Wavelet-based Principal Component Analysis Applied to Automated Surface Defect Detection

HONG-DAR LIN¹, CHUNG-YU CHUNG¹, WAN-TING LIN²

¹Department of Industrial Engineering and Management
Chaoyang University of Technology
168 Jifong E. Rd., Wufong Township, Taichung County, 41349
TAIWAN

²College of Business, University of Missouri-Columbia
213 Cornell Hall, Columbia, MO 65211
USA

Abstract: Automated visual inspection, a crucial manufacturing step, has been replacing the more time-consuming and less accurate human inspection. This research explores automated visual inspection of surface defects in a light-emitting diode (LED) chip. Commonly found on chip surface are water-spot defects which impair the appearance and functionality of LEDs. Automated inspection of water-spot defects is difficult because they have a semi-opaque appearance and a low intensity contrast with the rough exterior of the LED chip. Moreover, the defect may fall across two different background textures, which further increases detection difficulties. The one-level Haar wavelet transform is first used to decompose a chip image and extract four wavelet characteristics. Then, wavelet-based principal component analysis (WPCA) approach is applied to integrate the multiple wavelet characteristics. Finally, the principal component analysis of WPCA judges the existence of defects. Experimental results show that the proposed method achieves detection rates of above 93.8%, and false alarm rates of below 3.6%. A valid computer-aided visual defect inspection system is contributed to help meet the quality control needs of LED chip manufacturers.

Key-Words: Surface defect detection, Wavelet characteristics, Principal component analysis, Computer vision system.

1 Introduction
Quality control is designed to prevent defective products from reaching the customer. Visual inspection constitutes an important part of quality control in the industry. In most cases, quality control through visual inspection is still conducted by humans. However, difficulties exist in detecting defects by human eyes because inspectors are very likely to make erroneous judgments due to personal subjectivity or eye fatigue. Visual inspection determines product properties using visual information and is most often automated by employing machine vision techniques. Therefore, automated visual inspection of surface defects has become a critical task for manufacturers who strive to improve product quality and production efficiency [1-4]. In this study, we use machine vision techniques for automated surface inspection of light-emitting diode (LED) chips.

LED is a semiconductor device that emits visible light when an electric current passes through the semiconductor chip. Compared with incandescent and fluorescent illuminating devices, LEDs have lower power requirement, higher efficiency, and longer lifetime. Typical applications of LED components include indicator lights, LCD panel backlighting, fiber optic data transmission, etc. The basic structure of an LED consists of the light emitting semiconductor chip, a lead frame where the chip is actually placed, and the encapsulation epoxy which surrounds and protects the chip. To meet consumer and industry needs, LED products are being made in smaller sizes, which increase difficulties of product inspection.

![Fig. 1 LED chips images (a) normal chip (b) and (c) defective chips with water-spot defects](image-url)

Fig. 1 LED chips images (a) normal chip (b) and (c) defective chips with water-spot defects
With the popularity of LEDs, inspection of surface defects has become a critical task for manufacturers who strive to enhance LED product quality. Surface defects affect not only the appearances of LEDs but also their functionality, efficiency and stability. As inspecting surface defects by human eyes is ineffective and inefficient, this research aims to develop an automated vision system for detecting water-spot defects, which commonly appear on the surfaces of LED chips owing to the steam generated during the production process.

Automated inspection of a water-spot defect is difficult because the blemish has a semi-opaque appearance and a low intensity contrast with the rough exterior of the LED chip. With a width of 0.21 mm, an LED chip comprises an aluminum-pad (bonding pad) in the central area and a metal oxide semiconductor (emitting area) in the outer area, as shown in Fig. 1 (a). Texture of the central area has a random pattern while that of the outer area has a uniform appearance. A water-spot defect may fall across the two areas of significantly different textures, which complicates the defect detection procedure. Figures 1 (b)-(c) display the LED chip images with water-spot blemishes of different shapes.

Defect detection techniques compute a set of textural features in a sliding window and search for significant local deviations among the feature values. The detection techniques are generally classified into the spatial domain and the frequency domain. Siew et al. [5] applied the co-occurrence matrix method, a traditional spatial domain technique, to assess carpet wear by using two-order gray level statistics to build up probability density functions of intensity changes. For another spatial domain example, Latif-Amet et al. [6] presented wavelet theory and co-occurrence matrices for detection of defects encountered in textile images and classified each sub-window as defective or non-defective with a Mahalanobis distance.

As to techniques in the frequency domain, Chan and Pang [7] proposed a simulated fabric model based on Fourier transform for inspection of structural defects in fabric. Since a three-dimensional frequency spectrum is very difficult to analyze and defects occur mostly along the horizontal and vertical axes, the central spatial frequency spectrum approach has been proposed to increase efficiency of the analysis process. Seven significant characteristic parameters can be extracted from the central frequency spectrums for describing the defect types. Kumar and Pang [8] presented a new multi-channel filtering scheme for unsupervised fabric defect detection using a class of self-similar Gabor functions. Also, Lin [9] developed a novel approach that applies discrete cosine transform decomposition and cumulative sum techniques for the detection of tiny defects on passive component chips.

Regarding defect detection applications in the electronic industry, Lin and Chiu [10] used multivariate Hotelling $T^2$ statistic to integrate different coordinates of color models for MURA-type defect detection on Liquid Crystal Displays (LCD), and applied ant colony algorithm and back-propagation neural network techniques to develop an automatic inspection procedure. Lu and Tsai [11] proposed a global approach for automatic visual inspection of micro defects such as pinholes, scratches, particles and fingerprints. The Singular Value Decomposition (SVD) adopted by Lu and Tsai suits the need for detecting defects on the TFT-LCD images of highly periodical textural structures. Furthermore, in the recent decade, many vision systems have been developed for the inspection of surface defects on semiconductor wafers [12-14]. For instance, Fadzil and Weng [15] implemented a vision inspection system that achieves a 90% probability of accurately classifying defects, scratches, contamination, blemishes, off center defects, etc. in the encapsulations of diffused LED products.

The aforementioned techniques perform well in anomaly detection, but most of them do not detect defects with the properties of water-spot defects. This research has been motivated by the need for an efficient and effective technique that detects semi-opaque and low-intensity-contrast water-spot defects falling across two different background textures.

2 Proposed Method

To detect water-spot defects of LED chips, this research adopts the one-level Haar wavelet transform to conduct image transformation and extract wavelet characteristics. We apply the wavelet-based multivariate statistical approach to integrate multiple wavelet characteristics and then develop the principal component model of multivariate statistical analysis to individually judge the existence of water-spot defects in LED chip images.

2.1 Wavelet characteristics

The Haar wavelet transform is one of the simplest and basic transformations. Its base transform in the multiple-level scaling space can be implemented as:

$$v_{j,k} = \frac{v_{j+1,2k} + v_{j+1,2k+1}}{2}; w_{j,k} = \frac{v_{j+1,2k} - v_{j+1,2k+1}}{2}. \quad (1)$$
In this research, we apply a standard decomposition that covers wavelet row and column transfers to do the wavelet transform of a two-dimensional image. The Haar transform can be computed stepwise by the mean value and half of the differences of the tristimulus values of two contiguous pixels. We perform the 2-D wavelet transform by applying 1-D wavelet transform first on rows and then on columns. Based on the transfer concept of the one-dimensional space, the Haar wavelet transform can process a two-dimensional image of \((M \times N)\) pixels in the following way:

### Row transfer:

\[
g_r(p, q) = \frac{g(p, 2q) + g(p, 2q + 1)}{2}, \quad \frac{N}{2} \\
g_r(p, q + \frac{N}{2}) = \frac{g(p, 2q) - g(p, 2q + 1)}{2}, \quad \text{where } 0 \leq p \leq (M - 1), 0 \leq q \leq \frac{N}{2} - 1, \text{[1]} \text{ is Gauss symbol.}
\]

### Column transfer:

\[
g_c(p, q) = \frac{g_c(2p, q) + g_c(2p + 1, q)}{2}, \quad \frac{M}{2} \\
g_c(p + \frac{M}{2}, q) = \frac{g_c(2p, q) - g_c(2p + 1, q)}{2}, \quad \text{where } 0 \leq p \leq \frac{M}{2} - 1, 0 \leq q \leq (N - 1).
\]

In the above expressions (Eqs. (2)-(3)), \(g_r(p, q)\) represents an original image, \(g_r(p, q)\) the row transfer function of \(g(p, q)\), and \(g_c(p, q)\) the column transfer function of \(g(p, q)\). As \(g_c(p, q)\) is also the outcome of the wavelet decomposition of \(g(p, q)\), the outcomes of a wavelet transform can be defined as:

\[
A(p, q) = g_c(p, q); \quad D_1(p, q) = g_r(p, q + \frac{N}{2}); \\
D_2(p, q) = g_c(p + \frac{M}{2}, q); \quad D_3(p, q) = g_r(p + \frac{M}{2}, q + \frac{N}{2}); \quad \text{where } 0 \leq p \leq \frac{M}{2} - 1, 0 \leq q \leq \frac{N}{2} - 1.
\]

One level of wavelet decomposition generates one smooth sub-image and three detail sub-images that contain fine structures with horizontal, vertical, and diagonal orientations. An image is decomposed by wavelet transform into one approximation sub-image \(A\) and three detail sub-images \((D_1, D_2, D_3)\). These four sub-images, each of which has a size of \((M/2 \times N/2)\) pixels, form the wavelet characteristics. Wavelet transform provides a convenient way to obtain a multi-resolution representation, from which texture features can be easily extracted. The merits of using wavelet transform include local image processing, simple calculations, high speed processing and multiple image information [16-18].

### 2.2 Wavelet-based principal component analysis

Principal component analysis (PCA) is a popular technique for data compression and has been successfully used as initial step in many computer vision tasks [19-20]. The principal components of a set of process variables \(x_1, x_2, \ldots, x_p\) are just a particular set of linear combinations of these variables. Geometrically, the principal component variables \(y_1, y_2, \ldots, y_p\) are the axes of a new coordinate system obtained by rotating the axes of the original system (the \(p\)’s). The new axes represent the directions of maximum variability.

The basic intent of principal components is to find the new set of orthogonal directions that define the maximum variability in the original data, and this will lead to a description of the process requiring considerably fewer than the original \(p\) variables. The information contained in the complete set of all \(p\) principal components is exactly equivalent to the information in the complete set of all original process variables, but hopefully we can use far fewer than \(p\) principal components to obtain a satisfactory description [21].

Let the four random variables \(x_1, x_2, x_3, x_4\) be the four wavelet characteristics and be represented by a vector \(X = [A, D_1, D_2, D_3]^T\) with covariance matrix \(\Sigma\), and let the eigenvalues of \(\Sigma\) be \(\lambda_1, \lambda_2, \lambda_3, \lambda_4\). Then the constants \(e_i\) are simply the elements of the \(i\)th eigenvector \(e_i\) associated with the eigenvalue \(\lambda_i\).

Basically, if we let \(E\) be the matrix whose columns are the eigenvectors, then

\[
E^T \Sigma E = A
\]

where \(A\) is a \(4 \times 4\) diagonal matrix with main diagonal elements equal to the eigenvalues \(\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_4 \geq 0\). More specifically, the equation can be expressed by the eigenvalues and the eigenvectors as follows:

\[
[e_1^T, e_2^T, e_3^T, e_4^T] = \begin{bmatrix}
\sigma_2^2 & \sigma_{2,3} & \sigma_{2,4} & \sigma_{2,5} \\
\sigma_{3,2} & \sigma_3^2 & \sigma_{3,4} & \sigma_{3,5} \\
\sigma_{4,2} & \sigma_{4,3} & \sigma_4^2 & \sigma_{4,5} \\
\sigma_{5,2} & \sigma_{5,3} & \sigma_{5,4} & \sigma_5^2
\end{bmatrix} \begin{bmatrix} e_1^T \\
e_2^T \\
e_3^T \\
e_4^T \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\
0 & \lambda_2 & 0 & 0 \\
0 & 0 & \lambda_3 & 0 \\
0 & 0 & 0 & \lambda_4\end{bmatrix}
\]
The principal component analysis can be performed by computing the eigenvalues and eigenvectors [22]. The variance of the $i$th principal component is the $i$th eigenvalue $\lambda_i$. Consequently, the proportion of variability in the original data explained by the $i$th principal component is given by the ratio $\lambda_i / (\lambda_1 + \lambda_2 + \cdots + \lambda_d)$. Therefore, one can easily see how much variability (for instance, 80 to 90%) is explained by retaining just a few (say, $r$) of the $d$ principal components simply by computing the sum of the eigenvalues for those $r$ components and comparing that total to the sum of all $d$ eigenvalues.

Once the principal components have been calculated and a subset of them selected, we can obtain new principal component observations $y_{ij}$ (principal component (PC) scores) simply by substituting the original observations $x_{ij}$ into the set of $r$ retained principal components. After conducting many experiments, we find the first one principal component accounts for most of the variability in this study. If we have retained the first one (i.e. $r=1$) of the original four principal components, then the PC score $Y_{M(x,y)}$ of the multivariate processing unit $M(x,y)$ of a testing image can be defined as:

$$Y_{M(x,y)} = \left| e_i ' (X_{M(x,y)} - \bar{X}) \right|$$  \hspace{1cm} (7)

where

$$X_{M(x,y)} = \begin{bmatrix} A_1(x,y) \\ D_1(x,y) \\ D_2(x,y) \\ D_3(x,y) \end{bmatrix}_{4 \times 1}$$

$$\bar{X} = \begin{bmatrix} A_{max} \\ D_{min} \\ D_{min} \\ D_{min} \end{bmatrix}_{4 \times 1}$$

and $e_i ' = [e_{i1}, e_{i2}, e_{i3}, e_{i4}]$ is the first eigenvector of the $\Sigma$ of a testing image. Normal texture images are used to estimate the parameters of standard texture characteristics for bonding pad and emitting area, respectively. The sample expected matrix of the characteristics for bonding pad and emitting area, to estimate the parameters of standard texture $\Sigma$.

The principal component analysis can be performed by computing the eigenvalues and eigenvectors [22]. Consequently, the proportion of variability in the original data explained by the $i$th principal component is given by the ratio $\lambda_i / (\lambda_1 + \lambda_2 + \cdots + \lambda_d)$. Therefore, one can easily see how much variability (for instance, 80 to 90%) is explained by retaining just a few (say, $r$) of the $d$ principal components simply by computing the sum of the eigenvalues for those $r$ components and comparing that total to the sum of all $d$ eigenvalues.

Once the principal components have been calculated and a subset of them selected, we can obtain new principal component observations $y_{ij}$ (principal component (PC) scores) simply by substituting the original observations $x_{ij}$ into the set of $r$ retained principal components. After conducting many experiments, we find the first one principal component accounts for most of the variability in this study. If we have retained the first one (i.e. $r=1$) of the original four principal components, then the PC score $Y_{M(x,y)}$ of the multivariate processing unit $M(x,y)$ of a testing image can be defined as:

$$Y_{M(x,y)} = \left| e_i ' (X_{M(x,y)} - \bar{X}) \right|$$  \hspace{1cm} (7)

where

$$X_{M(x,y)} = \begin{bmatrix} A_1(x,y) \\ D_1(x,y) \\ D_2(x,y) \\ D_3(x,y) \end{bmatrix}_{4 \times 1}$$

$$\bar{X} = \begin{bmatrix} A_{max} \\ D_{min} \\ D_{min} \\ D_{min} \end{bmatrix}_{4 \times 1}$$

and $e_i ' = [e_{i1}, e_{i2}, e_{i3}, e_{i4}]$ is the first eigenvector of the $\Sigma$ of a testing image. Normal texture images are used to estimate the parameters of standard texture characteristics for bonding pad and emitting area, respectively. The sample expected matrix of the characteristics for bonding pad and emitting area, to estimate the parameters of standard texture $\Sigma$.

The performance evaluation indices, $(1-\alpha)$ and $(1-\beta)$, are used to represent correct detection judgments; the higher the two indices, the more accurate the detection results. Statistical type I error $\alpha$.
suggests the probability of producing false alarms, i.e. detecting normal regions as defects. Statistical type II error $\beta$ implies the probability of producing missing alarms, which fail to alarm real defects. We divide the area of normal region detected as defects by the area of actual normal region to obtain type I error, and the area of undetected defects by the area of actual defects to obtain type II error. The correct classification rate (CR) is defined as:

$$CR = \frac{(N_{cc} + N_{dd})}{N_{total}} \times 100\%$$  (9)

where $N_{cc}$ is the pixel number of normal textures detected as normal areas, $N_{dd}$ is the pixel number of ripple defects detected as defective regions, and $N_{total}$ is the total pixel number of a testing image.

By Otsu method | By WPCA method | By inspector

Fig. 2 Partial detection results of the Otsu, WPCA, and the professional inspector

The average detection rates (1-$\beta$) of all testing samples by the two methods are, respectively, 86.6% (Otsu method) and 93.8% (WPCA method). The proposed wavelet-based multivariate statistical approach has higher detection rates (1-$\beta$) and correct classification rates (CR) than does the traditional method applied to LED chip images. The WPCA method excels in its ability of correctly discriminating water-spot defects from normal regions. The average computation time for processing an image of 256 x 256 pixels is as follows: 1.84 seconds by the Otsu method, 2.26 seconds by the WPCA method.

As the decision threshold value changes, so do its false alarm rate ($\alpha$) and detection rate (1-$\beta$), both of which are used to describe the performance of a test according to hypothesis testing theory [24]. When various decision thresholds (Eq. (8)) are used, their pairs of false alarm rates and detection rates are plotted as points on a Receiver Operating Characteristic (ROC) curve. Figure 3 presents the two ROC curves of the WPCA approach, and the two points of the Otsu method for the bonding pad and the emitting area respectively. The upper-left corner of Fig. 3 is the optimum points, which have a 100% detection rate and a 0% false alarm rate. The more the ROC plot approaches the upper-left corner, the better the test performs. In industrial practices, a more than 90% detection rate and a less than 10% false alarm rate are a good rule of thumb for performance evaluation of a vision system. Accordingly, the proposed WPCA approach, with its ROC plots closer to the upper-left corner, outperforms the traditional method.

Fig. 3 ROC plots of the Otsu and MPCA methods for bonding pad and emitting area

4 Conclusion

Machine vision technology improves productivity and quality management, and provides a competitive advantage to industries that employ this technology. This research applies wavelet-based multivariate statistical approach combined with machine vision techniques to detect water-spot defects that fall across two different background textures of LED chips. The proposed approach uses the wavelet-based principal component analysis to judge the existence of water-spot defects through multivariate processes of combining image characteristics from wavelet decomposition of local image blocks.

Experimental results show that the WPCA approach achieves detection rates of above 93.8% and false alarm rates of below 3.6% in detecting water-spot blemishes across two different background textures. As indicated in the ROC curve analysis, the WPCA approach has lower false alarm rates and better detection rates than does the Otsu method. This research contributes a solution to a common surface defect detection problem of LED chips and offers a computer-aided visual defect inspection system to meet the inspection and quality control request.
Acknowledgements:
The authors thank the National Science Council of Taiwan (R.O.C.) for the financial support through the Grant NSC 95-2221-E-324-034-MY2.

References: