Using Wavelet Transform and Neural Network Approach to Develop a Wafer Bin Map Pattern Recognition Model

Shu Fan Liu*, Fei Long Chen, and An Sheng Chung

Abstract—**Recently, semiconductor manufacturing has emerged to one of the most important industries in Taiwan. To decrease cost, semiconductor manufacturing companies always aim for yield enhancement. The analysis of Wafer Bin Maps (WBMs) is important for yield improvement. WBMs are formed by process, and thus engineers can find clues from their patterns and then correlate them with specific process to discover root causes. Nowadays, the recognition of WBMs is performed manually. This paper presented a method helps to automatically recognize known WBMs patterns and to further enhance yield. With Spatial Signature Analysis, systematic patterns of WBMs are extracted and noises are eliminated. Moreover, Wavelet transform coefficients are used here to construct Neural Network model. Real data sets are collected from a famous semiconductor manufacturing company to verify the presented method. Four types of WBMs patterns, center, edge, local, and ring types are selected for verification. Experimental results showed that with adequate parameter settings, the method can successfully recognize the pattern types and distinguish between random and systematic WBMs. There were 17 testing samples, and 16 of them were recognized correctly. The accuracy was 94.12%.**

Index Terms— **Semiconductor, Neural Network, Wavelet Transform, Wafer Bin Maps**

I. Introduction

The semiconductor industry has played an important role in Taiwan. As how the data shows it has been one of the key branches of the economy. Semiconductor manufacturing is a complex and costly process involving hundreds of separate steps and lasting up to 12 weeks. For this reason, cost deduction has become one of the most major goals for semiconductor companies.

Yield improvement is the critical issue that has always been discussed and also the important element to ensure the profitability of semiconductor industry. In semiconductor manufacturing, manufacturing parameters are recorded automatically to provide information to engineers. But it is hard to find out the root cause of the problem from the enormous database. Also, the engineers cannot response quickly to when problems occur. Consequently, to build up the knowledge database by analyzing and transforming data into useful information has become the significant issue of yield improvement and management.

Generally, all dies one a wafer must go through the circuit probe (CP) test after all manufacturing processes. The test results determine the grade of each die on a wafer. The company will use different codes to represent each grade. The code distribution on the wafer is the so-called Wafer Bin Map (WBM).

WBMs provide clues to identify possible causes when low-yield situation happens. We can trace back to specific root cause depending on the different pattern types of WBMs. Nowadays, the recognition of WBMs is performed manually. The individual fatigue and emotion affect results of pattern recognition. Therefore, this research intends to develop an automatic WBMs pattern recognition system. Chen[1] presented to use neural networks ART1 to recognize the WBMs.

WBMs patterns can be classified into random and systematic types [3]. This research presented a high-efficiency solution to WBMs patterns. Real samples containing systematic WBMs patterns, center, edge, local, and ring patterns, are collected for testing. By applying the presented model for the recognition of systematic WBMs patterns, it is expected systematic patterns can be recognized correctly. Objectives of this research are summarized below:

1) To develop the mechanism of classifying the WBMs patterns.

2) To accurately recognize the four selected systematic pattern types: 1) center 2) edge 3) local 4) ring (as shown

Dr. S. F. Liu is with Department of Information Management, Yuanpei University of Technology, Hsichu County, 300, Taiwan(R.O.C) (phone: 886-3-538-1183; fax: 886-3-538-5353; e-mail: sfliu@mail.ypu.edu.tw).
Dr. F. L. Chen is with Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsichu County, 300, Taiwan(R.O.C) (phone: 886-3-571-7654; fax: 886-3-572-2685; e-mail: flchen@ie.nthu.edu.tw).
An Sheng Chung is with Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsichu County, 300, Taiwan(R.O.C) (phone: 886-3-571-7654; fax: 886-3-572-2685; e-mail: ouch26@hotmail.com).

in fig. 1)

3) To improve the automated level of the WBMs patterns recognition.

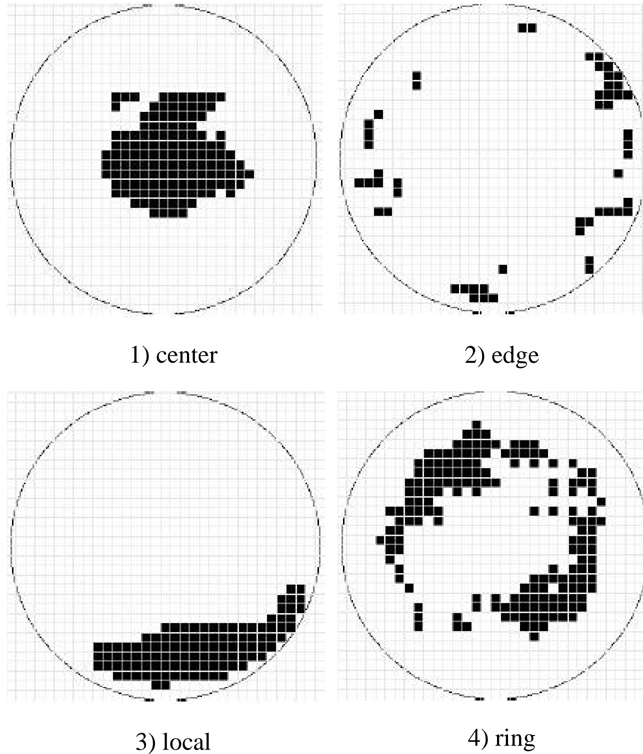

Fig 1. The systematic WBMs pattern types

Fig. 2 depicts the structure of the proposed recognition model starts from receiving WBMs raw data including the coordinate of bad-die point, the die size, the reticle size and the other related information. Details of the overall recognition procedure will be explained in Section II. Section III includes experimental results and discussions. Conclusions for the experiments results are given in Section IV.

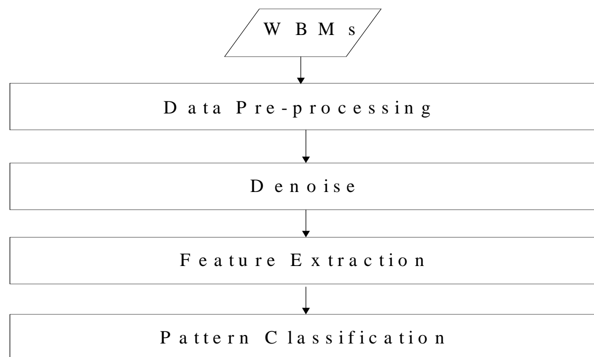

Fig. 2. Process of the proposed methodology

II. WBMs Pattern Recognition

This section describes the overall architecture of the proposed methodology.

A. Denoise

Our estimation procedure begins by reducing noises to obtain smooth wafer maps. To reduce noises of WBMs, Friendman [2] proposed a local kernel estimator be index. Firstly, Let W denotes a collection of chip sites that constitute a wafer map. The goal of our smoothing operation is to highlight regions with a large number of failed chips. Toward this end, for each $i \in W$, we let N_i denote the collection of neighboring chip locations surround chip i and let N_i denote their numbers. For each chip in $j \in Ni$, a weight $w_i(j)$ is assigned and the sum of the $w_i(\bullet)$ over the neighbors of chip i is equal to one.

Next, for each $i \in W$, we define a local kernel estimator of the proportion of failed chips on the wafer as equation (1).

$$\hat{p}_i = \sum_{j \in N_i} w_i(j) I_F(j)$$

(1)

Where for each $i \in W$, $I_F(i)$ is the indicator function of the event that the chip at position i is failed.

Keeping with the simple case of $w_i(j)=1/N_i$ for the moment, we observe that $\hat{p}_i$ has different variances under the hypothesis of complete spatial randomness for different overall yields, and which further complicates interpretation of the variation in these smoothed maps. To remove this dependence on yield we define a new map such that over each position i we may plot $\hat{s}_i$, where $\hat{s}_i$ is as equation (2).

$$\hat{s}_i = \sin^{-1}\sqrt{\hat{p}_i} \qquad (2)$$

It is well known that $\hat{s}_i$ has an approximately normal distribution with mean y and variance $1/(4N_i)$. More generally, if we allow the $w_i(\bullet)$ to define an arbitrary kernel over N_i, then the $\hat{s}_i$ will have an approximately normal distribution with mean y and variance

$$\sum_j [w_i(j)]^2 / 4. \qquad (3)$$

As a result, a mask of size n by n is required to calculate the estimator. Through try-and-error experiments, a 3×3 mask do meet our requirement.

After the local kernel estimator is obtained, a threshold value is setup for the denoising process.. Steps to setup this value are explained below:

Step1. We set the initial threshold value of 1 because the local kernel estimator is between 0 and 1.

Step2. Divide the WBMs into two parts as fig 3. The first part (top right) passes the threshold value while the second one doesn't (bottom right)..

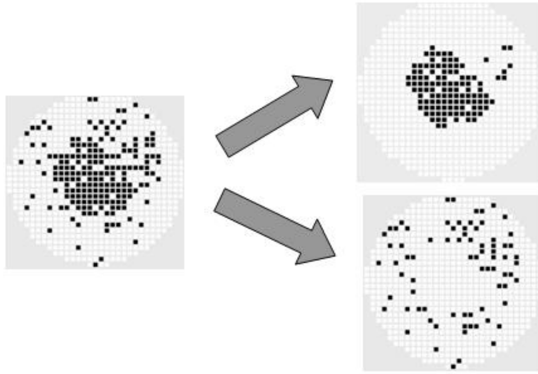

Fig 3. The process of denoising

Step3. If the second part in Step2 doesn't pass Spatial Randomness Test (this part is not Spatial Randomness), the threshold value is deducted by 0.01 and return to Step2. Log OR (Log odd ratio) is the testing estimator here [5].

Step4. The second part of Step2 shows the WBM after denoising.

B. Feature Extraction

In the context of pattern recognition, dimensionality reduction and determination of WBM features are very important for high recognition accuracy. In this study, we utilize Wavelet Transform to extract the features of WBMs. Wavelet Transform has a characteristic of multi-resolution, and that hepls to reduce the dimensionality of WBMs.

In other words, each WBM will be represented by a set of wavelet coefficients. Before executing the Wavelet Transform, we have to choose a wavelet function. Generally, wavelet functions can be classified into 4 categories [6]: crude wavelets, infinitely regular wavelets, orthogonal wavelets, and biorthogonal wavelets. We select Harr wavelet to be the wavelet function because it's simple and fast.

Each transformation decomposed signals into four coefficients, including one approximation coefficient and three high frequency coefficients. The author uses the approximation coefficient to be the feature of the WBMs because it's the most important signals of the WBM. Besides, The approximation coefficient can furthermore be decomposed into other four coefficients. Through experimental result, the author decided to choose 2-level wavelet coefficients to represent features of WBMs.

C. Pattern Classification

Pattern classification is the grouping of individuals in a population in order to discover structural information in the massive data. This research adopted LVQ neural network architecture for WBMs pattern classifications. LVQ was proposed by Kohonen in 1988, and is well-known for its fast learning algorithm and good classification accuracy.

The structure of LVQ neural network contains a competitive layer and a linear output layer. The concept of LVQ is that each category is given a core, and we adjust each core when next sample enters the model. The principle of core adjustments is to find the closest point through the function $q(X) = \min_{\forall j} \|X - w_j\|_2 \quad j = 1,2,...,M$. If the closest point is the correct category, the core will be adjusted through the function $w_q(k+1) = w_q(k) + \mu(k)[X - w_q(k)]$; otherwise, the core will be adjusted through the function $w_q(k+1) = w_q(k) - \mu(k)[X - w_q(k)]$.

III. EXPERIMENTAL RESULT

The system is developed with MATLAB 7.0 and C# language. A semiconductor company provided real WBM data to verify the performance of the presented system.

The proposed methodology includes denoising, feature extraction, and pattern classification. We adopt Wavelet Transform and LVQ neural network to feature extraction and pattern classification, respectively. Before the experiments, several parameters should be assigned for Wavelet Transform and LVQ. We test several parameter combinations to generate the best model. Table I shows the test result. In this table, *i-j-k* indicates the combinations of parameters.

$$i = \begin{cases} 2, 2\text{-}level\ wavelet\ transform \\ 3, 3\text{-}level\ wavelet\ transform \end{cases}$$

$$j = \begin{cases} 2, hidden\ nodes\ is\ double\ of\ input\ nodes \\ 3, hidden\ nodes\ is\ triple\ of\ input\ nodes \end{cases}$$

$$\mathrm{k} = \begin{cases} 1, learning\ rate\ is\ 0.05 \\ 2, learning\ rate\ is\ 0.1 \end{cases}$$

TABLE I The combinations of parameters

Model (i,j,k)	accuracy	model	accuracy
3-2-1	76.47%(13/17)	2-2-1	82.35%(14/17)
3-3-1	64.71%(11/17)	**2-3-1**	**94.12%(16/17)**
3-2-2	70.59%(12/17)	2-2-2	82.35%(14/17)

3-3-2	82.35%(14/17)	2-3-2	88.24%(15/17)

There are 17 real WBM samples for the experiments to test the accuracy of the model. As Table I shows, the highest accuracy among the combination of parameters is 94.12%. It indicates that 16 testing samples had been classified into correct category through the model. The only one sample that was classified incorrectly is shown as Fig. 4. This sample is ring type, but the model classified ii into center type. This situation really happened if there were ambiguities between different WBM patters. If some more real WBMs samples are available for the pattern classification procedure, the wrong classification situations can be improved.

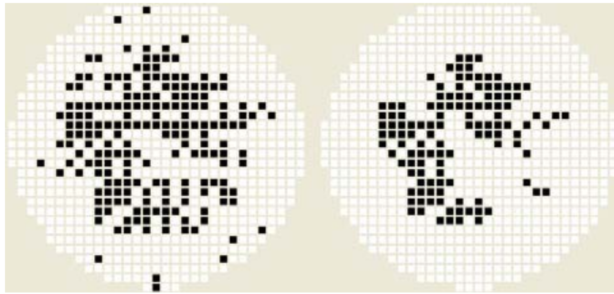

Fig. 4 the sample recognized incorrectly

We also compare with several combinations of feature selection and classification techniques to test the effectiveness of the presented model.

Invariant Moment is always used to extract features of image in previous researches. Lin [4] applied supervised neural network BP architecture to recognize systematic patterns. Different combinations of feature extraction and classification technique shown as Table II are designed for comparisons.

Table II Comparison of different methodologies

Model	training(sec)	epochs	accuracy
Wavelet-LVQ	**17.641**	**17**	**94.12%**
Wavelet-BP	54.425	100	88.24%
Invariant Moment-LVQ	73.008	100	52.94%
Invariant Moment-BP	75.068	100	76.47%

As experiments results shown in Table II, BP neural network can also get a satisfying accuracy but it need to adjust many parameters, as the amount of hidden layers, transfer function, and so on. Most of the parameters are continuous. Hence, it always wastes a lot of time to find the satisfying combination of these parameters. Besides, the output of BP neural network is continuous. We must set the stopping rule to terminate training and avoid overtraining at the same time.

Wavelet-BP doesn't get satisfactory accuracy level until enough training epochs. The models with Invariant Moment can't perform well. We believe that cooperate the other information of WBM will improve it.

IV. CONCLUSIONS

Keeping high yield and maintaining profitability are usually the way to lead in the fierce competition in semiconductor industrial. WBMs is one of the most important clues to response to the process variations quickly. Hence, constructing an automatic WBMs patterns recognition system is really important and necessary for yield enhancement. The presented pattern recognition methodology does help to recognize WBMs automatically and accurately. The results show that the accuracy of this methodology is up to 94.12% and it is also simpler than other methodologies for practical applications.

In the semiconductor manufacturing process, wafers of the same lot always pass through the same manufacturing process steps. Therefore, WBMs in one lot should be highly correlated. This is one of the interesting directions we expect to discover for the next step. Moreover, collecting much more WBMs pattern samples will help to generalize and improve efficiency of the presented approach and we are currently working on that.

REFERENCES

[1] Chen, F. L., and S. F. Liu, "A Neural-Network Approach To Recognize Defect Spatial Pattern in Semiconductor Fabrication," *IEEE Trans. on Semiconductor Manufacturing*, 13 (3), 366-373 (2000).

[2] Friedman, D. J., Hansen, M. H., Nair, V. N. And James, D. A.,"Model-Free Estimation of Defect Clustering in Integrated Circuit Fabrication," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 10, No.3, pp. 344-359, August, (1997).

[3] Kaempf, U., "The Binominal Test: A Simple Tools to Identify Process Problems," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 8, No. 2, pp. 160-166, (1995).

[4] Lin, G. T., "Patterned Recognition of Wafer Bin Maps", Master thesis, Taiwan University, (1998)

[5] Taam, W. and Hamada, M., "Defecting Spatial Effects from Factorial Experiments: An Application from Integrated-circuit Manufacturing," *Technometrics*, Vol. 35, No.2, (1993).

[6] Yeh Zong-Xing, "Hyperspectral Data Identification Using Wavelet Decomposition and RCE Neural Network," (2000).