowed object along the overlapping boundaries have gray-level readings defferent from what they would have had if there were no shadowing. The pose $(\overline{x}, \overline{y}, q_r)$ estimated in the foregoing vote generation phase is used to transform the boundary points of the model object onto the scene image and then we trace the estimated boundary of the scene object under study to find the shadowed points. Due to the pose deviation resulting from the proposed algorithm and quantization, the estimated pose (\bar{x}, \bar{y}, q_r) may not generate a prefect boundary in the scene image. In order to accommodate the distance error between the actual boundary points of the scene object and the mapped boundary points from the model object, we select the minimal intensity of a point (the darkest one) in the neighbor, defined by a small $W \times W$ window, of each boundary point of the model mapped on the scene image.

Assume all object surfaces under study have identical, fair intensity, and the objects are placed on a black background. The object on the top has many shadowed points along the overlapping boundaries in the image, but the object on the bottom will not show any shadowed points along its overlapping boundaries, ass seen in Figure 9. The average gray value and the number of shadowed points on the overlapping boundaries can, therefore, provide the cue to determine the top-bottom relation of overlapping objects. The detailed recognition procedure is described as follows.

Let $P = \{p_i = (x_i, y_i), i = 1, 2, ..., n\}$ be the boundary points of a model object, and $P' = \{p'_i = (x'_i, y'_i), i = 1, 2, ..., n\}$ the mapped boundary points from the model image to the scene image. Here,

$$P_i' = \underbrace{\begin{array}{c} & & \\ &$$

 $(\overline{x}, \overline{y})$ and q_r used in the transformation matrix above are the reference point and rotated angle of the scene object, which are determined in the vote generation phase. (x^*, y^*) is the reference point of the model object.

Also, let w=2k+1 represent a window size. and

$$N_w(x_i, y_i) = \{ (x_{i+j}, y_{i+m}), -k \le j \le k, -k \le m \le k \}$$

respect the neighboring points of (x_i, y_i) defined in the $W \times W$ window.

Define,

$$g(x_i, y_i) = \min\{f(x, y), \forall (x, y) \in N_w(x_i, y_i)\}, i = 1, 2, ..., n$$

where f(x,y) denotes the gray level of a point at (x,y). $g(x_i, y_i)$ records the minimal gray value of a point (i.e., a most likely shadowed point) in the neighbor of (x_i, y_i) , and this is equivalent to performing a simple gray-scale erosion operation along the mapped boundary in the image.

A crude way to detect shadowing is to compare the average gray values for the points on the overlapping boundaries. Denote this method by Gray Level Mean m_g . Let T_B be the binary threshold value used to segment the objects from the background. In this research, we used the within-group variance method [26] to determine the threshold T_B automatically. The Gray Level Mean m_g is defined by

$$\mathbf{m}_{g} = \sum_{g(x_{i}, y_{i}) \in D} g(x_{i}, y_{i}) / \sum_{g(x_{i}, y_{i}) \in D} 1$$

where D={ $g(x_i, y_i) | g(x_i, y_i) > T_B, i = 1, 2, ..., n$ }.

The denominator in eq. 12 is the number of elements in D. Recall that the object under study is placed on a black background. If $g(x_i, y_i) \leq T_B$, we assume that (x_i, y_i) is a boundary point connecting to the background. Therefore, the set D only includes those points on the overlapping boundaries. Given the object surfaces that have larger gray value than that of shadowed points, the object on the top will have smaller Gray Level Mean \mathbf{m}_{g} , compared to the \mathbf{m}_{g} of the object on the bottom. Let \mathbf{m}_{g1} and \mathbf{m}_{g2} represent the Gray Level Means of two overlapping objects 1 and 2, respectively. We conclude that if $\mathbf{m}_{g1} < \mathbf{m}_{g2}$, then object 1 is on the top and object 2 on the bottom, and vice versa.

Alternatively, we may accumulate the number of shadowed points with respect to the number of overlapping boundary points. Denote this method by Shadowed Points Ratio R_g value than that of the object on the bottom. Let $\overline{\mathbf{m}}$ represent the average gray value of object surfaces in the scene image. It is evaluated by f(x,y) is the

$$\overline{\boldsymbol{m}} = \sum_{f(x,y)\in s} f(x,y) / \sum_{f(x,y)\in s} 1$$
(13)
where $S = \{f(x,y) | f(x,y) > T_B, \forall (x,y) \},$

gray value at (x,y) in the scene image T_B is the binary threshold value as defined previously.

The shadowed POint Ratio R_g is defined by

$$R_{g} = \sum_{g(x_{i}, y_{i}) \in D} \frac{1}{2} / \sum_{g(x_{i}, y_{i}) \in D} \frac{1}{2}$$
(14)

where

$$D' = \{g(x_i', y_i') | T_B < g(x_i', y_i') < \overline{\mathbf{m}}, i = 1, 2, ..., n\}$$
$$D = \{g(x_i', y_i') | g(x_i', y_i') > T_B, i = 1, 2, ..., n\}$$

For the object on the top,the points on the overlapping boundaries are shaded and have lower gray value than the average gray value of object surfaces \overline{m} . Number of elements in $D^{'}$ is, therefore,close to the one in D, i.e., the shadowed Points Ratio R_g is approximate to 1. Let R_{g1} and R_{g2} represent the shadowed Points ratios of two overlapping objects 1 and 2, respectively. We conclude that if $R_{g1} > R_{g2}$, then object 1 is on the top and object 2 on the bottom, and vice versa. The performance of the Gray Level Mean and shadowed Points Ratio is evaluated in the section that follows.

5.EXPERIMENTAL RESULTS

All the experiments were implemented on a PC-486 personal computer using C. The image is 512×480 pixels wide and has 8-bit gray levels ranging from 0 to 255. The objects used in the experiments were also tested using the GHT method. However, the GHT method failed to detect and locate the object shapes involving only linear and circular segments in a reliable way, as seen in Figures 1.a and 1.b.

Figure 10.a shows the image of a prototype workpart, denoted by workpart A, used to create the reference table. Figure 10.b presents the line and circle primitives extracted from the outmost boundary of the workpart. The cross symbol shown in the figure marks the location of the center of a circle primitive. Figures 11.a, 11.b, and 11.c show on occluded version of workpart A in a scene image, the primitives extracted from the scene image, and the result that superimposes workpart A on the scene image, respectively. Figures 11.d, 11.e and 11.f show another occluded version and results of workpart A in the scene image. Note that even the main portion of workpart A is overlapped, the location of the workpart in the scene image is correctly evaluated.

In order to evaluate the reliability of the proposed Hough-clustering algorithm that determines the pose of an occluded object in the image and the performance of the Gray Level Mean and Shadowed Points Ratio that determine the top-bottom relation of overlapping objects, we used four workparts as shown in Figure 12 for testing. Denote the workparts in Figures 12.a, 12.b, 12.c and 12.d as workparts is 2mm. The average gray values of workpart surfaces and the background are 171.4 and 32.5, respectively. The light source of the image setting is provided by two fluorescent lights that illuminate the workparts under tset from upper-right and upper-left at the height of the camera. Five sets of workpart pairs are selected from workparts A, B, C and D to generate overlapping objects. The first set includes a workpart A and a workpart B, denoted by S(A, B). The remaining four sets are S(A,C), S(A,D), S(B,C) and S(C,D). We created 12 overlapping samples for each set of workpart pairs, six of which had one workpart on the top and the remaining six had

the other workpart on the top. The objects on the top were rotated by a random angle in between 0 and 2p, and were overlapped by a random coverage in between 30% and 70% in terms of the invisible portion of boundary length. A total of 60 samples were created from the five sets of workpart pairs. A 9×9 window that is used to search for the shadowed points along object boundaries is used in the experiment. Figure 13 demonstrates one sample of the set S(B,C). Figure 13.a shows the overlapping workparts B and c in the image. Figures 13.b and 13.c show the boundaries of workpart B and C, respectively, mapped on the scene images. Figure 13.d shows the result that workpart B is identified as the object on the top.

The pose of each workpart in the 60 samples is detected correctly, where every boundary point of the model workparts ins mapped within the 9×9 window from its actual location in the scene image. This result shows that the proposed algorithm is robust and reliabel to detect the workparts involving linear and circular segments in a complex image.

The averages of Gray Level Mean \mathbf{m}_{g} and Shadowed Points ratio R_{g} for each set of workpart pairs are summarized in Table 2. $\overline{\mathbf{m}}_{g,t}$ and $\overline{\mathbf{m}}_{g,b}$ shown in the table represent the averages of Gray Level Mean for the objects on the top and the objects on the bottom, respectively. Similarly, $\overline{R}_{e,b}$ and $\overline{R}_{e,t}$ denote the averages of Shadowed Points Ratio for the objects on the top and the objects on the bottom, respectively. It is observed from Table 2 that the average of Fray Level Mean for the objects on the bottom is about the same as the average gray value of object surfaces 171.4. The difference between $\overline{\mathbf{m}}_{g,t}$ and $\overline{\mathbf{m}}_{g,b}$ is not significant. Six of the 60 samples, the crude Gray LEvel Mean method failed to detect the correct top-bottom relations of overlapping workparts, i.e., only 90% of correct recognition is obtained. However, the Shadowed Points Ratio method has 100% correct recognition results for the 60 samples. It generates correct solution even the shadowed portion is hardly visible from the computer monitor by human eyes such as the sample shown in Figure 13.a. The objects on the top have their Shadowed Points Ratios close to 1. The defference between $\overline{R}_{g,b}$ and $\overline{R}_{g,t}$ is very significant, as seen in table 2. Columns 1 and 2 of Table 3 show the ratios of $\overline{\mathbf{m}}_{e,b}/\overline{\mathbf{m}}_{e,t}$ and $\overline{R}_{e,b}/\overline{R}_{e,t}$, respectively. The larger the ratio value, the more reliable for determining the top-bottom relation of overlapping objects. The ratio of $\overline{\mathbf{m}}_{g,b}/\overline{\mathbf{m}}_{g,t}$ is close to 1, which means the Gray Level Mean is a poor and unreliable measure to discriminate the difference. The average ratio of $\overline{R}_{g,b}/\overline{R}_{g,t}$ is two- or three-fold and ,therefore, the Shadowed Points Ratio is a robust and reliable method to determine the top-bottom for the objects with fair intensity readings on the surfaces.

Figures 14.a and 14.b show the images of two aluminum casting parts with an average gray value of 128. The shapes of these two parts are mirror-symmetric, and invalve linear, circular and some non-circular curves. The parts have an average depth of 10 mm, and a maximal depth of 20 mm. Figures 14.c and 14.d show the extracted primitives of the two casting parts. Figures 14.e and 14.f consist of the overlapping parts and the corresponding primitives extracted, respectively. The mapped boundaries of the two model parts are highlighted in the scene image, as shown in Figure 14.g. The Shadowed Points Ratios for the part on the top and the one on the bottom are 0.97 and 0.67, respectively, and are sufficient to discriminate the difference. The estimated locations of the aluminum casting parts are not as precise as the workparts A, B, C and D due to the depth of the casting parts and non-circular curves of shapes.

6.CONCLUSIONS

In this paper, object shapes are broken down into fragments, each representing either a line or a circular arc. These low level primitives of lines and circles are extracted in a reliabel way through dominant point detection, least-squares line/circle fitting followed by fragment partitioning. Recognition of partially occluded industrial parts is transformed to the assignment of primitive paris of secen objects to those of model objects. The binary relation of a primitive pair is used as a signature to the reference table of a medel object, and unary properties of primitives provide additional constraints to eliminate false matching. The reference point and rotated angle of a scene object are then computed from each assignable primitive pari and contribute a vote with such parameter values. All entries in the reference table are translation and rotation-invariant and, therefore, the vote generation process can be performed in one pass. Due to lighting, reflection, quantization, roundness of curved segments and preprocessin procedures (such as binary thresholding), the primitives extracted from an object shape may not be consistent from image to image. If more dominant points of a scene object are detected in the feature extraction phase, the resulting extra primitives can be treated as the segments of other objexts. If less dominant points of the scene object are detected, the missing primitives can be treated as the segments covered by other objects. A primitive with misestimated parameters values can be interpreted as either a segment of other objects or a missing segment of the scene object. This makes the proposed algorithm less rely on the results of feature extraction. Therefore, the proposed Hough-clustering procedure is useful for identifying and locating the occluded workparts that have no salient features on the boundaries or have their salient features completely covered by other objects.

This paper also present the methods to determine the top-bottom relation of overlapping objects in gray-scale image by tracing the shadowed points along the overlapping boundaries. The Shadowed Points Ratio that measures the number of shadowed points with respect to the number of overlapping boundary points is robust and reliable for objects with bright surfaces on the black background. The proposed methods were also used to examine dark objects (on average gray value of 75) on the white background. The Gray Level Mean and Shadowed Points Ratio performed unreliably (around 80% correctness) for such dark objects. This restricts the Shadowed Points Ratio in its current form to be useful for detecting workparts with bright surfaces.

R_{g}					
set Measure	1 S(A,B)	2 S(A,C)	3 S(A,D)	4 S(B,C)	5 S(C,D)
$\overline{\mathbf{m}}_{g,t}$	162.7	162.1	158.4	164.4	160.1
$\overline{\mathbf{m}}_{g,b}$	170.3	165.2	166.7	171.3	167.4
$\overline{R}_{g,t}$	0.97	0.99	0.96	0.97	0.98
$\overline{R}_{g,b}$	0.41	0.34	0.36	0.30	0.34

Table 2. The results of Gray Level Mean \mathbf{m}_{g} and Shadowed Points Ratio

Table 3. Ratios of Gray Level Mean (Bottom vs. Top) and Shadowed Points Ratio (Top vs. Bottom)

Set	1	2	3	4	5	
Ratio	S(A,B)	S(A,C)	S(A,D)	S(B,C)	S(C,D)	Average
$\overline{\boldsymbol{m}}_{g,b}/\overline{\boldsymbol{m}}_{g,t}$	1.05					
		1.02	1.05	1.04	1.05	1.04
$\overline{R}_{g,b}/\overline{R}_{g,t}$	2.37					
		2.91	2.67	3.23	2.88	2.81

Figure 9. Shadowed points on overlapping boundaries

(a) Original image of workpart A workpart A

(b) Primitives of

Figure 10. The image and primitive of a model workpart A

(a) Image of overlapping workparts workparts

(b)primitives extracted extracted

(d) Image of overlapping

(e)primitives

(f)Location of

(c) Location of workpart A

workpart A

Figure 11. Two occluded versions of workpart A

(a) Model workpart A workpart B

(b) Model

(c) Model workpart C workpart D

(d) Model

Figure 12. Four test workparts

(a) Overlapping workparts B and C workpart B

(b) Estimated boundary of

(c) Estimated boundary of workpart C

(d) Identifying workpart B as the object on the top

Figure 13. Top-bottom relation of overlapping workparts B and C

(a) Aluminum casting part A casting part B

(b) Aluminum

(c) Primitives of part A of part B

(d) Primitives

Figure 14. Detecting and locating two aluminum casting parts

(e) Overlapping parts A and B overlapping parts

(f) Primitives of

(g) The model boundaries mapped on the image

Figure 14. (continued)

REFERENCE

- T. Pavlidis and S. E. Horowitz, Segmentation of plane curves, IEEE trans. Comput., 23, 860-870, 1974.
- 2. L. P. Cordella, An O(N) algorithm for polygonal approximation, Pattern recognition Letters, 3, 93-97, 1985.
- B. Kr. Ray and K. S. Ray. A new approach to polygonal approximation, Pattern Recognition Letters, 12, 229-234, 1991.
- K. Wall and P. -E. Danielsson, A fast sequential method for polygonal approximation of digital curves, Computer Vison, Graphics, and Image Processing, 28, 220-227, 1984.
- 5. M. -H. Han and D. Jang, The use of maximum curvature points for the recognition of partially occluded objects, Pattern Recognition, 23, 21-33, 1990.
- 6. M. W. Koch and R. L. Kashyap, Using polygons to recognize and locate partially occluded objects, IEEE Trans. Pattern Anal. Mach. Intell., PAMI-9,483-494, 1987
- 7. E. H. Bouyakhf, An approach for recognizing and locating overlapping industrial parts, Int. J. Pattern Recognition and Artificial Intell., 2, 673-689, 1988.
- N. Ayache and O. D. Faugeras, HYPER: A new approach for the recognition and positioning of two-dimensional objects, IEEE Trans. Pattern Anal. Mach. Intell., PAMI-8, 44-54, 1986.
- J.T. Schwartz and M. Sharir, Identification of partially obscured object in two and three dimensions by matching noisy characteristic curves, Int. J. Robotics research, 6,29-44, 1987.
- S. Chaudhury, A. Acharyya and S. Subramanian, Recognition of occluded objects with heuristic search, Pattern Recognition, 23, 617-635, 1990.

- W. E. L. Grimson and T. Lozano-Perez, Localizing overlapping parts by searching the interpretation tree, IEEE Trans. Pattern Anal. Mach. Intell., PAMI-9, 469-482, 1987.
- C. C. Lin and R. Chellappa, Classification of partial 2-D shapes using Fourier descriptors, IEEE Trans. Pattern Anal. Mach. Intell., PAMI-9, 686-690, 1987.
- C. T. Zahn and R. Z. Roskies, Fourier descriptors for plane closed curves, IEEE Trans. Comput., C-21, 269-281, 1972.
- E. Person and K. -S. Fu, Shape discrimination using Fourier descriptors, IEEE Trans. Systems, Man and Cybernetics, SMC-7, 170-179, 1977.
- 15. O. R. Mitchell and T. A. Grogan, Global and partial shape discrimination for computer vision, Opt, Eng., 23, 484-491, 1974.
- J. W. Gorman, O. R. Mitchell and F. P. Kuhl, Partial shape recognition using dynamic programming, IEEE Trans. Pattern Anal. Mach. Intell., 10, 257-266, 1988.
- W. A. Perkins, A model based vision system for industrial parts, IEEE Trans. Comput., C-27, 126-143, 1978.
- J. L. Turney, T. N. Mudge and R. A. Volz, Recognizing partially occluded parts, IEEE Trans. Pattern Anal. Mach. Intell., PAMI-7, 410-421, 1985.
- 19. E. R. Davies, Occlusion analysis for object