Texture Retrieval Based on Nonlocal Singular Value Decomposition and Multiscale Transforms

J.W. Zhang, B. Liu

School of Information Science & Engineering, Lanzhou University, Lanzhou, China

Abstract—**Feature extraction based on the multiscale transforms is usually accomplished in the high frequency subbands. The features which are captured from the high frequency subbands mainly contain the edge changing information of images; meanwhile the structural information of images is mainly existed in the low frequency subband. In order to improve the effect of image processing, the structural features which are very important should be obtained from images. This paper adopts a nonlocal singular value decomposition (NL-SVD) algorithm to extract structural features from the low frequency subband. The structural features are combined with traditional statistical features such as means and standard deviations. These features are applied for texture retrieval. Experimental results show the retrieval ratio with these combined features is better and it proves that structural features are efficient texture features for images.**

Keyword-texture retrieval; NL-SVD; multiscale transforms; feature extraction

I   INTRODUCTION

Feature extraction is indispensable for texture retrieval. Texture features are the visual features which can reflect images' homogeneity phenomenon. They contain a variety of effective information of images. All kinds of methods which extract texture features are more and more popular in the field of image processing. Multiscale transforms are practical and efficient in these methods.

Multiscale transforms are widely used for feature extraction. With multiscale transforms, all the information of images is divided into two parts which are stored in the high frequency subbands and low frequency subbands. There are lots of statistical and geometrical features extracted from the high frequency subbands for showing the inherent characteristics of images [1]. These statistical features have been used for texture classification and applied in different applies [2]. Figure 1 (a), (b) show two texture images A1, A2 from Brodatz [3] and (c) (d) show the results of their four-level wavelet decomposition respectively. Traditional feature vector is formed with statistical features such as means and variances of the high frequency subband coefficients, and usually this feature vector represents the edge changing information (statistical features). But we can see from the two enlarge size low frequency subbands of two decomposed coefficients image of Fig.1, there are still remain sufficient visual information which are may useful for image processing. Because low frequency subband is an approximate of it original image under multiscale transform, so the low frequency subband has the same structural information with its original image but with reducing dimensions. For these reasons, we hope that the structural information can be mapped into structural features and can also be regard as a kind of texture features.

This paper adopts an intermediate algorithm called the nonlocal SVD (NL-SVD). With the nonlocal SVD algorithm we can capture structural features from the low frequency subband under different multiscale transform. The statistical features are captured from high frequency subbands. The feature vector of an image is combined with these two kind features. We use the Euclidean distance to measure the similarity among normalized feature vector for texture retrieval. The experiment results are relatively good, which prove that structural features extracted from the low frequency subband are useful and effective for texture retrieval.

This paper is organized as follows. Section II simply introduces the multiscale transforms we applied in our experiments. In section III, we focus on the nonlocal SVD algorithm. Section IV introduces the feature extraction and distance measuring methods in our texture retrieval scheme. Section V describes the experiments and the results are analyzed and compared. General conclusions are written in section VI.

II   MULTISCALE TRANSFORMS

This paper applies multiscale transforms to extract features for texture image retrieval. Multiscale transforms are considered to be an effective time-frequency methods for signal processing. They can analysis signal both in time and frequency domains simultaneously. Most multiscale transform can also discriminate direction information that exists widely in images. So multiscale transform is a effective tool for extracting features from images.

There are a lot of effective multiscale transforms that have been used for various image processing. In this paper we adopt discrete wavelet transform (DWT) [4-6], dual tree complex wavelet transform (DTCWT) [7], Contourlet transform (CT) [8] and shearlet transform (ST) [9] to evaluate the performance of our combing texture features. These transforms are all multiresolutions, time-frequency locality, good energy compactness, high direction selectivity and low redundancy. The shift-invariance and direction selectivity among them are different, and these differences lead to different performances in our experiments.

III NONLOCAL SVD ALGORITHM

In this section, we adopt a nonlocal SVD algorithm for texture retrieval. We first introduce the traditional SVD algorithm, and then we describe the nonlocal SVD algorithm in detail.

Given a matrix A of size $m_1 \times m_2$, there exists a factorization of the form:

$$A = U \times S \times V^T \qquad (1)$$

Where U is an $m_1 \times m_1$ orthonormal matrix, S is a $m_1 \times m_2$ diagonal matrix of singular values, and V is also an $m_2 \times m_2$ orthonormal matrix. The columns of V and the columns of U (respectively called the right and left singular vectors) are, respectively, the eigenvectors of the column-column correlation matrix $A^T A$ and the row-row correlation matrix AA^T. The singular values in S are the square roots of the eigenvalues of $A^T A$ (or AA^T). The singular values of natural images follow an exponential decay rule and the SVD bases have a frequency interpretation [10-12].

The above is the traditional SVD algorithm. The nonlocal SVD algorithm is based on the traditional one [13].

Given a matrix L of size $m \times m$, m tiled matrices L form a 3D matrix B (the size is $m \times m \times m$). First, we find a random matrix A_{rand} of size $m \times m$. Matrix U_{rand} and V_{rand} could be decomposed out by the SVD of A_{rand}. These parameters are defined as follows:

$$A_{rand}(i,j) = U_{rand}(i,j) \times S_{rand}(i,j) \times V_{rand}(i,j)^T \quad (2)$$

Where i, j are members of the ranks of matrices ($0 \le i, j \le m$).

Second, m matrices U_{rand} and V_{rand} are tiled to form 3D matrix U_1 and V_1 respectively, and then the coefficients of U_1 are weighted with a random 3D matrix F for a 3D matrix P. It can be computed via

$$P(i,j,n) = U_1(i,j,n) + 0.01F(i,j,n) \qquad (3)$$

Where $n = 1, 2, \cdots m$ (n is the third dimension). Similarly, a 3D matrix Q can be gotten with the weighted coefficients of V_1. According to the formula (1), we can obtained a 3D matrix $S = P^T B Q$, and S^k contains k larger singular values of $|S|$ ($k < m$). A 3D matrix C is computed via

$$C(i,j,n) = B(i,j,n) - P(i,j,n) \times S^k(i,j,n) \times Q^T(i,j,n) \quad (4)$$

We can get a transversal vector $N(1,n)$ which contains the square of diagonal values of C. The above matrices are used to construct two new 3D matrix $A_1 = \dfrac{B^T P S}{e^{0.012N}}$ and $A_2 = \dfrac{B Q S^T}{e^{0.012N}}$.

At last, the SVD of A_1 and A_2 can make out 3D matrix S_1 and S_2 respectively. S_1 is combined with S_2 for a 3D matrix:

$$S_3(i,j,n) = S_1(i,j,n) \times S_2^T(i,j,n) \qquad (5)$$

S_3 is obtained by the transform of original matrix B. It contains most of the eigenvalues of the matrix, which can effectively represent the feature of matrix.

We want to extracted structural features from low frequency subband, so we apply the nonlocal SVD algorithm in the low frequency subband.

IV FEATURE EXTRACTION AND CLASSIFIER

In this paper, we need to obtain structural features and structural features from images. We adopt structural features as kinds of texture features, combined with traditional statistical features such as means and standard deviations, and apply these features for texture retrieval. The Euclidean distance is used to measure the images' similarity. Here we introduce the feature extraction and classifier in detail.

A. Feature

In our method, three indices are used to represent an image separately. They are means, standard deviations and singular values. The three features are combined with each other to get better retrieval accuracy.

Through the previous introduction, a given texture image is decomposed into subbands by DWT with three levels. We can obtain the low frequency subband named LL. We use the NL-SVD method in the matrix which is composed of the coefficients of LL. We can get a 3D matrix S_3 from each block when the low frequency subband is divided into 1, 4 and 16 no-overlapped blocks respectively. The maximum value of S_3 of each block is the singular value (S_{NL-SVD}) which we need. Now we also need to obtain the means and the standard deviations from the high frequency subbands.

The means are defined as follows:

$$\mu = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} x(i,j) \qquad (6)$$

where $x(i,j)$ is the high frequency subband coefficients, and M , N are the members of row and column of each subband. All the means of three levels can be obtained by the formula (5). The standard deviations are computed via

$$\sigma = \sqrt{\frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [x(i,j) - \mu]^2} \qquad (7)$$

The standard deviations show the distribution of amplitude of the entire or local image approximately. It can be seen that μ represents the average intensity of the subband, and σ reflects the difference of intensity among the pixels inside this subband.

In each image, we form the means (μ), standard deviations (σ) and singular values (S_{NL-SVD}) to a vector and regard it as new image features. This is the way to get higher retrieval accuracy.

B. Classifier

The Euclidean distance is applied in this paper. The distance between two images features is defined as follows:

$$dis\tan ce(i,j)=\left|\frac{\mu_{mn}^{(i)}-\mu_{mn}^{(j)}}{\sigma(\mu)}\right|+\left|\frac{\sigma_{mn}^{(i)}-\sigma_{mn}^{(j)}}{\sigma(\sigma)}\right|+\left|\frac{S_{NL-SVD}^{(i)}-S_{NL-SVD}^{(j)}}{\sigma(S_{NL-SVD})}\right| \quad (8)$$

Where μ_{mn} is the mean and σ_{mn} is the standard deviation of the transform coefficients of subband n at scale m. Because the feature vector contains two kinds of eigenvalues, these eigenvalues are normalized to get better results [14]. $\sigma(\mu)$, $\sigma(\sigma)$ and $\sigma(S_{NL-SVD})$ are applied to normalize these eigenvalues. By calculating the Euclidean distance we can measure the similarity of images for classification.

V EXPERIMENT

The experiments use the texture images of Brodatz databases to test the new method. Each of the texture images (the size is 640×640) is spilt into 16 no-overlapping samples (the size is 160×160). We use 40 classes of Brodatz's texture images to get a dataset with 640 images, then our nonlocal SVD method is applied on the dataset to measure the performance of texture retrieval. We finish the experiments by MATLAB (R2012a).

In our experiments, each texture image in the dataset is decomposed into subbands by DWT, DTCWT, CT and ST with three scales. In DTCWT, due to its peculiar breakdown structure, there are two low-frequency subbands. We need to extract the singular values of each low frequency subband respectively. In DWT, CT and ST, we obtain the singular values from one low frequency subband. And these different features which are mentioned in section IV are combined with one another, for instance, statistical means (μ) combined with standard deviations (σ) with singular value (S_{NL-SVD}), means (μ) combined with singular value (S_{NL-SVD}).We finish these experiments with three situations (We divide the low-frequency subband into N =1, 4 and 16 no-overlapped blocks respectively. Every block is extracted a feature value). With the method we can reduce dimensions and the features of images can be extracted accurately. We also compare with the traditional SVD algorithm [15]. This kind of comparison can help us prove that the structure features of images are effective. The experimental data of texture retrieval are listed in Table I, Table II and Table III. These tables show the retrieval accuracy under different multiscale transforms and feature combinations.

From the above tables we can know that the retrieval ratio of S_{NL-SVD} with μ, σ or $\mu+\sigma$ is better than the one of μ and σ, and the retrieval ratio of S_{NL-SVD} is better than that of S_{SVD}. We can also know that the retrieval ratio of N =16 is better than the one of N =1, 4.

The time consumption of building the two feature vectors (S_{SVD} and S_{NL-SVD}) on our selected database is calculated. The time of S_{SVD} is slightly higher than that of S_{NL-SVD}.

Because of these results, we can know that the nonlocal SVD method is effective and structural features can improve the retrieval ratio in texture retrieval.

VI CONCLUSIONS

In this experiment, we apply the nonlocal SVD algorithm in the low frequency subband. The method describes the importance of extracting structural features. We combine them with statistical features such as means and standard deviations for getting higher retrieval accuracy. The results demonstrate that structural features are effective for texture retrieval.

TABLE I: THE RETRIEVAL ACCURACY WITH N=1.

Transform / Feature	DWT	DTCWT	CT	ST
μ	67.51%	68.75%	68.11%	66.73%
σ	64.53%	65.93%	64.15%	64.04%
S_{SVD}	18.05%	18.56%	18.12%	17.89%
S_{NL-SVD}	19.80%	19.98%	19.75%	19.54%
$S_{NL-SVD}+\mu$	75.78%	76.23%	75.35%	74.98%
$S_{NL-SVD}+\sigma$	78.41%	79.58%	77.93%	77.75%
$S_{NL-SVD}+\mu+\sigma$	73.59%	73.76%	72.97%	72.85%

TABLE II: THE RETRIEVAL ACCURACY WITH N=4.

Transform / Feature	DWT	DTCWT	CT	ST
μ	67.51%	68.75%	68.11%	66.73%
σ	64.53%	65.93%	64.15%	64.04%
S_{SVD}	19.16%	19.23%	19.18%	18.94%
S_{NL-SVD}	20.73%	21.02%	20.56%	20.23%
$S_{NL-SVD}+\mu$	76.77%	76.93%	75.95%	75.17%
$S_{NL-SVD}+\sigma$	78.16%	78.76%	77.43%	76.78%
$S_{NL-SVD}+\mu+\sigma$	74.78%	75.26%	73.89%	73.36%

TABLE III: THE RETRIEVAL ACCURACY WITH N=16.

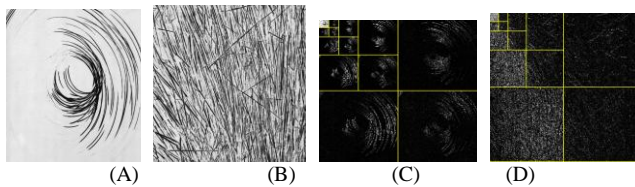| Transform / Feature	DWT	DTCWT	CT	ST
μ	67.51%	68.75%	68.11%	66.73%
σ	64.53%	65.93%	64.15%	64.04%
S_{SVD}	20.85%	21.54%	21.38%	19.98%
S_{NL-SVD}	21.75%	21.86%	21.43%	21.02%
$S_{NL-SVD}+\mu$	77.78%	77.86%	77.12%	76.87%
$S_{NL-SVD}+\sigma$	78.56%	80.03%	77.99%	78.75%
$S_{NL-SVD}+\mu+\sigma$	77.24%	78.98%	75.87%	74.56%

FIGURE I. IMAGES AND THEIR WAVELET SUBBANDS. (A) IMAGE A1, (B) IMAGE A2,(C) FOUR-LEVEL WAVELET DECOMPOSITION OF A1, (D) FOUR-LEVEL WAVELET DECOMPOSITION OF A2.

REFERENCES

[1] Yan Qiu Chen, Mark S.Nixon, David W.Thomas, "Statistical geometrical features for texture classification," Pattern Recognition Letters, Vol. 28, No. 4, pp. 537-522, April 1995.

[2] G.Feng, J.Jiang, "JPEG compressed image retrieval via statistical features," Pattern Recognition Letters, Vol. 36, No.4, pp. 977-985, April 2003.

[3] Trygve R. "Brodatz texture images," http://www.ux.uis.no/tranden/brodatz.html, September 2004.

[4] G.V.Wouwer, P. Scheunders and D.V.Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Trans. Image Process.*, vol. 8, no. 4, pp. 592–598, Apr. 1999.

[5] S.G.Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.11, No.7, pp. 674-693, July 1989.

[6] S.G.Mallat, "Multi-frequency channel decompositions of images and wavelet models," IEEE Transactions on Acoustics, Speech and Signal Processing, Vol.37, No.12, pp. 2091-2110, December 1989.

[7] I.W.Selesnick, R.G.Baraniuk, N.C.Kingsbury, "The dual-tree complex wavelet transform," IEEE Signal Processing Magazine, Vol.22, No.6, pp. 123-151, November 2005.

[8] M. N. Do, M. Vetterli, "Contourlet transform: An efficient directionalmultire solution image representation," IEEE Transaction on Image Processing, Vol.14, No.12, pp. 2091-2106, December 2005.

[9] G.R, Syst. Planning Corp. Arlington, VA , Labate, D., "Shearlet-Based Total Variation Diffusion for Denoising," IEEE Transaction on Image Processing, Vol.18, No.2, pp. 260-268, February 2009.

[10] H.Andrews and C.Patterson, "Singular Value Decompositions and Digital Image Processing," IEEE Trans. Acoustics, Speech, and Signal Processing, vol. 24, no. 1, pp. 26-53, Feb. 1976.

[11] A. Rangarajan, "Learning Matrix Space Image Representations," Proc. Int'l Conf. Energy Minimization Methods in Computer Vision and Pattern Recognition, pp. 153-168, 2001.

[12] C.Ding and J.Ye, "Two-Dimensional Singular Value Decomposition (2DSVD) for 2D Maps and Images, " Proc. SIAM Int'l Conf. Data Mining, pp. 32-43, 2005.

[13] Ajit Rajwade, Anand Rangarajan and Arunava Banerjee, "Image Denoising Using the Higher Order Singular Value Decomposition, "*IEEE Transsactions on Pattern Analysis and machine intelligence,* vol.35, no. 4, April 2013.

[14] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Texture features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[15] Srinivasan Selvan and Srinivasan Ramakrishnan, "SVD-Based Modeling for Image Texture Classification Using Wavelet Transformation," *IEEE Transactions On Image Processing,* vol.16, no. 11, Nov. 2007.