Novel visual sensing systems for real time corrective assembly motion in the presence of occlusion

Jin Young Kim*[a], Wan Soo Kim[b], Hyung Suck Cho[c]

[a]Dept. of Robot Systems Engineering, Tongmyong University of Information Technology, Korea
[b]Production Engineering Center, Samsung Electronics Co., Ltd., Korea
[c]Dept. of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Korea

ABSTRACT

In precision robotic assembly a small misalignment between parts to be assembled would always occur at the interface between mating parts due to positioning inaccuracy of robot, assembly fixtures and manufacturing tolerance of parts. This misalignment, however small can produce large contact forces, resulting in damage to parts or robots and thus lead to failure in assembly. Therefore the misalignment has to be detected and compensated accurately during mating period. For this purpose visual sensing technique has been widely used since it can detect rather large misalignment and also parts shape at distance. In case of adopting camera, this technique usually obtain local information due to the limited range of its field of view. Furthermore, it can not avoid self-occlusion generated by the invisible region occluded by objects to be viewed. This problem has been a major hindrance to the success of assembly action when assembly is executed by a vision-based technique. In this paper, two novel visual sensing methodologies are developed to avoid such criticisms. Both systems consist of four components: an inside mirror and outside mirror, a pair of plane mirrors and a camera with a collecting lens. The difference between the two is that the system A adopts conic mirror configuration, while the system B employs pyramidal one. Due to this configuration difference, the system A is shown to be capable of detecting two pi omni-directional image, while in addition to this functionality the system B can detect three dimensional measurement of objects with only one image capture. The measurement principles are described in detail and compared with each other. The image acquiring process is shown to easily detect the in-situ status of assembly action, while the recognition method is found to be effective to identify instantaneous misalignment between the peg and hole. The results obtained from a series of experiments show that the proposed visual sensing methods are an effective means of detecting misalignment between mating parts even in the presence of self-occlusion. The implication is that they will dramatically increase the rate of success when actually utilized in assembly process.

Keywords: visual sensing system, robotic assembly, occlusion, pyramidal mirror, conic mirror

1. INTRODUCTION

Assembly, which is one of the major steps in the manufacturing process, is to fit two or more components according to the pre-determined geometrical relations. Successful assembly can be achieved if it is possible to obtain complete a priori knowledge of the assembly environment, and if it is possible to perform the task with ideal positioning devices. In robotic assembly, however, misalignment is occurred by many factors such as the positioning error of robots or parts feeding devices. Therefore, successful assembly requires compensating for any possible misalignment between the parts to be assembled. To cope with such assembly problems, a variety of research has been made. The resulting developments can be classified as follows[1]: the passive method[2-4], the active method[5], the passive-active method[6].

Many techniques using force/torque sensors, optical fiber sensors, and pneumatic pressure sensors have been developed for misalignment detection and its compensation between mating parts[1,7]. These techniques are suitable for detecting relative local geometry between mating parts along their mating boundary. Also, visual sensing technique has been widely used since it can detect rather large misalignment and also parts shape at distance. In case of adopting camera, this technique usually obtain local information due to the limited range of its field of view. Furthermore, it can not avoid self-occlusion generated by the invisible region occluded by objects to be viewed. This problem has been a major hindrance to the success of assembly action when assembly is executed by a vision-based technique.

In this paper, two novel visual sensing methodologies are developed to avoid such criticisms. Both systems consist of four components: an inside mirror and outside mirror, a pair of plane mirrors and a camera with a collecting lens. The

* Correspondence: Email: kjy@tmic.tit.ac.kr; Telephone: +82-51-629-7236; Fax: +82-51-629-7249

difference between the two is that the system A adopts conic mirror configuration, while the system B employs pyramidal one. Due to this configuration difference, the system A is shown to be capable of detecting two pi omni-directional image, while in addition to this functionality the system B can detect three dimensional measurement of objects with only one image capture. The measurement principles are described in detail and compared with each other. The image acquiring process is shown to easily detect the in-situ status of assembly action, while the recognition method is found to be effective to identify instantaneous misalignment between the peg and hole. The results obtained from a series of experiments show that the proposed visual sensing are an effective means of detecting misalignment between mating parts even in the presence of self-occlusion.

This paper is organized as follows: Section 2 describes the configuration and specifications of the sensing system using pyramidal mirrors, and shows the experimental results for measuring part deformation and misalignment by using the system. Section 3 describes the configuration and specifications of the sensing system using conic mirrors, and shows the experimental results for measuring misalignment by using the system. Section 4 describes the comparison of pyramidal mirror system and conic mirror system. Finally, some conclusions are made in section 5.

2. A SENSING SYSTEM USING PYRAMIDAL MIRRORS

2.1. Configuration

Fig. 1(a) illustrates the basic configuration of the developed sensing system using pyramidal mirrors. It is composed of a camera, a pair of plane mirrors, and a pair of pyramidal mirrors. In order to measure three-dimensional deformation by using a camera, two views are necessary, as shown in Fig. 1(b). Fig. 1(c) illustrates an image of a peg and a hole pair. Because four images that are reflected from each face of the pyramidal mirrors are projected on the image plane of a camera, this system configuration is equivalent to that utilizes four cameras. This configuration allows the system to overcome self-occlusion[8].

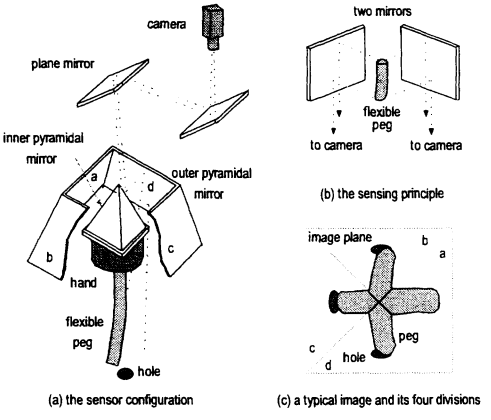


Fig. 1. The schematic of the sensing system using pyramidal mirrors

2.2. The field of view

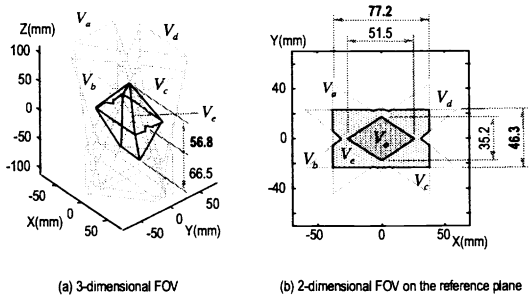
The field of view(FOV) of the system is obtained from the FOVs corresponding to each face of the pyramidal mirrors. Because the system uses stereo vision, the effective FOV V_e that can be seen through two or more faces of the pyramidal mirrors is actually available for three-dimensional measurement. Hereafter, what is referred to as the FOV means the effective FOV. The omni-directional FOV V_o is the region that can be seen through all faces of the pyramidal mirrors.

$$V_e = (V_a \cap V_b) \cup (V_a \cap V_c) \cup (V_a \cap V_d) \cup (V_b \cap V_c) \cup (V_b \cap V_d) \cup (V_c \cap V_d) \quad (1)$$

$$V_o = V_a \cap V_b \cap V_c \cap V_d \quad (2)$$

Because most objects have height, characteristics of the sensing system should be investigated in three-dimensional

space. It is also necessary to investigate characteristics and specifications on the two-dimensional plane, for example, the hole plane. Therefore, in consideration of a representative measurement height, we define the reference plane as the object plane that has a working distance of 508 *mm*. Fig. 2(a) shows the three-dimensional FOV, and Fig. 2(b) shows the two-dimensional FOV on the reference plane.



(a) 3-dimensional FOV (b) 2-dimensional FOV on the reference plane

Fig. 2. The field of view of the implemented sensing system

2.3. The depth of field and resolution

The depth of field D can be calculated by the following equation[9]:

$$D = \frac{2wacf(w-f)}{a^2f^2 - c^2w^2} \tag{3}$$

where w is the working distance, f is the focal length, a is the diameter of the aperture, and c is the allowable confusion diameter. In the implemented system, the depth of field is greater than the variation in working distance. Therefore, all points in the FOV can be focused within a pixel size on the image plane. On the other hand, resolution R is given by

$$R = \frac{V}{n_e} \tag{4}$$

where V is the size of the FOV, and n_e is the number of pixels in V. Resolution of the implemented system varies with the locations in the FOV. Therefore, we define that the standard resolution R_s is the smallest resolution in the inside region of the circle with the radius of 10 *mm* centered at the middle of the FOV. In the implemented system, R_s is about 0.2 *mm*.

2.4. Estimation of part deformation and misalignment

In order to assemble the flexible parts successfully, parts deformation in any direction and misalignments between mating parts need to be measured. In this section, the algorithm to estimate parts deformation and misalignments by using a visual sensing system in cylindrical peg-in-hole tasks is described. Parts deformation can be represented by the shape of the center-line of a peg. Misalignment between mating parts is defined as the relative error between the center of a hole and the center of the bottom of a peg.

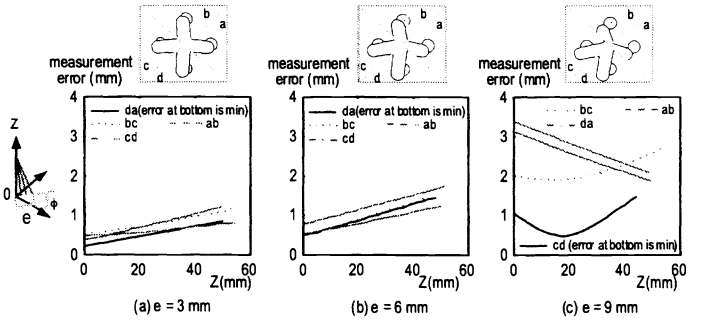
In the case of a cylindrical peg with a circular cross-section, it can be assumed that its center-line is projected to the center-lines of its projected images[10]. Thus, parts deformation can be obtained by estimating the center-lines in two-dimensional peg images.

2.4.1. Estimation of Part Deformation

To obtain parts deformation, pegs must first be recognized in edge images. Their edges are extracted from the images taken by the proposed system, and classified into the sides and the bottoms of peg. The two center-lines from the four images are then selected by using proper conditions. A center-line is composed of the midpoints between two side edges in one image. Next, corresponding points in two center-lines are found using epipolar constraint. Finally, from

corresponding points, the center-line of a peg is reconstructed in three-dimensional space.

Fig. 3 shows the error between original centerlines and measured centerlines of pegs in three-dimensional space when the actual misalignment e is equal to 3, 6, 9*mm* according to the inclination of pegs. The size of errors is varied, depending on which images are selected among four images. In assembly, the relative error between the center of the bottom of a peg and the center of a hole becomes the most important factor. The thick line shows the case when the error at the bottom of a peg is minimum.



(— : when the measurement error at the peg bottom is minimum
e : size of misalignment, ϕ : direction of misalignment)

Fig. 3. Experimental results of measurement of part deformation according to the selection of two views when $\phi = 45°$

2.4.2. Estimation of Misalignment

To estimate misalignments between the holes and their respective mating parts, the center of an occluded hole and that of the bottom of a peg are evaluated. First, pegs and holes in images are recognized. Next, the center of an occluded hole is found using epipolar constraint. Next, the center of the bottom of a peg is found by using the same method. Finally, misalignment is calculated from the difference between the two centers.

Fig. 4 shows the error between the measured misalignment and actual misalignment e when e = 3, 6, 9*mm*. This figure shows the maximum error, the minimum error, and the error of the case where a set of two images that have most in common is selected. The size of error is varied, depending on which images are selected among four images. From these results, we can confirm that the proposed system can be used more effectively in cases where occlusion occurs.

The measured deformation and misalignment information are used to generate a control action via an appropriate assembly method, whereas the control signal in turn actuates the manipulator. In other words, a vision sensor-based assembly strategy must compensate misalignment between mating parts from the measured information.

The relationship between the measured information and the corrective motion to be accomplished by a robot is very complex and nonlinear, and therefore, it is desirable to estimate the corrective motion from the measured information by using an AI-based method[11].

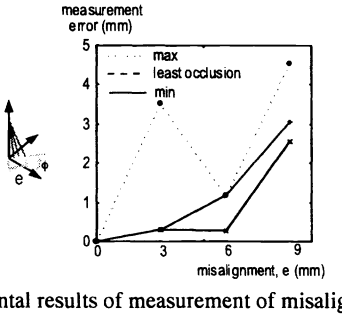


Fig. 4. Experimental results of measurement of misalignment views when $\phi = 45°$

3. A SENSING SYSTEM USING CONIC MIRRORS

3.1. Configuration

Fig. 5(a) illustrates the basic configuration of the sensing system. The system consists of four components: an inside

and an outside conic mirror, a pair of plane mirrors, a camera, and a gripper. The goal of the sensing system is to obtain a 2π view as described above. Let us assume that there are m-plane mirror patches, placed at interval $\varphi = 2\pi/m$ on the circumference of a circle with a radius R, and their oblique angles and widths are $\alpha/2$ and $\Delta W = \Delta\varphi R$, respectively.

This inside conic mirror is capable of reflecting the 2π figure of an object, encompassed by it, in its 2π mirror surface without self-occlusion. Therefore, the inside-conic mirror is used to obtain the 2π relative geometry between mating parts at once, without self-occlusion. However, an additional optical components are required to detect it by using a camera on off-axis, as shown in Fig. 5(a). At first, an outside-conic mirror with a mirror surface outside itself, placed co-axially on the central part of the inside-conic mirror, is used to collect the 2π image in the inside-conic mirror surface. In order to project again the collected 2π image onto the image plane of the camera, two plane mirrors are then used[12].

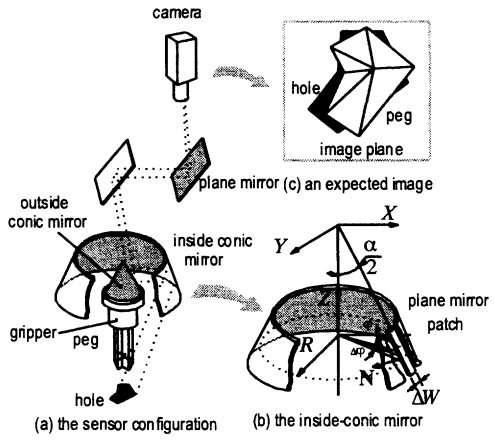


Fig. 5. The schematic of the sensing system using conic mirrors

3.2. The field of view

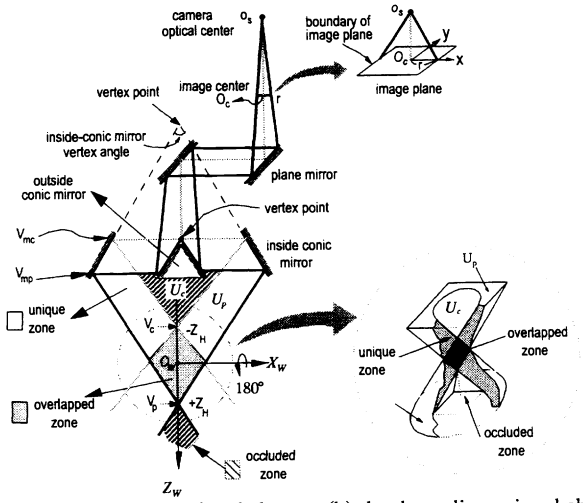
Fig.6 illustrates the sensing system field of view. Its vertical cross-section looks like a simple triangle, as shown in Fig. 6(a). In addition to its vertical cross-section, its three-dimensional shape consists of the conical and pyramidal volumes. These volumes are determined by shape of a sensing element such as the CCD cell and the changed optical path by mirrors. They are investigated by inversely tracing the rays of a light, projected onto the boundary of a sensing element and the image center O_c in the image plane, respectively.

First, when the rays, projected onto the rectangular boundary of the CCD cell, are inversely mapped onto the object space, it is natural that a rectangular shape should come out on the outside-conic mirror surface. The rays with the rectangular shape are reflected again on the inside-conic mirror surface. And they pass through a point V_p on the vertical axis, since the vertex angle of the inside conic mirror is less than 180°. Hence, these whole optical paths of the rays look like a pyramidal volume U_p, as shown in Fig. 6(b).

In addition to the case of the boundary of the CCD cell, when the rays of a light, projected onto the image center O_c in the image plane, are inversely mapped onto the vertex point of the outside-conic mirror, a horizontally circular shape comes out on a circumference of the vertex point. The rays with the circular shape are reflected again on the inside-conic mirror surface. And they pass through a point V_c, similarly as the case of the pyramidal volume. Then, these optical paths look like a conical volume U_c, as shown in Fig. 6(b). The rays generating these volumes are reflected on 2π horizontal lines through the different vertical positions V_{mc} and V_{mp} on the inside-conic mirror surface, respectively, and they pass through the different positions V_c and V_p. Therefore, they intersect each other and the lower structure of the field of view is symmetrically given as a shape rotating the conical and the pyramidal shape in 180° with respect to the X-axis at $Z = -Z_H$ and $Z = Z_H$, respectively.

The region of the field of view is classified into three zones according to mapping features: a unique zone, an overlapped zone, and an occluded zone. The unique zone, shaded with white in Fig. 6(a), has a feature of mapping a

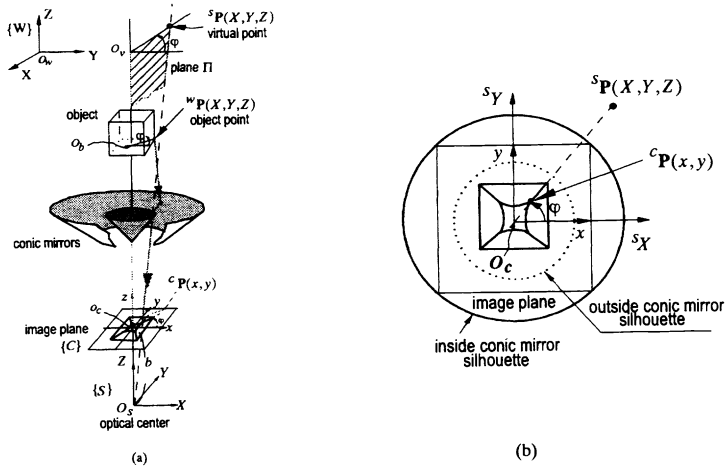
point object into the point shape. However, the overlapped zone, shaded with light gray, is the common region of the pyramidal volume before rotating in 180° and the conical volume after that. Since this zone is visible from all positions on the 2π surface of the inside-conic mirror, a point object, placed on the region, looks like a donut on the image plane. The occluded zone is out of the field of view, and it is natural that an object in this region should be invisible to the camera. As a result, the available zone is actually the unique zone; however, the overlapped zone is used within limits.



(a) the two-dimensional shape (b) the three-dimensional shape

Fig. 6. Configuration of the field of view

3.3. Invariability of Azimuth Angle



(a) (b)

Fig. 7. Invariability of azimuth angle: (a) a simplified model, (b) a mapping feature

Fig. 7(a) is a simply reconstructed configuration of the sensing system, as if the camera were placed coaxially above the outside conic mirror. This equivalent configuration is obtained, without loss of generality, by stretching the optical path of rays from the optical center O_S to a pair of plane mirrors with respect to the vertex axis through the vertex points of the conic mirrors. Suppose that a point ${}^{W}\mathbf{P}(X,Y,Z)$ on an object, defined with respect to the object coordinate {W} centered at $O_w - XYZ$, is projected onto the image plane. Then, its virtual point ${}^{S}\mathbf{P}(X,Y,Z)$, defined relative to a sensor coordinate {S} centered at $O_S - XYZ$, exists on the same plane Π, defined by the points O_b, O_S and azimuth angle φ. When the virtual point is projected onto a point ${}^{C}p(x,y)$ on the image plane, the following relationship is

obtained by the reflective law on the conic mirror surfaces and the perspective projection of the camera;

$$\tan\varphi = \frac{y}{x} = \frac{Y}{X}. \tag{5}$$

The result means that there is no variation in the azimuth angle during an image projection of the sensing system. All points with an azimuth φ in space, therefore, are mapped into the points on the radial line that passes through the image center O_C with the same azimuth φ on the image plane, as shown in Fig. 7(b).

3.4. Resolution

Resolution is one of the critical factors in evaluating sensing capacity of a sensor. In the sensing system, calculating the resolution e_r in consideration of the defocusing effect caused by depth variation, it is given as

$$e_r = \begin{cases} \dfrac{V}{M}, & for \quad D_b \le C \\ \left(\dfrac{D_b}{C}\right)\dfrac{V}{M}, & for \quad D_b \ge C \end{cases} \tag{6}$$

where a blurring circle diameter D_b is given by[9] $D_b = \dfrac{(w - w_r)D_a \times f \times V}{w_r(f - w) \times C \times M}$, w_r is the well focused measurement distance, or the reference measurement distance, w is measurement distance, f is lens focal length, C is size of pixel in image plane, M is nominal number of image pixels, V is field of view, and D_a is aperture diameter. The calculation of the sensing system resolution at the designed aperture diameter of 1.8 mm becomes about 0.088 mm.

3.5 Omni-directional image acquisition

Typically, the procedure of object recognition consists of four stages; obtaining a scene including an object, making feature extraction and description of the object, making models, and matching the description of the object with that of the model. From the viewpoint of a detector, the most important thing of the proposed sensing system among them is a sensing capability of a scene. Fig. 8 shows an experimental sensing example of a cylindrical object, with the characters, A, B, and C on its side surface, to obtain its omni-directional image. As shown in Fig. 8(b), the sensing system can immediately obtain an omni-directional image containing not only the object shape, but also the figures such as characters and holes on 2π side-surface of an object, without self-occlusion. On the contrary to other sensing systems capable of detecting local scenes, it needs no image composition in order to obtain an omni-directional object image of shape.

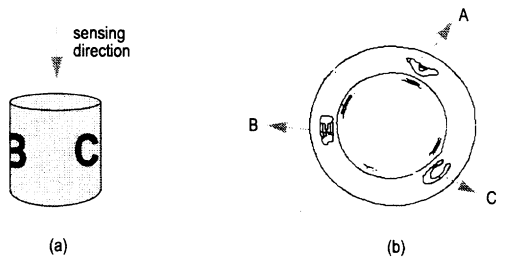
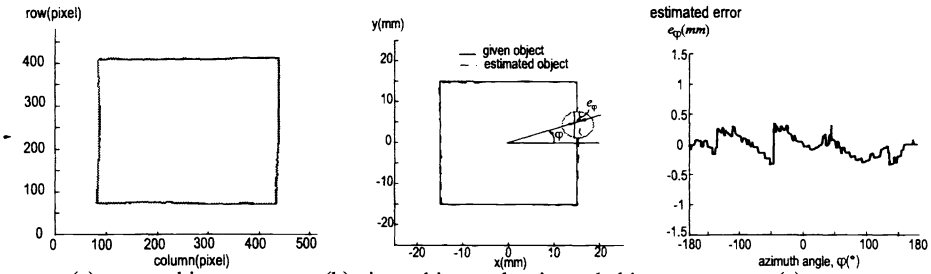


Fig.8. A sensing example of a cylindrical object with characters on its side:
(a) a cylindrical type object with characters on its side surface, (b) the detected omni-directional image

3.6. Estimation of misalignment

Some experiments have been executed so as to confirm the principle and show the feasibility of the sensing system. Fig. 9 shows an experimental result to validate the mapping relation between the object space and the image plane of

the camera, when an object is projected into the image plane through the proposed sensing system. Fig. 9(a) is a projection shape of a planar rectangle of $30mm \times 30mm$ placed on the measurement height $Z = Z_0$ from the optical center, through the sensing system. Fig. 9(b) is a compared result between the given rectangle and the inversely projected shape by using the inverse projection. Fig. 9(c) is an error analysis with respect to the result of Fig. 9(b). The result shows that the average error is within 2 pixels and the maximum error is about 4 pixels.

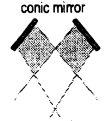


(a) captured image (b) given object and estimated object (c) error

Fig. 9. Comparison between a rectangular shape of $30 \times 30mm^2$ placed $Z_0 = 238mm$ and inversely projected object of captured image by using an inverse transformation model

4. COMPARISON OF PYRAMIDAL MIRROR SYSTEM AND CONIC MIRROR SYSTEM

Table 1. Comparison of pyramidal mirror system and conic mirror system

item	pyramidal mirror system	conic mirror system
principle	- multinocular system - effective FOV : field that can be seen through two or more faces of the pyramidal mirrors pyramidal mirror  common zone is useful	- monocular system - effective FOV : field corresponding to only one point of conical face conic mirror  unique zone is useful
fundamental purpose	- three-dimensional measurement (only one image is required) - additional funtion : omni-directional sensing	- omni-directional sensing - additional function : two images are required for three-dimensional measurement
application	- tasks that three-dimensional information is required - in assembly tasks : flexible parts assembly	- tasks that omni-dimensional information is required - in assembly tasks : assembly of parts with arbitrary cross-sectional shape

In this paper, two novel visual sensing systems for real time corrective motion were presented. As can be observed from the previous discussions, they are basically different from each other in fundamental purpose, sensing principle, and application areas, and so on, although both systems have similar exterior view. Therefore, it is difficult to evaluate which one is better by comparing the performance, cost and usefulness of acquired information. The main purpose of the pyramidal mirror system is a three-dimensional measurement, and the main purpose of the conical mirror system is an omni-directional measurement.

Because the pyramidal mirror system can acquire the three-dimensional information using only one image, image processing time is shorter than the conic mirror system. Also a plane mirror, which is a constituent unit of the pyramidal mirrors, is much cheaper than a conic mirror. Therefore, the pyramidal mirror system has an advantage to the conic mirror system in cost aspect.

Excepting a camera and the plane mirrors, the parameters to be calibrated in the pyramidal mirror system are the positions and orientations of the two pyramids. And the parameters in the conic mirror system are the positions and orientations of the two cones. Because the number of the parameters to be calibrated is the same, we can determine that the degree of difficulty of calibration is approximately the same in two systems.

The pyramidal mirror system can be applied to most tasks where three-dimensional information is required. In assembly tasks, it is particularly useful to flexible parts assembly. On the other hand, the conic mirror system is useful to assembly of parts with arbitrary cross-sectional shapes because it can effectively obtain 2π coaxial misalignment image. Table 1 shows comparison of pyramidal mirror system and conic mirror system.

5. CONCLUSIONS

In this paper, two novel visual sensing systems for real time corrective motion were presented. Both systems consist of four components: an inside mirror and outside mirror, a pair of plane mirrors and a camera with a collecting lens. The difference between the two is that the system A adopts conic mirror configuration, while the system B employs pyramidal one.

A conic mirror system has developed to particularly detect 2π misalignment between mating parts with asymmetrically complicated shapes without self-occlusion. This objective is implemented by using a pair of conic mirrors consisting of an inside-conic mirror and an outside-conic mirror. By using this configuration, it can immediately acquire not only an omni-directional image including enough information to fast recognize an object, but also a 2π coaxial misalignment image along the mating boundary interface between mating parts, without any self-occlusion. Therefore, the misalignment can be determined by utilizing pattern matching techniques without any image composition, no matter how complicated the shapes may be.

A pyramidal mirror system can detect three-dimensional measurement of objects with only one image capture. Therefore, this system can measure parts deformation in any direction and misalignments between mating parts without use of an analytical model.

This paper described the measurement principles of both systems in detail and compared with each other. The image acquiring process was shown to easily detect the in-situ status of assembly action, while the recognition method was found to be effective to identify instantaneous misalignment between the peg and hole. The results obtained from a series of experiments showed that the proposed visual sensing methods are an effective means of detecting misalignment between mating parts even in the presence of self-occlusion.

REFERENCES

1. H. S. Cho, H. J. Warnecke and D. G. Gweon, "Robotic Assembly: a Synthesizing Overview", *Robotica*, Vol. 5, pp. 153-165, 1987.
2. S. H. Drake, "Using Compliance in View of Sensory Feedback for Automatic Assembly", Ph. D Thesis, MIT,1977.
3. D. E. Whitney, "Quasi-Static Assembly of Compliantly Supported Rigid Parts", *ASME J. Dynam. Syst. Measur. Control*, Vol. 104, pp. 65-77, 1982.
4. K. W. Jeong and H. S.Cho, "Development of a Pneumatic Vibratory Wrist for Robotic Assembly", *Robotica*, Vol. 7, pp. 9-16, 1989.
5. Y. K. Park and H. S. Cho, "Fuzzy Rule-based Assembly Algorithm for Precision Parts Mating", *Mechatronics*, Vol. 3, pp. 433-450, 1993.
6. H. S. Cho and K. C. Koh, " The Development of a Flexible and Sensible Robot Wrist for Assembly Process", *Korean Society of Mechanical Engineers*, 8, No. 5, pp. 488-497, 1984.
7. N. Takanashi, H. Ikeda, T. Horiguchi and H. Fukuchi, "Hierarchical Robot Sensors Application in Assembly Tasks", *15th ISIR*, pp. 829-836, 1985.
8. J. Y. Kim, H. S. Cho and S. Kim, "A Visual Sensing System with Multiple Views for Flexible Parts Assembly" , *8th Int. Conf. on Advanced Robotics*, pp. 979-984, USA,1997.
9. E. Krotkov, "Focusing", *Int. Journal of Computer Vision*, Vol. 1, pp. 223-237, 1987.
10. N. Pillow, S. Utcke and A. Zisserman, "Viewpoint-invariant Representaton of Generalized Cylinders Using the Symmetry Set", *Image and Vision Computing*, Vol. 13, No. 5, June, pp. 355-365, 1995.
11. J. Y. Kim and H. S. Cho, "A Vision Based Error-Corrective Algorithm for Flexible Parts Assembly " , *1999 IEEE Int. Sym. on Assembly and Task Planning*, pp. 205-210, Portugal,1999.
12. W. S. Kim, H. S. Cho and S. Kim, "A New Omnidirectional Image Sensing System for Assembly(OISSA)", *Int. Conf. on IROS*, pp. 611-617, 1996.