A Statistical Shape Model and SVMs based Scheme for Visual Inspection of Microdrill Bits in PCB Production

Guifang Duan and Yen-Wei Chen
College of Information Science and Engineering, Ritsumeikan University
1-1-1 Noji-higashi, Kusatsu, Shiga 525-8577 JAPAN

Abstract

This paper proposes an automatic visual inspection scheme with phase identification of microdrill bits in printed circuit board (PCB) production. Our method mainly includes two procedures: firstly the statistical shape models of microdrill bit is built to get the shape subspace, and then the phase identification is performed in the shape subspace using some pattern recognition techniques. In this paper, we compared the performance of two statistical model methods, principal component analysis (PCA) and linear discriminate analysis (LDA) together with three classifiers, support vector machines (SVMs), neural networks (NNs) and k-nearest neighbors (kNN) respectively for phase identification of microdrill bits. The experimental results demonstrate that using low enlargement and resolution microdrill bit images the proposed method can measure up to high inspection accuracy, and provide an conclusion that the highest identification rates were obtained by PCA-SVMs.

1. Introduction

With the rapid growth of Printed Circuit Board (PCB) manufacturing industry, automatic inspection of microdrill bits has been more and more important. A worn-out microdrill damages the quality of PCB surface finish and the dimensions of the drilled hole [1]. Since the increasing circuit density brings about continuing microminiaturization of drill bits, the inspection has come into an enormous challenge. In the face of the microdrill bits with a diameter of just one-tenth or even one-hundredth millimeter, it is obviously impossible to satisfy the requirement by the naked eye. One solution is that the quality auditors work with micrometers and microscopes. Such kind of manual inspection is time-consuming, subjective and costly [3], thus, it is difficult to achieve the required inspection. Another solution, called automatic visual inspection, has been proved a promising way. It it is time-saving, objective and non-contact [3, 4,].

A lot of work has been done on the automatic visual inspection of microdrill bits by machine vision techniques. Some methods for geometric defect inspection have been presented by the authors in [2, 5, 6]. All these methods work in the similar scheme. Firstly edge detection or boundary following technique was employed to get the edge of cutting plane from a microdrill bit image, and then corner detection and curve fitting were implemented to measure some geometric quantities of the cutting planeOn the other hand, a scheme for flank wear measurement of cutting plane was reported in [1]. Using a set of pair points located on the raising edge and the falling edge, three quantities, flank wear area, average flank wear height, and maximum wear height, were calculated to measure the flank wear for tool life evaluation.

All the methods mentioned above adopt the strategy of "one-to-one" comparison. To guarantee measurement accuracy, the inspection instrument requires large enlargement lens (more than 300 times) and high resolution CCD (more than 640 by 480) [1, 2, 5, 6], which increases the cost of both system hardware and system computation. This paper in further develops a novel method that relatively lowers the requirement for imaging equipment of inspection instrument. The proposed method adopts the phases of microdrill bits lifecycle as an index for inspection. Generally during a microdrill bit lifecycle in PCB production there are six important phases which are denoted as N, ND, K1, K1D, K2 and K2D (examples are shown in Fig. 1). During inspection, if these phases can be identified, we can easily confirm its condition and evaluate the tool life. Thus, the inspection can be viewed as a six-class classification problem to identify the phases of microdrill bits. The basic idea of the proposed method is that using a set of training samples of microdrill bits, the statistical shape models are built first, and phase identification of microdrill bits follows in the shape subspace using a classifier. This paper demonstrates how the scheme works for phase identification and investigates the performance of two statistical model methods, principle component analysis (PCA) [8] and linear discriminate analysis (LDA) [9] together with three classifiers, support vector machines (SVMs) [10], neural networks (NNs) [11] and k-nearest neighbors (kNN) [11] respectively for phase identification of microdrill bits.

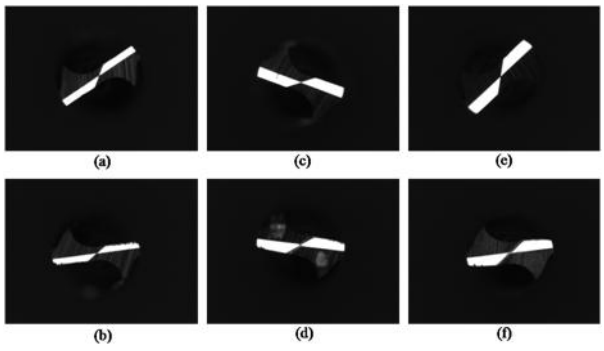


Figure 1. 0.5mm-diameter microdrill bit images of the six phases. (a) N, (b) ND, (c) K1, (d) K1D, (e) K2, (e). K2D, (f) K2.

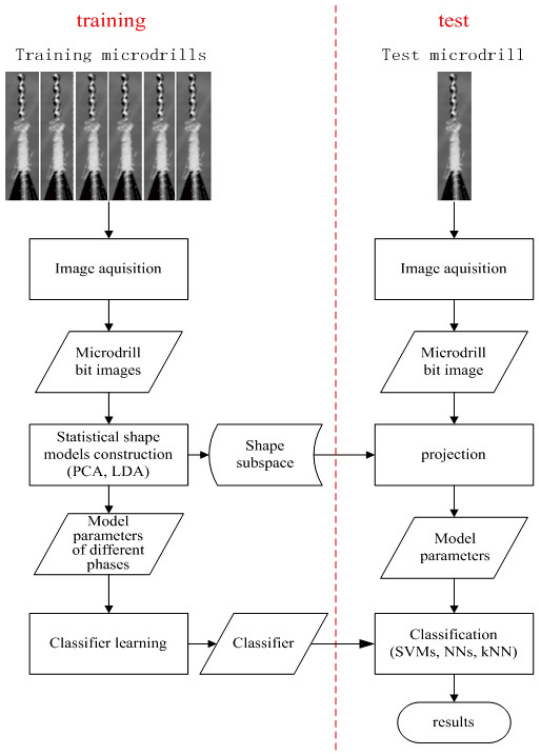


Figure 2. Flowchart of the phase indemnification of microdrill bits.

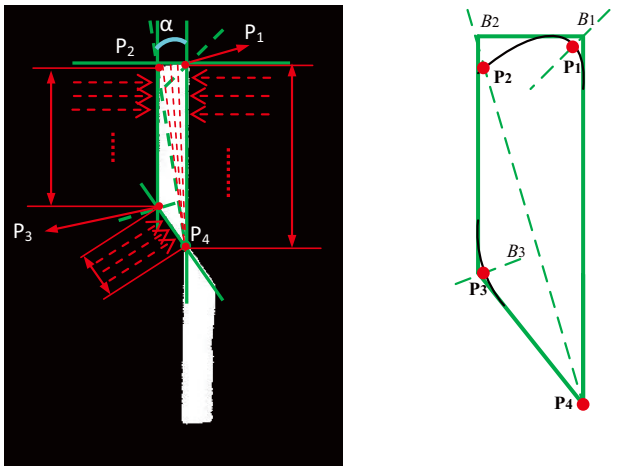
2. Overview of the Visual Inspection Scheme

The flowchart of our proposed scheme for phase identification of microdrill bits, which consists of training stage and test stage, is depicted in Fig. 2. Training stage includes image acquisition and statistical shape models construction. Images of microdrill bits are acquired by a special imaging system, and then they are used as training set to build the statistical shape models by PCA or LDA. For a test image, we project the shape of the microdrill bit to the shape subspace that we trained to obtain its model parameters. These parameters are taken as input of classifiers in the shape subspace for phase identification.

2.1. Image Acquisition

The image is captured by the CCD camera with a LED illumination system in frontal view of the drill bit in the scanner. For intuitional instruction, in the right side we give an example of the acquired image. The image is transformed into an eight-bit grayscale format (as shown in Fig. 1) and saved to the memory of the computer. Each digital image is 256 by 340 pixels in size. Both the magnification of lens and the resolution of CCD are much lower than those (mentioned above) commonly used in research and practice. The intensity value of each pixel ranges from 0 to 255. For the originally acquired images, an image thresholding according to histogram is implemented to extract cutting plane from the background. Then the cutting plane is aligned to horizontal direction.

2.2. Automatic Statistical Shape Models Construction



(a) (b)

Figure. 3. Cutting plane land marking

A) Automatic cutting plane land marking

For one of the cutting planes, as shown in Fig. 3 (a), the minimum quadrangle that can enclose the facet is determined first. The oblique bottom side of the quadrangle locates on the chisel edge, and the top side horizontally passes through the top of the facet. The right and left sides are a pair of vertical lines through the corresponding boundaries of the facet, respectively.

As shown in Fig. 3 (b), in the quadrangle mentioned above, the angular bisectors of the upper right vertex angle and the lower left vertex angle are termed as B_1 and B_3. The line, B_2, joins the two points, the upper left vertex and the lower right vertex. Then the intersecting points of B_1, B_2 and B_3 with the boundary of the cutting plane are defined as the basic points, P_1, P_2 and P_3, respectively. The lower right vertex is defined as the basic point P_4.

As shown in Fig. 3 (a), the intersecting points, of a set of lines that divide the angle α equally and the boundary of the cutting plane, are selected as the land marks between P_1 and P_2. The landmarks between P_1 and P_4 are defined by the vertically equal divisions of the boundary segment P_1P_4, and the landmarks on the boundaries P_2P_3 and P_3P_4 are defined in the same approach.

B) Statistical analysis

Given a shape of the cutting plane, it can be represented as a vector

$$\mathbf{s}_i = \left[x_1, y_1, \cdots, x_m, y_m\right]^T, i = 1, \ldots, n \qquad (1)$$

where (x_i, y_i) denotes the location of a landmark, m is the number of landmarks and n is the number of training examples. The shape model is then

$$\mathbf{s} = \bar{\mathbf{s}} + \mathbf{Pb}, \qquad (2)$$

where $\bar{\mathbf{s}}$ is the mean shape, $\mathbf{P} = [\mathbf{v}_1|\mathbf{v}_2|\ldots|\mathbf{v}_t]$, is the orthonormal transformation matrix which forms the shape subspace, and $\mathbf{b}$ is a vector of weights, called as model parameters. For a test drill bit shape $\mathbf{s}_{test}$, its model parameters is given by projecting it onto the shape subspace as

$$\mathbf{b}_{test} = \mathbf{P}^T\left(\mathbf{s}_{test} - \bar{\mathbf{s}}\right). \qquad (3)$$

In this paper, two popular methods, PCA and LDA are applied for dimensionality reduction and shape model construction.

i. PCA, to find a lower dimensional subspace corres-

ponding to the maximum variance direction of original multidimensional data space[12]. In PCA, the basis vectors are defined as eigenvectors of scatter matrix $\mathbf{\Phi}\mathbf{\Phi}^T$. $\mathbf{\Phi}$ is the normalized shape matrix

$$\mathbf{\Phi} = [(\mathbf{s_1} - \overline{\mathbf{s}})|\ldots|(\mathbf{s}_m - \overline{\mathbf{s}})]. \quad (4)$$

The orthonormal transformation matrix $\mathbf{P}$ is formed by the basis vectors corresponding to the t largest eigenvalues of $\mathbf{\Phi}\mathbf{\Phi}^T$. The principle to set t is

$$\frac{\sum_{i=1}^{t}\lambda_i}{\sum_{i=1}^{n}\lambda_i} \geq \gamma, \quad (5)$$

where γ is a threshold that defines the proportion of the total variation to be explained.

ii. LDA, also known as Fisher Discriminate Analysis (FDA), the same with PCA is also an orthogonal linear transformation. However, the difference from PCA is that LDA searches a lower dimensional subspace that best discriminates among classes [9], that is to say P is decided to make the ratio of the between-class scatter and the within-class scatter is maximized. The between-class scatter matrix and within-class scatter matrix are defined as

$$\mathbf{S}_B = \sum_{i=1}^{c} N_i \left(\overline{\mathbf{s}}_i - \overline{\mathbf{s}}\right)\left(\overline{\mathbf{s}}_i - \overline{\mathbf{s}}\right)^T, \quad (6)$$

$$\mathbf{S}_W = \sum_{i=1}^{c}\sum_{j=1}^{N_i}\left(\mathbf{s}_j^i - \overline{\mathbf{s}}_i\right)\left(\mathbf{s}_j^i - \overline{\mathbf{s}}_i\right)^T. \quad (7)$$

where c is the total class number, N_i is the training example number of class i, and $\overline{\mathbf{s}}_i$ is the mean vector of class i. Thus, P is chosen to maximize the ratio $det(\mathbf{S}_B)/det(\mathbf{S}_W)$, which contains a set of generalized eigenvectors of $\mathbf{S}_B$ and $\mathbf{S}_W$ corresponding to the t largest eigenvalues, i.e.,

$$\mathbf{S}_B\mathbf{p}_i = \lambda_i\mathbf{S}_W\mathbf{p}_i, \qquad i = 1,2,\ldots t. \quad (8)$$

where $\mathbf{p}_i$ is a generalized eigenvector corresponding to the eigenvalue λ_i. An upper bound on t is c-1, because there are at most c-1 nonzero generalized eigenvectors.

2.3. Phase Identification of Microdrill Bits

In our methods, the phase identification is implemented in the shape subspaces with model parameters, and SVMs are selected as the classifier. For benchmark comparisons with other classification methods, experiments were also performed with NN and kNN.

SVMs were originally proposed by Vapnik in 1995, and have been widely used in pattern classification applications [10]. It has been shown that SVMs can yield superior performance than traditional techniques for its higher ability in generalization [7]. In the following, we will give a brief introduction to SVMs.

SVMs belong to the class of maximum margin classifiers [10, 12]. From a training dataset, a binary classifier is constructed to perform pattern recognition between two classes. The purpose is to find a decision surface so that the margin of the nearest points which are termed as support vectors (SVs) in the training set is maximized. This decision surface is called the optimal separating hyperplane (OSH). Given a training dataset $\{\mathbf{x}_i, y_i\}$, $i = 1, 2,\ldots l$, $\mathbf{x}_i \in \mathbb{R}^n$, in two classes, and the label $y_i \in \{-1, +1\}$, the OSH has the form

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{l} y_i\alpha_i\mathbf{x}_i \cdot \mathbf{x} + b\right), \quad (9)$$

where b is bias, and α_i are coefficients which can be determined by solving a quadratic programming problem:

$$\text{minimize} \quad \frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\mathbf{Q}_{ij}\alpha_j - \sum_{i=1}^{l}\alpha_i \quad (10)$$

$$\text{subject to} \quad \sum_{i=1}^{l} y_i\alpha_i = 0, 0 \leq \alpha_i \leq C \text{ for } i = 1,\ldots,l. \quad (11)$$

where $\mathbf{Q}$ is an l by l semidefined matrix that depends on the training dataset and the SVMs function form, $\mathbf{Q}_{ij} = y_iy_j(\mathbf{x}_i \cdot \mathbf{x}_j)$. C is the upper bound, a user specified regularization perimeter, and a larger C means a higher penalty to the training errors.

However, there are a lot of non-separable cases for linear classifier in general pattern recognition problems, and it is also difficult to obtain a nonlinear decision function. A hopeful solution is to map the input dataset to a higher-dimensional space called feature space where the OSH can be found for linear separation. The feature space should be a dot product space, that is to say, by the mapping Φ, the dot production of two points $\Phi(\mathbf{x})\cdot\Phi(\mathbf{y})$ in the feature space can be described by a kernel function $K(\mathbf{x}, \mathbf{y})$:

$$K(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x})\cdot\Phi(\mathbf{y}). \quad (12)$$

According to Mercer's theorem [13], given a symmetric positive kernel K, there exists a mapping Φ that satisfies equation (13). Thereby we can construct a nonlinear SVM without treating the mapping Φ explicitly. The kernel replaces the inner production, then the decision function becomes

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{l} y_i\alpha_i K(\mathbf{x}_i \cdot \mathbf{x}) + b\right). \quad (13)$$

The solutions of α_i are still a quadratic programming problem.

In this work we adopted an important family of kernel functions, Gaussian radial basis function (RBF) as follows

$$K_{RBF}(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{1}{2\sigma^2}\|\mathbf{x} - \mathbf{y}\|^2\right). \quad (14)$$

3. Experimental Results

Forty sets of 0.5 mm-diameter drill bits are used for evaluation. The shape of a cutting plane is represented by 70 labeled points. Each set involves six samples corresponding to six phases. Totally we have 240 samples. The experiments are implemented in a leave-one-out strategy. Each time one set is left for test, and the other thirty nine sets are used as training sets. The experiments are repeated 40 times so that each set can be used as new test sample for evaluation.

We compared the mix identification rates of the three classifiers in the five cases (with different principal components number) in Fig. 4. The best identification rates of

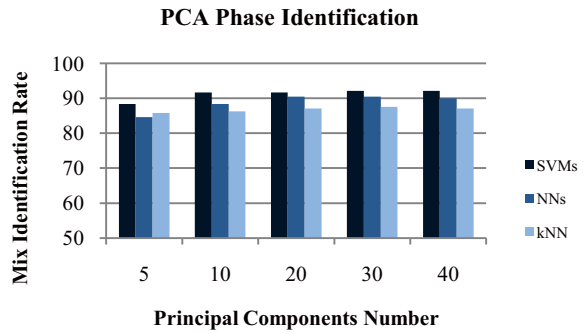


Figure 4. Mix identification rate using PCA for statistical shape construction with three classifiers, SVMs, NNs and kNN.

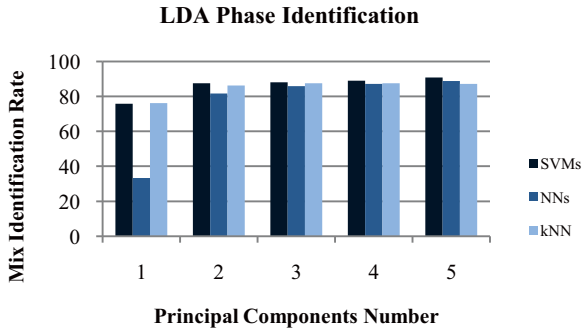


Figure 5. Mix identification rate using LDA for statistical shape construction with three classifiers, SVMs, NNs and kNN.

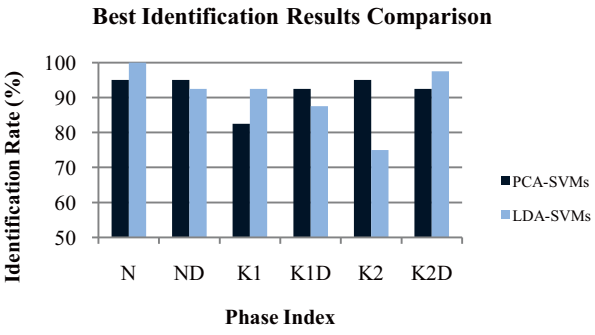


Figure 6. Identification rates using PCA-SVMs and LDA-SVMs

NNs and kNN are lower than that of SVMs. The overall best identification rate is achieved by SVMs with 30 components. Fig. 5 shows the mix identification rates of the three classifiers on the five conditions (with different dimension) by LDA. We found that the overall best identification rate is achieved by SVMs, which is the same with that using PCA for statistical shape model construction. Here, it can be said that in both the two shape subspaces, the best identification rate is achieved by SVMs. Fig. 6 shows the identification rate comparison of the two approaches.

It must be noted that the computation cost of training is high, but the test is fast. As to PCA-SVMs with 30 principal components, the training processing (include the training stage in shape model construction and the training stage in SVMs) time is 238 s and the test processing time is 0.96 s. So it is a real time inspection scheme.

4. Conclusion

An automatic visual inspection scheme with phase identification of microdrill bits in PCB production was presented. We proposed a new index, phase of lifecycle for microdrill bits inspection. Two methods (PCA and LDA) of statistical shape model construction together with three classifiers (SVMs, NNs and kNN) are implemented. The proposed scheme can identify the phase of a microdrill bit within 1s, which makes it practically in PCB production.

References

[1] J.C. Su, C.K. Huang, and Y.S. Tang, "An automatic flank wear measurement of microdrills using machine vision," *Journal of Materials Processing Technology* Vol.180, No. 1-3, pp. 328-335, 2006.

[2] F.-C. Tien, C.-H. Yeh, and K.-H. Hsieh: "Automated visual inspection for microdrills in printed circuit board production", International Journal of Production Research Vol.42, No. 12, pp.2477-2495, 2004.

[3] M. Moganti, F. Ercal, C.H. Dagli, and S. Tsunekawa, "Automatic PCB Inspection Algorithms: A Survey", *Computer Vision and Image Understanding*, Vol.63. No.2, pp.287-313, 1996.

[4] Eduardo and B. Corrochano, "Review of Automated Visual Inspection 1983 to 1993 Part I: conventional approaches," *SPIE - Intelligent Robots and Computer Vision XII* Vol.2055, pp.128-158, 1993.

[5] W.J. Zhang, D. Li, F. Ye, and H. Sun, "Automatic Optical Defect Inspection and Dimension Measurement of Drill Bit," *IEEE International Conference on Mechatronics and Automation*, pp.95-100, 2006.

[6] C.K. Huang, C.W. Liao, A.P. Huang, and Y.S. Tarng, "An automatic optical inspection of drill point defects for micro-drilling," *The International Journal of Advanced Manufacturing Technology*, Vol.37, pp.1133-1145, 2007.

[7] B. Schölkopf, K. Sung, C.J.C. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik, "Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function Classifiers," *IEEE Trans. Signal Processing*, Vol. 45, No. 11, pp. 2758-2765, 1997.

[8] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, Vol.3, No. 1, pp. 71-86, 1991.

[9] Peter N. Belhumeur, João P. Hespanha, David J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 711-720, 1997.

[10] S. Abe, *Suppot Vector Machines for Pattern Classification*, Springer-Verlag, London, 2005.

[11] Richard O. Duda, Peter E. Hart and David G. Stork, *Pattern Classification (2nd Edition)*, Wiley-Interscience, 2000.

[12] K.I. Kim, K. Jung, S.H. Park and H.J. Kim, "Support Vector Machines for Texture Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 11, pp. 1542-1550, 2002.

[13] Ulrich H.-G. Kreβel, "Pairwise classification and support vector machines," *Advances in kernel methods: support vector learning*, MIT Press, pp. 255 – 268, 1999.