Robust Estimation of Mixed-Type Wafer Map Similarity Utilizing Non-negative Matrix Factorization

Yukako Tanaka and Sho Saiki
{yukako.tanaka, sho.saiki}@toshiba.co.jp
TOSHIBA Memory Corporation
yukako.tanaka@toshiba.co.jp
Phone: +81 -59-330-1028 Fax: +81-59-330-1124

In this work, we propose a robust estimation method which calculates similarity between wafer maps. We focus on comparison of mixed-type wafer maps, that is, maps obtained in different inspections or quality tests in manufacturing processes, such as defect and failure bin maps. This work provides a novel approach to calculating similarity not with vectors of maps, but with 'activations' of maps for all observed wafer maps. The proposed method enables robust calculation of similarity between mixed-type maps.

Competitive strength in semiconductor manufacturing depends on yield analysis with big data. Among various informative data in a modern fabrication, wafer maps, which represent spatial patterns of failure, defect, or physical characteristics on wafers, are essential to yield analysis. Wafer maps often contain crucial information about how the failures or defects are caused, and detailed analysis of a large amount of maps leads to rapid yield enhancement.

Owing to the development of devices, fine maps are obtained in multiple manufacturing processes as shown in Figure 1. The maps represents various features on wafers in various resolutions, with various types of values. It is difficult even for skilled engineers to check a large amount of maps manually.

To support the comparison of maps, retrieval and classification using machine learning are known to be efficient (e.g. [1]). But sometimes such methods do not work especially when mixed-type maps are compared. Figure 2 illustrates the difficulty in comparing maps. The three maps obtained in different steps might be visually similar for engineers, but numerical vectors of the maps are different due to observational factors such as sensitivities and sampling rates. Comparison across mixed-type maps is one of difficult problems in yield analysis.

To address this problem, we propose a robust similarity estimation method utilizing Non-negative Matrix Factorization (NMF, [2]). Figure 3 represents an illustration of NMF. In [2], NMF is known to learn some basis images from all images and the weights which combine basis image to reproduce original images. Intuitively, when adapted to wafer maps, NMF provides hidden states (dictionaries) which represents the individual characteristic maps, and expresses each observed map by weighted sum of dictionaries (activation matrix) as shown in Figure 4.

We confirmed NMF works well for wafer maps by experiments with simulated failure maps. Figure 5 shows the results. We generated 500 failure bin maps for each class of Edge, Stripe, Center, and their combinations. Maps are factorized into three dictionaries and the activation matrix. In Figure 5, we can confirm dictionaries of NMF learn representative simulated maps, and the dictionaries are activated for corresponding 3000 simulated maps, as the top dictionary is rightly activated for wafers of 'Edge', 'Edge+Stripe' and 'Edge+Center'.

Our method fully utilizes the activation matrix calculated by NMF. Figure 6 shows the overview of the method. In Figure 6, instead of comparing the similarity between maps, our method compares the activations matrix of maps. The point is that, if a crucial failure happens in a process, its effect continually appears in succeeding inspections or tests on same wafers, and we can find the relation between maps through the activation. As stated above, it is difficult to compare mixed-type maps directly, and our method bridges a gap between maps with activation, leads to robust similarity estimation.

We adapted the proposed method to artificial maps. Figure 7 shows the scatter plot of activation. On the left, Map A and Map B are simultaneously activated on the same wafers, which means Map A and B are similar or related. On the right, to the contrary, Map C seems unrelated to Map B. By using this method, we expect to extract important relations to our target, Map B in this case, from all the maps obtained in multiple manufacturing processes. We estimate the method reduce 99% of engineer's laboring time for search of similar maps. It is remarkable that our activation-bridged estimation method is robust for observational factors of maps, and needs less parameters preset. Our activation-bridged similarity estimation will contribute to the comprehension of whole map relations, is promising for leading to rapid yield enhancement.

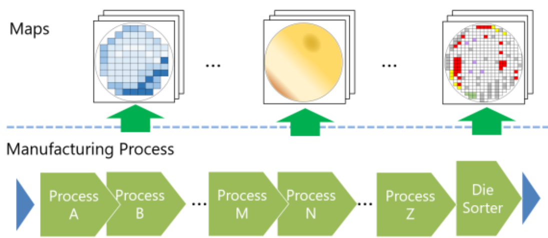


Fig. 1. Examples of Various Maps Obtained in Multiple Inspection or Quality Test Processes.

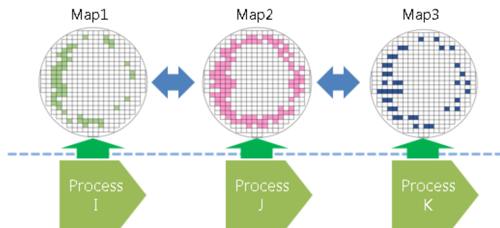


Fig. 2. An Illustration of Difficulty in Comparing Mixed-Type Maps.

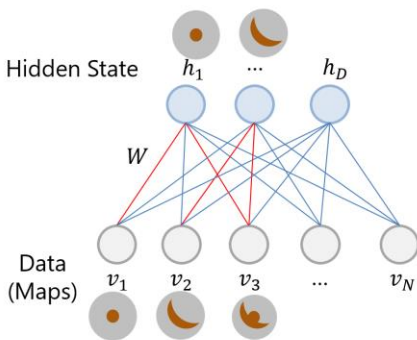


Fig. 3. An Illustration of Non-negative Matrix Factorization (NMF) adapted to Wafer Maps. Intuitively, Hidden states (dictionaries) represents the individual characteristic maps, and each data (an observed map) is expressed by weighted sum of dictionaries.

Number of Wafers (N)

Number of Chips (C)

$$\begin{pmatrix} 0\,1\,0 & \cdots & 0 \\ 1\,0\,0 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1\,0\,1 & \cdots & 1 \\ 0\,0\,1 & \cdots & 0 \end{pmatrix} \approx C \begin{pmatrix} w_{11} & \cdots & w_{1D} \\ \vdots & \ddots & \vdots \\ w_{C1} & \cdots & w_{CD} \end{pmatrix} \times D \begin{pmatrix} h_{11} & \cdots & h_{1N} \\ \vdots & \ddots & \vdots \\ h_{D1} & \cdots & h_{DN} \end{pmatrix}$$

Number of Dictionaries (D) N

Failure Maps Dictionaries Activation Matrix

Fig. 4. An Example of Factorization of Failure Maps by NMF. Factors of activation matrix w_{ij} represents the weight, at which observed map j is similar to the dictionary i.

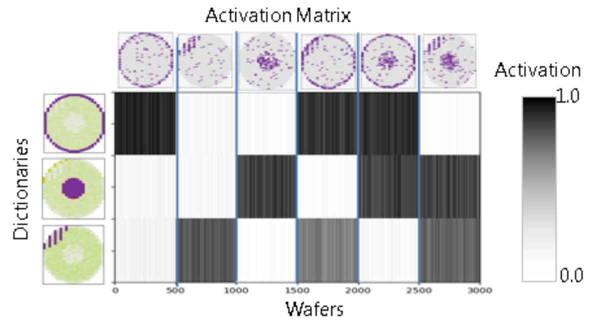
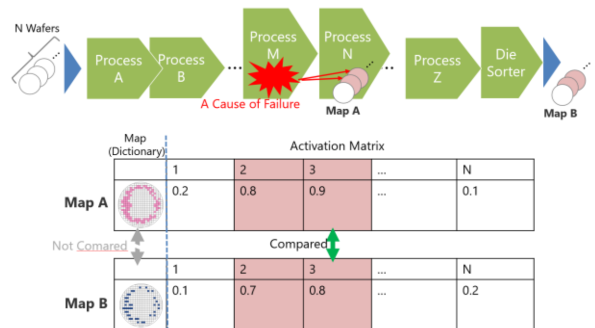


Fig. 5. Experimental Results of Factorization of Simulated Failure Maps.



	1	2	3	...	N
Map A	0.2	0.8	0.9	...	0.1
Map B	0.1	0.7	0.8	...	0.2

Fig. 6. The Overview of the Proposed Method.

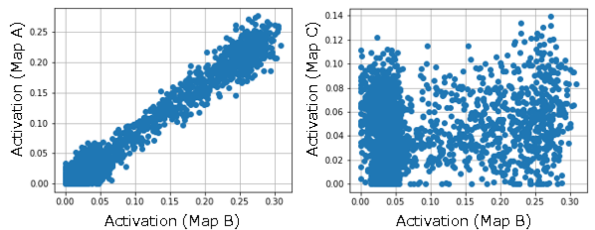


Fig. 7. Comparison of Artificial Maps. Each point represents one wafer and axes indicate activation values.

Reference:

[1] Wu, Ming-Ju, Jyh-Shing R. Jang, and Jui-Long Chen. "Wafer map failure pattern recognition and similarity ranking for large-scale data sets." IEEE Transactions on Semiconductor Manufacturing 28.1 (2014): 1-12.

[2] Lee, Daniel D., and H. Sebastian Seung. "Learning the parts of objects by non-negative matrix factorization." Nature 401.6755 (1999): 788.