PalArch's Journal of Archaeology of Egypt / Egyptology

DATA-DRIVEN PREDICTION MODEL OF COMPONENTS CHANGE IN SURFACE MOUNT TECHNOLOGY DURING REFLOW CYCLE

Hara Prasad Tripathy[1], Priyabrata Pattanaik[2], Sushanta Kumar Kamilla[3]

[1,2,3] Department of Management, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha

Email: [1]haraprasadtripathy@soa.ac.in, [2]priyabratapattanaik@soa.ac.in, [3]sushantakamilla@soa.ac.in

Hara Prasad Tripathy, Priyabrata Pattanaik, Sushanta Kumar Kamilla: Data-Driven Prediction Model Of Components Change In Surface Mount Technology During Reflow Cycle -- Palarch's Journal Of Archaeology Of Egypt/Egyptology 17(6). ISSN 1567-214x

Keywords: Electronic Packaging, Surface Mount Technology, Passive Chip Components Self-Alignment, Machine Learning Prediction Model, Support Vector Regression, Neural Network, Random Forest Regression.

ABSTRACT

In surface mount technology (SMT), installed components are subject to motion during the reflow cycle on soldered pads. This capacity is known as self-alignment, and is the product of molten solder paste's fluid dynamic behaviour. This capability is crucial in SMT, since inaccurate self-alignment causes defects such as overhanging, tombstoning, etc., whereas on the other side it can allow components to be perfectly self-assembled on or near the desire location. The goal of this study is to develop a machine learning model that predicts movement of the components in xx and yy-directions as well as rotation during reflow. The analysis consists of two steps: (1) experimental data are analysed to reveal the relationships between self-alignment and various variables, including component geometry, pad geometry, etc. (2) advanced machine learning prediction models are used to predict the distance and direction of components shifting using support vector regression (SVR), neural network (NN), and random regressi forest. As a result, RFR can predict components shifting with an average fitness of 99%, 99%, and 96% and an average prediction error of 13.47 (μmm), 12.02 (μmm), and 1.52 (deg.) for component shifts in xx, yy, and rotational directions respectively. This enhancement provides the future capability in the pick-and-place system to refine the parameters to monitor the best placement position and reduce the intrinsic defects caused by the self-alignment.

1. Introduction

Surface mount technology (SMT) is an improved electronic packaging system in which surface mounting components (SMCs) are mounted directly on the printed circuit boards (PCBs). In general, the surface mount assembly (SMA) line comprises three main operations such as stencil printing process (SPP), pick and position process (P&P), and reflow soldering process. At first point, solder paste is printed on PCB pads to create electrical connections, then picking and putting SMCs on their respective pads. Then, reflow oven is used by forming permanent solder joints to connect SMCs. Basically, PCB moves through many thermal areas, known as thermal profile, to form the solder joint. To start soaking, PCB is preheated first. Then, the temperature to melt deposited solder paste is increased by given peak point [1]–[3]. Finally, solder joint is created by lowering temperature in the refrigeration region. The amount of temperature and time PCB spends in each zone is predefined based on factors such as type of solder paste used, PCB material and thickness, number and SMC size, etc. In addition, solder paste's fluid dynamic behaviour causes SMCs to switch during the reflow cycle until the solder paste starts to melt. Mainly force arising from the surface tension allows components to travel in a direction that achieves the most magnificent symmetry with respect to its relative location with soldered pad. This capability is known as self-aligning SMCs, self-assembled, etc. There are no guarantees, however, that self-alignment always balances misplacing on the P&P cycle, and it can be used to optimize parameters for placement machining.

Despite numerous studies investigating the theoretical fundamentals of movement of components during reflow, there is no clear rationale in the real situation for addressing the practical challenges of self-alignment. At the one hand, there is no clear rule mentioned in the literature recommending which control variables contribute significantly to the movement of components in xx, yy-directions as well as rotation [4], [5]. On the other hand, the commonly used approaches are computational models followed by experimental findings along with simulation and numerical models. As for the benefit of data-driven techniques compared to traditional statistical methods, there are a few data-driven models for predicting movement of components as well as positioning of components after reflow cycle.

This study presented a one-layer neural network to predict the location of components after reflow with respect to their location before reflow in x, y and rotational directions, while neglecting the value of solder paste properties such as length, offset etc. [5]–[7]. It therefore enhances the necessities of: (1) gaining insight into the most important contributing factors in the shift of components, (2) outlining the machine learning algorithm that predicts movement of components during reflow with respect to those factors. In addition, a generalized prediction model addressing different type and size of movement of passive chip components may be a breakthrough in this field as all previous studies considered only one or two component types.

48 variables are selected to train the prediction models from the comprehensive work conducted to develop this work and domain information (see Figure 1). Such contributing factors are listed as components of geometry, pad geometry, inspection of solder paste, and inspection of location, relative factors of paste-pad, relative factors of location-post, and relative factors of placement-pad shown in Fig. 1. In addition, 3 targets are identified as components that are predicted to move in x, y direction, and rotation using a remarkable machine learning algorithm, including support for vector regression (SVR), deep neural network (NN), and random forest regression (RFR) [8]–[11]. The remainder of this paper is structured as follows: The paper discusses the literature relating to the self-alignment of components; the constructed prediction models are discussed followed by findings; the conclusions and future work of this research are also presented.

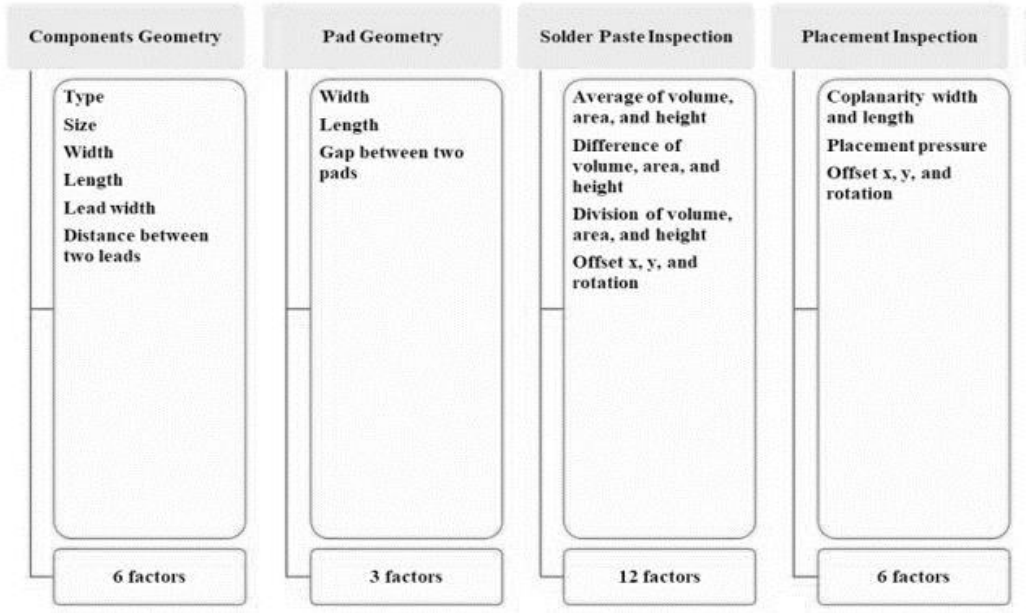

Figure 1. Contributing Factors in Components Movement during Reflow Process.

2. Literature Review

Many researchers studied the ability of components to self-align during the reflow cycle. The key focus of these researches is the study of experimental findings along with simulation and numerical estimation based on complex definition of fluid. Specific studies, however, discuss the theoretical aspect behind component shifts during reflow, with no significant data-driven analysis in that regard. During the reflow method, Ellis and Masada provided a dynamic model in respect of chip condenser behaviour. They studied the effects of different factors (for example, pad geometry, solder length, etc.). The model presented is based on computed forces acting on component and the moment of their actions. Liukkonen et al. studied the theoretical dimension of the self-alignment method, and compared the position of the components before and after reflow with experimental findings in both lead-free and tin-lead processes. However, they outlined that there was more self-alignment variance capability

in the lead-free process, but their experiment neglects other contributing factors other than the solder type. Kong et al. predicted the accuracy of soldered flip chip assemblies based on variations and distributions in chip mass, die tilt and solder volume. They applied a regression model which optimizes a chip's static equilibrium conditions and experimentally validates its outcome. They concluded that the variance in the solder volume had a greater effect on the chip standoff height (i.e., in the z-direction) than the lateral chip alignment accuracy (i.e. in x and y-direction). In addition, Dusek et al. examined self-alignment components with respect to three types of solders (one leaded and two lead-free) deposited with five different angle offsets and two different volumes that are reflowed with two different reflow technologies. Krammer investigated restore forces that are conceptually validated by a simulation model and experimental result acting on passive chip part (0603) during reflow.

3. Methodology

Within this section machine learning techniques has been addressed that are implemented after the reflow cycle to predict a change within passive chip components. A detailed overview of the experiment and collected data is provided at first, and contributing factors are added under this report. Finally, a brief summary is given of the applicable machine learning techniques and their rationale for this analysis.

1. Data and experiment description:

The experiment is designed to combine 6 types of passive chip components including 3 resistors and 3 condensers in 3 groups such as R1005 and C1005 (1 mm 0.5 mm), R0603 and C0603 (0.6 mm 0.3 mm) and R0402 and C0402 (0.4 mm 0.2 mm) respectively. The PCB is designed to position each of the above-named components on two corresponding pads (named pad1 and pad2), which are covered in SPP with a soldering paste. Illustration. 2 (a) describes the xx and y plane pads, soldering pastes, and components. The red dot inside the Fig. 2 (a) shows the middle of two pads that are called a reference point (0,0) in this analysis. Because all components are positioned horizontally, the component's longer and shorter sides are considered parallel to the xx and y directions, and rotation is characterized as a circular movement on the xx-axis (see Fig. 2(b)). Initially, lead-free solder paste is deposited on top of the pads, regarding the average volume of the paste, the difference in the volume of the paste, and the deliberate offset (x, y, rotation) as interest factors in SPP level. Then components are placed with mounter machine, considering intentional offset positioning pressure and placement (x, y, rotation) as factors of interest in the P&P stage. Lastly, all components are permanently attached to PCB via nitrogen reflow oven to minimize the effect of surface oxidation of PCB. In detail, 33 combinations of the above-mentioned interest factors are considered to position the 20 replications of 6 types of passive chip components.

During the reflow phase, the surface tension created by the liquid state of the soldering paste drags the component to the center of the pads where the component moves in xx and yy-direction and rotates around xx-axis. In

essence, the component shift is defined as the relative locational difference between the position of the component before and after the reflow stage. Notice that this sum of the change is considered as a predictive target for the research. Specifically, for the predictive modeling study all 48 variables into 7 classes were grouped as input features (see Figure 2). The first two of these 7 categories are obtained in respect of directory components and PCB architecture. It is worth noting that in this analysis certain specific component properties such as form and thickness of the component plating, weight, etc. are not considered. Paste testing and positioning testing is shown from solder paste inspection machine (SPI) and pre-reflow automated optical inspection system (Pre-AOI). The last three relative factors are determined on the basis of those categories listed above.

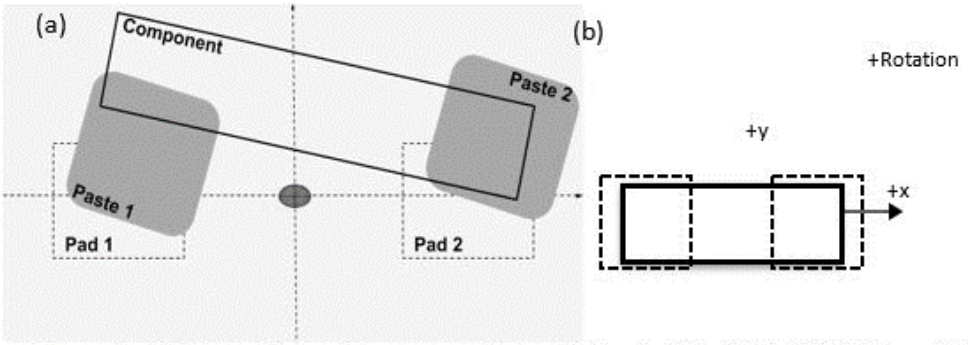

Figure 2. (A) Pads, Pastes Component Description in 2D; (B) A Definition of the Direction of Components.

2. Data pre-processing:

However, this experiment is designed to position 3940 components in 6 form categories, but some components are missing during the P&P and reflow soldering processes that could result from a given placement job file or floating components in reflow level. Since the cause of missing components is not within the scope of this analysis, these missing values are excluded in the research (i.e., components which are presently attached after reflow soldering were only considered). In addition, due to the large enough data size, outliers are omitted from the dataset, and the sample size is slowly decreased to 3917 records with 48 variables. In addition, component type is coded as component size for three size groups, and resistors and condensers are distinguished by categorical values -1 and 1 respectively to provide a generic prediction model for all types and sizes of passive chip components. The Spearman correlation coefficient is applied in terms of dimensionality reduction to resolve the nonlinear interaction of continuous and ordinal variables and to exclude variables that do not agree with prediction goals.

3. Support vector regression (SVR):

The general formulation of SVR and the concepts discussed in this analysis is based on literature review. For a given training set $\{(x_1, y_1), \ldots (x_n, y_n)\}$ in X

where X denotes the input function space, the SVR method fits hyperplane F(X) (Eq. (1)) where most data points fall on this plane.

$$F(x) = \langle w, x \rangle + b; \ \ where\ w \in X, b \in R \quad (1)$$

Wherever ..... is it? ,. Ê denotes the two vectors' dot product. In addition, the upper and lower boundary lines with the distance from the specified hyperplane are equipped in which the data points falling within this tolerance range are not penalized, while each point falling beyond this boundary will be penalized with the user-defined CC ratio with respect to the quantity of non-negative deviations between the two and the other. This restriction is known as the feature of loss insensitive to ÿ. The objective to solve the SVR estimation function is to minimize the hyperplane weight norm (w) with boundary line constraints given in Eq. (2) and (3) respectively;

$$\min \quad \frac{1}{2} \| w \|_2^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*) \quad (2)$$

$$s.t. \quad \begin{cases} y_i - \langle w, x_i \rangle - b & \leq \varepsilon + \xi_i \\ -y_i + \langle w, x_i \rangle + b & \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* & \geq 0 \end{cases} \quad (3)$$

The SVR's notable breakthrough is the use of support vectors, which make the model independent of XX input space dimensionality. It allows the algorithm to efficiently solve nonlinear problems. In this analysis, the function ÿ-insensitive loss is considered along with the linear kernel to address any possible linear relationship between the space and target values of the application. Basically, for the penalty ratio, the distance between boundaries from the hyperplane and the CC = 1 as $\pi = 0.1$ were set.

4. Deep neural network (NN):

For a given training set $\{(x_1, y_1), ..., (x_n, y_n)\}$ between X and R, independent variable X denotes the input space and dependent variable y denotes the output of a network that is linked to one or more hidden layers and corresponding hidden neurons through a graph shape path. Input neuron representing a feature carries an initialized weight that activates the summation of all input neurons and their corresponding weights on a hidden neuron with the mean of a predefined activation function. In addition, the NN will be fed sequentially and their output will be assessed with respect to a predefined loss function. The input weight and bias will be modified at each sequence by first derivation of the activation function (i.e. backpropagation) until the desire termination condition is met. NN's estimation function is specified in F(X, W) = f(W,X) + b, where b is a constant and f is the corresponding activation function shown as f(W,X) = f(n (wi ,xi). Normal initialized weight is considered to be i=1 for the proposed NN architecture that is enabled by the operation of the rectifier; Finally, the Adaptive Moment (Adam) Optimizer is applied to maximize the associated weight and training bias collection with respect to the statistical mean square error (MSE) and the real value of the y. Two hidden layers with the feature size and 100 hidden neurons respectively are proposed for this analysis.

5. Random forest regression (RFR):

Random forest is constructed on the basis of hierarchical tree-structured predictors that train each tree within the forest on a random split of features of input. Algorithm uses all input features along the top-to-dawn path from the root node of the tree to break nodes, and finally to leaves nodes. These features would be separated at each point until eventually, the best subset of features on terminal leaves would remain. In that respect; (1) both prediction and selection of features are considered simultaneously. (2) The dimensionality of the characteristic space shall not impact the predictive model. The general formulation of RFRs and the principles used in this analysis are based on literature review. For a given training set $\{(x_1, y_1), (x_n, y_n)\}$ in X, where X denotes the space of input characteristics, the RFR estimation function is: F(x) = 1 J fj(x), where j is the number of forest trees and fj(x) is the estimation function of each branch. For a J j=1 decision tree with MM splitting nodes, the space of the function will be split into MM regions as R_m where fj(x) = M m=1 b_m π(x, Rm); f_j = 1, J. A binary decision Ś(x, R_m) {0, 1} is made to decide whether xx is se input.

4. Results & Discussions

The determination coefficient (RR2) is used to calculate the fitness of the learning algorithm on a training set along with root-mean-squared error (RMSE) to determine the efficiency of prediction models on the test set. As stated in section 3.1, three continuous variables are identified as targets, called Shift XX, Shift YY, and Shift Rotation, in which models of predictions are trained separately on each target. RFR uses full mid-square error (MSE) training samples as a splitting consistency calculation. A check for all possible splits on each input variable decides the best split in RFR, and the decision tree is fully grown. Predictive model accuracy and learning algorithm fitness are evaluated through 10-fold cross validation to (1) ensure the model is well generalized with regard to any individual data set; (2) limit the overfitting problem.

Average and standard deviation of R2 and RMSE are considered simultaneously for comparing the performance of models. As far as training fitness (RR2) is concerned, all three models are well suited for Shift X and Y (see average and standard RR2 deviation for Shift XX and YY in Table 1), with above 90 percent average and zero standard RR2 deviation. NN is equipped with 90 percent average and 4 percent standard deviation RR2 when it comes to change rotation, while RFR outperforms with 96 percent average and zero standard exercise deviation. To conclude, RFR is better fitted for all three targets to the training dataset.

At the other hand, average and standard deviation of RMSE are illustrated for testing dataset to suggest overfitting of prediction models. Nonetheless, for all three target variables NN performed better in terms of average RMSE but its highest dispersion around average RMSE renders it unstable compared with other models. In addition, RFR offers better output with respect to average and standard RMSE variance of test dataset in which passive chip components are expected to change during reflow soldering whit 13.47 (μ), 12.02 (μ), and 1.52

(deg.) error in x, y, and rotational direction for this dataset. It is noteworthy that the motion of the components in xx and y-directions is responsive to the component scale. In addition to part size increases, the prediction error also increases in y -direction except for C0402 and R0402. In addition, the prediction error is smaller for condensers compared to their corresponding resistors except for C0603 and R0603 in x- direction, although prediction error is higher for condensers in y-direction. In respect to rotation, compounds rotate further together with the part size deduction. In addition, the shift rotation error estimation for condensers is smaller in comparison with their corresponding resistors except for C0402 and R0402. This may be attributed to the influence of body mass components on rotation of the components (Figure 3).

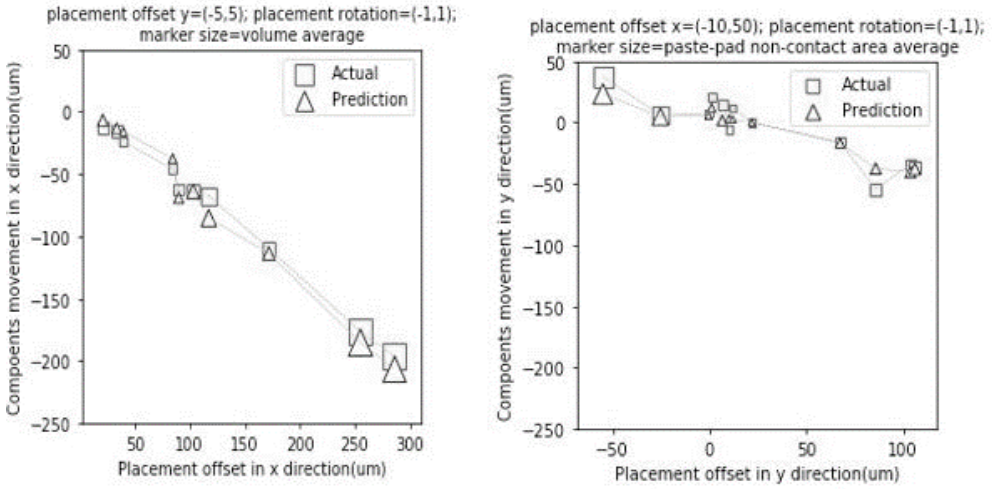

Figure 3. Prediction vs. Actual Value of Components Move In (a) X Direction; (b) Y Direction.

5. Conclusion

This study outlined 48 possible factors that lead to movement of components during reflow soldering and applied machine learning approaches such as SVR, NN, and RFR in order to predict how much and direction components shift in x, y, and rotational directions. Based on the analysis, RFR outlines other approaches in terms of average predictive losses and their standard deviations over 10-fold cross-validation. Indeed, it is worth mentioning that the RFR is well suited for all 3 goals with an average RR2 of 99 percent, 99 percent, and 96 percent with average prediction errors of 13.47(μ_m), 12.02 (μ_m), and 1.52 (deg.) for each X, Y, and rotation changes respectively. In addition, solder paste average length, paste-pad average non-contact area, and placement-paste average contact area are shown to be significant in the x, y and rotation part shifts respectively. This underlines the importance of solder paste status, relative offset of paste from pad, and relative offset of paste placement in explaining passive chip component self-alignment capability.

References

P. Lall, K. Patel, and V. Narayan, “Model for prediction of package-on-package warpage and the effect of process and material parameters,” in

Proceedings - Electronic Components and Technology Conference, 2013.

G. T. Ostrowicki, J. Williamson, V. Gupta, and S. P. Gurrum, "Thermal cycling reliability of lead free solder joints on multi-terminal passive components," in Proceedings - Electronic Components and Technology Conference, 2015.

I. Parviziomran, S. Cao, H. Yang, S. Park, and D. Won, "Data-Driven Prediction Model of Components Shift during Reflow Process in Surface Mount Technology," Procedia Manuf., 2019.

S. Cao, I. Parviziomran, H. Yang, S. Park, and D. Won, "Prediction of component shifts in pick and place process of surface mount technology using support vector regression," in Procedia Manufacturing, 2019.

565–582. doi:10.1016/j.jaridenv.2004.03.022 Wood, E.., Tappan, G.., Hadj, A., 2004. Understanding t Arid Environ. 59 et al., "Ecosystem Services Flows: Why Stakeholders' Power Relationships Matter," PLoS One, 2015.

T. (2010). M. in C. reservoirs. E. P. R. Icom/article/10.1007/s00027-008-8099-9 Watanabe, K., Monaghan, M., & Omura, T (2008). Longitudinal patterns of genetic diversity3820 Larssen et al., "RESTORING THE FISH FAUNA CONNECTIVITY OF THE HÂRTIBACIU RIVER-RETIŞ DAM STUDY CASE (TRANSYLVANIA, ROMANIA).," Acta Oecologica, 2017.

R. (200. R. from http://www.sciencedirect.com/science/article/pii/S0169555X12003819 Carley, J., Pasternack, G., Wyrick, J., & Barker, J. (2012). Significant decadal channel change 58–67years post-dam accounting for uncertainty in topographic change detection between et al., "RESTORING THE FISH FAUNA CONNECTIVITY OF THE HÂRTIBACIU RIVER-RETIŞ DAM STUDY CASE (TRANSYLVANIA, ROMANIA).," Acta Oecologica, 2017.

G. Chen, L. Liu, V. V. Silberschmidt, Y. C. Chan, C. Liu, and F. Wu, "Retained ratio of reinforcement in SAC305 composite solder joints: Effect of reinforcement type, processing and reflow cycle," Solder. Surf. Mt. Technol., 2016.

L. M. Lee, H. Haliman, and A. A. Mohamad, "Interfacial reaction of a Sn-3.0Ag-0.5Cu thin film during solder reflow," Solder. Surf. Mt. Technol., 2013.

S. Chung and J. B. Kwak, "Realistic warpage evaluation of printed board assembly during reflow process," Solder. Surf. Mt. Technol., 2015.

W. Xia, M. Xiao, Y. Chen, F. Wu, Z. Liu, and H. Fu, "Thermal warpage analysis of PBGA mounted on PCB during reflow process by FEM and experimental measurement," Solder. Surf. Mt. Technol., 2014.