Moiré interferogram phase extraction: a ridge detection algorithm for continuous wavelet transforms

Heng Liu, Alexander N. Cartwright, and Cemal Basaran

We present a procedure using continuous wavelet transforms (CWTs) to extract the phase information from moiré interferograms. The relationship between precise ridge detection of the two-dimensional CWT magnitude map and accurate phase extraction is detailed. A cost function is introduced for the adaptive selection of the ridge, and a computationally inexpensive implementation of the cost function ridge detection algorithm is explored with dynamic programming optimization. The results of the proposed ridge detection algorithm on actual interferograms are illustrated. Moreover, the resulting extracted phase is demonstrated to be smooth and accurate. As a result, the sensitivity of the moiré interferometry method is improved to obtain a pixel-by-pixel in-plane strain distribution map. © 2004 Optical Society of America

OCIS codes: 120.2650, 100.2000, 100.3010, 100.5070.

1. Introduction

With the rapid development of the semiconductor integrated circuit industry, electronics packaging has become more essential to improve the overall integrated circuit performance. As a great number of functions and components are integrated on a die or chip, the reliability of the packages is becoming even more important. It is well known that the major failure mode of the common multilayered structure used in electronics packaging is due to the coefficient of thermal-expansion mismatch of the different materials in the adjacent layers. In-plane strain of the multilayer packages that results from thermal loading, vibration loading, and high-density electrical loading is well documented. Moiré interferometry,[1] which was demonstrated as a high-sensitivity nondestructive tool to measure the in-plane deformation, has been widely used to study the mechanical reliability problem in the electronics packaging area.[2–4]

In moiré interferometry, a holographic diffraction grating pattern is applied to the cross section of a specimen. The two symmetric incoming laser beams are each diffracted from this grating normal to the specimen surface. The two diffracted beams thus form an interference fringe pattern. The deformation of the specimen, i.e., the deformation of the diffraction grating, is ascertained by the change of the observed fringe pattern. The analysis of the moiré fringe pattern enables us to calculate the in-plane strain distribution and understand the response of the package to external factors, such as temperature variation, environmental vibration, and high-density electrical current. The details of moiré interferometry have been addressed by Post *et al.*,[1] and the application to electronics packaging has been demonstrated by a number of groups.[2–4]

Generally, in any interferometer the acquired interferogram is readily expressed as a modulation of the phase of the intensity:

$$I = I_0(1 + \gamma \cos \varphi), \qquad (1)$$

where I_0 is the average image intensity, γ is the fringe visibility, and φ is the phase of the interferogram. More importantly, φ is directly related to the physical parameter of interest. The analysis of the mechanical problem with moiré interferometry is usually based on fringe analysis. According to the definition of the in-plane strain, one can easily derive the relationship between the strain and the fringe count.[1] The magnitude of the strain is directly proportional to the density of the fringes. The most common grating used in moiré interferometry has a pitch density of 1200 lines/mm, which leads to a deformation sensitivity of 0.417 μm/fringe.[1]

The authors are with the Electronic Packaging Laboratory, the State University of New York, Buffalo, New York 14260. A. N. Cartwright's e-mail address is anc@eng.buffalo.edu.

With the continuing reduction of the size of microelectronic components and their packages, such micrometer-level sensitivity cannot resolve the strain at the dimensions of interest to industry. Thus a number of methods have been introduced to improve the sensitivity of this technique. For example, spatial carrier fringes have been introduced.[5] In this case the interferometer is intentionally misaligned to produce a dense fringe pattern in the interference field, which are called carrier fringes. The existence of the spatial carrier fringes results in more fringe lines for the same amount of deformation. Therefore the sensitivity, i.e., the actual deformation represented by each fringe line, is improved. However, although the sensitivity is improved, there is an accompanying reduction in the accuracy when spatial carrier fringes are used.[6] Generally an error as large as one tenth of a fringe is present in the measurement of a fringe center. Therefore the more fringes that are involved, the more error is introduced.

A second method for improving the sensitivity is to directly determine the phase distribution of the interferogram. All the phase information is actually present in the interferogram. The fringe analysis mentioned above is just a simplified version of phase analysis. Specifically, with the fringe counting method, only locations where φ is a multiple of 2π are considered. In moiré interferometry[1] we use the density of the fringes to calculate the in-plane strain, resulting in the 0.417-μm/fringe sensitivity for a 1200-lines/mm grating. Therefore, if we can reconstruct the phase map out of the fringe patterns, the in-plane strains can be expressed as a function of the phase map:

$$\varepsilon_x = \frac{1}{2\pi f}\frac{\partial \varphi_u}{\partial x},$$

$$\varepsilon_y = \frac{1}{2\pi f}\frac{\partial \varphi_v}{\partial y},$$

$$\gamma_{xy} = \frac{1}{2\pi f}\left(\frac{\partial \varphi_u}{\partial y} + \frac{\partial \varphi_v}{\partial x}\right), \qquad (2)$$

where f is the virtual grating frequency; φ_u and φ_v are the phase of the horizontal and vertical interferograms, respectively; and ε_x, ε_y are normal strains and γ_{xy} is the shear strain to be measured.

A number of noise issues must be considered to precisely extract phase information from a fringe pattern taken by a monochromatic CCD camera. For example, spurious reflections, beam power fluctuations and nonuniformity, dust diffraction, and grating defects all lead to random variations of the background and fringe visibility. Because of the specific features of these localized noise types, a continuous wavelet transform- (CWT-) based filter that beats the conventional Fourier-transform-based filter is becoming a more accepted tool for interferogram image processing.[7–9] Basically, in fast Fourier transform, the frequency spectrum represents the

characteristic of the whole image and the local information is unavailable. On the contrary, CWT analysis eliminates the noise based on the local frequency distribution. As a result, CWT can be used to address more-complicated local noise. The details of CWT-based filtering are described in Section 2. As we show, CWT can fulfill the task of denoising as well as provide the ability to directly extract the phase information from fringe patterns. In the procedure to extract the phase, we apply CWT on each line of the interferogram image, obtain the two-dimensional transform magnitude map, select the ridge of the two-dimensional map, and track the phase of the ridge. We prove that the phase of the ridge is identical to the phase of the interferogram. The accuracy of the CWT process relies greatly on the ridge detection.[8,10] In this study we propose a new ridge detection algorithm based on dynamic optimization.[11] We demonstrate that these techniques are reliable for the determination of the ridge and can be implemented at a low computation cost.

2. Continuous Wavelet Transform Process

The CWT can be understood as a correlation of the signal and a pulse mother wavelet function:

$$S(a, b) = \frac{1}{a}\int_{-\infty}^{\infty} s(t) M\left(\frac{t-b}{a}\right)\mathrm{d}b, \qquad (3)$$

where a, b are the scaling parameter and the shifting parameter, respectively; $s(t)$ is the time and position domain signal; and $S(a, b)$ is the two-dimensional transform as a function of a and b. It is well known that the interferogram fringe patterns can be represented as a frequency-modulated sinusoidal function as in Eq. (1). Therefore we select the Morlet wavelet, which is a Gaussian-weighted complex function with sinusoidal characteristics,

$$M(x) = \exp\left(-\frac{x^2}{2}\right)\exp(j\omega_0 x), \qquad (4)$$

to use in moiré image filtering. In this study we set ω_0 to be 2π. The ridge of the CWT is defined as the location where the CWT magnitude reaches its local maximum along the scaling direction a.[8–10] By assuming a slowly varying fringe contrast γ and instantaneous spatial frequency ω_s, we can derive the ridge of the CWT magnitude map as

$$S(a_0, b) = \frac{1}{2} I_0 \gamma \sqrt{2\pi}\, \exp(\mp j\omega_s b). \qquad (5)$$

Obviously it can be seen from Eq. (5) that the ridge contains the phase information $\omega_s b$ of the original interference signal. In the case of a monotonically increasing (or decreasing) fringe-order distribution, the phase of the ridge is exactly the same as the phase of the signal. Otherwise, a phase-shifting method would need to be utilized to determine the sign of the gradient of the phase.
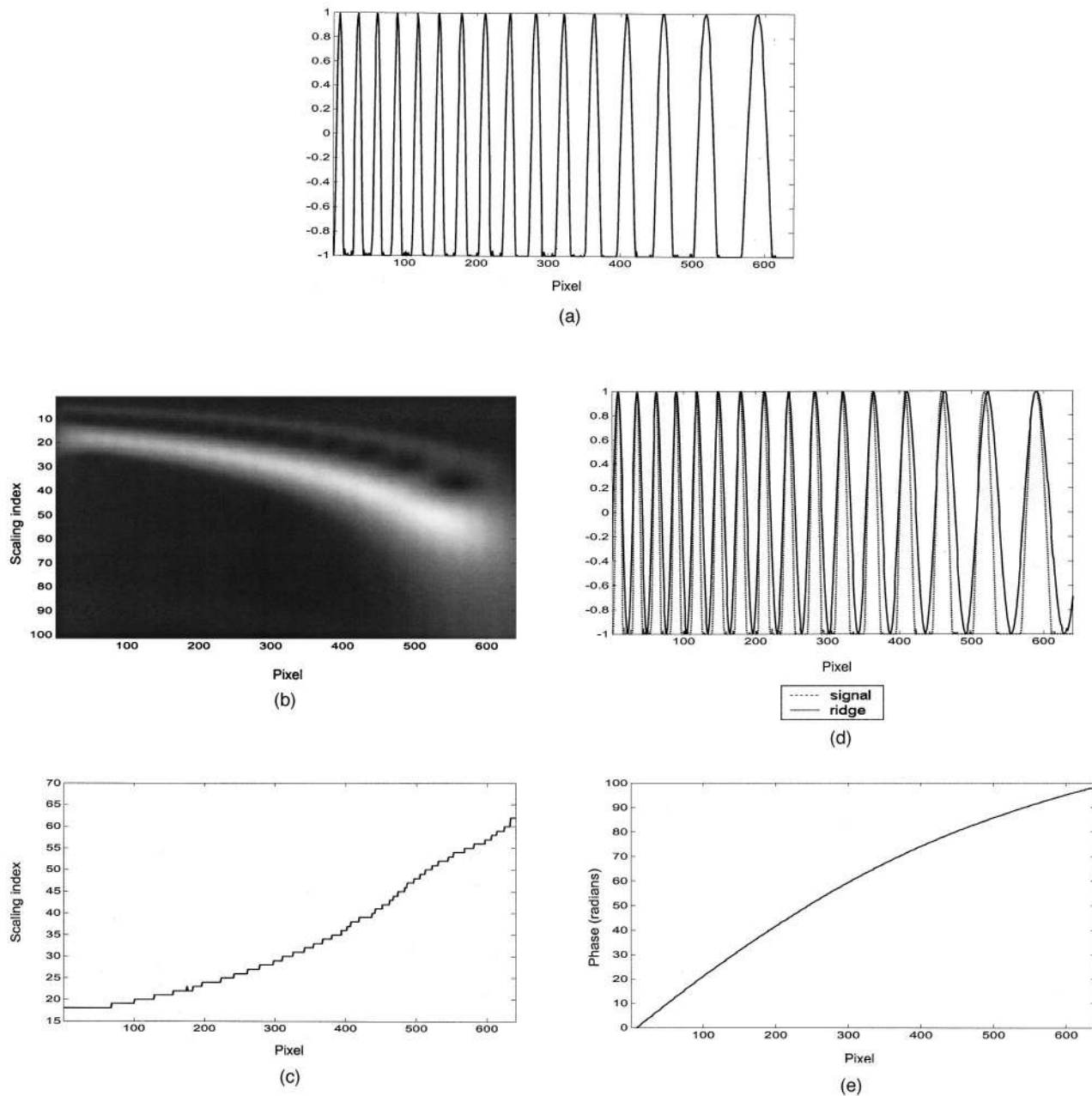
Fig. 1. Simulation of the direct ridge detection method: (a) original signal—a truncated chirped cosine signal, (b) the CWT magnitude, (c) the ridge index, (d) the real part of the detected ridge, (e) the phase extracted from the ridge.

3. Ridge Detection Algorithms

A. Direct Maximum Ridge Detection

As described in Section 2, once the image has been processed with CWT, the problem is reduced to the development of a reliable method to detect the ridge of the two-dimensional CWT. By definition, the ridge corresponds to the maximum of the CWT magnitude. Therefore the direct algorithm is to detect the maximum magnitude of all $S(a, b_i)$ curves for every b_i position. We refer to this method as the direct maximum ridge detection algorithm. An example of this method is shown in Fig. 1. Specifi-

cally, as shown in Fig. 1(a), the simulated signal is a truncated chirp sinusoidal signal without noise. As a result of the perfect signal quality, in the two-dimensional map of the CWT magnitude as shown in Fig. 1(b), the white part, which corresponds to the largest gray-scale intensity, clearly distinguishes itself from the dark part. Therefore the simple ridge detection equation

$$\text{ridge}(b_i) = \max(|S(a_j, b_i)|), \qquad (6)$$

where j equals all of the scaling parameters, works well. The selected ridge index k, where a_k is selected as the ridge, of all b_i positions is shown in Fig. 1(c).
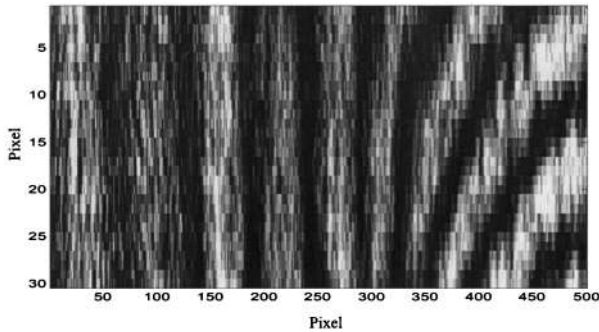
Fig. 2. Original noisy interferogram.

Figures 1(d) and 1(e) are the real part and the phase of the corresponding selected ridge, respectively. Obviously, the frequency of the resulting ridge shown in Fig. 1(d) follows the instantaneous frequency of the original signal closely, and the phase extracted from the ridge represents that of the signal.

Unfortunately, although the above result from a noiseless sinusoidal signal is satisfying, actual real-world signals will inevitably be affected by a number of noise sources. In this case, one would need to consider how the noise sources would contribute to the CWT magnitude. Actually, in moiré interferometry most noise sources result in higher frequencies than that of the signal in a local area. This serves as the basis for differentiating the interferogram signals from the coexisting noise. In the given range of the potential scaling parameters array a, the magnitude of the CWT is the result of both the noise and the signal. With good signal-to-noise ratio, i.e., if the human eye can easily resolve the fringes from the noise, the contribution of the signal will be much larger than that of the noise. As a result, the ridge detected by the direct maximum algorithm is clearly located at the signal frequency. However, in the case of a low signal-to-noise ratio, the magnitude of the CWT from the noise might be as strong as that from the signal. In this case, we will have difficulty in deciding which scaling position a corresponds to the signal frequency. If the image quality degrades further, the noise part of the CWT might even dominate. The direct maximum ridge detection will incorrectly extract the noise frequency as the signal frequency. Therefore, to process an image with less than perfect image quality, which is usually the case in an actual experiment as shown in Fig. 2, a more robust algorithm to fulfill ridge detection is needed.

B. Cost Function Ridge Detection

Fortunately, for the problem discussed above, there exists a solution based on the nature of interferometry. In our study, moiré interferometry is used to measure the in-plane displacement of electronic packages under vibration or temperature cycling. For a continuous solid medium, the in-plane displacement should be continuous. As a reflection of the deformation, the fringe spatial frequency should

change only gradually. Therefore we can utilize the continuity of the instantaneous spatial fringe frequency to reliably ignore the noise-induced strong CWT magnitude and accurately select the ridge in noisy interferograms. Specifically, a cost function is introduced as an effective tool for a new ridge detection algorithm:

$$\cos t[\phi(b),\, b] = -C_0 \int_b |S[\phi(b),\, b]|^2 \mathrm{d}b$$
$$+ C_1 \int_b \left| \frac{\partial \phi(b)}{\partial b} \right|^2 \mathrm{d}b, \qquad (7)$$

where $\phi(b)$ corresponds to the scaling parameter a as a parameter curve of the shifting parameter b; and C_0, C_1 are two positive constants that are the weights of the transform magnitude and the gradient of the parameter curve $\phi(b)$. Obviously, according to Eq. (7) there would be a small cost for signals with large magnitude and a smooth parameter curve. Therefore our task of searching for the ridge of the CWT magnitude reduces to minimizing the cost function for all possible parameter curves in the two-dimensional CWT magnitude map. For simplicity, we set the coefficient C_0 to 1. The selection of C_1 depends on the values of the CWT coefficients. Generally, the direct maximum ridge detection is applied to the obtained CWT coefficients. If the resulting ridge fringe pattern deviates from the original fringe pattern, the proposed cost function is applied. Subsequently, any pixel with large ridge detection error is analyzed to determine the coefficient C_1. C_1 is thus selected to equalize the weight of the scaling index discontinuity and the weight of the difference of the CWT coefficients at the noise-induced local maxima and the true local maxima for this pixel.

The most straightforward way to implement the cost function is to exhaustively search all possible curves. For an image with 30×500 pixels, as shown in Fig. 2, there are a total of 30^{500} possible curve combinations. Although we can constrain the search range to the pixels that have relatively large magnitude and ignore the small-magnitude pixels, the exhaustive searching is still computationally expensive. Inspired by the Viterbi decoding algorithm[12] that is widely used in telecommunications, in Subsections 3.B.1–3.B.3 we propose the dynamic optimization algorithm for cost function ridge detection.

1. Dynamic Optimization
For simplicity of explanation, we use $C_0 = C_1 = 1$ here. For this case, the cost function along the ridge can be written as

$$\text{cost} = \sum_{k=2}^{L} \{-|S[\phi(k),\, k]|^2 + |\phi(k) - \phi(k-1)|^2\}, \quad (8)$$

where L is the width of the interferogram; k is the discrete shifting parameter; and $\phi(k)$ is the discrete

version of $\phi(b)$ in Eq. (7) taking on values from $a(1)$ to $a(N)$, where N is the number of the scaling parameter array. Assuming that the optimal path goes through point (m, j), i.e., $\phi(j) = m$, then the cost can be revised as

$$
\begin{aligned}
\text{cost} = &\left(\sum_{k=2}^{j-1} \{-|S[\phi(k), k]|^2 + |\phi(k) - \phi(k-1)|^2\} \right) \\
&+ [-|S(m, j)|^2 + |m - \phi(j-1)|^2] \\
&+ \left(\sum_{k=j+1}^{L} \{-|S[\phi(k), k]|^2 + |\phi(k) \right. \\
&\left. - \phi(k-1)|^2\} \right).
\end{aligned} \tag{9}
$$

From Eq. (9) we can see that, at a given point (m, j), the optimization of the complete path can be divided into two parts—the optimization for path $\phi(1) \sim \phi(j)$ and the optimization for path $\phi(j) \sim \phi(L)$. This is actually an application of the principle of optimality defined by Bellman.[11] In another words, we can always decide the optimal path step by step by scanning from $k = 1$ to $k = L$.

Furthermore, if we have the knowledge of the beginning part of the optimal path ending at (m, j) and denote cost(m, j) as

$$
\begin{aligned}
\text{cost}(m, j) = \min_{\phi} &\left[\left(\sum_{k=2}^{j-1} \{-|S[\phi(k), k]|^2 + |\phi(k) \right. \right. \\
&\left. - \phi(k-1)|^2\} \right) + [-|S(m, j)|^2 \\
&\left. + |m - \phi(j-1)|^2] \right],
\end{aligned} \tag{10}
$$

we can thus express the cost of $(n, j + 1)$ as

$$
\begin{aligned}
\text{cost}(n, j+1) = \min_{m} &\{\text{cost}(m, j) - |S[\phi(n), j+1]|^2 \\
&+ |n - m|^2\}.
\end{aligned} \tag{11}
$$

As explained above, the minimization of the complete path is the minimization of the subpath ending at each node in each step. To determine the optimal subpath ending at $(n, j + 1)$, we need to compare the paths coming from N nodes of the previous step, the jth step. The subpath providing the minimum cost$(m, j) - |S[\phi(n), j + 1]|^2 + |n - m|^2$ is the survivor path with respect to node $(n, j + 1)$, and $-|S[\phi(n), j + 1]|^2 + |n - m|^2$ can be defined as the path cost.

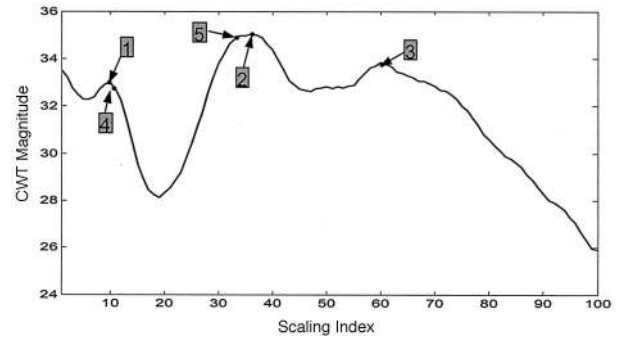Moreover, if the second derivative of the parameter function $\phi(k)$ is also included in the cost function to

Fig. 3. Typical magnitude curve ($b = 393$, fifth image row). Point 2 is the global maximum of this curve; points 1, 2, and 3 are three local maximums of the curve.

find the optimal path, we can simply modify the above algorithm to implement the optimization with

$$
\begin{aligned}
\text{cost} = \sum_{k=3}^{L} &\{-|S[\phi(k-1), k-1]|^2 + |\phi(k \\
&- 1) - \phi(k-2)|^2 + |2\phi(k-1) \\
&- \phi(k) - \phi(k-2)|^2\},
\end{aligned} \tag{12}
$$

$$
\begin{aligned}
\text{cost}(p, j+2) = \min_{m,n} &\left[\text{cost}(m, j) - |S(n, j+1)|^2 \right. \\
&+ |n - m|^2 + |2m - n - \phi(j-1)|^2 \\
&- |S(p, j+2)|^2 + |p - n|^2 \\
&\left. + |2n - m - p|^2 \right].
\end{aligned} \tag{13}
$$

Thus there are $N \times N$ subpaths that need to be considered to determine the optimal subpath ending at $(p, j + 2)$. Each subpath goes through nodes (m, j) and $(n, j + 1)$ before it reaches $(p, j + 2)$, i.e., we need to keep track of two steps back to decide cost$(p, j + 2)$.

2. Selecting Candidate Ridge Points

As we suggested above, the computational cost can be further reduced when we preselect the candidate ridge points to perform the optimization. Considering the competition between the strong CWT magnitude of the signal frequency and that due to image noise, we select all local maximums instead of the global maximum point alone to be the candidates of the ridge point of this column. This scales down the computation to a fraction $(N_j/N)^2$, where N_j is the number of local maximum points of the jth step. The local maximums feature the change from positive gradient to negative gradient. Figure 2 shows a typical two-dimensional moiré interferogram, and the CWT magnitude of pixel (5, 393) is plotted in Fig. 3. The curve in Fig. 3 is a typical curve of the CWT magnitude $|S(a, b)|$ versus scaling parameter array a; the local maximum points, points 1–3 are thus chosen as the candidate points of the ridge. In addition, the candidate list can be expanded to include the neighbors of these local maximum points to compensate for the slight error-induced difference between the neighboring columns. We can determine the neigh-
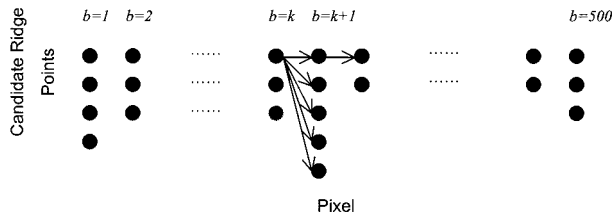
Fig. 4. Diagram of the cost function ridge detection algorithm for a typical image row. The black dots of each column represent the candidate ridge points for the kth pixel of that image row. The arrows are the paths from one candidate of the kth pixel to one candidate of the $(k + 1)$th pixel.

bors by simply setting up a threshold. For example, on the left side of the local maximum, the neighbors must have a positive gradient and a magnitude greater than 0.95 of the local maximum; on the right side of the local maximum, the neighbors must have a negative gradient and a magnitude greater than 0.95 of the local maximum. Points 4 and 5 in Fig. 3, for example, will be selected as the neighbors and thus be included in the candidate list.

Obviously from Fig. 3, the actual magnitudes of the local maximums, especially those of points 2 and 3, are close. Because there exists only one instantaneous spatial frequency at that image pixel, two of the local peaks and the corresponding neighbors are due to the image noise. However, it is difficult to determine which one is the correct fringe frequency when we look at this curve only. The local peaks from the current pixel have to be compared with the local peaks of its neighboring pixels to decide the correct fringe frequency. With the help of the cost function ridge detection and the dynamic optimization, not only is the absolute magnitude of the CWT important but also the gradient of the parameter curve is incorporated to determine which peak best fits the slowly varying instantaneous spatial-frequency assumption. This peak will provide the most reliable estimate of the ridge point at position $b = 393$.

With all the candidate ridge points stored, we can build a matrix as shown in Fig. 4. The elements of the matrix are the candidates of a typical image row. The nodes of the kth column in Fig. 4 represent the candidate ridge points for the kth pixel of that image row. The arrows indicate the paths connecting one candidate of the kth pixel to one candidate of the $(k + 1)$th pixel. Each node contains the information of the CWT magnitude and the scaling parameter of that magnitude. The implementation of the algorithm is presented in Subsection 3.B.3.

3. Implementation of the Cost Function Optimization

Assume that an optimized path of steps $1 \sim n$ is found. In this case the optimization for the best approach of steps $1 \sim (n + 1)$ will depend only on the information of the $(n + 1)$th step and the gradient from the nth step to the $(n + 1)$th step. Specifically
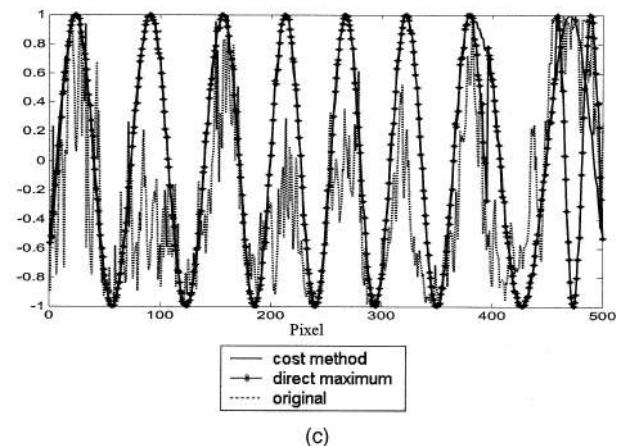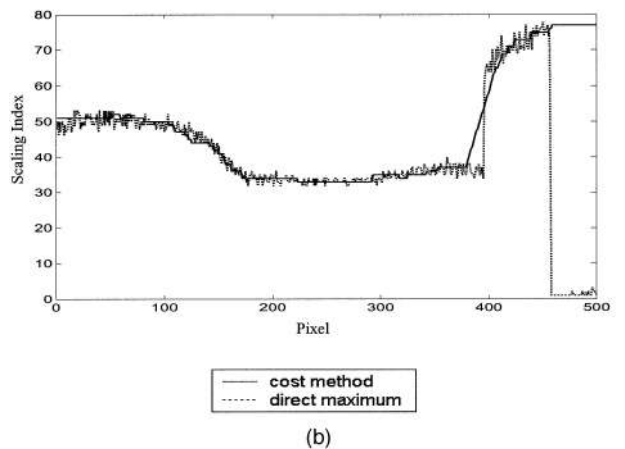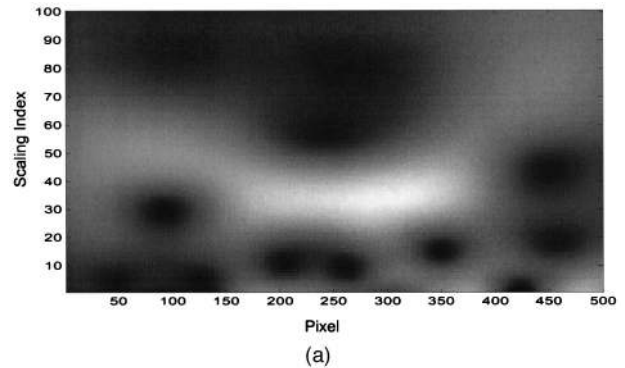

(a)

(b)

(c)

Fig. 5. Comparison of two ridge detection algorithms on an example interference signal (fifth row of Fig. 2). (a) Two-dimensional CWT magnitude map; the horizontal axis is the shifting parameter b and the vertical axis is the scaling parameter a. (b) The ridge selection by use of the direct maximum ridge detection and cost function ridge detection. (c) The reconstructed ridge of the direct maximum ridge detection and cost function ridge detection, as well as the original noisy interference signal.

in this study, for any point in the $k + 1$th column, the beginning of the path should come from a point in the kth column that contributes the least cost on this point in the $(k + 1)$th column. For example, as
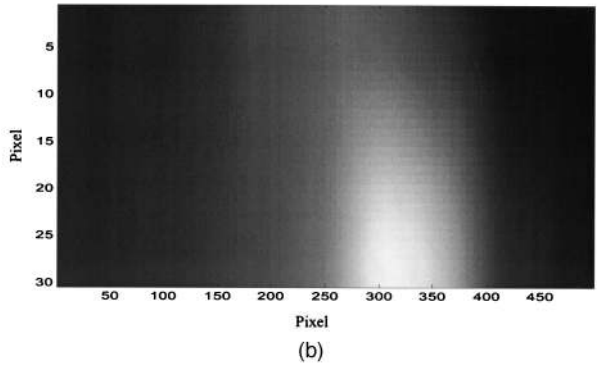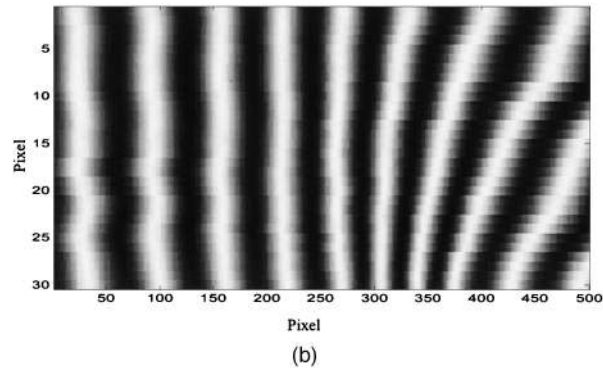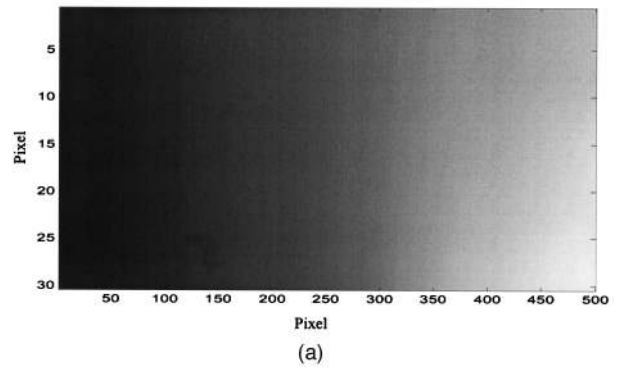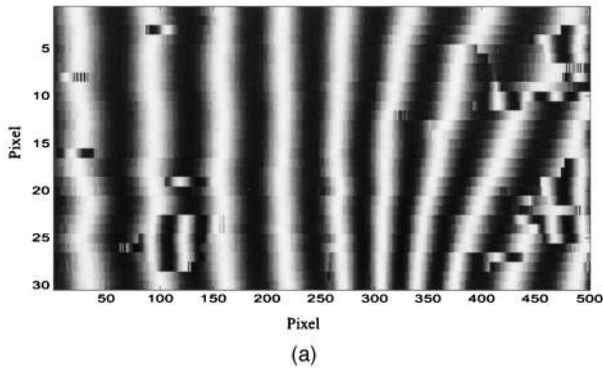
Fig. 6. Comparison of the direct maximum ridge detection and the cost ridge detection: (a) filtering with direct maximum ridge detection, (b) filtering with cost function ridge detection.

shown in Fig. 4, for the top point in column k, if the cost on that is available, we can calculate the additional costs, i.e., the path costs, from this point to all the candidate points in the next column, $b = k + 1$. Therefore the costs on all points of the $k + 1$ column are known for the paths starting from $(1, k)$. Similarly, performing the same calculation for the paths between $(2, k)$ and $(3, k)$ to $(1, k + 1)$, $(2, k + 1)$, $(3, k + 1)$, $(4, k + 1)$, $(5, k + 1)$, the three costs on each point in $(k + 1)$th column are determined. As stated in the beginning of this paragraph, the minimum of the three costs is the cost on the corresponding point in the $(k + 1)$th column. The path that contributes to this minimum cost should be selected as the path ending at that candidate point in $(k + 1)$th column. The other two paths fail to survive at this stage. Applying this procedure from the first column, $b = 1$, to the last one, $b = 500$, the final costs of three survived paths from the $b = 1$ to $b = 500$ will be available. Again, the path that contributes the minimum final cost is the one that should be chosen as the ridge of the two-dimensional CWT magnitude. Therefore the candidate points coming from localized noise will be eliminated from the final ridge selection because they will bring a sudden jump of the instantaneous spatial frequency and therefore add a large value of gradient magnitude in the cost function as described in Eq. (7).
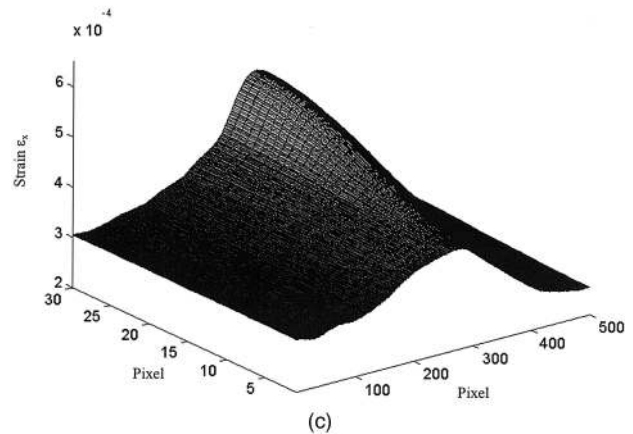

Fig. 7. (a) Phase reconstruction, (b) first derivative of the phase map in the horizontal direction, (c) the normal x strain of the image in Fig. 2.

4. Experiment

The experimental results of the proposed ridge detection algorithm applied to Fig. 2 are described in this section. As can be seen in Fig. 5(a), the two-dimensional CWT magnitude map of the fifth row of Fig. 2 has more than one local brightest point for most of the columns. One can imagine that, for such a CWT map, the direct maximum ridge detection will expose its shortcoming as a result of the existence of the noise. However, we should expect a reliable result from the cost function ridge detection algorithm that was introduced in the Subsection 3.B.2.

Clearly, comparing the ridge selection results of both algorithms as shown in Fig. 5(b), for the curves of the selected scaling parameter index we can observe a jump of the ridge index around $b = 400$ and $b = 460$ from the direct maximum detection algorithm; the corresponding index curve resulting from the cost function algorithm goes up smoothly. As a result, the real part of the selected ridge, shown in Fig. 5(c), from the direct maximum algorithm deviates from the envelope of the actual interferogram signal between these two positions. However, the ridge from the cost function ridge detection algorithm follows the original noisy signal closely, as shown in Fig. 5(c). The cost function ridge detection algorithm demonstrates its advantage in this example. The ridge reconstruction results of both algorithms on the entire interferogram image are shown in Figs. 6(a) and 6(b).

As was mentioned above, the ridge has the same phase as the interference signal. Once the smooth version of the ridge reconstruction is obtained with the cost function ridge detection algorithm, the two-dimensional phase distribution as shown in Fig. 7(a) is readily obtained. The phase map is smooth, reflecting the highest displacement at the right bottom corner and lowest displacement on the left, as was observed from the interferogram in Fig. 2. The smooth and accurate reconstruction of the phase map, i.e., the smoothness of the instantaneous spatial frequency, enables us to use the derivative of the phase to calculate the in-plane strains directly, as described in Eqs. (2). The first derivative of the phase map along the x direction is shown in Fig. 7(b), and the normal strain in the x direction ε_x is shown in Fig. 7(c). By applying similar calculations, we can calculate the normal strain in the y direction and the shear strain, thus completing the in-plane strain analysis.

5. Conclusion

A dynamic optimizing cost function ridge detection algorithm has been demonstrated to be effective to extract the phase information from an interferogram. Pixel-by-pixel in-plane strains were successfully extracted. This has a significant effect on the understanding of complicated mechanical problems. For a CCD with an 8-bit digitized output, i.e., a 256 gray-level output, the accuracy of the phase is $2\pi/256 = 0.02$ rad. The accuracy can be improved with more degrees of digitization or use of an analog output. For typical CCD pixel sizes of 9×9 μm^2 and an imaging system magnification of 2, we can obtain the average in-plane strain of an area as small as 20.25 μm^2 on the specimen. The information of a smaller area is available as long as we have smaller pixel sizes and larger magnification to ensure that the area of interest on the object is resolvable and occupies more than one pixel on the CCD sensor.

In this paper only the CWT magnitude and the gradient of the parameter curve are considered. One can also include the second derivative in the cost function[10] to further smooth the ridge selection. The implementation is described in Subsection 3.B.2. However, the inclusion of only the first derivative is sufficient for most actual interferograms. Therefore the second derivative can be neglected, which allows the computational expense to be minimized.

References
1. D. Post, B. Han, and P. Ifju, *High Sensitivity Moiré: Experimental Analysis for Mechanics and Materials* (Springer-Verlag, New York, 1994).
2. B. Han and P. Kunthong, "Micro-mechanical deformation analysis of surface laminar circuit in organic flip-chip package: an experimental study," J. Electron. Packaging **122,** 294–300 (2000).
3. Y. Zhao, C. Basaran, A. N. Cartwright, and T. Dishongh, "Thermomechanical behavior of micron scale solder joints under dynamic loads," Mech. Mater. **32,** 161–173 (2000).
4. H. Liu, A. N. Cartwright, C. Basaran, and W. Casey, "Moiré interferometry for microelectronic packaging interface fatigue reliability," in *International Conference on Compound Semiconductor Manufacturing Technology Digest* (GaAs MAN-TECH, St. Louis, Mo., 2002), pp. 165–168.
5. W. W. Macy, Jr., "Two-dimensional fringe-pattern analysis," Appl. Opt. **22,** 3898–3901 (1983).
6. K. Creath, "Phase-measurement interferometry techniques," in *Progress in Optics*, E. Wolf, ed. (Elsevier Science B. V., Amsterdam, The Netherlands, 1998), Vol. 26, Chap. 25, pp. 349–393.
7. M. Cherbuliez, P. Jacquot, and X. de Lega, "Wavelet processing of interferometric signals and fringe patterns," in *Wavelet Applications in Signal and Image Processing VII*, M. A. Unser, A. Aldroubi, and A. F. Laine, eds., Proc. SPIE **3813,** 692–702 (1999).
8. P. Tomassini, A. Giulietti, L. A. Gizzi, M. Galimberti, D. Giulietti, M. Borghesi, and O. Willi, "Analyzing laser plasma interferograms with a continuous wavelet transform ridge extraction technique: the method," Appl. Opt. **40,** 6561–6568 (2001).
9. H. Liu, A. N. Cartwright, and C. Basaran, "Sensitivity improvement in phase shifted moiré interferometry using 1D continuous wavelet transform image processing," Opt. Eng. **42,** 2646–2652 (2003).
10. R. A. Carmona, W. L. Hwang, and B. Torrésani, "Characterization of signals by the ridges of their wavelet transforms," IEEE Trans. Signal Process. **45,** 2586–2590 (1997).
11. R. Bellman, *Dynamic Programming* (Princeton University, Princeton, N.J., 1957).
12. A. J. Viterbi, "Error bounds for convolutional codes and a asymptotically optimum decoding algorithm," IEEE Trans. Inf. Theory **13,** 260–269 (1967).