Flip-chip alignment with a dual imaging system using a visual servoing method

Daljae Lee[*a], Xiaodong Tao[a], Hyungsuck Cho[a] and Youngjun Cho[b]
[a]Dept. of Mechanical Eng., Korea Advanced Institute of Science and Technology, 373-1, Guseong-dong, Yuseong-dong, Daejeon, Korea.
[b]Manufacturing System Research Center, Korea Institute of Industrial Technology, 35-3, HongChonRi, IbJangMyun, ChonAnsi, ChungNam 330-32, Korea.

ABSTRACT

IC chip has gradually become smaller and smaller, and thus it requires high packaging density. In chip packaging, accurate alignment of electronic components with respect to PCB is crucial for high quality packaging, especially in flip-chip assembly. In this paper, vision system is used to provide relative pose information between flip-chip and substrate. Based on this information, these two parts are aligned accurately using visual servoing. In order to achieve high accuracy alignment, a dual imaging system (DIS) is introduced in this work, which is composed of zoom lenses, beam-splitter, mirror, CCD, and LED illumination. It can simultaneously observe the solder bumps on flip-chip and the pattern of pads on substrate using one camera. Once the image frame containing flip-chip and substrate is obtained, their features are extracted from the preprocessed image. Extraction of the features enables us to obtain the position and orientation errors between the chip and the substrate. On the base of the measured errors, visual servoing method can determine the instantaneous velocity input of flip-chip at each servoing time and control the relative position and orientation precisely in an on-line manner. We carry out a series of experiments for various magnifications in order to evaluate the performance of the dual imaging system and the visual servoing algorithm as well.

Keywords: Flip-chip, Dual imaging system(DIS), Alignment, Feature extraction, Visual servoing

1. INTRODUCTION

Recently due to the development of IT technology and semi-conductor industry, electronic parts and products are becoming more and more complex, miniaturized and precise. As a result, current trend of IC packaging requires

[*]djsimple@lca.kaist.ac.kr; phone +82-42-869-3253; fax +82-42-869-3095; lca.kaist.ac.kr

Optomechatronic Systems Control, edited by Farrokh Janabi-Sharifi, Proc. of SPIE
Vol. 6052, 605203, (2005) · 0277-786X/05/$15 · doi: 10.1117/12.650315

miniaturization and high I/O density. To minimize IC chip, area of a chip should be almost the same as that of packaging. Especially, flip-chip is the smallest packaging which is possible to make[1]. This method uses solder bumps to connect flip-chip with pads on the substrate which face each other. The assembly of flip-chip is conducted by melting the solder bumps with the condition of high temperature and pressure after the bumps of the chip is aligned to the pattern on the substrate.

In this paper, in order to improve precision and speed of assembly, machine vision techniques are applied in flip-chip alignment for packaging. We developed an optical system to observe the solder bumps and the pattern on the substrate at the same time. We call this optical system as dual imaging system (DIS) which is composed of beam splitter, mirror, zoom-lens, CCD and illumination. We can get a composite image on CCD sensor, using the dual imaging system, where features are extracted from the solder bumps and the pads pattern on substrate. Using these features, we carry out precise alignment by visual servoing method. Here, the test flip-chip and pattern used in exprement are shown in Figure 1. The total size of flip-chip is $5.08mm \times 5.08mm$. The number of bumps is 88 and its diameter is $135\mu m$. The pitch is $204\mu m$.

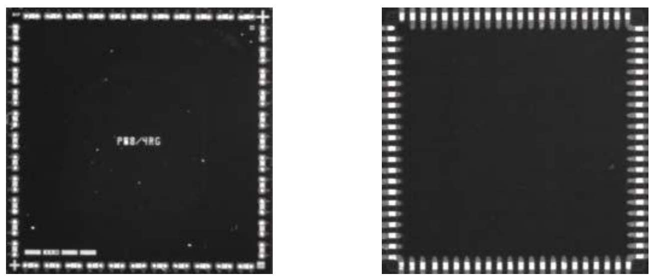
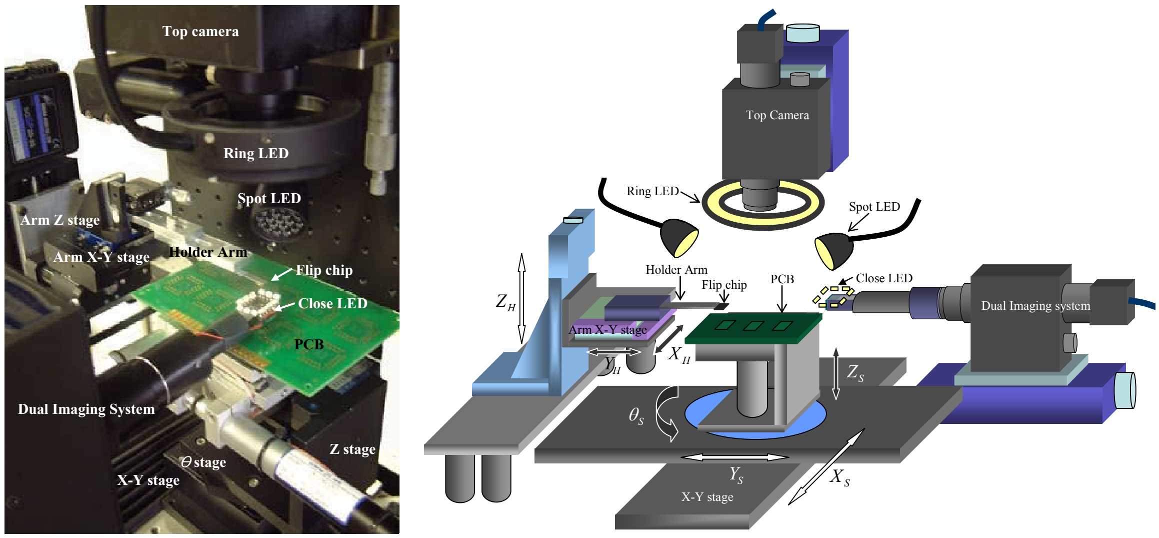


Figure 1: Patterns of flip-chip bumps and substrate

2. HARDWARE SETUP AND DUAL IMAGING SYSTEM

2.1 Configuration of visual servoing system

The hardware configuration of visual servoing system for flip-chip alignment is shown in Figure 2. It consists of vision unit, stage unit and accessory unit. In the vision unit, the top camera has large FOV (Field of view) which is used for coarse visual servoing, while dual imaging system is used for fine visual servoing. The stage unit is composed of two groups of stages. One is for moving substrate using step motors, the other is for moving flip-chip holder using DC motors. The former group has 4 axes (X_S, Y_S, Z_S, θ_S), and the latter group has 3 axes (X_H, Y_H, Z_H). Totally 7 axes are used for chip assembly. The accessory units include holder arm for griping the flip-chip, the ring type LED, the spot type LED, and the close LED.



(a) visual servioing system (b) schematic of visual servioing system

Figure 2: Configuration of visual servoing system

2.2 Dual imaging system

The dual imaging system (DIS) is defined as an optical system that can observe dual objects imaged in one image plane by different optical path. The Figure 3 shows the schematic of DIS designed in this paper. The rays from the flip-chip above the DIS and those from the substrate below the DIS are imaged on the CCD by passing through beam-splitter, mirror and lens respectively. Here, zoom lens is simplified to only one lens.

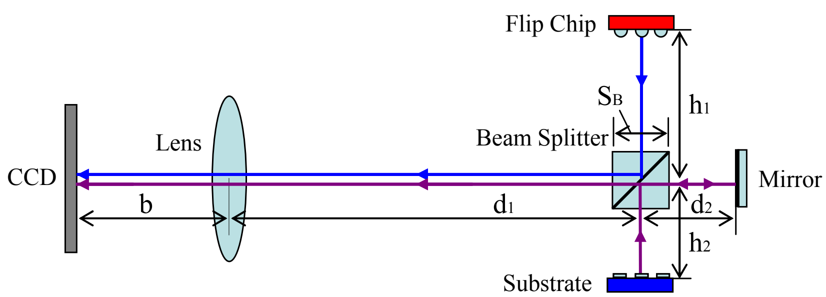


Figure 3: Dual imaging system (DIS)

In Figure 3, b is the distance between center of the lens and the CCD. The lengths of the optical path from center of lens to flip-chip and substrate are expressed by $d_1 + h_1$ and $d_1 + 2d_2 + h_2$ respectively. By applying thin lens model formulas to this system, following equation can be obtained.

$$\frac{1}{d_1 + h_1} + \frac{1}{b} = \frac{1}{f} \qquad (1)$$

$$\frac{1}{d_1 + 2d_2 + h_2} + \frac{1}{b} = \frac{1}{f} \tag{2}$$

where, f is the effective focal length of the zoom lens system. From equation (1), (2), the relation between h_1 and h_2 can be defined as following equation.

$$h_1 = 2d_2 + h_2 \tag{3}$$

The magnification, M of the zoom lens system is

$$M = \frac{b}{d_1 + h_1} = \frac{b}{d_1 + 2d_2 + h_2} \tag{4}$$

From Equation (1) and Equation (4), we can estimate the effective focal length.

$$f = \frac{M(d_1 + h_1)}{M + 1} \tag{5}$$

In zoom lens, the distance c between the object and the image is constant, which equals to $d_1 + h_1 + b = c(\text{constant})$. So equation (4) can be rewritten as

$$d_1 + h_1 = \frac{c}{1 + M} \tag{6}$$

From Equation (5) and Equation (6), the relation between magnification and focal length is defined as following equation.

$$f = \frac{M}{(M + 1)^2} c \tag{7}$$

The above equation shows that the effective focal length is determined by the magnification and the distance between the object and the image. The distance D between the flip-chip and the substrate is expressed as $D = h_1 + h_2 = 2(d_2 + h_2)$, which should be as short as possible for precise flip-chip assembly. If the size of the beam-splitter is S_B, we can choose minimum of the distance d_2 as $S_B / 2$. The distance h_2 includes the half size of the beam-splitter, $S_B / 2$. So h_2 must be larger than $S_B / 2$. Therefore, the distance D between the flip-chip and the substrate depends on the size of the beam-splitter, S_B. The size of the beam-splitter S_B also limits the maximum FOV (Field of view) of camera, which has a role as an aperture-stop of zoom lens. Here, we considered FOV, resolution and magnification of the DIS to determine the size of the beam-splitter, S_B.

In order to carry out precise alignment using visual servoing, the image resolution of DIS is a very important parameter. Considering on the pitch size of flip-chip bumps, proper resolution can be selected. Pitch size is various based on the type of flip-chip. Our test flip-chip's bump diameter is $135\mu m$, and pitch is $204\mu m$. In our system, we choose image resolution from 2 μm /pixel to 12 μm /pixel. For this resolution, we should choose adequate zoom lens and CCD sensor. The magnification of the zoom lens is determined by the relation between the size of CCD pixel plane and FOV of the camera.

$$M \times FOV = P \tag{8}$$

where, P is the size of CCD pixel plane. In our system, a SONY XC-HR300 CCD camera is used. The pixel number is

764×575, the size of CCD pixel plane is $6.4mm \times 4.8mm$ (1/2" type). Based on the image resolution and pixel number, the range of FOV of zoom lens is from 1.6mm to 9mm. From Equation (8), the magnification range is from $0.67\times$ to $4.3\times$.

Considering on FOV and the distance D between the flip-chip and the substrate, the size of beam-splitter can be chosen as $5mm \times 5mm \times 5mm$ cube type. The maximum of FOV of dual imaging system is $5mm$. The distance d_2 between the center of the beam-splitter and the mirror is $2.5mm$, the distance h_2 between the center of the beam-splitter and the substrate is about $4.5mm$. Therefore, the distance D between the flip-chip and the substrate is

$$D = 2(d_1 + h_2) = 2(2.5 + 4.5) = 14mm$$

Here, the beam-splitter is non-polarized. To decrease its weight and size, we use mirror coating on the surface of beam-splitter instead of attaching a mirror. The percents of transmitted light and reflected light are 50%. Only 25 % of the light from the substrate reaches CCD plane. In order to get brighter image, we use high luminance chip-LED near the beam-splitter. Figure 4 shows the internal shape of the dual imaging system with the zoom-lens, which can adjust magnification of image by changing the lens position using stepping motor.

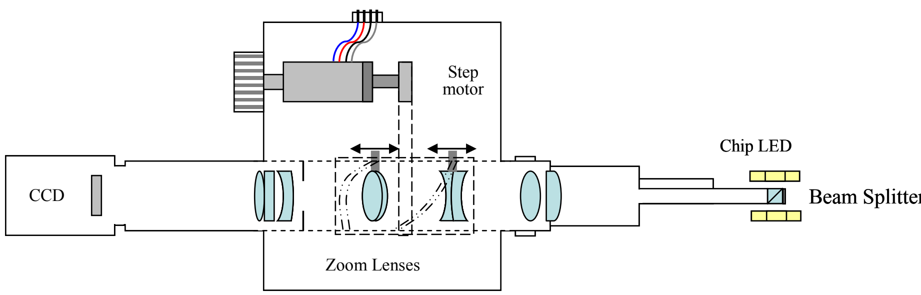
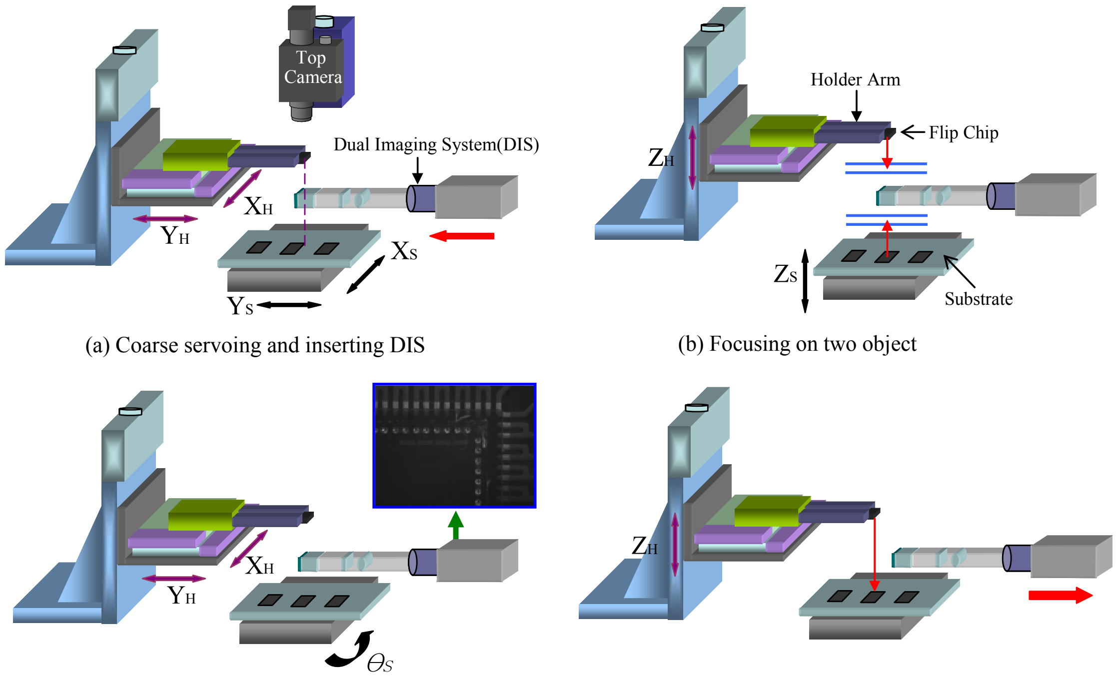


Figure 4: Schematic of dual imaging system

2.3 Assembly process using DIS system

Whole scenario of assembly is shown in Figure 5. In Figure 5(a), coarse visual servoing is conducted with the top camera. The flip-chip and the substrate move to the center of the image roughly. After finishing the coarse visual servoing, the dual imaging system (DIS) is inserted between the flip-chip and the substrate. Both Z_H and Z_S stages moves to focus two objects on the image plane in Figure 5(b). Then, fine visual servoing is conducted to align two objects. After finishing the alignment, the DIS moves out of the flip-chip and the substrate, and Z_H stage moves down to assemble the two objects as shown in figure 5(b). Because we use open-loop control to move the holder arm, kinematical errors in the stage can cause alignment errors. So the size of DIS would be as small as possible to decrease the space between the chip and the substrate.



(a) Coarse servoing and inserting DIS

(b) Focusing on two object

(c) Fine visual servoing

(d) Moving of Z_H axis to assemble

Figure 5: Assembly process of flip-chip

3. FEATURE EXTRACTION AND TRACKING

3.1 Feature extraction

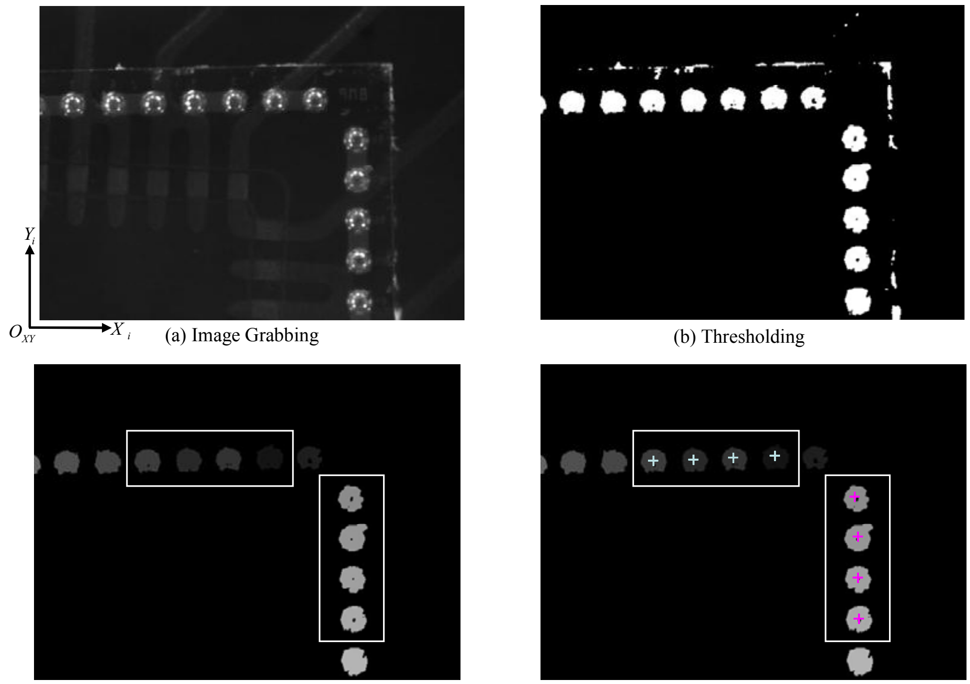
Based on visual servoing, features of flip-chip and substrate should be extracted for flip-chip alignment. These features represent the pose information of the flip-chip and the substrate. Here, we use center points of solder bump and pad on substrate to calculate pose information. As shown in Figure 1, the solder bumps are circle, while the pads on substrate are rectangle. Figure 6(a) shows the composite image of the flip-chip and the substrate using DIS. In this figure, the solder bump of flip-chip must be aligned to the pad on substrate. The figure only shows a part of object because high magnification zoom lens has small FOV, so the corner part of object is used for alignment. The procedures of feature extraction are (1) image grabbing, (2) thresholding, (3) Labeling and window setting (4) extraction of center point as shown in Figure 6(a)-(d). First of all, Thresholding is applied to the image to extract the object from background. Then labeling method is applied to process the binary image to count the number of objects seen in figure. Here only interested object is selected. So the noise is also suppressed. Finally, we can find the center points of the bumps and the patterns in these windows.

We use $(p+q)$ order moment m_{pq}, to find center points.

$$m_{pq} = \sum_i \sum_j x_i^p y_j^q, \qquad (x_i, y_j) \in R$$

$$x_c = \frac{m_{10}}{m_{00}}, \qquad y_c = \frac{m_{01}}{m_{00}}, \qquad (x_c, y_c) : \text{center point} \tag{9}$$

Here, R is region of object in pixel plane.



(a) Image Grabbing (b) Thresholding

(c) Labeling and Window Setting (d) Extraction of Center Point

Figure 6: Procedure of feature extraction

3.2 Definition of coordinate systems

The purpose of definition of coordinate systems is to get the relative position and orientation error between the flip-chip and the substrate. We can use these errors for visual servoing in next chapter.

In Figure 7, we define coordinate systems of the flip-chip and the substrate using feature points of each object in image coordinate (X_i-Y_i). To define coordinate systems, we extract at least two feature points in vertical and horizontal direction on the flip-chip and the substrate respectively. In Figure 7, we get vertical array of center points $\{ P_{v1}, P_{v2}, P_{v3}, \cdots, P_{vn} \}$ of the pattern on substrate, and fit a vertical line L_{Pv} passing through these points using least square method. Horizontal line L_{Pu} also can be obtained by horizontal array of center points $\{ P_{u1}, P_{u2}, P_{u3}, \cdots, P_{un} \}$ of the pattern on substrate. So we define coordinate system of the pattern on substrate using lines $\{ L_{Pu}, L_{Pv} \}$ and cross point (O_P). In the case of the solder bumps, the same method can be applied to obtain coordinate system of the flip-chip using lines $\{ L_{Bu}, L_{Bv} \}$ and origin point (O_B). Above explanation is expressed by the following equations.

- Center point of flip-chip's bump: $B_{ui} = \{x_{ui}^B, y_{ui}^B\}$, $B_{vi} = \{x_{vi}^B, y_{vi}^B\}$, $i = 1, 2, ..., n, \quad n \geq 2$
- Center point of PCB substrate: $P_{ui} = \{x_{ui}^P, y_{ui}^P\}$, $P_{vi} = \{x_{vi}^P, y_{vi}^P\}$, $i = 1, 2, ..., n, \quad n \geq 2$ (10)

So the coordinate system of flip-chip is

- Origin point of coordinate: O_B
- Axes: $L_{Bu} = \{B_{u1}, B_{u2}, B_{u3}, \cdots, B_{un}\}$, $L_{Bv} = \{B_{v1}, B_{v2}, B_{v3}, \cdots, B_{vn}\}$, $\quad n \geq 2$ (11)

The coordinate system of substrate also can be expressed by

- Origin point of coordinate: O_P
- Axes: $L_{Pu} = \{P_{u1}, P_{u2}, P_{u3}, \cdots, P_{un}\}$, $L_{Pv} = \{P_{v1}, P_{v2}, P_{v3}, \cdots, P_{vn}\}$, $\quad n \geq 2$ (12)

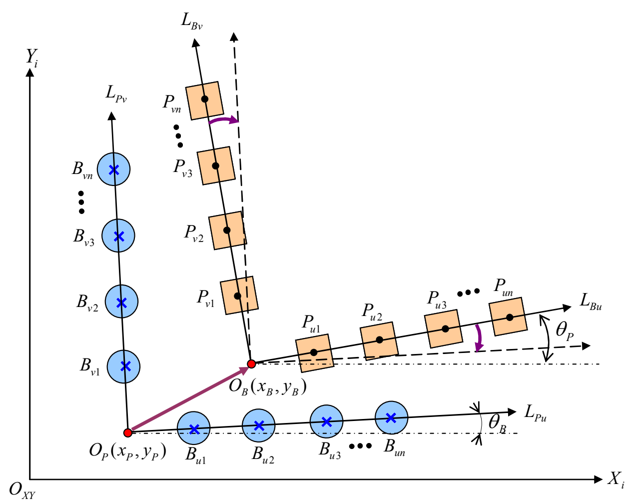


Figure 7: Coordinate systems of flip chip and substrate

3.3 Feature tracking

The extracted features must be tracked during visual servoing in order to obtain the positions of it on-line. We adopt a simple feature tracking method based on SSD (Sum of Squared difference)[3]. It generates feature template that contains feature point. We find next feature template by finding the minimum SSD value of two templates. Let us define a current feature point as $\mathbf{P}_i = (x_i, y_i)$, so the next feature point can be expressed as $\mathbf{P}_{i+1} = (x_{i+1}, y_{i+1})$. If we assume that the movement vector of feature point between current and next frame is $\mathbf{d} = (\Delta x, \Delta y)$ as shown in Figure 8, the coordinates of $\mathbf{P}_{i+1}$ will be $x_{i+1} = x_i + \Delta x$ and $y_{i+1} = y_i + \Delta y$. The vector $\mathbf{d} = (\Delta x, \Delta y)$ can be calculated by searching a position that has minimum SSD valve as following.

$$SSD(\mathbf{P_a}, \mathbf{d}) = \sum_{m,n \in N \times N} [I_i(x_i + m, y_i + n) - I_{i+1}(x_i + \Delta x + m, y_i + \Delta y + n)]^2 \tag{13}$$

where, $N \times N$ indicates the feature template around the feature, I_i and I_{i+1} is an intensity function of current and next frame. This SSD value has a minimum value when current and next SSD feature templates are matched each other.

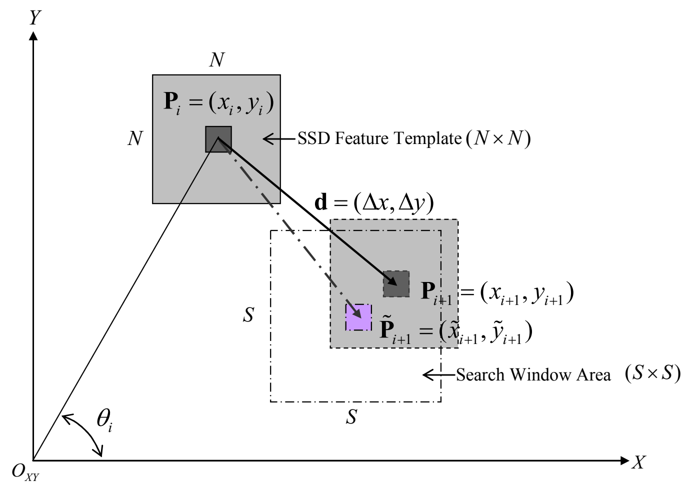


Figure 8: SSD template matching and estimation of search area

In order to reduce the searching time and improve accuracy, an estimation algorithm is used to search the region near in the next feature. In Figure 8, we estimate the next feature point $\tilde{\mathbf{P}}_{i+1} = (\tilde{x}_{i+1},\ \tilde{y}_{i+1})$ based on the similarity transformation and velocities ($\dot{x}_i$, $\dot{y}_i$, $\dot{\theta}_i$) of current feature. That is,

$$\begin{bmatrix} \tilde{x}_{i+1} \\ \tilde{y}_{i+1} \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\Delta\tilde{\theta} & -\sin\Delta\tilde{\theta} & \tilde{t}_x \\ \sin\Delta\tilde{\theta} & \cos\Delta\tilde{\theta} & \tilde{t}_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \tag{14}$$

where, $\tilde{t}_x$ and $\tilde{t}_y$ is the estimated translation distance. $\Delta\tilde{\theta}$ is the estimated rotation angle. They can be expressed as following equation,

$$\tilde{t}_x = \dot{x}_i \cdot \Delta t\,,\ \ \tilde{t}_y = \dot{y}_i \cdot \Delta t\,,\ \ \Delta\tilde{\theta} = \dot{\theta}_i \cdot \Delta t \tag{15}$$

Based on the next estimated feature point $\tilde{\mathbf{P}}_{i+1}$, we define a search window area($S \times S$) around it. In this search window area, we can find next SSD feature template which has minimum SSD value.

4. VISUAL SERVOING FOR FLIP-CHIP ALIGNMENT

In generally, visual servoing uses the error between the target feature point and the current tracking feature point to compute velocity screw[4]. But in this paper, we define two coordinate systems with features of flip-chip and substrate. The errors of orientation and position can be obtained from these coordinate systems. In Figure 7, the substrate rotates to align the coordinate of substrate to the coordinate of flip-chip. Origin point of flip chip's coordinate system is used for translation velocity control. Translation and rotation can be done separately or simultaneously. Figure 9 shows the flow chart of visual servoing for flip-chip alignment.

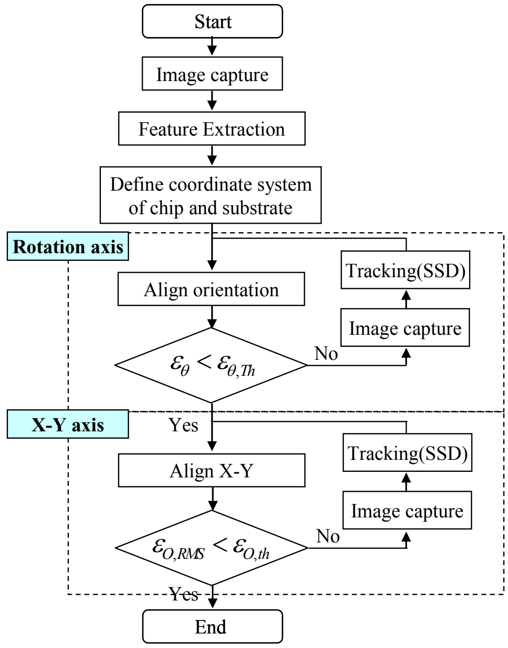


Figure 9: Flow chart of visual servoing for flip-chip alignment

We derive orientation angles θ_B and θ_P from gradients of line L_{Bu} and L_{Pu} based on the image coordinate(X_i-Y_i), so orientation error is $\varepsilon_\theta = \theta_P - \theta_B$. If the position of O_P is $\{x_P, y_P\}$ and the position of O_B is $\{x_B, y_B\}$ in image plane, the desire position of visual servoing is $O_P\{x_P, y_P\}$, and the control position is $O_B\{x_B, y_B\}$. The position errors of two coordinate systems are $\varepsilon_{Ox} = x_P - x_B$ and $\varepsilon_{Oy} = y_P - y_B$. The RMS (Root Mean Square) error of position is

$$\varepsilon_{O,RMS} = \sqrt{(x_P - x_B)^2 + (y_P - y_B)^2} \tag{16}$$

We adopt velocity control using above position errors for translation. The velocity of $O_B\ \{x_B, y_B\}$ can be expressed by $\dot{x}_B$ and $\dot{y}_B$ in pixel coordinate, and each velocity in the world coordinates is V_X, and V_Y. So the relationship between feature velocity in pixel coordinate and that in the world coordinate is image Jacobian, as following equation.

$$\begin{bmatrix} \dot{x}_B \\ \dot{y}_B \end{bmatrix} = \mathbf{J}_\mathrm{V} \begin{bmatrix} V_X \\ V_Y \end{bmatrix}, \quad \mathbf{J}_\mathrm{V} = \begin{bmatrix} K_x M & 0 \\ 0 & K_y M \end{bmatrix} \tag{17}$$

where, K_x and K_y is the scaling factors of the pixel coordinates to the image coordinates in x, and y direction, respectively. M is the magnification of DIS. Based on the above Jacobian matrix, we get velocities input for stage using proportional control law in the world coordinates.

$$\begin{bmatrix} V_X \\ V_Y \end{bmatrix} = \mathbf{K}\mathbf{J}_\mathrm{V}^{+} \begin{bmatrix} x_P - x_B \\ y_P - y_B \end{bmatrix} \tag{18}$$

Here, $\mathbf{K}$ is the proportional gain. The holder arm will be controlled based on Equation (18).

Rotation velocity is derived using orientation error of two coordinate systems. Angular velocity of all feature point is same in pixel plane and world coordinate. Because angular velocity is not affected by other parameters such as lens magnification, we get angular velocity (ω_Z) input for stage by multiplying gain K_W to the relative orientation error of two coordinate system, that is $\omega_Z = K_W(\theta_{Pu} - \theta_{Bu})$. In Figure 5(c), the rotation stage will be controlled based on this equation.

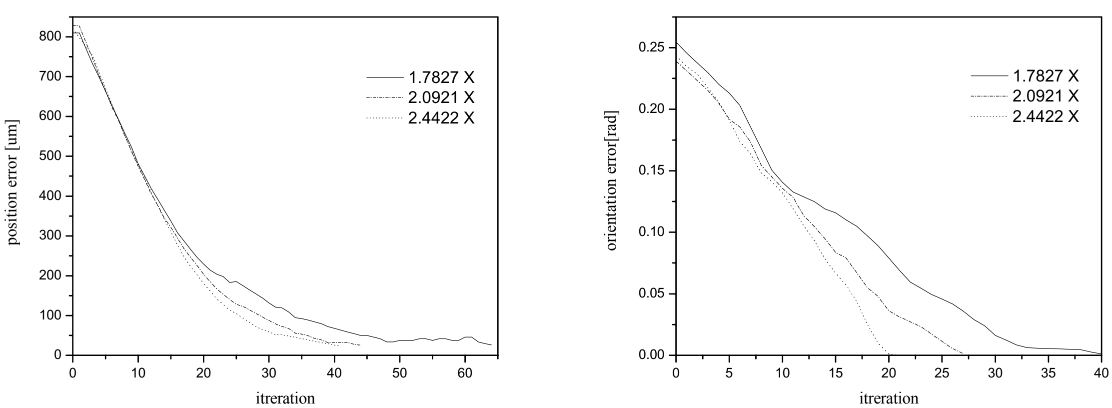
5. EXPERIMENTAL RESULTS

5.1 Effects of magnification

In order to evaluate the effects of magnification to the visual servoing process, a series of experiment in three different magnifications, $1.7827\times$, $2.0921\times$ and $2.4422\times$ is implemented. The initial position of all cases is same. The iteration time is 348*msec*. The visual servoing process stops when the position error in pixel coordinate calculated from Equation (16) is smaller than 7 pixels. The final real errors at different magnification are 26.34 μm at magnification of $1.7827\times$, 25.40 μm at magnification of $2.0921\times$, 22.79 μm at magnification of $2.4422\times$. The final angular errors at different magnification are 0.0018 rad, 0.0014 rad, 0.0012rad respectively. Figure 10 shows graph of the position and orientation error of visual servoing at different magnifications.

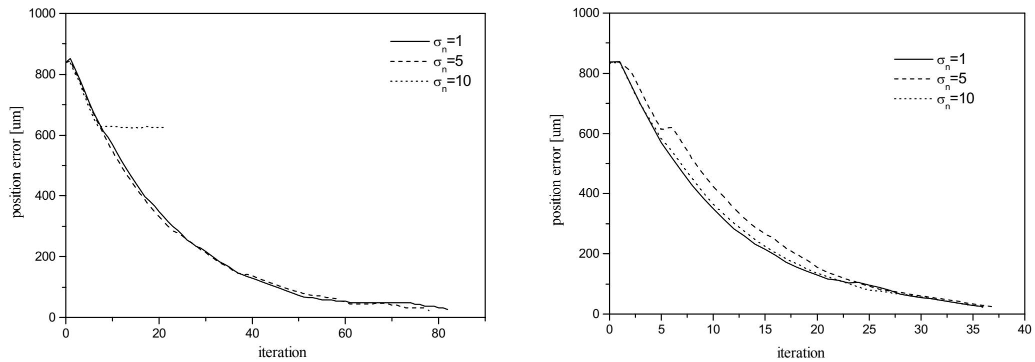
5.2 Effects of size of SSD search window

In chapter 3, we proposed an algorithm for SSD tracking and estimation of feature point. We add Gaussian noise with different standard deviation σ_n to the estimation value of next feature point. Figure 11 shows the result of position error at different size of search window and different noise level. It shows that the large searching area is robust to the noise influence. When the searching area is 15X15 pixels, visual servoing fail when σ_n equals to 10. But the large searching area will also increase searching time. So we should make a treadoff between robust and speed.



(a) Position error vs. iteration (b) Orientation error vs. iteration

Figure 10: Position and orientation error vs. iteration at different magnification



(a) 15X15 searching window (b) 25X25 searching window

Figure 11: Position error vs. iteration at different size of search window and different noise level

6. CONCLUSIONS

In this paper, we design the dual imaging system (DIS) for flip-chip alignment which can observes the bumps of flip-chip and the pads on substrate in one image. We can extract the features of the bumps of flip-chip and the pads on substrate on the same time. In order to get accurate pose information of flip-chip and pads on substrate, we define two coordinate systems for flip-chip and substrate, respectively. The propose feature tracking method apply the similarity transformation to predicate the position of SSD feature template, which improve the accuracy and speed of tracking. In the future, the dynamic performance of the system will be researched. More general and fast tracking algorithm will also be researched.

7. ACKNOWLEDGMENTS

This research was supported by KITECH(Korea Institute of Industrial Technology) from June, 2003~September, 2005.

REFERENCES

1. K. Boustedt, E. J. Vardaman, *Tomorrow's Packaging - Chip Scale Packaging vs Flip Chip*, Microelectronics international: journal of ISHM - Europe, vol. 14, no. 3, pp. 31-32, 1997.

2. John H. Lau, *Flip Chip Technologies*, McGraw-Hill., 1995.

3. K. Nickels, S. Hutchinson, *Estimating uncertainty in SSD-based feature tracking*, Image and vision computing, vol. 20, no. 1, pp.47-58, 2002.

4. G. D. Hager, Seth Hutchinson, *A Tutorial on Visual Servo Control*, IEEE Transactions on Robotics and Automation, Vol 12, No. 5, pp. 651-670, 1996.

5. Patra, S. K., and Y. C. Lee, *Modeling the Self Alignment Mechanism in Flip Chip Soldering. II Multichip Solder Joints*, Proceedings of 41st Electronic Components and Technology Conference, pp. 783-788, 1991.