Vision Based Bin Picking Method Using Hierarchical Image Analysis

Kyekyung Kim[1], Sangseung Kang[1], Jaeyeon Lee[1] and Jinho Kim[2]
[1]Intelligent Cognitive Technology Research Department, ETRI, Daejeon, Korea
[2]Department of Electronics, Kyungil University, Daegu, Korea
[1]{kyekyung, kss, leejy}@etri.re.kr http://www.etri.re.kr
[2] ho@kiu.ac.kr http://www.kiu.ac.kr

ABSTRACT

In this paper, we describe vision based bin-picking system for robot application using multiple local features, which are extracted from single camera. Multiple features are extracted from texture on object surface that has used to estimate surface rotation angle and distance to object to be picked. Challenging problem is to estimate accurate picking point and distance to object. It is difficult to solve aforementioned problem because the distorted image affected by illumination has caused reflection on the image surface or feature data loss of texture. In this paper, we proposed hierarchical analysis using multiple cues by multi-resolution images to estimate picking points of piled objects in the bin. The estimation of object location by coarse image and picking point by fine image have processed. We have tested to evaluate performance on ETRI database, which have captured under various lighting condition in the pilot system, which is constructed like industrial environment.

KEYWORDS

Bin-picking, Object detection, Pose estimation, Hierarchical analysis, Multi-resolution images

1 INTRODUCTION

The industrial robot appeared in factory automating systems has been used as an effective means for solving challenges in industrial sites such as automation task, production cost saving, etc. As vision sensor technology has developed in a manufacturing line, object recognition method has applied to factory automation.

In recent years, due to introduction of cell production methods for producing various types of products in one line, need to estimate a type and a posture angle of a component and accurately mount the component is being gradually increased. Bin picking method of pick out piled object has studied intensively to perform works using industrial robot at cell manufacturing system because of cost saving by removing component feeding equipment, enhancing competitiveness of flexible manufacturing cell production systems.

Even in a machine component process had a low rate of automation, industrial robots are introduced to perform works at a manufacturing line instead skilled workers using various IT technologies of a control or machine and elements technologies such as a vision sensor, a force sensor, etc. As a sensing technology among element technologies applied to an industrial robot is enhanced and thus performance in a 3D visual sensor is also significantly improved, a need to develop an intelligent robot capable of performing a bin picking work in which a needed component is recognized and picked from a stacked pile is being increased. However, a vision based bin-picking method or technology has not developed sufficiently to apply to manufacturing line and has low recognition performance.

Bin-picking can be mainly applied to pick or assembly process of components using the detection of object location and pose estimation of x, y, z axis with 2D or 3D vision sensors [1-5]. 2D vision based bin-picking methods using stereo camera or single camera has extracted features from shape of object and estimated object pose by calculating distance from extracted features [1-2]. Meanwhile, 3D vision based bin-picking method [3-4] has captured depth information of object using laser sensors or structure lighting pattern. An existing bin-picking method has extracted simple cue such as a circle or a rectangle from object and estimated object location and pose. A method using a camera image and a CAD model, and a method for modeling a 2D curved surface with both laser measurement and images to recognize a position and direction of a 3D object.

However, it is actually difficult to estimate accurate angle of component disposed at a variety of angles in piles and illuminations changed according to an actual production environment. Therefore, bin-picking task has considered non-trivial task because those problem have low reliability in performing task such as location detection or pose estimation. Many researchers have studied steadily to solve aforementioned problem and to get more accurate performance result because of fits to many application fields, which were be needed high demand of production automation

In this paper, bin picking method using 2D vision sensor has proposed that uses multiple local features extracted from multi-resolution image and hierarchical analysis to estimate picking point of piled components. Candidate picking components have detected with low resolution image, which selected by extracting simple cue such as a rectangle. And then, multiple local features are extracted to estimate pose or distance of a picking component with high resolution image. An accurate pose or distance has estimated using multiple local features calculated from transformed geometric data according to x, y, z axis. To evaluate performance of position detection and pose estimation of picking component, we have tested bin picking algorithm using image data, which has captured on various lighting condition by time passing. We have constructed a pilot system like a real working environment and tested picking task using robot by providing calibration coordinate to drive robot.

2 VISION BASED BIN PICKING SYSTEM

The vision based bin picking system has composed of several processes such as image processing, feature extraction, candidate component selection, object location detection and pose estimation [6-11]. Bin picking target object are various kinds of material components, which includes reflection object like a metal, plastic wrapped object. A polarizing filter has used and image filtering technique has strengthen because of a reflection object. Bin picking pilot system and target object has shown in Figure 1.

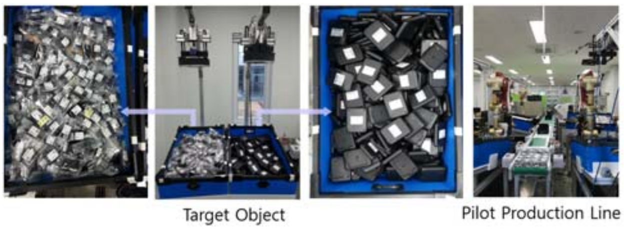

Figure 1. Bin picking pilot system and target object

2.1 Image Processing

An image processing includes image smoothing, edge detection, binarization, noise remove steps. An image preprocessing configured to extract a plurality of edges of a picking component from an input image. The edge detection includes a second derivative calculation by applying Gaussian filter to the component image. Blobs in the input image are detected by applying the local adaptive binarization technique [6] to the component image. The image preprocessing unit combines the detected second component edge and detected component by analyzing intensity of local region in input image. The combination result builds up component, which keeps more features and are saved to detect bin picking component. It is robust to illumination variation that is very important factor to apply into real environment condition.

A polarizing filter has installed to reduce illumination effect for atypical bin picking object such as reflection material surface component, plastic wrapped component. Figure 2 shows the example of edge detection and local adaptive binarization image.

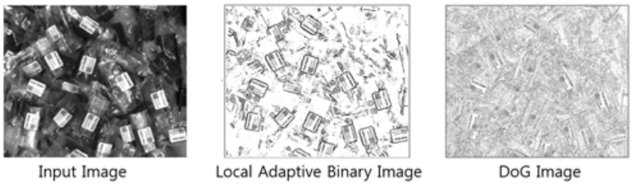

Figure 2. Image processing results

2.2 Candidate Object Detection

Object detection has made progress using hierarchical image analysis with multiple resolution images. Candidate positions of components to be picked have detected with low resolution image, and then more detailed features are extracted to estimate accurate component position, pose and distance from vision sensor to camera with high resolution image.

Candidate picking objects are selected by extracting simple cue feature like a rectangle, which offers to a reference position to pick a component and is used to calculate center point of a picking component. The simple cue for picking position by extracting features in interior component is detected with low resolution image. A rectangle feature in label has extracted even the feature has transformed or distorted due to illumination.

The rectangle has also used to estimate distance from camera to a picking component. Camera position and distance to component can be estimated using area or length for four sides of rectangle. Real distance for rectangle in the real world is calculated by matching geometric information of rectangle and related real distance in the world [7].

Final decision of picking component has fixed through calculation of area, geometric transform factor according to x, y, z axis, respectively. The optimal component to be picked has chosen by verifying whether the component is in advantageous picking location by examining the component located on the top or less rotation to each axis. The center point of rectangle is selected as picking point of component.

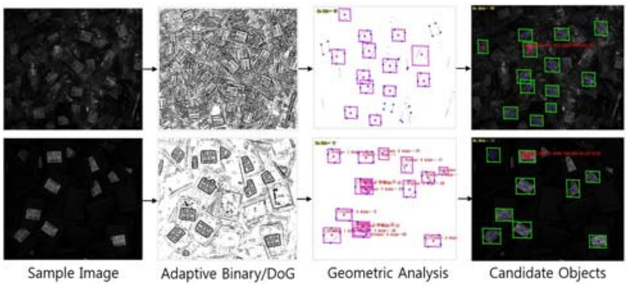

Figure 3. Candidate object detection using simple cue in low resolution image

2.3 Pose estimation

More features are extracted to estimate accurate component location, pose and distance to pick component. Some characters and symbols are appeared in rectangle on label those are extracted in high resolution image. However, some characters and symbols are missed or distorted because of illumination effect. Therefore, interior features are trained to extract more features even missing or distorted features are included in rectangle area. Hierarchical image analysis using multiple image resolution has processed to estimate accurate picking point decision as shown in Figure 4.

It is very difficult to get a surface orientation of a rotated component according to x, y, z axis. Therefore, several processes are applied to estimate pose of picking component. A pose estimation has processed by detecting maximal axis of component that has investigated by recognizing multiple features of component in a label. The multiple features include the direction information of component, which is used to assemble with right position.

Another pose estimation has processed to detect surface orientation of x, y, z axis. An extracted rectangle has used to analyze several factors, which are defined to examine geometric transformed parameter of rectangle and used to estimate surface orientation.

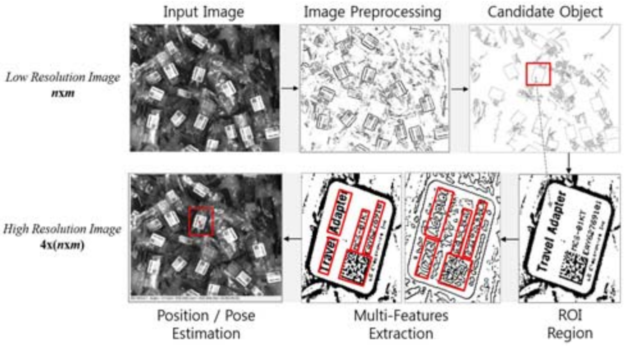

Figure 4. Hierarchical image analysis with multi-resolution images

Scale, length and angle of four sides of rectangle, b={b1, b2, b3, b4, b5}, are calculated. An accurate pose of component is estimated as shown in Figure 5.

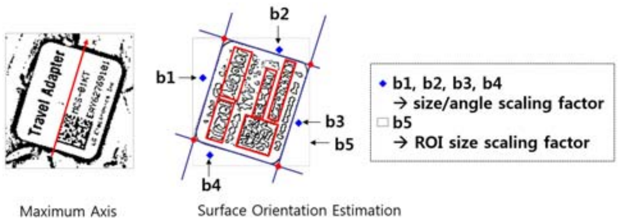

Figure 5. Pose estimation using multiple feature extraction and surface orientation factors

Multiple features are trained using neural network to estimate accurate pose of component and component position, as shown in figure 6.

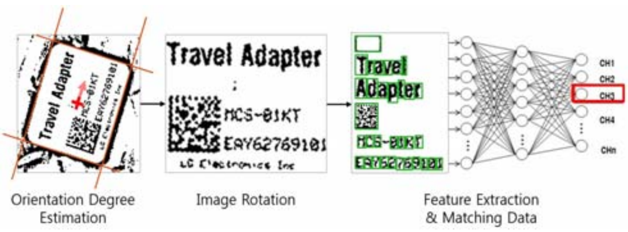

Figure 6. The training process for pose estimation using neural network

3 EXPERIMENTS

Experiment environment has set to evaluate proposed bin picking method which includes pilot system like real factory for bin picking and working table for capturing taring or testing data. Working table has consisted of supplying bin of 600x400x200(mm), LED ring light of 200x200x25(mm) / 21W, vision sensor of Matrix Vision mvBlueCOUGAR-X225G. Vision sensor has mounted within 1m distance to supplying bin. Bin picking robot has located at the center of supplying bins. Figure 7 shows bin picking working environment and target components to be picked. Bin picking components have randomly piled in a bin and input images of 2248x2050 grey 8bit BMP have captured. Bin picking components include a plastic wrapped component and formal shape component. Training and testing database have built under various lighting condition and pilot system during 3 years.

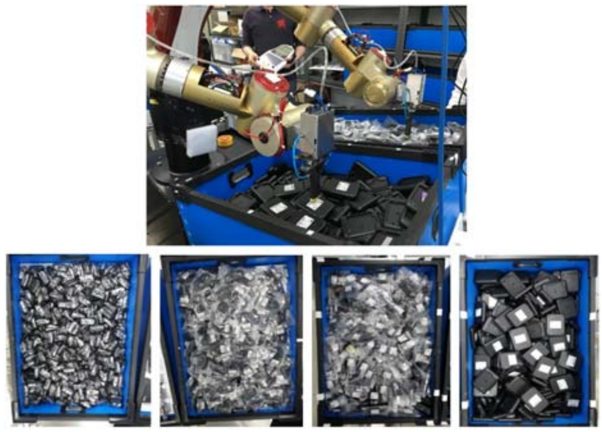

Figure 7. Bin picking using dual arm robot

Surface orientation estimation data have captured and calculated the length of four sides of rectangle and matched real distance in the real world with database captured according to different distance as shown in Figure 8. Codebook has created with database captured by different distance level that has picking distance from camera to component, x, y, z orientation degree.

Figure 8. Database captured under different distances for codebook generation

Performance evaluation of object detection and pose estimation has processed with 400 image data acquired from images in ETRI and pilot line in KIMM. 7 kinds of components have used to test object recognition performance such as position detection and pose estimation of object and bin picking robot has linked to experiment. Table 1 shows the detection rates and pose estimation of picking component in a bin. 1 to 4 components are randomly piled in a bin.

Table 1. The detection rate and pose estimation of 4 kinds of picking components

Kind of comp.	Testing no. of sample image	No. of detection image	Comp. detection rate(%)	Pose estimation (°)
Set1	100	97	97	0.9
Set2	100	98	98	0.7
Set3	100	96	96	0.9
Set4	100	95	95	1.2
Total	400	386	96.5	0.92

It is difficult to estimate accurate pose, distance to component and surface rotation to x, y, z axis, respectively. To solve the challenges problem, multiple features are extracted and trained to estimate pose of picking component. A heterogeneous features such as SIFT can be combined to enhance pose and distance by generating feature vector.

A different kinds of features in a label have extracted by SIFT algorithm using stereo images. The features extracted from stereo images are used to get a close distance and surface orientation angle. Distances are calculated by matching features of candidate components from stereo image. Stereo based 3D restoration has processed and surface normal vector has generated to each component.

4 CONCLUSIONS

In this paper, we have proposed vision based bin picking using multiple local features with multi-resolution images for automation factory using industrial robot. The working environment setting for bin picking, the detection of object position, pose estimation, distance measurement and surface orientation angle detection steps have included in bin picking system. Hierarchical image analysis has processed to estimate picking point of piled components in a bin. Candidate picking components have detected with low resolution image and multiple local features are extracted to estimate accurate pose or distance of a picking component with high resolution image. Multiple local features have trained to detect component position and estimate pose and distance to picking component. An accurate pose or distance has estimated from transformed geometric data according to x, y, z axis.

For evaluating the proposed bin picking method, we have tested on database, which have captured in various lighting condition and pilot system. In addition to, the vision based bin picking method has had linked testing with dual arm robot.

In the future, heterogeneous local features will be combined to get more precise object position, distance to object, surface orientation to each axis. And also, the research to enhance bin picking performance will be processed by combining sensors, optimizing lighting condition, increasing reliability and analyzing error factors.

ACKNOWLEDGMENTS

This work was supported by the R & D program of MOTIE & KEIT [10038660, Development of dual-arm robot system based on multi-robot cooperation for cell manufacturing process of IT products]

REFERENCES

[1] J. K. Oh, S. H. Lee and C. H. Lee, “Stereo Vision Based Automation for a Bin-Picking Solution,” International Journal of Control, Automation, and Systems, vol. 10, no. 2, pp. 362-373, 2012.

[2] K. Rahardja, and A. Kosaka, “Vision-based bin-picking : Recognition and localization of multiple complex objects using simple visual cues,” IEEE Proc. of International Conference on Intelligent Robots and System, vol. 3. pp. 1448-1457, 1996.

[3] O. Kazuya, H. Toshihiro, F. Masakazu, S. Nobuhiro, S. Mitsuharu, "Development for Industrial Robotics Applications, IHI Engineering review," vol . 42, no. 2, pp. 103-107, 2009.

[4] S. Lee, J. Kim, M. Lee, K. Yoo, L. G. Barajas and R. Menassa, "3D Visual Perception System for Bin Picking in Automotive Sub-Assembly Automation," 8th IEEE International Conference on Automation Science and Engineering, pp. 706-713, 2012.

[5] K. Ikeuchi, B. K. P. Horn and S. Nagata, "Picking up an Object From a Pile of Object," Artificial Intelligence Lab. of the Massachusetts Institute of Technology, A.I. Memo, no. 726, pp. 1-26, 1983.

[6] F. Shafait, D. Keysers and T. M. Breuel, "Efficient Implementation of Local Adaptive Thresholding Technique Using Integral Images," Document Recognition and Retrieval XV, Proceedings of the SPIE, vol. 6815, pp. 681510-681510-6, 2008.

[7] J. Lee, "A New Solution for Projective Reconstruction Based on Coupled Line Camera," ETRI Journal, vol. 35, no. 5, pp. 939-942, 2013.

[8] D. Lee and M. S. Nixon, "Vision-based finger action recognition by angle detection and contour analysis," ETRI Journal, vol. 33, no. 3, pp. 415-422, 2011.

[9] P. F. Felzenszwalb and J. Schwartz, "Hierarchical matching of deformable shapes," Computer Vision and Pattern Recognition, pp.1-8, 2007.

[10] C. Lu, N. Adluru, H. Ling, G. Zhu, L. J. Latecki, "Contour based object detection using part bundle," Journal of Computer Vision and Image Understanding, vol. 114, issue 7, pp. 827-834, 2010.

[11] V. Ferrari, L. Fevrier, F.Jurie, and C. Schmid, "Groups of adjacent contour segments for object detection," IEEE Trans. on Pattern Anal. Mach. Intel., vol 30, no. 1, pp.36-51, 2008.