Simultaneous Photometric Correction and Defect Detection in Semiconductor Manufacturing

Yijiang Shen and Edmund Y. Lam

Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong

ABSTRACT

This paper reports on an image processing algorithm for simultaneous photometric correction and defect detection in semiconductor manufacturing. We note that this problem has some resemblance to change detection in real time image analysis. In particular, the changes between the two images are analogous to the defects in our machine vision system. We therefore applied several detection methods and examined their applicability to defect detection. We first performed a sub-pixel image registration, using a phase correlation method together with a singular value decomposition factorization of the correlation matrix to compute the necessary alignment. We then tested a few change detection methods, including the shading model, derivative model, statistical change detection, linear dependence change detector and Wronskian change detection model. We subjected this system to our collection of raw data acquired from an industrial system, and we evaluated the different methods with respect to the detection accuracy, robustness, and speed of the system. We have promising results at this stage, especially in detecting the blob and line defects that are most commonly found, and when the lighting variation is within a certain threshold.

Keywords: image registration, Phase Correlation Method (PCM), change detection, shading model, derivative model, statistical change detection, linear dependence change detector, Wronskian change detection model

1. INTRODUCTION

In semiconductor processing, contaminations and other patterning errors inevitably exist that would affect the yield of the integrated circuits. Defect detection is therefore a necessary step to identify possible problems with the dies before they undergo further processing.[1] There have been numerous efforts in the area using machine vision techniques. A common difficulty is the existence of photometric variation between the two images. Even when they are properly registered, i.e. proper scaling, translation or rotation have been performed, we cannot simply subtract one image from another and examine the difference. As these images are taken under different lighting conditions, the intensity value may differ even though they represent the same positions on the object.

One possible approach is to develop a linear model to approximate the spectral properties of the surfaces and illuminants with respect to a collection of sensing devices under different illumination conditions. Marimont and Wandell chose the linear-model basis functions by minimizing the error in approximating sensor responses for collections of surfaces and illuminants,[2] i.e. a low-dimension linear model was developed to approximate $S(\lambda)$ and $E(\lambda)$ with

$$I = e * n \int [S(\lambda) * E(\lambda) * F(\lambda)] d\lambda, \tag{1}$$

where λ is the wavelength, I is the pixel intensity, e is the orientation of the incident light, n is the normal of the viewed surface, $S(\lambda)$ is the surface reflectance, $E(\lambda)$ is the illumination spectrum, and $F(\lambda)$ is the camera response function. Healey and Benites analyzed the use of linear models for infrared spectral reflectance functions.[3] Other evaluations of linear models of surface spectral reflectance can be referenced to Ref. 4, Ref. 5 and Ref. 6. Finlayson and Hordley showed that there exists a single invariant color coordinate, a function of R, G, B that depends only on surface reflectance, and developed a log-chromaticity differences image out of the original image which is illumination invariant.[7] The linear model algorithms needed sufficient knowledge of the viewed surface, and

Further author information: (Send correspondence to Yijiang Shen)
Yijiang Shen: E-mail: yjshen@eee.hku.hk, Telephone: +852 63036167

Machine Vision Applications in Industrial Inspection XIV, edited by Fabrice Meriaudeau, Kurt S. Niel,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6070, 60700F, © 2006 SPIE-IS&T · 0277-786X/06/$15

the color constancy algorithms only focus on color images, whereas in the case of photometric correction of the die images in semiconductor manufacturing, the die images are gray scale images.

A common method for defect detection is known as die-to-die comparison. In this approach, an image of a known good die is taken, which must be defect free. Then we capture another image of the die under investigation. The two dies must contain identical circuit patterns. By comparing the two images, we can locate any discrepancy between them and identify it as a defect. But as mentioned earlier, even if the images are properly registered, image subtraction cannot be performed because of the difference in lighting conditions. We notice the resemblance of the photometric correction under defect detection and change detection in the sense that they both need to detect regions of change under different illuminations; the difference is that common change detection detect regions of change out of two images without registration problems, while die images may be subject to scaling, translation and rotation and thus change detection must be applied after an image registration process. In this paper, we generally deal with a reference image which is defect free, and a test image taken under different lighting conditions which needs to be detected for defects.

Many methods have been developed to measure the translational bulk displacement between similar images. The phase correlation method is popularly used because of its robust performance and computational simplicity.[8] Foroosh et al. developed an extended image registration algorithm based on Phase Correlation Method (PCM) which overcomes the disadvantage that PCM only gives out integral pixel level translational displacement.[9] This method requires an integral pixel displacement identification before the sub-pixel registration. Hoge further explored the PCM algorithm and developed a very attractive extension which combines non-integral pixel displacement identification, robustness to noise and limited computational complexity.[10] Other image registration methods can be found in Ref. 11. We will use the methods in Ref. 10 to align the test image with the reference image before applying change detection.

Change detection plays a very important role in our approach to detect defects. In Ref. 12, several change detection techniques which are robust to varying illumination are proposed, including the shading model (SM), the derivative model (DM), the statistical change detection (SCD), the linear dependence change detector (LDD) and the Wronskian change detection model (WM).[12] We will apply some of the methods to detect possible defects in the test image. Further exploration on change detection can be found in Ref. 13.

This paper is organized as follows. In the next section, we will describe the sub-pixel image registration method we use to align the reference image and the test image. In section 3, different change detection techniques will be discussed. Experimental results are given in section 4, with raw data acquired from an industrial system. And in section 5, some concluding remarks and future insights are provided.

2. SUB-PIXEL IMAGE REGISTRATION

The phase correlation method is based on the Fourier shift property. Specifically, a shift in the coordinate frame of two images $f(x, y)$ and $g(x, y)$ results in a linear phase difference in the Discrete Fourier Transform of the two images

$$\mathcal{F}_f(k, l) = \mathcal{F}_g(k, l)exp\{-j(ka + lb)\}, \tag{2}$$

where $\mathcal{F}_f(k, l)$ and $\mathcal{F}_g(k, l)$ are the Discrete Fourier Transform of two images $f(x, y)$ and $g(x, y)$ respectively; $f(x, y)$ and $g(x, y)$ are related with some translational shift. k, l are the discrete Fourier domain coordinates, and a, b are the magnitude of the horizontal and vertical displacements that occur between $f(x, y)$ and $g(x, y)$.

We can compute a normalized cross power spectrum between $g(x, y)$ and $f(x, y)$ to form a phase correlation matrix

$$\mathcal{Q}(k, l) = \frac{\mathcal{F}_f(k, l)\mathcal{F}_g(k, l)^*}{|\mathcal{F}_f(k, l)\mathcal{F}_g(k, l)^*|}, \tag{3}$$

where * is the conjugate transpose. Once the phase correlation matrix $\mathcal{Q}$ is computed, the inverse Fourier transform of $\mathcal{Q}$ is performed to identify a, b. If $g(x, y)$ and $f(x, y)$ are in fact related with some translational shift, then the inverse Fourier transform of $\mathcal{Q}$ would be a delta function

$$Q(x, y) = \mathcal{F}^{-1}(\mathcal{Q}(k, l)) = \delta(x - a, y - b), \tag{4}$$

where the peak of the function identifies the translational shift a, b.

However in case of images, the PCM algorithm will only give out integral translational displacement. In Ref. 10, a sub-pixel image registration method based on PCM and SVD factorization of the correlation matrix is proposed. In Ref. 10, the problem of finding the exact lateral shift between two images is recast as finding the rank one approximation of the normalized phase correlation matrix $\mathcal{Q}$. A straight forward way of finding this rank one approximation is by using singular value decomposition of $\mathcal{Q}$. From the left and right dominant singular vectors, we can identify independently the linear phase coefficients, thus we can estimate the horizontal and vertical translation shifts. To identify the linear phase coefficients in each right and left dominant singular vector, a least square fit to the phase component of the dominant singular vectors is used. Before the least square fit of the phase component, we will first correct the radian phase angles in the dominant singular vector by adding multiplies of $\pm 2\pi$ when absolute jumps between consecutive vector elements are greater than π radians in order to get a smooth increasing or decreasing feature of the radian phase angles of the singular vector. For an image $g(x, y)$ of size $M \times N$, let $k = 2\pi x/M$ and $l = 2\pi y/N$. Let μ and c be the slope and abscissa of the fitted line; for a singular vector v, by constructing the set of normal equations

$$R \begin{bmatrix} \mu \\ c \end{bmatrix} = \mathcal{V}, \tag{5}$$

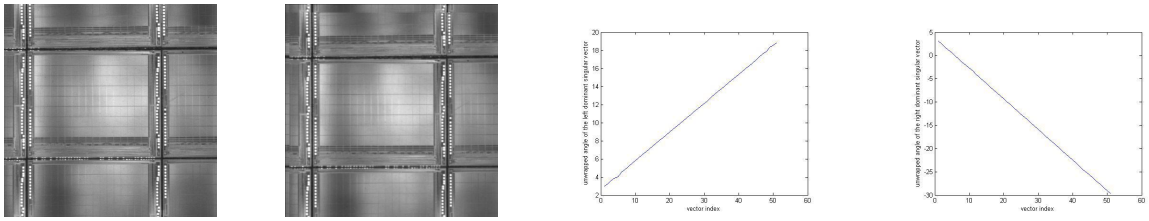
where the rows of R are equal to $\begin{bmatrix} r & 1 \end{bmatrix}$ for $r = \{0, 1, 2 \ldots (s-1)\}$ with s equal to the length of v; $\mathcal{V}$ is the corrected radian phase angles of $\angle v$, and $\angle$ returns the radian phase angles of vector v.

The slope of the fitted line, μ , maps to the translational shift. μ and c can be solved by

$$\begin{bmatrix} \mu \\ c \end{bmatrix} = (R^T R)^{-1} R^T \mathcal{V}, \tag{6}$$

specifically, $a = \mu(M/2\pi)$ when v is the left dominant singular vector and $b = \mu(N/2\pi)$ when v is the right dominant singular vector.

In addition, in order to restrict spectral component corrupted by aliasing from the shift estimation, a mask can be used to capture the components of Q with a magnitude larger than a threshold within a radius from the spectrum origin. Another advantage of the method is that it is robust to the presence of noise. We show a simple example of the method in Fig. 1. The image alignment method yields 10.4 and 20.5 pixels translation along horizontal and vertical directions respectively.



(a) (b) (c) (d)

Figure 1. Image registration example. (a) The original image. (b) The shifted image, 10.5 and 20.5 pixels along horizontal and vertical directions respectively. (c) The fitted line to the corrected radian phase angles of the left dominant singular vector. (d) The fitted line to the corrected radian phase angles of the right dominant singular vector.

3. CHANGE DETECTION TECHNIQUES

In this section, we will discuss some of the change detection techniques which are robust against illumination conditions. Change detection plays a very important role in real-time image analysis. In the case of photometric correction and defect detection, which detect defects from die images taken under different lighting conditions,

it is quite straight forward to apply change detection techniques to the reference image and the test image to detect regions of changes in the images and the changes would be defects.

In Ref. 12, several change detection methods are described, which are robust to illumination variations. The shading model method, the derivative model method and the statistical change detection method are categorized as region based methods, while the linear dependence change detector method and the Wronsian change detection model methods are categorized as vector based methods. We will describe the methods below and apply the methods to the raw data images.

3.1. The Shading Model (SM)

The model was first introduced by Oppenheim[14] and is commonly used in computer graphics. In this model,

$$I_p = I_i S_p, \tag{7}$$

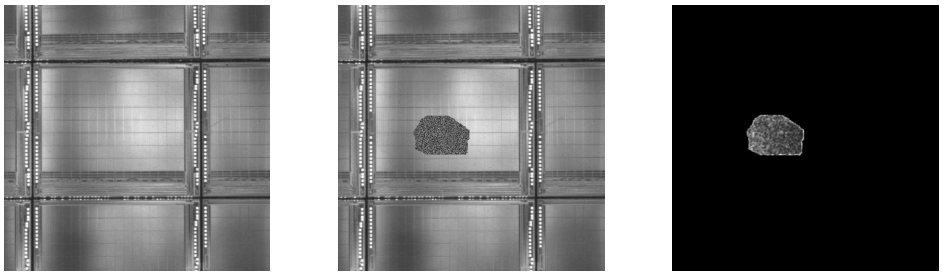
where I_p is the intensity value on a given point, I_i is the illumination and S_p is the shading coefficient.

The main point of the model is that if there is no physical change between two given areas of the images, then the intensity ratio of (7) becomes

$$\frac{I_{p1}}{I_{p2}} = \frac{I_{i1}}{I_{i2}}, \tag{8}$$

where I_{p1} and I_{p2} are intensity values of two images, I_{i1} and I_{i2} are the illuminations.

The intensity ratio in (8) is not dependent on the shading coefficients. For change detection with the SM, the reference die image and the test die image are cut into fixed block areas of interest (AOI). The variance of the intensity ratio described in (8) for each intensity value in the corresponding AOIs of the reference image and the test image is computed. If the value of the variance exceeds a certain predetermined threshold, the test image is marked as changed and a defect is detected. We will use a reference image and a test image with defect to show a example of the SM methods in Fig. 2. The size of the threshold window is 4. From the test, we can see that without the problem of image alignment, the SM method gives out good results; the defect area is well detected. The speed of the SM method is slow.



(a) (b) (c)

Figure 2. SM method on the reference image on the test image. (a) The reference image with a defect. (b) The test image (a defected area in the reference image(a)). (c) The result of the SM method.

3.2. The Derivative Model (DM)

The main point of DM is that the intensity value of a given point in the AOIs is described as a second-order bivariate polynomial of the pixel coordinates of the AOI window. If the AOI window is of size 3×2, the gray-value surface function $s_i(x, y)$ in the window is defined as

$$s_i(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy + a_{20}x^2 + a_{02}y^2, \tag{9}$$

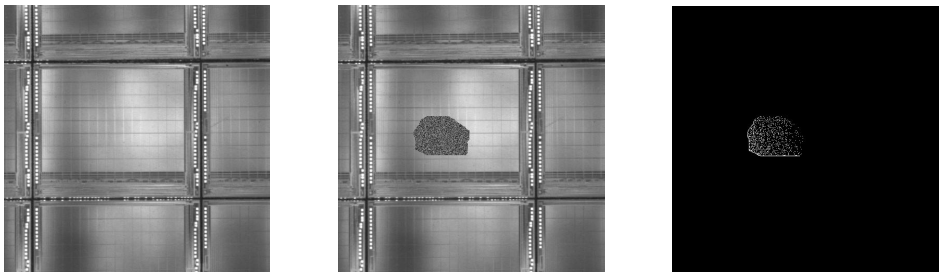
where $\{a_{00}, a_{10}, a_{01}, a_{11}, a_{20}, a_{02}\}$ are the six unknown coefficients. The change detection difference metric for every AOI A is

$$D = \sum_{x,y \in A} m_1(x,y) - m_2(x,y), \tag{10}$$

with m_i defined as:

$$m_i(x,y) = \frac{\partial s_i(x,y)}{\partial(x)} + \frac{\partial s_i(x,y)}{\partial(y)}, \quad for\ i = 1, 2. \tag{11}$$

When the difference of $m_1(x,y)$ and $m_2(x,y)$ in a AOI is bigger than a predetermined threshold, the area is marked with changes and a defect is detected. We will still use the reference image and the test image in Fig. 2 to test the DM method.



(a) (b) (c)

Figure 3. DM method on the reference image and the test image. (a) The reference image without a defect. (b) The test image with a defect. (c) The results of the DM method.

From Fig. 3, we can see that the DM method has difficulties in detecting semantic objects, and the boundary points detected are discrete. The computation speed of DM method is faster than the SM method.

3.3. The Statistical Change Detection (SCD)

The main idea behind the SCD method is that SCD assumes the change detection could be detected as a statistical hypothesis testing with: H_0: no change occurred at a pixel; H_1: a change occurred at a pixel. Another assumption of SCD is that the noise can be represented by either a Gaussian or a Laplacian distribution. The hypothesis testing is done by

$$\Delta_i = \frac{\sqrt{2}}{\sigma_0} \sum_{x,y \in w_i} |d(x,y)|, \tag{12}$$

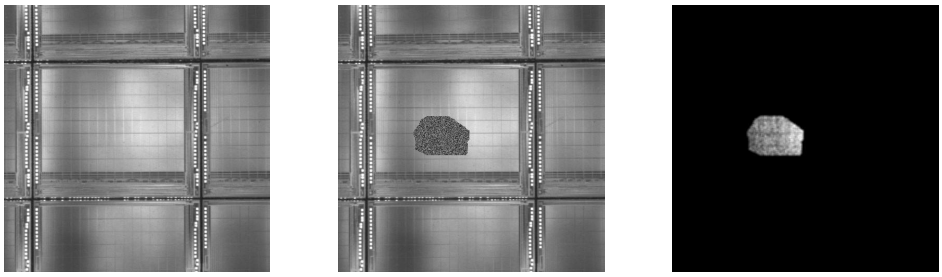
where Δ_i is the sum of absolute differences within a sliding window w_i, of size N pixels and centered at position i, σ_0 is the noise standard deviation of the gray-level differences $d(x,y)$ and is assumed to be constant over space; $d(x,y)$ is the gray-level difference at space locations (x,y), and $d(x,y)$ is defined in

$$d(x,y) = f(x,y) - g(x,y), \tag{13}$$

where $f(x,y), g(x,y)$ are the images under consideration. The significance value α is defined to be the probability that $\Delta_i > T$(T is a predetermined threshold) under the hypothesis H_0 that d_k is due to noise but not a significant change:

$$\alpha = Prob(\Delta_i > T | H_0) \tag{14}$$

Again, we will test the methods using the same reference image and the test image as in SM and DM. From Fig. 4, we can see that the boundary points and the interior of the area are detected. The computation complexity of the SCD method is moderate. From Ref. 12, the SCD method has difficulties with changing illumination conditions.



(a) (b) (c)

Figure 4. SCD method on the reference image and the test image. (a) The reference image without a defect. (b) The test image with a defect area. (c) The result of the SCD method.

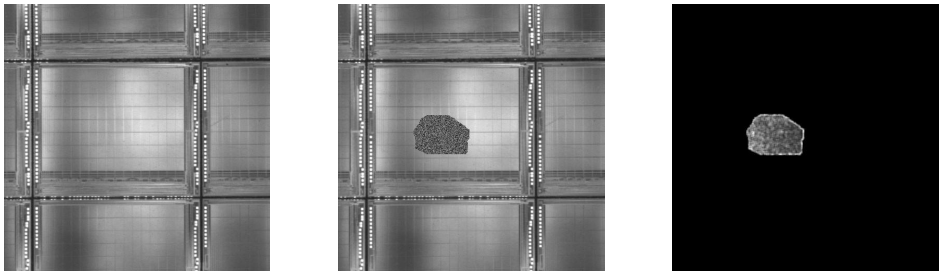
3.4. The Linear Dependence Change Detector (LDD)

Each pixel of an image with its neighboring area is the region of support for the corresponding vectors. The center pixel of a region of support is replaced by the vectors. For example, if the region of support is 3×3, the center pixel of the region is replaced a 9×1 vector. Further more, let $\vec{X} = (x_1, ..., x_n), \vec{Y} = (y_1, ..., y_n)$ be vectors of two corresponding regions of support from a reference image and a current image, then the ratio of their components is constant, i.e., $x_1/y_1 = \cdots = x_n/y_n = k$, and thus the variance σ^2 where

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\frac{x_i}{y_i} - \mu)^2 \qquad with \qquad \mu = \frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{y_i} \tag{15}$$

must be zero.

From (15), a criterion for change between a reference and a current image can be introduced for the Linear Dependence Change Detector: $\sigma_2 = 0$; $\vec{X}$ and $\vec{Y}$ are linearly dependent, i.e., no change has occurred;$\sigma_2 > 0$; $\vec{X}$ and $\vec{Y}$ are linearly independent, i.e., change has occurred. We will still use the reference image and the test image to test the LDD method. The computation complexity of the LDD method is not high.



(a) (b) (c)

Figure 5. LDD method on the reference image and the test image. (a) The reference image without a defect. (b) The test image with a defect area. (c) The result of the LDD method.

3.5. The Wronskian Change Detector Method (WM)

The Wronskian $W(\{f_i\}_{i=1}^n; x)$ is of a set of functions $\{f_i(x)\}_{i=1}^n$, each of which possesses derivatives of order $n-1$, is defined as:

$$W(\{f_i\}_{i=1}^n; x) := \begin{vmatrix} f_1(x) & f_2(x) & ... & f_n(x) \\ f_1^{'}(x) & f_2^{'}(x) & ... & f_n^{'}(x) \\ ... & ... & ... & ... \\ f_1^{n-1}(x) & f_2^{n-1}(x) & ... & f_n^{n-1}(x) \end{vmatrix} \tag{16}$$

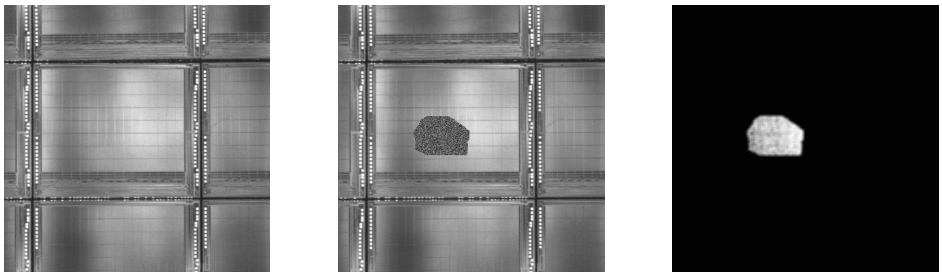
Given two components $x(E)$ and $y(E)$ of the vectors $\vec{X}$ and $\vec{Y}$ from a reference and a current image, $x(E)$ and $y(E)$ are functions of the illuminance E. As mentioned in section 3.4, if $\vec{X}$ and $\vec{Y}$ are linearly independent, change has occurred. To detect the changes between two images, we initially assume that there has been no change, i.e., $x(E) = y(E) = E$, and consider the linear combination of $x(E)$ and $y(E)$: $x(E)k_1 + y(E)k_2 = 0$, thus we have $\frac{x(E)}{y(E)} \cdot k_1 + 1 \cdot k_2 = 0$, if k_1 and k_2 are equal to zero, by which means $x(E)$ and $y(E)$ are linear dependent and no change has occurred, the Wronskian of functions $\frac{x(E)}{y(E)}$ and 1 has to be zero; the Wronskian of them is computed as

$$0 = \begin{vmatrix} \frac{x}{y} & 1 \\ (\frac{x}{y})' & 0 \end{vmatrix} = -(\frac{x}{y})' = -\frac{x'}{y} - x(\frac{1}{y}) = \frac{x}{y^2} - \frac{1}{y} = \frac{x^2}{y^2} - \frac{x}{y}, \tag{17}$$

with $x' = dx/dE$. Applying it to all components of $\vec{X}$ and $\vec{Y}$ we have

$$W(\frac{x_i}{y_i}) = \frac{1}{n}\sum_{i=1}^{n}\frac{x_i^2}{y_i^2} - \frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{y_i} = 0. \tag{18}$$

We use equation (18) to detect the changes in the reference image and the test image to test the results of WM: WM is computed in a window of a certain size (7×7 in our case), if the WM value in the window is bigger than a certain threshold, a change is detected. The WM method is of low computation complexity. Another advantage of the WM method is that it detects not only the contours of the changed areas, but also the inside of the changed areas. We will use WM method to detect the defects after the reference image and the test image are aligned.



(a) (b) (c)

Figure 6. WM method on the reference image and the test image. (a) The reference image without a defect. (b) The test image with a defect area. (c) The result of the WM method.

4. EXPERIMENTAL RESULTS

We have described the image registration technique in section 2, and change detection techniques in section 3. To be able to test the performance of the algorithm, now we will apply the techniques to detect possible defects in a test image with reference to a reference image, which are captured under different lighting conditions. The detect procedure are as follows: firstly, we will use the image registration technique to identify the translational shifts between the images; secondly, we will align the test image with the reference image; then we will use the change detection methods mentioned in section 3 to detect the changes, i.e., possible defects.

Two examples are given in Fig. 7 and Fig. 8. Fig. 7(a) and Fig. 8(a) are the reference images with no defect from an industrial system. Fig. 7(b) and Fig. 8(b) are the test images with defects of different shape; the defects are added to another defect-free image. The reference images and the test images are taken under different illuminations, and the dies in the images are shifted. Fig. 7(c), Fig. 7(d), Fig. 7(e), Fig. 7(f), Fig. 7(g) are the results of defect detection when applying SM, DM, SCD, LDD and WM methods respectively. From Fig. 7(g),

we can see that the WM method detects both the boundaries and the inside parts of the defected area; the boundaries of the detected defected area are closed and correspond to those of the real defected area in Fig. 7(b). In Fig. 7(c), the SM method has similar results as the WM method, but it is more time consuming than the WM method; in Fig. 7(d), the DM method has difficulties detecting the inside part of the defect area and the boundary of the detected area is discrete; in Fig. 7(e), the result is not good as some pixels in the defected area are not detected while some pixels which are not defected are detected, it is referred from Ref. 12 that the SCD method can not handle changing illumination conditions in general and shadows or reflections in particular; in Fig. 7(f), the LDD method yields similar results as the WM method and the computational complexity is low. It is natural we use the LDD method and the WM method to detect the defected area in our approach. Fig. 8(c) and Fig. 8(d) are the results of defect detection applying the LDD and the WM methods respectively. We can also increase the size of the windows to improve performance, but the computational complexity increase as well.

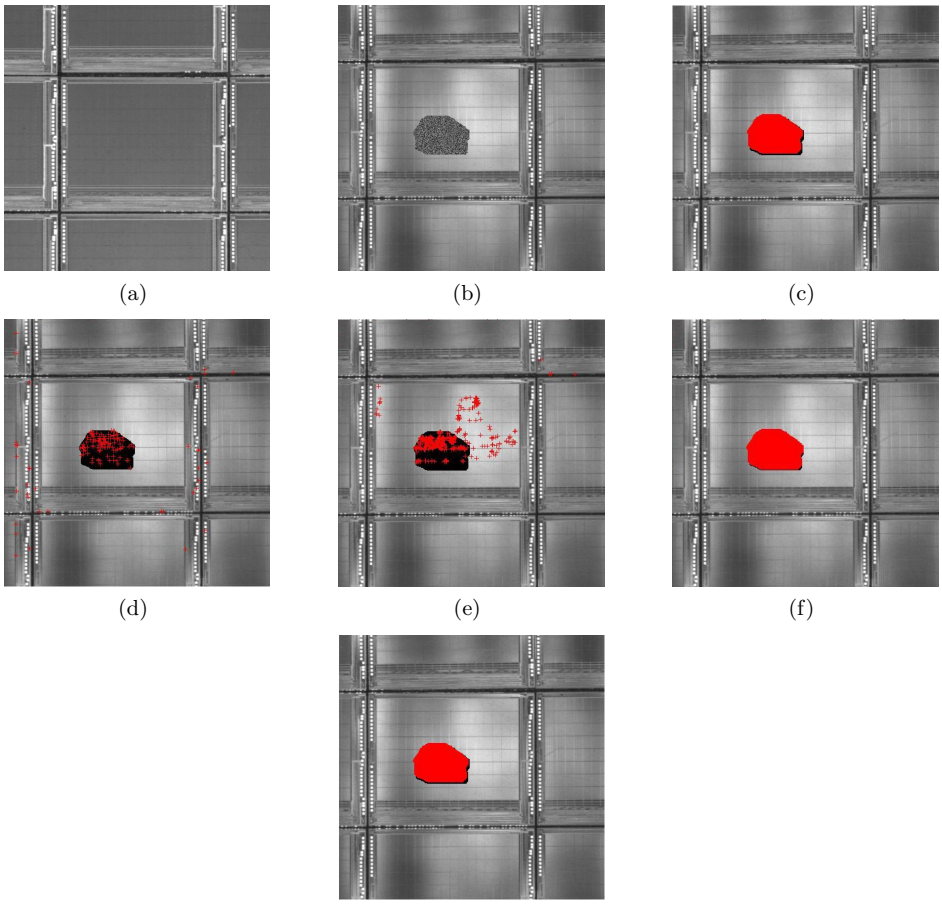
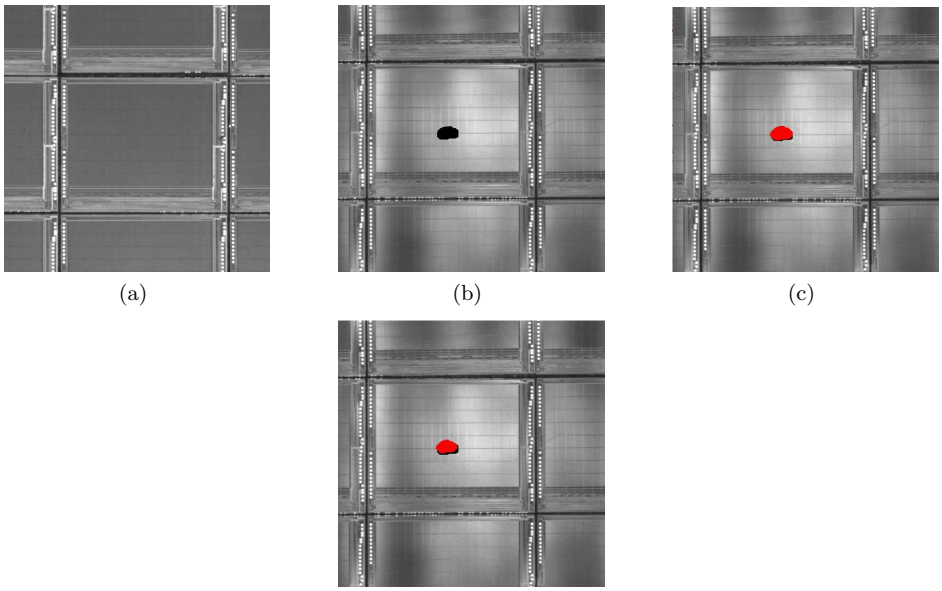


Figure 7. Defect detection on the reference image and the test image. (a) The reference image without a defect. (b) The test image with a defect area. (c) The result of the SM defect detection. (d) The result of the DM defect detection. (e) The result of the SCD defect detection. (f) The result of the LDD defect detection. (g) The result of the WM defect detection.



(a) (b) (c)

(d)

Figure 8. Defect detection on the reference image and the test image. (a) The reference image without a defect. (b) The test image with a defect area. (c) The result of the LDD defect detection. (d) The results of the WM defect detection.

5. CONCLUSIONS AND FUTURE INSIGHTS

We have discussed in this paper a new machine vision technique to detect defects on die images under varying illuminations. The technique first performs a sub-pixel registration described in section 2 to compute the image alignments, and then applies change detection methods (LDD and WM) to detect the defects. No prior knowledge is required of the viewed surface and lighting condition; the computation complexity of the technique is low.

REFERENCES

1. E. Y. Lam, "Robust minimization of lighting variation for real-time defect detection," *Real-Time Imaging* **10**(6), pp. 365–370, 2004.
2. D. Marimont and B. Wandell, "Linear models of surface and illuminant spectra," *J. Opt. Soc. Am. A* **9**(11), pp. 1905–1913, 1992.
3. G. Healey and L. Benites, "Linear models for spectral reflectance functions over the mid-wave and long-wave infrared," *J. Opt. Soc. Am. A* **15**(8), pp. 2216–2277, 1998.
4. L. T. Maloney, "Evaluations of linear models of surface spectral reflectance with small number of parameters," *J. Opt. Soc. Am. A* **3**(10), pp. 1673–1683, 1986.
5. W. A. Richards, "Lightness scale from image intensity distributions," *Applied Optics* **21**(14), pp. 2569–2582, 1982.
6. R. O. Dror, E. H. Adelson, and A. S. Willsky, "Estimating surface reflectance properties from images under unknown illuminations," in *Human Vision and Electronic Imaging IV*, *Proc. SPIE* **4299**, 2001.
7. G. D. Finlayson and S. D. Hordley, "Color constancy at a pixel," *J. Opt. Soc. Am. A* **18**(2), pp. 253–264, 2001.
8. C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," *Prot. Int. Conf. on Cybernetics and Society* , pp. 163–165, 1975.
9. H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE. Trans. Image Processing* **11**(3), pp. 188–200, 2002.

10. W. S. Hoge, "A subspace identification extension to phase correlation method," *IEEE. Trans. Medical Imaging* **22**(2), pp. 277–280, 2003.
11. B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing* **24**, pp. 977–1000, 2003.
12. E. Durucan and T. Ebrahimi, "Change detection and background extraction by linear algebra," *Proc. IEEE* **89**(10), pp. 1368–1381, 2001.
13. R. J. Radke, S. Andra, O. Al-kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Processing* **14**, pp. 294–307, 2005.
14. A. V. Oppenheim, R. W. Schafer, and T. G. Stockham, "Nonlinear filtering of multiplied and convolved signals," *Proc. IEEE* **56**, pp. 1264–1291, 1968.