Characterization of Low-temperature SU-8 Negative Photoresist Processing for MEMS Applications

(Invited Paper)

Gary S. May[a]
School of Electrical and Computer Engineering, Georgia Institute of Technology, 791 Atlantic Dr. NW. Atlanta, 30332, GA

Seung-soo Han
Department of Information Engineering, Myongji University, Nam-dong, Yongin-si, Gyeonggi 449-728, Korea

Sang Jeen Hong
Department of Electronic Engineering and Nano-Bio Research Center, Myongji University, Nam-dong, Yongin-si, Gyeonggi 449-728, Korea

[a] E-mail: gary.may@ece.gatech.edu

(Received August 16 2005, Accepted August 22 2005)

In this paper, negative SU-8 photoresist processed at low temperature is characterized in terms of delamination. Based on a 3^3 factorial designed experiment, 27 samples are fabricated, and the degree of delamination is measured for each. In addition, nine samples are fabricated for the purpose of verification. Employing the neural network modeling technique, a process model is established, and response surfaces are generated to investigate degree of delamination associated with three process parameters: post exposure bake (PEB) temperature, PEB time, and exposure energy. From the response surfaces generated, two significant parameters associated with delamination are identified, and their effects on delamination are analyzed. Higher PEB temperature at a fixed PEB time results in a greater degree of delamination. In addition, a higher dose of exposure energy lowers the temperature at which the delamination begins and also results in a larger degree of delamination. These results identify acceptable ranges of the three process variables to avoid delamination of SU-8 film, which in turn might lead to potential defects in MEMS device fabrication.

Keywords : MEMS, SU-8, Design of experiment, Neural networks, CTE mismatch

1. INTRODUCTION

Among numerous polymers being used in the development and fabrication of MEMS devices, the popularity of SU-8 has increased because of its mechanical stability, biocompatibility, and suitability for fabricating high aspect ratio features[1-3]. SU-8 is a negative near–UV photoresist designed to produce uniform thick films in a single spin-coating step. Vertical sidewalls and high aspect ratio features result from the product of photochemical and thermal cationic processes. The exposed and subsequently cross-linked portions of the film are rendered insoluble to liquid developers. SU-8 has low optical absorption, thus allowing the patterning of very thick films.

However, standard recipes suggested for SU-8 processing have proven in practice to be very sensitive to process conditions, and the parameter values described in the literature have varied over a wide range[3,4]. For these reasons, previous efforts at characterization and optimization of SU-8 processing have employed statistically designed experiments[5,6], and the results have suggested optimal processing parameters for various thickness of SU-8 films.

Despite the advantages of SU-8, previous studies have reported delamination of the SU-8 microstructures and films. The failure of microposts in[1] was due to an interfacial fracture at the base, as no failure occurred in the micropost bodies. Mechanical delamination of SU-8 has also been observed in MEMS drug delivery devices

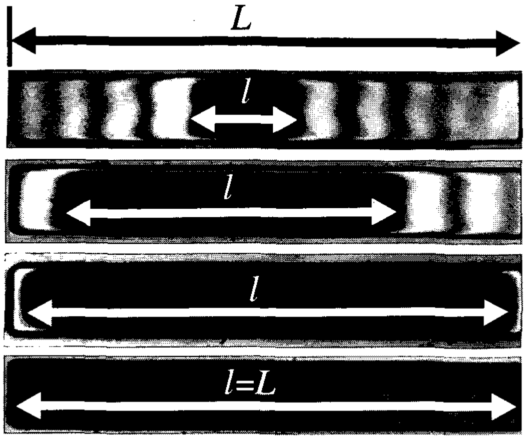

Fig. 1. Examples of different degrees of delamination (photos taken with Olympus Vanox microscope).

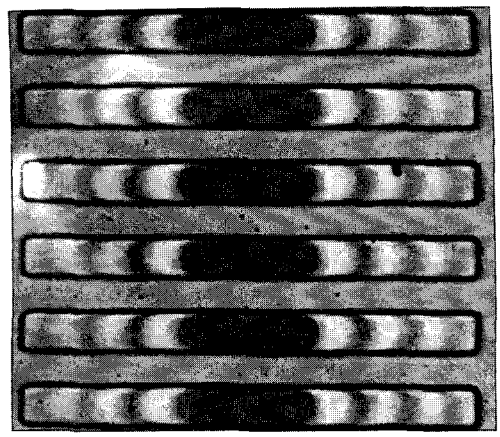

Fig. 2. Bar patterns used for the measurement of degree of delamination.

[2]. Brunet *et al.* also reported the delamination of the SU-8 microstructures during developing in the development of high aspect ratio magnetic coils. Thick layers of SU-8 experienced more stress, and the structures tended to delaminate more quickly than thin layers[3].

For multi-layer MEMS fabrication, which is currently under investigation, delamination associated with stress has led to concerns about defective SU-8 fabrication. In an effort to reduce the amount of stress on SU-8 microstructures, a low-temperature process with prolonged bake time has been investigated. By trial and error, delamination was reduced. However, it is necessary to perform a more systematic characterization experiment to clarify the relationship between process parameters and identify suitable ranges for process variables to ensure the fabrication without delamination. Therefore, this paper investigates the variation of low-temperature SU-8 processing, with the ultimate goal of minimizing delamination, using response surfaces generated from neural network models.

The paper is organized as follows: Section 2 describes how the experiment was performed for the 100 μm thick SU-8 films. Section 3 provides background information on neural network modeling. Results are provided in Section 4, followed by a summary and discussion of future work in the final section.

2. EXPERIMENT

2.1 Statistical experimental design

If a process has more than a very small number of steps whose possible values have a large range, the number of experiments needed for process characterization can be prohibitively large. In addition, the role of each step in determining the final outcome is generally not clear. The traditional method of collecting large quantities of data by holding each factor constant in turn until all possibilities have been tested is an approach that quickly becomes impossible as the number of factors increases. Statistical experimental design is a systematic and efficient alternative methodology for characterization and modeling using a relatively small number of experiments[7].

In this study, five parameters at three levels each were initially considered. The parameters were soft baking temperature/time, exposure energy, and post exposure baking (PEB) temperature/time. Since cross-linking takes place after exposure, the variables in the soft baking step were later omitted. Instead, the soft baking step was performed in a consistent manner for all samples. Develop time after PEB plays an important role in adhesion to the surface. Extended development time may increase the chance of delamination of the exposed area from the substrate, but insufficient time may negatively impact the lithographic resolution[8]. In this research, a development time that allowed decent lithographic resolution was consistently used in order to avoid any additional complexity in characterization. The process variables and their ranges appear in Table 1. A 3^3 factorial design requiring 27 experiments was conducted, and this design was further augmented with nine randomly selected experiments for model verification purposes.

2.2 Sample fabrication and measurement

SU-8 was spin coated on 4" silicon wafers to a thickness of 100 um, and the samples were soft baked at 70 °C on a hot plate to drive off solvents. Based on the designed experiment, all possible orthogonal combinations of three parameters were applied. All samples were developed for a fixed time, and the degree of delamination was measured. The degree of delami-

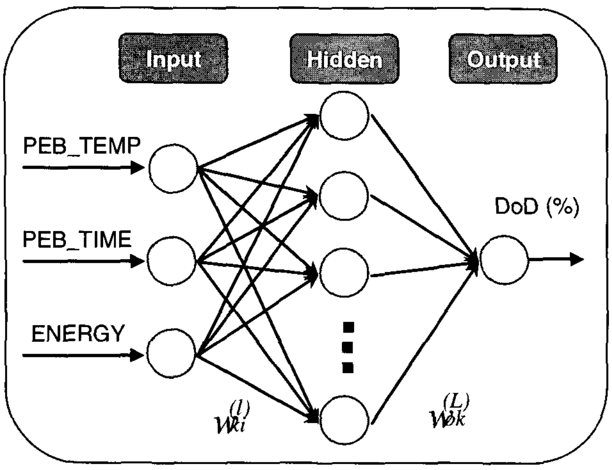

Fig. 3. An illustration of a multilayer perceptron neural network.

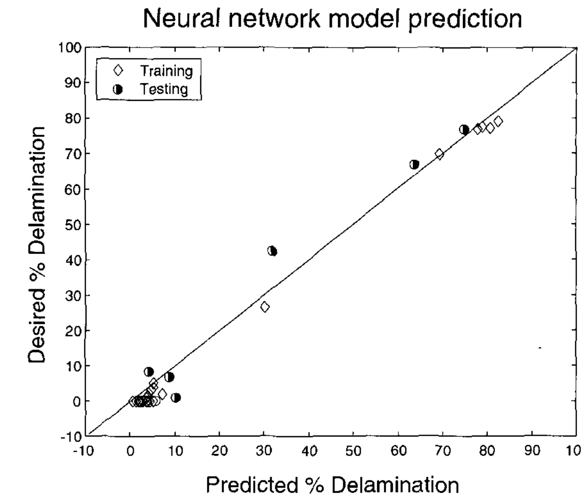

Fig. 4. Performance of neural process model. Straight line represents 100 % accuracy.

nation (DoD) was quantified by the following expression:

$$\mathrm{DoD} = \left(\frac{(L-l)}{L} \right) \times 100\ (\%) \tag{1}$$

where L is the length of the original bar pattern, and l is the length of bar pattern that remained on substrate (see Fig. 1). To minimize measurement error, l was averaged over eight bar patterns in one location as shown in Fig. 2.

3. NEURAL NETWORKS

Neural networks have become useful tools in process modeling and demonstrated the capability of learning complex relationships between groups of related parameters[9]. A neural network is a structured interconnection of computational nodes called *neurons* that contribute to parallel computation in a manner similar to the human brain. The interconnection of neurons establishes knowledge that is acquired by the network through a learning process, and that knowledge is stored in the form of inter-neuron connection strengths known as *weights*. Each neuron contains the weighted sum of its inputs filtered by a sigmoidal "squashing" function, providing neural networks with the ability to generalize with an added degree of freedom that is not available in statistical regression techniques.

The learning algorithm used in this study is the error back-propagation (BP) algorithm. A typical back-propagation neural network structure is depicted in Fig. 3. In the BP learning algorithm, a single iteration consists of two parts: a forward and a backward propagation. In the forward propagation, the outputs from the i^{th} layer are weighted and summed, and the weighted sum is filtered through a sigmoid function. The outputs of neurons in j^{th} layer become inputs to the neurons in the next layer k. The forward propagation is described by the following equations:

$$s_j = \sum_{i=1}^{n_i} w_{ji} o_i \tag{2}$$

$$o_j = \frac{1}{1+\exp(-s_j)} \tag{3}$$

where s_j are weighted sum input to neurons in layer j, o_i is output from neurons in layer i, n_i is the number of neurons in the i^{th} layer, and w_{ji} is the weight connecting neurons i and j. In the same manner, y_k, s_k, o_i, and s_i also can be derived. In backward propagation, weights are updated in the direction that minimizes an error function defined by:

$$E_k = \frac{1}{2}(\hat{y}_k - y_k)^2 \tag{4}$$

where $\hat{y}_k$ is a target, and y_k is the actual output value of the last layer k. The *generalized delta-rule* based on *gradient descent* approach is applied to minimize the error function. The weights are initially randomized, and forward propagation is performed. Once the outputs of the last layer are calculated, weights are updated by the weight changes for each node calculated from the output layer (layer k) and back-propagated to the input layer (layer i). The *generalized delta rule* is:

Table 1. Process parameters and ranges.

Step	Parameters	Abbrev.	Ranges	Units
Exposure	Energy	ENERGY	440-580-720	mJ/cm^2
PEB	Temp.	PEB TMP	60-70-80	°C
	Time	PEB TIME	20-30-40	min.

$$\Delta w_{ji}(n+1) = \eta \delta_j (n+1) o_i (n+1) \qquad (5)$$

$$w_{ji}(n+1) = \eta \Delta w_{ji}(n+1) + \alpha \Delta w_{ji}(n) \qquad (6)$$

where n is the number of iteration, η is the *learning rate* and α is the *momentum*. The learning rate is a constant that represents the rate at which a weight will be changed along its slope to the minimum error. The momentum is a constant that includes a portion of the previous weight change to the current weights.

Utilizing *ObOrNNs*[10], a custom neural network simulation package, neural network based response surface models of the SU-8 fabrication process were derived. Initially, neural networks were trained with the data generated from the 3^3 factorial designed experiments, and these models were verified with data that was not previously introduced to the networks during training. Inputs to the networks were the three parameters of interest (exposure energy, post-bake temperature, and post-bake time), and the output of the networks was the degree of delamination. "Hidden" neurons (neurons in the middle layers) extract nonlinear features from the data, and several networks with different numbers of hidden neurons were constructed and tested. The average RMS error in training was 2.57 %, and that in testing was 4.87 %. Model performance is depicted graphically in Fig. 4.

4. RESULTS AND DISCUSSION

Once the neural process model was established, response surfaces were generated to illustrate the relation-

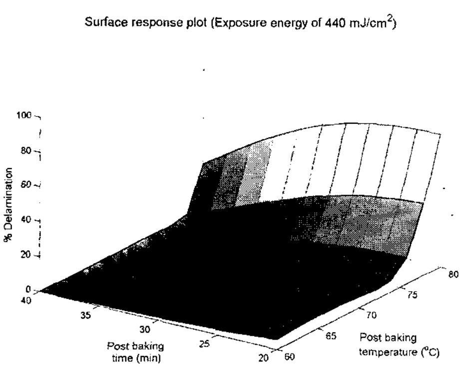

Fig. 5. Response surface plot: fixed exposure energy at 440 mJ/cm^2.

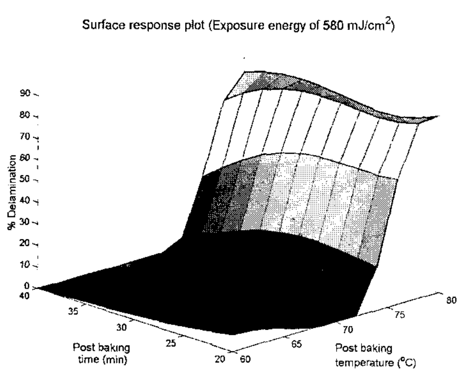

Fig. 6. Response surface plot: fixed exposure energy at 580 mJ/cm^2.

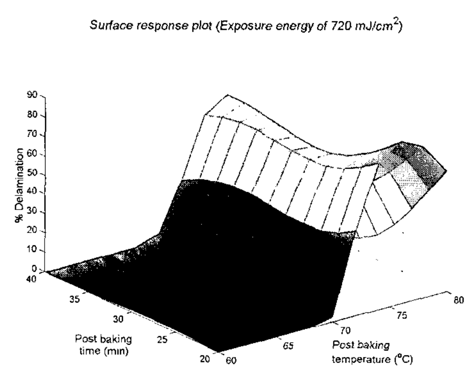

Fig. 7. Response surface plot: fixed exposure energy at 720 mJ/cm^2.

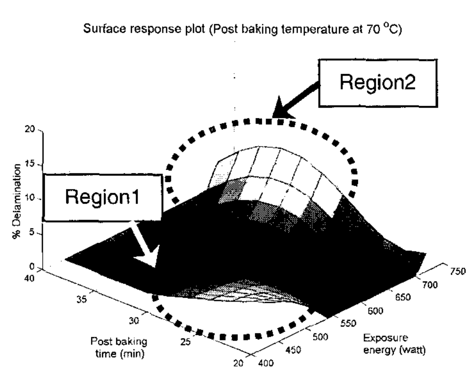

Fig. 8. Response surface plot: fixed post exposure baking temperature at 70 °C.

ships between the process parameters and degree of delamination. Any two of the three process parameters were simultaneously varied within the ranges in Table 1, while the remaining parameter was set at its mid-range level. The predictions of the neural process model were then graphed using 3-D contour plots (see Figs. 5-8).

Figure 5 illustrates the effect of PEB time and temperature on the degree of delamination when the exposure energy is fixed at 440 mJ/cm^2. The energies are set at 580 mJ/cm^2 and 720 mJ/cm^2 in Figs. 6 and 7, respectively. In this experiment, PEB temperature appeared to be the most critical factor affecting delamination under the condition that the dose of energy is fixed. To ensure cross-linking of SU-8 after exposure, enough PEB time is required at the proper temperature. It is observed, however, that the higher the temperature, the larger the degree of delamination. In addition, PEB time also somewhat affects the degree of delamination at a given temperature. The shorter the PEB time, the less cross-linking, increasing the degree of delamination after a certain temperature. The high degree of delamination at temperatures above 70-75 °C is primarily due to the coefficient of thermal expansion (CTE) mismatch between SU-8 and the silicon wafer with native oxidation.

As the exposure energy increases, the degree of delamination increases, while the temperature at which the delamination starts to occur decreases. Higher exposure energy tends to increase cross-linking of the polymer in the exposed area, and consequently, this increases film stress due to volume changes. The effect of exposure energy on delamination can be observed clearly in Fig. 8. By setting the temperature at a certain level, the degree of delamination caused by CTE mismatch can be neglected. At a fixed PEB temperature of 70 °C and a reasonable PEB time of 25-40 minutes, the degree of delamination increases with exposure energy. Region 1, where the PEB time is less than 30 minutes and the exposure energy is less than 520 mJ/cm^2, showed some degree of delamination due to incomplete cross-linking. Region 2, where the PEB time is longer than 25 minutes and the exposure energy is larger than 650 mJ/cm^2, shows approximately 5 % delamination due to the stress induced by volume changes.

5. CONCLUSION

To summarize, two significant parameters associated with SU-8 delamination were investigated, and their effects on delamination determined found from the response surfaces generated from neural network models. Higher PEB temperatures at a fixed PEB time result in more delamination due to CTE mismatch. In addition, a greater dose of exposure energy lowers the temperature at which delamination starts to occur and increases degree of delamination. The response surfaces generated also identify suitable ranges of process conditions that avoid SU-8 delamination, which can ultimately cause defects in MEMS devices.

REFERENCES

[1] H. Khoo, K. Liu, and F. Tseng, "Mechanical strength and interfacial failure analysis of cantilevered SU-8 microposts", J. Micromech. Microeng., Vol. 13, p. 822, 2003.

[2] G. Voskerician, M. Shive, R. Shawgo, H. Recum, J. Anderson, M. Cima, and R. Langer, "Biocompatibility and biofouling of MEMS drug delivery devices", Biomaterials, Vol. 24, p. 1959, 2003.

[3] M. Brunet, T. O'Donnell, J. O'Brien, P. McCloskey, and S. Mathuna, "Thick photoresist development for the fabrication of high aspect ratio magnetic coils", J. Micromech. Microeng., Vol. 12, p. 444, 2002.

[4] M. Despont, H. Lorenz, N. Fahrni, J. Brugger, P. Renaud, and P. Vettiger, "High-aspect ratio ultrathick, negative tone near-UV photoresist for MSMS applications", in Proc. IEEE, Tenth Anuual Internaltional Workshop in Micro Eletro Mechanical Systems, Nagoya, Japan, p. 518, 1997.

[5] J. Zhang, K. L. Tan, and H. Q. Gong, "Characterization of the polymerization of SU-8 photoresist and its applications in micro-electro-mechanical systems", Polymer Testing, Vol. 20, p. 693, 2001.

[6] B. Eyre, J. Blosiu, and D. Wiberg, "Taguchi optimization for the processing of Epon SU-8 resist", in Proc. IEEE, The eleventh Annual International Workshop on Micro Eelectro Mechanical Systems, p. 25, 1998.

[7] G. Box, W. Hunter, and J. Hunter, "Statistics for Experimenters", New York: Wiley, 1978.

[8] A. Wong and D. Linton, "Application of SU-8 in flip chip bump micromachining for millimeter wave applications", in Proceedings of 3rd Electronics Packaging Technology Conference (EPTC 2000), p. 204, 2000.

[9] S. Hong, G. May, and D. Park, "Neural network modeling of reactive ion etch using optical emission spectroscopy data", IEEE Trans. Semi. Manufac., Vol. 16, No. 4, p. 1, 2003.

[10] C. Davis, S. Hong, R. Setia, R. Pratap, T. Brown, B. Ku, G. Triplett, and G. May, "A java_based objected-oriented neural network simulator for semiconductor manufacturing applications", submitted to the 8th World Multi-conference on Systemics, Cybernetics and Informatics, Orlando, FL, 2004.