Restructured Eigenfilter Matching for Novelty Detection in Random Textures

A. Monadjemi, M. Mirmehdi, B. Thomas
Department of Computer Science
University of Bristol, Bristol, BS8 1UB, England
{monadjem,majid,barry}@cs.bris.ac.uk

Abstract

A new eigenfilter-based novelty detection approach to find abnormalities in random textures is presented. The proposed algorithm reconstructs a given texture twice using a subset of its own eigenfilter bank and a subset of a reference (template) eigenfilter bank, and measures the reconstruction error as the level of novelty. We then present an improved reconstruction generated by structurally matched eigenfilters through rotation, negation, and mirroring. We apply the method to the detection of defects in textured ceramic tiles. The method is over 90% accurate, and is fast and amenable to implementation on a production line.

1 Introduction

Surface inspection and quality classification of tiles is an essential stage in the tile manufacturing industry. Due to the high cost of human inspection, speed of production line and repetitious nature of the activity, development of an automatic inspection and defect detection system would have an impressive impact on the overall performance of a tile production plant. Typically, defects (e.g. a crack, a colour drop or a random pattern density change) affect the expected texture of the tile and hence can signify a 'textural abnormality'. Figure 1 shows normal and defective tiles for two different patterns. We are particularly interested in randomly textured tiles where defects are more difficult to discern.

One major advantage in applications such as tile manufacturing is that one always has significant constraints, e.g. during a production run, one will always see one type of tile, albeit with random textures. Hence, a training stage can concentrate on one particular tile type rather than several different ones. However, the goal would still be for the reuse of the developed algorithms for application to different tile patterns at training and testing stages.

A typical approach to defect classification comprises feature extraction from normal and abnormal samples, then training a classifier, and subsequently applying the trained classifier to classify unknown samples [8]. By employing subtle pre and post processing steps, efficient features, and powerful classifiers, we can expect good classification performance. However, this well-established approach can suffer from practical drawbacks. Firstly, in real cases the number of abnormal training samples is usually much

BMVC 2004 doi:10.5244/C.18.66

lower than the number of normals. Secondly, defects are diverse and unpredictable, so even a large set of negative samples may not necessarily cover all the possible forms of defects. In other words, in a given N-dimensional feature space, whereas the normal samples build well-separated clusters, we may see many ill-defined scattered regions of abnormal samples. However, the abnormal samples should show a common characteristic: a considerable distance to normal clusters. This distance is the basis of a different approach to normal-abnormal sample classification called *novelty detection* or *concept learning* [5, 13].

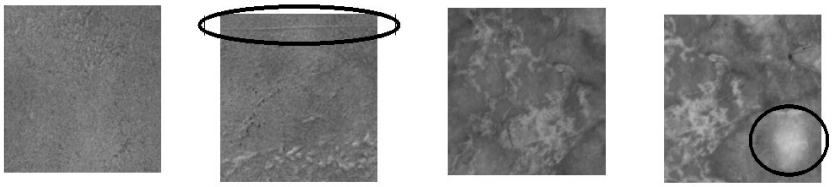


Figure 1: Normal and abnormal tile samples from two different tile types.

In this paper we report a new eigenfilter based novelty detection scheme to classify abnormal textured tiles. The method is based on the double reconstruction of the test image, once by a subset of its own eigenfilter bank and once by a subset of the template tile's eigenfilter bank, and measuring the reconstruction error as the level of novelty. Eigenfilters are an orthogonal set, therefore any reconstruction using a complete set would be error free and result in the original image. Hence, we examine (a) how to determine a relevant subset of them only for the reconstruction stages and (b) how to determine the optimum threshold on the reconstruction error to label a tile defective or not. An improved method which rearranges the eigenfilters by rotation, negation, and mirroring is then developed for better matching between the test and template eigenfilter sets. We start with reviewing some relevant background work and outlining our experimental data set in Section 2. Then the proposed method and the results of the first series of experiments are detailed in Section 3. The improved structure-based eigenfilter matching method is discussed in section 4. The paper is concluded in Section 5.

2 Background

Texture analysis techniques such as co-occurrence matrices or frequency domain-based methods can be computationally expensive for the demands of a real-time defect inspection system, but nevertheless they have been exploited in many studies as highly accurate techniques. One pioneering work was by Kruger et *al.* [7] on the diagnosis of coal workers' Pneumoconiosis using texture features extracted via both co-occurrence matrices and Fourier domain analysis. In [12], a vision system was developed for fabric inspection, exploiting multi-scale Wavelet representation to obtain basis images. Bernoulli's rule of combination was then used to recompose the images to highlight the edges and defective regions. In [3], Boukouvalas et *al.* used optimal filters to detect abnormal lines and spots in tiles. They also used the Wigner distribution to combine the advantages of both spatial and spatial frequency domains to detect cracks. Furthermore in [6], the authors presented a method for detecting random texture tile defects consisting of K-means clustering, fol-

lowed by perceptual merging of clusters in Luv space and morphological analysis. This was computationally expensive although a promising approach.

Karhunen-Loeve transform (KLT) and eigenfiltering are popular in texture classification and defect detection works, for example [2, 14, 9]. Unser [14] tested different local linear transforms such as KLT, discrete cosine, sine and Hadamard transforms (DCT, DST, and DHT respectively) for texture classification and found KLT as the best algorithm. Also Ade et *al.* [2] compared Laws filters, KLT, DCT and DHT for textile defect detection. In their experiments, the KLT performance, particularly on larger window sizes, was amongst the best. In [9], Kumar and Pang developed an eigenfilter-based optimal filtering scheme for texture defect detection yielding promising results for fabric defect detection. They employed linear FIR filters with optimised energy separation for both supervised and unsupervised defect detection. For optimisations, they focused on selection of eigenvectors based on a new maximisation function and the selection of the filter size.

It is common in novelty detection studies to apply auto-associative neural networks (e.g. Worden [16] for cracked beam defect detection) or self organising maps (e.g. Iivarinen et *al.* for web surface inspection [4]). However, both schemes are rather structurally complicated and need a subtle training stage as well. Here we avoid any network-type structure or a complex training phase and only have a brief training stage to determine optimum values for two parameters for our proposed method.

Our data set contains over 1500 tiles of eight diverse types of textures. Each tile is represented as 256×256 pixels at a resolution of 4 pixels/mm. Samples of each type were divided into four non-overlapping categories of *abnormal*, *normal*, *template* (or *reference*) and *parameter estimation* (or *training*). The number of samples in the normal and abnormal categories was kept equal and all the remaining good samples were used to build the template and the training sets. Typically for each tile type in this study, the template set contained a few hundred and the training set contained a few tens of images.

3 Proposed Method

Consider the process of breaking an image G into its constituent eigenimages which can also be used to reconstruct the image. Using sliding windows, neighbourhood $\mathcal{N}$patches of size $n \times n$ in G are extracted and rearranged as different observations of data into a $k \times n^2$ matrix, where k is the number of patches. The patch size is typically in the range of $\mathcal{N}$=3×3, 5×5 and 7×7. The covariance matrix C, is then computed and the eigenvectors and eigenvalues are obtained:

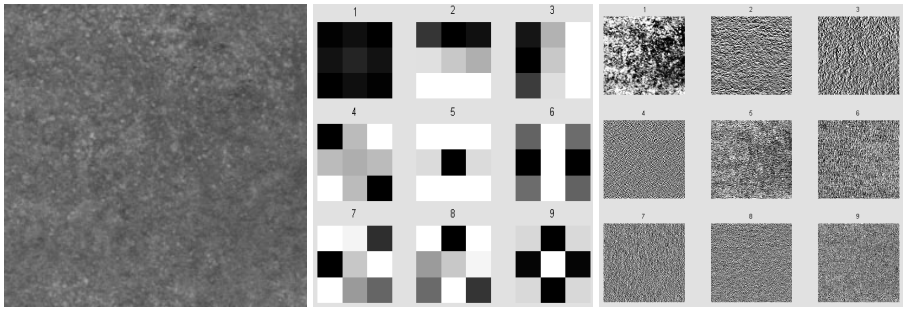
$$C(x) = E[(x - \overline{x})(x - \overline{x})^T] \tag{1}$$

$$(C(x) - \lambda_x I)e = 0 \tag{2}$$

where $\overline{x}$ is the mean value, I is the unit matrix, λ is the eigenvalue and e is the eigenvector matrix. A $n \times n$ rearrangement of the eigenvectors could be interpreted as a bank of adapted filters of the same size, which optimally cover all $n \times n$ relations of the test image pixels. Then basis images are obtained by 2D spatial domain convolution of the test image by the members of the eigenfilter bank:

$$D_i^G = G \otimes F_i \ , \quad i = \{1, ..., \mathcal{N}\} \tag{3}$$

where F_i and D_i are the i^{th} eigenfilter and basis image from a list of n^2 basis images of each input [1, 11]. Fig. 2 shows a tile, its 3×3 eigenfilter bank, and basis images. The filters' orthogonality is of importance as it builds up uncorrelated basis images ordered by their role as constituents of the original image. These together subtly describe and can reconstruct the image.



(a) (b) (c)

Figure 2: (a) Original tile, (b) its 3×3 eigenfilter set, (c) detailed/basis images.

Initially, we compute and store the eigenfilters of all the reference images in our *template* set T. Then, the eigenfilters of an unseen tile image G are computed (e.g. see Fig. 2(b)). We search the eigenfilters of the template set to find the image M with the most similar set of eigenfilters to G. For the search, the widely applied (e.g. in [15]) χ^2 distance function is used. Given A and B as any two eigenfilters that are to be compared, then:

$$\chi^2(A,B) = \sum_{i=1}^{\mathcal{N}} \frac{(A_i - B_i)^2}{|A_i| + |B_i|} \tag{4}$$

Next, the test image G is reconstructed twice, once by a subset of its own eigenfilters (R_G), and once again by a subset of the matched template M's eigenfilters (R_M). The number of filters in the subset is naturally important and is dealt with in Section 3.2. Reconstruction could be carried out by simple addition of the filter responses or by Bernoulli's rule of combination [12]. In our experiments, Bernoulli's rule did not show any advantage to simple addition, hence subsequently only addition-based reconstruction was used:

$$R_G = \sum_i D_i^G \quad and \quad R_M = \sum_i D_i^M \qquad i = \{1,..,\mathcal{N}\} \tag{5}$$

An error between the reconstructed pair larger than a given threshold Υ is then considered as an indication of abnormality of test image G:

$$(\Delta E = |R_G - R_M|) > \Upsilon \quad \longrightarrow \text{ DEFECT} \tag{6}$$

The second column in Table 1, shows the best classification accuracy (CA) for three different neighbourhood sizes, $\mathcal{N}$=3×3, 5×5 and 7×7 for this reconstruction. The computation for $\mathcal{N}$=3×3 took $0.501s$ on a 700 MHz PC in Matlab. At the expense of slightly more elaborate computation, instead of the eigenfilters, we can search for the best matching filter responses, i.e. basis images, before reconstructing the images. The

χ^2 search distance comparison then becomes:

$$\chi 2(D^G, D^M) = \sum_{i=1}^{\mathcal{N}} \frac{(D_i^G - D_i^M)^2}{|D_i^G| + |D_i^M|} \tag{7}$$

$\mathcal{N}$	**Eigenfilters Only CA**	**Using Basis Images CA**	**No. of Filters**
3×3	81.12%	85.32%	6
5×5	77.62%	81.25 %	12
7×7	74.22%	79.90%	32

Table 1: Classification performance for the MBE approach

This significantly increases accuracy, as shown in the third column of Table 1, at a very marginal increase in computation time to $0.538s$. The number of filters used in the reconstruction stage in each case are shown in the final column. We refer to this method as *Matching by Eigenfilters* or MBE.

If we were to ignore the reconstruction stage and detect the defects by only examining the distance amongst the eigenfilter sets for the test image and the template, a reduced accuracy of only 72.41% would be achieved. This vindicates the use of reconstructed images to perform the classification. However, the best result is still not satisfactory. After dealing with the Υ parameter and the best number of filters in the next two subsections, we show how the results can be improved through the restructuring of eigenfilter subsets in reconstructing G.

3.1 Finding optimum Υ

The choice of Υ is determined through a simple training or parameter estimation stage depending on the type of tile texture. Initially, we apply the proposed algorithm on the training set P (which contains only 'good' samples) and obtain the reconstruction errors. Then the mean (μ_P) and the standard deviation (σ_P) of the reconstruction errors are computed. The optimum threshold Υ is assumed to be an α weighted deviation from the mean:

$$\Upsilon = \mu_P + \alpha\sigma_P \tag{8}$$

Thus, any unseen tiles with reconstruction error $\Delta E > \Upsilon$ will be considered as abnormal. Furthermore, we continue the parameter estimation stage to determine the optimum value for α using k-fold cross validation. Here a 4-fold cross validation was employed, where 75% of samples were used for parameter estimation, and 25% for testing the performance. The result is taken as the average of four iterations of that procedure on non-overlapped subsets.

As an example, Fig. 3 depicts the distributions of ΔE for normal, abnormal and template sets of a specific type of tile. In this case, the template's ΔE statistics are $\mu_P = 0.24$, and $\sigma_P = 0.11$ (Fig. 3(top)). The test data reconstruction errors are plotted in Fig. 3(bottom). The cross validation algorithm estimates the optimum separation parameter set as $\Upsilon = 0.36$, from $\alpha = 1.12$, on the normalised ΔE axis. The subsequent correct classification rate for this example was 95.0%.

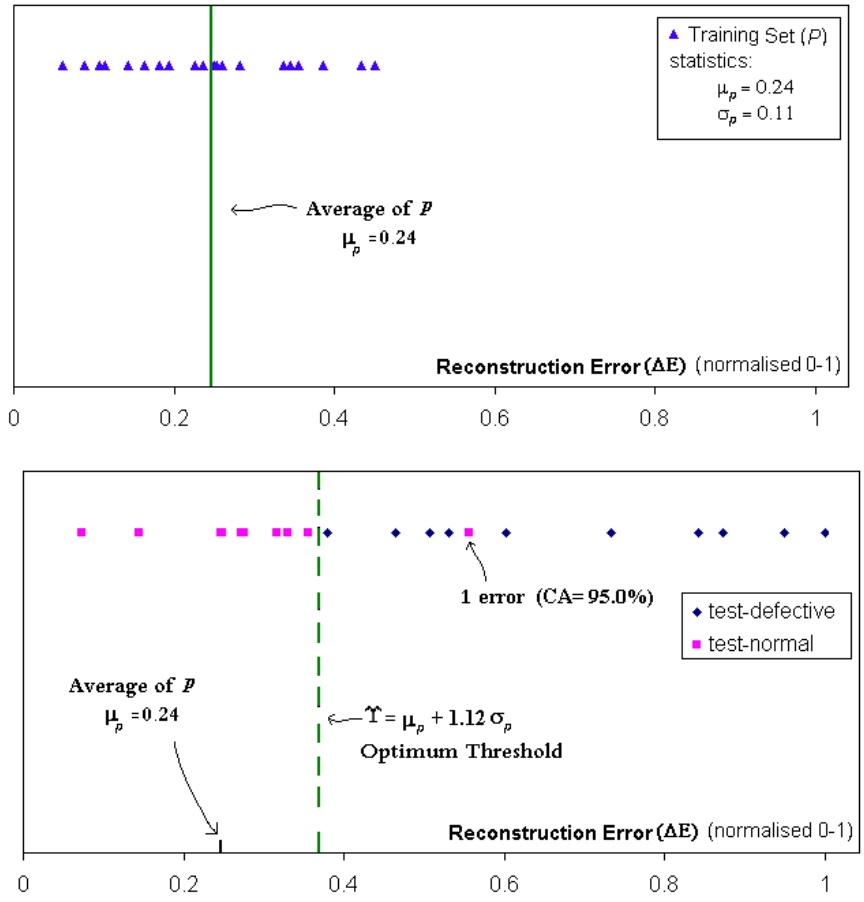


Figure 3: Reconstruction error (ΔE) distribution for (top) training set P, (bottom) normal and abnormal test samples. The training set parameters (μ_P, σ_P) are used in computing the optimum threshold for the test samples. The ΔE axis has been normalised to lie in the range [0-1].

Next we describe how the near-optimal number of eigenfilters are selected to reconstruct an image, whether for routine comparison of unseen tiles against the template set T or to determine the optimum value of Υ as just described.

3.2 The Subset of Eigenfilters

Choosing the proper subset of eigenfilters is fundamental to obtaining the most accurate classification results. Randomly selecting a number of the eigenvectors corresponding to the highest eigenvalues does not necessarily capture the defective area characteristics since such areas are relatively small in the image. We obtained an average of only 72.92% accuracy using this approach.

For the experimental results reported in Table 1 we measured the χ^2 distance between corresponding eigenfilters of the test image G and the template image M and then used the measures to reorder the eigenfilter pairs into a list, from minimum distance to maximum. We then assumed pairs at the maximum distance end of the list are novelty indicators that convey the differences, i.e. abnormality information, whereas pairs at the minimum distance end of the list convey the similarities. To determine how many pairs in the list should be used we exhaustively tried all possible values from $\{2,...,\mathcal{N}-1\}$ in the training stage to determine the optimum number. This was then used throughout the testing stage. In all the tile types, we found the optimum value was around $\mathcal{N}/2$, so the search for the optimum when testing a new tile type can be limited to the vicinity of that number for the given neighbourhood size.

In summary, the eigenfilter based approach proposed so far manages a respectable classification accuracy of 85.32% with parameters that can be easily determined through a simple training stage for the tile texture under production. Next, we introduce an improvement to increase this accuracy.

4 Improvement through matching by structure

When we examine the 3×3 eigenfilter sets for two sample images, as shown in Fig. 4, it becomes apparent that some corresponding filters may be negations and/or rotations of each other (for example see pairs 2, 3 and 9), meaning that the computed overall distance between two textures would not be reliable. In fact, even a simple rotation of the same tile will result in rotated eigenfilters and disrupt the distance measurements. The figure also demonstrates that finding similarities or relationships amongst larger 5×5 filters is much more involved than smaller 3×3 ones, mainly due to the considerably higher diversity of 5×5 matrices. Therefore the MBE scheme would be even less effective for larger neighbourhood sizes. Non-smooth surfaces of many tiles and slanted lighting of the image grabber system, (which is essential for visibility of some sorts of defects), amplify the effects of rotation on the texture. Varma and Zisserman have considered this phenomenon as a 3D effect on textures [15].

We therefore develop a more effective and rotation-invariant scheme by matching the filters' structure. This will allow all filters of a template image M to compete as the possible corresponding i^{th} filter of the test image G, regardless of their associated eigenvalues. Hence, the minimum distance δ between two specific filters is computed as:

$$\delta_{A_i,B_j} = min(\ \chi^2(A_i,B_j)\ ,\ \chi^2(A_i,B_j^{\rightleftarrows})\ ,\ \chi^2(A_i,B_j^{\uparrow\downarrow})\ ,\ \chi^2(A_i,\overline{B_j})\ ,\ \chi^2(A_i,B_j^{\theta})\ ) \quad (9)$$

where $\chi^2(A_i,B_j^*) = \frac{(A_i-B_j^*)^2}{|A_i|+|B_j^*|}$, $\theta = \{90^\circ, 180^\circ, 270^\circ\}$, $(\overset{\rightleftarrows}{\cdot})$ and $(\cdot\ \uparrow\downarrow)$ are vertically and horizontally mirrored (i.e. 'flipped') matrices, $(\bar{\cdot})$ is the complement (i.e. 'negative') of the matrix, and M_j^θ indicates the θ degrees rotated version of the input. Mirroring is implemented by swapping the columns or rows of M. Complementing is performed by using the mean value of the filter μ_X as the origin:

$$\overline{X} = -(X-\mu_X)+\mu_X \quad (10)$$

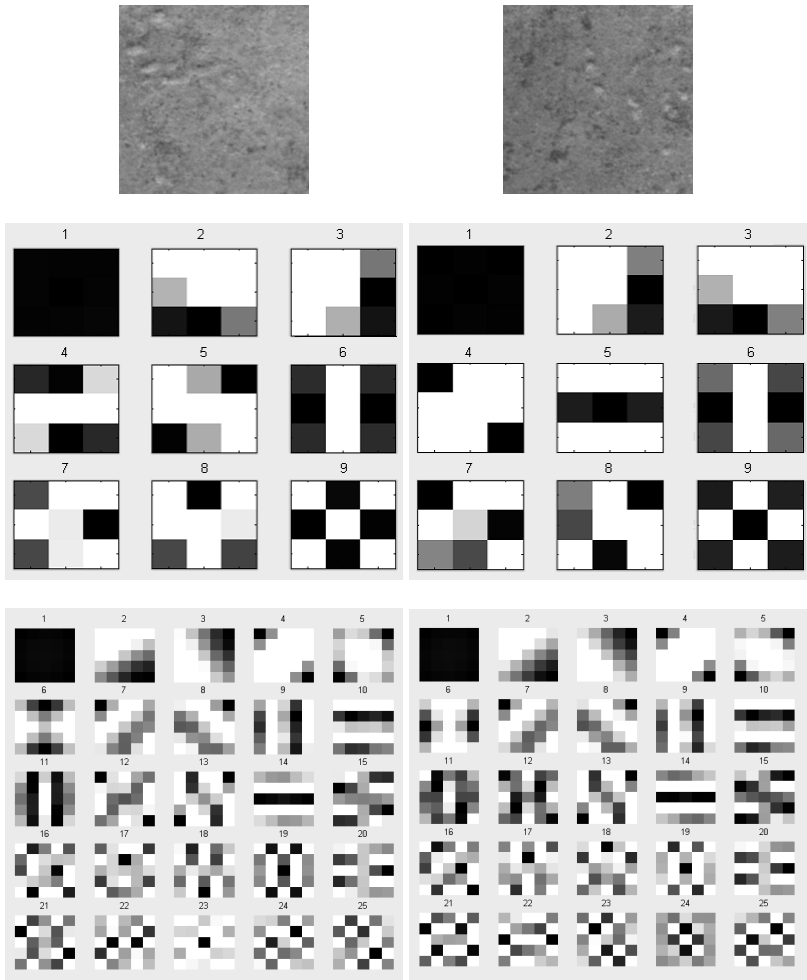


Figure 4: (top) Two original tiles, (middle) their 3×3 filter banks, (bottom) their 5×5 filter banks.

We refer to this improved approach as *matching by structure* or MBS. The number of filters in the subset used for computing the reconstructed image is then worked out as before (see Section 3.2). Table 2 presents the performance of the MBS method, including accuracy, sensitivity (SNS), and specificity (SPC), averaged across our data set, showing a slight improvement over MBE for the smaller neighbourhood size of $\mathcal{N}$=3×3, and considerable improvements for larger sizes $\mathcal{N}$=5×5, 7×7, and 9×9 with the best result at 91.46% overall accuracy for $\mathcal{N}$=7×7 (cf. Table 1). This seems reasonable as a 3×3 neighbourhood may be less likely to capture the characteristics of the texture than a slightly larger one. Unfortunately however, the combined demands of the increased number of χ^2 comparisons and the larger neighbourhoods result in an exponential increase in computation time: 0.740s, 2.55s, 11.4s, and 69.2s respectively for the four $\mathcal{N}$s tested. The table also shows the average number of filters used in the reconstruction stages.

Table 3 presents the detailed results of the MBS method for different $\mathcal{N}$for all our tile types. Except for case S1, both $\mathcal{N}$=5×5 and 7×7 always achieve better classification accuracy than $\mathcal{N}$=3×3. By the time we get to a 9×9 neighbourhood, a decline in the accuracy can be observed, thus showing $\mathcal{N}$=7×7 as the optimal window on average. However, the $\mathcal{N}$=5×5 case achieves a very close average to the $\mathcal{N}$=7×7 case, while also getting better individual accuracies for L1 and S1 tile textures, all at much lower computational cost. Perhaps optimising $\mathcal{N}$for a given type in the training stage could improve the overall classification performance even more.

Neighbourhood Size	**CA%**	**SNS**	**SPC**	**No. Filters Involved**
3×3	86.71	0.898	0.835	5
5×5	91.19	0.969	0.853	14
7×7	91.46	0.972	0.855	26
9×9	90.74	0.966	0.850	43

Table 2: Classification performance using matched-by-structure filters

	$\mathcal{N}$=3×3			$\mathcal{N}$=5×5			$\mathcal{N}$=7×7			$\mathcal{N}$=9×9		
Type	**CA%**	**SNS**	**SPC**	**CA%**	**SNS**	**SPC**	**CA%**	**SNS**	**SPC**	**CA%**	**SNS**	**SPC**
A1	82.45	0.852	0.797	83.72	0.874	0.800	87.54	0.875	0.875	84.56	0.832	0.859
A2	76.29	0.850	0.675	87.24	0.919	0.825	89.84	0.965	0.831	87.33	0.943	0.804
D1	100	1	1	100	1	1	100	1	1	99.81	0.998	0.998
D2	79.29	0.792	0.792	100	1	1	100	1	1	99.81	0.999	0.998
K1	88.00	0.899	0.861	93.55	0.978	0.893	97.81	1	0.956	94.27	0.998	0.887
L1	85.29	0.934	0.771	86.76	1	0.735	81.72	1	0.624	81.97	0.976	0.712
S1	92.22	0.922	0.922	85.34	1	0.706	80.80	0.969	0.646	85.30	0.983	0.682
S2	90.15	0.935	0.868	92.89	0.988	0.869	93.98	0.968	0.910	92.88	0.998	0.859
μ	86.71	0.898	0.835	91.19	0.969	0.853	91.46	0.972	0.855	90.74	0.966	0.850
σ^2	0.006	0.004	0.010	0.004	0.002	0.012	0.006	0.002	0.022	0.004	0.003	0.013

Table 3: Classification accuracy of different tile types for different neighbourhood sizes $\mathcal{N}$using MBS. μ and σ^2 are mean and variance.

The results presented in Table 3 can be improved upon by resorting to a standard training and testing classification paradigm. We applied a backpropagation neural network (BPNN) classifier trained on Gabor filter features and obtained an average accuracy of 96.57%. However, unlike the proposed novelty detection scheme, this approach required hours of training for each tile type making it less practical for realistic implementation. It also can suffer from having to deal with previously unseen defects as described in Section 1. For more comparative results and analysis, the reader is referred to [10].

5 Conclusion

We introduced a method of novelty detection for tile inspection based on eigenfilter analysis. Initially the MBE approach was outlined consisting of the reconstruction of a given texture twice using a subset of its own eigenfilter bank and a subset of a reference (template) eigenfilter bank. A measure of the reconstruction error indicated the existence of defects. We then presented the MBS scheme comprising an improved reconstruction generated by structurally matched eigenfilters through rotation, negation, and mirroring. This

made eigenfilter comparison more resilient and meaningful in terms of matching a template's eigenfilter against an unseen tile's. Results were presented for a large selection of tile textures.

Like other novelty detection schemes, the most important advantage of the proposed method is its relatively low dependence on negative samples. The MBS approach is rotationally invariant and capable of more accurate normal/abnormal classification. The method has full potential for real-time implementation and we plan to exploit this by further code optimisation.

Acknowledgments

This work was supported by EC project G1RD-CT-2002-00783-MONOTONE.

References

[1] F. Ade. Characterization of textures by eigenfilters. *Signal Processing*, 5:451–457, 1983.

[2] F. Ade, N. Lins, and M. Unser. Comparison of various filter sets for defect detection in textiles. In *ICPR*, volume I, pages 428–431, 1984.

[3] C. Boukouvalas, J. Kittler, R. Marik, M. Mirmehdi, and M. Petrou. Ceramic tile inspection for colour and structural defects. In *AMPT95*, pages 390–399, 1995.

[4] J. Iivarinen, K. Heikkinen, J. Rauhamaa, P. Vourimaa, and A. Visa. A defect detection scheme for web surface inspection. *IJPRAI*, 14(6):735–755, 2000.

[5] N. Japkowicz, C. Myers, and M. Gluck. A novelty detection approach to classification. In *IJCAI-95*, pages 518–523, 1995.

[6] J. Kittler, R. Marik, M. Mirmehdi, M. Petrou, and J. Song. Detection of defects in colour texture surfaces. In *IAPR Conference on Machine Vision Apllications*, pages 558–567, 1994.

[7] R. Kruger, W. Thompson, and A. Turner. Computer diagnosis of pneumoconiosis. *IEEE T-SMC*, 4(1):40–49, 1974.

[8] A. Kumar. Neural network based detection of local textile defects. *Pattern Recognition*, 36(7):1645–1659, 2003.

[9] A. Kumar and G. Pang. Defect detection in textured materials using optimized filters. *IEEE T-SMC*, 32(5):553–570, 2002.

[10] A. Monadjemi. *Texture Classification and Abnormality Detection*. PhD thesis, Department of Computer Science, University of Bristol, October 2004.

[11] T. Randen and J.H. Husoy. Filtering for texture classification: A comparative study. *IEEE T-PAMI*, 21(4):291–310, 1999.

[12] H. Sari-Sarraf and J. S. Goddard. Vision system for on-loom fabric inspection. *IEEE T-IA*, 35(8):1252–1259, 1999.

[13] L. Tarassenko, P. Hayton, N. Cerneaz, and M. Brady. Novelty detection for the identification of masses in mammograms. In *ICANN*, pages 442–447, 1995.

[14] M. Unser. Local linear transform for texture measurements. *Signal Processing*, 11:61–79, 1986.

[15] M. Varma and A. Zisserman. Classifying images of materials:achieving viewpoint and illumination independence. In *ECCV*, pages 255–271, 2002.

[16] K. Worden. Structural fault detection using a novelty measure. *Journal of Sound and Vibration*, 201(1):85–101, 1997.