# SURF BASED FAULT IMAGE DETECTION FOR PRINTED CIRCUIT BOARD INSPECTION

**[1]EUN-HYE YUK, [2]SEONGWON JANG, [3]SEUNGHWAN PARK, [4]CHEONG-SOOL PARK, [5]JUN-GEOLBAEK**

[1,2,3,4,5]Department of Industrial Management Engineering, Korea University
E-mail: [1]eunhyeyuk@korea.ac.kr, [2]jsw900806@korea.ac.kr, [3]udongpang@korea.ac.kr, [4]dumm97@korea.ac.kr, [5]jungeol@korea.ac.kr

**Abstract—** It is possible to collect, manage and analyze large amounts of data in real time because of development of fusion technology and spread of Internet of Things(IoT) in modern society. With the advanced technology, the high-tech product manufacturers began to offer differentiated services by producing fewer customized products. According to this, PCB (Printed Circuit Board, PCB), key component of the digital products, is also produced in small quantity batch production. Current PCB inspection systems require information about the normal image because it detects faults by comparison with the non-defective image. Therefore, this means that the test is not possible without the normal image and needs to have the all reference image. It is a major cause of reducing the efficiency of the PCB inspection system. This paper proposes a method for detecting the fault in PCB without normal image by learning the pattern of abnormal image. As a result of this methodology, it is expected to check more effectively the defects in the system to produce a variety of products and bring the time and cost savings in PCB inspection.

**Keywords—** PCB Inspection, Speeded Up Robust Feature, Random Forest, Kernel Density Estimation.

## I. INTRODUCTION

While IoT(the Internet of Things, IoT) era begins with the rapid development of semiconductor industry and communication technology, high-tech products such as smart phones and wearable devices have been widely spreading in Our lives. Thus, the high-tech product manufacturers are focusing on customized production for consumers by using the IoT technology. These products are composed with PCB, semiconductor parts and display parts, etc.

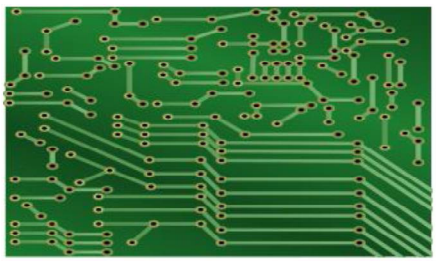


**Fig.1. An example of PCB Image**

In particular, the PCB is a thin plate made by printing a conductive circuit having an electrically conductive on an insulator and the form of PCB is like fig.1.It is one of the important key components due to connect the part electrically by converting the analog signal into a digital signal. Since PCB is typically produced in order production system, the importance of small quantity batch production is being emphasized in the PCB manufacturers in accordance with the high-tech production system. Generally the circuit board manufacturing consists of many steps: Cutting, Inner Layer Etching, AOI(Automatic Optical Inspection, AOI), Lay-up, Lamination, Etching, Drilling, Solder mask, Routing, BBT(Bare Board Test, BBT), Quality Control, Packing and Shipping in order. Faults generating in the PCB manufacturing process are various and typical faults are non-etching, over-etching, open circuit and scratches etc. Especially, Scratch defect is one of the serious defects because it can change the electrical properties of the product and preclude operation of finished product. PCB manufacturers previously detect faults by inspecting before main process to maximize production efficiency. The method of PCB inspection is divided into a method utilizing reference image and method not using. There are image subtraction and template matching etc in methods utilizing reference image and morphological techniques, RLE(Run-Length Encoding) in method not using. Image subtraction operation compares the reference image and the inspected image pixel-by-pixel using XOR logic operator. The result gets the image of part only defective (Chauhan et al, 2011).Therefore, image subtraction operation is positively necessary a same size of reference image and has the risk positioning problems. Template matching method detects parts that compare and match the feature points extracted from the reference and inspected image. This method requires high-quantity storage that can store all information about reference image and has to be provided exact information for compare the two images. This method requires a mass storage device that can store all the information about the reference image and it has to be provided to exact information to compare the two images (Moganti et al, 1996). Morphological techniques process and analyze a type of image. It is commonly used in PCB inspection but has disadvantages that should be applied in different pre-processing algorithm for each defect type (Ye et al, 1988).

Currently, most PCB manufactures inspect PCB through AOI. The AOI is non-contact inspection

system that detects the faults such as non-etching, surface defects and circuit open etc. It utilizes the image comparison operation using the reference image and the grayscale comparison methods using the GRAY-LEVEL. Image comparison operation compares inspected image with reference image and confirms a similar degree after setting a non-defective image in the reference image. The reference image is necessary because the form of PCB is slightly different depending on the supplier. Thus it means that the PCB faults detection is difficult without reference image. In addition, the performance of the AOI is affected by the location or rotation angle of PCB. It can be formalized by drawing guidelines on the conveyor belt or installing fixed frame to the products, but there is no guarantee that a fully standardized. So the research on robust system that can detect PCB faults in rotated state is necessary (Hwang et al, 2012).

This paper proposes a methodology that can detect faults without being affected by environmental changes such as size, rotation and location of PCB. The proposed methodology first extracts robust features utilizing image processing technique and detects faults using efficient classification technique in the high dimensional data. And then draws the weighted kernel density estimation map on the inspected PCB image to the predicted probability value and compares actual fault position with predicted faults section to evaluate the performance of the method.

The paper is organized as follows. In section 2, the theory applied to proposed methodology is presented and the inspection methodology based on SURF and Random Forest is offered. Section 3 describes the result of faults detection using the real PCB images and measures the performance of suggested methods. Finally, section 4 shows the conclusions and direction of further research.

## II. EXPERIMENT DESIGN

### *2.1. Speeded Up Robust Feature*

SURF proposed by Bay and Herbert is one of the image processing techniques. It detects robust features and creates distinctive descriptors on the noise. In addition, it reduces the computation time more than SIFT (Scale Invariant Feature Transform) algorithm (Bay et al, 2006). SURF relies on integral images to improve the detection speed and it is called 'Fast-Hessian' detector. Thus SURF extracts the robust feature and descriptor exploiting Fast-Hessian detector.

The SURF algorithm is mostly composed of three steps. Interest points on the image are explored using integral images in the first step. And then find the features using second differentiation filter approximated in the generated integral images. Finally, it extracts the descriptor by calculating the intensity and direction component of the local interest points.

### *2.2. Random Forests*

Random Forests is a classification method proposed by Breiman and it is a kind of ensemble learning method by constructing a multitude of decision trees. It combines the predictions about the multiple decision trees. Classification tree is built by applying the bootstrap method. When splits on each node in the process of growing the individual tree, it investigates for the optimal properties aimed at the only subset, which is selected randomly by predetermined number without considering the specified entire set(Ohn et al, 2013). Eventually each tree gives a classification and the forest chooses the classification having the most. Random Forests grow many classification trees without pruning, which means that can obtain low variation tree.

Accordingly the Random Forests can generate relatively accurate classifier at fast rate in the high-dimensional data and even prediction performance is good at noisy data.

### *2.3. Weighted Kernel Density Estimation (WKDE)*

In the results of Random Forest, we can get the probability to be classified of each feature points extracted from the inspected image into false class. Exploiting the value, Draw a graph of WKDE (Weighted Kernel Density Estimation, WKDE) and predict a position of faults.KDE is one of the density estimation methods using a kernel function and is a non-parametric way even applicable in high dimensional data (Fotheringham et al, 2000).The kernel function is symmetric at the origin and its integral value is non-negative. The representative kernel function is Gaussian and Uniforms etc. The kernel density estimate values are calculated like (1). Generate the kernel function centered on each of data at first. Then divide by the total number of data after adding all the kernel functions.

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right) \quad (1)$$

To do this, the continuous probability density function is obtained by the adding kernel function value of respective data (Anderson et al, 2009).

$$\hat{f}(x) = \frac{1}{h}\sum_{i=1}^{n} w_i K\left(\frac{x-x_i}{h}\right) \quad (2)$$

As our object is to detect faults area, we utilize the probability to be classified each of feature into False class. In other words, the predicted probability is given as a weight to the (2). Then the red area in graph of weighted kernel density estimation function is the predicted regions as false class by the detection model in this paper.

Finally draw the graph of WKDE on the actual PCB image and confirm the performance of inspection model by checking whether the red zone matches the real-defective area.

### *2.3. Fault Detection Model Procedure*

This experiment use the data generated by the inspection step of the PCB manufacturing. Especially the data is occurred by dust or scratch before plating. This type of defect appears in the scratch form or the stain shape such as Fig.2, Fig.3. It is one of the major types of faults that influences on the quality of the PCB.

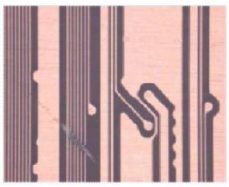


**Fig.2. Scratch Form**

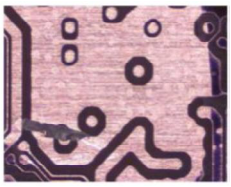


**Fig.3. Stain Shape**

The process of this methodology is like Fig.4 as follows: First, extract the features of PCB image using SURF. Second, learn the faults pattern utilizing Random Forest. Finally, draw a WKDEusing the probability that each feature is classified as true or false. Then we can predict the defective areas through red color region.

The experiment is carried out using a variety of image processing techniques to select the optimal detection model. Among the various techniques, SURF, SIFT (Scale Invariant Feature Transform), ORB (Orientation by Intensity Centroid) technique is used. And they are well known for extracting the robust features by environmental factors. It is compared with the identical PCB image that the detection performance of the model by implemented by the respective techniques (SURF, SIFT, ORB). As a result, the performance of the model using SIFT is not different with SURF. But SIFT is not able to extract the features of the small sized defective portion. And ORB only extracts the feature points of the curve or circle part rather than straight line of the circuit. Therefore, since the SURF gives strong performance regardless of the faults size or shape of the circuit, the SURF is adopted in the feature extraction method of inspection model. Because the Random Forest can create relatively fast and accurate classifier in the high-dimensional data and even shows a good predictive performance in noisy data, this research uses the Random Forest as classification model.

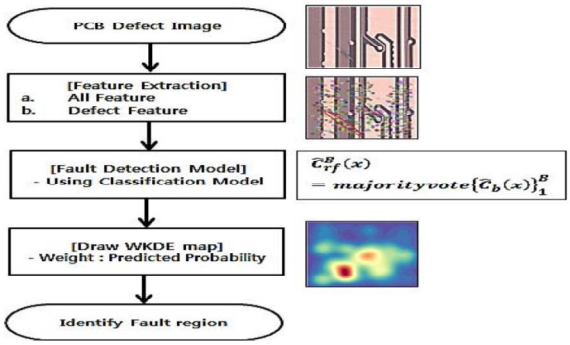



**Fig.4. Flow Chart of proposed PCB inspection model**

## III. EXPERIMENTAL RESULTS

The experiment is done with the defective PCB images caused by dust and scratch before plating. The experiment proceeds 10 times with the 10 PCB images. The train data sets 9 of 10, and the remaining one in the test.

When extract the features of Fig.2 using the SURF, the descriptors are extracted, it represents the coordinate and robust degree of keypoints, local direction component and the intensity. In this paper, we used descriptors of extended dimension 128 to obtain more detailed descriptors. In addition, the lower threshold and the larger octave, the more features are produced by the detector. But there is a possibility that noise is included. A good default threshold value could be from 300 to 500 and octave is 5.In this experiment, to find the feature points in the small size of the faults, the threshold sets to 150 and octave is 5. The result of extracted feature point is shown in Fig.5

Following this, set the defective region in the ROI (Region of Interest, ROI) to re-extract the feature points and the results is like Fig.6False class is assigned to the features in the ROI and the rest were designated as true class.

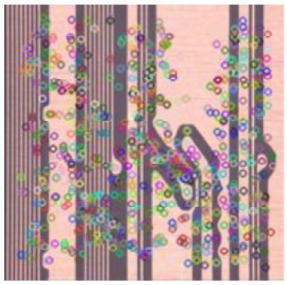


**Fig.5. Result of extracting the feature of Defect ①**

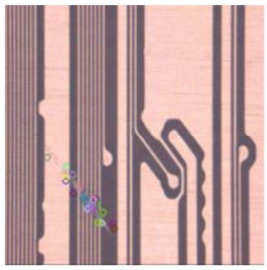


**Fig.6. Result of extracting the feature of only Defect ① ROI**

Before learning of random forests, visualize the distribution of features belonging to false class using t-SNE (t-distributed Stochastic Neighbor Embedding, t-SNE).PCA(Principal Component Analysis, PCA) is a representative technique of data dimension reduction. PCA, however, has the disadvantage that it is inappropriate to the non-linear data. On the other hand, t-SNE is superior in non-linear data, and also shows strong performance in comparison with other techniques (Van der Maaten et al, 2008). So it is confirmed that the distribution of defective features

of each image is shown as Fig.7 using t-SNE. A red dot is the features of small black spots in Fig.17 and Fig.18. Black dot represents the features of large black spots likeFig.19, and green dot is the features of the scratch form as shown in the other images.
Although the engineers designate with the same kind of faults, this actually can be assumed that there are faults with different properties. Therefore, faults with different properties have the potential to not be properly learned.

**Table 1: Probability of the defective features to being classified in F/T class**

| | Predicted Probability in F-class | Predicted Probability in T-class |
|---|---|---|
| 1 | **0.060000** | 0.940000 |
| 2 | **0.060000** | 0.940000 |
| 3 | **0.060000** | 0.940000 |
| 4 | **0.240000** | 0.760000 |
| 5 | **0.046667** | 0.953333 |
| 6 | **0.233333** | 0.766667 |
| 7 | **0.200000** | 0.800000 |
| 8 | **0.246667** | 0.753333 |

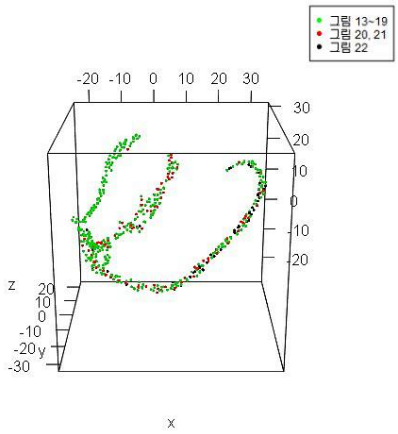



**Fig.7. Visualization of defective data using t-SNE**

And then, Calculate the probability which the features of inspected image to be in false class by learning the Random Forests.Table 1 is the probability of the features to be classified as False or True class. At this time, the WKDE is weighted by probability to be classified as false class and the graph is drawn as Fig.8

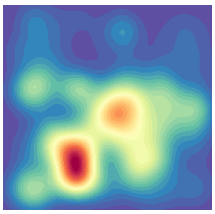


**Fig.8. WKDE Map of Defect ①**

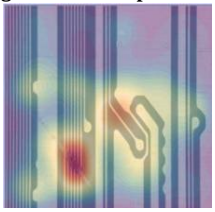


**Fig.9. Overlapped WKDE map on Defect ① image**

The Fig.8 is overlapped on top of the original image(Fig.2) to validate the performance of the proposed detection model. The result is shown as Fig.9 and the red area corresponded with the actual defect section.Other PCB images are experimented in the same manner as above. As a result, red section expected is consistent with the actual region of fault in the scratch type as 7 pictures (Fig.10, Fig.11, Fig.12, Fig.13, Fig.14, and Fig.15, Fig.16). However, the faults patterns of black spots don't match the expected position such as Fig.17, Fig.18, and Fig.19. This, as previously confirmed, means that there are faults with different properties even though the faults are same type. For this reason, it can be seen that the detection performance of Fig.17, Fig.18 and Fig.19 are not good.

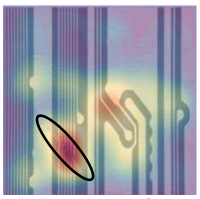


**Fig.10.Defect ①**

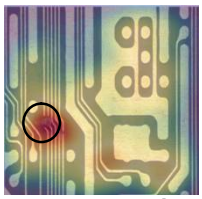


**Fig.11.Defect ②**

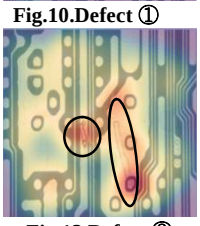


**Fig.12.Defect ③**

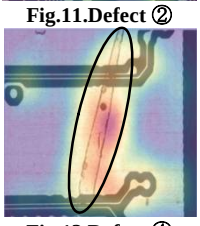


**Fig.13.Defect ④**

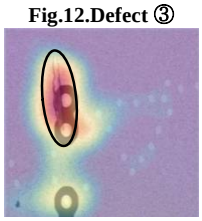


**Fig.14.Defect ⑤**

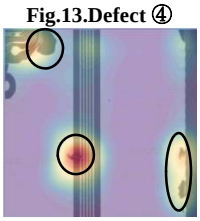


**Fig.15.Defect ⑥**

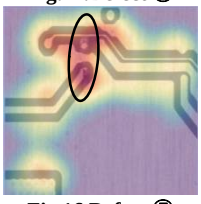


**Fig.16.Defect ⑦**

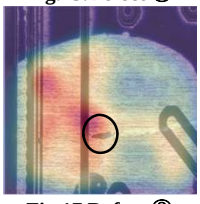


**Fig.17.Defect ⑧**

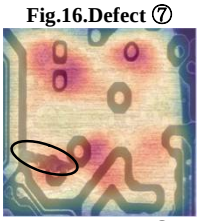


**Fig.18.Defect ⑨**



**Fig.19.Defect ⑩**

## CONCLUSIONS

The research on the PCB defect detection has been actively conducted. A new kind of PCB, however, has a limit that inspection is impossible without the non-defective image because most studies require reference images. Therefore, this paper presents the

PCB faults detecting method. And it does not affected by environmental factors and only utilizes the information about the defect type not reference image.In the future, we will apply the proposed inspection model with other PCB image having different defect pattern in addition to scratch fault. We also will try to cluster of all the features belonging to the defect class regardless of the type of fault. Then re-define the class based on the results and apply to the methodology.

## ACKNOWLEDGMENTS

This work was supported by the BK21 Plus (Big Data in Manufacturing and Logistics Systems, Korea University).
This work was supported by the Samsung electronics.

## REFERENCES

[1] Chauhan, Ajay Pal Singh, and Sharat Chandra Bhardwaj, "Detection of bare PCB defects by image subtraction method using machine vision", Proceedings of the World Congress on Engineering. Vol. 2. 2011.
[2] Moganti, Madhav, et al, "Automatic PCB inspection algorithms: a survey", Computer vision and image understanding 63.2 (1996): 287-313.
[3] Ye, Qin-Zhong, and Per E. Danielsson, "Inspection of printed circuit boards by connectivity preserving shrinking", Pattern Analysis and Machine Intelligence, IEEE Transactions on 10.5 (1988): 737-742.
[4] Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool, "Surf: Speeded up robust features", Computer vision–ECCV 2006. Springer Berlin Heidelberg, 2006. 404-417.
[5] Hwang D-D., Shin S-W., and Lee G-S.(2012), "Method of PCB Short Circuit Detection using SURF", Journal of the Korea Academia-Industrial cooperation Society, 5471-5478
[6] Bay, Herbert, et al, "Speeded-up robust features (SURF)", Computer vision and image understanding 110.3 (2008): 346-359.
[7] Shin J-S., and Kang D-S.(2012), "A Study on Moving Object Tracking Algorithm Using Depth Information and SURF Algorithm", Journal of Advanced Information Technology and Convergence, 144-148
[8] Breiman, Leo, "Random forests", Machine learning 45.1 (2001): 5-32.
[9] Ohn S-Y., Ji S-D., and Han M-Y.(2013), "Feature Selection for Classification of Mass Spectrometric Proteomic Data Using Random Forest", Journal of The Korea Society for Simulation, 22(4), 139-147
[10] Lindeberg, Tony, "Scale invariant feature transform", Scholarpedia 7.5 (2012): 10491.
[11] Prasad, Anantha M., Louis R. Iverson, and Andy Liaw, "Newer classification and regression tree techniques: bagging and random forests for ecological prediction", Ecosystems 9.2 (2006): 181-199.
[12] Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton, "Quantitative geography: perspectives on spatial data analysis", Sage, 2000.
[13] Anderson, Tessa K, "Kernel density estimation and K-means clustering to profile road accident hotspots", Accident Analysis &Prevention 41.3 (2009): 359-364.
[14] Van der Maaten, Laurens, and Geoffrey Hinton, "Visualizing data using t-SNE", Journal of Machine Learning Research 9.2579-2605 (2008): 85.

★★★