AN AUTOMATIC POSITION RECOGNITION TECHNIQUE FOR LSI ASSEMBLY

Michihiro Mese, Takafumi Miyatake, Seiji Kashioka, Masakazu Ejiri

Central Research Laboratory, Hitachi Ltd.

Isamu Yamazaki

Equipment Developing Department, Semiconductor and Integrated Circuit Division, Hitachi Ltd.

Toshimitsu Hamada

Production Engineering Research Laboratory, Hitachi Ltd.

Abstract

A fully automatic LSI wire-bonding system utilizing image processing techniques is described. The position recognition system developed is composed of measuring stations, a recognition device, a minicomputer, and cassette tape-recorders. The position of the bonding pad is found, generally in real-time, by the recognition device and computer. The computer calculates positional data of all bonding pads and stores them on a cassette tape for a subsequent wire-bonding process. The position recognition methods discussed are composed of a thresholding algorithm and a position-finding algorithm. These methods are evaluated from the viewpoint of thresholding stability, recognition rate and positional accuracy.

1, Introduction

Recognition of object position and posture has been recognized as a vital technique in the machine perception field. Many studies have been conducted to date in this field and have led to the development of practical techniques such as bolt-tightening robots(l), printed-eircuit board drilling machines(2), wheel-fitting machines(3), as well as others(4)(5).

One critical area in which machine perception research is expected to prove especially valuable is automated semiconductor assembly. This is because semiconductor devices have gotten steadily more sophisticated and have been miniaturized to the point where human operation is becoming increasingly problematic. Along this line, success has already been realized in developing a new transistor assembly system(6), the first practical application of a group-controlled vision system.

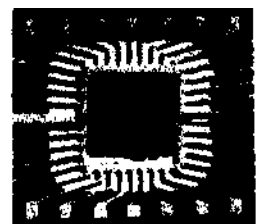
This paper deals with an approach to the assembly of other semiconductor devices, especially LSI chips. It has been developed in parallel with the transistor assembly system by the same machine perception group. To date, almost all LSI assembly processes have already been automated, the notable exception being the position measuring station for wire-bonding. The position recognition technique in this study is designed to rectify this situation, opening the way for fully automated LSI wire-bonding systems and, consequently, more reliable LSI products.

2. Outline of LSI Wire-bonding

An example of an LSI lead-frame prior to assembly is shown in Figure 1(a). The close-up photograph in Figure 1(b) shows bonding pads distributed in the peripheral portion of the LSI chip. The wire-bonding process is where gold or aluminum wires are bonded and stretched between the bonding pads on the chip and the outer leads on the lead-frame. This process is carried out by numerically-controlled high-speed wire-bonding machines at a speed of 0.2 seconds/wire by using previously-measured positions of bonding pads. However, the measurement is still made by operators; either by manually pointing out a typical corner of the chip on a projected image screen or by manually aligning the chip image with the two cross-lines on a magnified view.

The position recognition technique developed here uses two TV cameras to scan images of two different peripheral portions of a single LSI chip fed automatically into the measuring station. The position of a bonding pad in each field of vision is then determined approximately by a special image processing hardware and a minicomputer. The computer then calculates positional data of all bonding pads and records them on magnetic cassette tape for subsequent bonding.

(a) LSI lead-frame



(b) LSI chip (Bonding-pad are distributed in the peripheral portion of the chip.)

Fig, 1 LSI lead-frame and bonding pads,

Initial positional errors, due mainly to previous die-bonding processes, of +150μrn in the X and Y directions and ±30' in an angular direction are inevitable and must be tolerated at the measuring station. The position recognition technique must compensate for the above loading error. It must also have a position finding accuracy of less than ±10 μm to ensure reliable wiring at all bonding pads. Furthermore, the recognition time and recognition rate should be less than 0.3 seconds/chip and more than 99.9 %, respectively, to accommodate various production condition.

3. Position Recognition Method

The image input system for LSI chip position recognition is shown in Figure 2. In order to assure adequate resolution, two TV cameras and microscope lenses were installed above the LSI chip, which was illuminated by a lamp. Shading and geometric distortions were almost completely eliminated by compensation circuitry. The size of each visual field was set to 800 urn horizontally and 600 μrn vertically. For our purposes, the size of an LSI chip ranged from 3 to 6 mm square. To accommodate various LSI types and sizes, two independent XY tables were actuated to adjust the camera positions and distances whenever product changoovers, e.g. from 3 to 6 mm chips, become necessary.

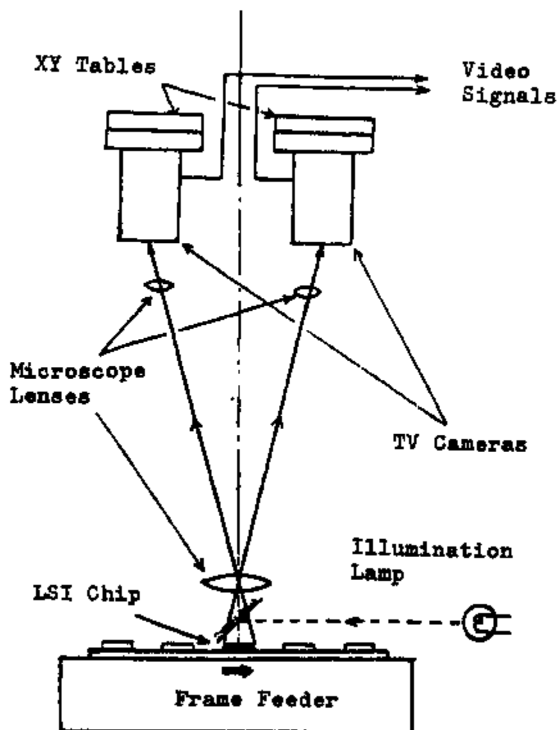


Fig. 2 Image input system for position recognition of LSI chip.

In the process, a frame-feeder feeds LSI chips one by one, into the system's field of vision. An analog video signal containing target pad information can be obtained from each TV camera. The video signal from each TV camera is sampled and the effective scanning field is divided into 320 x 240 picture elements. Thus, one picture element corresponds to *2.5* μm square. The image processing technique adopted here has two major functions.

One stably thresholds the analog video signal into a binary image. This thresholding is particularly important in practical image processing. This is because a very simple and high-speed image processor can be realized if stable binary video signals can be assured in spite of a wide range of illumination intensities, reflection ratio changes for different chips and other variations.

The other function detects the position of the bonding pad in the image with high accuracy and reliability. This position detecting algorithm consists of two successive steps i.e. (1) macroscopic and (2) microscopic. In the first step, the positions of the bonding pads are detected approximately by using 80 x 60 picture elements sampled in the X and Y directions. More accurate positioning is accomplished in the second step with the original picture element resolution. Finally, a geometric check is performed for the positional data of an arbitrary pair of pads between the visual fields. When a reasonable pair is found, all other pad positions are calculated from known geometric relations of the pads in the chip.

4. Thresholding Algorithm

(a) Analog image (b) Binary image

Fig. 3 Images of peripheral part of LSI chip.

The binary image obtained by the "balanced threshold method", developed here, is shown in Figure 3(b). The aluminum bonding pad and other parts are discriminated clearly in the figure.

The "balanced threshold method" determines a threshold by successive approximations. First, an arbitral threshold level is chosen in advance, then the average values higher and lower than the threshold level are calculated. Secondly, utilizing these average values, the threshold level is modified to a new one. This process is continued N times until an approximate equilibrium threshold level is attained; where N is a constant. This process is expressed as the following iteration equations:
For n=0,1,2,...,N,

$$\theta^{(n+1)} = \theta^{(n)} + \beta \Delta f_W^{(n)} + (1-\beta)\Delta f_B^{(n)} \qquad (1)$$

$$\Delta f_W^{(n)} = \int_{t\in T_W^{(n)}} (f(t) - \theta^{(n)})dt \Big/ \int_{t\in T_W^{(n)}} dt \qquad (2)$$

$$\Delta f_B^{(n)} = \int_{t\in T_B^{(n)}} (f(t) - \theta^{(n)})dt \Big/ \int_{t\in T_B^{(n)}} dt , \qquad (3)$$

where $f(t)$ is the analog video signal from a TV camera, $\theta(n)$ is the n-th threshold level, $\Delta f_W(n)$ and $\Delta f_B(n)$ are the n-th average values higher and lower than $\theta(n)$, β and $1-\beta$ are constants that relate to the weight of $\Delta w(n)$ and $\Delta f\theta(n)$, and $T_W(n)$ and $T\theta(n)$ are time periods where the image signal is greater and less than $\theta(m)$.

This algorithm can be explained more concretely by using a simple signal model to illustrate the meaning of these equations. A signal model with only two brightness levels, f_W and f_S, whose time periods are T_W and T_B respectively is shown in Figure 4. Note that those time intervals are closely related to areas of the scanned image. When $f9<\theta(n)<f_W$ and when the second and third terms of Equation (1) are rewritten as:

$$\Delta\theta^{(n)} = \Delta f_W + (1-\beta)\Delta f_B , \qquad (4)$$

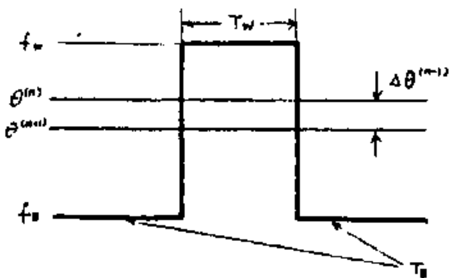


Fig. *k* Signal model with two brightness levels.

$\Delta\theta^{(n)}$ for this particular signal model becomes:

$$\Delta\theta^{(n)} = \beta f_W + (1-\beta) f_B - \theta^{(n)}. \qquad (5)$$

Let θ^* be the equilibrium threshold, that is, $\theta^{(n)}$ so that $\Delta\theta^{(n)}=0$ in Equation (5). Solving Equation (5),

$$\theta^* = \beta f_W + (1-\beta) f_B . \qquad (6)$$

If $\Delta\theta(n)$ becomes comparatively small after several iterations, the threshold given by Equation (I) approximates to the equilibrium threshold.

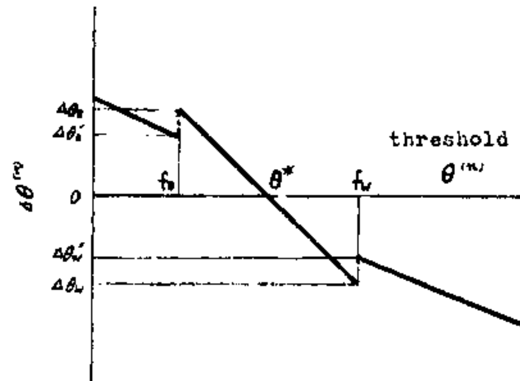


Fig. 5 Relation between *B(*'* and *AO^* for simple signal model shown in Fig, *k*

The relation between 0(~'and *A9{v)foT* this model is illustrated in Figure 5. Among the values *A6y* , J#B , *AQ^* and *bO'^* at the discontinuous points in the figure, the following equations hold:

$$\Delta\theta_W = -(f_W - f_B)\beta$$

$$\Delta\theta_B = (f_W - f_B)$$

$$\Delta\theta_W' = -\frac{T_B}{T_W + T_B}(f_W - f_B)\beta \qquad (7)$$

$$\Delta\theta_B' = \frac{T_W}{T_W + T_B}(f_W - f_B)\beta .$$

Thus, the following relation can be obtained from Equation (7):

$$\Delta\theta_W < \Delta\theta_W' < 0 < \Delta\theta_B' < \Delta\theta_B . \qquad (8)$$

Therefore, *0* is unique regardless of the initial threshold value, as shown in Figure 5.

However, the images of LSI chips are not as simple as the model because they usually include various brightness levels. Suppose that the image of an LSI chip involves m major brightness levels. Then it can be shown that, at most, m-1 equilibrium thresholds exist.

However, illumination and reflection are stable enough that the brightness of the aluminum

is always between an upper bound ξ and a lower bound η . Thus the time periods $T_W^{(n)}$ and $T_B^{(n)}$ should be changed into $T_W^{(n)}$ and T , where $\xi \geq f(t) \geq \theta^{(n)}$ and $\eta \leq f(t) \leq \theta^{(n)}$, respectively.

In this case, the video signal for an LSI chip can be approximated by the model in Figure 4. A stable binary video signal, which can cope with the various changes in practical production conditions e.g. illumination changes, reflection ratio changes for each LSI chip, sensitivity drift of the TV cameras, can be obtained by the above "balanced threshold method".

In order to apply the above algorithm to this system, Equations (2) and (3) are converted into sums:

$$\Delta f_W^{(n)} = \sum_{i \in T_{W\xi}^{(n)}} (f_i - \theta^{(n)}) / \sum_{i \in T_{W\xi}^{(n)}} \quad (9)$$

$$\Delta f_B^{(n)} = \sum_{i \in T_{B\eta}^{(n)}} (f_i - \theta^{(n)}) / \sum_{i \in T_{B\eta}^{(n)}} , \quad (10)$$

where $T_{W\xi}^{(n)}$, $T_{B\eta}^{(n)}$ are the digitized time periods for $\xi \geq f_i \geq \theta^{(n)}$ and $\eta \leq f_i \leq \theta^{(n)}$, respectively.
The numerators and denominators in Equations (9) and (10) are easily obtained by counting picture elements as image scanning proceeds. The division and multiplication steps in Equations (1), (9) and (10) are executed by software within the blanking period.

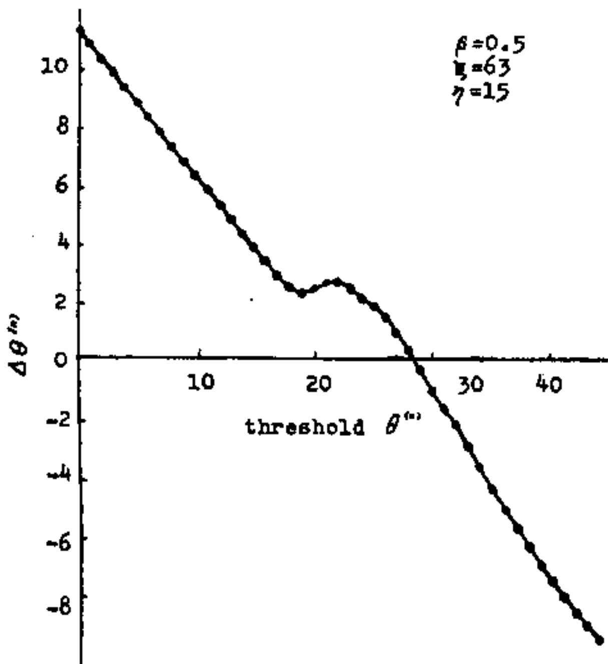


Fig. 6 Relation between $\theta^{(n)}$ and $\Delta\theta^{(n)}$ for an actual LSI chip.

An actual example of the relation between $\theta^{(n)}$ and $\theta\Delta^{(n)}$ for an LSI chip is shown in Figure 6. Here, the video signal is quantized into 64 levels and typical parameters for this case are shown.

This thresholding method is essentially equivalent to the method which utilizes the histograrn(7) of the LSI chip brightness distribution. However, the processing speed is much faster than that of the histogram method.

5. Position Detection Algorithm

The bonding pad of an LSI chip is typically 120 um square. This corresponds to 48 x 48 picture elements in its image. Adjacent to the square portion, the pad has a protruding part on one side edge that joints to inner patterns. In addition the pad usually has a scratch on its surface, as shown in Figure 3, caused by the needle probe used in the previous wafer inspection process.

The position detecting algorithm should accommodate these typical features of bonding pads and stably detect the position of the pads. The algorithm consists of two parts: the macroscopic and microscopic processes. In the macroscopic process, the roughly-sampled image is used, and a pattern matching method is applied. In the microscopic, several rectangular domains are generated, and the white areas in the domains are calculated to determine the position more accurately.

5.1 Macroscopic processing

In the macroscopic, the binary video signal of an LSI chip is sampled every four elements to yield an 80 x 60 element image. In this case, a pad corresponds to a 12 x 12 element. A preprocessing noise-elimination step is performed on this roughly-sampled image. A template for this procedure is illustrated in Figure 7. When the number of white picture elements among P_0 , P_1 , P_2 , . . . , P_8 is less than four, P_0 is regarded as black, otherwise it is regarded as white. Thus a simple binary image is obtained in which uneven boundaries and isolated noises are eliminated.

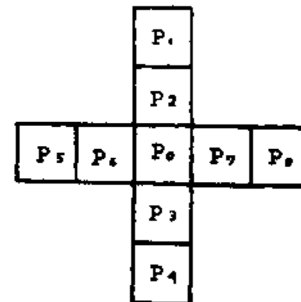


Fig. 7 Template for noise elimination.

Next, the position of the bonding pad is detected by a pattern matching method, as illustrated in Figure 8. Four templates *kl* , A2' A3 and A4 are scanned in parallel over the input image pattern to determine whether the input pattern is matched. The size of each template is 5 x 5 picture elements. Its standard pattern is equal to each corner of the bonding pad to be detected. Let Uij be a standard pattern, Dij be a "don't-care" pattern and Iij be an input pattern corresponding to a domain Ak (where k=l,...,4). The matching result Mk is given as follows:

$$M_k = \bigwedge_{(i,j)\in A_k} (\overline{U_{ij} \odot I_{ij}} \vee D_{ij}). \tag{11}$$

This equation means that the Mk in the template Ak is "1" when Uij matches Iij except for Dij="l", otherwise "0".

However, at the preceding wafer inspection stage every bonding pad is scratched by a needle probe. Consequently, the pattern matching does not always hold for all four corners. Thus, regarding logic variable Mk numerically, the following numerical sum M is calculated:

$$M = \sum_{k=1}^{4} M_k. \tag{12}$$

If M> 3, the pattern is regarded as a bonding pad. Therefore, the pad is easily detected, even if one of the four coners lacks its normal shape as shown in Figure 8.

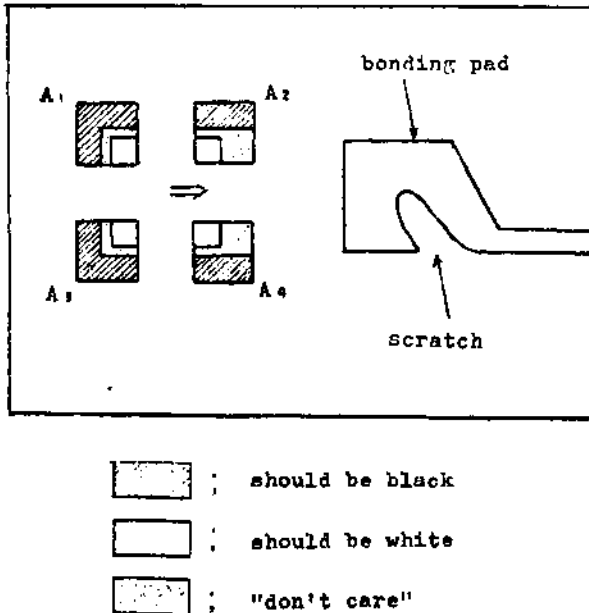


Fig. 8 Position finding method by pattern matching.

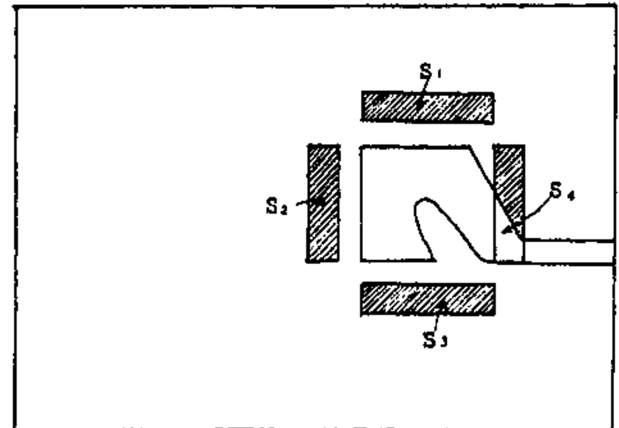
5.2 Microscopic processing

Strictly speaking, the macroscopic processing finds only pad candidates. This is because there can be other patterns in the image which resemble pads.

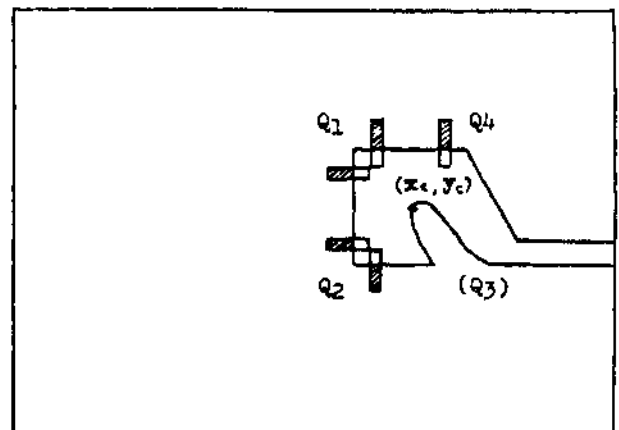
The first microscopic processing step checks and classifies the type of pads by investigating the joint terminal direction of the candidate for every picture element. In the process four rectangular domains are generated around the candidate, as illustrated in Figure 9(a). In the process, four rectangulrr domains are generated around the candidate. White areas S1 , *S2* , S3 , and S4 are then measured. In the example of the figure, the following conditions are checked:

$$S_1, S_2, S_3 < S_B > S_u \quad (1?)$$

where SB and Sw are appropriate threshold values. Only when relation (13) holds, the candidate is accepted as a bonding pad.



(a) Direction finding of joint terminal of candidate.



(b) Position modification of every corner.

Fig. 9 Microscopic processing.

The second step detects the position of the pad. For this purpose, several minor rectangular domains are set around the corners as illustrated in Figure 9(b), depending on the results of the previous step. Thus, the X and Y positions of the corners are determined by calculating the white areas in the domains at the resolution of the original picture. For example, when area Sy is detected at a domain whose center position in the Y direction is YMas shown in Figure 10, the precise Y coordinate becomes:

$$y_m = Y_M + c - S_Y / w. \qquad (14)$$

Thus, the position of the corners, except for any missed corners, are determined.

Lastly, the center of the bonding pad is calculated. Here, the joint terminal and the position of the scratch inside the pad are taken into consideration. For example in Figure 9, the center coordinates (xc,yc) of the bonding pad are calculated as follows:

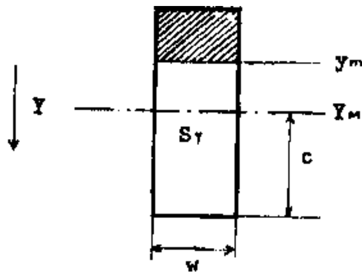


Fig. 10 Example of position modification.

$$xc = \frac{x_1 + x_2}{2} + 24 \qquad (15)$$

$$yc = \frac{1}{2}\left(y_2 + \frac{y_1 + y_4}{2}\right) \qquad (16)$$

Here (x1 , *y1*) and (x2 , y2) are coordinates of corners Q1 and Q2 , and y4 is the Y coordinate of corner Q4 in the figure. The second term in Equation (15) means l/2 of the pad size, in picture element.

6. System Configuration and Results

6.1 Recognition system

A position detection system based on these method has been constructed (see Figures 11, 12).

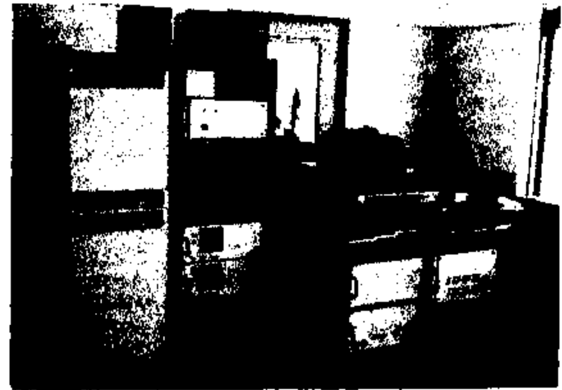


Fig. 11 General view of position recognition system for LSI chips.

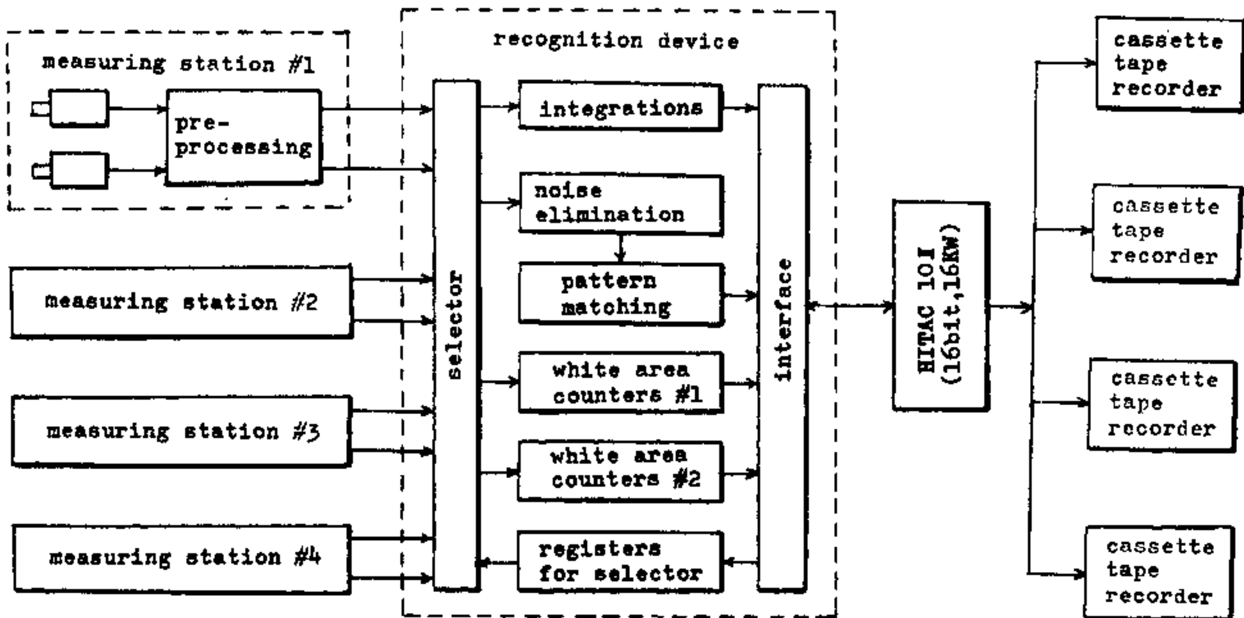


Fig. 12 Block diagram of position recognition system for LSI chip.

It consists of four measuring stations involving image input TV cameras, a recognition device, a mini-computer HITAC-10 II (16 bit, 16 KW) and cassette tape-recorders, as shown in Figure 12. A maximum of four measuring stations can be controlled in parallel by a single recognition device and the mini-computer.

In the recognition device, an integration circuit for thresholding calculates the four integrated values in the numerators and denominators in Equations (9) and (10). This circuit basically consists of an AD converter and four counters. The noise elimination circuit and the pattern matching circuit are the macroscopic processing circuits, and send positions of pad candidates to the computer. In the pattern matching circuit, several line memories are used to pick up a local image in the real-time mode.

In accordance with the positional data of the candidates, the computer sends several sets of rectangular domain positions to the two groups of white area counters which do the microscopic processing. The counted data are sent back to the computer, which checks the terminal direction and calculates the position of the bonding pad. These information transfers are executed by the interface circuit using Direct Memory Access (DMA) as well as multi-level interrupting functions. The selector is used to determine the appropriate measuring stations, TV cameras, and image processing circuits depending on the processing sequence. Some of these are selected in parallel to enable simultaneous processing.

A time chart of the image processing is shown in Figure 13. As illustrated, three fields are devoted to thresholding. Pattern matching is done by one field. Direction finding of the joint terminal is accomplished by one field. At the same time, position adjustment of each corner is effected by one field. One field is equal to 16.7 ms. Thus, total image processing time becomes 133 ms. In addition, however, there is some lag due to the diminishing after-image effect of the TV camera and other software computing time. Therefore, overall position locating time becomes approximately 200 ms at maximum. This is well within the initial target time of 300 ms. The results of each image processing sequence is shown in Figure 14(a) to (d).

6.2 Test result

The recognition methods discussed here are evaluated from three points of view: (1) effectiveness of the thresholding method, (2) recognition rate, and (3) positional accuracy.

(1) In practical production lines, various disturbances must be considered, especially the effects of:
(a) intensity changes of the illumination,
(b) positional change in the image field, and
(c) lens focus.

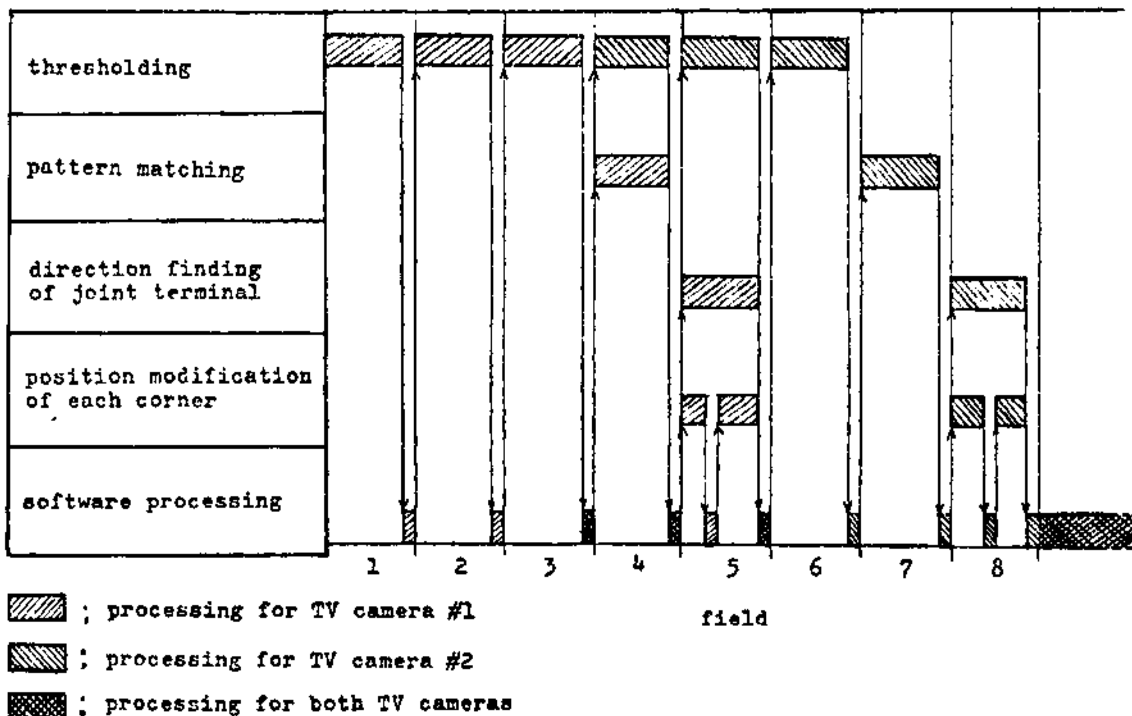
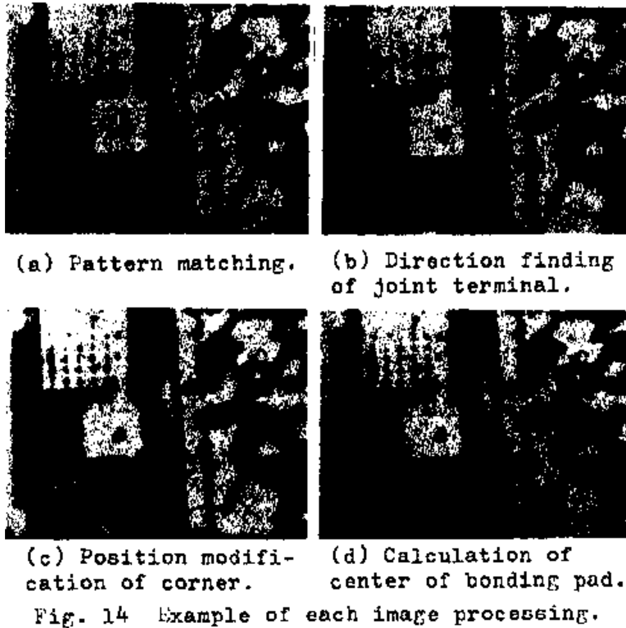


Fig. 13 Time chart of image processing.

The system was tested by changing illumination ± 15 % from its normal intensity. To determine positional dependence, position are altered t] 50µ m along both X and Y directions. To investigate the effect of lens focus, the vertical position of the chip is varied +120µm along the optical axis.

The results are shown in Figure 15. Each cross-sectioned region gives correct operation of the recognition algorithm as long as a threshold value in the region is adopted. Thus, each median axis of the region corresponds to the threshold with the highest margin for pad recognition. As a result of the evaluation, the optimum thresholding parameters are follows: N=3, a=64, b=15, when analog video signals are quantized into 64 intensity levels. The optimum value of *ft* is determined to be 0.55 by finding the minimum standard deviation from the median axis. The relation between *ft* and the corresponding standard deviation is shown in Pipure 16.



(a) Pattern matching. (b) Direction finding of joint terminal.

(c) Position modification of corner. (d) Calculation of center of bonding pad.

Fig. 14 Example of each image processing.

(2) After determining the thresholding parameters, the recognition rate was measured for several types of LSI chips. The results shown in Table 1 demonstrate that the expected recognition rate of 99.9 % has been achieved. The single misrecognition was due to the defective chip whose large white domain happened to resemble a bonding pad.

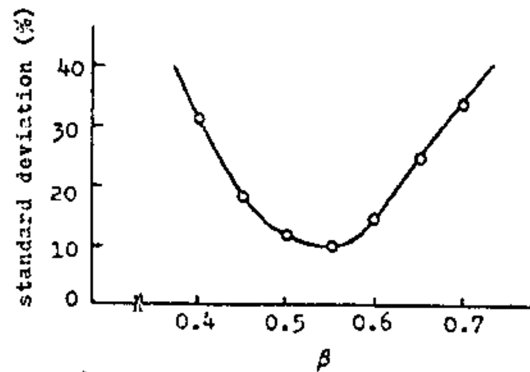
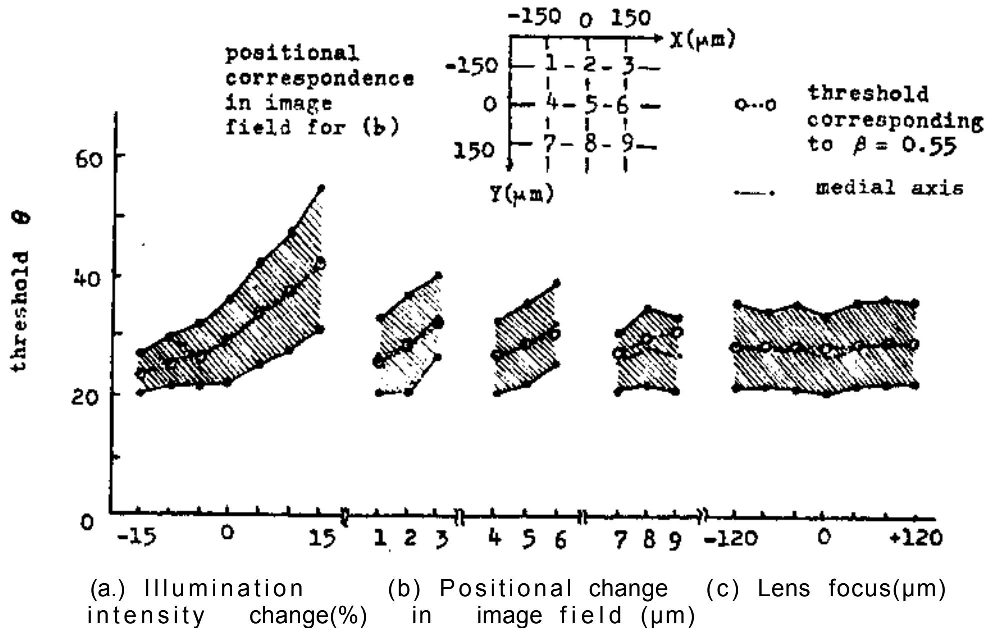


Fig. 16 Relation between β and standard deviation from the medial axis.



(a.) Illumination intensity change(%) (b) Positional change in image field (µm) (c) Lens focus(µm)

Fig. 15 Threshold change due to various disturbance factors.

(3) For positional accuracy testing, three representative types of LSI chips were selected. The results are shown in Table 2. The recognition accuracy was ±5.9 μm, well within the initial target of ± 10 urn.

Table 1 Recognition rate.

	cases	%
total	895	100
correct	894	99.9
rejected	0	0.0
mis-recognition	1	0.1

Table 2 Positional accuracy.

accuracy measured (μm)	r.m.s.	max.
	2.5	5.9

8. Conclusion

Position recognition methods for LSI chips were discussed, as applications of image processing techniques. The methods developed here depends on a stable thresholding algorithm and a two-step position finding algorithm. The recognition rate and accuracy for these methods were well within the targpt ranges. The completion of this system makes high-speed automatic assembly of LSI packages feasible, and avoids human errors encountered in previous positioning techniques.

9. Acknowledgment

The authors wish to thank Jun Suzuki, Department Manager of Equipment Developing Department in Semiconductor and Integrated Circuit Division, Hitachi Ltd., for his insights invaluable arrangements. Thanks are also due to Dr. Hiroshi Watanabe, General Manager, Dr. Jun Kawasaki, Department Manager, of Central Research Laboratory, Hitachi Ltd., as well as many others who contributed to the realization of this system, for their help and encouragement in carrying out this study.

10. References

(1) T.Uno, M.Ejiri and T.Tokunaga: A method of real-time recognition of moving objects and its application, Pattern Recognition, Vol.8, pp.201-208, 1976

(2) J.A.G.Hale and P.Saraga: Control of a PCB drilling machine by visual feedback, Proc. of 4th IJCAI, p.775, Tbilisi, 1975

(3) J.T.Olsztyn et al.; An application of computer vision to a simulated assembly task, Proc. of 1st IJCPR, Washington D.C., 1972

(4) P.Saraga and D.R.Skoyles: An experimental visually controlled pick and place machine for industry, Proc. of 3rd IJCPR, Coronado, 1976

(5) M.L.Baird: An application *of* computer vision to automated *IC* chip manufacture, Proc. of 3rd IJCPR, Coronado, 1976

(6) S.Kashioka, M.Ejiri and Y.Sakamoto: A transistor wire-bonding system utilizing multiple local pattern matching techniques, Trans. of IEEE on Systems, Man. and Cybernetics, Vol. SMC-6, No.8, pp.562-570, Aug. 1976

(7) A.Rosenfeld: Picture Processing by Computer, Academic Press, New York, 1969