Computers & Industrial Engineering 142 (2020) 106358

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Similarity matching of wafer bin maps for manufacturing intelligence to empower Industry 3.5 for semiconductor manufacturing

Chia-Yu Hsu[a,b,*], Wei-Ju Chen[b], Ju-Chien Chien[b,c]

[a] Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei 10608, Taiwan
[b] Artificial Intelligence for Intelligent Manufacturing Systems (AIMS) Research Center, Ministry of Science & Technology, Hsinchu 30013, Taiwan
[c] Department of Computer Science, National Tsing Hua University, Hsinchu 30013, Taiwan

ARTICLE INFO

Keywords:
Wafer bin map
Yield enhancement
Defect diagnosis
Similarity
Manufacturing intelligence
Industry 3.5

ABSTRACT

Yield improvement is increasingly important as advanced fabrication technologies are complicated and interrelated for semiconductor manufacturing. Wafer bin maps (WBM) present specific failure patterns which provide crucial information to track the process excursions to empower intelligent manufacturing for wafer fabrication. In practice, WBM identification is still subjective relied on domain knowledge and human-eye. As the semiconductor industry continuously migrates for advanced nano technologies, many rare defect patterns are also generated by different pattern, pattern size, noise degree, pattern density, pattern shift, and wafer rotation. Existing studies regarding WBM focus on classification and lack of capability to detect a rare pattern. In order to overcome the shortage of WBM classification, the similar WBMs provide useful information of WBM identification. Following Industry 3.5 as a hybrid strategy between Industry 3.0 and to-be Industry 4.0, this study aims to develop a novel approach to measure the similarity of defect patterns of WBMs to enhance decision quality for fault detection and defect diagnosis effectively and efficiently. In particular, the proposed approach applied a mountain clustering algorithm to enhance the defect features depending on clustering density. Then, Weighted Modified Hausdorff Distance (WMHD) is employed to measure the similarity level. Furthermore, a decision support system embedded the developed algorithms is constructed. An empirical study of WBM clustering was conducted in a fab for validation. The results have shown practical viability of the proposed approach.

1. Introduction

The manufacturing process of in semiconductors fabrication involves hundreds of steps, and amounts of big data including the wafer lot history, process recipe, inline metrology measurement, equipment sensor value, defect inspection, and electrical test data are automatically generated and recoded. The challenge in semiconductor companies not only integrate big data from various sources into a platform or data warehouse, but also lack intelligent analytic solutions to extract useful manufacturing intelligence and support decision making regarding production planning, process control, equipment monitoring, and yield enhancement (Chien, Dou, & Fu, 2018; Dou, Chien, Kacem, & Hsu, 2018; Dou, He, & Hsu, 2018).

Wafer bin map (WBM) is outcome of circuit probe (CP) process, which records the spatial distribution of defect dies on the wafer. WBM spatial patterns contain potentially useful information, in which specific pattern refers to potential failure of a specific manufacturing process. Thus, analyzing spatial patterns of WBMs can provide useful manufacturing intelligence for engineers to track the problem in specific manufacturing processes for fault detection and yield enhancement.

In practice, most of semiconductor manufacturing companies still rely on domain knowledge and experience in which online engineers employ visual inspection and personal judgments to identify and classify the patterns for the wafer bin maps in batch. However, existing judgement by human-eye may be subjective, inconsistence, and lack of justification, while is also time consuming.

Indeed, a number of studies have been done for WBM analysis and pattern recognition that can divided into two major categories: supervised approaches for defect classification and unsupervised approaches for WBM clustering (Chien, Hsu, & Chen, 2013). However, supervised approaches have the advantage for quickly detecting the existing patterns, yet with the limitation for extending the classification rules in light of new defects from technology migration. On the other hand, unsupervised approaches have the advantage for analyzing all of the patterns, yet with the limitation for recognizing specific patterns

* Corresponding author at: Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei 10608, Taiwan.
E-mail address: chiayuhsu@mail.ntut.edu.tw (C.-Y. Hsu).

https://doi.org/10.1016/j.cie.2020.106358

Available online 19 February 2020
0360-8352/

such as the ring patterns (Hsu & Chien, 2007). With the development of big data analytic and machine learning methods, adaptive resonance theory (ART) neural network (Liu, Chen, & Lu, 2002; Hsu & Chien, 2007; Liu & Chien, 2013), decision tree-based (Piao, Jin, Lee, & Byun, 2018), support vector machine (SVM) (Baly & Hajj, 2012; Liao, Hsieh, Huang, & Chien, 2014; Wu, Jang, & Chen, 2015), and joint local and nonlocal linear discriminant analysis (Yu & Lu, 2016) were applied for WBM classification. Saqlain, Jargalsaikhan, and Lee (2019) extracted 66 features including density-based, geometry-based, and radon-based features from raw wafer images and four kinds of classifiers including logistic regression (LR), random forests (RFs), gradient boosting machine (GBM), and artificial neural network (ANN) were used for model training. The ensemble result is better than individual classifier. However, relying on generated features in advance are not enough to cover all kind of WBM failure patterns. Convolutional neural network (CNN) was also applied for WBM classification (Nakazawa & Kulkarni, 2018; Yu, Zheng, & Liu, 2019) and segmentation (Nakazawa & Kulkarni, 2019).

Various failure patterns are useful to facilitate rapidly identifying the associate root causes of low yield. As the semiconductor industry continuously migrates for advanced nano technologies, the involved complexity of product design and manufacturing process is exponentially increased. Many rare defect patterns are generated by different pattern shape, pattern size, noise degree, pattern density, pattern shift, and wafer rotation. However, existing studies of WBMs focus on WBM classification and lack of capability to detect a rare pattern. In order to help engineers to recognize the rare defect patterns for root case identification as possible, the information about measuring the similarity between two WBMs are needed.

As most of the industries may not be ready for migration, Industry 3.5 was proposed as hybrid strategy between the existing Industry 3.0 and to-be Industry 4.0 to empower industrial transformation (Chien, Chou, & Yu, 2016; Chien, Hong, & Guo, 2017). While the big data is accumulated automatically during semiconductor manufacturing process, how to exploit the useful rule and extract the intelligence from amounts of equipment sensor data, inspection data, and production data to assist operational decision-making has been changed the paradigm of semiconductor manufacturing. In particular, this study addressed the WBM analysis problem and hybrid various methods to support the matching of defect patterns for root case identification. That is, similar defect patterns may provide the clues for the engineer to track potential root causes and narrow scope for trouble shooting. But, the similarity between two WBMs are difficult to determine by a mathematical equation because the "similarity" perceived by individual engineers will be different subject to visual sensibility, physical or mental condition. Moreover, the main factors including pattern size, pattern noise, different pattern density, location shift and wafer rotation will lead to generate various kind of WBMs and increase the difficulty of WBMs similarity matching

To fill the gaps and overcome the factors which effect the precision of analysis results, this study aims to develop a framework for WBM similarity matching and thus construct integrated WBM similarity measurement system to support engineers for smart production in real settings. In particular, the proposed framework consists of four phases to prepare the data, formulate the similarity measurement models, and effectively derive similarity indices to support the decision with a decision support system. Using data preparation method to enhance data quality and provide consistent analysis baseline. Next, the spatial distribution of defective dice is transformed by mountain function. The similarity of two WBMs is measured by weighted modified Hausdorff distance (WMHD). To validate the effectiveness of the proposed framework of WBM similarity matching, an empirical study from a semiconductor company was conducted.

To compare with the existing literature, the major contributions of this paper are as follows. First, the proposed method of WBM similarity matching can be applied to search the similar defect patterns among the existing detect pattern without any pre-defined features and classes. The proposed method can be used for WBM identification with considering the various factors including pattern size, noise degree, pattern density, pattern shift, and wafer rotation. It overcomes the shortage of existing classification-based methods in WBM analysis. Second, the proposed method can be used for not only the rare defect pattern but also the pre-defined defect patterns without collection of large amounts of training data and training a classification model.

The remainder of this paper is organized as follows. Section 2 reviews the related studies on manufacturing intelligence and wafer bin map analysis. Section 3 presents the proposed framework for WBM similarity matching. Section 4 presents an empirical study in a worldwide leading semiconductor manufacturing company for validation. Section 5 concludes with discussions of contribution and future research directions.

2. Fundamental

After wafer fabrication, the wafers must go through serial pass-or-fail functional tests, so-called the circuit probe (CP) yield test, to test the electrical function of dies whether is well or fail (Hsu & Chien, 2007). The CP test is a significant index of yield, the CP yield improvement can divide into two parts, one is the base line yield improvement, improving the tool performance and reducing failure by tuning the process recipe; another is the low yield trouble shooting, monitoring and diagnosing the factor lead to failure (Chien, Hsu, & Chen, 2013; Liao et al., 2014).

The CP test determines the grade of each die on a wafer, the better grade would be assigned if the die is well-functional, and vice versa. The grade is presented by specific bin code, differs on different product types and company (Liu et al., 2002). There are hundreds of testing items in CP test. Fig. 1 illustrate a WBM with mixed defect bin code. For example, test die is well-functional, the test bin code is denoted as 1. If a die is failed to certain testing item, it will be denoted a corresponding bin code. The example in Fig. 1 use green or "1" to denote well-functional dies, and other color or "0" to denote defective dies. The quality test of CP test can be illustrated in spatial distribution, and result in wafer bin map (WBM). With different quality level, WBM is multi-dimensional and have complex structures. To assist visualization and analysis, WBM is usually transformed into a binary map that represents it using binary code.

Typical WBM failure patterns consist of three categories: (I) Random defect, (II) Systematic defect and (III) Mixed defect (Hsu & Chien, 2007; Chien, Hsu, & Chen, 2013; Liu & Chien, 2013).

(1) Random defect: the defective chips are randomly distributed in the map. Random defects are usually caused by the manufacturing environment. Even in a near-sterile environment, the particles cannot be removed completely.
(2) Systematic defect: the positions of defective chips in the wafer show the spatial correlation, for example, ring, edge-fail, checkerboard, "Bull-eyes" and scratch shapes. These frequently-seen patterns usually represent different sources of process failure causes. This spatial information is an important clue for engineers to trace the process problem.
(3) Mixed defect: a combination of systematic defects and random defect in one map.

Mixed defect patterns that are common in the fab. It is important for WBM analysis to separate random noise to identify root causes of the systematic defects for yield enhancement. The random defect might impact the identification of systematic defect, and cause difficulty to extract the feature of the main defect. However, the mixed defect is the most common type of defect in the fab. To analyze such type of defect, the engineer required to eliminate the defects caused by random noises, since only the systematic defect can lead the engineer to discover the major process problem (Hsu & Chien, 2007).

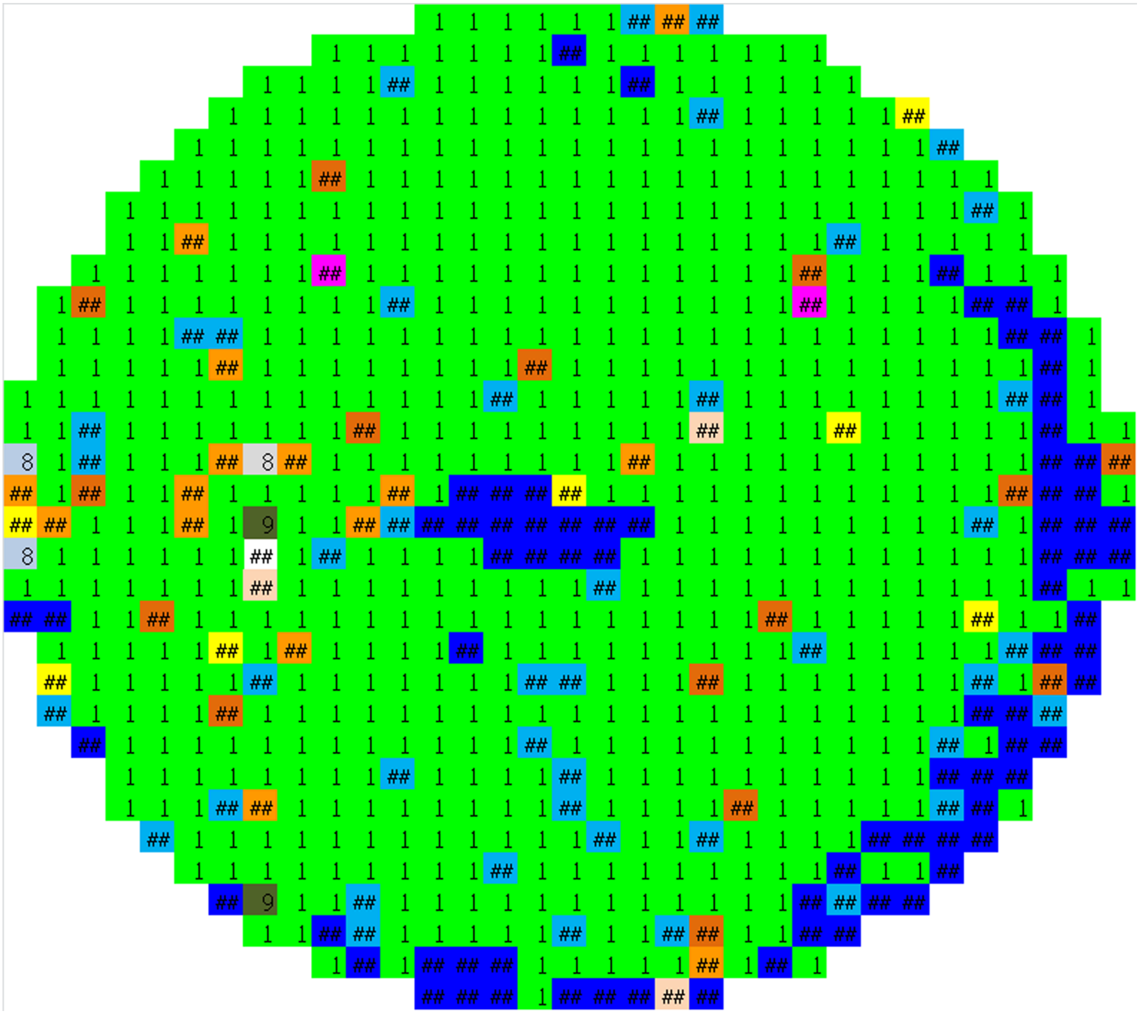


Fig. 1. Illustration of wafer bin map with mixed defects.

Most of existing WBMs studies are focus on diagnosing systematic defects, pattern recognition and classification (Chao & Tong, 2009; Hsu & Chien, 2007; Li & Huang, 2009; Liu et al., 2002; Taam & Hamada, 1993; Wang, 2008; Wang, Kuo, & Bensmail, 2006; Yuan, Bae, & Park, 2010; Yuan, Kuo, & Bae, 2011). Few researchers proposed statistical model to detect whether there is specific defect patterns (Yuan et al., 2010; Yuan et al., 2011). Statistic model can precisely recognize symmetric or complex patterns, however need to well and specifically define the pattern feature. Also less flexibility: for those rare and which doesn't have significant feature, developing a statistic model may consider an inefficient way.

Machine learning techniques including ART neural network (Hsu & Chien, 2007; Liu et al., 2002), decision tree (Hsu & Chien, 2007), single-class SVM (Liao et al., 2014), multi-class SVM (Chao & Tong, 2009; Baly & Hajj, 2012), self-organizing map (SOM) (Li & Huang, 2009), and CNN (Nakazawa & Kulkarni, 2018) was also applied for WBM classification. However, machine learning techniques may need large amounts of training data to learn the complex function and focus on common pattern classification such as center, donut, edge-local, edge-ring, local, and scratch (Wu et al., 2015).

Big data analytics and data mining approaches have been employed to explore huge data to extract useful rules to support smart production for semiconductor manufacturing (Chien & Chuang, 2014; Chien, Liu, & Chuang, 2017; Khakifirooz, Chien, & Chen, 2018). Manufacturing intelligence can be derived to enhance decision quality and operation efficiency of decision problems involved in smart production that are characterized by uncertainty and a need for tradeoff among various objectives and justification for the decisions in short time (Chien & Hsu, 2006; Chien & Hsu, 2011; Chien, Chen, & Peng, 2010; Kuo, Chien, & Chen, 2011; Chien, Hsu, & Hsiao, 2012; Chien, Hsu, & Chen, 2013; Chien & Hsu, 2014). In particular, based on domain knowledge, wafer fabrication with a lengthy process from a large number of correlated variables is structured. The information and intelligence are extracted by multivariate statistical approaches as well as big data analytics to provide an aid for fault detection and defect diagnosis to eliminate the causes and thus improve the performance of the process and enhance the yield (Chien, Hsu, & Chen, 2013).

Due to continuous migration of semiconductor industry, many rare defect patterns that never been seen before will cause the limitation of existing approaches to fulfill the needs in real settings. Rather than developing an intelligent solution to replace domain experts, this study focused on develop a decision support system with similarity matching to provide an effective solution to empower human judgments of the engineers, as proposed in Industry 3.5.

3. WBM similarity matching

The proposed framework for WBM similarity matching consists of the following four steps as illustrated in Fig. 2. In the first step, a candidate WBM for similarity matching is selected first and denoted as “selected WBM”. In the second step, the raw data of CP test are turned into binary map, in which the dice pass all the testing items are denoted as 1 and the dice fail any testing items are denoted as 0. Then, the selected WBM is applied enhance the signal and remove the noise (ESRN) method (Hsu & Chien, 2007) to strength the feature of defect

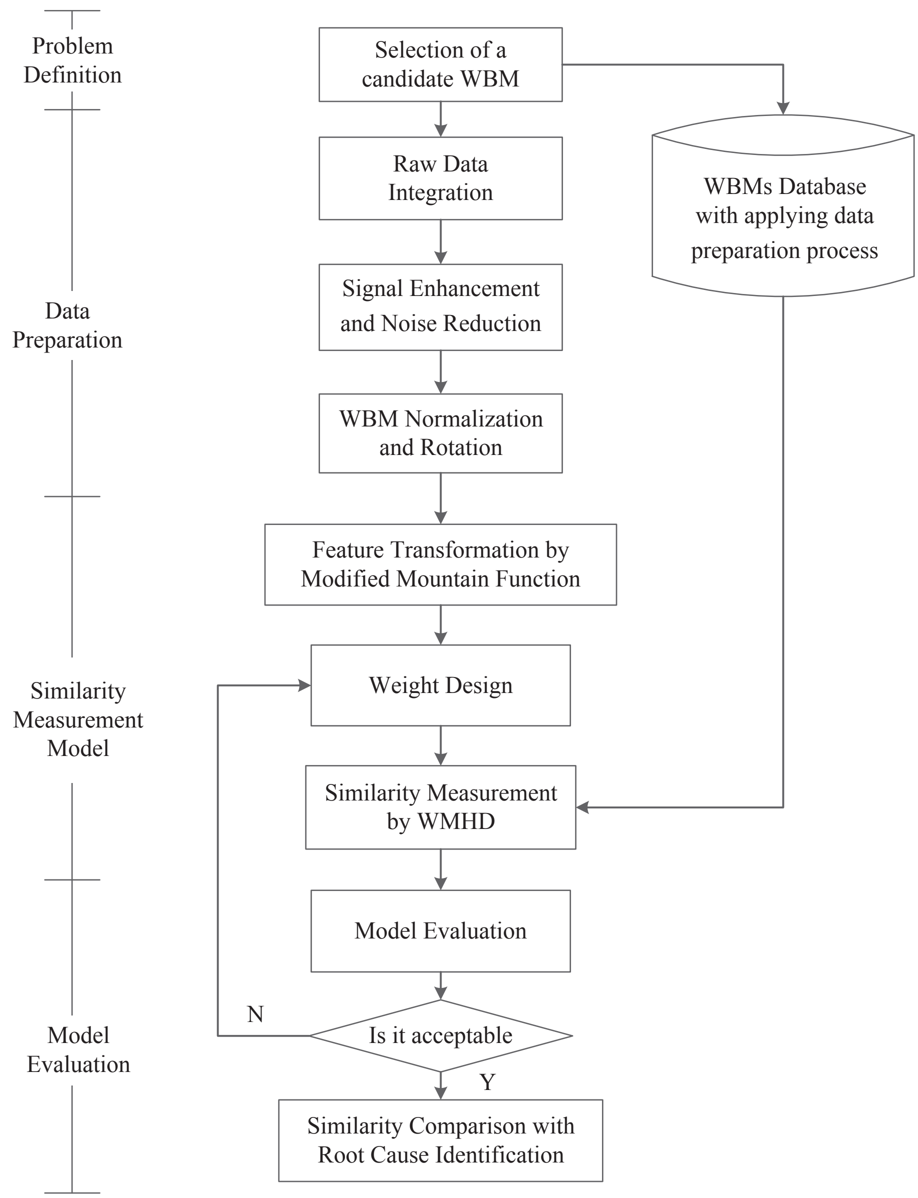


Fig. 2. Research framework.

pattern and remove the unnecessary random defect. In order to consider two similar WBMs with rotation, we also perform Cartesian-Polar coordinate transformation. The third step is to build the similarity measurement model by using modified mountain function to enhance the feature of WBMs. The value of modified mountain function is also used as a weight for similarity measurement. Then, the similarity between WBMs are determined by WMHD and the weight design by modified mountain function. Final step is to evaluate the performance and validate the effectiveness of the proposed method. The similarity results are used for failure analysis database with further application.

The terminology and notation used in this study are summarized as follows:

d_{ij}: the distance between die i and the defective die j.
d_{jc}: distance between defect die j and wafer centroid c
$m\beta$: kernel width in mountain function
m: parameter of determining kernel width
β: constant of kernel width, $\beta = \left(\frac{\sum_j d_{jc}}{N}\right)^{-1}$
A: the dataset of defective die on selected wafer
B: the dataset of defective die on compared wafer
$a(x, y)$: coordinate (x, y) for the selected wafer
$b(x, y)$: coordinate (x, y) for the compared wafer
ω_{ab}: the weight of defective die a belong to **A** and defective die b belong to **B**
N_d: the number of defective die on a wafer
N: the number of die per wafer
N_A: the number of defective die on the selected wafer
N_B: the number of defective die on the compared wafer
N_o: the number of defective die on a selected wafer in which the spatial distance between the selected wafer and the compared wafer are large than the maximum matching tolerance
$M(i)$: the mountain function
M_a: the mountain value of defective die a from the selected wafer
M_b: the mountain value of defective die b from the compared wafer
S_{out}: the maximum matching tolerance of two defective dice a and b
$h(\mathbf{A},\mathbf{B})$: the WMHD between the selected WBM and the compared WBM
$S(\mathbf{A},\mathbf{B})$: the similarity between the selected WBM and the compared WBM

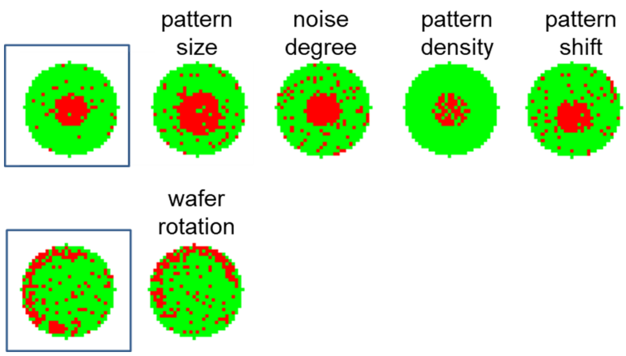


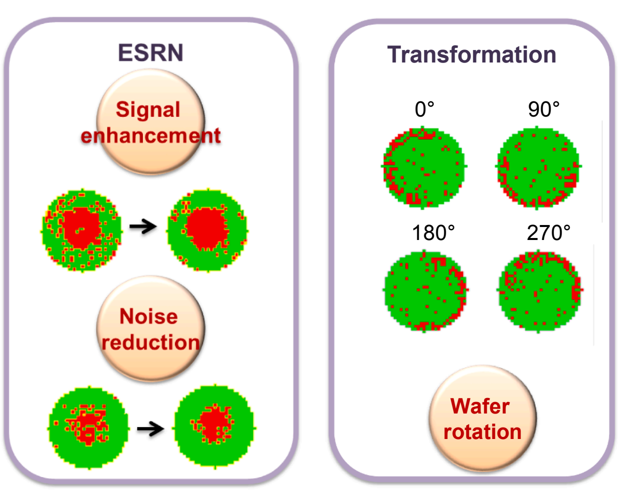
Fig. 3. Illustration of considered factors.

3.1. Problem definition

Contribute to the variation in manufacturing process and different setting of parameters, even same type of pattern result in different level of shape change. Fig. 3 illustrates the factors need to be considered in similarity matching such as pattern size, noise degree, pattern density, pattern shift, and wafer rotation. The pattern size denotes that the WBMs have same defect pattern but the defective size is different. The noise degree represents the same defect pattern with different random defective die on a WBM. The pattern density denotes that the WBMs with same defect pattern but the distribution of defective dies is different. The pattern shift means the same defect pattern but the locations of pattern on a wafer are different. The wafer rotation denotes that the WBMs have the same defect pattern with the different angles.

3.2. Data preparation

Fig. 4 illustrates the process of data preparation for WBMs. Data Cleaning: To eliminate the effect cause by random defect (noise), using Enhance the signal and remove the noise (ESRN) (Hsu & Chien, 2007) method to enhance the signal of cluster and remove the noise. Weighing adjacent dies in King-Move neighborhood for each die on the wafer. Calculating the proportion of the adjacent neighbors, if the proportion of well-functional die achieved the defined criteria, transfer the central die into 1 (good); on the other hand, if the proportion of bad die achieved the defined criteria, transfer the central die into 0 (failure). Moreover, we also perform Cartesian-Polar coordinate transformation to solve wafer rotation problem due to the machine alignment.



(a) ESRN (b) Transformation

Fig. 4. Illustration of WBM by data preparation.

3.3. Similarity measurement model

3.3.1. Feature transformation

The mountain method is a simple and effective heuristic algorithm to estimate the approximate cluster centers (Yager & Filev, 1994). The mountain function is similar to Parzen window, which estimates the probability density function of the feature. Yang and Wu (2005) modified the mountain clustering algorithm by defining on data points (defective die) instead the number of die on a wafer. Assume we have N_d defective dice, $j = 1, 2, \cdots, N_d$, and the number of die per wafer is N, $i = 1, 2, \cdots, N$, the mountain function $M(i)$ is defined as:

$$M(i) = \sum_{j=1}^{N_d} \exp(-m\beta d_{ij}), \quad i = 1, 2, \cdots, \mathrm{N} \tag{1}$$

where $\beta = \left(\frac{\sum_j d_{jc}}{N}\right)^{-1}$ denotes as the defective die j to the wafer center c. The variable d_{ij} is the distance between die i and the defective die j. Parameter β is the normalization factor for the distance between defective die j and the wafer centroid c. Parameter m is a constant. Parameter $m\beta$ determines the approximate density shape of the wafer. For example, Fig. 5 shows an example of a mountain value (z-axis) derived from wafer bin map. The x-axis and y-axis represent the coordinate of a wafer (x, y).

Through mountain method, the feature of clustering can be amplified. Also, the mountain value can provide profile of specific pattern as shown in Fig. 6. In particular, Fig. 6(a) and Fig. 6(b) represent that the different mountain value (z-axis) of different pattern. Fig. 6(a) and Fig. 6 (c) represent that the different mountain value of different pattern density. Fig. 6(a) and Fig. 6(d) show that the different mountain value of different pattern density. Through modified mountain function for each WBM, it can show the difference among the WBM with different factors pattern size, pattern density, and pattern shift.

3.3.2. Similarity calculation

In the shape matching of WBMs, Hausdorff distance could be simply applied to determine the quality of matching between two patterns. Through Hausdorff is distance is simple in calculating, it is sensitive to degradation such as noise and occlusions. The modified Hausdorff distance (MHD) have been proposed to decrease the impact from outliners, such as partial Hausdorff distance (Huttenlocher & Rucklidge, 1992) and weighted Hausdorff distance (WHD) (Lu, Tan, Huang, & Fan, 2001).

Assume the set of defective die on the selected wafer is **A**, and the set of defective die on the compared WBM is **B**. The distance between two points $a \in \mathbf{A}$ and $b \in \mathbf{B}$ is defined as $\|a - b\|$ in city block distance. Thus the weighted MHD (WMHD) can be determined as follows:

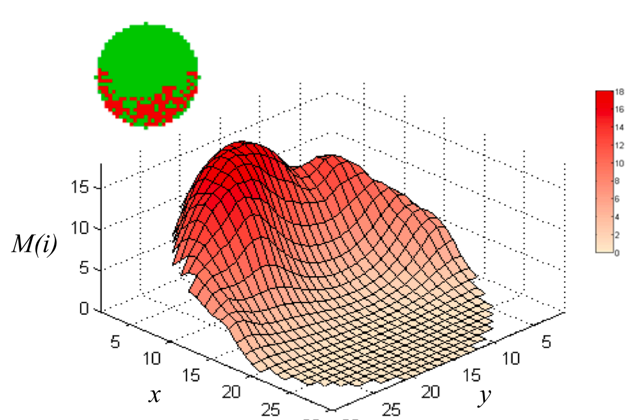
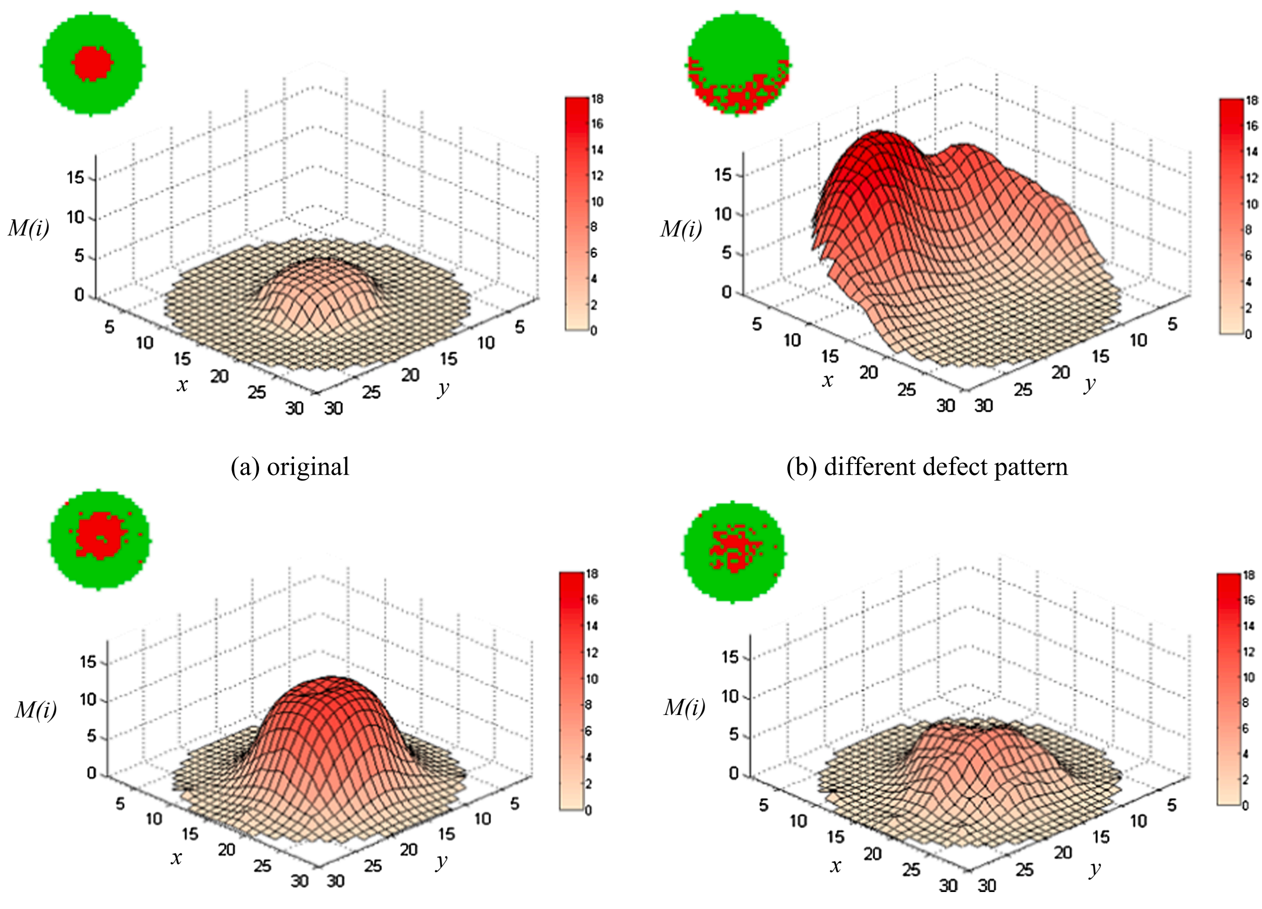


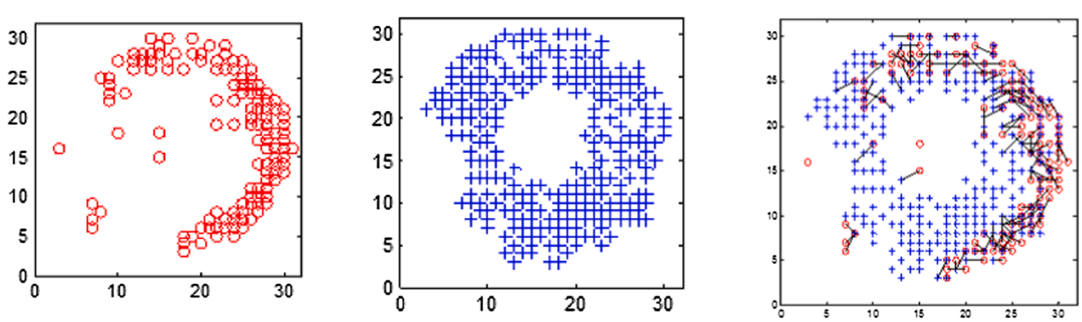
Fig. 5. Example of a mountain value derived from wafer bin map.



(a) original (b) different defect pattern

(c) different size (d) different density

Fig. 6. Example of a mountain value with different defect pattern, size, and density level.



(a) The selected WBM (b) The compared WBM (c) The matching process

Fig. 7. Illustration of matching process between selected WBM and compared WBM.

Table 1
The setting of weight and outlier score in similarity measurement.

	Outliner score	
Weight	**Without(O_0)**	**With(O_1)**
No weight(N)	N-O_0	N-O_1
Weight 1(W_1)	W_1-O_0	W_1-O_1
Weight 2(W_2)	W_2-O_0	W_2-O_1

Table 2
Confusion Matrix.

	System prediction	
Actual label	**Similar**	**Dissimilar**
Similar	True Positive (TP)	True Negative (TN)
Dissimilar	False Positive (FP)	False Negative (FN)

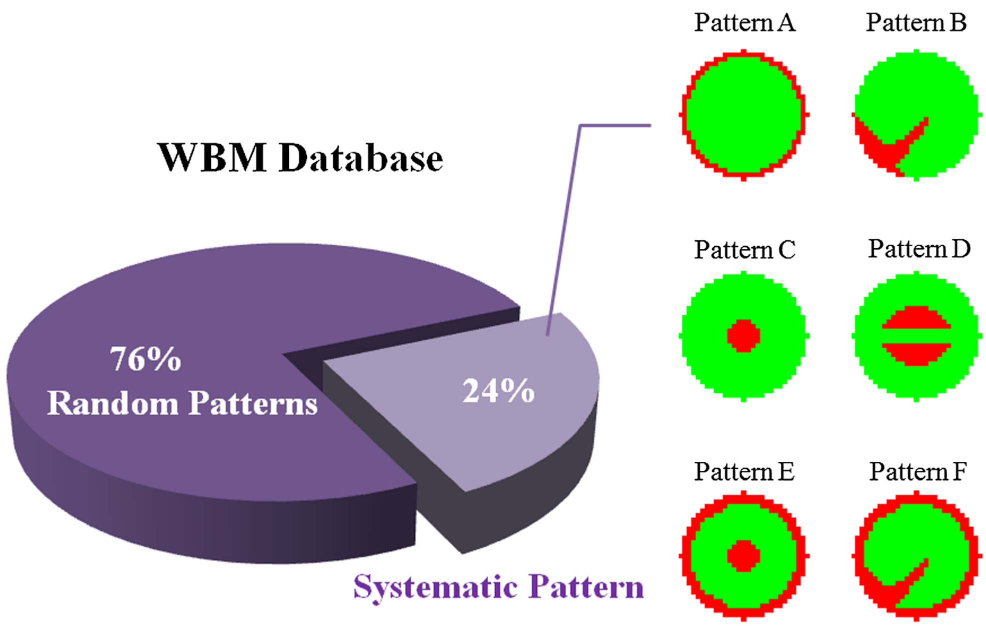


Fig. 8. WBM database composition.

Table 3
Simulation design.

Pattern	Yield high/low	Noise (+)	Density (−)	Shift	Rotation	trail	Total
Random (R)	0.6–1.0	N/A	N/A	N/A	N/A	190	760
A	H:0.82 L:0.70	5%, 10%, 15%	10%, 20%, 30%	N/A	N/A	4	48
B	H:0.87 L:0.68	5%, 10%, 15%	10%, 20%, 30%	N/A	0° 75°	2	48
C	H:0.94 L:0.75	5%, 10%, 15%	10%, 20%, 30%	0%, 5%	N/A	2	48
D	H:0.82 L:0.74	5%, 10%, 15%	10%, 20%, 30%	0%, 5%	0° 75°	1	48
Complex E,F	H:0.64 L:0.57 H:0.63 L:0.59	5%, 10%, 15%	10%, 20%, 30%	N/A	N/A	1	48

$$h(\mathbf{A}, \mathbf{B}) = \frac{1}{N_A} \sum_{a=1}^{N_A} \min_{b \in B} (\omega_{ab} \times \| a - b \|) \tag{2}$$

where the parameter ω_{ab} is a weight for WMHD. Considering the features of WBM, this study proposed three types of weight design as follows:

(a) No weight: Only use spatial distance as similarity measurement.

$$\omega_{ab} = 1 \tag{3}$$

(b) Weight design 1: The mountain value is applied here as weight ω_{ab} and ω_{ba}. Denote mountain value of a is M_a, mountain value of b is M_b:

$$\omega_{ab} = \frac{|M_a - M_b|}{\max(M_a, M_b)} \tag{4}$$

(c) Weight Design 2: Naive matching kernel (the closest value) will certainly fail and results into many false matches. Hence not only the interest point but the corresponding close local features in the image space, in which "the neighborhood" should also be considered.

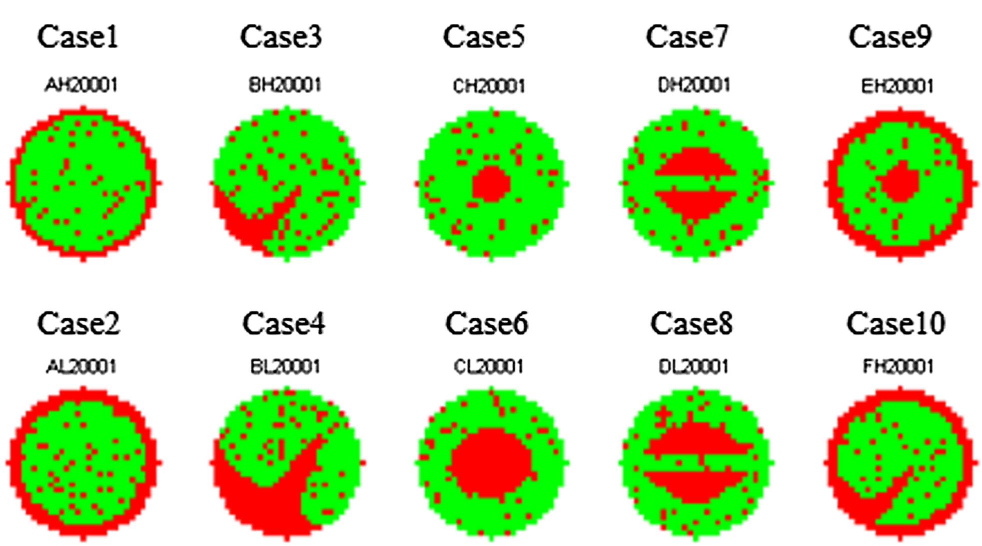


Fig. 9. 10 Selected WBMs.

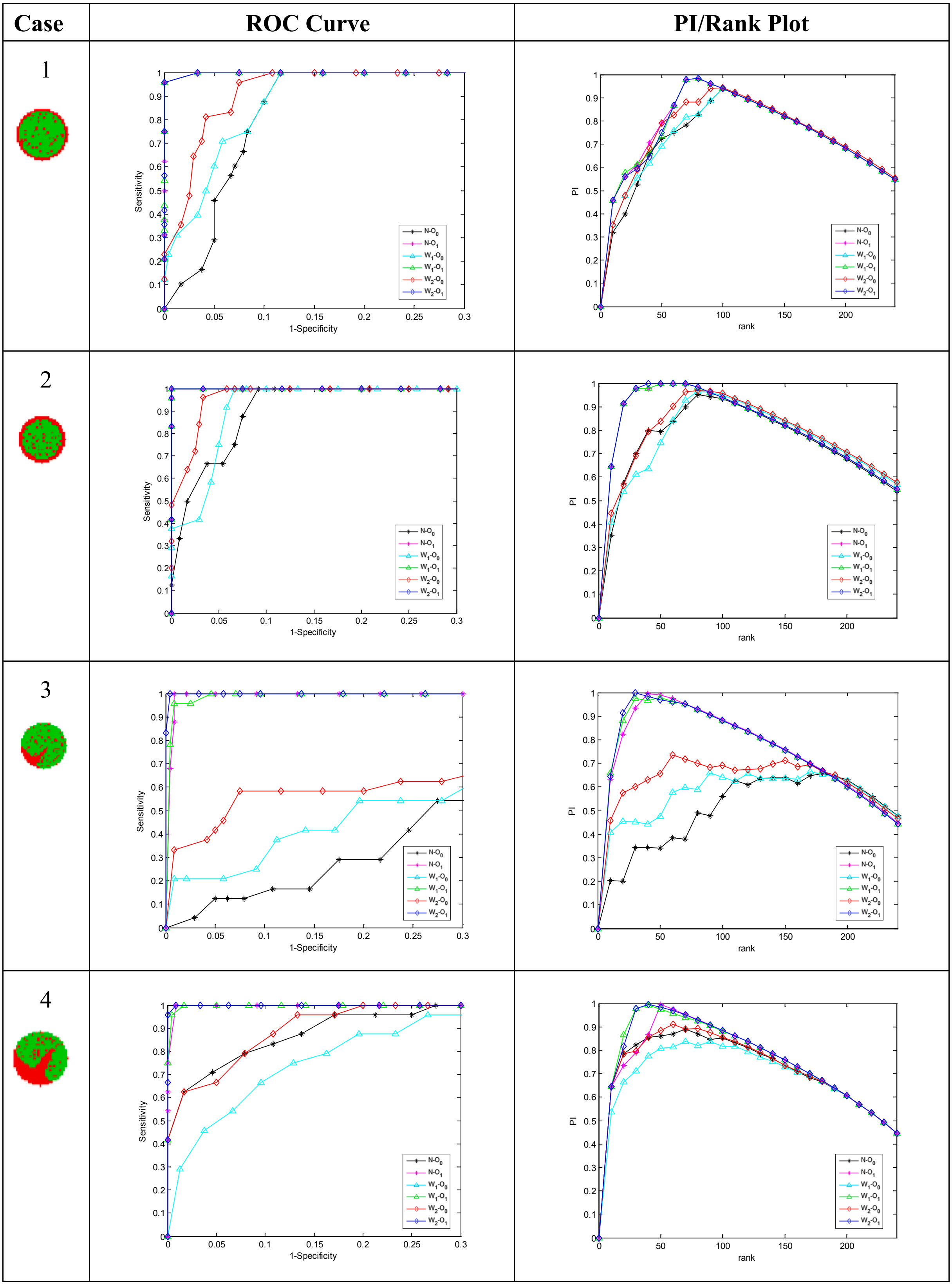


Fig. 10. Similarity matching comparison result for various selected WBMs.

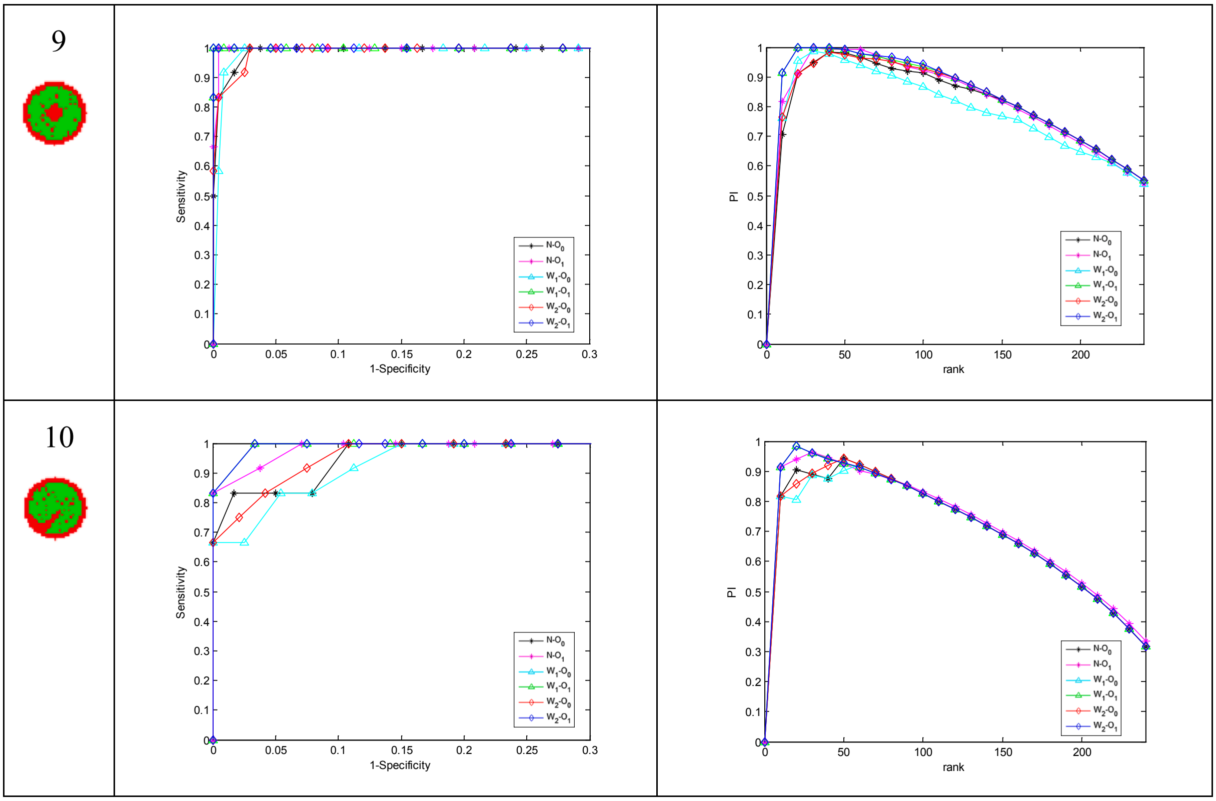


Fig. 10. (*continued*)

Table 4
Maximum PI for each selected WBM.

	Without outliner detection			With outliner detection		
	W_0	W_1	W_2	W_0	W_1	W_2
Case1	0.962	0.968	0.983	1.000	1.000	1.000
Case2	0.942	0.946	0.962	1.000	1.000	1.000
Case3	0.669	0.672	0.735	0.996	0.987	0.998
Case4	0.779	0.900	0.812	0.987	0.994	1.000
Case5	0.854	0.779	0.911	1.000	0.998	1.000
Case6	0.726	0.865	0.772	0.987	0.992	0.987
Case7	0.935	0.866	0.979	1.000	0.979	0.979
Case8	0.965	0.964	0.955	0.935	0.954	0.965
Case9	0.935	0.829	0.979	1.000	1.000	1.000
Case10	0.922	0.860	0.933	0.940	0.964	0.970
average	0.869	0.865	0.902	0.985	0.987	0.990

Table 5
Sensitivity under specificity = 90%.

Specificity = 90%						
	Without outliner detection			**With outliner detection**		
	W_0	W_1	W_2	W_0	W_1	W_2
Case1	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Case2	79.2%	81.3%	100.0%	100.0%	100.0%	100.0%
Case3	16.7%	25.0%	58.3%	100.0%	100.0%	100.0%
Case4	79.2%	66.7%	87.5%	100.0%	100.0%	100.0%
Case5	83.3%	75.0%	54.2%	100.0%	100.0%	100.0%
Case6	100.0%	100.0%	100.0%	95.8%	100.0%	100.0%
Case7	91.7%	83.3%	91.7%	91.7%	100.0%	100.0%
Case8	35.3%	85.3%	100.0%	100.0%	100.0%	100.0%
Case9	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Case10	100.0%	83.3%	100.0%	100.0%	100.0%	100.0%
average	78.5%	80.0%	89.2%	98.8%	100.0%	100.0%

Table 6
Specificity under sensitivity = 90%.

Sensitivity = 90%						
	Without outliner detection			**With outliner detection**		
	W_0	W_1	W_2	W_0	W_1	W_2
Case1	89.6%	89.6%	92.5%	100.0%	100.0%	100.0%
Case2	92.5%	94.2%	97.1%	100.0%	100.0%	100.0%
Case3	48.3%	46.7%	52.5%	99.2%	99.6%	100.0%
Case4	83.3%	74.2%	87.5%	100.0%	99.6%	100.0%
Case5	95.4%	94.6%	94.6%	92.1%	95.8%	97.5%
Case6	59.2%	73.3%	69.6%	100.0%	100.0%	100.0%
Case7	92.1%	86.7%	92.9%	92.9%	95.0%	95.8%
Case8	85.0%	85.8%	94.2%	94.2%	95.0%	99.2%
Case9	99.2%	99.6%	98.8%	99.6%	100.0%	100.0%
Case10	90.8%	90.0%	93.3%	97.5%	100.0%	100.0%
average	83.5%	83.5%	87.3%	97.6%	98.5%	99.3%

$$\omega_{ab} = \sum_{p\in(-1,0,1),j\in(-1,0,1)} \frac{|M_{a(x+p,y+q)} - M_{b(x+p,y+q)}|}{\max(M_{a(x+p,y+q)}, M_{b(x+p,y+q)})} \tag{5}$$

where (x,y) is the coordinate on a wafer, the parameters p and q denote as the adjacent location of die (x,y) by King-move (Hsu & Chien, 2007).

To magnify the difference between two WBMs with outlier of defective die, we define the maximum matching tolerance as S_{out}. Two defect dies be able to match only if the spatial distance is less than S_{out}, otherwise offer S_{out} as punishment. Fig. 7 illustrates the matching process between selected WBM and the compared WBM. Fig. 7(a) is the selected WBM and Fig. 7(b) is the compared WBM. In particular, not all defective die on the selected are used for similarity measurement and only spatial distance is within the matching tolerance will be counted as shown in the Fig. 7(c). Assuming there are N_o detective dice with large difference. Therefore, the similar between two WBMs is the summary of WMHD $h(\mathbf{A},\mathbf{B})$ and outliner score $(N_o \times S_{out})$.

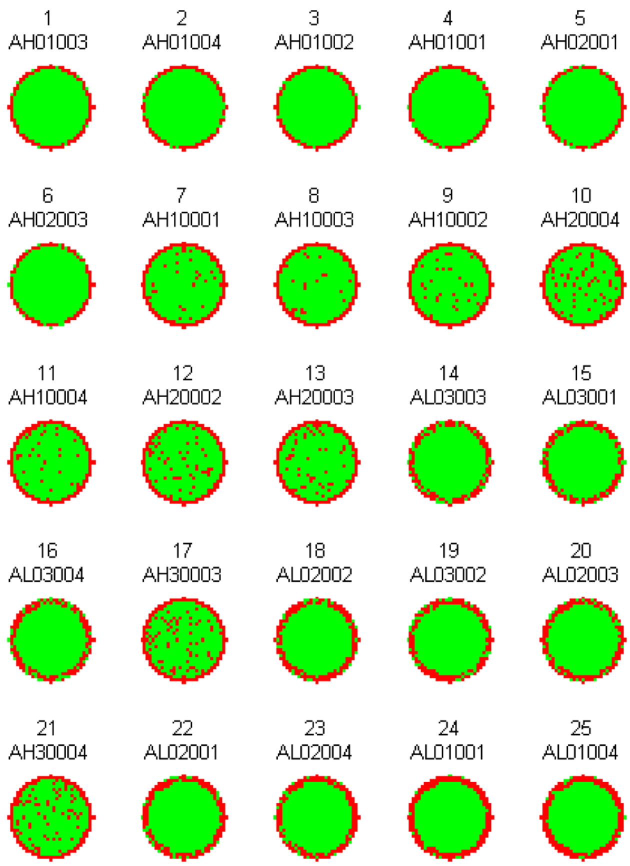


Fig. 11. Similarity matching result of Case 1.

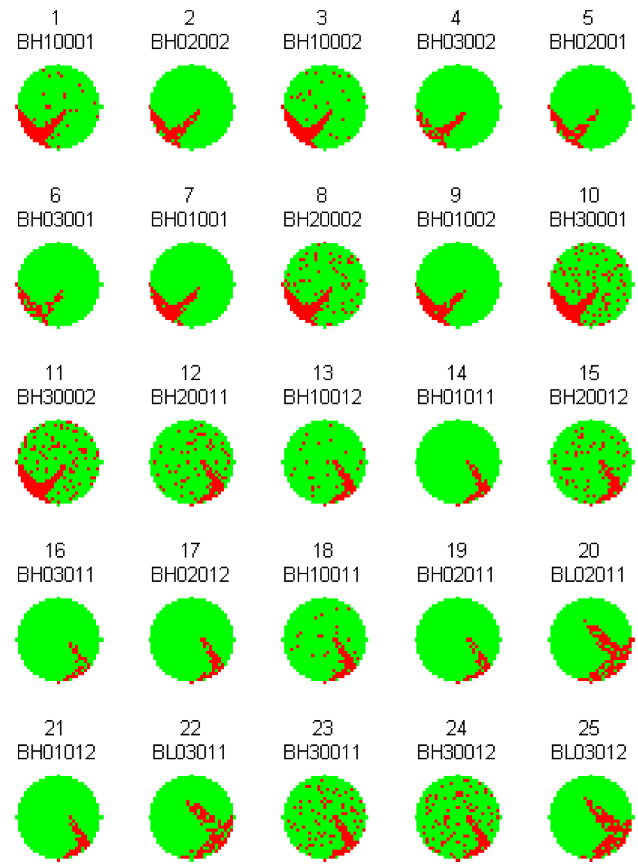


Fig. 12. Similarity matching result of Case 3.

$$S(\mathbf{A}, \mathbf{B}) = h(\mathbf{A}, \mathbf{B}) + N_o \times S_{out} \tag{6}$$

Table 1 summarizes the calculation strategy for weight design and outlier detection. The dissimilar score is the standard to measure two wafer bin map. Notice that, the smaller the dissimilar score presents two wafer bin maps is more similar.

3.4. Model evaluation

Rank the WBMs by dissimilar score in ascendant order. Based on the similar and dissimilar result between system predict and actual label, true-positive (TP), false-positive (FP), true-negative (TN), and false-negative (FN) are calculated as shown in Table 2. Sensitivity and Specificity are used to evaluate the model.

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \tag{7}$$

$$Specificity = \frac{TN}{FP + TN} \times 100\% \tag{8}$$

The performance index (PI) is defined as geometric average between Sensitivity and Specificity:

$$PI = \sqrt{Sensitivity \times Specificity} \tag{9}$$

With sensitivity as longitudinal axis and 1-specificity as horizontal axis, we are able to plot the receiver operating characteristic curve (ROC) curve.

4. Empirical study

4.1. Problem structuring

To validate the effectiveness of the proposed framework, but also obey the confidential clause with cooperate company, the data is simulated and adjust form real manufacturing data. The database consists of 1000 WBMs and 6 major patterns as shown in Fig. 8. To approximate the real situation, the database consists of 76% amounts of random patterns (yield form 60%–100%), and 24% amounts of symmetric patterns.

Since each pattern has different features, the considered factors including different number of defective die, noise degree, pattern density, degree of pattern shift and wafer rotation. The features can be described as edge-locate, symmetry, and complex. Consider different level of yield, noise, density, shift and rotation angle, the detail of simulation design for compared WBMs is shown in Table 3.

To be able to fully demonstrate influence of factors for the similarity matching between WBMs including pattern size, noise degree, pattern density, pattern shift, and wafer rotation, we generate 10 selected WBMs with two pattern size in high yield and low yield, and 10 random defective die as shown in Fig. 9. Cases 1 and 2 represent the different pattern size. The similar idea also uses for cases 3 and 4, cases 5 and 6, cases 7 and 8. Cases 9 and 10 are used as complex WBMs with two defect patterns on a WBM.

4.2. Model evaluation with different settings of weight and outlier score

Fig. 10 shows the ROC curves and PI/Rank plots of each selected WBM, the interval is every 10 rank, to evaluate the result form 6 WMHD design as shown in Table 1. According to the curve of PI in different rank, the optimal threshold of similar WBM can be determined. Based on the result of ROC curves in these 10 cases, the performance of WBM similarity matching (i.e., N-O_0, W_1-O_0, W_2-O_0) by using no outlier detection factor S_{out} are worse than the method using outlier detection factor (N-O_1, W_1-O_1, W_2-O_1).

According to Table 4, the WMHD which combines with outliner

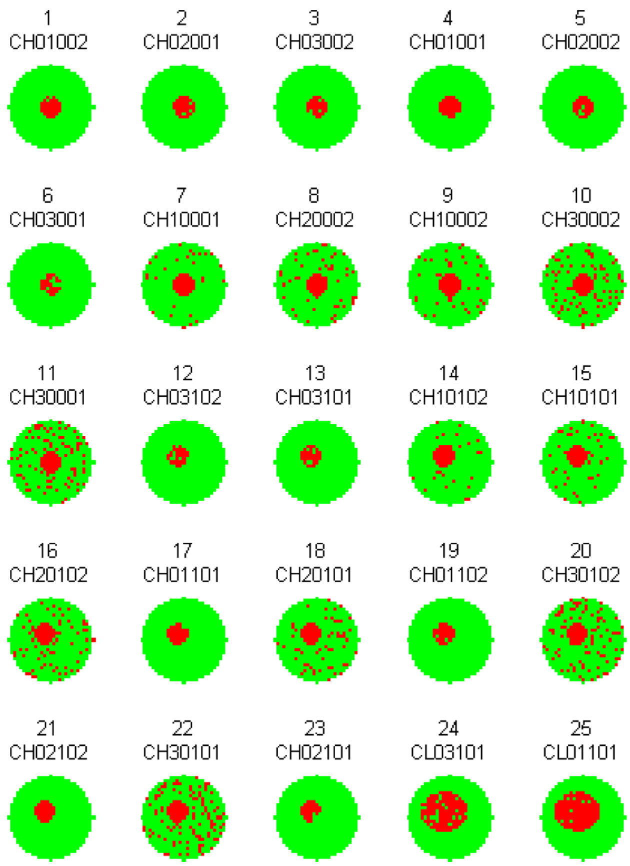


Fig. 13. Similarity matching result of Case 5.

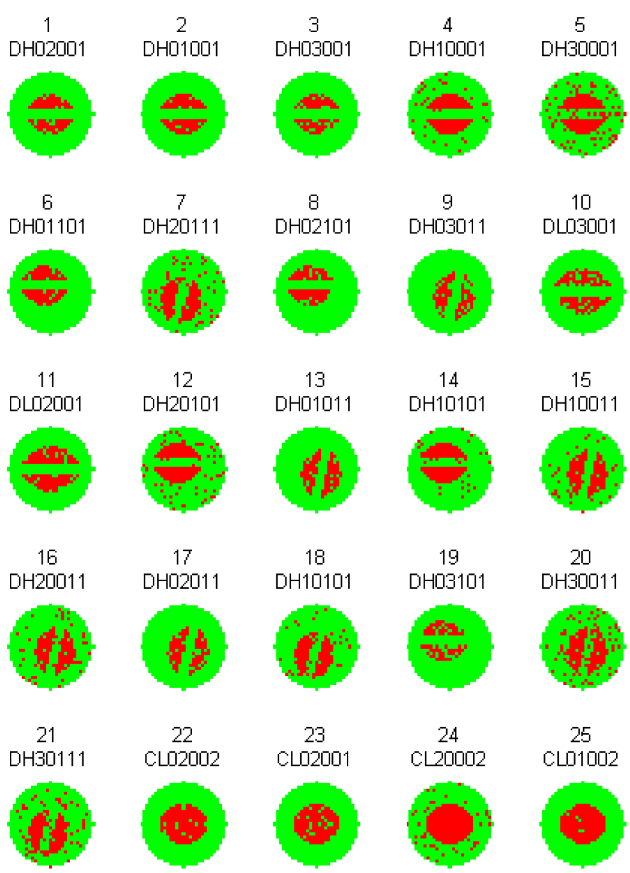


Fig. 14. Similarity matching result of Case 7.

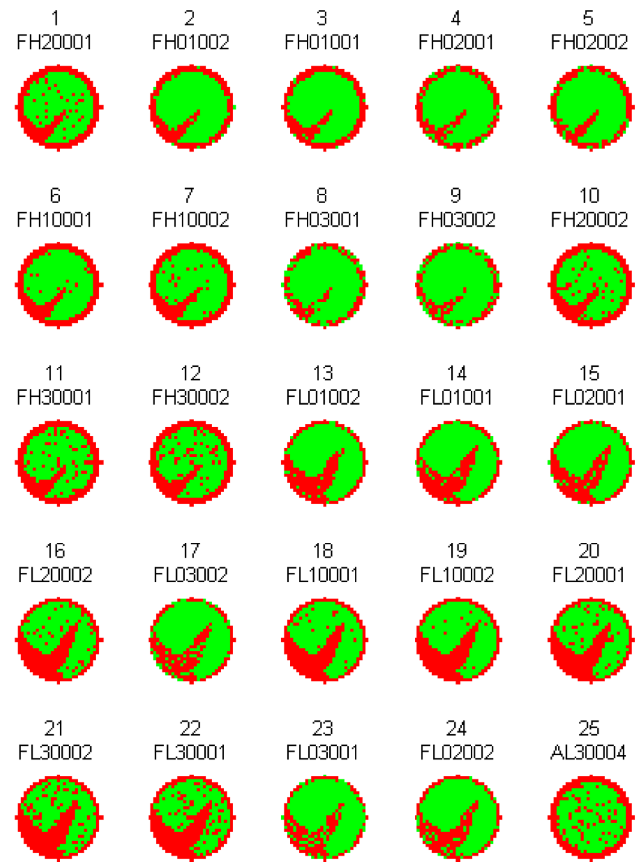


Fig. 15. Similarity matching result of Case 10.

detection performs much better than which not detect the outliner. Hence, the outliner detection is beneficial in Sensitivity and Specificity enhancement. Another advantage of combining outliner detection is high ranking consistency. On the other hand, without outliner detection, the performance deviate from different testing case, especially the case be impact with location change (shift, rotation) factor.

Table 4 shows the maximum PI value that reveals the weight design have great influence of system performance. Different weight design also results in different system validity. Overall evaluation, the weight design 2 (W_2-O_1) is outperform the weight design 1 and no weight. Even though W_2-O1 is not the best on case 5 and 6, the PI still reach 0.95. Therefore, it can conclude that the W_2-O_1 is the optimal WMHD.

The defect area of selected WBM have considering influence of system performance. Same pattern, the larger the defect area (lower the yield), the better the analysis result. The reason is the spatial distance is one of measurement standard of the similarity level. In addition, the defect area is also impact in the scale of mountain value. The larger the defect area is, the larger the deviation of spatial distance and weight is. The difference of two wafer bin maps can be emphasized. As the result, the defect pattern which yield low have better performance.

Without outliner detection, the spatial distance dominates the similarity score, and have much greater impact than weight. Thus the performance of system is case by case. By combing the outliner detection, the distance can be normalized and the performance of system is relatively stable and consistence.

Table 5 shows the analysis result of WBM similar matching under fixing the Specificity on 90%, examine the Sensitivity of system, with outliner detection, it's able to reach 90% of sensitivity. Table 6 presents the analysis result of WBM similar matching under fixing the Sensitivity on 90%. Table 6 also examines the Specificity, which the error is under 5% with outliner detection and weight design 2 (W_2-O_1). The results indicates that system can precisely detect 90% of similar WBMs in the

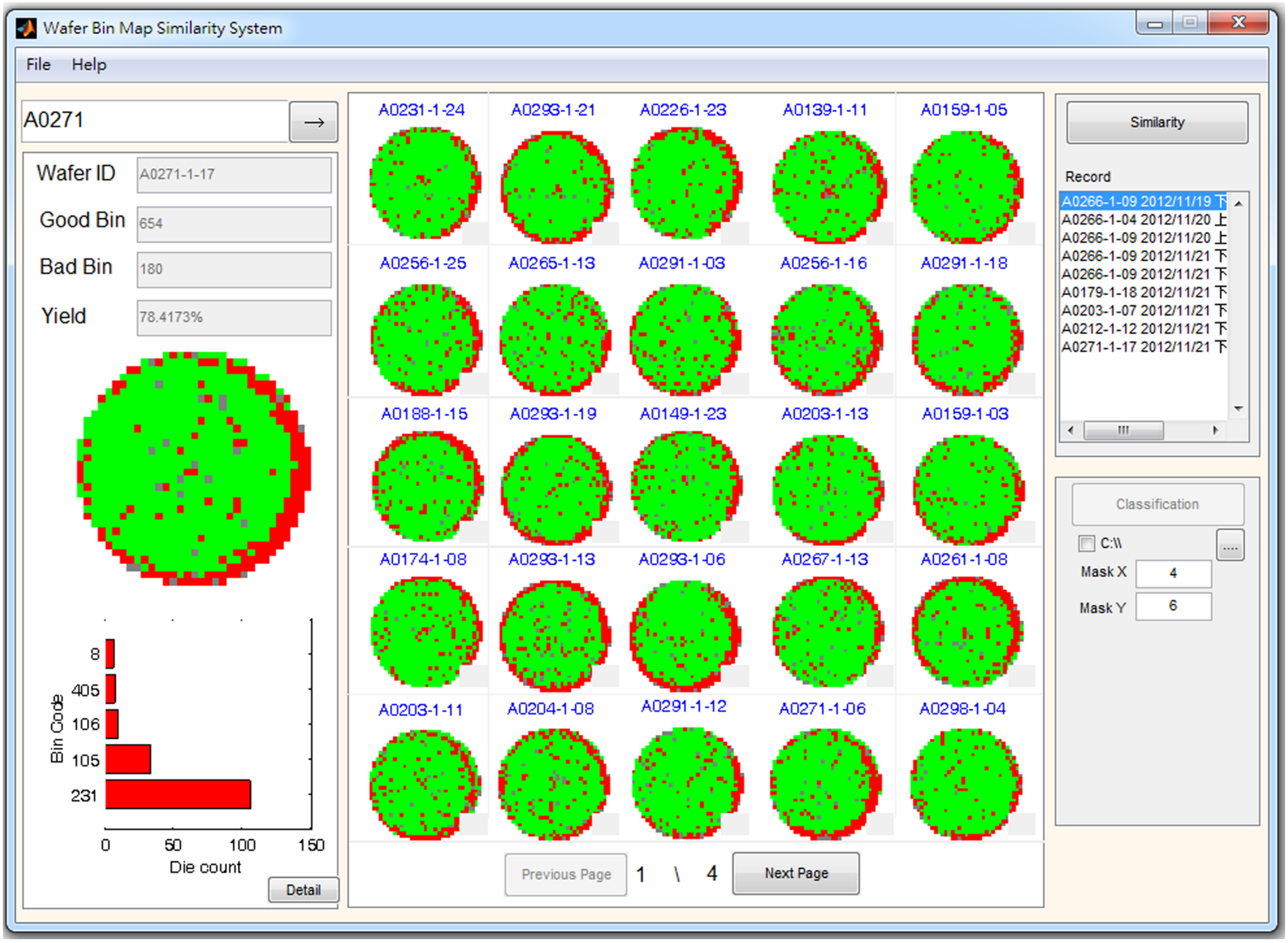


Fig. 16. Interface of intelligent WBM similarity matching system.

database when 10% of false alarm occurs, make under 5% error when the system detect 90% of similar WBMs.

4.3. Discussion of WBM similarity matching

This section demonstrates and discuss the impact of each factor under different level by using visual inspection. We only present the W_2-O_1 analysis result and 5 representative cases (Case1, Case3, Case5, Case7, and Case10) in rank 25.

Fig. 11 shows the similarity matching result of Case 1. The proposed method is able to precisely identify the similar WBMs with various pattern size, pattern density, and noise degree (ex. #1AH01003, #11AH1004, and #22AL02001). Fig. 12 shows the similarity matching result of Case 3 and demonstrate that the proposed method not only can handle noise and density (ex. #1BH10001 and #12BH20011), but also able to deal with wafer rotation. The no rotation WBMs (i.e., rotation angle is 0°) appeal more similar to system than those rotated. The similarity matching result of Case 5 is shown in Fig. 13. It shows that the proposed method can detect not only the shifting defect patterns (ex. #12CH03102 and #22CH30101) but also the large pattern size with the same defect pattern (ex. #24CL03101 and #25CL01101). In particular, the defect patterns with no shift still have smaller rank than those WBMs with shift defect pattern. Fig. 14 shows the similarity matching result of Case 7. It demonstrates that the performance of proposed method under the defect pattern with pattern shift and wafer rotate. The first 21 WBMs in Fig. 14 are similar with Case 7. However, there is some false-alarm at #22 to #25. The reason is the shape features of selected WBM (Case 7) and the compared WBMs (#22 to #25) have the similar pattern of defective die which are both located in the middle of wafer. Additionally, the shape seems like a round circle. Case 10 is used to evaluate the capability of identifying similar complex patterns. Fig. 15 shows that the proposed method can identify the most of similar WBMs except for #25AL30004. Since the complex pattern is composed by two different defect patterns, which were caused by two kinds of failure causes during the manufacturing process. In root cause identification by WBM analysis in practice, detecting only one of the patterns still can provide useful information for engineer to find the potential root cause.

Simulation results demonstrate that the proposed model can provide precisely similarity measurement between similar and dissimilar pattern. The similarity matching system proposed by this study have following features:

(1) The proposed WBM similarity matching method is able to identify the similar WBM by considering these five factors (pattern size, noise degree, pattern density, pattern shift, and wafer rotation). The degree of pattern size and pattern shift have large impact for WBM similarity matching. The pattern density and noise degree have small impact for similarity matching.
(2) Two different patterns which have approximate defect size and location might incorrectly determine as similar patterns. It might express that the proposed similarity matching method can process considering non-rigid deformation.
(3) The proposed WBM similarity matching method is efficiently

screened out the random patterns while identifying the similar WBM with low false alarm.

This study design the database and testing case considering the patterns which are common to fab (Patterns A and C), the patterns which are very rare in amount and have no symmetry appearance (Patterns B and D), and relatively rare complex patterns (Patterns E and F). Therefore, the system is capable to analysis most of realistic situations.

Furthermore, the proposed WBM similarity matching has been embedded into a decision support system and has been implemented in a semiconductor company in Taiwan. Fig. 16 shows the interface of the developed system of WBM similarity matching. This decision support system of WBM similarity matching can assist engineers to identify defect patterns efficiently and thus diagnose the root cause of defect pattern effectively to improve quick response for smart production.

5. Conclusion

To address the needs to deal with new defect patterns, this study developed an approach of measuring the similarity among wafer bin maps that can accurately determine the similar level despite size, noise, density, shift, and rotation issue. A decision support system embedded with the developed similarity matching algorithms is also constructed to provide an effective method to empower human judgments of the engineers, as proposed in Industry 3.5. Comparing with existing approaches for WBM analysis, the proposed similarity matching can achieve higher performance on failure analysis than pattern recognition and classification, especially in advanced technologies foe semiconductor manufacturing. Moreover, the developed system does not need to pre-define the pattern categories, nor need to train the model with huge historical data. Moreover, the similarity also provides a clue to define a newly discovered pattern that may be closer to some of existing patterns to support root cause diagnosis.

Future research can be done to enforce the data linkage between the developed system and manufacturing process data, increase the efficiency and automatic level of failure analysis. Further study should be done to improve the functions and control rules generated from the proposed framework through lessons learned from implementation in real settings. Similar approaches can be done for other problems to empower the engineers for analysis and decision making involved for smart production.

CRediT authorship contribution statement

Chia-Yu Hsu: Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Resources, Supervision, Project administration. **Wei-Ju Chen:** Investigation, Validation, Writing - original draft. **Ju-Chien Chien:** Data curation, Visualization, Formal analysis.

Acknowledgements

This research is supported by Ministry of Science and Technology, Taiwan ((MOST106-2628-E-027-002-MY3; MOST108-2813-C-027-017-E).

References

Baly, R., & Hajj, H. (2012). Wafer classification using support vector machines. *IEEE Transactions on Semiconductor Manufacturing, 25*(3), 373–383.
Chao, L.-C., & Tong, L.-I. (2009). Wafer defect pattern recognition by multi-class support vector machines by using a novel defect cluster index. *Expert Systems with Applications, 36*(6), 10158–10167.
Chien, C. F., Dou, R., & Fu, W. (2018). Strategic capacity planning for smart production: Decision modeling under demand uncertainty. *Applied Soft Computing, 68*, 900–909.
Chien, C.-F., & Hsu, C.-Y. (2006). A novel method for determining machine subgroups and backups with an empirical study for semiconductor manufacturing. *Journal of Intelligent Manufacturing, 17*(4), 429–439.
Chien, C.-F., & Hsu, C.-Y. (2011). UNISON analysis to model and reduce step-and-scan overlay errors for semiconductor manufacturing. *Journal of Intelligent Manufacturing, 22*(3), 399–412.
Chien, C.-F., & Hsu, C.-Y. (2014). Data mining for optimizing IC feature designs to enhance overall wafer effectiveness. *IEEE Transactions on Semiconductor Manufacturing, 27*(1), 71–82.
Chien, C.-F., & Chuang, S.-C. (2014). A framework for root cause detection of sub-batch processing system for semiconductor manufacturing big data analytics. *IEEE Transactions on Semiconductor Manufacturing, 27*(4), 475–488.
Chien, C.-F., Chen, Y.-J., & Peng, J.-T. (2010). Manufacturing intelligence for semiconductor demand forecast based on technology diffusion and product life cycle. *International Journal of Production Economics, 128*(2), 496–509.
Chien, C.-F., Chou, C.-W., & Yu, H.-C. (2016). A Novel route selection and resource allocation approach to improve the efficiency of manual material handling system in 200-mm wafer fabs for industry 3.5. *IEEE Transactions on Automation Science and Engineering, 13*(4), 1567–1580.
Chien, C.-F., Hong, T.-Y., & Guo, H.-Z. (2017). An empirical study for smart production for TFT-LCD to empower Industry 3.5. *Journal of the Chinese Institute of Engineers, 40*(7), 552–561.
Chien, C.-F., Hsu, C.-Y., & Hsiao, C. (2012). Manufacturing intelligence to forecast and reduce semiconductor cycle time. *Journal of Intelligent Manufacturing, 23*(6), 2281–2294.
Chien, C.-F., Hsu, C.-Y., & Chen, P.-L. (2013). Semiconductor fault detection and classification for yield enhancement and manufacturing intelligence. *Flexible Services and Manufacturing Journal, 25*(3), 367–388.
Chien, C.-F., Hsu, S.-C., & Chen, Y.-J. (2013). A system for online detection and classification of wafer bin map defect patterns for manufacturing intelligence. *International Journal of Production Research, 51*(8), 2324–2338.
Chien, C.-F., Liu, C.-W., & Chuang, S.-C. (2017). Analysing semiconductor manufacturing big data for root cause detection of excursion for yield enhancement. *International Journal of Production Research, 55*(17), 5095–5107.
Dou, R., Chien, C. F., Kacem, I., & Hsu, C. Y. (2018). Industry applications of computational intelligence: Preface. *International Journal of Computational Intelligence Systems, 11*(1), 803–804.
Dou, R., He, Z., & Hsu, C. Y. (2018). Foreword: Smart manufacturing, innovative product and service design to empower Industry 4.0. *Computers & Industrial Engineering, 125*, 514–516.
Hsu, S.-C., & Chien, C.-F. (2007). Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing. *International Journal of Production Economics, 107*(1), 88–103.
Huttenlocher, D. P. & Rucklidge, W. J. (1992). A multi-resolution technique for comparing images using the Hausdorff distance. In: Cornell University.
Khakifirooz, M., Chien, C.-F., & Chen, Y.-J. (2018). Bayesian inference for mining semiconductor manufacturing big data for yield enhancement and smart production to empower industry 4.0. *Applied Soft Computing, 68*, 990–999.
Kuo, C.-J., Chien, C.-F., & Chen, C.-D. (2011). Manufacturing intelligence to exploit the value of production and tool data to reduce cycle time. *IEEE Transactions on Automation Science and Engineering, 8*(1), 103–111.
Li, T.-S., & Huang, C.-L. (2009). Defect spatial pattern recognition using a hybrid SOM–SVM approach in semiconductor manufacturing. *Expert Systems with Applications, 36*(1), 374–385.
Liao, C.-S., Hsieh, T.-J., Huang, Y.-S., & Chien, C.-F. (2014). Similarity searching for defective wafer bin maps in semiconductor manufacturing. *IEEE Transactions on Automation Science and Engineering, 11*(3), 953–960.
Liu, C.-W., & Chien, C.-F. (2013). An intelligent system for wafer bin map defect diagnosis: An empirical study for semiconductor manufacturing. *Engineering Applications of Artificial Intelligence, 26*(5–6), 1479–1486.
Liu, S., Chen, F., & Lu, W. (2002). Wafer bin map recognition using a neural network approach. *International Journal of Production Research, 40*(10), 2207–2223.
Lu, Y., Tan, C. L., Huang, W., & Fan, L. (2001). An approach to word image matching based on weighted Hausdorff distance. *Proceedings of sixth international conference on document analysis and recognition* (pp. 921–925). IEEE.
Nakazawa, T., & Kulkarni, D. V. (2018). Wafer map defect pattern classification and image retrieval using convolutional neural network. *IEEE Transactions on Semiconductor Manufacturing, 31*(2), 309–314.
Nakazawa, T., & Kulkarni, D. V. (2019). Anomaly detection and segmentation for wafer defect patterns using deep convolutional encoder–decoder neural network architectures in semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing, 32*(2), 250–256.
Piao, M., Jin, C. H., Lee, J. Y., & Byun, J. Y. (2018). Decision tree ensemble-based wafer map failure pattern recognition based on radon transform-based features. *IEEE Transactions on Semiconductor Manufacturing, 31*(2), 250–257.
Saqlain, M., Jargalsaikhan, B., & Lee, J. Y. (2019). A voting ensemble classifier for wafer map defect patterns identification in semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing, 32*(2), 171–182.
Taam, W., & Hamada, M. (1993). Detecting spatial effects from factorial experiments: An application from integrated-circuit manufacturing. *Technometrics, 35*(2), 149–160.
Wang, C.-H. (2008). Recognition of semiconductor defect patterns using spatial filtering and spectral clustering. *Expert Systems with Applications, 34*(3), 1914–1923.
Wang, C.-H., Kuo, W., & Bensmail, H. (2006). Detection and classification of defect patterns on semiconductor wafers. *IIE Transactions, 38*(12), 1059–1068.
Wu, M. J., Jang, J. S. R., & Chen, J. L. (2015). Wafer map failure pattern recognition and similarity ranking for large-scale data sets. *IEEE Transactions on Semiconductor Manufacturing, 28*(1), 1–12.
Yager, R. R., & Filev, D. P. (1994). Approximate clustering via the mountain method. *IEEE Transactions on Systems, Man, and Cybernetics, 24*(8), 1279–1284.

Yang, M.-S., & Wu, K.-L. (2005). A modified mountain clustering algorithm. *Pattern Analysis and Applications, 8*(1–2), 125–138.

Yu, J., & Lu, X. (2016). Wafer map defect detection and recognition using joint local and nonlocal linear discriminant analysis. *IEEE Transactions on Semiconductor Manufacturing, 29*(1), 33–43.

Yu, J., Zheng, X., & Liu, J. (2019). Stacked convolutional sparse denoising auto-encoder for identification of defect patterns in semiconductor wafer map. *Computers in Industry, 109*, 121–133.

Yuan, T., Bae, S. J., & Park, J. I. (2010). Bayesian spatial defect pattern recognition in semiconductor fabrication using support vector clustering. *The International Journal of Advanced Manufacturing Technology, 51*(5–8), 671–683.

Yuan, T., Kuo, W., & Bae, S. J. (2011). Detection of spatial defect patterns generated in semiconductor fabrication processes. *IEEE Transactions on Semiconductor Manufacturing, 24*(3), 392–403.