Proceedings of the World Congress on Engineering and Computer Science 2013 Vol II
WCECS 2013, 23-25 October, 2013, San Francisco, USA

An Accurate Hybrid-Similarity Technique for User-Defined Wafer Fail-Map Pattern Detection

Sunho Song, Bongseok Kim, Sunjae Lee, Youngkyun Jeong, Hanjig Cho and Junghee Kim

***Abstract*—A Hybrid-Similarity technique is proposed for improving the matching accuracy in wafer fail-map pattern detection compared with Cosine-Similarity and Jaccard-Similarity. This is good for gathering the failure data in engineer pre-defined patterns. The adapted low pass filtering and cutting-off technique make the matching calculation simpler and faster than conventional methods. From 5 kinds of typical fail bin maps, this Hybrid-Similarity technique has achieved 87.21% of accuracy which is 8.85% and 17.57% higher than Cosine-Similarity and Jaccard-Similarity, respectively. Execution time of Hybrid-Similarity is 4.05 milliseconds that is 250 times faster than technique of Neural Network.**

***Index Terms*—Fail Bin Map (FBM), Neural Network, Scale Invariant Feature Transformation (SIFT), Yield, Yield Management System (YMS)**

I. Introduction

THe semiconductor industry has developed in terms of mass speed and low power. Currently, the number of CPU transistors in Intel is integrated to more than 2.9 billion [1]. This essential point of progress of the semiconductor industry is miniaturization of transistor size and as it becomes smaller, semiconductor manufacturing process have become more and more complicated. Also the process is automatically controlled and monitored. However, failures have been explosively increasing by the complicated manufacturing process and the amount of failure analysis operation has been increased significantly [2].

In general, the various analyses for each process are conducted during the manufacturing process. First of all, the yield trend is checked to find abnormal wafer fail-map pattern for analyzing the root cause of severe yield drop. During this step, the only numerical numbering information of yield trend data brings many restrictions and confusions to define the exact problems. Thus, if additional information on failed wafers is provided to engineers such as fail-map shapes or patterns, higher accuracy and reduced time-spending of root-cause analysis will be achieved. There are many kinds of wafer failure analysis techniques in the semiconductor industry. These techniques examine the whole wafer, spending a lot of time and occupying the memory storage of machines. Then the classified data is matched and picked-up manually by an engineer, or additional steps must be done until a match appears. These time consuming processes increase the total loss of manufacturing cost due to the long low-yield period. However, when the company has similar failure experience and mass-data to be referenced, engineers screen and pick-up specific patterned wafers, manually classifying the whole wafers and saving the data.

Various methods for pattern classification were previously proposed that have been automated in order to solve the problems above. Classical methods in [2-4] are based on the well-known Neural Network. It is possible to use the variables in several types of inputs and outputs, and then the results are provided to a combination of non-linear. This method brings excellent prediction as an advantage. On the other hand, there are several disadvantages. All inputs and outputs must be converted into 0 and 1. Too much information is required to calculate, such as the node and group count of spatial patterns. In addition, the node complexity is increased by type of patterns and eventually the operation speed becomes very slow. Techniques based on Neural Network takes approximately 1 second per each FBM pattern match [4]. This problem limits the performance of analysis used in situ during mass production.

In this paper, a Hybrid-Similarity technique is proposed in order to overcome the disadvantages of Neural Network. This new method has a simple algorithm and does not require the pre-learning pattern. It derives only the matched output with what FBM patterns engineer wants to detect. It also does not set any undefined variable parameter.

For better understanding, common data classification methods are explained in section II, and then the study of Hybrid-Similarity and experimental results are described in section III and IV, respectively.

II. Common Types of Classification

There are many pattern classifying methods, which are consist of Neural Network, Scale Invariant Feature Transformation (SIFT), and Engineer's experience.

A. Neural Network-based Classification

Neural Network based classification methods consist of the number of nodes, the classified Fail Bin Map (FBM) data, and interconnecting weight values between segmented nodes. The weight values are from the initial classification, and are transferred to the other nodes as inputs, and then each node modifies the weights [2-4]. The general configuration of the Neural Network is described in Fig. 1. When all the weights

Sunho Song is with Samsung Electronics co., Ltd, and Samsung Institute of Technology, Yongin-City Gyonggi-Do, Korea (corresponding author to provide e-mail: sunho.song@samsung.com).

Bongseok Kim (e-mail: bongdol.kim@samsung.com) and Sunjae Lee (e-mail: sunjae76.lee@samsung.com) is with Samsung Electronics co., Ltd, Yongin-City Gyonggi-Do, Korea.

Youngkyun Jeong is with Samsung Electronics co., Ltd, and Samsung Electronics co., Ltd, Yongin-City Gyonggi-Do, Korea (e-mail: youngk.jeong@samsung.com).

Hanjig Cho (e-mail: hanjig.cho@samsung.com) and Junghee Kim (e-mail: whitecos.kim@samsung.com) is with Samsung Electronics co., Ltd, Yongin-City Gyonggi-Do, Korea.

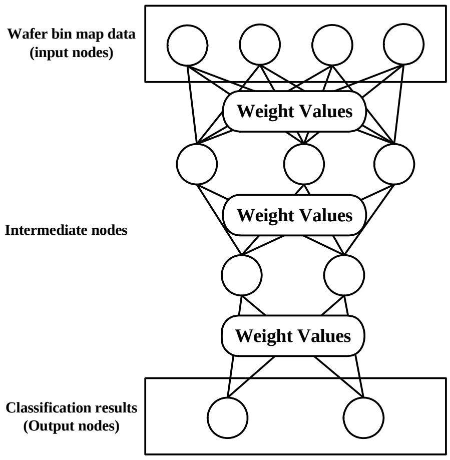


Fig. 1. Common configuration of Neural Network.

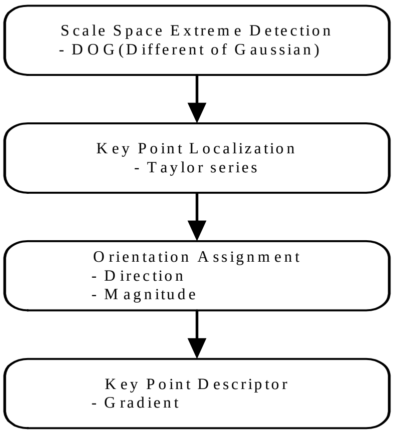


Fig. 2. Conceptual procedure and structure of SIFT

in the system are adjusted, this system is ready to classify the newly incoming pattern data.

On the other hand, new patterns cannot be correctly classified because all the weight values must finish adjusting before any new pattern classification. Also, the configuration of input, output, and intermediate nodes is too complicated to be recognized, so that the classified data by the Neural Network based method does not work well in engineering analysis.

B. *SIFT-based Classification*

SIFT is the method of extracting a specific key point which is not influenced by the change of shade, rotation and size in the wafer fail-map. With the extracted key point, a descriptor which is barely influenced by the change of the fail-map is created and the similarities among the fail-maps are judged by counting the number of similar characteristic spots [5]. Fig. 2 is the overall structure of the SIFT. The interest points which are not likely to be influenced by the scale and orientation are extracted first. Then, the second interest points will be extracted by the key point localization, and the points which have higher changing possibility will be re-sorted by using Taylor series. In the orientation assignment and key point descriptor, the direction, magnitude and gradient of key points are calculated. At last, the matching is started with the previously calculated values.

While the change of the scale of a FBM pattern is proportional to Gross Die, the key point which is not changed

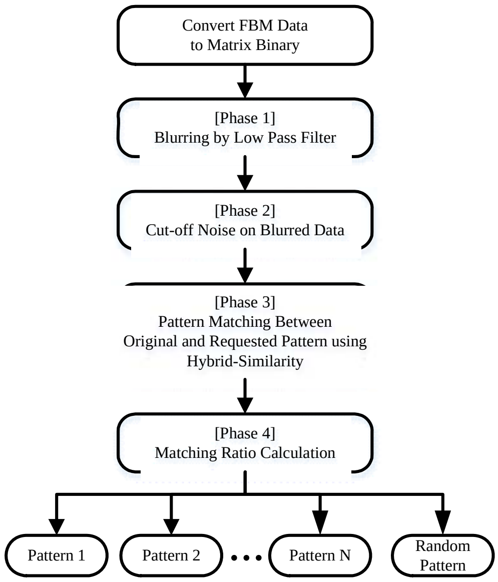


Fig. 3. Overall process of Hybrid-Similarity FBM classification

by the size and rotation is found and compared. So SIFT based FBM classification has many restrictions [6]. Also, it is not suitable for real-time pattern matching due to the long calculation time despite its high reliability to pattern matching regarding scale and illumination.

C. *Experience based Classification*

This method is based on the practical experience of the field engineers. The procedure is similar to the SIFT-based classification, but all the steps are done manually. It is very time consuming and the accuracy of results depends on the level of the engineer's skill.

III. HYBRID-SIMILARITY

The overall flow of FBM pattern classification using Hybrid-Similarity is explained in Fig. 3 and consists of four stages.

A. *Blurring and Cut-off Noise*

In Phase1, uncontrollable high frequency components will be filtered out by a Low Pass Filter (LPF). In the case of FBM data, non-patterned and patterned data exist at the same time. Usually the non-patterned data has an ungrouped shape and higher frequency components compare to the patterned data. Therefore, in order to enhance the classification precision, the high frequency in the non-patterned data should be removed by LPF [7]. Moreover, applying the LPF to input cluster data has higher efficiency than applying to the entire input matrix data [8]. In the proposed Hybrid-Similarity, the data cluster size is set to 7 before applying the LPF. However, some of none patterned data will remain after applying the LPF in phase 1; thus, the remains will be removed in Phase 2 secondly.

In Phase2, as shown in Fig. 4 (a), the previously filtered matrix data will be cleaned again in order to remove the remaining non-patterned data to achieve clearer data patterns. These unnecessary data is remain in more than one place of FBM, and there is a possibility to make errors and lower the accuracy in the final results. This uncertainty is higher in the

ISBN: 978-988-19253-1-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

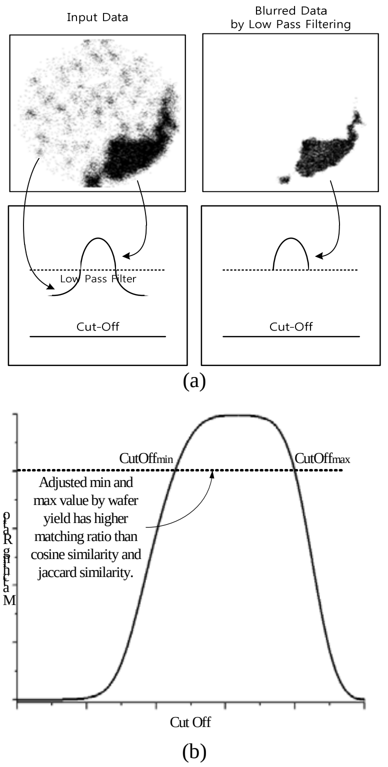


Fig. 4. (a) Cut-off non-pattern data on blurred data and (b) selecting plot for cut-off range limitation.

FBM of a low-yield wafer than that of a high-yield wafer, so that the cut-off setting value must be set as high for the low-yield wafer testing. The equation to choose the cut-off values is derived in equation (1). Because the cut-off values must be changed on the wafer yield value, the result will be changed adaptively by the wafer yield value.

$$CutOff = CutOff_{\min} + \left(1 - \frac{Yield}{100}\right) \cdot (CutOff_{\max} - CutOff_{\min}) \qquad (1)$$

$CutOff_{min}$ and $CutOff_{max}$ values stand for the low setting limit for the high-yield wafer testing and the high setting limit for the low-yield wafer testing. The values should be decided at the expecting accuracy of matching ratio, which is from the initial cut-off sweep testing as seen in Fig. 4 (b).

B. Pattern Matching and Calculation

In Phase 3, the processed matrix data after Phase 2 will be compared with the user defined FBM shape, and this user shape for Hybrid-Similarity is provided as a newly calculated matrix data format in an equation (2).

$$Hybrid\ \ Similarity = \sqrt{\frac{A \cdot B}{\|A\| \cdot \|B\|} \times \frac{A \cap B}{A \cup B}} \qquad (2)$$

$$\cos(\theta) = \frac{A \cdot B}{\|A\| \cdot \|B\|} \qquad (3)$$

$$J(A, B) = \frac{A \cap B}{A \cup B} \qquad (4)$$

In this equation, A and B are two different input matrix data in Phase 3, and it is the mathematical product of Cosine-Similarity in equation (3) and Jaccard-Similarity in equation (4). Jaccard-Similarity calculates the correlation of the range. Its result is obtained by the intersection of two vectors divided by the union value [9].

When the FBM pattern and the user defined pattern are compared, the information of location and area should be compared simultaneously as shown in Fig. 5, after which an accurate similarity will be obtained.

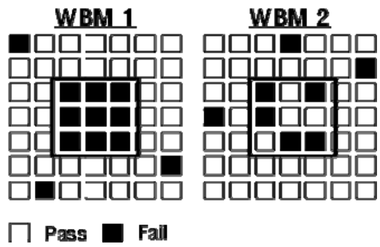


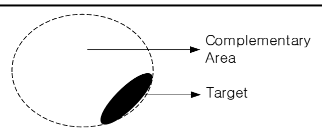
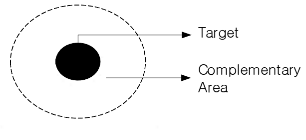
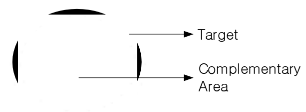
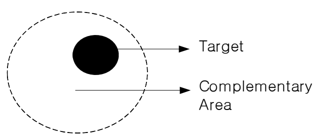
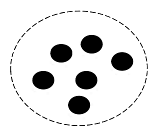
Fig. 5. Pattern matching on matrix data.

In Phase 4, the highest matching score from Phase 3 between actual FBM pattern and user defined pattern is decided, and then the processed data will be classified.

The validity of Hybrid-Similarity demonstrated through tests as shown in Table 1. For tests, 5 typical FBM patterns were selected and 250 field data were used. Cut-off values for effective deletions of non-patterned data need to be set for the validation of logic of Hybrid-Similarity. The validation tests were conducted with cut-off values set in the range of 0.1 to 0.9 referenced to 50 of FBM data of the center group. The results of tests are shown in Fig. 6. The cut-off values of 0.46 and 0.8 were chosen for the minimum and maximum cut-off values ($CutOff_{min}$ and $CutOff_{max}$), respectively, by selecting 80% as a criterion.

The matching ratio of the criterion, 80%, is higher than one of Cosine's, and the criterion can fully reflect the range of fluctuation of many patterns. Then a minimum threshold

TABLE I
CLASSIFIED 5 TYPICAL FAIL BIN MAP PATTERNS

Name	Pattern
Apple Group	
Center Group	
Edge Group	
Shift Center Group	
Atypical	

ISBN: 978-988-19253-1-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

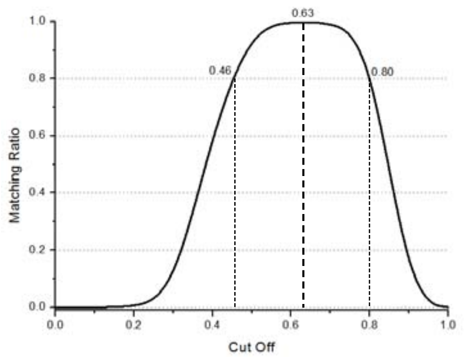


Fig. 6. Distribution chart between cut-off value and matching ratio.

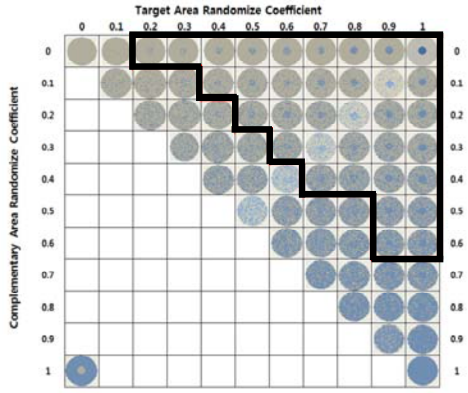


Fig. 7. Engineer's pattern matching boundary on FBM.

value which is a criterion for pattern was chosen with the cut-off values. FBM with the center group pattern was combined with noise and it was set as the Complementary Area Randomize Coefficient and Target Area Randomize Coefficient. Then, the matching result by Hybrid-Similarity method and the result of manual matching by an engineer were compared. The result of the comparison is shown as boundary values in Fig. 7. With the boundary values an average of Complementary Area Randomize Coefficient and Target Area Randomize Coefficient was calculated as 0.35. If the similarity is smaller than the average, a typical pattern is matched in Hybrid-Similarity. The total 250 of FBM patterns were tested using Cosine-Similarity, Jaccard-Similarity, and Hybrid-Similarity methods, respectively, with the final threshold values.

IV. Experimental Results

The test results to real FBM data using 3 types of similarity methods are shown as confusion matrix values in Table 2, Table 3, and Table 4.

The result by Hybrid-Similarity was 8.85% higher and 17.13% higher than ones of Cosine-Similarity and Jaccard-Similarity, respectively. The precision of Hybrid-Similarity was 13% higher and 41% higher than

TABLE 2
Confusion matrix for a Cosine-Similarity

		Predicted Pattern	
		True	False
Actual Pattern	True	17.6	16.4
	False	36.4	173.6

The overall accuracy and precision rate of Cosine-similarity is 78.36% and 51.76%.

Cosine-Similarity and Jaccard-Similarity, respectively.

As shown in table 5, execution time of each pattern matching for our suggested method is 250 times faster than the technique of Neural Network. This speed of pattern matching is acceptable in real mass production system.

TABLE 3
Confusion matrix for a Jaccard-Similarity

		Predicted Pattern	
		True	False
Actual Pattern	True	12.6	40.4
	False	34.4	162.6

The overall accuracy and precision rate of Jaccard-similarity is 70.08% and 23.77%.

TABLE 4
Confusion matrix for a Hybrid-Similarity

		Predicted Pattern	
		True	False
Actual Pattern	True	34.2	18.6
	False	13.4	184

The overall accuracy and precision rate of Hybrid-similarity is 87.21% and 64.77%.

TABLE 5
Execution time of hybrid-similarity and neural network

	Neural Network	Hybrid-Similarity
Execution Time per each Matching	1,009.6 ms [4]	4.05 ms

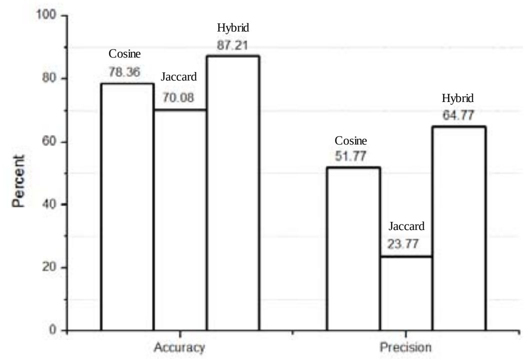


Fig. 8. Accuracy and precision chart each of Cosine, Jaccard and hybrid similarity.

V. Conclusion

There are many kind of pattern matching methods for fail bin map pattern. However, these methods are too complicated or require long calculation times to be applied in the field. In this paper, a simple Hybrid-Similarity method with low pass filter and cut-off technique is suggested. The accuracy of Hybrid-Similarity method was 87.21% which is 8.85% and 17.57% higher than Cosine-Similarity and Jaccard-Similarity, as shown in Fig. 8. This accuracy is good enough to be applied in the mass production field. Utilizing the, Hybrid-Similarity method we could save much time in detecting a cause of defects.

References

[1] Megan Langer, "Intel Developer Forum Manufacturing Keynote Disclosures", http://download.intel.com/pressroom/kits/events/idffall_2009/pdfs/IDFDay1_Baker_FactSheet.pdf, IDF2009, September, 2009.
[2] F. Palma, G.Nicolas and G.Miraglia, E.Pasquinetti and F.Piccinini, "Unsupervised spatial pattern classification of electrical-wafer-sorting

maps in semiconductor manufacturing", Pattern Recognition Letters 26, 1857-1865, 2005.

[3] T.Li and C.Huang, "Defect spatial pattern recognition using a hybrid SOM-SVM approach in semiconductor manufacturing", Expert System with Applications 36, 374-385, 2009.

[4] S. Hsu and C. Chien, "Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing", International Journal of Production Economics 107, 88-109, 2007.

[5] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.

[6] Kaiyang Liao, Guizhong Liu and Youshi Hui, "An improvement to the SIFT descriptor for image representation and matching", Pattern Recognition Letters 34, pp. 1211-1220, 2013.

[7] Gregory A. Baxes, "Digital Image Processing", John Wiley & Sons, Inc, 88-91, 1994.

[8] Liu, C.-W., Chien, C.-F., "An intelligent system for wafer bin map defect diagnosis", An empirical study for semiconductor manufacturing. Eng. Appl. Artif. Intel.(2012), http://dx.doi.org/10.1016/j.engappai.2012.11.009.

[9] Roberto J. Bayardo, Yiming Ma and Ramakrishnan Srikant, "Scaling Up All Pairs Similarity Search", ACM 978-1-59593-654-7/07/005, 2007.