ROTATION INVARIANT TEXTURE CLASSIFICATION USING DIRECTIONAL FILTER BANK AND SUPPORT VECTOR MACHINE

Hong Man, Ling Chen and Rong Duan

Department of ECE, Stevens Institute of Technology
Hoboken, NJ 07030,
{hman, lchen, rduan}@stevens-tech.edu

ABSTRACT

This paper presents a rotation invariant texture classification method using a special directional filter bank (DFB) and support vector machine (SVM). This method extracts a set of coefficient vectors from directional subband domain, and models them as multivariate Gaussian densities. Eigen-analysis is then applied to the covariance metrics of these density functions to form rotation invariant feature vectors. Classification is based on SVM, which only takes non-rotated images for training and uses images at various rotation angles for testing. Experimental results have shown that this DFB is very effective in capturing directional information of texture images, and the proposed rotation invariant feature generation and SVM classification method can in fact achieve relatively consistent classification accuracy on both non-rotated and rotated images.

1. INTRODUCTION

Texture classification is a fundamental building block of image analysis that is frequently applied in a verity of important applications, such as target recognition, robotic vision, image/video indexing and retrieval, and data mining etc. Texture classification has been an active research topic for several decades. However effective and efficient classification of rotated texture images remains to be a challenge. A number of methods for rotation invariant texture classification have been proposed [1], and most research interests lie on rotation invariant feature extraction. Madiraju and Liu [2] proposed a method using eigen-analysis of local covariance of image blocks to obtain six rotation invariant features representing roughness, anisotropy and other high-order texture characteristics. Charalampidis and Kasparis [3] also introduced roughness features in directional wavelet domain based on fractal dimension (FD). The directional wavelet is implemented as linear combination of two orthogonal wavelets, which is referred to as "steerable wavelet". Steerable wavelet is also studied by Do and Vetterli in [4], in which a Gaussian hidden Markov tree (HMT) is used to model cross-scale wavelet coefficients. Rotation invariance is achieved by replacing covariance matrices in HMT parameter set with matrices of eigenvalues. Porter and Canagarajah [5] introduced a wavelet domain feature using circularly symmetric Gaussian Markov random field (GMRF) model for rotation invariance.

In this paper, we present a texture classification method based on the special properties of a unique directional filter bank (DFB) for feature generation. This DFB was developed by Bamberger [6] and improved by Park [7] to obtain visualizable subband domain representation.

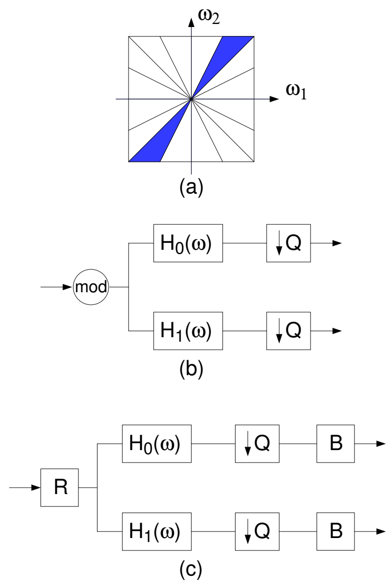

Fig. 1. Directional filter bank structure.

This paper is organized as follows. In section 2, we briefly review the operators and properties of the DFB. In section 3, we discuss the multi-dimensional Gaussian distribution as well as eigen-analysis that can produce rotate invariant features. In section 4, we introduce the classification menthod based on support vector machine. In section 5, we present the experimental results. Conclusion is provided in the last section.

2. DIRECTIONAL FILTER BANK

The directional filter bank [7] is able to partition the frequency plane into a set of equal-sized wedge-shaped passbands, as shown in Fig 1(a). It can be implemented efficiently through a series of two-band subband decompositions. At each stage of the two-band decomposition, two complementary fan filters are applied, and a special downsampling matrix Q is used to take samples lying on a quincunx lattice for the output. The fan filters at different stages are implemented through two different procedures. For the first two decomposition stages, the structure of the filter bank is shown in Fig 1(b), and for the following stages, the structure of the filter bank takes the form shown in Fig 1(c). In these structures, $H_0(\omega)$ is a diamond filter and $H_1(\omega)$ is its complement; Q represents a quincunx downsampling matrix. MOD in Fig 1(b) corresponds to a modulation of the input by π on either n_1 or n_2 direction in spatial domain, which shifts the diamond shape passband to fan shape passbands. In Fig 1(c), R represents a unitary frequency resampling (skewing) matrix that can reshape a diamond passband into different parallelogram passbands, and together with the passbands of the previous stages these will produce wedge-shaped passbands. After quincunx downsampling Q each directional subband takes shape of a rectangle. In traditional filter bank decomposition, each subband maintains the original image structure. With this DFB, the resampling matrix and the quincunx downsampling matrix causes the content of each subband skewed and rotated. A backsampling matrix B is therefore introduced in Fig 1(c) [7] to compensate this distortion and rearrange the subband coefficients so that each subband becomes visually proportional to the original image with only exception of ration aspect. Fig 2 provides examples of an eight-band DFB decomposition of the image STRAW at two rotation angles.

In texture analysis, discriminative information usually resides in high frequency regions. Although this DFB provides good directional resolution, it does not provide frequency resolution. Each subband covers the whole frequency spectrum. To avoid negative impact of low frequency variations, we apply a highpass prefiltering before the DFB.

3. FEATURE GENERATION IN DIRECTIONAL SUBBAND DOMAIN

Texture classification based on the directional filter bank was first reported in [8], in which the distribution of directional subband coefficients are modelled as zero-mean Gaussian density, and a variance is extracted from each subband to form the feature vector. For an eight-band decomposition, each image is represented by a vector of eight variance values. The conditional distribution of the feature vectors from each class is assumed to be a multivariate Gaussian density, with a mean vector and a covariance matrix that can be calculated from training feature vectors. The classification of a test feature vector is based on the minimum Bayes distance to each of the class distributions. This method is able to achieve very high classification accuracy on the non-rotated Brodatz texture images when large number of training images were generated using overlapping partitioning (e.g. 100 images per class).

The work in [8] was focused on general texture classification without concerning rotation invariance. In this work, we present a new feature generation method that can take advantage of directional resolution provided by the DFB and formulate a rotation invariant feature for texture classification.

As oppose to the assumption made in [8], we consider that the probability distribution of coefficients from different directional subband are somewhat correlated. We model these coefficients as a single multivariate Gaussian density. More specifically, we first use down-sampling to split each rectangular subband into two or more smaller subbands with the original aspect ratio. The purpose is to unify the size of all subbands. Coefficients are just regrouped, and no coefficient is missed in the down-sampling procedure. We then take one coefficient from each resulting subband at the same location to form an observation vector. When all coefficients within each subband are scanned, a sequence of observation vector is generated. This vector sequence is used to estimate the covariance matrix of the multivariate Gaussian density. Because of the prefiltering, we assume this Gaussian is zero mean, i.e. the mean vector contains all zeros. The covariance matrix not only describes the distributions for individual subbands, but also indicates the correlation among the distributions of different subbands. Therefore we attempt to use the covariance matrix as the feature vector for each image, and we assume the covariance matrices of different images belonging to the same class will cluster in a high dimensional space.

When the original image is rotated, we expect that all the coefficients inside each subband are collectively rotated by the same angle, i.e. the subband domain image will have the same orientation as the rotated original image. However the magnitude of coefficients inside each specific subband may change. For example, if an image has a strong directional component, its energy in directional subband domain will be mostly concentrated in one or two subbands corresponding to that direction. After a rotation, this energy concentration still exists, but will be in different subbands. This represents the special energy compaction property of this DFB.

A simple example is shown in Fig 3, in which we assume that two subbands (N=2) are obtained from the DFB decomposition. While the original input image produces a large energy concentration in one of the subbands, as shown in Fig 3 (a), a 90-degree rotation of the input will shift the energy concentration to the other subband and therefore rotate the bi-variate Gaussian density, as shown in Fig 3 (b). In this case, the two subbands just exchange their density functions.

Based on this observation, we realize that the principal axes of the multivariate Gaussian density can be a good candidate for rotation invariant feature. The lengths of these principal axes are the eigenvalues of the covariance matrix, and if these eigenvalues are sorted according to their values, they can form a feature vector that will not be affected by any rotation of the multivariate Gaussian density.

An N-dimensional multivariate Gaussian density function has the form of

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{C}|^{1/2}} exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T\mathbf{C}^{-1}(\mathbf{x}-\mu)\right), \quad (1)$$

where $\mathbf{x}$ is the observation vector, μ is the mean vector, and $\mathbf{C}$ is the covariance matrix. Then according to eigen decomposition theorem, the covariance matrix can be decomposed into the form

$$\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T, \quad (2)$$

where columns of $\mathbf{U}$ are the normalized eigenvectors of $\mathbf{C}$, and $\mathbf{\Lambda}$ is a diagonal matrix containing the corresponding eigenvalues λ_i for $i = 1, \ldots, N$. The eigenvalues are sorted in descending order, and each of them representing the variance of the multivariate density along a principle axis determined by the corresponding eigenvector. The column vector $\mathbf{v} = [\lambda_1, \lambda_2, \ldots, \lambda_N]^T$ is then used as the feature vector to represent one particular image.

A further advantage of this feature vector is that its size can be easily reduced by keeping only a few largest eigenvalues, i.e. principle components. This approach can effectively reduce the computation at both training and testing phases.

4. SVM CLASSIFICATION METHOD

Support vector machine (SVM) is a linear classification method frequently used in pattern recognition applications. For binary

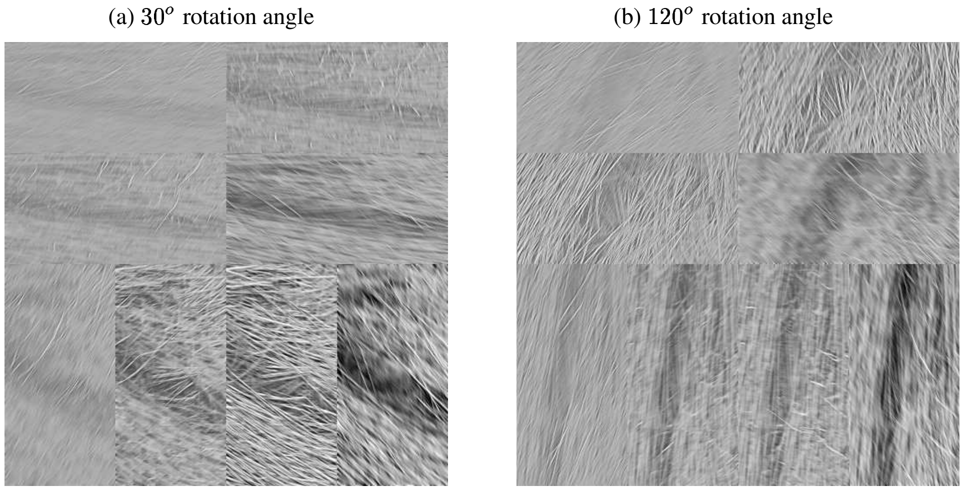

Fig. 2. Examples of 8-band directional subband decomposition of image STRAW.

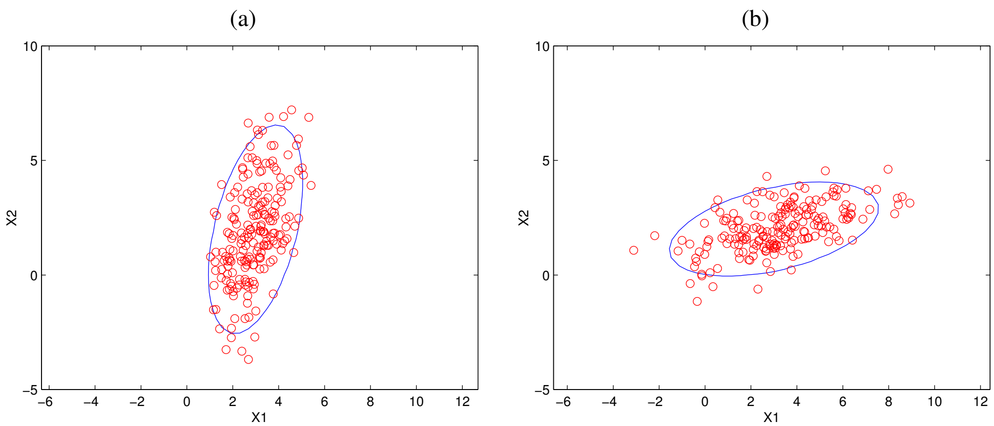

Fig. 3. Example of a bivariate Gaussian distribution with energy shift caused by rotation.

classification, SVM is to find a decision hyperplane that separates the training samples belonging to two classes with the largest margin. Let $\mathbf{x}_i \in \Re^n$, i=1,2,...,N, be a feature vector of the training set. For each $\mathbf{x}_i$, a class indicator is denoted as $y_i \in -1, 1$, which classifies $\mathbf{x}_i$ to one of the two classes. For linearly separable data set, the optimal separating hyperplane can be expressed as

$$f(\mathbf{x}) = \sum_{i=1}^{N} \lambda_i y_i \mathbf{x_i} \cdot \mathbf{x} + b, \tag{3}$$

where λ_i and b are the solutions of a quadratic optimization problem that maximizes the separating margin, and

$$\sum_{i=1}^{N} \lambda_i y_i = 0. \tag{4}$$

Given this decision hyperplane, any new test feature vector $\mathbf{x}$ will be classified to one of the two classes based on the sign of $f(\mathbf{x})$.

For non-linearly separable data, one option is to project the feature vectors to a high dimensional space where a linear separating hyperplane can be found. SVM training and testing in this high dimensional space is based on $\Phi(\mathbf{x_i})$ and $\Phi(\mathbf{x})$ where $\Phi(\cdot)$ is the mapping function. If the mapping function satisfies

$$\Phi(\mathbf{x_i}) \cdot \Phi(\mathbf{x}) = \mathcal{K}(\mathbf{x_i}, \mathbf{x}), \tag{5}$$

where $\mathcal{K}(\cdot)$ is a kernel function, The decision surface cab be rewritten as

$$f(\mathbf{x}) = \sum_{i=1}^{N} \lambda_i y_i \mathcal{K}(\mathbf{x_i}, \mathbf{x}) + b. \tag{6}$$

Therefore the calculation remains in the $\Re^n$ feature space, which avoids explicit implementation of the mapping function $\Phi(\cdot)$. There are a number of kernel functions that have been found to have good generalization capacities.

The binary SVM classification can be extended to multi-class classification in a pair-wise fashion. For an M-class problem, a total of $M(M-1)/2$ SVMs can be trained. At each step of the testing phase, the input feature vector is classified into one of the two possible classes, one of which is resulting from the previous classification step.

5. EXPERIMENTAL RESULTS

The Brodatz texture dataset [9] is used in our experiments. It contains thirteen classes of texture images with size of 512×512.

texture	SVM-16	SVM-8	var. vec.[8]
bark	0.8614	0.8416	0.5644
brick	0.0792	0.0891	0.0300
bubble	0.8911	0.8911	0.6039
grass	0.8911	0.8911	0.6634
leather	0.7327	0.5743	0.3267
pigskin	0.5149	0.6436	0.1584
raffia	0.4554	0.4059	0.1683
sand	0.8416	0.8416	0.9010
straw	0.7525	0.6337	0.4158
water	0.7426	0.5644	0.0396
weave	0.9505	0.9406	0.0396
wood	0.7030	0.6832	0.0495
wool	0.7525	0.7030	0.1188

Table 1. Classification performance of rotated texture images averaged over all seven different angles.

Each class was from a single texture photo, which was digitized at each of the seven rotation angles, i.e. 0^o, 30^o, 60^o, 90^o, 120^o, 150^o and 200^o. In order to obtain sufficient number of training and testing images, we partition each image into sixteen non-overlapping small images with size of 128×128. Therefore for each of the thirteen classes we have $16 \times 7 = 112$ images, out of which eleven training images are randomly chosen solely from the 0^o (or non-rotated) images (i.e. the first 11), and all the rest are used as testing images.

Each 128×128 images is first passed through a highpass filter to eliminate the effect of DC component. The highpass filter is a complement of a 9×9 rotationally symmetric Gaussian low-pass filter with $\sigma = 2$. The resulting image is then filtered by an eight-band DFB, producing eight subbands with sizes of 64×32 or 32×64 as shown in Figure 2. In order to unify the sizes of subbands, we split each subband into two 32×32 small subbands using subsampling. So totally there are sixteen such subbands for each input image. Then a 16-dimensional feature vectors is calculated according to our feature generation method in Section 3.

Tow linear SVM classifiers are implemented. The first SVM is based on the normal 16-dimensional feature vectors. The second SVM uses the simplified 8-dimensional feature vectors, each of which contains the largest eight components (or principle components) from a 16-dimensional feature vector.

For comparison, we also implemented the feature generation and classification method introduced in [8], and we applied the same test conditions as used in our methods.

Table 1 shows the average classification results from various methods on images over all seven different angles. And Table 2 shows the average classification results for seven different angles over all thirteen classes. It is clear that our proposed methods achieve significant improvement on rotated images. For textures with strong directional structures, e.g. bark, grass, and weave, our methods work remarkably well on all different orientations.

It should be pointed out that we used all the images at all the rotation angles in these tests. We did not discard any image containing heterogeneous texture (e.g. "straw") or insufficient representative structure (e.g. "brick"), as did in some of previous works. Also the focus of this paper is to study rotation invariant feature generation. The classification performance can be further improved by refining the operators at several stages of our method.

Rotation	SVM-16	SVM-8	var. vec.[8]
0^o	0.8154	0.7846	0.8000
30^o	0.7500	0.7067	0.3400
60^o	0.6058	0.5577	0.2010
90^o	0.7067	0.7212	0.2500
120^o	0.6538	0.6394	0.2060
150^o	0.6587	0.6010	0.3120
200^o	0.8221	0.7548	0.4180

Table 2. Classification performance of images at different rotation angles averaged over all thirteen classes.

6. CONCLUSION

In this paper we have presented a new rotation invariant feature generation ad classification method based on the directional filter bank and support vector machine. For each image, the subband coefficients are modelled as a multi-variate Gaussian density, and the feature vector contains the eigenvalues of the covariance matrix. Two linear SVM classifiers are implemented based on different feature vectors. Experimental results clearly demonstrated the possible performance improvement over an existing feature generation method based on the same filter bank.

7. REFERENCES

[1] S. R. Fountain, T. N. Tan, and K. D. Baker, “Comparative study of rotation invariant classification and retrieval of texture images,” in *Proc. British Machine Vision Conf.*, 1998.

[2] S. V. R. Madiraju and C.-C. Liu, “Rotation invariant texture classification using covariance,” in *IEEE International Conference on Image Processing*, vol. 2, (Austin, TX), pp. 655–659, Nov 1994.

[3] D. Charalampidis and T. Kasparis, “Wavelet-based rotational invariant roughness features for texture classification and segmentation,” *IEEE Trans. on Image Processing*, vol. 11, pp. 825–837, Aug. 2002.

[4] M. Do and M. Vetterli, “Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden markov models,” *IEEE Trans. on Multimedia*, vol. 4, pp. 517–527, Dec. 2002.

[5] R. Porter and N. Canagarajah, “Robust rotation invariant texture classification,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, (Munich, Germany), pp. 3157–3160, Apr 1997.

[6] R. H. Bamberger and M. J. T. Smith, “A filter bank for the directional filter decomposition of images,” *IEEE Trans. on Signal Processing*, vol. 40, pp. 882–893, Apr. 1992.

[7] S.-I. Park, M. J. T. Smith, and R. M. Mersereau, “A new directional filter bank for image analysis and calssification,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, (Chicago, IL), pp. 1286–12190, Oct 1999.

[8] J. Rosiles and M.J.T.Smith, “Texture classification with a biorthogonal directional filter bank,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, (Salt Lake City, UT), pp. 1549 – 1552, May 2001.

[9] S. Uni. Southern California and I. P. Institute, “Rotated textures.” http://sipi.usc.edu/services/database/Database.html.