# Saliency-Based Defect Detection in Industrial Images by Using Phase Spectrum

Xiaolong Bai, Yuming Fang, Weisi Lin, *Senior Member, IEEE*, Lipo Wang, *Senior Member, IEEE*, and Bing-Feng Ju

***Abstract*—For computer vision-based inspection of electronic chips or dies in semiconductor production lines, we propose a new method to effectively and efficiently detect defects in images. Different from the traditional methods that compare the image of each test chip or die with the template image one by one, which are sensitive to misalignment between the test and template images, a collection of multiple test images are used as the input image for processing simultaneously in our method with two steps. The first step is to obtain salient regions of the whole collection of test images, and the second step is to evaluate local discrepancy between salient regions in test images and the corresponding regions in the defect-free template image. To be more specific, in the first step of our method, phase-only Fourier transform (POFT), which is computationally efficient for online applications in industry, is used for saliency detection. We provide the theoretical justification for POFT to be effective to attenuate the normal regions and amplify the defects in multiple test images, which are usually arranged in a matrix format in industrial practice. By comparing with four other popular methods, the proposed algorithm can efficiently accommodate small variations (inevitable in practice) in test chips or dies, such as the spatial misalignments and product variations. Experimental results on a large-scale database including 1073 images, 94 of which are defective, show that our method performs much better than the other methods in terms of precision, recall, and F-measure.**

***Index Terms*—Computer vision, defect detection, Fourier transform, phase spectrum, saliency, surface defects, template matching.**

## I. Introduction

COMPUTER vision-based defect detection in electronic chips, integrated circuits (ICs), or other electronic components is highly demanded in industry [1]–[8]. Specifically, the following two stages of inspection are important in semiconductor industry [1]. First, before the packaging process, each die is cut from the wafer by using the dicing saw. All the separated dies are placed on a thin film in a matrix array for inspection using scanning electronic microscopes (SEM), cameras, or other imaging instruments [4], [5]. Then, the defective dies are discarded and the defect-free dies are placed in a chip tray for packaging. However, the packaging process may introduce various kinds of defects to the dies, such as voids, inclusions, delaminations, etc. Second, the packaged chips (or ICs) are aligned in a matrix array for final inspection, in which scanning acoustic microscope (SAM), X-rays, or other nondestructive imaging systems are employed to obtain images of the interior structures of the electronic chips (or ICs) [6], [7], [14].

Traditionally, all the test images of the chips or dies are processed one by one, and a template image which is defect-free is used for comparison with the test images [5]. Image comparison-based methods include image subtraction [8], template matching, and phase-only methods [9]. There are also some model-based methods for detection of different kinds of defects. For certain types of defects of solder joints in electronic chips, Acciani *et al.* extracted the features of the regions of interest in test images and then built multilayer neural networks for defect detection [10]. Wavelet features were extracted by comparing wavelet coefficients of the test images with those of the template image. For inspection of light emitting diode (LED) wafers, Li *et al.* used regional formulation to improve the segmentation of test images and then adopted the general features for defect detection [11]. However, this method is specifically designed and cannot be easily applied to inspection of other different types of electronic chips or dies. For defect detection in the liquid crystal display (LCD) wafer images, Kim *et al.* extracted the features from the template and test images using corner detectors, based on which accurate template matching was achieved [12]. Most researchers select one or a few defect-free image(s) as template(s) from the test images, or generate the template image indirectly using other methods for comparison with all the test images [13].

Most prevalent methods consider the defect detection as a template-matching problem, based on the dissimilarity between the template image and test images. In the test images, the regions with the dissimilarity exceeding a predefined threshold are identified as defects. The normalized cross correlation (NCC) is a classical metric measuring the degree of the dissimilarity [14], [15]. Some new methods have been developed based on NCC-extended dissimilarity metrics, such as partial information correlation coefficient (PICC) [16]. NCC-based

Manuscript received January 10, 2014; revised July 25, 2014; accepted September 09, 2014. Date of publication September 22, 2014; date of current version November 04, 2014. This work was supported in part by the National Natural Science Foundation of China, Project 51175467; in part by the National Basic Research Program of China, 973 Program, under Grant 2011CB706700; in part by the National High Technology Research and Development Program of China, 863 Program, under Grant 2012AA040405; in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LZ13E050001; and in part by the Fundamental Research Funds for the Central Universities and the Scholarship Award for Excellent Doctoral Students granted by the China Scholarship Council under Grant CSC 201306320002. *(Corresponding author: Bing-Feng Ju.)*. Paper no. TII-14-0037.

X. Bai and B.-F. Ju are with the State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, Hangzhou 310027, China (e-mail: mbfju@zju.edu.cn).

Y. Fang is with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, Jiangxi, China (e-mail: fa0001ng@e.ntu.edu.sg).

W. Lin is with the School of Computer Engineering, Nanyang Technological University, 639798 Singapore (e-mail: wslin@ntu.edu.sg).

L. Wang is with the School of Electrical and Electronic Engineering, Nanyang Technological University, 639798 Singapore (e-mail: elpwang@ntu.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2014.2359416

1551-3203 © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

template-matching methods and their extended versions have been used extensively in the inspection of printed circuit board (PCB), and electronic chips or dies.

Wang *et al.* matched the test image with multiple templates simultaneously, and the template with the largest NCC value was selected as the optimal template [17]. Crispin *et al.* first localized multiple objects in PCBs to generate the templates for different chips, and then used NCC as the template-matching metric for defect detection [18]. Wang *et al.* used PICC which was more robust to spatial misalignment than NCC [16]. Some fast NCC-based algorithms were proposed to improve computational efficiency [19], [20].

The NCC-related template-matching methods rely heavily on the accurate alignment of the template and test images [12], [14], [18]. Tsai *et al.* proposed a shift-tolerant measure for surface defect detection in LED wafer images [14]. Using the defect-free template image as the reference, the magnitude of optical flow of each pixel in the test image is calculated, which can be used as a metric for defect detection in LED images. However, it is not effective in inspecting complicated components such as electronic chips (or ICs), in which the optical flow magnitudes of some defective regions are not large enough. Actually, for inspection of electronic chips or dies, most aforementioned methods based on image comparison with the template image suffer certain limitations, one of which is the sensitivity to spatial shift between the template and test images, which is inevitable in practice because chips or dies should be placed in the chip tray for the aforementioned two-stage inspection and positions between the chips or dies are not always kept exactly the same [5]. Therefore, it is much desired to develop a fast defect-detection method which is not only robust to the minor spatial misalignment, such as shift, but also able to be easily applied to inspection of different types of industrial images.

This research justifies and investigates a new paradigm of thinking into industrial defect detection in the production line, in order to address the abovementioned shortcomings in the existing approaches. The defect is usually sparse and unpredictable (i.e., as unusual happenings), and it can be considered as abnormal regions and formulated as salient regions in test images, with inspiration of human visual attention modeling [21]–[26]. In this paper, we propose a new method for fast defect detection in images of electronic chips or dies based on the devised salient detection model. The proposed formulation of industrial defect detection into a saliency detection problem enables the following differentiation: different from the aforementioned existing image comparison or model-based defect detection methods, in which chips or dies are inspected one by one, we treat a collection of test images arranged in an array as a whole as the input image for detection of salient regions, removing as many normal regions as possible by utilizing the self-similarities among the multiple images. Defects are then extracted from salient regions. Thus, the proposed method for defect detection in industrial images, e.g., the images of electronic chips or dies, consists of two steps: first, salient regions are fast extracted from an array of test images; second, defects are identified from the salient regions using template matching based on a self-defined spatial misalignment-tolerant metric. Our method provides a hybrid way of saliency detection and image comparison. Most normal regions are efficiently removed in the first step, so that only a few regions need to be further analyzed in detail. In the second step, spatial misalignments of test images are well accommodated by using template matching.

In this study, the salient detection model based on phase-oOnly Fourier transform (POFT) [22] is adopted for saliency extraction due to its efficiency. We provide the theoretical justification of POFT with application to defect detection.

The rest of this paper is therefore organized as follows. In Section II, the principle and complete theoretical justification of the salient detection model based on POFT for defect detection are given. In Section III, the proposed method is presented, with experimental results demonstrated in Section IV. Finally, conclusion is made in Section V.

## II. Saliency Detection Based on POFT

During the past decades, various visual attention models have been developed for saliency detection in images/videos [22]–[26]. The effectiveness of POFT as a saliency detection model for images/videos was experimentally discovered by Lei *et al.* [22]. That model is also fast in saliency detection due to the computational efficiency of fast Fourier transform. By POFT, the phase spectrum of the original image is retained with the amplitude spectrum set to a constant in the Fourier domain and then inverse Fourier transform is applied to get the reconstructed image, in which the salient regions are expected to pop out. The hypothesis behind that model is that the phase spectrum rather than amplitude spectrum obtained from Fourier transform can detect salient regions in images.

However, there is no complete theoretical justification of POFT for the saliency detection. Aiger *et al.* [27] used POFT to extract surface defects with a textured background. The phases of the texture and defect are considered separately after Fourier transform and the study proves that the phase integral excursion of the texture is almost zero, while that of a random defect is nonzero but does not mathematically explain why the defect can pop out as salient regions (with amplified amplitudes) after amplitude normalization and inverse Fourier transform of the mixed signal of texture and defects. They resorted to simulations of one-dimensional (1-D) signals and two-dimensional (2-D) images to show the effectiveness of POFT. In the rest of this section, we will present our findings as a complete theoretical justification of POFT with application to defect detection in a collection of test images of electronic chips or dies.

Given multiple test images aligned in a $N_x \times N_y$ matrix array in the 2-D spatial space $(x, y)$, defects are always irregular, sparse, and unpredictable, and therefore can be regarded as salient regions in the whole collection of test images as an array. So, with this arrangement, a collection of test images contains two parts: the defective regions as the foreground and the defect-free regions as the background. Defects are sparse 2-D signals, while the defect-free regions are 2-D periodic signals in the spatial domain due to the matrix alignment of multiple test chips or dies.

The arrangement of visual data above essentially enables turning of the problem under attack into a saliency (attention) detection formulation. The defects can be seen as nonperiodic signal in the 2-D spatial space $(x,\ y)$ and modeled as

$$f_F(x,y) = \sum_{m=1}^{M} f_{Fm}(x,y) = \sum_{m=1}^{M} f_m \cdot \mathrm{rect}_{Fm}(x-x_m, y-y_m) \quad (1)$$

where $M$ is the total number of defects in the whole collection and $f_m$ is the signal amplitude of the $m$th defect whose location is denoted by $(x_m, y_m)$ and

$$\mathrm{rect}_{Fm}(x,y) = \begin{cases} 1, & -\frac{W_{Fmx}}{2} < x < \frac{W_{Fmx}}{2} - \frac{W_{Fmy}}{2} < y < \frac{W_{Fmy}}{2} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

is a 2-D rectangular window function with narrow widths of $W_{Fmx}$ and $W_{Fmy}$.

The defect-free regions which contain multiple test images aligned in an array are modeled as

$$f_B(x,y) = \sum_{r=-\infty}^{+\infty} \sum_{s=-\infty}^{+\infty} f_{B0}(x - rX_B, y - sY_B) \quad (3)$$

where $r$ and $s$ denote the number of duplicated chips or dies along $x$ and $y$ axes, respectively; $f_B(x,y)$ can be seen as a 2-D periodic signal with periods of $X_B$ and $Y_B$; $f_B(x,y)$ is further expanded as

$$f_B(x,y) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} a_{kl} \exp(i(ku_B x + lv_B y)) \quad (4)$$

using Fourier series, where

$$a_{kl} = \frac{1}{X_B \cdot Y_B} \int_{-X_B/2}^{X_B/2} \int_{-Y_B/2}^{Y_B/2} f_{B0}(x,y) \cdot \exp(-i(ku_B x + lv_B y)) \mathrm{d}x\mathrm{d}y \quad (5)$$

$i = \sqrt{-1}$ is the imaginary unit; $u_B = 2\pi/X_B$, $v_B = 2\pi/Y_B$; $k$ and $l$ are integral numbers.

The mixed 2-D signal which is a linear superposition of the defects $f_F(x,y)$ and defect-free regions $f_B(x,y)$ can be represented as

$$f(x,y) = f_F(x,y) + f_B(x,y) = \sum_{m=1}^{M} f_m \cdot \mathrm{rect}_{Fm}(x-x_m, y-y_m) + \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} a_{kl} \exp(i(ku_B x + lv_B y)). \quad (6)$$

The 2-D Fourier transform of $f(x,y)$ is denoted as

$$F(u,v) = \mathrm{FT}[f(x,y)] = \mathrm{FT}[f_F(x,y)] + \mathrm{FT}[f_B(x,y)] \quad (7)$$

in which

$$\mathrm{FT}[f_F(x,y)] = \sum_{m=1}^{M} A_{Fm}(u,v) \cdot \exp(-i\varphi_{Fm}(u,v)) \quad (8)$$

where FT[ ] denotes the 2-D Fourier-transform operation; $A_{Fm}(u,v)$ and $\varphi_{Fm}(u,v)$ are the amplitude and phase spectrum of $f_{Fm}(x,y)$ representing the $m$th defect, and

$$\mathrm{FT}[f_B(x,y)] = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} \mathrm{FT}[a_{kl} \cdot \exp(i(ku_B x + lv_B y))] = \frac{4\pi^2}{X_B Y_B} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} A_B(u,v) \cdot \exp(-i\phi_B(u,v)) \cdot \delta(u - ku_B, v - lv_B) \quad (9)$$

where $A_B(u,v)$ and $\varphi_B(u,v)$ are the amplitude and phase spectrum of $f_{B0}(x,y)$, respectively, and

$$\delta(u - ku_B, v - lv_B) = \begin{cases} +\infty, & \text{if } u = ku_B, v = lv_B \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

is the 2-D Dirac delta function.

Based on (8) and (10), the $F(u,v)$ in (7) is given by

$$F(u,v) = F_F(u,v) + F_B(u,v) = \sum_{m=1}^{M} A_{Fm}(u,v) \cdot \exp(-i\phi_{Fm}(u,v)) + \frac{4\pi^2}{X_B Y_B} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} A_B(u,v) \cdot \exp(-i\phi_B(u,v)) \cdot \delta(u - ku_B, v - lv_B). \quad (11)$$

After normalization of the amplitude of $F(u,v)$ to unity over the entire 2-D spatial frequency space $(u,v)$, $F(u,v)$ in (7) can be transformed to

$$F'(u,v) = \exp(-i\phi_F(u,v)) + \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} [\exp(-i\phi_{BF}(u,v)) - \exp(-i\phi_F(u,v))] \cdot \delta_0(u - ku_B, v - lv_B) \quad (12)$$

where

$$\delta_0(u - ku_B, v - lv_B) = \begin{cases} 1, & \text{if } u = ku_B, v = lv_B \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

and $\varphi_{BF}(u,v)$ is the resultant phase containing the phase information of both the defect and defect-free regions when $u$ is the integer multiples of $u_B (u = ku_B)$ and $v$ is the integer multiples

of $v_B (v = lv_B)$, while $\varphi_F(u,v)$ is dominated by the phase of the defect.

Applying the 2-D inverse Fourier transform to these two parts in the right-hand side of (12), we can get

$$f_1'(x,y) = \text{IFT}\left[\exp(-i\phi_F(u,v))\right] = \sum_{m=1}^{M} \delta(x - x_m, y - y_m) \tag{14a}$$

$$\begin{aligned} f_2'(x,y) &= \text{IFT}\left[\sum_{k=-\infty}^{+\infty}\sum_{l=-\infty}^{+\infty}\left[\exp(-i\phi_{BF}(u,v)) - \exp(-i\phi_F(u,v))\right]\delta_0(u-ku_B, v-lv_B)\right] \\ &= \text{IFT}\left[\sum_{r=-\infty}^{+\infty}\sum_{s=-\infty}^{+\infty}(a_{rs} - b_{rs}) \cdot \exp(i(rX_B u + sY_B v))\right] \end{aligned} \tag{14b}$$

where IFT[ ] denotes the 2-D inverse Fourier transform operation, and

$$a_{rs} = \frac{1}{4\pi^2 u_B v_B}\int_{-u_B/2}^{u_B/2}\int_{-v_B/2}^{v_B/2} \exp(-i(\phi_{BF}(u,v) + rX_B u + sY_B v)) \cdot \delta_0(u,v)\mathrm{d}u\mathrm{d}v \tag{15a}$$

and

$$b_{rs} = \frac{1}{4\pi^2 u_B v_B}\int_{-u_B/2}^{u_B/2}\int_{-v_B/2}^{v_B/2} \exp(-i(\phi_F(u,v) + rX_B u + sY_B v)) \cdot \delta_0(u,v)\mathrm{d}u\mathrm{d}v. \tag{15b}$$

The function $\delta_0(u,v)$ in (15a) and (15b) can be seen as a rectangular window function with infinitesimally small window width and amplitude of one. So its integral over the 2-D space $(u,v)$ is equal to zero, and the results of integrals in (15a) and (15b) are also equal to zero, and we have $a_{rs} = 0$, $b_{rs} = 0$, and $f_2{}'(x,y) = 0$.

Combining (14) and (15), the inverse Fourier transform of $F'(u,v)$ is obtained as

$$f'(x,y) = \sum_{m=1}^{M} \delta(x - x_m, y - y_m) \tag{16}$$

which consists of multiple 2-D Dirac delta functions with extremely large amplitudes at the locations of the defects. The information of the defect-free regions is eliminated in (16). In practice, it is impossible to have an infinite number of the test images aligned for defect detection, so $r$ and $s$ in (3) will have finite lower and upper limits, which means the 2-D signal of defects $f_F(x,y)$ and defect-free regions $f_B(x,y)$ are truncated by a window. The Dirac delta functions in (16) will be replaced by *sinc* functions, but still with very large amplitude at the locations of the defects. Thus, POFT can greatly amplify the defects and attenuate the defect-free regions. The defects will automatically pop out as salient regions in the image after POFT is applied.

## III. Proposed Method

According to the analysis in Section II, if the spatial intervals between some chips or dies in the same chip tray are not exactly the same, some regions in the chips may not have the common period shared by other regions in the chips. These regions will also pop out as salient regions after POFT are applied to the whole collection of test images. So the salient regions may include both the defective regions and some defect-free regions resulted from the spatial shift between the chips and other variations. Further processing is needed to obtain the defects from the extracted salient regions. Therefore, the proposed method contains two steps: the first step is to use POFT for saliency detection, while the second step is to detect defects from the extracted salient regions.

### A. Step 1: Saliency Detection Based on POFT

In this step, we first apply POFT to the whole collection of test images to get the saliency map. Then, a 2-D Gaussian filter

$$G_f(x,y) = \frac{1}{2\pi\sigma_f^2}\cdot\exp\left(-\frac{x^2+y^2}{2\sigma_f^2}\right) \tag{17}$$

is applied to eliminate the noisy peaks in the saliency map, where $\sigma_f$ is the standard deviation of the Gaussian distribution.

From a statistical point of view, it is assumed that the gray values in the image after POFT have a Gaussian distribution and the pixels whose gray values exceed the tolerance interval can be seen as salient pixels. So, after calculating the mean value $\mu_s$ and standard deviation $\sigma_s$, the pixels with gray values larger than $\mu_s + C_s \cdot \sigma_s$ can be seen as unusual or salient ones. In practical applications [28], $C_s$ is usually set as 3.

So, the binary saliency map is given by

$$B(x,y) = \begin{cases} 0, & \text{if } s(x,y) \geq \mu_s + C_s \cdot \sigma_s \\ g_{\max}, & \text{otherwise} \end{cases} \tag{18}$$

where $g_{\max}$ is the largest gray level of the image, being equal to 255 if the gray level is displayed in 8 bits.

As an illustration, a collection of images of chips which are aligned in a $4\,\text{chip} \times 4\,\text{chip}$ array is shown in Fig. 1(a). These images are acquired by a SAM which is capable of imaging the interiors of chips using high-frequency ultrasound. The chips are not very accurately aligned. There exist small variations in the test images and the spatial intervals between the adjacent test images. There is a crack (a defect), marked by a red circle in Fig. 1(a). The binary saliency map $B(x,y)$ of Fig. 1(a) is illustrated in Fig. 1(b). It can be seen that the defect (shown in the

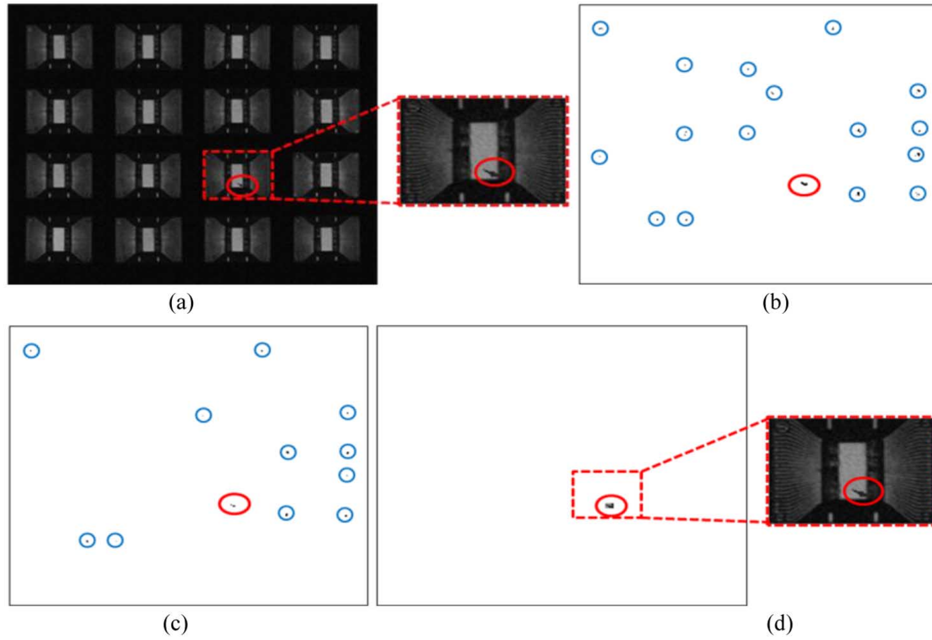



Fig. 1. Original image with its saliency maps and detection results. (a) Image of a collection of 4 × 4 test chips, one of which has a defect. (b) Binary saliency map obtained from the Step 1 of the proposed method. (c) Binary map of top 20% salient pixels in (b). (d) Final defect -etection result.

red circle) together with a few defect-free regions (shown in the blue circles) pop out as salient regions. The inevitable spatial shift between the chips and minor product variations make these defect-free regions salient.

Then, the salient pixels are used for further processing in Step 2. Optionally, after sorting all the salient pixels according to their saliency values obtained in Fig. 1(b) in descending order, we can select a portion of the salient pixels for processing. The selected top 20% salient pixels are shown in Fig. 1(c), from which we can see that the salient pixels representing the defects (shown in the red circle) do not disappear, revealing their high saliency. So, it is possible that not all the salient pixels are necessarily selected for further processing to improve the computational efficiency in Step 2.

### B. Step 2: Defect Extraction From the Salient Regions

In this step, we extract defects from the salient regions calculated in Step 1. For each salient pixel in the test images, template matching is applied to calculate its similarity with the corresponding pixel in a template image which is defect-free, as shown in Fig. 2. Assume that the $i$th salient pixel ($i = 1, 2, \ldots, N_s$) is located at the coordinate $(x_i,\ y_i)$ in the collection of test images, and this salient pixel belongs to the $(j,\ k)$th test image ($j = 1, 2, \ldots, N_x,\ k = 1, 2, \ldots, N_y$), where $N_s$ denotes the number of the salient pixels obtained from Step 1, and $N_x$ and $N_y$ are the numbers of the test images under inspection along the $x$ and $y$ axes, respectively. Before template-matching, a neighborhood window $W_i$ around the center $(x_i,\ y_i)$ in the test image and the corresponding window $W_i'$ around the center $(x_i',\ y_i')$ in the template image are defined. The size of both the two windows is $w_n \times w_n$. $x_i' = x_i - jD_x$ and $y_i' = y_i - kD_y$, where $D_x$ and $D_y$ are the average spatial intervals between two adjacent test images along $x$ and $y$ axes, respectively, as indicated in Fig. 2.

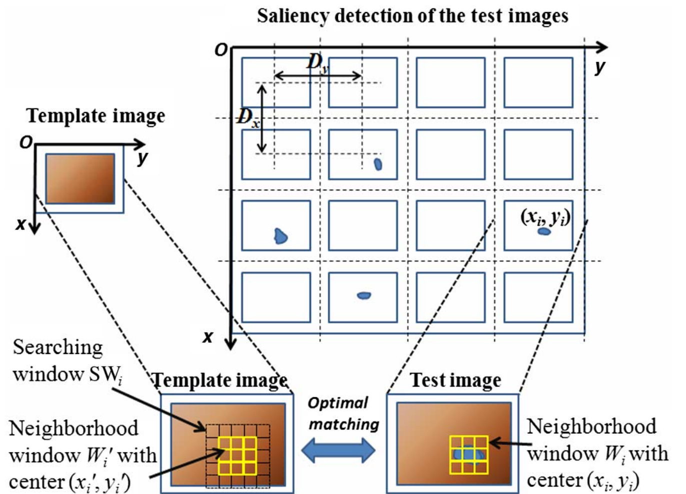



Fig. 2. Second step of the proposed method, which involves matching the test image with the template at locations of the salient pixels.

Subsequently, the similarity (or dissimilarity) between the pixels in $W_i$ and $W_i'$ is computed. Different metrics can be applied here to evaluate the degree of local similarity, such as energy difference, NCC, local binary pattern (LBP), etc. For simplicity, we apply energy difference between the pixels in these two windows to calculate the similarity between these two windows as

$$s(x_i, y_i) = \frac{1}{w_n^2 \cdot g_{\max}^2} \cdot \sum_{W_i, W_i'} [g_i(x, y) - g_i(x', y')]^2 \quad (19)$$

where $g_i(x, y)$ and $g_i(x', y')$ are the gray levels of pixels inside $W_i$ and $W_i'$, respectively, and $g_{\max}$ is the largest gray level

as shown in (18). The value of energy difference $s(x_i, y_i)$ is between 0 and 1. Due to the misalignment between the template image and the test image, a searching window $\mathrm{SW}_i$ with size of $w_s \times w_s$ around the center $(x_i', y_i')$ in the template image is defined to allow the location of the window $W_i'$ to have shift of $\Delta x_i$ and $\Delta y_i$ along $x$ and $y$ axes $(-(w_s - w_n)/2 \leq \Delta x_i \leq (w_s - w_n)/2$, and $(-(w_s - w_n)/2 \leq \Delta y_i \leq (w_s - w_n)/2)$, as shown in Fig. 2. The energy differences between the neighborhood window $W_i$ and each possible window $W_i'$ bounded by the searching window $\mathrm{SW}_i$ are calculated. To tolerate rotational misalignment, the $(j, k)$th test image is rotated with different angles $\theta_l$ $(-\theta_{\max} \leq \theta_l \leq \theta_{\max}, l = 1, 2, \ldots, N_\theta)$, and then the energy differences are calculated.

Here, the minimum value of the energy difference is defined as the misalignment-tolerant local dissimilarity measure (MTLDM) which is robust to horizontal, vertical, and rotational misalignment. MTLDM at location $(x_i, y_i)$ is denoted as $d(x_i,\ y_i)$ which is used as a measure of the extent to which the $i$th salient pixel may represent a defect, with the shift of $W_i{}'$ being $\Delta x_i^*$ and $\Delta y_i^*$, and the rotational angle of the test image being $\theta_i^*$. With a larger value of $d(x_i,\ y_i)$, the $i$th salient pixel is more likely to represent a defect, and vice versa.

Local adaptive threshold is required to decide whether the salient pixel represents a defect. We define $t(x, y)$ as the threshold at each pixel location $(x, y)$ in the template image. Then the $i$th salient pixel in the collection of test images will represent the defect if the MTLDM at this pixel

$$d(x_i, y_i) > t(x_i' + \Delta x_i^*, y_i' + \Delta y_i^*) \tag{20}$$

with the rotational angle of the $(j, k)$th test image being $\theta_i^*$.

In order to determine the local adaptive threshold $t(x,\ y)$ of the template image, a set of $M$ defective-free test images $\{I_1, I_2, \ldots, I_m, \ldots, I_M\}$ is used as the training set. For the template image, the MTLDM at each pixel location $(x,\ y)$ is obtained as $d_m(x,\ y)$ after comparison with each training image $I_m$, in which the searching window SW is applied. Then, the local adaptive threshold of the template image is given by

$$t(x, y) = \max_{m=1,2,\ldots,M} d_m(x, y) \tag{21}$$

to accommodate small product variations in the chips.

Based on the process presented above, the defect in the test images in Fig. 1(a) is extracted and shown in Fig. 1(d), with salient pixels which represent the defect-free regions discarded in Fig. 1(d) after comparison with the template image.

### C. Analysis of Computational Complexity

By using fast Fourier transform in Step 1, it requires $2 \cdot N_x \cdot N_y \cdot N \cdot \log(N_x \cdot N_y \cdot N)$ operations, where $N_x$ and $N_y$ are the numbers of the test images along the $x$ and $y$ axes in Fig. 2, respectively, and $N$ is the number of pixels in each test image. In Step 2, for each salient pixel, the computation of MTLDM needs $(w_s - w_n + 1)^2 \cdot w_n^2 \cdot N_\theta \cdot N_s$ operations. Hence, it totally takes $(2 \cdot N \cdot \log(N_x \cdot N_y \cdot N) + (w_s - w_n + 1)^2 \cdot w_n^2 \cdot N_\theta \cdot N_s/(N_x \cdot N_y))$ operations per image.

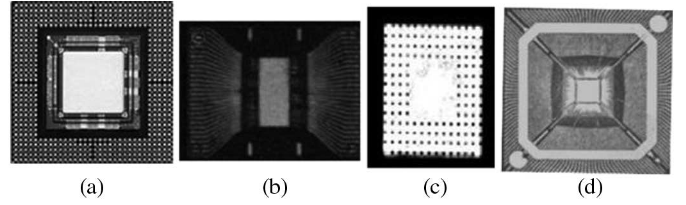


Fig. 3. Images of four different electronic chips. (a) Type A. (b) Type B. (c) Type C. (d) Type D.

TABLE I
DETAILS OF THE TEST IMAGES USED IN THE EXPERIMENTS

| Type | Number of total test images | Number of defective images | Size of the collection of test images (pixels by pixels) | Number of test images in each collection | Size of each test image (pixels by pixels) |
|---|---|---|---|---|---|
| A | 385 | 23 | 854 × 650 | 35 (5 × 7) | 122 × 130 |
| B | 128 | 14 | 708 × 604 | 16 (4 × 4) | 177 × 151 |
| C | 350 | 18 | 637 × 520 | 35 (5 × 7) | 91 × 104 |
| D | 210 | 39 | 616 × 510 | 42 (6 × 7) | 88 × 85 |

As shown in Fig. 1(b) and (c), after Step 1, very few pixels are left for further processing in Step 2, and in this example, the number of salient pixels $N_s$ is only 298, accounting for only 0.07% of the total number of pixels $N_x \cdot N_y \cdot N$. In most experiments, $N_s$ is far smaller than $N_x \cdot N_y \cdot N$, and the computational time of Step 2 is usually an order of magnitude less than that of Step 1. Therefore, the computational complexity of our method is mainly determined by Step 1, which is approximated as $O(N \cdot \log(N_x \cdot N_y \cdot N))$, when that of Step 2 is negligible.

## IV. Experimental Results

In this section, we evaluate the performance of the proposed method by using 1073 images of four typical kinds of electronic chips with different structures, in which 94 are defective. The defect-free template images of these four chips provided by different semiconductor manufacturing companies and laboratories are shown in Fig. 3. The detailed descriptions of the test images used in the experiment are provided in Table I. The proposed algorithm is implemented by C++ and compiled in Microsoft Visual Studio 2010, in a computer with Intel 2.2 GHz CPU.

### A. Preparation of Experiments

The sizes of the neighborhood windows $W$ and $W'$ introduced in Step 2 are set as $5 \times 5$ which is suitable for defect detection in test images with approximate image size of $100 \times 100$ pixels according to [11]; and the size of the searching window SW is set as $9 \times 9$ to make the spatial shift within 4 pixels well tolerated, and the number of rotational angles $N_\theta$ and $\theta_{\max}$ is set as $7^\circ$ and $3^\circ$, respectively (the test image is rotated $1^\circ$ each time, with rotational angle ranging from $-3^\circ$ to $3^\circ$), which are enough for semiconductor inspection [16].

We use 25 defect-free test images as training set to determine the local adaptive threshold $t(x, y)$ for each type of image

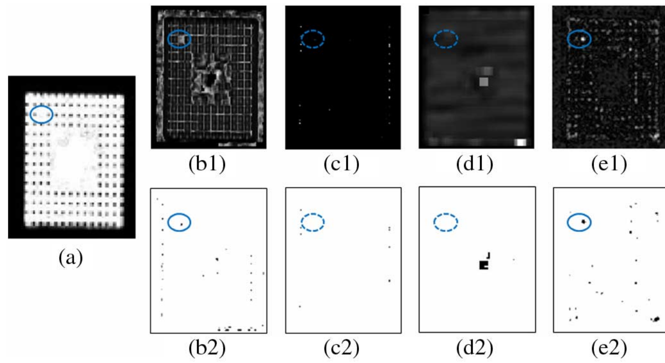



Fig. 4. First test image and defect detection results using the four comparison methods. (a) Defective test image. (b1), (c1), (d1), and (e1) are the calculated values of NCC, partial NCC, optical flow magnitude, and Mahalanobis distance. (b2), (c2), (d2), and (e2) are the detection results using the four methods, with dark pixels representing defects.

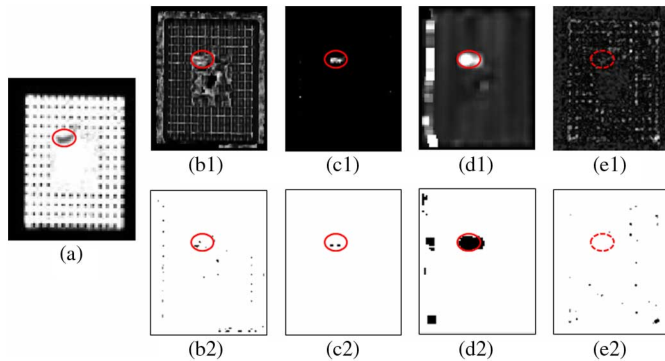



Fig. 5. Second test image and defect detection results using the four comparison methods. (a) Defective test image. (b1), (c1), (d1), and (e1) are the calculated values of NCC, partial NCC, optical flow magnitude, and Mahalanobis distance. (b2), (c2), (d2), and (e2) are the detection results using the four methods, with dark pixels representing defects.

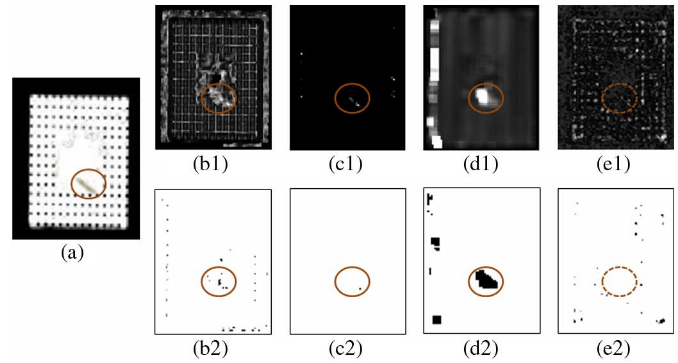



Fig. 6. Third test image and defect detection results using the four comparison methods. (a) Defective test image. (b1), (c1), (d1), and (e1) are the calculated values of NCC, partial NCC, optical flow magnitude, and Mahalanobis distance. (b2), (c2), (d2), and (e2) are the detection results using the four methods, with dark pixels representing defects.

following the procedures in Step 2 in Section II. To compute the $D_x$ and $D_y$ in Step 2, the projections of the whole collection of images are made along $x$ and $y$ axes, respectively, and the gray levels of each row and column are accumulated. The regions between each two adjacent test images usually have dark gray levels [as shown in Figs. 1(a) and 7(a)], which lead to local minima in the projected data along the two axes. So the $D_x$ and $D_y$ are calculated as the average distances between adjacent local minima in the two projected data, respectively.

By using our method, the average processing times for each test image of types A, B, C, and D are 36, 42, 23, and 22 ms, respectively. In this experiment, we select top 50% salient pixels in Step 1 for further processing in Step 2. However, this selection is not necessary when the computational efficiency is not badly needed. Using all the salient pixels in Step 1 for processing only makes the processing time of each image about 2 ms longer.

## *B. Performance Evaluation*

In this study, we use four methods, including the zero-mean NCC-based template matching [14], [20], partial NCC-based template matching [16], optical flow-based image comparison [14], and Aiger and Talbot's method [27] to compare with our method experimentally. Twenty-five defect-free images for each type of test images are used as the training set to determine the thresholds. According to [14] and [20], local adaptive threshold $T(x, y) = \mu(x, y) + C \cdot \sigma(x, y)$ is used for NCC and optical flow-based methods, in which $\mu(x, y)$ and $\sigma(x, y)$ are the mean and standard deviation of NCC value (or optical flow magnitude) at pixel location $(x, y)$ from the training images, and the optimal control constant $C$ is set at $-4$ and 4 for the NCC-based template matching and the optical flow-based methods, respectively. For a test image under inspection, the pixels with NCC values (or optical flow magnitudes) at the location $(x, y)$ smaller (or larger) than $T(x, y)$ represent the defect. As suggested by [16] and [27], a constant correlation threshold of 0.2 is used in the partial NCC-based method, and a Mahalanobis distance of 4 is used as the threshold to obtain the defect in Agier and Talbot's method.

As an illustration, we show the defect-detection results of some test images of type C. Three defective test images and one defect-free test image, shown in Figs. 4(a)–7(a), together with the detection results using these four methods are shown in Figs. 4–7, respectively. The gray levels in Figs. 4(b1)–7(b1) are inversely proportional to the NCC values, and those in Figs. 4(c1)–7(c1) are inversely proportional to the partial NCC values, while those in Figs. 4(d1)–7(d1) are proportional to the optical flow magnitudes, and those in Figs. 4(e1)–7(e1) are proportional to the Mahalanobis distances calculated using Aiger and Talbot's method.

In the traditional NCC-based template-matching method, the NCC value is calculated for each pixel within a $5 \times 5$ neighborhood window as in [14]. The final detection results are shown in Figs. 4(b2)–7(b2) for the four test images, and the dark pixels indicate defects. This method successfully detects the true defects (marked by a circle). But it is sensitive to the spatial shift between the template and test images. Some defect-free regions are detected as defects due to the spatial misalignment in Figs. 4(b2)–7(b2).

In the partial NCC-based template-matching method, the size of the neighborhood window is set as $3 \times 3$ as in [16]. The

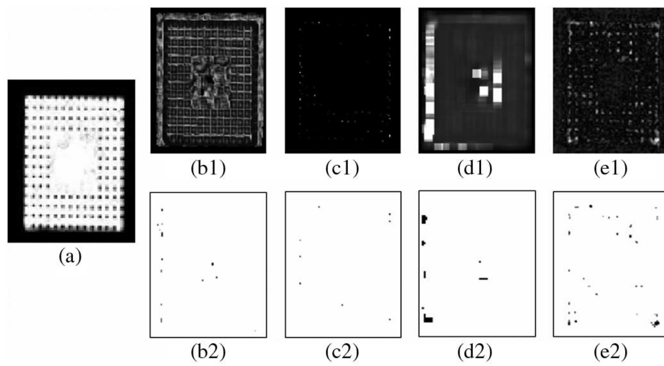


Fig. 7. Fourth test image and defect detection results using the four comparison methods. (a) Test image which is defect-free but with spatial misalignment. (b1), (c1), (d1), and (e1) are the calculated values of NCC, partial NCC, optical flow magnitude, and Mahalanobis distance. (b2), (c2), (d2), and (e2) are the detection results using the four methods, with dark pixels representing defects.

final detection results are shown in Figs. 4(c2)–7(c2) for the four test images, and the number of dark pixels representing the defects decreases when comparing to the NCC-based detection results, giving less false positives. Due to the use of dominant pixels with a large gray level difference with respect to the template image to calculate the NCC values, this method is more robust to spatial misalignments and successfully detects the two defects in Figs. 5(a) and 6(a). However, it fails to detect the small defect shown in Fig. 4(a) and gives false positives when applied to a normal test image with large spatial misalignment as shown in Fig. 7(a).

In the optical flow-based method, the magnitude of the optical flow of each pixel within a $5 \times 5$ neighborhood window is calculated [14]. The final detection results are shown in Figs. 4(d2)–7(d2). This method is suitable for detecting the defect which is a sparse signal with a uniform background, where the defect can be modeled as a "sudden motion" given the template image as the previous frame of the test image. The "sudden motion" is represented as a large optical flow magnitude so that the defect can be segmented. However, it is not so successful in defect detection in electronic chips which have more complex structures. It cannot detect the defect shown in Fig. 4(a), whose optical flow is not large enough but is suitable to detect the more evident defect shown in Figs. 5(a) and 6(a). The optical flow-based method has more tolerance to the spatial shift of the test images than the traditional NCC-based method. However, it still cannot fully eliminate the effect of misalignment and mistake some normal regions for defects.

Aiger and Talbot's method is unsupervised and directly use POFT to each single test image to find the defect. Its detection results are shown in Figs. 4(e2)–7(e2). It can successfully detect the defect as shown in Fig. 4(e2). However, as pointed out in [27], it cannot detect defects structured as a scratch or line, which may be seen as regular patterns in a single test image. Seen from the detection results in Figs. 5(e2) and 6(e2), it fails to detect the two defects in Figs. 5(a) and 6(a). Furthermore, this method erroneously identifies some normal but irregular regions as defects.

In our method, a collection of 35 test images which are not aligned well is taken as an input image as shown in Fig. 8(a),

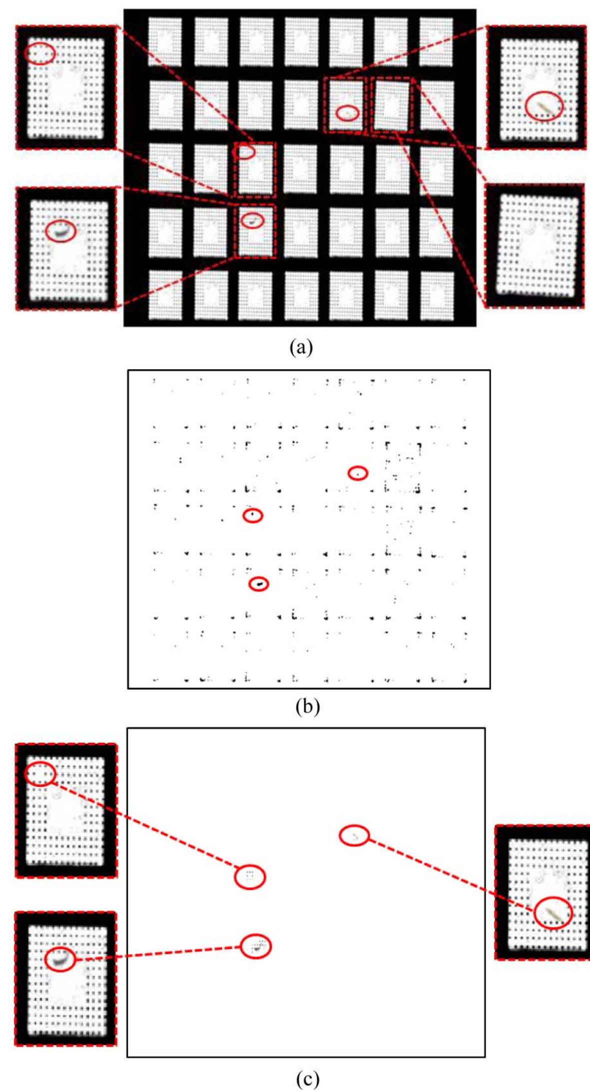


Fig. 8. Input image collection and defect detection results using our method. (a) Collection of $7 \times 5$ test images. (b) Binary saliency map obtained in Step 1 of our method ($C_s = 3$). (c) Final detection result obtained in Step 2 of our method.

in which the four test images shown in Figs. 4(a)–7(a) are included. Some images have horizontal, vertical, and rotational misalignments. The salient regions calculated from Step 1 of the proposed method are shown in Fig. 8(b), with the salient regions representing the defects marked by red circles. We can see that the pixels representing the defects have large salient values. After Step 2, the defects are successfully detected as shown in Fig. 8(c). Comparing to the former four methods, no false positive is detected with the existence of minor spatial shift between the test images and other tiny product variations.

All the methods presented above are applied to all the test images of type A, B, C, and D. Here, the precision, recall, and F-measure are used to evaluate the performance. Precision $P$ is computed as the ratio of the number of the defective pixels which are correctly detected to that of all the pixels which are detected as defective. Recall $R$ is defined as the ratio of the number of the detected defective pixels to that of all the

TABLE II
EXPERIMENTAL PERFORMANCE OF DIFFERENT METHODS

| Methods | | Precision rate | Recall rate | F-measure | Computational complexity per image | Average computational time per image (ms) |
|---|---|---|---|---|---|---|
| NCC-based template matching [14, 20] | | 0.56 | 0.91 | 0.61 | $O(w^2 \cdot N)$ (without fast algorithm)<br>$O(N)$ (with fast algorithm) | 77 (without fast algorithm)<br>21 (with fast algorithm) |
| Partial NCC-based template matching [16] | | 0.79 | 0.85 | 0.80 | $O(w^2 \cdot N)$ | 62 |
| Optical flow-based method [14] | | 0.66 | 0.81 | 0.69 | $O(w^2 \cdot N)$ (without fast algorithm)<br>$O(N)$ (with fast algorithm) | 87 (without fast algorithm)<br>25 (with fast algorithm) |
| Aiger and Talbot's method [27] | | 0.49 | 0.74 | 0.53 | $O(N \cdot \log(N))$ | 26 |
| Our method | ($C_s = 1$) | 0.95 | 0.99 | 0.96 | $O(N \cdot \log(N_x \cdot N_y \cdot N))$<br>($N_x \cdot N_y$ is the number of test images in the collection) | 73 |
| | ($C_s = 2$) | 0.97 | 0.98 | 0.97 | | 40 |
| | ($C_s = 3$) | 0.97 | 0.98 | 0.97 | | 30 |
| | ($C_s = 4$) | 0.98 | 0.94 | 0.97 | | 29 |

defective pixels. F-measure which combines the precision and recall is given by

$$F_\alpha = \frac{(1+\alpha) \cdot P \cdot R}{\alpha \cdot P + R} \tag{22}$$

where $\alpha$ is a weight coefficient, which is set as 0.3. The results of the precision, recall, and F-measure are shown in Table II for these five compared methods, from which we can see that the NCC, partial NCC, and the optical flow-based method have a relatively high recall rate but low precision rate because of mistaking the defect-free pixels as defective ones. The partial NCC-based method has a better precision rate than NCC-based, optical flow-based methods and Agier and Talbot's method, and it is more robust to spatial alignment. Agier and Talbot's method performs worse than the other methods because many normal regions are wrongly identified as defects, and its recall rate is not satisfactory because of its insensitivity to some types of defects such as lines or scratches. Our method achieves much better performance than the other four methods in terms of precision, recall, and F-measure. Additionally, with different parameter $C_s$ (varies from 1 to 4), the performance of the proposed method is similar.

The computational complexity and average computational time per test image of all these methods are also summarized in Table II, where $w$ is the size of the neighborhood window and $N$ is the number of pixels in the test image. In terms of computational efficiency, our method is comparable to the fast NCC-based, optical flow-based methods and Agier and Talbot's method when $C_s$ is set as 3 or 4. There is much influence for the value of parameter $C_s$ on the computational complexity, especially when it is set as 1 and many pixels are required for processing in Step 2.

### C. Further Analysis and Discussion

In Step 1, we fast locate candidate-defective regions using saliency detection model, and in Step 2, the true defects are extracted. Comparing with Agier and Talbot's method [28] which directly uses POFT in each single test image to find the defect, our method is different in that we use a collection of images together as the input for processing and utilize the self-similarities among the multiple images, making themselves form a global periodic pattern, which can be removed by POFT efficiently. So the proposed method has two major advantages over Agier and Talbot's method.

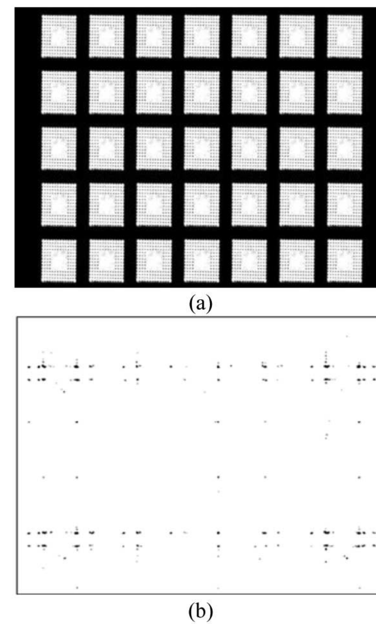



Fig. 9. (a) Collection of well aligned and defect-free test images of Type C. (b) Binary saliency map obtained in Step 1 by using the proposed method.

1) The defects such as scratches or lines which Agier and Talbot's method cannot be detected are successfully detected by the proposed method, as these types of defects are no more regular when compared to the global periodic pattern.
2) In Agier and Talbot's method, many normal regions with high contrast or irregular patterns, such as corners, are misclassified as defects.

In the proposed method, due to the effect of spatial misalignment, some normal regions still pop out in Step 1 as shown in Fig. 8(b). The template-matching-based comparison at each salient pixel in Step 2 ensures that only the true defects are detected, as shown in Fig. 8(c).

Ideally, if the test images are all detect-free and aligned well, all these normal regions will be automatically removed in Step 1 because of the global pattern. In practice, only a finite number of test images are available for processing. Due to the effect of window function in Fourier transform, some normal regions still pop out as salient regions in Step 1, even though there is no spatial misalignment. For example, there is a collection of well-aligned and defect-free images shown in Fig. 9(a). The binary

saliency map is obtained as Fig. 9(b). The proposed method will not detect any defect finally because of the image comparison used in Step 2. (The detection result is not shown here as it is a totally blank image.)

Another issue is how to determine the proper array size $N_x \times N_y$ of the collection of test images. $N_x \times N_y$ should be at least $2 \times 2$ to form a global periodic pattern among the multiple test images. In practice, it should be larger than $2 \times 2$ to include more defect-free test images into the collection. Consider an extreme case in which we only have $2 \times 2$ test images and all the four test images have defects at the same location. These four defects will form a periodicity which makes them difficult to be detected by POFT. More defect-free test images are needed to avoid such cases. Thus, it is preferred to have a relatively large array size. Meanwhile, given an image acquisition system, the image resolution of each test image will drop as the array size increases, so some tiny defects may not be captured. To ensure a good resolution for each test image, the array size should not be too large. Overall, for simplicity, the proper array size is set a little smaller or equal to an upper bound which is subject to some limitations, such as resolution requirement for each test image.

## V. Conclusion

Effective and efficient defect detection is important in semiconductor manufacturing and electronic production environment. Before packaging process, the dies are cut down from the wafer and aligned in an array for defect detection, and then the packaged electronic chips are delivered for further inspection. In this paper, we take a collection of test images aligned into an array as the input image for defect detection and have proposed a new method for fast defect-detection in images of electronic chips or dies from semiconductor production line. First, we apply the POFT to the whole collection of test images to obtain the salient regions. We theoretically prove that POFT can efficiently and effectively eliminate the periodic signals corresponding to the duplicated normal regions in multiple test images and greatly amplify the nonperiodic and sparse signals corresponding to the defects in the collection. However, there are spatial misalignment between the aligned test chips, which may also lead to salient regions. In the second step, the defects are extracted from the salient regions by matching the local image statistics around the salient pixels with that in the template image, in order to exclude false saliency. Experimental results have shown better performance of the proposed method compared with the relevant existing ones.

## References


[1] C. H. Yeh, F. C. Wu, W. L. Ji, and C. Y. Huang, "A wavelet-based approach in detecting visual defects on semiconductor wafer dies," *IEEE Trans. Semicond. Manuf.*, vol. 23, no. 2, pp. 284–292, May 2010.

[2] H. Gao, C. Ding, C. Song, and J. Mei, "Automated inspection of E-shaped magnetic core elements using K-tSL-center clustering and active shape models," *IEEE Trans. Ind. Informat.*, vol. 9, no. 3, pp. 1782–1789, Aug. 2013.

[3] W. C. Li and D. M. Tsai, "Defect inspection in low-contrast LCD images using Hough transform-based nonstationary line detection," *IEEE Trans. Ind. Informat.*, vol. 7, no. 1, pp. 136–147, Feb. 2011.

[4] R. Nakagaki, T. Honda, and K. Nakamae, "Automatic recognition of defect areas on a semiconductor wafer using multiple scanning electron microscope images," *Meas. Sci. Technol.*, vol. 20, pp. 075503–12, 2009.

[5] C.-S. Chen, C.-L. Huang, and C.-W. Yeh, "A hybrid defect detection for in-tray semiconductor chip," *Int. J. Adv. Manuf. Technol.*, vol. 65, pp. 43–56, 2013.

[6] L. Su, T. Shi, Z. Xu, X. Lu, and G. Liao, "Defect inspection of flip chip solder bumps using an ultrasonic transducer," *Sensors*, vol. 13, pp. 16281–16291, 2013.

[7] D. M. Tsai, S. C. Wu, and W. Y. Chiu, "Defect detection in solar modules using ICA basis images," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 122–131, Feb. 2013.

[8] N. G. Shankar and Z. W. Zhong, "Defect detection on semiconductor wafer surfaces," *Microelectron. Eng.*, vol. 77, pp. 337–34, 2005.

[9] D. E. B. Lee and P. D. Henshaw, "Printed circuit board inspection—A novel approach," in *Proc. SPIE Autom. Inspection Meas.*, 1986, pp. 164–173.

[10] G. Acciani, G. Brunetti, and G. Fornarelli, "Application of neural networks in optical inspection and classification of solder joints in surface mount technology," *IEEE Trans. Ind. Informat.*, vol. 2, no. 3, pp. 200–209, Aug. 2006.

[11] C. H. Li, C. Y. Chang, and M. Jeng, "Applying regional level-set formulation to postsawing four-element led wafer inspection," *IEEE Trans. Syst. Man, Cybern. C., Appl. Rev.*, vol. 41, no. 6, pp. 842–853, Nov. 2011.

[12] H. W. Kim and S. I. Yoo, "Defect detection using feature point matching for non-repetitive patterned images," *Pattern Anal. Appl.*, vol. 10, pp. 7–19, 2012.

[13] J. Kim, "Template-based defect detection of a brazed heat exchanger using an x-ray image," *Opt. Eng.*, vol. 52, p. 036501, 2013.

[14] D. M. Tsai, I. Y. Chiang, and Y. H. Tsai, "A shift-tolerant dissimilarity measure for surface defect detection," *IEEE Trans. Ind. Informat.*, vol. 8, no. 1, pp. 128–137, Feb. 2012.

[15] E. Elboher and M. Werman, "Asymmetric correlation: A noise Robust similarity measure for template matching," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3062–3073, Aug. 2013.

[16] C. C. Wang, B. C. Jiang, J. Y. Lin, and C. C. Chu, "Machine vision-based defect detection in IC images using the partial information correlation coefficient," *IEEE Trans. Semicond. Manuf.*, vol. 26, no. 3, pp. 378–384, Aug. 2013.

[17] D. Z. Wang, C. H. Wu, A. Ip, C. Y. Chan, and D. W. Wang, "Fast multi-template matching using a particle swarm optimization algorithm for PCB inspection," *Lect. Notes Comput. Sci.*, vol. 4974, pp. 265–370, 2008.

[18] A. J. Crispin and V. Rankov, "Automated inspection of PCB components using a genetic algorithm template-matching approach," *Int. J. Adv. Manuf. Technol.*, vol. 35, pp. 293–300, 2007.

[19] Z. Yang, "Fast template matching based on normalized cross correlation with centroid bounding," in *Proc. Int. Conf. Meas. Technol. Mechatron. Autom.*, Mar. 2010, vol. 2, pp. 224–227.

[20] S. Mattoccia, F. Tombari, and L. Stefano, "Efficient template matching for multi-channel images," *Pattern Recognit. Lett.*, vol. 32, pp. 694–700, 2011.

[21] G. Li, J. Shi, H. Luo, and M. Tang, "A computational model of visual attention for inspection of surface quality in production line," *Mach. Vis. Appl.*, vol. 24, pp. 844–835, 2013.

[22] G. Lei, Q. Ma, and L. Zhang, "Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.

[23] L. M. Zhang and W. Lin, *Selective Visual Attention: Computational Models and Applications*. Hoboken, NJ, USA: Wiley, 2013.

[24] L. Itti and A. Borji, "Selective visual attention: Computational models and applications," *Cognitive Neuroscience: The Biology of the Mind*, 5th ed. Hoboken, NJ, USA: Wiley, 2014, pp. 1–10.

[25] X. Hou, J. Harel, and C. Koch, "Image signature: Highlighting sparse salient regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 194–201, Jan. 2012.

[26] Y. Fang *et al.*, "Bottom-up saliency detection model based on human visual sensitivity and amplitude spectrum," *IEEE Trans. Multimedia*, vol. 14, no. 1, pp. 187–198, Feb. 2012.

[27] D. Aiger and H. Talbot, "The phase only transform for unsupervised surface defect detection," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, 2010, pp. 295–302.

[28] S. M. Rose, *Introduction to Probability and Statistics for Engineers and Scientist*, 4th ed. Amsterdam, The Netherlands: Elsevier, 2009.

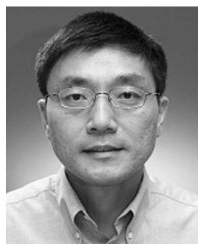
**Xiaolong Bai** received the Bachelor's degree in mechatronic engineering from Zhejiang University, China, in 2010, where he is currently pursuing the Ph.D. degree.

His research interests include ultrasonic imaging technique and instrumentation, image processing, and pattern recognition.

**Yuming Fang** received the B.E. degree from Sichuan University, Sichuan, China, in 2006; the M.S. degree from the Beijing University of Technology, Beijing, China, in 2009; and the Ph.D. degree from Nanyang Technological University, Singapore, in 2013, all in computer science.

From October 2011 to January 2012, he was a Visiting Ph.D. student at National Tsinghua University, Hsinchu, Taiwan. He is currently a Faculty Member with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China. From September 2012 to December 2012, he was a Visiting Scholar with the University of Waterloo, Waterloo, ON, Canada. He was also a (Visiting) Post-Doctoral Research Fellow with the IRCCyN Laboratory, PolyTech' Nantes & University of Nantes, Nantes, France; University of Waterloo; and Nanyang Technological University. His research interests include visual attention modeling, visual quality assessment, image retargeting, computer vision, and 3-D image/video processing.

Dr. Fang was a Secretary of Joint Conference on Harmonious Human Machine Environment (HHME) 2013. He was also a Workshop Organizer for the IEEE International Conference on Multimedia and Expo (ICME) 2014, and a Special Session Organizer for the IEEE Conference on Visual Communication and Image Processing (VCIP) 2013 and the International Workshop on Quality of Multimedia Experience (QoMEX) 2014.



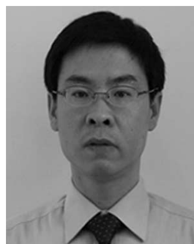
**Weisi Lin** (SM'98) received the Ph.D. degree in computer science from King's College London, London, U.K., in 1992.

He is an Associate Professor in Computer Engineering, Nanyang Technological University, Singapore. He was the Laboratory Head with the Institute for Infocomm Research, Singapore. His research interests include image processing, perceptual multimedia modeling and evaluation, and visual signal compression and communication. He has published more than 100 refereed journal papers and more than 190 conference papers, and held seven patents.

Dr. Lin has been on the editorial boards of the IEEE TRANSACTIONS ON MULTIMEDIA, IEEE SIGNAL PROCESSING LETTERS, and the *Journal of Visual Communication and Image Representation*, since 2011–2013. He currently chairs the IEEE Multimedia Communications Technical Committee Interest Group (MMTC IG) on Quality-of-Experience. He has been elected as an Asia Pacific Signal and Information Processing Association (APSIPA) Distinguished Lecturer (2012/2013). He is a Technical-Program Chair for the IEEE International Conference on Multimedia and Expo (ICME) 2013, the Pacific-Rim Conference on Multimedia (PCM) 2012, and the International Workshop on Quality of Multimedia Experience (QoMEX) 2014. He is a Fellow of the Institution of Engineering Technology and an Honorary Fellow of the Singapore Institute of Engineering Technologists.

**Lipo Wang** (SM'98) received the Bachelor's degree from the National University of Defense Technology, Changsha, China, and the Ph.D. degree from Louisiana State University, Baton Rouge, LA, USA, in 1983 and 1988, respectively, both in computer science.

He is a (co-)author of over 270 papers, of which more than 90 are in journals. He holds a U.S. patent in neural networks and a Chinese patent in Very Large Scale Integration (VLSI). His research interests include neural network systems and smart techniques with applications to communications, image/video processing, biomedical engineering, and data mining.



**Bing-Feng Ju** received the B.S. and Ph.D. degrees in mechanical engineering from Zhejiang University, Hangzhou, China, 1994 and 1999, respectively.

He was a Post-Doctoral Research Fellow with the Centre for Mechanics of Micro-Systems, Nanyang Technological University, Singapore, from 2000 to 2002. For a year (2002–2003), he was a Research Scientist with the Defence Science Organization (DSO) National Laboratories, Singapore. Since 2003, he has been a Foreign Researcher with the Japan Society for the Promotion of Science (JSPS), Tohoku University, Sendai, Japan. In 2005, he became a Faculty Member at Tohoku University. In April 2007, he joined Zhejiang University as a Full Professor. His research interests include ultraprecision metrology.

Dr. Ju is the recipient of several world famous scholarships and awards, including the Alexander von Humboldt Scholarship and the Rockwell Automation Award.