Automatic Stencil Design Aperture Openings for BGAs to Mitigate Reflow Defects Caused by Chip Warpage using Numerical Shadow Moiré Data

Hamzeh Alshraida, Daryl Santos
Systems Science and Industrial Engineering Department
Watson Institute for Systems Excellence (WISE)
Binghamton University – State University of New York (SUNY)
NY, USA
halshra1@binghamton.edu, santos@binghamton.edu

Joseph Lee
FOXCONN Industrial Internet (FII)
CA, USA
joseph.lee@fii-na.com

Iman Dabous, Talbert Aguilar
NVIDIA
CA, USA
idabous@nvidia.com, taguilar@nvidia.com

Neil Hubble
AKROMETRIX
GA, USA
nhubble@akrometrix.com

ABSTRACT
The growing need for dependable assembly procedures in electronic manufacturing, driven by the trend toward miniaturization, necessitates effective measures to counteract reflow defects arising from chip warpage in Ball Grid Array (BGA) packages. Out-of-plane deformation is common in BGAs, attributed to varying thermal expansion coefficients, and eventually results in opens, Head-in-Pillow (HIP), or bridging defects. This work addresses the shortcomings of existing stencil design guidelines, which often overlook chip warpage, leading to common reflow defects. We propose a novel approach utilizing the Shadow Moiré technique to numerically quantify chip warpage at all pin locations, enabling the automatic optimization of stencil apertures. By processing this data with an advanced algorithm, we design apertures considering the following parameters: localized chip warpage, stencil thickness, Cu pad size, and pitch. This automated method reduces the reliance on engineer experience and eliminates the need for extensive experimentation. In this research, we have substantiated a previously suggested framework by applying emulated warpage data and evaluating its effectiveness using three commercial BGA packages susceptible to warping during reflow. Through empirical experiments and utilizing the proposed framework, we successfully designed the optimal aperture opening for each Cu pad under the chip, automatically, eliminating defects like opens, bridging, and HIP during reflow. The robustness and reliability of this approach have been verified for addressing challenges associated with chip warpage in the Surface Mount Technology (SMT) reflow process. The outcomes underscore significant improvements in defect prevention and overall assembly quality.

Key words: Electronics Manufacturing, Defect Reduction, DOE, BGA chip warpage, Dynamic warpage, Shadow Moiré, Stencil aperture optimization, Reflow defects, Head-In-Pillow

INTRODUCTION

Surface Mount Technology (SMT) has become a crucial method in assembling electronic devices and is known for its space-efficiency design, affordability, and enhanced electrical functionality (Prasad, 2013). BGA packages, among different types utilized in SMT, have become highly favored for their dense pin configuration, superior performance, and outstanding heat management (Liu et al., 1995). Despite these benefits, the assembly of BGAs presents specific challenges, especially in addressing the defects arising from chip warping during the reflow process.

Chip warping, often encountered in BGA packages, results from the differing Coefficients of Thermal Expansion (CTE) among various materials such as the base substrate, copper

(Cu) pads, and silicon (Si) die. This discrepancy in CTE causes significant deformations, typically in the tens of micrometers range (Suhir and Segelken, 1990), which can lead to various reflow-related defects, as shown in Figure 1. These issues, including incomplete soldering, solder bridging, stretching, and HIP problems, have the potential to negatively affect the dependability and functionality of the final electronic products.

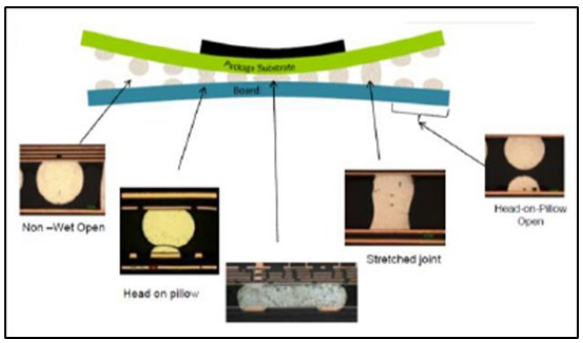

Figure 1. Possible chip warpage related defects (Huang et al., 2021)

In response to these issues, the industry has developed stencil design guidelines, such as those recommended by the Institute for Printed Circuits (IPC-7525A) in 2003. These guidelines typically suggest matching the stencil opening to the Cu pad size or reducing it by a specific percentage. However, these standards do not fully address the issue of chip warping, and thus are not completely effective in averting associated defects. Consequently, many Contract Electronic Manufacturing (CEM) companies have adopted a strategy of optimizing stencil aperture openings through a multi-zone approach. This technique involves creating several zones, usually between three and five, based on warpage data gathered using the Shadow Moiré method, particularly the TherMoiré technique. The Shadow Moiré method provides precise measurements of chip warping by analyzing the chip surface's topography under a temperature profile similar to the SMT reflow process (Wang and Hassell, 1997).

While commonly used, the multi-zone stencil optimization technique suffers from a lack of standardization and largely depends on the expertise of SMT engineers. Despite its prevalence, there is a lack of comprehensive literature on this topic. This gap underscores the urgent need for a more systematic and efficient method to optimize stencil aperture openings. Such a solution should aim to reduce reliance on empirical knowledge and decrease the necessity for extensive Design of Experiments (DOEs).

In their study, Huang et al. (2021) investigated reflow defects linked to warpage in BGA packages during their integration onto printed circuit boards (PCBs). They utilized the Shadow Moiré technique for examining the dynamic warpage of both the BGA substrate and the PCB. Their objective was to pinpoint high-risk pin locations on the chip that are more likely to experience defects, thereby enabling the optimization of stencil aperture openings for these specific pins. The study specifically revealed that pins located at the corners of the chip were more prone to defects such as open or stretched joints. To address this issue, the researchers suggested enlarging the stencil openings at these vulnerable points, a strategy depicted in Figure 2. This approach is in line with methods commonly employed by CEMs. However, it is important to acknowledge that such strategies are highly dependent on the skill and experience of SMT engineers and require thorough validation through DOEs.

	Pad area	Diameter (mm)		
		PCB pad	Stencil 1	Stencil 2
BGA U1 opening	Red	0.53	0.635	0.64
	Orange	0.51	0.635	0.62
	Green	0.48	0.635	0.60
	White	0.48	0.41	0.58

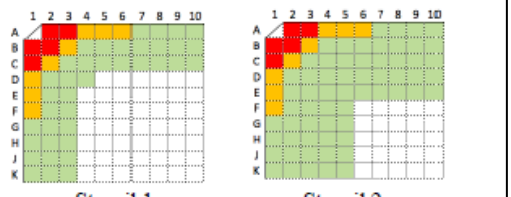

Figure 2. Stencil optimization at the high-risk locations (Huang et al., 2021)

This paper introduces a new approach that utilizes the Shadow Moiré technique to measure chip warping and automatically determines optimal stencil aperture openings (Alshraida et al., 2023). Our method involves processing the Shadow Moiré data with a sophisticated algorithm, which designs the aperture openings based on factors like the degree of chip warping in the z-direction, copper pad pitch, and size. This proposed system offers an automated framework that reduces dependency on the experience of SMT engineers and minimizes the need for extensive DOEs, making the process more streamlined, efficient, and cost-effective.

This automatic framework presents a more systematic and efficient alternative to the existing multi-zone stencil optimization practice. It features optimized stencil aperture openings across N-zones, each corresponding to a Cu pad. This setup ensures precise solder deposition, effectively reducing reflow defects caused by chip warping. By optimizing stencil aperture openings using the proposed method and comparing them with current techniques, we showcase the efficacy of our approach in minimizing defects and enhancing the overall quality of the assembly process.

Our method, which automates stencil optimization using Shadow Moiré numerical data, offers a solid and dependable way to tackle issues related to chip warping in the SMT reflow process. This innovation boosts the reliability and effectiveness of electronic devices, significantly enhancing the overall quality of electronic assemblies.

METHODOLOGY

Figure 3 illustrates our proposed methodology, showcasing an automated, data-driven approach to fine-tuning stencil aperture openings for BGAs. By integrating sophisticated algorithms, machine learning techniques, and meticulous control of aperture dimensions, our method significantly enhances the reliability and performance of electronic devices equipped with BGAs. It effectively addresses reflow defects caused by chip warping. This system determines the optimal aperture size by a specific function that considers variables such as Cu pad dimensions, pitch, center coordinates, z-warpage value, and stencil thickness, as detailed in Equation 1.

$$Aperture\ Size_i = f(\emptyset, P, x_{ci},\ y_{ci}, dz_i, TH) \qquad (1)$$

Where $\emptyset$ is the Cu pad diameter, P is the BGA pitch, $(x_{ci},\ y_{ci}, dz_i)$ are the deformation matrix, and TH is the thickness of stencil.

Our model starts by taking Shadow Moiré data as input, which captures the out-of-plane deformation, or z-warpage, at each (x, y) location on the chip. In addition, we use the Gerber file of the BGA copper pads. The Gerber file is converted into a DXF_ASCII text file to make this data compatible with Python processing. This conversion produces a matrix detailing all BGA copper pads' centers $(x_{ci},\ y_{ci})$, a crucial step for the following analyses.

Subsequently, we employ a prediction model known as Random Forest (RF) to determine the z-warpage at each copper pad center. RF, a supervised machine learning algorithm, is commonly used for regression challenges like ours. It operates by constructing multiple decision trees on various data samples and then using a majority voting system for the final prediction (Segal, 2004). For our purpose, the RF model is trained using the Shadow Moiré data, which indicates dz (z-warpage) at every (x, y) point. This training allows the model to accurately predict dz_ci at the Cu pad centers $(x_{ci},\ y_{ci})$. The selection of RF is based on its proven high accuracy in prediction, as evidenced by its high R^2 value and low mean square and absolute error (MSE, MAE) values.

After determining the z-warpage values at each copper pad center, we then identify potential soldering defects associated with the varying z-warpage measurements. A high value of z-warpage suggests a likelihood of defects such as open, insufficient, stretch, or Head-in-Pillow in the solder joints. Conversely, Near Zero z-warpage value indicates a potential risk of bridging defects in the soldering process.

To counteract the effects of warpage and avert soldering defects, precisely regulating the volume of solder paste deposited on each copper pad is essential. This control

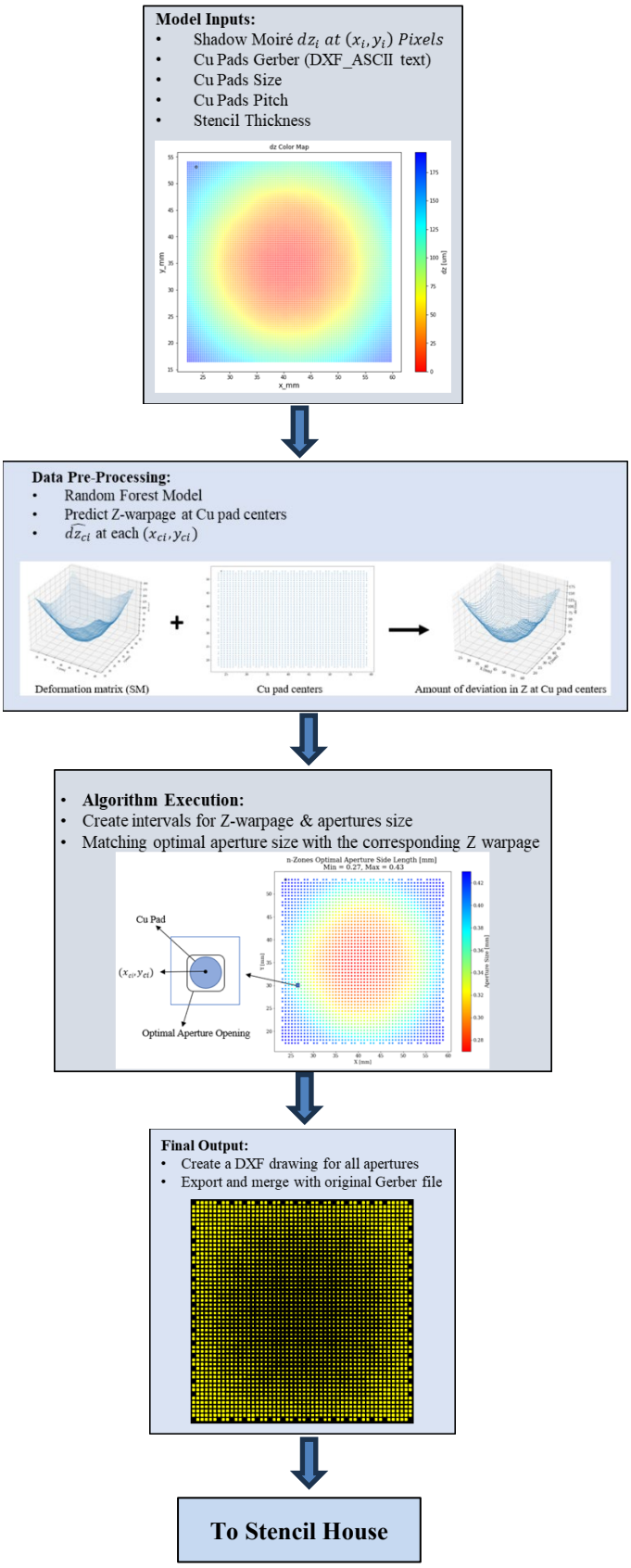

Figure 3. Process flow for optimizing stencil aperture openings

is attained by designing optimal stencil aperture openings. Our algorithm automates the design of apertures for each individual Cu pad. It considers a range of factors, including the degree of warpage at the pad's center, the pad's size and pitch, stencil thickness, and the height of the solder bump. This tailored approach ensures effective mitigation of potential defects due to warpage.

Our algorithm systematically automates the design of stencil apertures for each copper pad. It begins by establishing an interval based on z-warpage: one from the maximum value of z-warpage to zero. This interval is then subdivided into 'N' equally spaced distances or ranks, in which N equals the number of Cu pads in the BGA components. Next, the algorithm considers factors like the Cu pad size, pitch, area ratio, and stencil thickness to determine the maximum and minimum sizes for the aperture openings. Another interval is created: from the allowable minimum aperture size to the allowable maximum aperture size. Similar to the warpage interval, these are divided into the same 'N' ranks.

As illustrated in Figure 4, the final step is to match each aperture size rank with its corresponding warpage rank. Consequently, Cu pads with the maximum warpage are assigned the largest aperture size to prevent open defects. Conversely, pads with the minimum (Near Zero) warpage receive the smallest openings to avoid bridging. This ensures every Cu pad receives an aperture tailored to its specific warpage characteristics, as depicted in Figure 1.

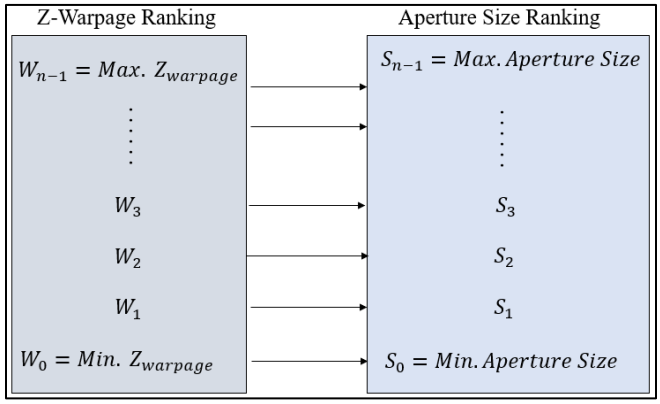

Figure 4. Matching optimal aperture Size with the corresponding amount of warpage

The value of 'N', which is consistent across all the intervals, is set to be the number of aperture zones intended to be created. The objective is to have a distinct opening for each Cu pad. This approach ensures that, as 'N' increases, the algorithm facilitates a smooth and gradual transition in aperture sizes across all Cu pads, accommodating the different warpage levels.

This aspect of the algorithm is particularly crucial. It effectively addresses the limitations inherent in the multi-zone approach commonly used by most CEMs, which can result in optimal aperture sizes due to varying degrees of warpage. Our method's nuanced approach to aperture size variation ensures a more precise and effective response to warpage across all Cu pads.

Addressing the next challenge involves creating a Gerber file with optimized stencil openings. By this stage, the algorithm has successfully generated a data frame that includes the center coordinates (x_{ci}, y_{ci}) of each Cu pad, along with the respective optimal aperture size for each. Given that some chips have as many as 12,000 Cu pads, manual drawing of these apertures is not feasible. Therefore, automation becomes essential.

To efficiently execute this task, we utilize the "ezdxf" library in Python, noted for its powerful functionality. This library can quickly generate a DXF drawing, completing the process in a remarkably short time - from a few seconds to several minutes. After the optimal stencil aperture opening DXF is prepared, the final step is to merge it with the original Gerber file of the PCB. This integration can be easily done using any standard commercial PCB design software, making the file ready for further processing and stencil cutting. This streamlined process ensures that the stencil is accurately and efficiently prepared, ready for use in the subsequent stages of electronic assembly.

EXPERIMENTAL SETUP

In this section, we delve into a case study focused on a 38mm x 38mm chip equipped with copper pads that have a pitch of 0.7mm and a size of 0.3mm, as depicted in Figure 5. This example will illustrate the practical application of our innovative methodology for designing optimal stencil aperture openings. We will provide a comprehensive walkthrough of each phase in the process, covering model input, data preprocessing, the execution of our algorithm, and the final output generation. The objective of this case study is to demonstrate the efficiency and effectiveness of our approach in addressing the challenges posed by chip warping during the reflow process. We aim to show how our method can significantly reduce defects, highlighting its practical utility in electronic device assembly.

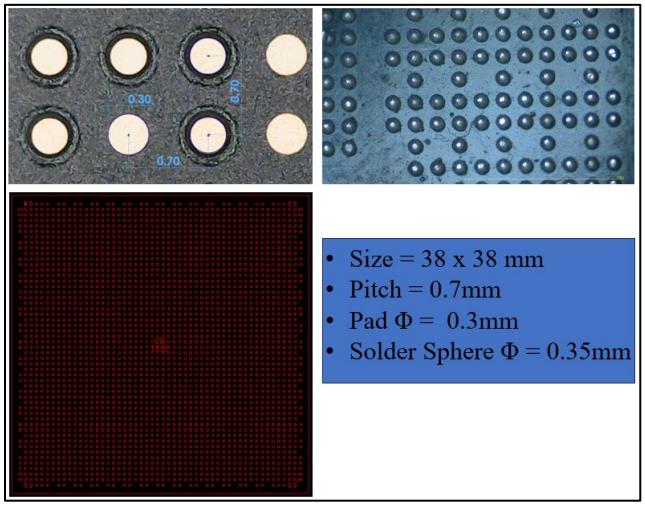

Figure 5. Chip dimensions

Model Input

In this model, the first input is the Gerber file of the Cu pads on the PCB. This file is converted into a DXF_ASCII text format using Python, a crucial step as it enables the extraction of the center coordinates (x_{ci}, y_{ci}) of each Cu pad. These coordinates are essential for predicting warpage at the center of each pad and for the automated drawing of optimal aperture openings.

The second key input is the deformation matrix obtained from the Shadow Moiré test, which includes the (x_i, y_i) pixels on the chip along with their respective out-of-plane deviations (dz_i). For this case study, the deformation matrix was sourced from an AKROMETRIX TherMoiré machine. Here, the 38mm x 38mm chip was subjected to a thermal condition emulating a real-life reflow profile, enabling us to derive the chip's warpage surface as a deformation matrix (x_i, y_i, dz_i), as illustrated in Figure 6. The resulting out-of-plane values from this emulation ranged from zero to 205.8 µm at the peak temperature of 206ºC, indicating the extent of warpage across the entire chip surface under emulated reflow conditions. These values are crucial for determining the range for maximum and minimum aperture openings, for more details, see Arellano (2018).

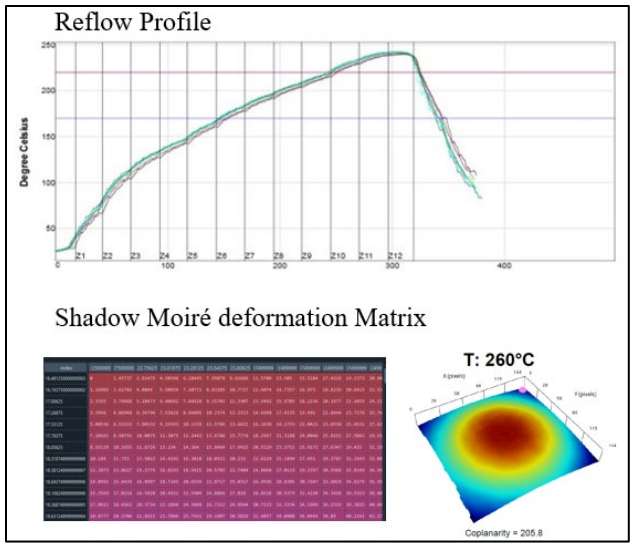

Figure 6. Shadow Moiré Chip Warpage Test

Additionally, other necessary inputs include the pitch (0.7mm), Cu pad size (0.3mm), and stencil thickness (4mils). These parameters are indispensable as they define the range of possible aperture opening sizes for the optimization process. By integrating all these inputs, our algorithm is able to precisely adjust the aperture sizes for each individual Cu pad, taking into account its unique warpage characteristics and geometrical factors. This comprehensive approach ensures the optimization of each Cu pad for the best possible soldering outcomes.

Data Preprocessing

During the data preprocessing phase, our primary objective is calculating the out-of-plane deviation at the center of each copper pad (dz_{ci}). For this, we utilize the Random Forest regression model, inputting the (x_i, y_i, dz_i) data to predict the (dz_{ci}) values. The RF model showcases outstanding accuracy, evidenced by its performance metrics: a Mean Square Error (MSE) of just 0.35µm, a mean absolute error of 0.26µm, and an impressive R^2 value of 99.99%. These statistics highlight the RF model's precision and reliability in accurately determining the warpage at the centers of the Cu pads. Figure 7 illustrates the performance of the RF model, confirming its efficacy in accurately representing warpage patterns across the chip's surface. This step is crucial as it ensures the subsequent optimization of aperture sizes is based on precise and reliable warpage data, critical for minimizing defects in the reflow soldering process.

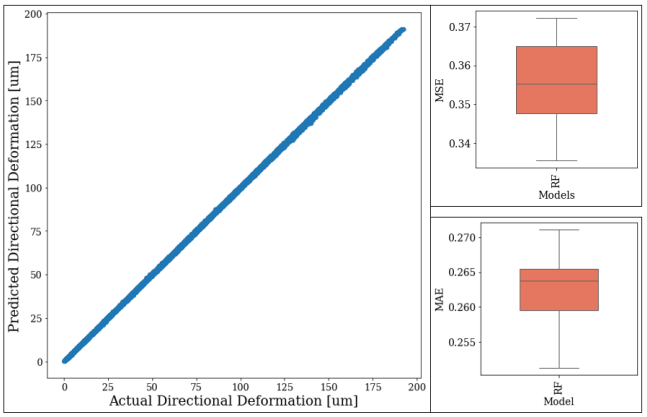

Figure 7. Random Forest performance measures [µm]

Algorithm Execution

At this stage, we have obtained the z-warpage (dz_{ci}) at each Cu pad center, leading to the identification of the following interval: (0, 205.8µm) warpage. Regarding the minimum and maximum aperture size in this particular case study, the intervals of side length of a square-round aperture shape span from 0.27mm to 0.43mm. The algorithm's approach involves selecting a zone per Cu pad, dividing the interval into N equally spaced segments, thereby resulting in N different aperture sizes. For instance, a Cu pad with a maximum warpage of 205.8µm will correspond to an aperture opening edge length of 043mm. Following the algorithm's execution, we are left with a new matrix comprising $(x_i, y_i,$ optimal aperture edge length), which will be used in the final output.

Model Final Output

In the Final Output phase, our process culminates with assembling a detailed matrix. This matrix includes all the copper pad center coordinates and their respective optimal aperture side lengths. To convert this matrix into a practical form, we create a DXF drawing. This is done using the ezdxf Python library, which facilitates the crafting of square apertures with the calculated optimal side lengths around their corresponding centers (x_{ci}, y_{ci}). To enhance the aperture shapes, fillets with radii equal to 20% of these optimal edge lengths are integrated, refining the geometries of each aperture. This process is replicated across all 2645

aperture openings, a substantial number in our study. The resulting DXF drawing, which includes all these apertures, is then exported in DXF format. Figure 8 illustrates this DXF file, which is seamlessly merged with the original Gerber file for the stencil house's cutting process. Notably, this file demonstrates the smooth transition in aperture sizes corresponding to the varying amounts of warpage.

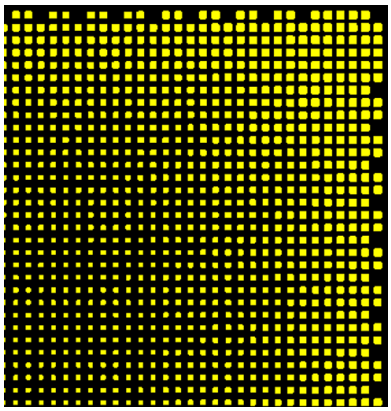

Figure 8. Optimal aperture openings in Gerber format

Automating this complex phase is both beneficial and essential, given the impracticality of manually drawing such a large number of unique apertures. Automation enables the efficient conversion of the entire matrix into drawings in just a few minutes, showcasing the efficiency and effectiveness of our approach in addressing the intricacies of stencil design for optimal solder paste application in electronic assembly.

RESULTS AND DESCUSSION

To demonstrate the effectiveness of our proposed methodology over the common practice of using three to five aperture opening zones, Figures 9-a and 9-b present comparative plots that illustrate the relationship between out-of-plane deformation and optimal aperture size for the existing practice and our proposed method, respectively. These figures highlight the distinctions between the two approaches. In Figure 9-b and 9-d, our method's capability to finely tailor aperture openings in response to varying amounts of warpage is showcased. This demonstrates how our approach aligns each aperture size precisely with the corresponding z-warpage level, ensuring a more accurate and effective solder deposition.

On the other hand, Figures 9-a and 9-c reveal the shortcomings of the current practice. Here, it is clear that different z-warpage levels are often assigned the same aperture sizes. This can lead to inconsistencies, especially at the borders of different aperture zones, where adjacent apertures with nearly the same z-warpage might have significantly different sizes, resulting in abrupt transitions in aperture size.

Although the current approach might be functional, the incorrect volume of solder deposition can stress the joints, affecting their long-term durability. This highlights how our proposed method ensures the precise amount of solder is deposited on each individual Cu pad. This approach offers a dual advantage: it not only eliminates reflow defects during the assembly process but also has the potential to enhance the long-term performance of the solder joints. This aspect underscores the need for further research to fully understand the long-term benefits of our methodology.

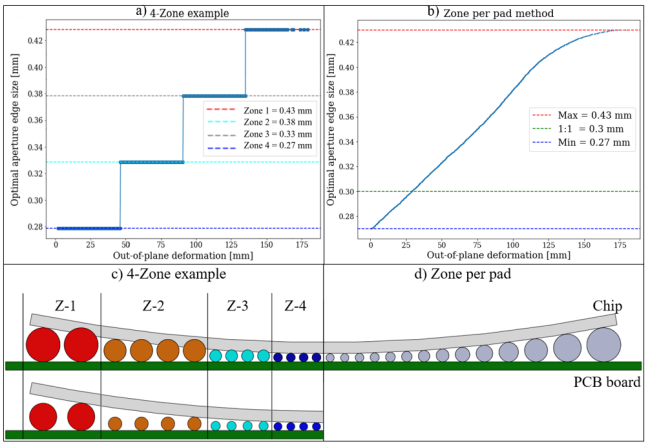

Figure 9. Out-of-plane deformation vs. optimal aperture size for both methods

While the prevailing method typically requires numerous DOEs to identify optimal aperture sizes, a process that is both time-consuming and resource-intensive, our proposed methodology offers a more efficient alternative. With the automation of this process, once the Shadow Moiré data is obtained, it only takes minutes to generate the Gerber file with the optimal aperture openings. This streamlined, automated feature of our method represents a significant advancement, reducing the time and resources needed for stencil design in electronic assembly.

CONCLUSION

As electronic devices continue to shrink in size, the demand for flawless assembly processes to combat reflow defects caused by chip warping in BGA packages has become increasingly crucial. This study has highlighted the limitations of existing stencil aperture design guidelines that fail to adequately address the issue of chip warping, leading to common reflow problems such as open soldering, solder bridging, and HIP defects. In response, we have developed an innovative solution that leverages the Shadow Moiré technique, machine learning, and process automation to revolutionize stencil aperture optimization. By intricately processing Shadow Moiré data through an advanced algorithm, we have crafted stencil aperture designs that accurately consider the extent of warpage, Cu pad pitch, and individual pad dimensions. This groundbreaking approach reduces reliance on SMT engineer expertise and extensive DOEs, thereby setting new standards in efficiency and accuracy. Our method uses the Shadow Moiré technique to map chip surface warpage accurately during emulated SMT reflow conditions. This data enables us to customize aperture sizes and shapes for each Cu pad across N-zones, ensuring

precise solder deposition and reducing reflow defects related to chip warping.

To further validate the effectiveness of this proposed methodology, future research will focus on extensively analyzing the long-term performance of solder joints. This next phase will include a series of reliability tests, such as thermal cycling and shock assessments. The results of these tests will be crucial in determining the long-term viability and durability of our proposed method, both in the SMT assembly process and throughout the service life of the electronic components.

REFERENCES

1. Huang, X., Ou, C., Zhu, S., & Huang, Y. (2021, September). Warpage measurement of substrates and printed circuit boards with Shadow Moiré. In 2021 22nd International Conference on Electronic Packaging Technology (ICEPT) (pp. 1-4). IEEE.
2. IPC - Association Connecting Electronics Industries. (2003). Stencil Design Guidelines (IPC-7525A). IPC - Association Connecting Electronics Industries.
3. Prasad, R. (2013). Surface mount technology: principles and practice. Springer Science & Business Media.
4. Liu, J. J., Berg, H., Wen, Y., Mulgaonker, S., Bowlby, R., & Mawer, A. (1995). Plastic ball grid array (PBGA) overview. Materials chemistry and physics, 40(4), 236-244.
5. Segal, M. R. (2004). Machine learning benchmarks and random forest regression. Center for Bioinformatics and Molecular Biostatistics, University of California, San Francisco. https://escholarship.org/uc/item/35x3v9t4
6. Suhir, E. & Segelken, J. M. (1990). Mechanical behavior of flip-chip encapsulants. Journal of Electronics Packaging, 112(4), 327-332.
7. Wang, Y. Y. & Hassell, P. (1997, October). Measurement of thermally induced warpage of BGA packages/substrates using phase-stepping shadow moiré. In Proceedings of the 1997 1st Electronic Packaging Technology Conference (Cat. No. 97TH8307) (pp. 283-289). IEEE.
8. Alshraida, H., Santos, D., Lee, J., Chen, W., and Zhang, A. (2023). Automatic Optimization of SMT Aperture Openings for BGAs to Mitigate Reflow Defects Caused by Chip Warpage using Shadow Moiré Data. Proceedings of the 12th Annual World Conference of the Society of Industrial and Systems Engineering (SISE), Munich, Germany, October 9-10.
9. Arellano, J., Alvarez, E., Perng, S., Hubble, N., & Mexico, J. (2018). Thermal Shadow Moiré to Cross-Section Correlation Study. In Proceedings SMTA International.