Temperature Prediction of PCB Assembly in Reflow Process Using Encoder-decoder Networks

Yangyang Lai, Jun Kataoka, Ke Pan, Jonghwan Ha, Karthik A Deo, Jiefeng Xu, Seungbae Park

Department of Mechanical Engineering  
The State University of New York at Binghamton  
Binghamton, NY, USA  
ylai9@binghamton.edu

***Abstract—*** **Commercial reflow ovens contain multiple heated zones, which can be individually controlled for temperature. PCB assemblies travel through each zone at a controlled rate to achieve the desired reflow profile. The reflow profiles of all the components on the board are supposed to fall within a safe range. The minimum reflow temperature should be reached for the largest component. Meanwhile, the temperature cannot exceed the threshold temperature that may damage the smallest components.**  **CFD model is widely used to simulate the reflow soldering process. To hit the target reflow profiles, computational expense is required to seek the optimal boundary conditions, which are the preset temperatures of heating zones. The number of zones in the reflow oven determines the complexity of the combination of boundary conditions. To ease the computation cost, the encoder-decoder networks based on the CFD model are employed to predict the reflow profile of a bulky BGA package. The networks are demonstrated to rapidly predict transient temperatures of the BGA in seconds and provide an average error below 0.6%.**

***Index Terms—*Surface-mount technology, reflow soldering process, computational fluid dynamics (CFD), artificial neural network (ANN) CNN.**

# I. INTRODUCTION

With the increasing complexity of PCB assembly, the reflow profiling must consider the dimension of the board, the density of the components on the board, and the mix between large components and small ones. Temperature distribution should be studied for all new board designs using thermocouples at multiple locations on the components. If there are varied sizes of devices on the board, reflow profiles should be checked at various locations across the board. The minimum reflow temperature should be reached for the largest component. Meanwhile, the temperature cannot exceed the threshold temperature that may damage the smallest components. Hence, the solder temperature at the center of the biggest BGA package and the board surface temperature are regarded as the interests of the study.

Numerical methods, such as computational fluid dynamics (CFD) models are widely used to simulate the reflow soldering process. Shao employed the CFD model to predict the temperature distribution of the 2.5D package and mapped it to the FEA model to get the package warpage(Shao *et al.*, 2018). Lai used the CFD model to investigate the temperatures of solder balls of varied-sized components in the reflow soldering process(Lai *et al.*, 2021). Deng used the CFD model to simulate the temperature distribution of a system in package assembly in the reflow soldering process(Deng *et al.*, 2014).

In this work, the solder joint temperatures of the biggest component on the PCB were computed using the CFD model. The CFD model was validated with the results of field measurements. To hit the target reflow profile, computational expense is required to seek the optimal boundary conditions, which are the preset temperatures of heating zones. The number of zones in the reflow oven determines the complexity of the combination of boundary conditions. To ease the computation cost, the encoder-decoder networks based on the physics model are employed to predict the reflow profile of the bulky BGA package.

In the past, machine learning models have been widely used to optimize the reflow soldering process. Zhancheng employed an SVM model to investigate the real process temperature at any point of the board. Tsai developed a neuro-computing approach to study the second-side thermal profile of PCB assemblies. Lam used a non-contact temperature prediction of PCB during the reflow process. Li employed the neural network to predict the thermal profile of a PCB board, which achieved 97.2% accuracy.

The machine learning model used in this study is based on the CFD model. The hybrid method replacing repetitive parts of the physical model with efficient networks can provide reflow profiles subjected to varied boundary conditions rapidly. The heat transform problem is transformed to an image-to-image task. The inputs to the thermal problem are the dimensions of the parts inside the BGA and the temperatures of the reflow oven. Outputs are the temperature contours of the BGA. Inspired by the convolutional neural network in the digital image process, the value of each pixel feature is affected by all the other pixels in the receptive field at the previous convolution stage, with the largest contributions coming from pixels near the center of the receptive field. This matches with my study, where temperature contours for a pixel are most affected by the features in the same pixel, and partially by features in nearby pixels, with decreasing importance for those that are farther away. The transient temperature contours of the BGA can be regarded as three-dimension tensors. Physical information of the package including density, thermal conductivity, and specific heat can be fabricated into the matrixes with the same size of temperature contours. The in-plane dimensions can be defined by their corresponding positions in the board matrix. Oven temperatures are defined using the same shape matrixes. Encoder-decoder based generative (EDGe) networks can translate the input matrixes into the temperature contours rapidly, which is always time-costly for numerical methods. The performance of neural networks is evaluated using error analysis, which gives an average error below 0.6%.

II. MODELING DEVELOPMENT AND VALIDATION

*A. Field measurement*

As shown in Fig.1, thermocouples were embedded into the front and reverse sides of the PCB board to monitor the board temperature. The other thermocouples were embedded under the BGA of a bulky package. All the thermocouples were bonded to the PCB using thermal cement. The multi-channel thermal profiler was used to collect transient temperature data in the reflow process.

A picture containing text, electronics, circuit

Description automatically generated

Fig. 1. PCB with the BGA Package

A 7-zone reflow oven as shown in Fig. 2 is used in the field measurement. The board is transported through the seven heating zones and finally one cooling zone. The temperature of each zone was set as per Table I. The dwell time of each zone is determined by conveyor speed, which is 45 cm/min. In this study, the duration from room temperature to peak temperature is maintained at 240 seconds.

Icon

Description automatically generated

Fig. 2. Schematic diagram of the reflow oven

Table I. Temperature setting I

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Zone | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| T (°C) | 120 | 150 | 180 | 220 | 260 | 280 | 300 |

*B. CFD Model*

The structure of the BGA package is shown in Fig. 3. The dimension of every part of the BGA package is listed in Table II. Small features such as leads, thermal interface material, copper via, and underfill, are not included in the compact model. Because to use such detailed models in a board-level simulation is not practical.

Shape, rectangle

Description automatically generated

Fig. 3. The schematic of the BGA package

Table II. Dimensions of the BGA package

|  |  |
| --- | --- |
| Material | Dimensions / mm |
| PCB | 140 x 90 x 3.1 |
| IHS | 47.5 x 47.5 x 1.43 |
| Solder joint | Height: 0.45, Pitch: 1.0 |
| Die | 35 x 25 x 0.8 |
| Substrate | 47.5 x 47.5 x 1.9 |

As shown in Fig.4, a quarter symmetric model is used to save the computation cost. SAC305 solder balls are the interconnection between the package and PCB. The space around solder balls is vacuumed because the dominant mechanism of the BGA is heat conduction.

Chart, radar chart

Description automatically generated

Fig. 4. The schematic of the CFD model

Polyhedral meshing was used in ANSYS-FLUENT to achieve more efficiency compared with other formats, such as tetrahedral or hybrid meshes. Meshing independence of the CFD model was conducted for both the solid and fluid portions. Finer mesh granules should be used until achieving stable results. Meanwhile, the residual error should be maintained below 10-3 for the transient model.

The thermal properties of air were set in polynomial equations of temperature as shown in Table III. Because the solder paste experiences phase transform, the latent heat is involved in the simulation. The pure solvent melting heat of SAC 305 is 53.08 J/g. Solidus and liquidus temperatures of the solder are 217 ℃ and 220 ℃, respectively. The material properties of the other materials are in Table IV [3], [6], [8], [21].

Table III. Thermal Properties of Air

|  |  |
| --- | --- |
| Thermal property | Polynomial equations of temperature, T (K) |
| *K*  *(W/m.K)* | *K = 6.79E-05T+ 5.51E-03* |
| *Cp*  *(J/kg.K)* | *Cp = 4.00E-04T2 - 2.06E-01T + 1.03E+03* |
| *(kg/m3)* | *= 4.99E-06T2 - 6.33E-03T + 2.61E+00* |

Table IV. Dimensions of the BGA package

|  |  |  |  |
| --- | --- | --- | --- |
| Material | *K*  *(W/m.K)* | *Cp*  *(J/kg.K)* | *d*  *(kg/m3)* |
| Silicon | 98.4 | 721 | 2300 |
| PCB | 0.5 | 1405 | 2252 |
| SAC solder | 58 | 232 | 7380 |
| Substrate | 0.2 | 920 | 1700 |
| Copper (IHS) | 392 | 387 | 8940 |

*C. Model Validation*

The temperatures of the board and solder joint at the center of the BGA were used to fit the simulation results as shown in Fig.5. The peak temperature and time above liquidous of the measurement results are extracted to compare with the simulation results as shown in Table V. The error of three parameters shows that the CFD model is compatible with the measurement data. The maximum error is approximately 1 percent. Hence, the reflow profile simulated by the CFD model is feasible for the 7-zone convection reflow oven.

Fig. 5. Experiment and Simulation Results

Table V. Simulation results versus measurement results

|  |  |  |
| --- | --- | --- |
| Comparison of the linear ramp profile based on Table I | | |
| Parameter | Difference | Error |
| Board peak temperature | 2.35 °C | 1% |
| Center Solder\_Peak temperature | 1.606 °C | 0.67% |
| Center Solder\_Time above liquidous | 0.5 second | 0.60% |

III. Encoder-decoder Network

*A. Problem Formulations*

Temperature of the solder joints depends on the heat conducted from the package and board. The reflow profile of the BGA is associated with the physical information of the assembly and reflow recipe, which is the preset temperatures of the reflow oven. In this study, the dimensions of the PCB, substrate, and die combined with the reflow recipe are treated as the input. A sequence of temperature contours across the PCB is regarded as the output, which is used to define the reflow profile. Each input is formatted into a 2-D matrix, which is in the same shape of the output images. we map these problems to standard ML networks as shown in Fig. 6.

Diagram

Description automatically generated with medium confidence

Fig. 6. Data representation: Mapping the board physical information and reflow recipe into temperature contours.

*B. Network Development*

Encoder networks are successful in extracting 2D information for image classification and labeling tasks, which have low-dimensional outputs (class or label). In this study, the required outputs are high dimensional distributions of temperature contours, where the dimensionality corresponds to the pixel number of the contour and the pixel number is proportional to the size of the electronic package. This calls for a generative network that can translate the extracted low-dimensional data from an encoder network back into high-dimensional data.

As shown in Fig. 7, the Encoder-decoder network consists of two parts. Encoder network uses a sequence of 2-D convolution and max pooling layer pairs that capture key features from the high-dimensional input data set. The convolution layer conducts a weighted sum using a sliding window across the input image, and the max pooling layer reduces the dimension of the input data by extracting the maximum value from the sliding window. In Fig. 7, data dimension is reduced by 50% at each stage, and after 3 times such operations, the compressed, low-dimensional feature representing the input is achieved. Through these operations, the encoder helps understand the magnitude of the package dimensions and reflow recipe in the input matrixes but tends to be imprecise with the package structure. Decoder network is responsible for retrieving the layout of the package structure, which is blurred in data encoding. The decoder network uses the transpose convolution and upsampling layers, which are functionally the opposite of pooling layers, to increase the dimension of the input matrixes by replicating columns and rows.

Diagram

Description automatically generated

Fig. 7. Encoder-decoder network Architecture

*C. Network Training*

Regarding to the effect of the package structure on reflow profiles, out of plane dimensions are the most significant parameters. The thickness of the PCB, substrate and die are selected as the input factors as shown in Table VI. Regarding to the impact of the recipe, we know that the more backward the zone is, the more effect it brings. The preset temperatures of zone 4-7 are selected as the other four input factors. Complete cross combination was conducted to generate 972 sets of input data.

Table VI. Geometry parameters

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Package dimensions (mm)** | | | **Reflow recipe (°C)** | | | |
| **PCB thickness** | **Substrate thickness** | **Die thickness** | **Zone 4** | **Zone 5** | **Zone 6** | **Zone 7** |
| 3 | 1.9 |  | 180 | 220 | 260 | 280 |
| 2 | 0.9 |  | 210 | 250 | 280 | 300 |
| 1 | NA | NA | 240 | 280 | 300 | 320 |

The data was split for training, validation, and test sets. Before training the model, the raw data is normalized by subtracting the mean and dividing by the standard deviation and is used to train the network using an ADAM optimizer where the loss function is a pixelwise mean square error (MSE). To avoid overfitting, ReLU activation is applied after encoder layer and transpose convolution layer. Training time are xxx m using the NVIDIA xxx.

*D. Evaluation of Network Performance*

A comparison between the CFD model results and the neural network results are listed in Table VII. The runtime of each study case is 1s approximately. As show in Fig. 8, twelve uniformly distributed spots were chosen for the error statistic. The network has an average error of 0.63 °C and a maximum error of °C. Consider the low error percentage, the networks performance is acceptable for this application.

A picture containing application

Description automatically generated

Fig. 8

Table VII. Result predicted by the network

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| #1 |  |  |  |  |  |  |  |  |  |  |  |  |
| #2 |  |  |  |  |  |  |  |  |  |  |  |  |
| #3 |  |  |  |  |  |  |  |  |  |  |  |  |
| #4 |  |  |  |  |  |  |  |  |  |  |  |  |
| #5 |  |  |  |  |  |  |  |  |  |  |  |  |

Chart

Description automatically generated

Fig. 9 Temperature contour for case#1 (a) CFD simulation result (b) network predicted result

Chart

Description automatically generated

Fig. 10 Histogram of prediction error

The graphical view of network predicted temperature contour for test case#1 is shown in Fig. 9. Small discrepancy is observed between the numerical model calculated result and network predicted result. The histogram of prediction error in Fig. 10 shows the error level concentrated at 0-1%, The maximum error is xxx %

V. CONCLUSION

This paper presents a methodology to predict the reflow profile for the bulky BGA package. The CFD model used to simulate the reflow process was validated using a profiling board. The simulation results subjected to several sets of boundary conditions were organized and transferred to the network as the training data. The temperature contour under the BGA in 15 time slots can be output by the network in seconds. Instead of the conventional methodology, measuring the temperature of the tiny solder joint difficultly, the reflow profile predicted by the network ease the expense of experiment trials, which is benefit to reflow recipe optimization.

References

[25] Y. Lai, K. Pan, Y. Cen, C. Cai, P. Yin, J. Yang, and S. Park, “An Intelligent System for Reflow Oven Temperature Settings Based on Hybrid Physics-Machine Learning Model,” Soldering *and Surface Mount Technology*, 2022, in press.

[26] Y. Lai, K. Pan, C. Cai, P. Yin, J. Yang, and S. Park, “Smarter Temperature Setup for Reflow Oven to Minimize Temperature Variation Among Components,” *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2022, "in press".