A NEW SUBSPACE LEARNING METHOD IN FOURIER DOMAIN FOR TEXTURE CLASSIFICATION

Shu Liao and Albert C.S. Chung

Lo Kwee-Seong Medical Image Analysis Laboratory,
Department of Computer Science and Engineering,
The Hong Kong University of Science and Technology, Hong Kong.

ABSTRACT

This paper proposes a new texture classification approach. There are two main contributions in the proposed method. First, input texture images are transformed to the composite Fourier domain (CFD) by using both the local and global Fourier transforms. The composite Fourier domain is rotation invariant and preserves the contextual information for the texture images in the original spatial domain. Second, the null-space based linear discriminant analysis (nLDA) is adopted to find the optimal representations of the texture images in the composite Fourier domain. This paper proposes a systematic way to cooperate subspace learning methods for texture classification in the frequency domain, which cannot be directly applied in the spatial domain for texture classification. The proposed method is evaluated on both the Brodatz and CUReT databases and compared with several state-of-the-art texture classification approaches. Experimental results show that the proposed method achieves the highest classification rate among all the compared methods.

***Index Terms*—** Texture Classification, Composite Fourier Domain, Null-space Based LDA

1. INTRODUCTION

Texture classification plays an important role in many real world applications such as content based image retrieval, remote sensing, and material categorization. In general, there are two main steps for texture classification: feature extraction and classification. In this paper, we mainly focus on the feature extraction step.

In the literature, various kinds of texture features for texture classification were proposed: Chellappa *et al* [1] proposed the Gaussian Markov random field (GMRF) to model the texture appearance and the optimal parameters of the GMRF model are estimated by least squared method and adopted as features for texture images. Laine and Fan [2] used the wavelet packet signatures to represent texture images. Varma and Zisserman [3] proposed the texton histogram features based on the local exemplar image patches. An essential property for texture features is rotation invariance, i.e., as texture images belonging to the same class should be recognized despite of the absolute orientations under which they are taken. Therefore, Deng and Clausi [4] proposed the anisotropic circular Gaussian MRF (ACGMRF) model to make the MRF feature rotation invariant. Porter and Canagarajah [5] removed the absolute orientation by combining the LH and HL wavelet subband coefficients of the traditional wavelet based method [2]. In recent years, Ojala *et al* [6] proposed the uniform local binary pattern (LBP) approach to extract features from texture images. LBP is a rotation and monotonic gray-level transformation invariant feature which can also be computed efficiently.

In this paper, we propose a new feature extraction method for texture classification. The proposed method is a subspace learning method in a composite Fourier domain (CFD). The CFD is constructed based on both the local and global Fourier transforms of the texture images. The main contributions of the paper lie in the following aspects. (1) The composite Fourier domain is introduced, which provides rotation invariant texture representations in the frequency domain; (2) The most discriminant subspace in the composite Fourier domain is then estimated such that the projection coefficients on such subspace in the composite Fourier domain are served as the final features to represent each texture image.

The composite Fourier domain is constructed by two steps. First, the local Fourier transform is applied to the original texture images such that each pixel is represented by a rotation invariant signature. The global multi-dimensional Fourier transform is then applied to the local Fourier transformed image to obtain the multi-dimensional frequency domain coefficients for the image, which is the resulting composite Fourier domain (CFD) representation. The null-space based linear discriminant analysis (nLDA) [7] is applied to the images transformed to the CFD to find the optimal subspace representation in CFD and adopted as the final features for classification. The support vector machine (SVM) with the Gaussian radial basis function (RBF) kernel is used as the classifier in this paper.

The proposed method is evaluated on both the Brodatz [8] and CUReT [9] databases and compared with several widely used texture classification approaches. Experimental results show that the proposed method achieves the highest classification accuracy among all the compared methods.

2. THE COMPOSITE FOURIER DOMAIN

In this section, details about the transformation from input texture images to the composite Fourier domain (CFD) are given.

2.1. The Local Fourier Transform Signatures for each Pixel

The first step to construct CFD is to apply the local Fourier transform to each pixel such that each pixel is represented by a rotation invariant signature which can also reflect its anatomical properties. Such procedure is illustrated in Figure 1:

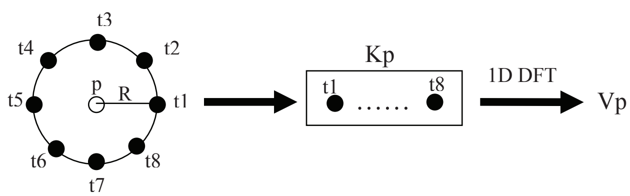

Fig. 1. Illustration of applying the local Fourier transform with respect to each pixel p.

As shown in Figure 1, for each referencing pixel p in the texture image, a circularly symmetric neighborhood system is defined with radius R and number of neighboring samples N. In Figure 1, the number of neighboring samples N is 8. Radius R controls the scale of interest. Given the values of R and N, N samples are uniformly taken on the circle centered at p with radius R as shown in Figure 1, for samples which do not fall exactly on the image grid, their intensities are obtained by using the bilinear interpolation. We denote the first neighboring sample as the pixel to the right of the referencing pixel p with distance R as shown in Figure 1, other neighboring samples' indices are determined based on an anti-closewise order from the first neighboring sample. In Figure 1, the eight neighboring samples' intensities are denoted by $t_1, ..., t_8$ respectively. Then, $t_1, ..., t_8$ are organized as a one dimensional row vector in order, denoted as K_p in Figure 1.

The 1D discrete Fourier transform is then applied to $K_p = [t_1, t_2, ..., t_N]$ as shown in Figure 1, the resulting one dimensional vector now is $V_p = [v_1, v_2, ..., v_N]$, with v_i $(i = 1, ..., N)$ calculated by Equation 1.

$$v_i = \sum_{n=1}^{N} t_n \cdot \exp(\frac{-2\pi i}{N} \cdot i \cdot n). \quad (1)$$

It should be noted that v_i $(i = 1, ..., N)$ are complex fourier coefficients. We then denote $|V_p| = [|v_1|, |v_2|, ..., |v_N|]$

Algorithm 1 Perform the Global Multi-Dimensional Fourier Transform

Input: The local Fourier transformed image $LFT(G)$. The number of dimension N of $|V_p|$ for each $p \in G$.

Output: The transformed image X in the composite Fourier domain.

1. Set $X = LFT(G)$.
2. FOR i = 1 to N
3. Perform 1D Fourier transform along dimension i with Equation 1 to X.
4. END FOR
5. Return X.

where each element i of vector $|V_p|$ is the magnitude of v_i. The $|V_p|$ signature for pixel p is rotation invariant because image rotation actually results in the circular shift of vector K_p, and magnitude vector of the 1D discrete Fourier transform corresponding to a shifted vector remains the same. Therefore, the column vector $|V_p|$ is used as the rotation invariant signature for each pixel p. $|V_p|$ also reflects the anatomical properties around pixel p based on its neighboring samples.

The value of radius R shown in Figure 1 reflects the scale of interest. In this paper, three different values of R: $R_1 = 1$, $R_2 = 2$, $R_3 = 3$ are used to achieve the multi-resolution analysis. The numbers of neighboring samples corresponding to each radius are: $N_1 = 8$, $N_2 = 16$, $N_3 = 24$. Therefore, after the local Fourier transform procedure, each pixel now is represented by a 48 $(8 + 16 + 24)$ dimensional feature vector. Suppose the original texture image is G, we denote the output vector image after performing the local Fourier transform to G as $LFT(G)$.

2.2. The Global Multi-Dimensional Fourier Transform

After the local Fourier transform procedure, each pixel $p \in G$ now is represented by its own rotation invariant feature signature $|V_p|$. In order to analyze the frequency component configuration of $LFT(G)$, the next step is to perform the global multi-dimensional Fourier transform to $LFT(G)$ to obtain the coefficients of $LFT(G)$ in each frequency band. The global multi-dimensional Fourier transform can be summarized in Algorithm 1.

After applying the global multi-dimensional Fourier transform to $LFT(G)$ as shown in Algorithm 1, the resulting image X is said to be in the composite Fourier domain (CFD). It should be noted that each multi-dimensional index of X now represents the coefficient of a particular multi-dimensional frequency band to reconstruct $LFT(G)$.

3. NULL-SPACE BASED LINEAR DISCRIMINANT ANALYSIS IN THE COMPOSITE FOURIER DOMAIN

After transforming the original texture image G to the composite Fourier domain (CFD), the null-space based linear discriminant analysis (nLDA) [7] is adopted to find out the optimal discriminant subspace in CFD. The nLDA [7] method is derived from the conventional LDA [10] which can avoid the small sample size problem. The nLDA formulation is described as follows.

Suppose there are in total m texture images $G_1, G_2, \ldots, G_m$ belonging to c different classes with n_r rows and n_c columns for each image. Each class D_i $(i = 1, ..., c)$ has C_i images. For each texture image G_j $(j = 1, ..., m)$, let K_j denote its corresponding vector image in the composite Fourier domain, and let d denote the number of dimensions of the composite Fourier domain.

In this paper, we only make use of the magnitude of the transformed image K_j in CFD. Each transformed image in CFD K_j, $j = 1, 2, \ldots, m$, is represented as a column vector $\vec{H}_j$, which contains $n_r \times n_c \times d$ components. $\vec{H}_j$ is denoted as,

$$\vec{H}_j = \left[mag(\vec{K}_j(1,1)), \ldots, mag(\vec{K}_j(n_r, n_c))\right]^T, \quad (2)$$

where $mag(\vec{F})$ denotes the magnitude of each element of vector $\vec{F}$. nLDA aims at finding a projection matrix $\mathbf{W}$ such that the original feature vector $\vec{H}_j$ is reduced to a lower dimensional feature vector by using Equation 3,

$$\vec{Y}_j = \mathbf{W}^T \vec{H}_j. \quad (3)$$

$\mathbf{W}$ consists of a set of column vectors $\vec{q}$ which are computed based on the maximization of the following generalized Fisher's discriminant criterion,

$$J_{\text{nLDA}}(\vec{q}) = \frac{\vec{q}^{\,t}\mathbf{S_B}\vec{q}}{\vec{q}^{\,t}\mathbf{S_W}\vec{q} + \vec{q}^{\,t}\mathbf{S_B}\vec{q}}, \quad (4)$$

where $\mathbf{S_B}$ and $\mathbf{S_W}$ are the between class scatter matrix, and the within class scatter matrix respectively.

$$\begin{aligned}
\mathbf{S_B} &= \sum_{i=1}^{c} C_i(\vec{\mu_i} - \vec{\mu_g})(\vec{\mu_i} - \vec{\mu_g})^T, \\
\mathbf{S_W} &= \sum_{i=1}^{c} \sum_{\vec{H} \in D_i} (\vec{H} - \vec{\mu_i})(\vec{H} - \vec{\mu_i})^T,
\end{aligned}$$

where $\vec{\mu_i}$ is the $n_r \times n_c \times d$ dimensional mean vector of the class D_i, i.e. $\vec{\mu_i} = \frac{1}{C_i}\sum_{\vec{H}\in D_i}\vec{H}$. $\vec{\mu_g}$ is the $n_r \times n_c \times d$ global mean vector defined as $\vec{\mu_g} = \frac{1}{m}\sum_{j=1}^{m}\vec{H}_j$.

The above problem is in turn transformed to a generalized eigenvalue problem,

$$\mathbf{S_B W} = \lambda \mathbf{S_W W}. \quad (5)$$

In the case that $\mathbf{S_W}$ is not a singular matrix, Equation 5 is formulated as,

$$(\mathbf{S_W})^{-1}\mathbf{S_B W} = \lambda \mathbf{W}. \quad (6)$$

In which, the first k eigenvectors $\vec{V_1}, \vec{V_2}, \ldots, \vec{V_k}$ corresponding to the first k largest eigenvalues of the matrix $(\mathbf{S_W})^{-1}\mathbf{S_B}$ are selected to form the projection matrix. Therefore, $\mathbf{W} = [\vec{V_1}, \vec{V_2}, \ldots, \vec{V_k}]$. Since the rank of $\mathbf{S_B}$ is at most $c-1$. The transformed feature vector $\vec{Y_j}$ can have at most $c-1$ dimensions. When the number of samples m is smaller than the number of dimensions of $\vec{H}_j$, the small sample size problem occurs. In this case $\mathbf{S_W}$ becomes singular and Equation 6 cannot be applied. In this case, the singular value decomposition is performed on $\mathbf{S_W}$ such that $\mathbf{S_W} = \mathbf{U\Sigma V^t}$, suppose the rank of matrix $\mathbf{S_W}$ is r and $\mathbf{U} = [\vec{U}_1, \ldots, \vec{U}_r, \ldots, \vec{U}_{n_r \times n_c \times d}]$, where $\vec{U}_{r+1}, \ldots, \vec{U}_{n_r \times n_c \times d}$ denote the eigenvectors which have zero eigenvalues.

The projection matrix $\mathbf{Q} = [\vec{U}_{r+1}, \ldots, \vec{U}_{n_r \times n_c \times d}]$ can then transform the input features into the null-space of $\mathbf{S_W}$. The between class scatter matrix $\mathbf{S_B}$ is transformed to the null-space of $\mathbf{S_W}$ and given as,

$$\tilde{\mathbf{S_B}} = \mathbf{QQ}^t\mathbf{S_B}(\mathbf{QQ}^t)^t. \quad (7)$$

Finally, certain number of eigenvectors corresponding to the largest eigenvalues of matrix $\tilde{\mathbf{S_B}}$ are used to form the projection matrix $\mathbf{W}$. The number of eigenvectors, m', is chosen according to the criterion,

$$\frac{\sum_{i=1}^{m'}\lambda_i}{\sum_{s=1}^{c-1}\lambda_s} \geq 0.95, \quad (8)$$

where λ_i is the ith largest eigenvalues. After obtaining the projection matrix $\mathbf{W}$, then the original feature vectors H_j, $j = 1, 2, \ldots, m$, are transformed into m' dimensional feature vectors using the Equation 3. The minimum number of m' that satisfies Equation 8 is used as the number of the dimensions of the transformed feature space.

It is worth pointing out that the nLDA procedure cannot be directly applied to the texture images in the spatial domain similar to the issue of face recognition [10] because facial images are topological objects and can be aligned to similar positions before performing nLDA. However, texture images usually contain repetitive patterns which are difficult or even impossible to be aligned to the same spatial domain. The important role of global Fourier transform in Section 2.2 is to "align" the texture images to the common composite Fourier domain.

4. EXPERIMENTAL RESULTS

The proposed method was evaluated on two texture databases: (1) 24 textures selected from the Brodatz album [8]; (2) 61 textures selected from the CUReT database [9]. The proposed method is also compared with several widely used texture features. In all the experiments, the support vector machine (SVM) with the Gaussian radial basis function (RBF) kernel was adopted as the classifier. The optimal parameters of SVM are found by performing grid search over the parameter space.

For the Brodatz database, each texture class contains 25 64×64 images. Each time 13 images were randomly selected from each class and used as training images, the rest of the images were served as the testing images. To test the robustness of different approaches against image rotation, training and testing images were also rotated by a randomly generated angle between 0 to 360 degrees. The experiments were repeated 100 times and Table 1 lists the mean classification accuracy of different approaches for the Brodatz database for the original and rotated images.

	Classification Accuracy (in $\%$)	
Methods	Original Image	Randomly Rotated Image
1. Wavelet [2]	97.43	82.03
2. ACGMRF [4]	94.18	92.74
3. LBP [6]	98.15	97.69
4. nLDA + CFD	**99.32**	**98.86**

Table 1. The mean classification accuracy of different methods on the Brodatz database. The last row represents the proposed method.

It is observed in Table 1 that under the original texture image (i.e. no rotation) condition, Wavelet [2], LBP [6] and the proposed method achieve high classification accuracy (i.e. above 95%). When the texture images are randomly rotated, the classification accuracy of Wavelet [2] drops significantly as it is not rotation invariant. Under both conditions, the proposed method achieves the highest classification accuracy among all the compared methods.

For the CUReT database [9], there are in total 61 classes of textures. Each class has 92 texture images taken under different illumination conditions and viewing angles with resolution 200×200. For each time 46 texture images were randomly selected from each class and used as the training set, the rest of the images were served as the testing set. The experiments were repeated for 100 times. Table 2 lists the mean classification accuracy for different approaches.

	Wavelet [2]	ACGMRF [4]	LBP [6]	nLDA + CFD
CA (in $\%$)	62.75	83.34	96.38	**98.67**

Table 2. The mean classification accuracy of different methods on the CUReT database. The last column represents the proposed method. CA denotes the classification accuracy.

It is observed in Table 2 that the proposed method also has the highest classification accuracy among all the compared methods. Therefore, the robustness of the proposed method is strongly implied.

5. CONCLUSION

In this paper, a new texture classification method is proposed. The proposed method can be factorized into two stages. First, the original texture images are transformed to the composite Fourier domain by applying both the local Fourier transform and global Fourier transform in order. It is theoretically proved in this paper that the local Fourier transform can assign each pixel a rotation invariant anatomical signature. Then, the coefficients of different frequency bands are obtained via the global Fourier transform. Finally, the null-space based linear discriminant analysis (nLDA) is performed in the composite Fourier domain to extract the most discriminant features for each texture image. The proposed method was evaluated on both the Brodatz and CUReT databases and compared with three widely used texture classification approaches. Experimental results show that the proposed method consistently achieves the highest classification accuracy among all the compared methods, which illustrates the robustness of the proposed method.

6. REFERENCES

[1] R. Chellappa and S. Chatterjee, "Classification of textures using Gaussian Markov random fields," *IEEE ASSP*, vol. ASSP-33, no. 4, pp. 959–963, 1985.

[2] A. Laine and J. Fan, "Texture Classification by Wavelet Packet Signatures," *IEEE PAMI*, vol. 15, no. 11, pp. 1186–1191, 1993.

[3] M. Varma and A. Zisserman, "A Statistical Approach to Material Classification Using Image Patch Exemplars," *IEEE PAMI*, vol. 31, pp. 2032–2047, 2009.

[4] H. Deng and D.A. Clausi, "Gaussian MRF Rotation-Invariant Features for Image Classification," *IEEE PAMI*, vol. 26, no. 7, pp. 951–955, 2004.

[5] R. Porter and N. Canagarajah, "Robust rotation-invariant texture classification: wavelet, Gabor filter and GMRF based schemes," *IEE Proceedings of VISP*, vol. 144, no. 3, pp. 180–188, 1997.

[6] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE PAMI*, vol. 24, no. 7, pp. 971–987, 2002.

[7] L.F. Chen, H.Y.M. Liao, M.T. Ko, J.C. Lin, and G.J. Yu, "A new LDA-based face recognition system which can solve the small sample size problem," *Pattern Recognition*, vol. 33, pp. 1713–1726, 2000.

[8] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, New York: Dover Publications, June 1966.

[9] K. Dana, B. Ginneken, S. Nayar, and J. Koenderink, "Reflectance and Texture of Real-World Surfaces," *ACM Transactions on Graphics*, vol. 18, no. 1, pp. 1–34, January 1999.

[10] Peter N. Belhumeur, Joao P. Hespanha, and David J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE PAMI*, vol. 19, pp. 711–720, 1997.