OPTIMAL ITERATIVE LEARNING FEEDFORWARD CONTROL OF WAFER TEMPERATURE UNIFORMITY IN VACUUM REFLOW SOLDERING

Min Sig Kang, Won Ho Jee
Dept. of Mechanical & Automotive Engineering Gachon University, Nano Solution Tech.
Bokjung-Dong, Sujung-Gu, Sungnam, Kyunggi-Do/KOREA
mskang@gachon.ac.kr

ABSTRACT
As device dimensions shrinks to submicrometer range, wafer level vacuum packaging and integration is an emerging method to resolve the requirements from industries. Addressing this issue and in consideration of nearly identical environmental conditions of each soldering process and difficulties in measuring temperature of wafer in real soldering process, in this paper, an optimal run-to-run iterative learning gain scheduled feedforward control of wafer temperature uniformity in vacuum reflow soldering process has been investigated. The proposed technique has been applied to a vacuum reflow soldering equipment for 8-in wafer. Along with run-to-run sequential experiments, gradual improvement of temperature uniformity across a wafer surface has been shown. The wafer was heated to desired temperature recipe and the temperature uniformity was maintained as desired level by applying the proposed control.

KEY WORDS
Vacuum reflow soldering, Wafer level packaging, Temperature uniformity, Thermal control, Optimal iterative learning control, Gain scheduled feedforward control

1. Introduction
Recently, wafer level vacuum packaging gradually substituted the traditional packaging technologies because of its significant advantages such as: higher packaging efficiency, smaller size, lighter weight, lower cost, higher productivity, shorter lead length, environmental-friendly, etc[1].
Solder bonding is an essential process of wafer level packaging. A common solder bonding for MEMS and LED industries is a BGA(ball grid array). Strong mechanical rigidity, low electrical resistance, and low thermal resistance are important functions to be provided by solder joint. However voids in solder joint weaken mechanical stiffness of interconnection, increase electric resistance, and increase thermal resistance. Voids are generated very easily during soldering of BGA[2].

Vacuum soldering is an effective process to reduce voids in the joint. Moreover, during soldering, it is required to maintain temperature uniformity over wafer to guarantee uniform quality of joint[2]. To maintain uniformity of temperature, spatial density of energy flux reaching to wafer should be controllable arbitrary.
A suggested soldering temperature range for the commonly used lead-free solders such as SnAg, SnAgCu, SnCu is between 230 ~ 250°C. The ramp rate of 1 2/o Cs is suggested in the aspect of minimizing defects in solder joint [1].
Recently, RTP(rapid thermal processing) RTP systems attempt very high rates of ramping, e.g. 100 ~ 200°C/s to increase throughput and control the temperature nonuniformity accurately during ramp-up and steady state. In the control aspect, to achieve the desired uniformity, multi-variable distributed temperature control is necessary. Many studies addressing on control of RTP systems employed various forms of MIMO control strategies [3-12]. Basically, these controls adopt various forms of real-time feedback and/or feedforward controls, which require multiple real-time measurements of wafer temperature. Compared to the common RTP systems, the vacuum reflow soldering addressed in this paper requires less strict performances such as slower ramp rate and wider bounds of permissible nonuniformity. Furthermore each soldering process has nearly identical environmental conditions.
Motivated form the above, we proposed an optimal run-to-run iterative learning gain scheduled feedforward control technique. Under the similar environment conditions and the fixed desired temperature profile, the power trajectories to achieve temperature uniformity over wafer are also similar to each other, from run to run. This control is simple in structure and cost effective than other controls. The run-to-run iterative control is a feasible technique typically for repetitive processes [13-18].
The proposed control has been applied to a vacuum reflow soldering equipment. The feasibility of the proposed control technique is demonstrated along with some experimental results.
This paper is organized as follows. In section 2, we discuss the configuration of an experimental vacuum reflow soldering equipment. In section 3, we describe the control design technique. Some experimental results and
discussions are detailed in section 4. Concluding remarks are followed in section 5.

2. Vacuum Reflow Soldering Equipment

A photo of the vacuum reflow soldering equipment under consideration is shown in Fig. 1. It was designed and manufactured for lead-free, void-free and fluxless 12-in wafer level vacuum reflow soldering process. As can be seen in Fig. 1, the soldering equipment consists of a vacuum pump, a soldering chamber, and a fast cooling system by means of $N_2$ gas. The vacuum pump is used to vent and maintain the vacuum level in the soldering chamber under 1 torr. Soldering in vacuum environment is crucial to reduce void in solder joints. The fast cooling by $N_2$ gas is used to improve throughput.

As shown in Fig. 4, the front wall of the chamber has an access door to deliver and draw out wafers, and the vacuum pump and inlet of coolant $N_2$ gas are connected through the top cover of the chamber. Inside of all walls, the coolant water is circulated along the path in the wall to keep off overheating.

![Figure 1. Photo of the vacuum reflow soldering equipment](image1)

The vacuum reflow equipment in Fig. 1 is specially designed for wafer level soldering of BGA as shown in Fig. 2 and Fig. 3. In this soldering process, an array of solder balls is on a wafer and chips to be bonded are on the solder balls.

![Figure 2. Photo of BGA on wafer](image2)

As shown in Fig. 5 schematically, the heating mechanism consisted of bank of lamps, a cylindrical chuck, and a wafer. A flat reflector is located under the bank of lamps. To achieve temperature uniformity across the wafer over a range of soldering condition, heat flux distribution radiating to the wafer can be varied according to the desired soldering recipe.

Even though infinite number of independent heat source are necessary to achieve perfect temperature uniformity across the whole wafer surface, we selected 7 concentric circular ring type tungsten-halogen lamps as heating sources in the consideration of the level of permissible nonuniformity and its cost. As a matter of course the output power of each lamp can is controlled independently.

The wafer is put on a circular array of pins and the gap between the wafer and the chick is approximately 0.5mm. The chuck is made of graphite which has low thermal expansion rate, high thermal conductivity, high strength, etc. Its thickness is 5mm. Evidently, energy transfer from the lamps to the wafer is not direct because the chuck blocks the path of radiation.
from the lamps. The radiation energy flux from the lamps transferred to the chuck and then the radiation energy flux from the heated chuck absorbed by the wafer. In addition, conductive heat transfer and additional radiative heat transfer from the chamber wall take place in the wafer.

It can be expected that this indirect heating mechanism is not efficient for fast heating response, but it can give benefits in the aspect of improving uniformity across the large wafer. The chuck makes even the sharp change of temperature distribution of the wafer.

Since soldering is carried out in the vacuum and nitrogen environment, convection through the spaces between the lamps and the chuck, and the chuck and the wafer can be neglected.

A thermocouple is bonded to the chuck at its center to measure its temperature. The measured temperature signal is used to control the ramp and dwell periods defined in the given temperature recipe for soldering.

3. Temperature Control Design

The thickness of the typical wafer under consideration is in the range of $0.5 \sim 0.7 \text{mm}$. Thus we can assume that the temperature is uniform across the thickness of the wafer.

As mentioned in the section 2, energy flow from the heaters to the wafer is indirect. This gives more complex energy transfer mechanism than direct heating mechanism. The heat flux reaching to the wafer is a function of power of individual lamps, shapes and positions of lamps, position of the wafer, chuck position and its geometry, chamber geometry, reflectance of the reflector, pressure in the chamber, etc. Thus the governing thermal model is seriously complex and has thoroughly nonlinear characteristics.

The equation describing the heat transfer characteristics between the N-sensor positions on the wafer and the M-lamps can be represented as follows[17]:

$$\hat{T}(T) = F(T) + A(T)H$$

(1)

where $T$ denotes an $N \times 1$ vector and each entity refers to temperature at each sensor position. $\hat{T}$ denotes the time derivative of $T$. $F(T)$ denotes an $N \times 1$ vector and each entity is a function describing heat transfer. $A(T)$ is an $N \times M$ matrix and its element $a_{ij}$ represents the influence coefficient of the power of the $j$-th lamp to the rate of temperature at the $i$-th sensor position. $H$ is an $M \times 1$ vector and each entity denotes power of each lamp. In eq.(1), the notation denoting time is dropped for representational convenience.

Even though, $F(T)$ and $A(T)$ in Eq.(1) are highly nonlinear and highly dependent upon status of environment, they can be represented by a series of linearized models. Furthermore, with vacuum reflow soldering process, the environmental conditions surrounding the wafer are almost identical to each other, from run to run. From these discussions, we can rewrite the governing equation in Eq.(1) as the following form:

$$\hat{T}_i = E_i + A_iH, \; i = 1, 2, ..., n$$

(2)

where $n$ denotes the number of equation which is linearized within the consecutive temperature range $T_{i-1} \leq T < T_i$.

Based on the assumption in the above that the model in Eq.(2) is almost identical to the one in the previous run, we can design a gain scheduled feedforward control by using a run-to-run iterative learning technique for temperature uniformity across the wafer surface.

Figure 6 illustrates the design steps proposed in this paper and detailed procedure is as follows:

---

Step 1: Model identification
Identifying the matrix $A_i(T), \; i = 1, 2, ..., n$ composed of the influence coefficients between the lamp powers and the temperature rates of the wafer through sequential experiments. For this experiment, a test wafer which is the same as the actual wafer is prepared. On the test wafer, N-thermocouples are bonded. After sitting the test wafer on the chuck, the chamber is closed and vented, and then one of the lamps is excited. The temperature responses are measured by the thermocouples on the test wafer. The same procedure is repeated by changing lamps. Finally, temperature range is divided into n-zones from rate analysis of the whole temperature responses achieved from M-experiments. In each zone, the matrix $A(T)$ is fitted into a constant matrix $A_i$. The elements of $A_i$ can be determined from an optimal fit of temperature rate.

Step 2: First trial of feedforward control input
Since $E_i$ is not identified in the step 1, the first trial feedforward control input of each lamp is determined from the following sub-optimization.
Find \( H_i, i=1, 2, ..., n \) which minimizes the performance index \( J_i, i=1, 2, ..., n \) defined by

\[
J_i = (S_i^* - AH_i)^T (S_i^* - AH_i), \quad i=1, 2, ..., n \quad (3)
\]

where \( S_i^* \) is an \( N \times 1 \) vector which denotes the given desired rate of temperature in the \( i \)-th temperature zone. In the optimization in Eq. (3), the limited power of each lamp and non-negativity of input power should be encountered. This thus optimization is a constrained optimization problem.

Next, apply the input vector \( H_i \) determined in the above optimization to each lamp and measure the temperature profile. During the procedure, the dwell periods are controlled by referencing the temperature from the thermocouple bonded on the chuck.

Step 3: Iterative modification of feedforward control input
From the temperature profiles measured in the previous run, find the temperature rates and its errors \( \Delta S_i, i=1, 2, ..., n \) in each zone defined by

\[
\Delta S_i = S_i^* - S_i^{exp}, \quad i=1, 2, ..., n \quad (4)
\]

where \( S_i^{exp} \) is an \( N \times 1 \) vector which represents the calculated temperature rate from the measurements.

To compensate the rate errors, the feedforward control input powers can be updated from the following optimization:

Find \( \Delta H_i, i=1, 2, ..., n \) which minimizes the performance index \( \Delta J_i, i=1, 2, ..., n \) defined by

\[
\Delta J_i = (\Delta S_i - AH_i)^T (\Delta S_i - AH_i), \quad i=1, 2, ..., n \quad (5)
\]

Then, the updated feedforward control input powers \( H_{i,new}, i=1, 2, ..., n \) are determined as

\[
H_{i,new} = H_{i,old} + \Delta H_i, \quad i=1, 2, ..., n \quad (6)
\]

where \( H_{i,old} \) is the vector of control input power applied in the previous run.

As a matter of course, the constraints on the power limit of each lamp should be encountered in the optimization in Eq. (5).

Applying the newly updated control input \( H_{i,new} \) to the system and measure the temperature profiles.

The above step 3 will be repeated until the temperature uniformity across the wafer falls within the permissible level.

### 4. Experimental Results

To evaluate the control performances suggested in the above, the control has been applied to the vacuum reflow soldering equipment illustrated in section 2. On the test wafer, 5 thermocouples were bonded on the front side of the wafer surface. The thermocouples were equally spaced along a radial line from its center to the edge. The thermocouples were numbered by ascending order from the center to the edge.

According to the stepwise procedure proposed in the section 3, the identification procedure illustrated in the step 1 has been carried out.

In Fig. 7, a set of wafer temperature responses is shown. In this experiment, only the lamp #1 is excited by 70% of its maximum power. The number by each plot in the figure represents the number of the thermocouples. During this experiment, the other lamps except #1 were not excited.

As can be seen in Fig. 7, because of large inertia of chamber wall, the temperature increases continuously. The indirect heat transfer characteristic of this system is another main reason of slow response. Evidently, the temperature of the thermocouple #1 shows faster increase than others because the thermocouple #1 is closer than others to the lamp #1. This exhibits the feature of radiational heat transfer.

![Figure 7. Temperature of wafer driven by Lamp #1 with output power ratio 70%](image)

Table 1. Temperature rate when driven by Lamp #1.

<table>
<thead>
<tr>
<th>Lamp</th>
<th>~100°C</th>
<th>~170°C</th>
<th>~220°C</th>
<th>~270°C</th>
<th>270-300°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1</td>
<td>4.12</td>
<td>2.92</td>
<td>1.90</td>
<td>1.43</td>
<td>1.00</td>
</tr>
<tr>
<td>s2</td>
<td>3.01</td>
<td>2.00</td>
<td>1.60</td>
<td>1.21</td>
<td>0.87</td>
</tr>
<tr>
<td>s3</td>
<td>2.22</td>
<td>1.76</td>
<td>1.34</td>
<td>1.04</td>
<td>1.04</td>
</tr>
<tr>
<td>s4</td>
<td>1.75</td>
<td>1.66</td>
<td>1.24</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>s5</td>
<td>1.65</td>
<td>1.49</td>
<td>1.15</td>
<td>0.93</td>
<td>0.93</td>
</tr>
</tbody>
</table>

From the responses in Fig. 7, we divided the temperature range from the ambient temperature to 300°C into 5-zones from rate analysis, and linearized rates in all zones were determined from an optimum fitting calculation. In this linearization, we considered allowable fitting errors. The fitted rates of temperature in each zone are...
demonstrated in Table 1. Where \( s_i, i = 1, 2, \ldots, 5 \) refers to the calculated rate of temperature from the response measured by the thermocouple \#i.

The same experimental procedures were carried by changing driving lamp. Figure 8 shows the measured temperature profile when only the lamp \#4 was derived. The fitted rates from the response in Fig. 8 are given in Table 2.

Table 2. Temperature rate when driven by Lamp \#4

<table>
<thead>
<tr>
<th>( s_i )</th>
<th>(-100^\circ C)</th>
<th>(-170^\circ C)</th>
<th>(-220^\circ C)</th>
<th>(-270^\circ C)</th>
<th>(270^\circ C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1</td>
<td>2.24</td>
<td>3.03</td>
<td>2.07</td>
<td>1.53</td>
<td>1.34</td>
</tr>
<tr>
<td>s2</td>
<td>2.54</td>
<td>2.93</td>
<td>1.91</td>
<td>1.80</td>
<td>1.30</td>
</tr>
<tr>
<td>s3</td>
<td>2.27</td>
<td>2.58</td>
<td>2.24</td>
<td>1.57</td>
<td>1.36</td>
</tr>
<tr>
<td>s4</td>
<td>2.00</td>
<td>2.57</td>
<td>2.11</td>
<td>1.62</td>
<td>1.21</td>
</tr>
<tr>
<td>s5</td>
<td>1.77</td>
<td>2.42</td>
<td>1.92</td>
<td>1.56</td>
<td>1.16</td>
</tr>
</tbody>
</table>

The responses when only the lamp \#7 was driven are illustrated in Fig. 9 and the fitted rates of temperature are summarized in Table 3.

Table 3. Temperature rate when driven by Lamp \#7.

<table>
<thead>
<tr>
<th>( s_i )</th>
<th>(-100^\circ C)</th>
<th>(-170^\circ C)</th>
<th>(-220^\circ C)</th>
<th>(-270^\circ C)</th>
<th>(270^\circ C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1</td>
<td>1.99</td>
<td>2.03</td>
<td>1.75</td>
<td>1.54</td>
<td>1.17</td>
</tr>
<tr>
<td>s2</td>
<td>2.70</td>
<td>2.60</td>
<td>2.06</td>
<td>1.75</td>
<td>1.33</td>
</tr>
<tr>
<td>s3</td>
<td>3.15</td>
<td>2.73</td>
<td>2.38</td>
<td>1.87</td>
<td>1.50</td>
</tr>
<tr>
<td>s4</td>
<td>4.32</td>
<td>3.10</td>
<td>2.31</td>
<td>1.93</td>
<td>1.60</td>
</tr>
<tr>
<td>s5</td>
<td>4.48</td>
<td>3.15</td>
<td>2.44</td>
<td>1.99</td>
<td>1.57</td>
</tr>
</tbody>
</table>

The reference temperature recipe and the dwell periods are given in Table 4 and Table 5, respectively.

Table 4. Reference for temperature rate.

<table>
<thead>
<tr>
<th>Temperature range</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-100^\circ C)</td>
<td>(2^\circ C/s)</td>
</tr>
<tr>
<td>(100^\circ C \sim \sim 220^\circ C)</td>
<td>(1.5^\circ C/s)</td>
</tr>
<tr>
<td>(220^\circ C \sim \sim 330^\circ C)</td>
<td>(1.0^\circ C/s)</td>
</tr>
</tbody>
</table>

Table 5. Reference for dwell periods

<table>
<thead>
<tr>
<th>Temperature range</th>
<th>Dwell duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>(100^\circ C)</td>
<td>10 s</td>
</tr>
<tr>
<td>(165^\circ C)</td>
<td>15 s</td>
</tr>
<tr>
<td>(220^\circ C)</td>
<td>20 s</td>
</tr>
<tr>
<td>(285^\circ C)</td>
<td>25 s</td>
</tr>
</tbody>
</table>

According to the procedure illustrated in the step \#2, as given in Table 6, the first trial input to each lamp was determined from the optimization. The inputs to the outer lamps are larger than the inner lamps due to the edge effect.

Table 6. Lamp output rate for the first trial

<table>
<thead>
<tr>
<th>( h_i )</th>
<th>(-100^\circ C)</th>
<th>(-170^\circ C)</th>
<th>(-220^\circ C)</th>
<th>(-270^\circ C)</th>
<th>(270^\circ C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>0.200</td>
<td>0.212</td>
<td>0.308</td>
<td>0.279</td>
<td>0.373</td>
</tr>
<tr>
<td>h2</td>
<td>0.200</td>
<td>0.200</td>
<td>0.267</td>
<td>0.261</td>
<td>0.421</td>
</tr>
<tr>
<td>h3</td>
<td>0.200</td>
<td>0.200</td>
<td>0.234</td>
<td>0.260</td>
<td>0.472</td>
</tr>
<tr>
<td>h4</td>
<td>0.305</td>
<td>0.238</td>
<td>0.326</td>
<td>0.297</td>
<td>0.432</td>
</tr>
<tr>
<td>h5</td>
<td>0.516</td>
<td>0.333</td>
<td>0.418</td>
<td>0.323</td>
<td>0.417</td>
</tr>
<tr>
<td>h6</td>
<td>0.642</td>
<td>0.420</td>
<td>0.462</td>
<td>0.343</td>
<td>0.382</td>
</tr>
<tr>
<td>h7</td>
<td>0.660</td>
<td>0.512</td>
<td>0.553</td>
<td>0.376</td>
<td>0.348</td>
</tr>
</tbody>
</table>
To improve the uniformity, the optimal iterative learning algorithm illustrated in the step #3 was employed and the control inputs to the lamps were modified. Figure 11 illustrates the effect of the input update. The maximum uniformity error was reduced to ±8.1°C.

As repeating the update of control inputs, the maximum uniformity error was decreased gradually, and the results from the 5-th run are shown in Fig. 12. Where the temperature uniformity during ramp and dwell was improved and its maximum error was reduced to be ±4.0°C. This error meets the permissible error ±5.0°C. From the successive experiments, a temperature nonuniformity of less than ±2.0°C was achieved, but further reduction of nonuniformity was not attained.

Consequently, from the above experimental results, we can conclude that the proposed optimal iterative learning gain scheduled feedforward control is effective for temperature uniformity in vacuum reflow soldering process.

5. Conclusion

Addressing a wafer level packaging, an optimal run-to-run iterative learning gain scheduled feedforward control of wafer temperature uniformity in vacuum reflow BGA soldering process has been investigated in consideration of nearly identical environmental conditions of each soldering process and difficulties in measuring temperature of wafer in real soldering process. From the experimental results, we verified the feasibility of the proposed control for temperature uniformity. Along with run-to-run sequential experiments, gradual improvement of temperature uniformity across a wafer surface has been found, and we achieved a temperature nonuniformity of less than ±2.0°C.

Acknowledgement

This work was supported by grant no (No. 2010-0008252) from the Basic Research Program of the Korea Science & Engineering Foundation.

References


