Segmentation Algorithm for Wafer ID using Active Multiple Templates Model

In-Mo Ahn*, Dong-Joong Kang**, and Yoon-Tack Chung***

* Division of Computer and Electrical Engineering, Masan College, Masan, KN, Korea
(Tel : +82-55-230-1328; E-mail: imahn@masan.ac.kr)
**Department of Mechatronics Engineering, Tongmyong University, Busan, Korea
(Tel : +82-51-610-8361; E-mail: djkang@tit.ac.kr)
*** Division of Internet Business, Masan College, Masan, KN, Korea
(Tel : +82-55-230-1318; E-mail: ytchung@masan.ac.kr)

Abstract: This paper presents a method to segment wafer ID marks on poor quality images under uncontrolled lighting conditions of the semiconductor process. The active multiple templates matching method is suggested to search ID areas on wafers and segment them into meaningful regions and it would have been impossible to recognize characters using general OCR algorithms. This active template model is designed by applying a snake model that is used for active contour tracking. Active multiple template model searches character areas and segments them into single characters optimally, tracking each character that can vary in a flexible manner according to string configurations. Applying active multiple templates, the optimization of the snake energy is done using Greedy algorithm, to maximize its efficiency by automatically controlling each template gap. These vary according to the configuration of character string. Experimental results using wafer images from real FA environment are presented.

Keywords: wafer ID, segmentation, low-quality image, active multiple template, NGC, Greedy algorithm, snake algorithm

1. INTRODUCTION

Since our industrial structure has been highly developed, the need for automatic inspection of quality control has increased as traditionally it was done manually. In order to develop a vision inspection system that can be applied in the general production process, it is necessary to develop an algorithm which executes high speed pattern matching with stability.

Semiconductor wafers are usually provided with identification information such as product numbers, manufacturing numbers and lot numbers by markings. Such identification numbers are read, recognized, and used for various purposes, such as automatic control of manufacturing procedures, sorting of products, and quality control.

Image segmentation is separating meaningful region from an image. This is regarded as essential and very important stage in computer vision. Until now, many researches about image segmentation has been conducting up to date and various applications in real image is also being carried out[1-5].

In general, there are three methods in image segmentation using computer vision, region based segmentation, edge based segmentation, and segmentation using histogram [1-4].

The method "region based segmentation" is to categorize pixels with similar value to guarantee uniformity in a given area in order to segment the image. The method "edge based segmentation" is to segment the image by area with guaranteed connectivity of the area using the edge elements that rapidly vary the pixel value.

The split and merge method [2], which is typical type of the region based segmentation, divides the image to certain unit, then compares the similarity to the neighboring pixel. If this appears to be the same area then merges the pixels and if it is different, then it split the two areas. Repetition of this process eventually segments the image. Such a method may enable one to obtain relatively accurate results of image segmentation, but takes rather longer time to complete.

Region based segmentation can be done by tracking the contour [3]. From the pixel gradient value and direction information of input images, the image can be divided into contour and region surrounded the contour by following the contour step by step.

Edge based segmentation can segment the image into contour and the inner component without additional processing. In the case when any noise was found, the method may follow a false edge. Thus this thorough separation process is required.

Segmentation by using histogram [4] is mainly used when the distribution of the gray level is simple. The method makes image segmentation easy by carrying out quantization on the whole image as the highest two gray level after obtaining the histogram of all gray levels. Thus this method may have a good segmentation effect when the distribution of histograms is concentrated on the two gray levels which are background and object respectively. Especially, in the automation process, the algorithm used in segmenting the inspection area can be divided into the histogram projection and Hough transformation. Most of such existing algorithm may show a good result in clean images with good visualization whereas the segmentation of test area may be a serious problem in the difficult situation of controlling the lighting.

Such difficulty is caused by vague distribution characteristics of image brightness due to noise and irregular lighting condition.

This paper aims to segment automatically the character area for recognizing the wafer ID under poor visualization condition.

The segmentation of inspection area is pre-processing procedure for recognition and the successful execution of this task will determine the success of whole process.

In low visualization situation, none of the features value may be detected with reliability. Thus applications of the conventional histogram projection or binarization methods become difficult to be applied. Therefore, the character area can be extracted by using template matching. But the method is also difficult to apply for the reason of frequent occurrence of fault matching in poor visualization image.

Thus this piece of work seeks to suggest the multiple template method of how to extract the character area automatically by applying NGC matching to the detection of the ID mark area. This may be inspired by the fact that, the ID part consists of a couple of continuing mark normally and define this as a matter of optimization by searching the patterns.

2. CHARACTERISTICS OF WAFER ID

In the image including wafer identification numbers, contrasts of background and characters are not constant in most case. The several variations in wafer surface cause various changes in the edge of ID number images, in each process, the position of contour of ID numbers also slightly changed. Furthermore, random noises come from the influence of contamination through the transportation mechanism. On the other hand, the position of ID is different, according to the kinds of wafer. The gaps among each character in ID are different and changed above all [5].

The characteristics of wafer ID in the field of view are summarized as follows:

- Position of ID is not fixed
- Orientation by 2° is possible.
- Patterns on background are very complex.
- Gaps between characters are not same.
- Neighbor image is very noisy.
- Shape of characters is comparatively stable, due to a regular font.

3. PATTERN MATCHING ALGORITHM

Vision industries and researchers have employed a mathematical procedure known as normalized gray-scale correlation (NGC) to locate the reference pattern within an image under consideration.

Over the years, *NGC* has been proved to be a robust and reliable method [6]. If sub-pixel accuracy is required, the correlation surface may be interpolated and accuracy of better than 1/16th of a pixel can be achieved. This conventional *NGC* is invariant to linear changes in brightness, but that is where its invariance ends. That is, conventional *NGC* has very little tolerance to any changes in rotation, scale, perspective distortion, non-linear changes in brightness, and multilayer buildup in wafer manufacturing. In addition, *NGC* is computationally very expensive ($O(n^4)$ to find translated patterns). The formula for calculating the *NGC* score *NC* is:

$$NC = \frac{\sum_{x,y}(I-\bar{I})(T-\bar{T})}{\sqrt{\sum_{x,y}(I-\bar{I})^2\sum_{x,y}(T-\bar{T})^2}} \qquad (1)$$

where $\bar{I} = \sum I / N$ and $\bar{T} = \sum T / N$

T and *I* are template and image data, respectively, and *N* is the number of template pixels [7].

It is possible to compare the patterns by normalizing the intensity of two patterns. The *NGC* equation used in this paper is eq. (2), where *N* is the pixel number of model image, *I* is the gray level of input image pixels, and *M* is the gray level of model pixels.

$$C = \frac{N\sum IM - (\sum I)\sum M}{\sqrt{[N\sum I^2 - (\sum I)^2][N\sum M^2 - (\sum M)^2]}} \qquad (2)$$

Recognition rule is as follows.

At first, the system calculates the similarity of an input pattern and *L* reference patterns and then arrange in order of value. The recognition result is to be the reference patterns that have maximum similarity. Minus value is excluded, because the calculation of square root increases the processing time. Final matching value, square of C, *Score* becomes following as eq. (3).

$$Score = [\max(C,0)]^2 \qquad (3)$$

Fig. 1 is the process of comparison of character patterns using normalized grayscale correlation method. Input image is "3" and 0~9 templates are applied to it repeatedly, similarity are checked.

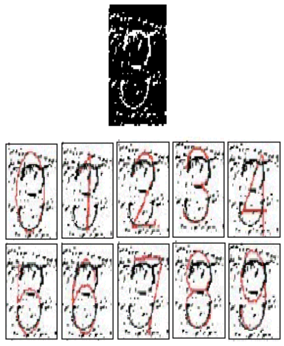


Fig. 1 Comparison process of patterns

4. SEGMENATION BY MULTIPLE TEMPLATES

Prerequisite preprocessing is segmentation of a character region from its background. But in the bad visualized image, preprocessing is difficult. Fig. 2 is an example of a bad visualization image of wafer ID. Wafer has the characteristics that the surface is shinny and the patterns on the background are complex and noisy.

As image quality is sensitive to the variation of inspection conditions and lighting conditions, system performance is affected.

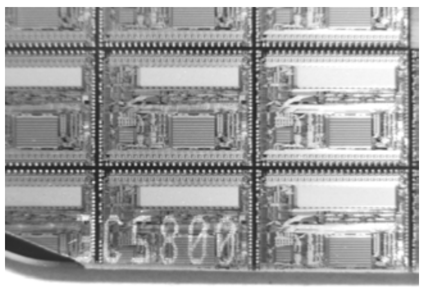


Fig. 2 Example of low-quality wafer image

As you may guess in this example, in the case of the recognition of the character and mark similar to this one in a factory automation environment, segmentation using conventional methods such as region segmentation, contour tracking and other binarization, must be impossible.

4.1 Snake algorithm

Contour based technique has been proven to be an effective approach in object recognition. Active Contour Models (also called snakes), which optimize/minimize an energy function, has become quite popular for boundary detection in the past few years. The energy function is generally obtained from internal sources with different weighting coefficients [8].

After the original snake model was proposed, Greedy algorithm that searched local neighborhoods instead of performing global optimization was developed to reduce the computational complexity of the snake models. We present an adaptive algorithm based on Greedy algorithm to solve optimal energy in multiple template models.

A snake is an elastic curve $(v(s))$ that evolves from its initial shape and position as a result of combined action of external and internal forces. The internal forces model the elasticity of the curve, whereas external forces push the snake towards features of the image. The conventional energy function in the snake is defined as follows:

$$E_{snake} = \int [E_{\text{int}}(v(s)) + E_{ext}(v(s))]ds \tag{4}$$

where E_{int} and E_{extt} are the internal spline energy and the external image energy respectively.

Discrete form is defined as eq. (5).

$$E_{snake} = \sum_{i=1}^{N} (E_{\text{int}}(i) + E_{ext}(i)) \tag{5}$$

where $E_{int}(i)$ and $E_{ext}(i)$ are the internal spline energy and the external image energy at the ith position of the contour (called snaxel s_i), respectively.

The image forces can be due to various events, e.g., lines, edges and terminations. The internal energy is written as

$$\begin{aligned} E_{\text{int}}(i) &= \alpha_i |s_i - s_{i-1}|^2 + \beta_i |s_{i-1} - 2s_i + s_{i+1}|^2 \\ E_{ext}(i) &= -|\nabla f(s_i)| \end{aligned} \tag{6}$$

where $f(s_i)$ is the original pixel value at s_i. The weighting factors in (6), α_i and β_i are defined to control the relative importance of the distance gaps and NGC values for ith snaxel, respectively. The external energy is set to be the negative magnitude of the image gradient as given in (6), so that the snake is attracted to the regions with low external energy [8].

The optimum locations of the snaxels are searched so that the snake energy E_{snake} can be minimized. There are many ways to obtain new snaxel positions with lower snake energy, e.g., greedy algorithm, gradient-descent methods [9] and dynamic programming methods [10,12,13].

We formulate the snake model of active multiple template as followings.

$$E_{snake} = \sum_{i=1}^{N} (E^i_{\text{int}} + E^i_{\text{ext}}) \tag{7}$$

$$E^i_{\text{int}} = \frac{1}{2}\alpha \cdot (d(v_i, v_{i-1}) - \bar{d})^2 \tag{8}$$

$$E^i_{ext} = \frac{1}{2}\beta \cdot Corr(v_{i-1}) \tag{9}$$

$d(v_i, v_{i-1})$ is the distance generated from the matching position of two templates neighboring in the ith and $i+1$th step, the difference between average distance $\bar{d}$ generate internal energy. The distance between two neighboring templates approaches to the average distance and the internal energy approaches to zero. The longer the distance between two neighboring templates, the more the internal energy will be increased as if spring would be pressed or pulled excessively. The external energy $E^i{}_{ext}$ of snake model in eq. (8) will be NGC matching value. Eq. (8) means that the coefficient at the position of which NGC value is maximum. It is called to external energy because image data plays a role to pull the template to the position of corresponding node. In this paper, we search the node that the value E_{snake} of chain node of the combination of successive multiple template is minimum using Greedy algorithm.

4.2 Multiple templates

Suppose that it is impossible to extract the character region from the background in the image which binarization is difficult. Even if binarization has been applied, to extract a character area from background is difficult. We suggest the searching method using template matching. Instead of binarization, we make templates composed with each character included in the image. Because all characters in input images are included in these template categories, at least each patterns are to character in this categories. Searching the (6)
character area with this template is to being recognized the character. That is, it is possible to search character area and recognize the character using template matching.

However, in the case of low-quality images, this matching is also impossible and not reliable. Matching failure results to recognition failure.

In order to solve this problem, we suggest multiple templates, which is composed of N multiple single templates as Fig. 3. Multiple templates increase the possibility of matching because the correlation value will be increasing N times than that of a single template.

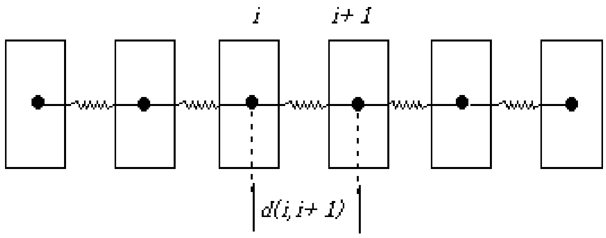


Fig. 3 The structure of multiple templates

Fig. 4 shows the multiple template models. In designing multiple templates, there are some restrictions as follows.

(1) The height(vertical length) of each template must be same respectively.

(2) No rotation in the character in input image.

(3) No variation of scale in the character in input image.

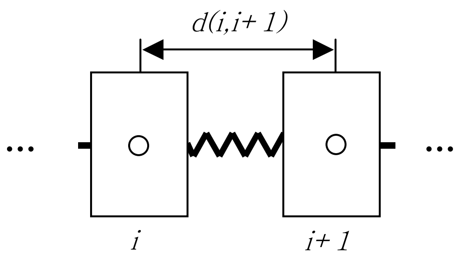


Fig. 4 Multiple templates model

In order to satisfy the above conditions, we normalized the size of template and process matching with the normalized pattern using MBR(minimum boundary rectangle) of each template model. Serial chain of multiple templates is configured for NGC matching using prepared M model templates. The number of serial chains of multiple templates is the same as the number of the characters in input image. Fig. 4 shows only two chains of whole serial chain. The M resulting values of NGC matching at *i*th position are generated at near *i*th position. The M resulting values of NGC matching at *i+1*th position are generated at near *i+1*th position. The number of configuration of distance generated by two neighbor templates is M^2. If the numbers of serial chain are N, the configuration numbers are M^{N-1}.

4.3 Analysis by Greedy algorithm

One method to solve the optimization problem is Greedy algorithm that selects the pattern of maximum matching value in present each time, without something to do with earlier or next decision [11].

Greedy algorithm doesn't guarantee global optimum but it is much faster than the dynamic programming, one of the global solutions [12,13].

Each node of inter-connected chain of multiple templates will be a candidate of recognition character. The positions of each node are different on the same horizontal line, the connection of *i*th and *i+1*th node will be a distance between two neighbor templates. The automatic segmentation of patterns in input image using multiple templates is accomplished as following step.

① Acquire the optimal matching point of the first node of multiple templates varying the position to the horizontal direction only.
② Repeat the matching in the position shifted by the width of a template.
③ After searching to the horizontal direction, the sum of matching coefficient and matching position are saved, a set of multiple template is acquired.
④ Repeat the steps ①~③ and acquire the multiple templates set so that the sum of matching coefficient is maximum.

Because acquired multiple templates set have each matching point which matching coefficients are maximum of matching value of each template in input image, it is possible to segment characters from background even in the low-quality image.

Fig. 5 illustrates the process of algorithm deciding individual template position in the active multiple template model.

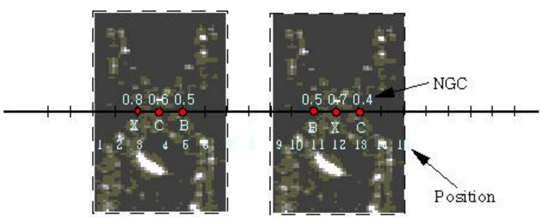


Fig. 5 Decision of individual template position

For example, the number of template is 3, "B", "C", "X", and string is composed of only 2 characters(N=2), the number of all path is 9. In the first area, *NGC*s of "B", "C", "X" are 0.5, 0.6 and 0.8, and the positions are 5, 4 and 3 respectively. In the second area, *NGC*s of "B", "C", "X" are 0.5, 0.4 and 0.7, and the positions are 11, 13 and 12 respectively.

Snake model is as eq. (10).

$$E^i_{snake} = \frac{1}{2}\left\{a \cdot (d(i,i+1) - \bar{d})^2 + b \cdot Corr(v(i-1))\right\} \quad (10)$$

where *Corr(v(i-1))* is a correlation value in *i*th x position.

The result of each case is shown in table 1. In this example, *a*=0.10, *b*=-0.51, $\bar{d} = 8$.

Comparing the result value in table 1, minimum energy of the snake model $E_{snake}(\min)$ is acquired at matching the template "X-X".

$$E_{snake}(\min) = E^1{}_7 : (X - X)$$

Table 1 Calculation example of snake energy.

Template		Internal energy E^i_{int}	External energy E^i_{ext}	Snake energy E^i_{snake}	
Area 1	Area 2				
B	X	0.05	-0.383	E^1_1	-0.256
B	B	0.2	-0.255	E^1_2	-0.055
B	C	0	-0.230	E^1_3	-0.230
C	X	0	-0.332	E^1_4	-0.332
C	B	0.05	-0.281	E^1_5	-0.231
C	C	0.05	-0.255	E^1_6	-0.205
X	X	0.05	-0.383	E^1_7	-0.333
X	B	0	-0.332	E^1_8	-0.332
X	C	0.2	-0.306	E^1_9	-0.106

Matching point to "X" in the first area and matching point to "X" in the second area are decided to be optimal positions. Thus, position 3 and 12 is resulting center of the X-axis of character area in fig. 6.

5. EXPERIMENTAL RESULTS AND CONSIDERATIONS

In order to verify usefulness of the proposed methodology, various experiments are conducted for a few kinds of low-quality wafer ID images.

Fig. 6 is an example of wafer ID image of semiconductor which contains noisy complex patterns on the background and any discriminative characteristics relative to the position of ID number is not shown in the histogram profile.

This says that conventional segmentation method such as histogram projection or Hough transform cannot be used.

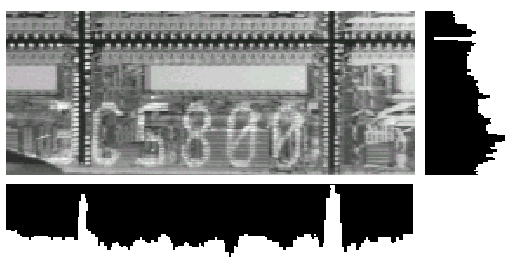
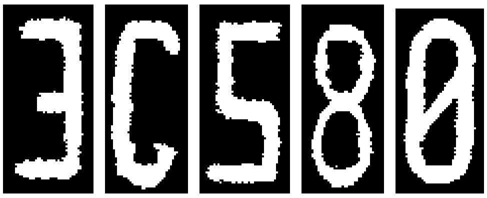


Fig. 6 Histogram profile of wafer ID image

Fig. 7 are examples of template models made from real wafer identification numbers. Normalized template models are used in matching the frontal laser marking.



(a) Laser mark template

(b)Normalized (80*32) laser mark template

Fig. 7 Template model of binary pattern

Fig. 8 shows the result of segmentation using single template matching. We could know that characters could not be segmented due to the noise and breakage in the area of characters. Fig. 8 says that that the segmentation using NGC searching and matching with single template is difficult.

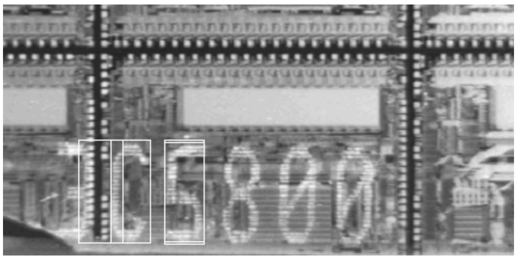


Fig. 8 Result of segmentation using single template

Fig. 9 shows the segmentation results in the case of applying active multiple templates for the low-quality image of wafer ID recognition.

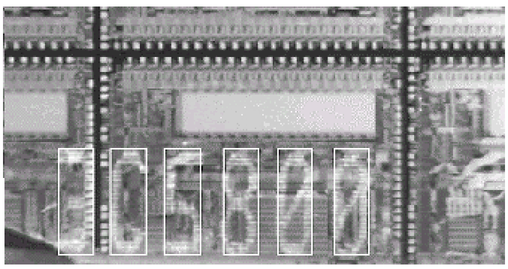


Fig. 9 Segmentation result by the proposed method

We can see accurate segmentation results even in low-quality image. It is impossible to segment consistently each character by conventional methods because of bad visualization and many noises.

In spite that the quality of input image is very low, we can segment each character effectively using Greedy algorithm with an active multiple template.

6. CONCLUSION

This paper proposed a method to segment wafer ID marks on poor quality images under uncontrolled lighting conditions of a semiconductor process. The key idea is based on the assumption as follows. If the characteristics of certain patterns appear consistently and repeatedly in an image, the possibility of being character region is high, even though the low-quality ID marks do not have complete characteristics of a pattern.

Because the state of visualization is very poor, none of the feature values may be detected with reliability. Thus the conventional histogram projection or binarization methods are difficult to be applied. Therefore, we tried to template matching to extract the character area and segment individual characters. But the method is also difficult to apply for the reason of frequent occurrence of fault matching in very poor visualization image. So we designed multiple templates model. It has the structure that several templates are connected by spring, in order to adapt the variation of gaps between characters. This model will be matched to snake model.

We applied the Greedy algorithm to solve optimization problem of active multiple template model and checked the result of segmentation in 8” wafer image. As a result, suggested methodology will be able to broaden the area of pattern recognition of low-quality images.

REFERENCES

[1] R. Haralick and L. Shapiro, "Survey: Image segmentation technique," *CVGIP*, vol. 29, pp 100-132, Jan. 1985.

[2] S. Horowits and T. Pavlidis, "Picture segmentation by a tree traversal algorithm," *J. CM*, vol 23, pp. 368-388, Apr. 1976.

[3] S. M. Lee, "Low rate video coding using 3-D segmentation with two change detection masks," ISO/IEC/JTC1/SC29/WG11 MPEG93/941, Nov, 1993.

[4] C. K. Chow and T. Kaneko, *Boundary detection of radiographic images by a thresholding method*, *in Frontiers of Pattern Recognition* (S. Watanabe, Ed.), pp. 61-82, Academic Press, New York, 1972.

[5] Fumihiko Saitoh, “Identification Number Recognition System for a Semiconductor Wafer using the Generalized Hough Transform,” *Journal of Japan Precision Engineering*, Vol. 61, No. 10, 1995.

[6] A. Rosenfeld and A.C. Kak, *Digital picture processing*, Academic Press, New York, 1976.

[7] S. Manickam, S.D. Roth, T. Bushman, "Intelligent and Optimal Normalized Correlation for High-Speed Pattern Matching, Datacube Technical Paper," Datacube Incorporation, 2000.

[8] Hyun Wook Park, Todd Schoepflin, and Yongmin Kim, “Active Contour Model with Gradient Directional Information: Directional Snake,” *IEEE Trans. On Circuit and Systems for Video Technology*, Vol. 11, No. 2, 2001.

[9] M. Kass, W. Within and D. Tetzopoulos, “Snake:active contour model,” *Int J. Computer Vision*, Vol. 1 (4), pp. 321-333, 1987.

[10] Amini, A., Weymouth, T.E., Jain, R.C., “ Using dynamic programming for solving variational problems in vision,” *IEEE Trans. Pattern Analysis and Machine Intelligence,* 12(9), 1991.

[11] R.E. Neapolitan and K. *Naimipour, Foundations of algorithms using C++ pseudocode*, 2ndedition, Jones and Barlett Publishing, 1988.

[12] D.J. Kang, "A fast and stable algorithm for medical images," *Pattern Recognition Letters* 20, pp.507-512, 1999.

[13] D.J. Kang, J.E. Ha, and I.S. Kweon, "Fast Object Recognition using Dynamic Programming from Combination of Salient Line Groups," *Pattern Recognition*, Vol. 36(1), pp.79-90, 2003.