



<b>Title</b>	<b>Discriminative fabric defect detection using adaptive wavelets</b>
<b>Author(s)</b>	<b>Yang, XZ; Pang, GKH; Yung, NHC</b>
<b>Citation</b>	<b>Optical Engineering, 2002, v. 41 n. 12, p. 3116-3126</b>
<b>Issued Date</b>	<b>2002</b>
<b>URL</b>	<b><a href="http://hdl.handle.net/10722/42921">http://hdl.handle.net/10722/42921</a></b>
<b>Rights</b>	<b>Copyright 2002 Society of Photo-Optical Instrumentation Engineers. This paper was published in Optical Engineering, 2002, v. 41 n. 12, p. 3116-3125 and is made available as an electronic reprint with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.</b>

# Discriminative fabric defect detection using adaptive wavelets

Xue Zhi Yang

Grantham K. H. Pang

Nelson H. C. Yung

The University of Hong Kong

Department of Electrical and Electronic  
Engineering

Pokfulam Road

Hong Kong

E-mail: xzyang@eee.hku.hk

**Abstract.** We propose a new method for fabric defect detection by incorporating the design of an adaptive wavelet-based feature extractor with the design of an Euclidean distance-based detector. The proposed method characterizes the fabric image with multiscale wavelet features by using undecimated discrete wavelet transforms. Each nonoverlapping window of the fabric image is then detected as defect or nondefect with an Euclidean distance-based detector. Instead of using the standard wavelet bases, an adaptive wavelet basis is designed for the detection of fabric defects. Minimization of the detection error is achieved by incorporating the design of the adaptive wavelet with the design of the detector parameters using a discriminative feature extraction (DFE) training method. The proposed method has been evaluated on 480 defect samples from five types of defects, and 480 nondefect samples, where a 97.5% detection rate and 0.63% false alarm rate were achieved. The evaluations were also carried out on unknown types of defects, where a 93.3% detection rate and 3.97% false alarm rate were achieved in the detection of 180 defect samples and 780 nondefect samples. © 2002 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.1517290]

Subject terms: defect detection; undecimated discrete wavelet transform; adaptive wavelets; discriminative feature extraction.

Paper 020021 received Jan. 22, 2002; revised manuscript received May 6, 2002; accepted for publication May 6, 2002.

## 1 Introduction

Automatic fabric defect detection is becoming an attractive alternative to human visual inspection in the modern textile industry. Based on advances in image processing and pattern recognition, automatic fabric defect detection can potentially provide an objective and reliable evaluation on fabric production quality. In general, a fabric defect detection system consists of a feature extractor and a detector. Much research has been focused on the design of the feature extractor, which aims at yielding features with powerful discrimination between the defect region and the defect-free fabric texture. The design of the feature extractor can be divided into two categories. One is based on statistical texture analysis,<sup>1-4</sup> which discriminates the defect region in terms of the statistical textural properties of the fabric image. The second category is the transform-based approaches, where feature extraction is based on Fourier transforms<sup>5,6</sup> Gabor transforms<sup>7,8</sup> or wavelet transforms.<sup>9-15</sup> Since Fourier bases do not have local support in the spatial domain, the features extracted from Fourier transforms are not as effective in detecting small local defects. As fabric images have regular periodic texture patterns produced during manufacturing, fabric defects either appear to be singularities in the homogeneous background, or textures whose primitives are different from the background texture in scale and orientation. Based on this viewpoint, wavelet transforms and the similar Gabor transforms, which can provide localized spatial-frequency analysis of the fabric image at multiscale and multiorientation, are

more capable for the detection of fabric defects than the traditional methods that rely on the statistical texture analysis in a single scale.

In wavelet-based defect detection, the fundamental problem remaining to be solved is the selection of the wavelet basis. Normally, the wavelet basis is heuristically selected from the standard wavelet bases, which may not yield features with optimal discrimination for the detection of all kinds of fabric defects. A solution to this problem is to design the wavelet basis adapting to the detection of the fabric defects. In Refs. 13 and 14, orthogonal adaptive wavelets are derived directly from the texture data of defect-free fabric images, such that the wavelet filter gives a close-to-zero response to this texture, while disturbances in the texture due to noise and defects are expected to produce a distinct nonzero output. The wavelets derived in such a way only achieve the optimal representation of the defect-free fabric texture, but not the optimal discrimination between the defect region and the defect-free fabric texture. Moreover, the orthogonal wavelet transform, due to its shift variant property, is not suitable for fabric defect detection.

To obtain shift-invariant representation and more flexibility in the wavelet design, Yang, Pang, and Yung<sup>5</sup> derived adaptive wavelets on the basis of undecimated wavelet transforms. For each type of fabric defect, an adaptive wavelet was designed to enhance the defect at one selected channel of the wavelet transform. Compared to Daubechies wavelets, the designed adaptive wavelets largely enhance the defect region, where good detection of these fabric de-

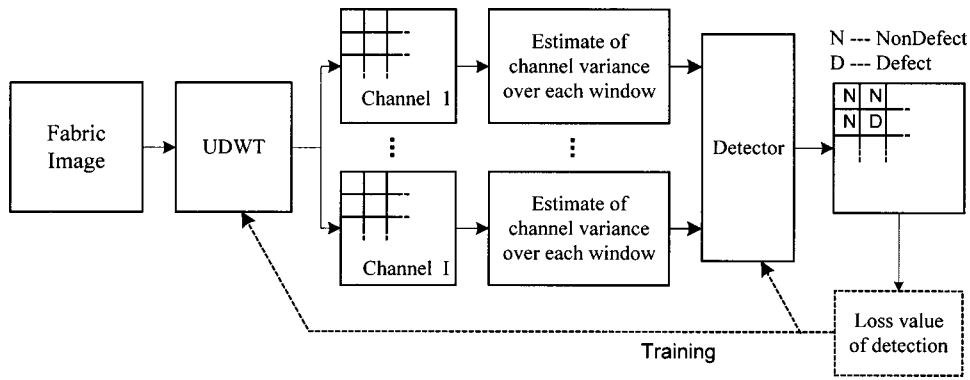


Fig. 1 The proposed fabric defect detection method.

fects was achieved by using a threshold detector. However, multiple adaptive wavelet bases are used in feature extraction, which is computationally expensive. In this work, a single adaptive wavelet has been designed for the fast detection of multiple types of defects. Moreover, by using a discriminative feature extraction (DFE) training method,<sup>16</sup> the design of the adaptive wavelet is incorporated with the design of the detector parameters for the objective of minimum error rate in the detection. Traditionally, the design of the feature extractor and the detector in a defect detection system are loosely linked, which may not yield appropriate interactions between the feature extractor and the detector. By using the DFE training method, the inconsistency between the feature extractor and the detector is alleviated, which leads to better performance on the defect detection. The proposed defect detection method has been evaluated on 480 defect samples from five types of defects, and 480 nondefect samples, where a 97.5% detection rate and a 0.63% false alarm rate were achieved. The evaluations were also carried out on these types of defects, which were unknown to the designed feature extractor and detector. A 93.3% detection rate and 3.97% false alarm rate were achieved in the detection of 180 defect samples from three types of defects and 780 nondefect samples. Compared to the standard wavelet bases, the adaptive wavelet enables the defect detection to achieve similar performance with fewer scales of wavelet features, which leads to substantial computational savings for the detection.

This work is organized as follows. In the next section, the proposed defect detection method is presented. The feature extraction module and detection module in the defect detection are described first. Then we describe how to incorporate the wavelet design with the design of the detector parameters by using the DFE method for achieving the objective of minimum error rate in the defect detection. The evaluation results of the proposed method are reported in Sec. 3. Section 4 concludes this work.

## 2 Detection-Centric Adaptive Wavelet Design

### 2.1 Overview

Figure 1 illustrates the block diagram of the proposed fabric defect detection method. The defect detection essentially consists of two modules: the feature extraction module followed by the detection module. In the feature extraction module, an undecimated discrete wavelet trans-

form (UDWT) has been used to yield multiscale representation of the fabric image. Feature vectors consisting of channel variances at the output of the wavelet transform were extracted to characterize each nonoverlapping window of the fabric image. In the detection module, an Euclidean distance-based detector was used. Minimization of the detection error was achieved by using the DFE training method, which is illustrated in Fig. 1 using dashed lines.<sup>17</sup> In the DFE framework, defect detection on a set of training images is evaluated by using a loss value that is consistent with the detection error rate. The loss value is then minimized by the design of the adaptive wavelet in the feature extraction module and the design of the detector parameters in the detection module.

### 2.2 Feature Extraction Based on Wavelet Transform

Figure 2<sup>18</sup> illustrates the 1-D undecimated discrete wavelet transform implemented by an undecimated octave band filter bank, where  $H(z)$  and  $G(z)$  denote the  $z$  transform of the low-pass filter  $h[n]$  and high-pass filter  $g[n]$ , respectively. Compared to the critically sampled wavelet transform, the undecimated wavelet transform achieves translation invariance, which is desirable for fabric defect detection. The cascade of the filtering operations in the undecimated wavelet transform shown in Fig. 2 can be described in the following equivalent filter sequences<sup>19</sup>

$$H_r(z) = H(z^{2^{r-1}})H_{r-1}(z) = \prod_{k=0}^{r-1} H(z^{2^k}), \quad \text{for } r = 1, \dots, J, \tag{1}$$

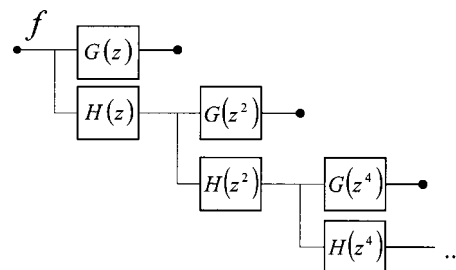
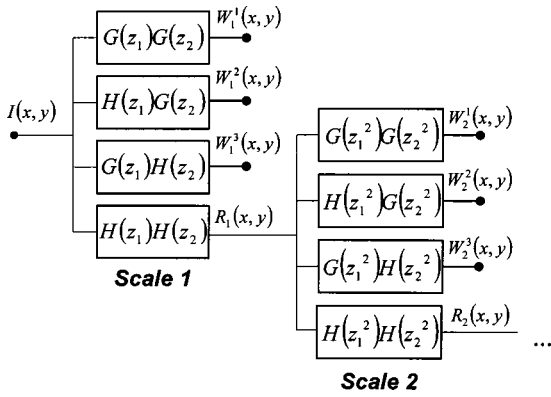


Fig. 2 Filter bank implementation of a 1-D undecimated wavelet transform.



**Fig. 3** Filter bank implementation of a 2-D undecimated wavelet transform.

$$G_r(z) = G(z^{2^{r-1}})H_{r-1}(z) = G(z^{2^{r-1}}) \prod_{k=0}^{r-2} H(z^{2^k}),$$

for  $r = 2, \dots, I,$  (2)

where  $H_0(z) = 1$  and  $G_1(z) = G(z)$ .  $r$  is the scale index of the wavelet transform and  $I$  denotes the depth of the wavelet transform.

The 2-D undecimated wavelet transform is obtained by using the tensor products of the 1-D version.<sup>20</sup> The filter bank implementation of the 2-D undecimated wavelet transform is illustrated in Fig. 3. In this figure,  $I(x,y)$  denotes an image and  $(x,y)$  is the spatial indices.  $\{W_r^1(x,y), W_r^2(x,y), W_r^3(x,y)\}$  denote the wavelet coefficients at scale  $r$ , with diagonal, vertical, and horizontal orientation, respectively, and  $R_r(x,y)$  represents the residue signal at scale  $r$ .

Fabric defects have various sizes and most of the defects only occupy a small portion of the fabric image. To locate the defect, the fabric image is divided into nonoverlapping windows with size  $N_w \times N_w$ , and the defect detection is performed on each image window. To characterize each image window, the channel variances<sup>19</sup> at the outputs of the wavelet transform are used. As it is shown in Ref. 19, channel variances are able to provide efficient discriminations among different kinds of textures. Therefore, these features are employed here for the discrimination between defect-free fabric textures and defective textures. Corresponding to a window in the fabric image, the channel variances are estimated as the mean energy of the wavelet coefficients in the window<sup>19</sup>

$$w_r^d = \text{Mean}_{(x,y) \in \text{window}} [W_r^d(x,y)]^2, \text{ for } d = 1, 2, 3. \tag{3}$$

The channel variances at each channel of the wavelet transform form the feature vector to characterize the image window

$$\mathbf{F} = [w_1^1, w_1^2, w_1^3, \dots, w_I^1, w_I^2, w_I^3], \tag{4}$$

where  $I$  is the decomposition depth of the wavelet transform, and the total number of features in the feature vector  $\mathbf{F}$  is  $3I$ .

### 2.3 Detection Algorithm

Based on the Euclidean distance similarity measure, the discriminant function  $g_j(\mathbf{F}; \mathbf{\Lambda})$  for class  $C_j$  is defined as follows.

$$g_j(\mathbf{F}; \mathbf{\Lambda}) = \|\mathbf{F} - \mathbf{m}_j\|^2, \tag{5}$$

where  $j = 1, 2$  denote defect and nondefect, respectively, and  $\mathbf{\Lambda} = \{\mathbf{m}_j\}_{j=1,2}$  are the reference vectors representing defect and nondefect, respectively. The reference vectors  $\mathbf{\Lambda}$  are estimated by using the DFE training method described in the next subsection. The decision rule of the detector is

$$\mathbf{F} \in C_i \text{ if } i = \underset{j}{\text{argmin}} g_j(\mathbf{F}; \mathbf{\Lambda}), \tag{6}$$

which assigns the feature vector  $\mathbf{F}$  into the class whose reference vector has minimum Euclidean distance to  $\mathbf{F}$ .

### 2.4 DFE Using Adaptive Wavelets for Defect Detection

In wavelet-based defect detection, feature extraction using different wavelet bases yields wavelet features with different discriminations between defect and nondefect, which indicates that the selection of the wavelet basis is closely related to the performance of the defect detection. Standard wavelet bases, e.g., Daubechies wavelets, Haar wavelets, etc., are not necessarily the best candidates. A better alternative to the standard wavelets is the custom-designed wavelet, which is adapted to the detection of the fabric defects. The performance of defect detection is determined by not only the design of the adaptive wavelet-based feature extractor, but also the design of the detector. To achieve appropriate interactions between the feature extractor and the detector, the DFE training method is used to perform the overall design of both the feature extractor and the detector. The DFE training method stems from the minimum classification error (MCE) training method, which was developed by Juang and Katagiri<sup>21</sup> for the classifier design. Biem et al.<sup>16,22</sup> further extended the MCE training method from the back-end classifier to the front-end feature extractor for the design of the overall pattern recognizer. In our approach, the DFE training method incorporates the design of the adaptive wavelet with the design of the reference vectors of the detector, so that the error rate in the defect detection is minimized. In this subsection, before the implementation of the DFE training for the defect detection, we first describe how to parameterize the wavelet filters in a form suitable to be designed with an optimization approach.

#### 2.4.1 Parametrization of wavelet filters

Based on the undecimated octave band filter bank, we consider the design of the wavelet filters under the constraint that  $H(z)$  and  $G(z)$  are power complementary,<sup>18</sup> i.e.,  $H(z)$  and  $G(z)$  satisfy the following condition



$$H(z)H(z^{-1}) + G(z)G(z^{-1}) = 1, \quad (7)$$

we also impose the following constraints on  $H(z)$  and  $G(z)$ , that  $H(-1) = 0$  and  $G(1) = 0$ , where  $H(-1) = 0$  is imposed for meeting a regularity requirement. By choosing  $G(1) = 0$ , we are able to construct a wavelet with one vanishing moment, which would yield multiscale edge representation of the fabric image.

To design the wavelet filters with an optimization strategy, we parameterize the wavelet filters  $H(z)$  and  $G(z)$ , which satisfy the prior constraints by using lattice structure factorization.<sup>23</sup> Lattice structure factorization performs a cascade-form factorization for the power complementary pairs  $[H(z)G(z)]$

$$\begin{bmatrix} H(z) \\ G(z) \end{bmatrix} = R_m \Lambda(z) R_{m-1} \Lambda(z) \dots R_1 \Lambda(z) \begin{bmatrix} \cos \theta_0 \\ \sin \theta_0 \end{bmatrix}, \quad (8)$$

where

$$R_k = \begin{bmatrix} \cos \theta_k & \sin \theta_k \\ -\sin \theta_k & \cos \theta_k \end{bmatrix} \text{ and } \Lambda(z) = \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}.$$

Here  $H(z)$  and  $G(z)$  have equal filter length  $m + 1$ . For satisfying the constraints  $H(-1) = 0$  and  $G(1) = 0$  in this factorization structure, the lattice coefficients  $\{\theta_k\}_{0 \leq k \leq m}$  should satisfy the following relations

$$\theta_m = \frac{\pi}{4} + \theta_0 - (\theta_2 + \theta_4 + \dots + \theta_{m-2}),$$

$$\theta_{m-1} = -\frac{\pi}{4} - (\theta_1 + \theta_3 + \dots + \theta_{m-3}),$$

when  $m$  is even, (9)

or

$$\theta_m = \frac{\pi}{4} - (\theta_1 + \theta_3 + \dots + \theta_{m-2}),$$

$$\theta_{m-1} = -\frac{\pi}{4} + \theta_0 - (\theta_2 + \theta_4 + \dots + \theta_{m-3}),$$

when  $m$  is odd. (10)

That is, in designing the wavelet filters with length  $m + 1$ , we can freely choose from a set of  $m - 1$  lattice coefficients  $\Theta = \{\theta_k\}_{0 \leq k \leq m-2}$ . Since power complementary property is structurally satisfied, the design of the wavelet filters turns out to be an unconstrained optimization of the lattice coefficients. This advantage makes the lattice structure suitable for the wavelet filter design with an optimization approach.

#### 2.4.2 Implementation of DFE training for defect detection

In the defect detection shown in Fig. 1, the adjustable parameters of the feature extractor are the lattice coefficients  $\Theta$ , which determine the wavelet basis. The adjustable pa-

rameters of the Euclidean distance-based detector are the reference vectors  $\Lambda$ . The total set of adjustable parameters in the defect detection is  $\mathbf{T} = \{\Theta, \Lambda\}$ . DFE training on the parameter set  $\mathbf{T} = \{\Theta, \Lambda\}$  is implemented as follows.<sup>16</sup>

Given a set of training samples

$$\Gamma = \{\mathbf{F}_n\}_{n=1}^N,$$

where each sample is labeled as defect or nondefect, an incorrect detection measure  $d_n$  is defined for each training sample  $\mathbf{F}_n$  as

$$d_n = 1 - \frac{g_j(\mathbf{F}_n; \Lambda)}{g_i(\mathbf{F}_n; \Lambda)}, \quad \text{if } \mathbf{F}_n \in C_i, \quad (11)$$

where  $i \neq j$ . According to the decision rule defined in Eq. (6),  $d_n \leq 0$  indicates a correct detection, while  $d_n > 0$  indicates otherwise. By incorporating the decision rule in this incorrect detection measure,  $d_n$  enumerates how likely the sample  $\mathbf{F}_n$  is incorrectly detected.

Based on the incorrect detection measure  $d_n$ , a loss function is then used to evaluate the detection performance. The loss function is defined as the smoothed zero-one function of the incorrect detection measure

$$l_n = \frac{1}{1 + \exp(-\alpha d_n)}, \quad (12)$$

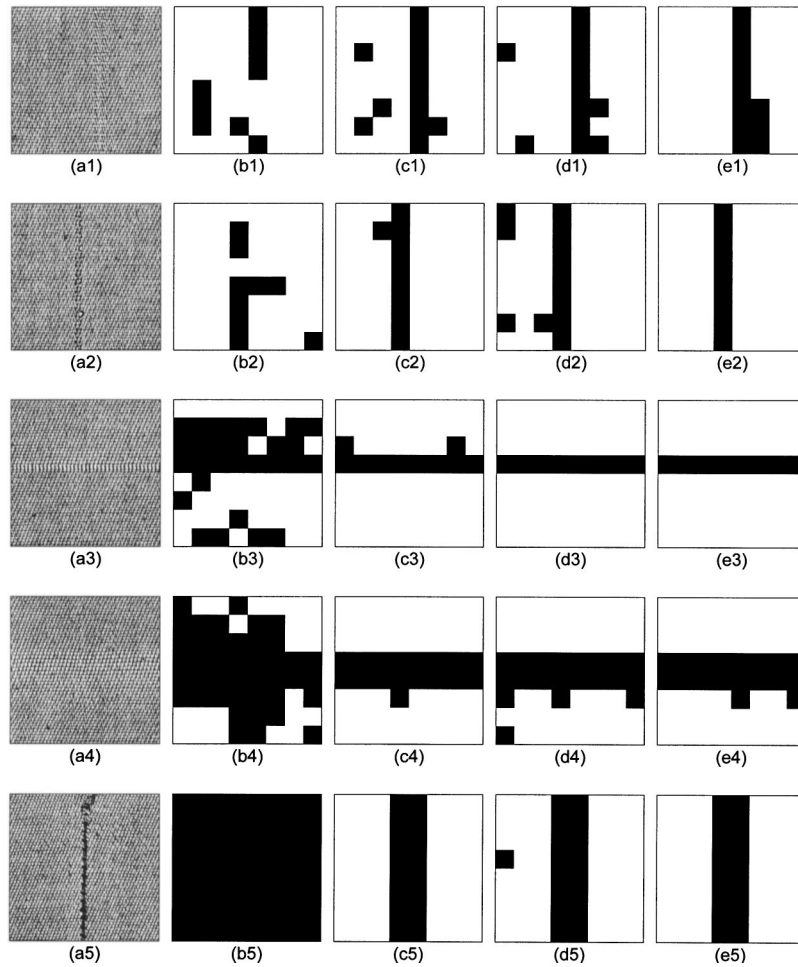
where  $\alpha (> 0)$  determines the smoothness of this loss function. Obviously, the loss value incurred by a correct detection ( $d_n \leq 0$ ) is close to zero, while the loss value incurred by an incorrect detection ( $d_n > 0$ ) is close to one.

Finally, the empirical average cost for the total set of training samples  $\Gamma$  is defined as

$$L = \frac{1}{N} \sum_{n=1}^N l_n. \quad (13)$$

By minimizing this empirical average cost with respect to the set of parameters  $\mathbf{T} = \{\Theta, \Lambda\}$ , both the adaptive wavelet-based feature extractor and the detector are designed for the minimum error rate in the defect detection. Compared to the DFE training method, the MCE training method has the same procedure in the implementation, but the empirical average cost is minimized only with respect to the detector parameters  $\Lambda$ . In that way, only the detector is trained for the objective of minimum error rate. The steepest gradient descent algorithm is normally employed by the DFE to minimize the empirical average cost. However, this optimization algorithm is quite inefficient, especially when the calculation of the gradient of the empirical average cost  $L$  with respect to the lattice coefficients  $\Theta$  is time consuming. To perform the optimization more efficiently, a quasi-Newton optimization method<sup>24</sup> is used instead. The calculation of the gradient of the empirical average cost  $L$  with respect to the parameter set  $\mathbf{T} = \{\Theta, \Lambda\}$ , as required by the quasi-Newton method, is given in Sec. 5.

The complete algorithm of DFE training for the defect detection is summarized as follows.



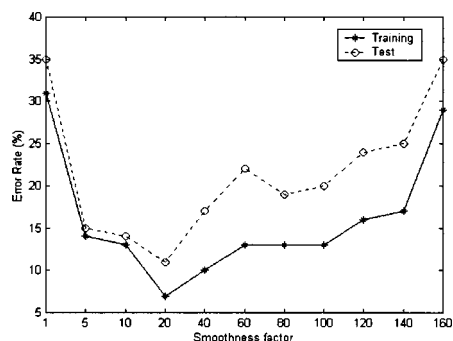
**Fig. 4** (a1)–(a5) Fabric images containing defect Brokenend, Slackend, Mispick, Thinbar, and Netmultiples. (b1)–(b5) Detection results using Haar wavelets and 2-scale wavelet features. (c1)–(c5) Detection results using Haar wavelets and 3-scale wavelet features. (d1)–(d5) Detection results using adaptive wavelets and 2-scale wavelet features. (e1)–(e5) Detection results using adaptive wavelet and 3-scale wavelet features.

1. Select a set of fabric images for training. Label the image windows as defect or nondefect.
2. Initialize the lattice coefficients  $\Theta$  and the reference vectors  $\Lambda$ . The maximum number of iterations in the quasi-Newton optimization procedure is set as a stopping criterion for the DFE training.
3. Corresponding to the lattice coefficients  $\Theta$ , the wavelet filters  $H(z)$  and  $G(z)$  are derived using Eq. (8).
4. Perform the wavelet transform on the training fabric images using the derived wavelet filters. The feature vectors of the training samples are obtained using Eq. (4).
5. Calculate the empirical average cost using Eq. (13). Calculate the gradients of the empirical average cost with respect to  $\Theta$  and  $\Lambda$ .
6. Update  $\Theta$  and  $\Lambda$  using the quasi-Newton optimization method to reduce the empirical average cost.
7. If the stopping criterion is satisfied, the DFE training is terminated. Otherwise, go to step 3 and continue the training procedure.

### 3 Evaluations

#### 3.1 Data Collection

The proposed defect detection method was evaluated on the detection of five types of representative fabric defects. These fabric defects are Brokenend, Slackend, Mispick, Thinbar, and Netmultiples [Figs. 4(a1) to 4(a5)]. The fabric images are  $256 \times 256$  pixels in size with 256 gray levels. 25 fabric images containing the five types of defects were captured for the evaluation. Feature vectors were extracted to characterize image windows  $32 \times 32$ . In each fabric image, 24 feature vectors with defect were collected, which correspond to 24 windows containing different parts of the defect region. 24 feature vectors without defects were also collected in the same fabric image, which correspond to 24 nonoverlapping and defect-free windows in the fabric image. Five fabric images containing the five types of defects, respectively, were used for training, where 120 defect samples and 120 nondefect samples were collected. The remaining 20 fabric images were used as tests, where 480 defect samples and 480 nondefect samples were collected.



**Fig. 5** The effect of the smoothness factor  $\alpha$  on the performance of the defect detection using the DFE training method.

## 3.2 Evaluation Conditions

### 3.2.1 Evaluation criteria

The overall detection performance is evaluated by the error rate in the detection, which is defined as the percentage of samples incorrectly detected. Detailed evaluations of the detection performance adopt two criteria: detection rate and false alarm rate. The detection rate is defined as the percentage of defect samples that are correctly detected as defect, and the false alarm rate is defined as the percentage of nondefect samples that are incorrectly detected as defect.

### 3.2.2 Decomposition depth of the wavelet transform

When the decomposition depth of the wavelet transform is increased, the feature vector  $\mathbf{F}$  [Eq. (4)] includes more features that are extracted from the channels at the increased scales of the wavelet transform. In our evaluation, wavelet transform with decomposition depths 1, 2, and 3 were investigated, where 1 scale features (3 features), 2 scales features (6 features), and 3 scales features (9 features) of the wavelet transform were used for the defect detection, respectively.

### 3.2.3 Length of the wavelet filters

The length of the wavelet filters  $H(z)$  and  $G(z)$  determines the number of free parameters in designing the wavelets basis. On one hand, a long filter allows more design freedoms. However, too many free parameters will reduce the generalization capacity of the detection method. On the other hand, a short filter can provide more accurate location of the defect region and lower computational costs than a long filter. Wavelet filters with different lengths (from 3 to 20) were evaluated in the defect detection. From our work, filters with lengths between 8 and 16 usually gave better detection performance than filters with other lengths. In the following evaluations, a filter length of 10 was used.

### 3.2.4 Smoothness factor of the loss function

The smoothness factor  $\alpha$  in the loss function [Eq. (12)] controls the loss value from the training samples. Consequently, the selection of  $\alpha$  affects the detection performance. Figure 5 illustrates the error rate in the defect detection, where different  $\alpha$  were used in the DFE training. Here the decomposition depth of the wavelet transform was 1. When the decomposition depth was 2 or 3, similar relations were obtained. As depicted in Fig. 5, the detection

**Table 1** The effect of the window size on the performance of the defect detection.

Window size	Error rate (%)		
	1-scale features	2-scale features	3-scale features
$16 \times 16$	14.5	10.5	7.3
$32 \times 32$	10.5	4.2	1.6
$64 \times 64$	6.9	3	0.7

performance substantially decreases if  $\alpha$  is either too large or too small. When  $\alpha$  is too large, the loss values from most of the training samples are close to zero or one. As a result, the final cost surface is not smooth and is unsuitable for a gradient algorithm. Moreover, the loss value cannot efficiently reflect how likely a sample is incorrectly detected. However,  $\alpha$  with too small of a value also results in poor performance, since the empirical average cost contains too much loss from those training samples that can be well detected. In our evaluations, the value of  $\alpha$  was chosen to be 20.

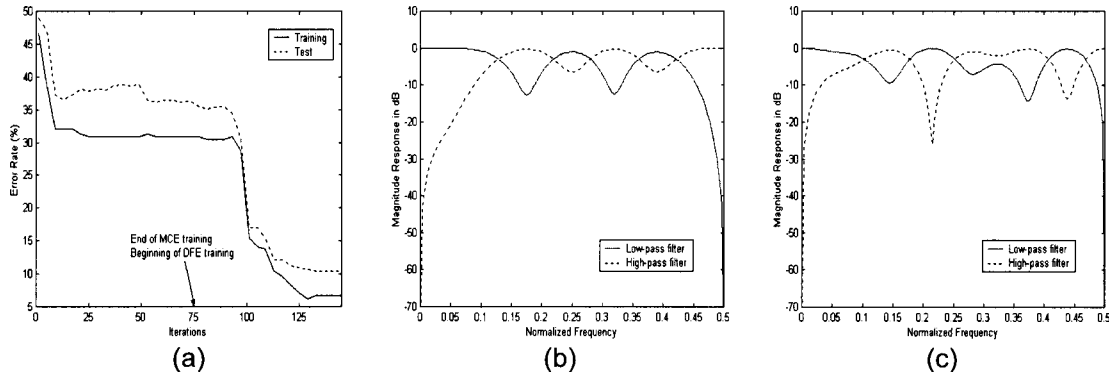
### 3.2.5 Effect of using different window size

For an accurate location of the defect, the size of the image window should be small. However, a small window cannot preserve well the texture property of the defect-free fabric texture, and could lead to poor discrimination between the window containing the defect and the window without the defect. Obviously, the suitable window size is determined by the resolution of the fabric image. Based on the fabric images we have captured, DFE training was performed on image windows of sizes  $16 \times 16$ ,  $32 \times 32$ , and  $64 \times 64$ , respectively. The corresponding error rate in the detection of test samples are summarized in Table 1.

As shown in Table 1, the largest window,  $64 \times 64$ , achieved the smallest error rate. However, the accuracy in locating the defect along with such a window size is not satisfactory. On the other hand, the smallest window with size  $16 \times 16$  yielded the largest error rate, although it corresponds to the best accuracy in locating the defects. In this work, evaluation results based on the  $32 \times 32$  window are presented.

### 3.2.6 Initialization of the DFE

Since the implementation of the DFE is based on gradient descent optimization (quasi-Newton method), the performance of the defect detection using the DFE training depends on the initialization of the parameter set  $\mathbf{T} = \{\Theta, \Lambda\}$ . However, it is difficult to predetermine a set of lattice coefficients  $\Theta$  that would yield good detection performance after DFE training. Several sets of randomly generated  $\Theta$  were used for the initialization of the lattice coefficients and the best results are presented. After the lattice coefficients were initialized, the MCE training method was used to initialize the reference vectors  $\Lambda$  of the detector. Corresponding to the initial wavelet, the detector parameters were initialized for the minimum error rate in the detection. The MCE training itself also needed reasonable



**Fig. 6** DFE training using 1-scale wavelet features: (a) learning curves, (b) frequency responses of the wavelet filters before DFE training, and (c) after DFE training.

initialization on  $\Lambda$ , which was implemented by a maximum likelihood method (using class-dependent mean vectors).

### 3.3 Evaluation Results

The training procedure for the defect detection is divided into three steps. At each step, the detection performance is evaluated.

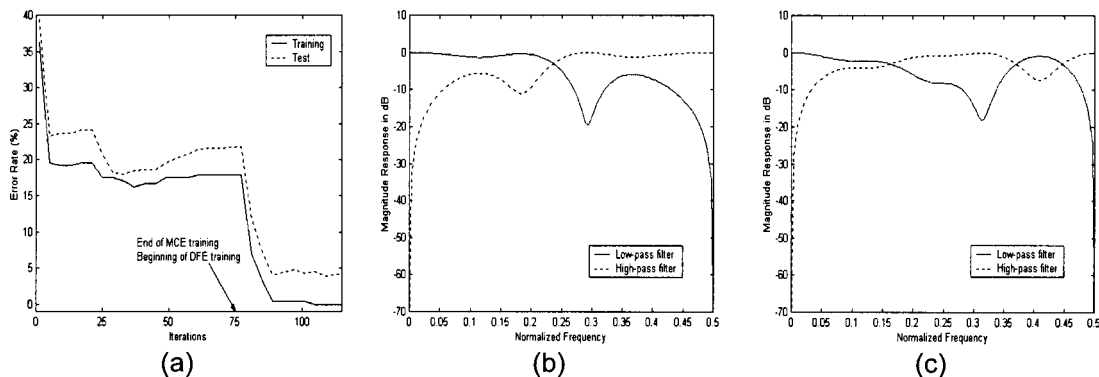
1. Initialize the lattice coefficients  $\Theta$  with random values. The reference vectors  $\Lambda$  of the detector are obtained by using a maximum likelihood method, where the reference vector for each class is calculated as the class-dependent mean vectors.
2. MCE training on the reference vectors  $\Lambda$ .
3. DFE training on  $\mathbf{T} = \{\Theta, \Lambda\}$ .

At step 1, after the initialization of the lattice coefficients, the reference vectors  $\Lambda$  are initialized as the class-dependent mean vectors for the MCE training at step 2. Subsequently, the reference vectors  $\Lambda$  are trained by using the MCE method, which initializes  $\Lambda$  for the DFE training in step 3.

When using 1-scale, 2-scale, and 3-scale wavelet features, learning curves of the DFE training procedure are illustrated in Fig. 6(a), Fig. 7(a), and Fig. 8(a), respectively. At each step of the training, the error rate in the detection of training samples and test samples are summarized in Table 2. In step 1, the poor detection performance indicates

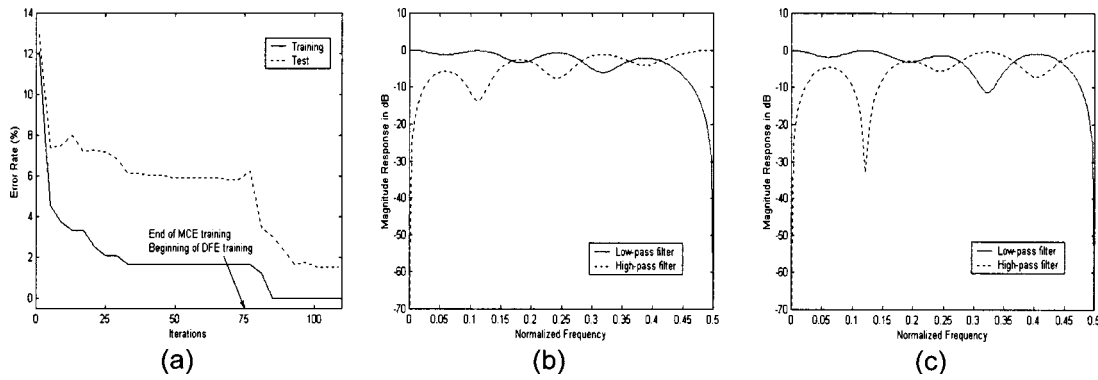
that the reference vectors estimated using the class-dependent mean vectors cause large decision bias. In step 2, the MCE method results in 12.7, 17.7, and 7.1% reduction of the error rate in the detection of test samples, where 1-scale, 2-scale, and 3-scale wavelet features are used, respectively. By using the MCE training method, the decision bias caused by the estimation on the detector parameters is reduced and better detection performance is achieved.

So far, the training only focuses on the back-end detector, and the detection performance is limited by the discrimination between defect and nondefect, which is determined by the front-end feature extractor. In step 3, it can be seen that the error rate in the detection of test samples substantially decreases after the DFE training. Compared to the MCE training, the defect detection using DFE training further achieves 25.8, 17.5, and 4.3% reduction on the error rate in the detection of test samples. Such improvements on the detection performance are primarily due to the design of the wavelet basis, which substantially increases the discriminative power between defect and nondefect. Moreover, the large improvement on the detection performance on test samples also indicates that the design of the feature extractor is an efficient approach to increase the generalization capacity of the defect detection to unknown samples. Corresponding to the use of 1-scale, 2-scale, and 3-scale wavelet features, the wavelet filters before and after the DFE training are illustrated in Figs. 6(b) and 6(c), Figs. 7(b) and 7(c), and Figs. 8(b) and 8(c), respectively.



**Fig. 7** DFE training using 2-scale wavelet features: (a) learning curves, (b) frequency responses of the wavelet filters before DFE training, and (c) after DFE training.





**Fig. 8** DFE training using 3-scale wavelet features: (a) learning curves, (b) frequency responses of the wavelet filters before DFE training, and (c) after DFE training.

Detection rate and false alarm rate are used for the detailed evaluation on the performance of the defect detection method. Corresponding to the minimum error rate (1.6%) achieved by using adaptive wavelets with 3-scale wavelet features, the detection rate and false alarm rate are 97.5 and 0.625%, respectively.

### 3.4 Comparison between Adaptive and Standard Wavelets

For comparative study, several commonly used wavelet bases have been evaluated in defect detection. These wavelet bases are Haar wavelets, Daubechies wavelets with length 12, and Battle-Lemarie wavelets of order 1.

As shown in Table 3, the detection performance of the simple Haar wavelet always outperforms the other two types of standard wavelets, and the minimum error rate in the detection of test samples is 5.7%. It is worth noting that such a detection performance is only obtained by using 3-scale wavelet features, while adaptive wavelets obtains less error rates (4.2%) by using 2-scale wavelet features.

Figures 4(b) and 4(c) illustrate the detection results on one test image of each type of defects using 2-scale and 3-scale wavelet features, respectively, where the wavelet basis is a Haar wavelet. Compared to the Haar wavelet, the detection results on the same images using an adaptive wavelet are illustrated in Figs. 4(d) and 4(e). From these results, it can be seen that the Haar wavelet yields poor detection performance when using 2-scale wavelet features, especially on the detection of defect Netmultiples, where all the nondefect samples are incorrectly detected as defect. Note here that the fabric containing the defect Netmultiples is slightly different from the fabric containing the other

four types of defects. When using 3-scale wavelet features, the discrimination between defect and nondefect is increased and better results are obtained. Compared to the Haar wavelet, the adaptive wavelet always performs better when the same scales of the wavelet features are used. Satisfactory detection results are achieved when using adaptive wavelet and 3-scale wavelet features. The results also show that adaptive wavelets using 2-scale wavelet features perform slightly better than Haar wavelets using 3-scale wavelet features. Since feature extraction from more scales of the wavelet transform requires more computational cost, the use of the adaptive wavelet enables a faster detection than the standard wavelets.

### 3.5 Additional Evaluations on Unknown Types of Defects

The detection performance of the proposed method has also been evaluated on unknown types of defects. The unknown types of defects refer to those that have not been used in the DFE training and are unknown to the DFE-trained feature extractor and detector. Three types of defects, Misyarn [Fig. 9(a1)], Broken [Fig. 9(a2)] and Stain [Fig. 9(a3)], are used for the evaluations. A total of 15 fabric images containing the three types of defects were captured. From these 15 fabric images, 180 defect samples and 780 nondefect samples were collected. Performance of the defect detection on these unknown types of test samples are summarized in Table 4. Examples of the defect detection of each type of defect are shown in Fig. 9.

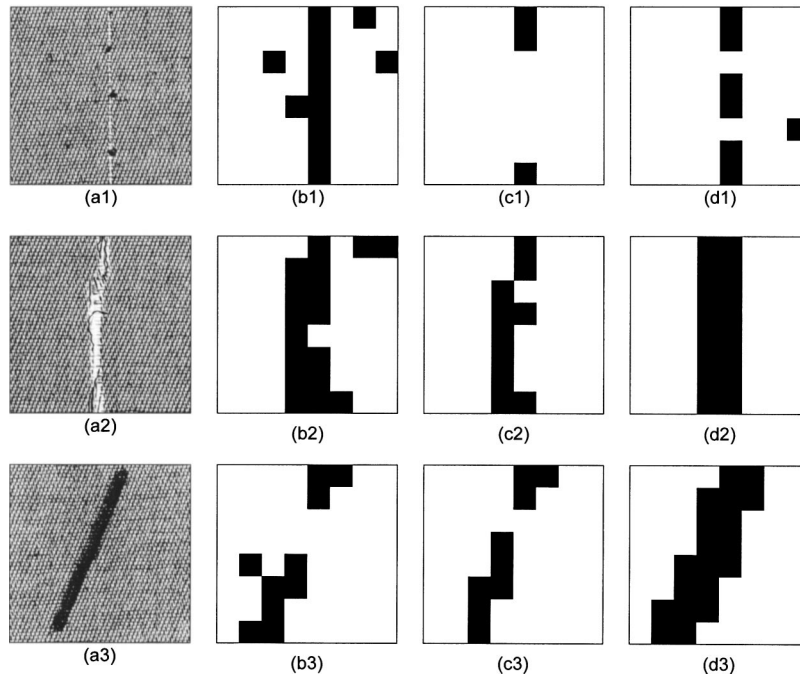
In contrast to the detection performance on the five known types of defects, the error rate in the detection of unknown types of defects increase 3, 4.6, and 2.9%, respec-

**Table 2** Error rate of the defect detection using DFE training with adaptive wavelets.

DFE training procedure	Error rate (%)					
	1-scale features		2-scale features		3-scale features	
	Train	Test	Train	Test	Train	Test
Step 1	46.7	49	36.3	39.4	12.1	13
Step 2	30.9	36.3	18	21.7	1.7	5.9
Step 3	6.7	10.5	0	4.2	0	1.6

**Table 3** Error rate of the defect detection using MCE training with standard wavelets.

Standard wavelet bases	Error rate (%)					
	1-scale features		2-scale features		3-scale features	
	Train	Test	Train	Test	Train	Test
Haar	23.8	33.4	16.3	24.5	0.5	5.7
Daubechies	26.3	33.9	19.6	26.6	0.5	8.7
Battle-Lemarie	25	34.6	21.7	27	0.9	7.8



**Fig. 9** (a1)–(a3) Fabric images containing defects Misyarn, Broken, and Stain. (b1)–(b3) Detection results using Haar wavelets and 3-scale wavelet features. (c1)–(c3) Detection results using adaptive wavelets and 2-scale wavelet features. (d1)–(d3) Detection results using adaptive wavelets and 3-scale wavelet features.

tively. On the detection of small defects like Misyarn, adaptive wavelets with 3-scale wavelet features perform similarly to adaptive wavelets with 2-scale wavelet features and Haar wavelets with 3-scale wavelet features, but much better on the detection of large defects Broken and Stain. On average, adaptive wavelets with 2-scale wavelet features perform similarly to Haar wavelets with 3-scale wavelet features. These results further demonstrate the advantages of using the adaptive wavelet. Corresponding to the minimum error rate (4.5%) achieved by using adaptive wavelets with 3-scale wavelet features, the detection rate and false alarm rate are 93.3 and 3.97%, respectively.

**4 Conclusions**

A new method that incorporates the design of an adaptive wavelet-based feature extractor with the design of an Euclidean distance-based detector is proposed for fabric defect detection. In the wavelet-based fabric defect detection, the selection of the wavelet basis is closely related to the

detection performance. As a major contribution of this work, we show how to design the wavelet basis by adapting it to the detection of fabric defects. By using the DFE training method, the design of the adaptive wavelet is incorporated with the design of the detector parameters to minimize the error rate in the defect detection. In this way, appropriate interactions between the feature extractor and the detector have been achieved. Compared to the standard wavelet bases, the adaptive wavelet enables the defect detection to achieve similar performance with fewer scales of wavelet features, which leads to substantial computational savings for the detection.

**5 Appendix A**

The gradients of the empirical average cost  $L$  with respect to the parameter set  $\mathbf{T}=\{\Theta, \Lambda\}$  are derived as follows:

$$\frac{\partial L}{\partial \mathbf{m}_j} = \frac{1}{N} \sum_{n=1}^N \frac{\partial l_n}{\partial \mathbf{m}_j} \quad \text{and} \quad \frac{\partial L}{\partial \theta_k} = \frac{1}{N} \sum_{n=1}^N \frac{\partial l_n}{\partial \theta_k}. \tag{14}$$

Given that the loss value  $l_n$  is caused by training sample  $\mathbf{F}_n \in C_i$ , the gradient of  $l_n$  with respect to the reference vectors are derived based on Eqs. (5), (11), and (12),

$$\frac{\partial l_n}{\partial \mathbf{m}_i} = \frac{2 \alpha g_j \exp(-\alpha d_n)(\mathbf{m}_i - \mathbf{F}_n)}{g_i^2 [1 + \exp(-\alpha d_n)]^2}, \tag{15}$$

and

$$\frac{\partial l_n}{\partial \mathbf{m}_j} = \frac{2 \alpha \exp(-\alpha d_n)(\mathbf{F}_n - \mathbf{m}_j)}{g_i [1 + \exp(-\alpha d_n)]^2}, \tag{16}$$

**Table 4** Error rate of the defect detection on unknown test samples.

Defect type	Error rate (%)		
	Haar wavelet with 3-scale wavelet features	Adaptive wavelet with 2-scale wavelet features	Adaptive wavelet with 3-scale wavelet features
Misyarn	6.8	6.5	6.5
Broken	10.3	8.4	3.4
Stain	8.7	11.2	3.4
Average	8.7	8.8	4.5

where  $i \neq j$ .

The gradient of  $l_n$  with respect to the lattice coefficients is derived as follows:

$$\frac{\partial l_n}{\partial \theta_k} = \frac{2\alpha g_j \exp(-\alpha d_n)}{g_i^2 [1 + \exp(-\alpha d_n)]^2} \times \sum_{p=1}^P \left( \{[F_n^{(p)} - m_i^{(p)}]g_j - [F_n^{(p)} - m_j^{(p)}]g_i\} \frac{\partial F_n^{(p)}}{\partial \theta_k} \right), \quad (17)$$

where  $P$  denotes the dimension of the feature vector  $\mathbf{F}_n$ , and  $F_n^{(p)}$  denotes the  $p$ 'th component of the feature vector  $\mathbf{F}_n$ . According to Eq. (4), the feature component  $F_n^{(p)}$  corresponds to one component  $w_r^d$ . From Eq. (3), we have

$$\frac{\partial w_r^d}{\partial \theta_k} = \text{Mean}_{(x,y) \in \text{window}} \left[ 2W_r^d(x,y) \frac{\partial W_r^d(x,y)}{\partial \theta_k} \right]^2, \quad (18)$$

where

$$\frac{\partial W_r^d(x,y)}{\partial \theta_k} = \frac{\partial \psi_r^d}{\partial \theta_k} * I(x,y), \quad (19)$$

In Eq. (19),  $(x,y)$  are the spatial indices and  $*$  denotes the 2-D convolution.  $\psi_r^d$  is the cascade of the filter sequences in the wavelet transform corresponding to the output of  $W_r^d(x,y)$ . By using the chain rule for differentiation, the calculation of  $\partial \psi_r^d / \partial \theta_k$  depends on the calculation of  $\partial H(z) / \partial \theta_k$  and  $\partial G(z) / \partial \theta_k$ , which can be derived based on the lattice structure as given in Eq. (8)

$$\begin{aligned} \begin{bmatrix} \partial H(z) / \partial \theta_k \\ \partial G(z) / \partial \theta_k \end{bmatrix} &= R_m \Lambda(z) R_{m-1} \Lambda(z) M(z) \\ &\times \begin{bmatrix} -\sin \theta_k & \cos \theta_k \\ -\cos \theta_k & -\sin \theta_k \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} N(z) \\ &+ \partial \theta_m / \partial \theta_k \begin{bmatrix} -\sin \theta_m & \cos \theta_m \\ -\cos \theta_m & -\sin \theta_m \end{bmatrix} \\ &\times \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} R_{m-1} \Lambda(z) W(z), \end{aligned} \quad (20)$$

where  $m+k$  is even, or

$$\begin{aligned} \begin{bmatrix} \partial H(z) / \partial \theta_k \\ \partial G(z) / \partial \theta_k \end{bmatrix} &= R_m \Lambda(z) R_{m-1} \Lambda(z) M(z) \\ &\times \begin{bmatrix} -\sin \theta_k & \cos \theta_k \\ -\cos \theta_k & -\sin \theta_k \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} N(z) \\ &+ R_m \Lambda(z) \partial \theta_{m-1} / \partial \theta_k \\ &\times \begin{bmatrix} -\sin \theta_{m-1} & \cos \theta_{m-1} \\ -\cos \theta_{m-1} & -\sin \theta_{m-1} \end{bmatrix} \\ &\times \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} W(z), \end{aligned} \quad (21)$$

where  $m+k$  is odd.

Here  $N(z)$  and  $M(z)$  are  $2 \times 1$  and  $2 \times 2$  matrices of polynomials whose coefficients are functions of  $\{\theta_j\}_{0 \leq j \leq k-1}$  and  $\{\theta_j\}_{k+1 \leq j \leq m-2}$ , respectively.  $W(z)$  are  $2 \times 1$  matrices of polynomials whose coefficients are functions of  $\{\theta_j\}_{0 \leq j \leq m-2}$ .

## References

1. C. Neubauer, "Segmentation of defects in textile fabric," *Proc. IEEE Intl. Conf. Patt. Recog.* **1**, 688–691 (1992).
2. Y. F. Zhang and R. R. Bresee, "Fabric defect detection and classification using image analysis," *Text. Res. J.* **65**(1), 1–9 (1995).
3. A. Bodnarova, J. A. Williams, M. Bennamoun, and K. K. Kubik, "Optimal textural features for flaw detection in textile materials," *Proc. IEEE TENCON*, pp. 307–310 (1997).
4. F. S. Cohen, Z. Fan, and S. Attali, "Automated inspection of textile fabrics using textural models," *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(8), 803–808 (1991).
5. C. H. Chan and G. Pang, "Fabric defect detection by Fourier analysis," *IEEE Trans. Ind. Appl.* **36**(5), 1267–1276 (2000).
6. C. Ciambellini, F. Francini, P. Sansoni, and B. Tiribilli, "Defect detection in textured materials by optical filtering with structured detectors and self-adaptable masks," *Opt. Eng.* **35**(3), 838–844 (1996).
7. A. Kumar and G. Pang, "Defect detection in textured materials using Gabor filters," *Proc. IEEE Intl. Conf. Industrial Application* **2**, 1041–1047 (2000).
8. J. Escofet, R. Navarro, and M. S. Millan, "Detection of local defects in textile webs using Gabor filters," *Opt. Eng.* **37**(8), 2297–2307 (1998).
9. H. Sari-Sarraf and J. S. Goddard, "Vision system for on-loom fabric inspection," *IEEE Trans. Ind. Appl.* **35**(6), 1252–1259 (1999).
10. G. J. Vachtsevanos, M. Mufti, and J. L. Dorrity, "Method and apparatus for analyzing an image to detect and identify defects," U.S. Patent, No. 5815198 (1998).
11. S. Kim, M. H. Lee, and K. B. Woo, "Wavelet analysis to fabric defects detection in weaving processes," *Proc. IEEE Intl. Symp. Industrial Electron.* **3**, 1406–1409 (1999).
12. G. Lambert and F. Bock, "Wavelet methods for texture defect detection," *Proc. IEEE Intl. Conf. Image Process.* **3**, 201–204 (1997).
13. W. J. Jasper, S. J. Garnier, and H. Potlapalli, "Texture characterization and defect detection using adaptive wavelets," *Opt. Eng.* **35**(11), 3140–3149 (1996).
14. Y. N. Gong, "Study on image analysis of fabric defects," PhD dissertation, China Textile University, Shanghai, China (1999).
15. X. Z. Yang, G. K. H. Pang, and N. H. C. Yung, "Fabric defect detection using adaptive wavelet," *Proc. IEEE Intl. Conf. Acoustics, Speech Sig. Process.* **6**, 3697–3700 (2001).
16. A. Biem, S. Katagiri, E. McDermott, and B. H. Juang, "An application of discriminative feature extraction to filter-bank-based speech recognition," *IEEE Trans. Speech Audio Process.* **9**(2), 96–110 (2001).
17. S. Katagiri, B. H. Juang, and C. H. Lee, "Pattern recognition using a family of design algorithms based upon the generalized probabilistic descent method," *Proc. IEEE* **86**(11), 2345–2373 (1998).
18. Z. Cvetkovic and M. Vetterli, "Discrete-time wavelet extrema representation: design and consistent reconstruction," *IEEE Trans. Signal Process.* **43**(3), 681–693 (1995).
19. M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Process.* **4**(11), 1549–1560 (1995).
20. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(7), 674–693 (1989).
21. B. H. Juang and S. Katagiri, "Discriminant learning for minimum error classification," *IEEE Trans. Signal Process.* **40**(12), 3043–3054 (1992).
22. A. Biem, S. Katagiri, and B. H. Juang, "Pattern recognition based on discriminative feature extraction," *IEEE Trans. Signal Process.* **45**(2), 500–504 (1997).
23. P. P. Vaidyanathan and Z. Doganata, "The role of lossless systems in modern digital signal processing: a tutorial," *IEEE Trans. Educ.* **32**(3), 181–197 (1989).
24. R. Fletcher, *Practical Methods of Optimization*, 2nd ed., Wiley and Sons, Inc., New York (1987).



**Xue Zhi Yang** received his BEng degree from Anhui University, China, in 1992, and MEng degree from Hefei University of Technology, China, in 1995, all in electronic engineering. Currently, he is working toward the PhD degree in the Department of Electrical and Electronic Engineering at the University of Hong Kong. His research interests include pattern recognition, multi-scale image/signal processing using wavelets, and automated visual inspection system.



**Grantham K. H. Pang** obtained his PhD degree from the University of Cambridge in 1986 for research in multivariable control system design and expert systems. He was with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, from 1986 to 1996, and joined the Department of Electrical and Electronic Engineering at The University of Hong Kong in 1996. He has published more than 130 technical papers and has authored or coauthored six books. His research interests include

expert systems for control system design, intelligent control, intelligent transportation systems, neural networks, control theory, and computer-aided design.



**Nelson H. C. Yung** received his BSc and PhD degrees from the University of Newcastle-Upon-Tyne in 1982 and 1985, respectively. He was lecturer at the same university from 1985 until 1990. From 1990 to 1993, he worked as a senior research scientist at the Department of Defence, Australia. He joined the University of Hong Kong in late 1993 as Associate Professor. He leads a research team in digital image processing and intelligent transportation systems, and a number of research collaborative projects with various companies and organizations in the HKSAR, China, Japan, and Australia. He is the founding Director of the *Laboratory for Intelligent Transportation Systems Research* at HKU. He has coauthored a computer vision book, and has published more than 100 journal and conference papers in the areas of digital image processing, parallel algorithms, visual traffic surveillance and autonomous vehicle navigation.