Detection and Recognition of Non-Occluded Objects using Signature Map

Sangbum Park, Youngjoon Han, Hernsoo Hahn
College of Information Technology
Soongsil University
Sangdo-5Dong, Dongjak-Gu, Seoul
Republic of KOREA
http://visionlab.ssu.ac.kr

Abstract: - For constructing a flexible bin picking system where parts can be provided with arbitrarily stacked in a workspace, detection and recognition of non-occluded objects are essential process. For implementing such process, this paper proposes a new algorithm which determines whether an object is occluded or not and at the same time which object it is in the DB of object models. It is based on a signature map which is constructed by detecting the objects in an input image and drawing the signatures of the whole image with reference to individual objects. Thus the number of signature maps is equal to the number of the objects detected in the input image. A signature map shows the outer contour and inside edge features. Occlusion by other objects appears as distortions in the outer contour of the signature map. The inside edge features are used for discerning the objects having the same outer contour by different inside shape. To make the manipulator pick up a selected part, a pose estimation method for elliptical objects is also proposed. The performance of the proposed algorithm has been tested with the task of picking the top or non-overlapped object from a stack of arbitrarily located objects. In the experiment, a recognition rate of 98% has been achieved.

Key-Words: - Bin-picking, Signature map, Contour tracing, Object detection, Object recognition

1 Introduction

In many automatic assembly systems, the known parts are palletized and thus recognition process is not needed. However, if the parts can be delivered with unsorted, the palletizing devices are not needed but a complex recognition and pose estimation processes are required. For this reason, how efficiently represent 3D objects using the features that can be easily extracted has drawn a great attention from many researchers working for developing automatic assembly systems. However, detection of the top object on a stack of arbitrarily located objects still remains as the core problem in construction of the bin picking system without pallets.[1,2]

Many researchers have proposed different solutions for bin-picking problem that a manipulator grasps an object located in a pallet constraining the pose of the object or on a conveyer belt without constraints. To exactly locate the object's pose, an object representation scheme which may provide such information from its 2D image should be developed. Most object representation schemes proposed for this purpose construct have been based on the shape from contour method.[3,4] The features extracted from the 2D image are used for matching with the object models in the recognition module. When objects are overlapped in the image, the approaches based on the shape from contour method generate many problems. One of them is that it requires an excessive computation time since all models in the database should be compared with a test object. The other one is that incorrect matching results may occur if some parts of the object features are occluded.[5] The approaches using T-junction analysis for recognizing overlapped objects are suitable for discerning the objects having the linear contours. When contours are curved ones then it does not work appropriately.[6]

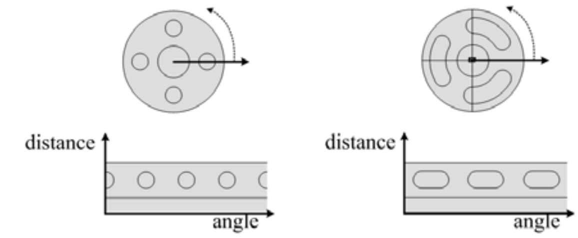
For solving these problems, this paper propose a new signature map technique for extracting geometrical features of 3D objects using a 2D image to detect and recognize the top object on a stack of 3D objects. The proposed method consists of two parts: the generation of signature maps of an input image and the detection and localization of the top object. To generate the signature map, we use an effective and robust edge detector to reduce the noise or to reconstruct the lost portion of edges. For this reason, a contour tracing method which is based on a cubic spline interpolation is used in the preprocessing module. It is in order to include curved edges as well as linear edges in the application domain of the proposed method. Based on the proposed representation scheme, this paper proposes a new algorithm to find overlapped object using bending point information. The algorithm memorizes the number of bending points of the object models and compares it with that of a test object. If that of a test object is bigger than that of the object model, then the object model is selected as the model of the test object. Finally, the depth information included on the signature map is extracted for estimating the pose of the object. The depth information of the objects in the signature map is generated from contours of object. This signature map also describes the object features located inside the boundary. This information is very useful to discern the object having the same outer boundary but different inner shapes.

The rest of this paper is constructed as follows: Section 2 explains how to generate a signature map and how to use it for object recognition. Section 3 illustrates how to detect overlap of objects and select the top object. Section 4 shows how to determine the pose of a detected object and Section 5 illustrates the performance of the proposed algorithm by the experiments.

2 Object Recognition using Signature Map

2.1 Construction of a signature map

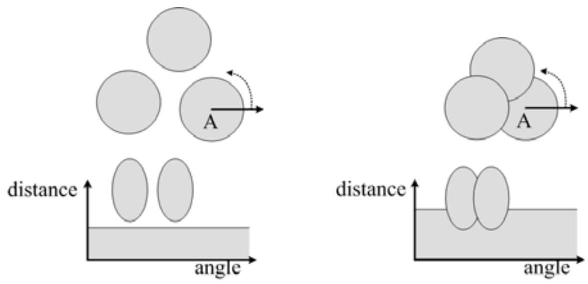
The signature of a 3D object is generated from its 2D image. It draws the distances of the boundary points from the center of the object in a 2D plane by increasing the angle to the counter-clock direction as shown in Fig. 1. Thus the accuracy of a signature depends on how clearly the boundary is detected. To extract the boundary for this purpose, contour tracing or boundary tracking method is used in general. These methods are basically edge tracking techniques. They search all neighboring pixels around the reference edge to test the connectivity with it. Thus, they are very sensitive to noise or disconnection of an edge since such things generate uncertainty in determining the direction to follow. To solve this problem, the signature map proposed in this paper describes the internal and external edge features. Thus, even in the case where there are objects having the same outer contour shape but different internal edge features as shown in Fig. 1, the proposed algorithm can separate them clearly. Fig. 1(a) shows the case of having circular and Fig. 1(b) does the case of having rectangular carves, inside the outer boundary. Although both objects have the same outer contour in signature map, the proposed method easily discerns one object from another using the internal edge features.



(a)Circular holes included (b)Rectangular carves included

Fig. 1 The signature maps for the objects having the same outer contour but different inside edge features

When multiple objects are given in an image, the signature maps corresponding to individual objects are generated as shown in Fig. 2. If an object is not occluded by any other object, then the signature map generated with reference to the center of this object may generate the same signature as that of the object model as shown in Fig. 2(a). However, if an object is occluded by other objects as shown in Fig. 2(b), then the signature map generated with reference to the object becomes distorted and it also shows which object is occluding the reference object.



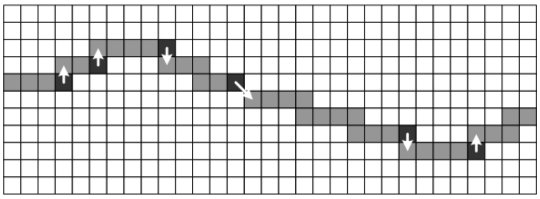
(a)Not-occluded case (b)Occluded case

Fig. 2 The signature maps in the cases of occluded and not-occluded

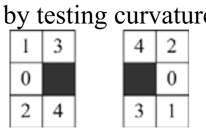
If a signature map is used, then a recognition task can be defined as a process of testing if the signature of a test object is matching with which object model's signature. In order to simplify the signature comparison process, the proposed algorithm uses the signature features such as the number of poles(N_p), the angles between the poles(θ_i), the number of contours(L) and the distance between contours(d_i).

2.2 Feature extraction from a signature map

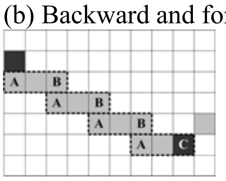
The signature basically shows the shape of an object. Therefore, a pole in a signature corresponds to a convex vertex of an object in a 2D image and the number of poles tells the number of corners. The pole is detected analyzing the variation of signature as shown in Fig. 3(a). For a given signature, a chain code is used to find a pole. However, a conventional chain code representation is easily contaminated by noise[7]. Therefore, this paper uses a false pole elimination technique using the forward and backward detection window given in Fig. 3(b).

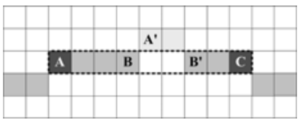


(a) Pole detection by testing curvature variation



(b) Backward and forward detection window





(c) Pole candidate detection (d) False pole elemination

Fig. 3 Pole extraction from a signature

Let's assume that a signature as shown in Fig. 3(c) is given and the tracing has begun from the leftmost top pixel shaped in the figure to generate an index sequence 30030030030030. In the sequence, the pixel marked as B are considered as a pole candidate since the index change occurs at the pixels. However a new edge is not generated from pixel B and the pattern is repeated, they are considered as a part of the edge and continues until pixel C. Since a new index 2 is appeared at pixel C, it is considered as a pole candidate and a new tracing starts from it. If a new sequence is long enough to be considered as an edge, then pixel C is determined as a pole pixel. However, if the length of a new edge is too short as shown in Fig. 3(d) then it is considered as a part of an edge. Backward detection window is used to find a branching point. When the poles are detected in the signature map, the angle between the neighboring poles is same as the distance between poles in the signature map and it should be larger than a threshold

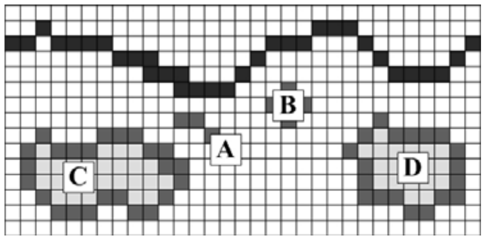


Fig. 4 The internal features of the object

Differently from the outer contour, the detection of inside edge features requires more attention to eliminate false features generated by noise. Shadows or reflection caused by light sources may generate noisy features. To eliminate such features, every feature is tested if it has the area larger than a threshold. If it is smaller than the threshold, then it is considered as a noisy region.

2.3 Model registration and decision rule for matching

The signature map of each model(M_i) is registered in the database as a feature vector, M_i={contour feature(F_{out}), internal feature(F_{in})}. F_{out} and F_{in} are defined as F_{out}={the number of pole(N_p), the angle between poles(θ_i), the number of contour(L), the distance between contours(d_i)} and $F_{in}=\{f_{in}|f^i_{in}=$ {Shape(S_i), Position(P_i), average intensity (I_{avg})}}, respectively.

The matching function T in Eq.(1) is utilized to decide for if two feature vectors A and B are coincide. The Boolean function T returns TRUE(1) if the difference between two features A and B is smaller than the error threshold(α). Otherwise, it returns FALSE(0). The error threshold (α) is decided by the experiments.

$$T\left(A,B,\alpha\right)=BOOL\left(0<\frac{A\times\alpha}{\left\|A-B\right\|}\right) \qquad (1)$$

Using T, the decision rule comparing two feature vectors is defined as Eq. (2) which tests the match of individual elements of the feature vectors.

$$T_{out}^{cm}=T\left(\theta_m,\theta_c,\alpha\right)\cdot T\left(\mu_m,\mu_c,\alpha\right)\cdot T\left(L_m,L_c,\alpha\right) \cdot T\left(d_m,d_c,\alpha\right) \qquad (2)$$

In Eq. (2), the subscript m and c are the indices of the model and test objects. If the number of model objects matched with the test object is more than 2, the internal features are compared using Eq.(3).

$$T_{in}^{cm}=T\left(C(S_m),C(S_c),\alpha\right)\cdot T\left(p_m,p_c,\alpha\right) \cdot T\left(I_m,I_c,\alpha\right) \qquad (3)$$

In Eq. (3), C is the function of testing how close to a circle, defined as Eq. (4).

$$C(S) = \frac{n_p^2}{\pi\left(L^s/2\right)^2} \tag{4}$$

where n_p^2 is the number of pixel in the area S, L^s is the distance of long axis

3 Overlap Decision and Recognition

3.1 Contour tracing in the signature map

Since the contour of a circular object does not change sharply, overlap of objects can be easily detected by contour tracing. Fig. 8 illustrates the contour detecting method used in this paper when the contour is contaminated by noise or disconnection. If a contour has a disconnected part, the algorithm estimates the direction of contour and calculates connective strength to apply the spline interpolation.

This spline interpolation assumes that the contour is differentiable at every points and its curvature does not change a lot. Therefore, the algorithm select those points which are rendering a curvature smaller than a threshold among the neighboring points. For example, in Fig. 5, when disconnected point is detected at the n-th pixel in section A, the algorithm searches the nearest edge in the shaded region. If found, then a sectional polynomial-function which is differentiable in the section $[x_0, x_n]$ is inserted to connect the contour given in section A and the newly found edge. The connected edge region is defined as a new section A'. In the same way, this contour tracing continues until it cannot find the connected edges any more. Section B' shows the reconstructed region by the interpolation process

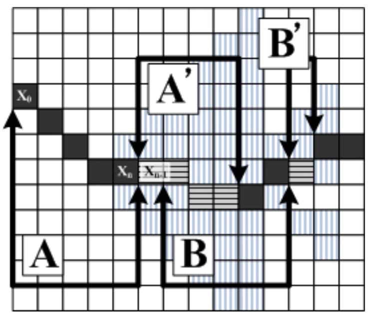


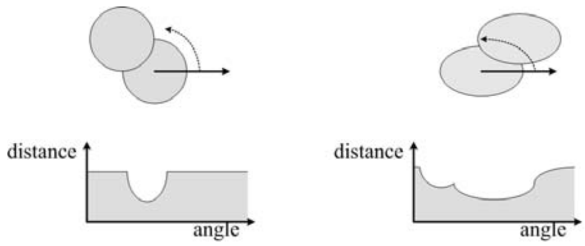
Fig. 5 The contour tracing method

3.2 Overlap estimation

If the target object is overlapped, the signature becomes distorted as shown in Fig. 6. This distortion can be detected by searching the points where the curvature changes abruptly. In the process of contour tracing, a candidate for an occlusion point is selected by testing the following conditions given Eq. (5) which tests the curvature variation range.

$$\begin{array}{ccccc} d_{\min}^2 & \le & \|p-p^+\| & \le & d_{\max}^2 \\ d_{\min}^2 & \le & \|p-p^-\| & \le & d_{\max}^2 \\ 0 & \le & |\alpha| & \le & \alpha_{\max} \end{array} \tag{5}$$

In Eq. (5), $\| p - p^{\pm} \|$ is a linear distance between two points, $\alpha \in [-\pi, \pi]$ is the angle between p and $p^{\pm}$, d_{min}, d_{max} and α_{max} are the reference values to be used as the thresholds. $|\alpha(p)|$ at an occlusion point can be represented as the inclination between the edges, $\beta(p)=\pi - |\alpha(p)|$. Since d_{min} and d_{max} define the angle difference, they are determined to be inversely proportional to the object size



(a) Circular objects (b) Ellipse objects

Fig. 6 Distorted signature by occluding objects

4 Pose Estimation of The Top Object

To control a manipulator to pick up the target object, its pose must be estimated with error smaller than the tolerable error. Fig. 7 shows configuration of the pose estimation process where O is the world frame, C is the camera frame, and T is the object frame.

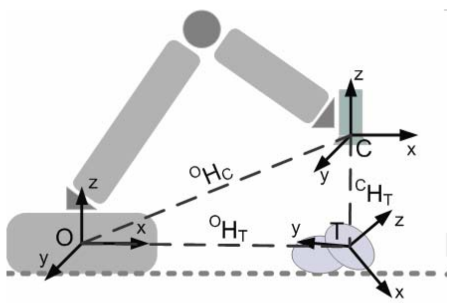


Fig. 7 The configuration of the pose estimation process

To control the manipulator, the relation, ${}^{o}H_{c}$, between the world frame and the camera frame should be known. Once ${}^{o}H_{c}$ is known, then the object frame in the world frame can be extracted through measuring the object frame with reference to the camera frame. Then, ${}^{o}H_{c}$ can be expressed by Eq. (6).

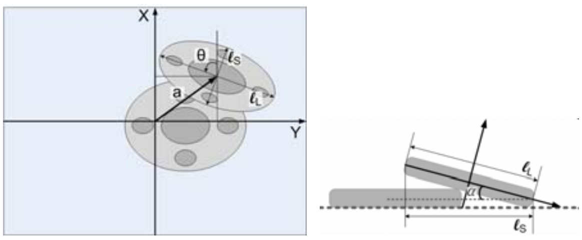
$$
{}^{O}H_{T} = {}^{O}H_{C} \cdot {}^{C}H_{T}
$$

$$
{}^{O}H_{T} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & a \\ \sin\theta\cos\alpha & \cos\theta\cos\alpha & -\sin\alpha & -\sin\alpha d \\ \sin\theta\sin\alpha & \cos\theta\sin\alpha & \cos\alpha & \cos\alpha d \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6)
$$

Under these relationships among the frames, as shown in Fig. 8(a), the object pose can be represented by two rotation angles about X axis and Z axis and translation in the (X,Y) plane. The position of the object frame with reference to the camera image frame (x,y) is determined first. After locating the camera so that the origin of the camera image frame and the object center, the camera is rotated about Z axis of the camera image frame by θ degrees to coincide the X axes of the camera image frame and the object frame. Then it is rotated once more about the X axis of the camera image frame by α degrees to coincide the Z axes of the camera image frame and the object frame. α can be by Eq.(7) using the lengths of the circular object measured from Fig. 8(b).

$$
\alpha = \cos^{-1}\frac{l_L}{l_S} \quad (7)
$$

By applying these pose parameters to Eq. (6), the pose of the object in the world frame can be extracted.



(a) Pose parameters (b) Definition of α

Fig. 8 Definition of pose parameters

5 Experimental Results

To evaluate the performance of the proposed algorithm, the algorithm has been implemented on the 6-DOF manipulator having a single camera mounted on the gripper. It is asked to detect and recognize and pick up the top or non-occluded object among arbitrarily stacked multiple objects. Six different types of objects are used to make a stack of objects. Their shapes and signature features are given in Table 1.

5.1 Recognition and detection of the top object

To test the recognition accuracy, each object model is located at an arbitrary position in a 2D plane and the algorithm is applied. This test has been repeated 100 times per each object model. Since the result will be same if an object is located separately from other objects and it is segmented successfully, the recognition performance is tested in this way. The experimental results testing the recognition process are summarized in Table 2. The process extracts the signature map with reference to the test object and matches with those of the 6 object models, to find the best matching object model.

Table1 The DB of the 6 object models.

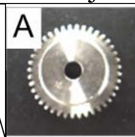
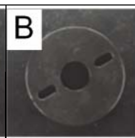
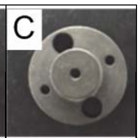
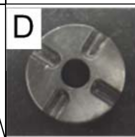
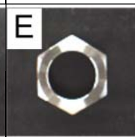
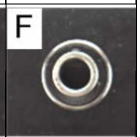
Type / Features				
External Features	N_p/θ_i	48/8	0/0	0/0
	L/d_i	2/6,36	2/10,38	3/3,20,38
Internal Features	S_i	0.0	0.6	1.0
	p_i	0	180	90
	I_{avg}	0	7	8
Type / Features				
External Features	N_p/θ_i	0/0	6/60	0/0
	L/d_i	2/10,38	2/18,27	4/12,17,25,27
Internal Features	S_i	0.35	0.0	0.0
	p_i	90	0	0
	I_{avg}	68	0	0

Table 2 The recognition rate [%].

Type / Decision	A	B	C	D	E	F
Correct	99	98	100	98	99	99
Incorrect	0	1	0	1	0	0
Not-available	1	0	0	1	1	1

As can be noticed, the recognition rate depends on the shape of a test object. In the cases of object model B and D having the same outer contour, recognition errors are occurred when the match process is performed with only the outer contours. However, if the inner features are also used, they can be successfully discerned. In the cases of object model A, E and F which have distinctive shapes of outer boundaries, the recognition is also successful but there are chance to be failed since there exists no inside edge feature. In contrast, the object model C which has the inside edge features can be easily recognized even when other objects have the same outer boundary, since it has enough inside edge features.

The top object detection process is tested with the 100 images in which 3 to 10 objects of 6 types are arbitrarily stacked. The task is to segment all not-occluded objects in the image. It is confirmed by the experiments that the successful detection rate is over 98%.

5.2 Pose estimation of the top object

To test the accuracy of the pose estimation algorithm, the poses of the detected not-occluded objects are estimated. Here the pose resolution of the manipulator is assumed less than 1.0[mm]. The rotation is expressed by the angles rotated with reference to the X and Z axes of the manipulator, and the position of the object center is expressed by the (X,Y) coordinates in the manipulator frame. The experimental results are summarized in Table 3. The error included in rotation estimation about the X axis increases as rotation angle gets larger, but the error included in rotation estimation about the Z axis does not change depending on the rotation angle, as can be expected. Translation error also increases if the object is located far from the origin of the camera image frame. As shown in Table 3, both rotation and translation errors are small enough to be used for guiding the manipulator to pick up an object.

Table 3 The errors included in the pose estimations.

Contents			Average Error
Rotation [degree]	X axis	+30	3.12
		+15	2.34
		-15	2.29
		-30	2.94
	Z axis	+90	2.12
		+45	2.30
		-45	2.22
		-90	2.51
Translation [mm]	100	50	2.50
	50	100	2.71
	50	50	1.53
	-100	-50	2.76
	-50	-100	2.81
	-50	-50	1.72

6 Conclusions

We have presented a new algorithm for detection and recognition of the top object on a stack of arbitrarily located objects, using a signature map. A signature map is the signature of whole image generated with selecting the center of a reference object in the image as the center of the signature. In a signature map, since the signature of the reference object is not distorted if it is not occluded, it can be determined of the reference object is on the top or not. Of course, if an object is located without occlusion, it is also considered as the top object that can be grasped. The experimental results have show that the proposed method has several advantages in detecting and recognizing the top object. The first one is that the method simultaneously detects and recognizes the top object by analyzing the signature map. The second one is that the method can discern those objects having the same boundaries but having different inside features. Above all, the can detect and recognize the object even when the boundary of a reference object is not completely connected. To obtain these advantages, one weakness that must be enhanced is that the time complexity increases as the number of objects in the input image increases.

References:

[1] Berger. M, Bachler. G, Scherer. S, "*Vision Guided Bin Picking and Mounting in a Flexible Assembly Cell*," IEA/AIE 2000, pp.109-118, 2000.

[2] Dmitry Chetverikov, "*A Simple and Efficient Algorithm for Detection of High Curvature Points in Planar Curves*," Computer Analysis of Image and Pattern 2003, pp.746-753, 2003.

[3] Adnan A.Y. Mustafa, "*Boundary Signature Matching for Object Recognition*," VI2001 Vision Interface Annual Conference, pp.72-79, 2001.

[4] Adnan A.Y. Mustafa, "*Matching Incomplete Object Using Boundary Signatures*," Proceedings of the 4th International Workshop on Visual Form, pp.563-572, 2001.

[5] Ovidiu Ghita, Paul F. Whelan, and John Mallon, "*A bin picking system based on depth from defocus*," Journal of Electronic Image, Vol.14, Issue 2, pp.234-244, 2003.

[6] Kirkegaard.J, Moeslund.T.B, "*Bin-Picking based on Harmonic Shape Contexts and Graph-Based Matching*," ICPR06, pp.581-584, 2006.

[7] Fu Chang and Chun-Jen Chen, "*A Component -Labeling Algorithm Using Contour Tracing Technique*," IEEE Document Analysis and Recognition, pp.741-745, 2003.