Subband SD And Kurtosis Featured NSST Texture Retrieval

Cheng Wan

College of Electronic and Information Engineering,
Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Abstract
The selection of features extracted for image retrieval is quite important. We in this paper choose standard deviation and kurtosis to character texture. Non-subsampled shearlet transform is used to derive texture features. Shearlet is a new sparse representation tool of multidimensional function, which provides a simple and efficient mathematical framework. We firstly decompose the source images on various scales and in different directions with non-subsampled shearlet transform. The standard deviation and kurtosis of each subband are cascaded as texture features of the images. The similarity measurement of the image retrieval system is achieved with the Average Euclidean Distance. The experiments are executed on four different multidimensional transforms: DWT, Contourlet, NSCT and NSST. The results show the proposed retrieval method is superior to that of the traditional methods.
Key words: NON-SUBSAMPLED SHEARLET TRANSFORM, IMAGE RETRIEVAL, MULTIDIMENSIONAL TRANSFORM, SIMILARITY MEASUREMENT

1. Introduction

Content–based image retrieval (CBIR) is an active research area in computer vision. The destination of image retrieval systems is to find relevant images in the database which are similar to the query image. CBIR aims to develop an efficient Visual Content Based technique to search, browse and retrieve relevant images from large scale digital image collections[1,2]. CBIR adopts feature extraction and similarity measurement for image retrieval process.

Image content features including color, texture and shape are extracted by computer and measured the distance between the query image and dataset images to derive efficient retrieval. In these content features, the texture features are quite effective models and are utilized frequently in CBIR systems [3], since texture can efficiently reflect the structural, directional, granularity, or regularity differences of diverse regions in a visual image.

Multi-resolution analysis tools such as pyramid transform and Discrete Wavelet Transform (DWT) have been used to extract the image texture features [4-6] to realize image retrieval. DWT has many advantages such as localization and direction, and it can well represent the detail information of the image. However, the directional selectivity of DWT is very limited and the it is not shift-invariant. Contourlet transform [7,8] is a new image analysis tool, which is anisotropic and has good directional

selectivity. The sub-sampling operation is employed in the implementation of Contourlet transform. So it is not shift-invariant and sub-band spectrum aliasing phenomenon takes place, which weakens the directional selectivity of Contourlet transform. To overcome this disadvantage, Cunha et al. [9] proposed non-subsampled Contourlet transform (NSCT) which is an improved version of Contourlet transform. But the computational efficiency of NSCT is low. In 2005, Labate et al. [10] proposed a new multi-scale geometric analysis tools: shearlet, which is optimally sparse in representing images. The decomposition of shearlet is similar to Contourlet transform, but an important advantage of shearlet over Contourlet transform is that there are no restrictions on the number of directions for the shearing. In addition, the inverse shearlet transform only requires a summation of the shearing filters rather than inverting a directional filter banks, thus the implementation of shearlet is more efficient computationally. Due to these advantages with the shearlet, it has been applied in many image processing fields, such as image denoising, edge detection, and image fusion. The shearlet transform also uses the subsampling operation, so it is not shift-invariant and it will produce pseudo-Gibbs phenomenon around the singular point. The non-subsampled shearlet transform (NSST) [11] by cascading of non-subsampled pyramid filter banks and shearing filter banks has all the advantages of the shearlet transform and does not need the implementation of the sub-sampling operation.

We in this research adopt non-subsampled shearlet transform (NSST) to decompose the images and compute the standard deviation and kurtosis of NSST subbands to describe the image texture features. The similarity measurement strategy is to make use of Average Euclidean distance [12]. The MIT vision texture database (640 images) and the Brodatz texture database (1280 images) were used to verify the retrieval performance. Experimental results show that the retrieval performance can be improved by using the proposed method comparing with some existing feature extraction and similarity distance measurement methods.

2. Texture feature extraction

2.1. Non-subsampled shearlet transform (NSST)

In dimension n = 2, the affine systems of shearlet transform is defined as follows:

$$\mathbf{M}_{AB}(\psi)=\{\psi_{j,\ell,k}(x)=|\det \mathbf{A}|^{j/2}\,\psi(\mathbf{B}^{\ell}\mathbf{A}^{j}x-k), j,\ell\in\mathbb{Z},k\in\mathbb{Z}^2\} \quad (1)$$

where ψ is a collection of basis function and satisfies $\Psi\in L^2(R^2)$. A denotes the anisotropy matrix for multi-scale partitions, B is a shear matrix for directional analysis. i, j, k are scale, direction and shift parameter respectively. A, B are both 2×2 invertible matrices and $\det|\mathbf{B}|=1$. For each $a>0$ and $b\in\mathbf{R}$, the matrices of A and B are represented as:

$$\mathbf{A}=\begin{pmatrix} a & 0 \\ 0 & \sqrt{a}\end{pmatrix},\ \mathbf{B}=\begin{pmatrix} 1 & b \\ 0 & 1\end{pmatrix} \quad (2)$$

These two matrices are important roles in the process of shearlet transform. The former dominates the scaling of shearlet, and the latter controls the orientation of shearlet. When a = 4, s =1, (2) is written as follows:

$$\mathbf{A}=\begin{pmatrix} 4 & 0 \\ 0 & 2\end{pmatrix},\ \mathbf{B}=\begin{pmatrix} 1 & 1 \\ 0 & 1\end{pmatrix} \quad (3)$$

For $\forall\xi=(\xi_1,\xi_2)\in\widehat{\mathbb{R}}^2$, $\xi_1\neq 0$, $\widehat{\psi}^{(0)}$ for ST can be described as:

$$\widehat{\psi}^{(0)}(\xi)=\widehat{\psi}^{(0)}(\xi_1,\xi_2)=\widehat{\psi}_1(\xi_1)\widehat{\psi}_2(\frac{\xi_2}{\xi_1}) \quad (4)$$

Here, $\widehat{\psi}$ is the Fourier transform of ψ, $\mathrm{supp}\widehat{\psi}_1\subset[-\frac{1}{2},-\frac{1}{16}]\cup[\frac{1}{16},\frac{1}{2}]$, $\mathrm{supp}\widehat{\psi}_2\subset[-1,1]$.

Assume that

$$\sum_{j\geq 0}|\widehat{\psi}_1(2^{-2j}\omega)|^2=1,\quad |\omega|\geq\frac{1}{8} \quad (5)$$

$$\sum_{\ell=-2^j}^{2^j-1}|\widehat{\psi}_2(2^{-2j}\omega-\ell)|=1,\quad |\omega|\leq 1, j\geq 0 \quad (6)$$

We can see $\psi_{j,\ell,k}(x)$ has the frequency support as follows.

$$\mathrm{supp}\widehat{\psi}^{(0)}_{j,\ell,k}\subset\{(\xi_1,\xi_2):\xi_1\in[-2^{2j-1},-2^{2j-4}]\cup[2^{2j-4},2^{2j-1}],|\frac{\xi_2}{\xi_1}+\ell 2^{-j}|\leq 2^{-j}\} \quad (7)$$

Hence, $\widehat{\psi}_{j,\ell,k}$ is supported on a pair of trapeziform zones, whose sizes all approximate to $2^{2j}\times 2^{j}$.

The shearlet transform has the following main properties: well localizing, parabolic scaling, highly directional sensitivity, spatially localizing and optimally sparse. The nonsubsampled shearlet transform (NSST), which combined the non-subsampled Laplacian pyramid transform with several different combinations of the shearing filters, is the shiftinvariant version of the shearlet transform. The NSST differs from the shearlet transform in that the NSST eliminates the down-samplers and up-samplers. The NSST is a fully shiftinvariant, multi-scale, and multi-direction expansion. Consequently, introduction of NSST into image retrieval could make use of the good characters of NSST in effectively extracting features from original images.

2.2. Texture feature extraction by NSST

In this paper, we compute the standard deviation σ and kurtosis K of each NSST subband to build texture feature vector. The mean μ and standard deviation σ standard deviation are computed as follows.

$$\mu_{s,d} = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|W_{s,d}(m,n)\right| \quad (8)$$

$$\sigma_{s,d} = \left[\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|\left|W_{s,d}(m,n)\right| - \mu_{s,d}\right|^2\right]^{1/2} \quad (9)$$

Here, s and d are the scale and direction of the subband respectively. M and N are the number of the rows and columns of the subband coefficient matrix.

Kurtosis K is another vital texture feature which is defined as follows.

$$K_{s,d} = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\frac{\left|\left|W_{s,d}(m,n)\right| - \mu_{s,d}\right|^4}{{\sigma_{s,d}}^4} \quad (10)$$

Assume NSST gives n subbands, we can derive 2n parameters of texture feature. The texture feature vector $\mathbf{V}_t$ is described as:

$$\mathbf{V}_t = [\sigma_1, K_1, \sigma_2, K_2, \cdots \sigma_n, K_n] \quad (11)$$

Here σ_i and K_i are standard deviation and kurtosis of ith subband respectively.

2.3. Distance between texture vectors

The similarity measurement on texture features between two images can be figured out by the statistical features of NSST subbands. The texture feature distance between the query image and the database image is calculated by Average Euclidean distance (AED), which is represented as follows.

$$D(\mathbf{V}_t, \mathbf{V}_t') = \sqrt{\sum_{i=1}^{c}\left(\frac{V_{ti} - V_{ti}'}{\left|V_{ti}\right| + \left|V_{ti}'\right|}\right)^2} \quad (12)$$

where $\mathbf{V}_t$ is the texture feature vector of the query image, $\mathbf{V}_t'$ is the texture feature vector of the image in the database, and c is the number of vector elements.

3. Computational results and comparisons

As demonstrated above, we proposed a texture-based retrieval system by extracting the texture features with NSST, and measuring Average Euclidean distance. The retrieval task is to search the top N images that are similar to one query image within a large database of total M unlabeled images. In our retrieval system, each image is decomposed into multiple directional subbands. Then, we compute the statistical features $\{\sigma, K\}$ of each subband, which are the extracted image texture features. Average Euclidean Distance between the query image and each database image is measured. The top N database images that have the smallest Average Euclidean distance

Table 1. Overall recognition rate (%) with 6 features and 2 distances across 4 different transforms in Vistex database

Transform	Distance			Feature			
		$\{\mu\}$	$\{\sigma\}$	$\{K\}$	$\{\mu, \sigma\}$	$\{\mu, K\}$	$\{\sigma, K\}$
DWT	ED	70.332	69.1211	33.0859	73.3887	65.1758	67.7051
	AED	73.8574	72.7734	36.9824	76.875	76.875	77.8516
Contourlet	ED	74.4727	71.9824	32.5391	75.2344	65.8789	67.207
	AED	76.8848	74.7559	41.1426	78.1543	77.6758	77.5977
NSCT	ED	74.3164	73.7012	41.8359	76.9141	69.3555	71.6992
	AED	77.0801	75.4297	46.416	79.0527	79.0527	79.9023
NSST	ED	76.2305	75.2051	41.1426	77.8418	70.3027	72.6758
	AED	79.0039	77.3730	46.9629	80.3613	80.8301	81.4355

Table 2. Overall recognition rate (%) with 6 features and 2 distances across 4 adifferent transforms in Brodatz database

Transform	Distance			Feature			
		$\{\mu\}$	$\{\sigma\}$	$\{K\}$	$\{\mu, \sigma\}$	$\{\mu, K\}$	$\{\sigma, K\}$
DWT	ED	71.3867	71.416	42.7295	78.2813	73.1738	77.0117
	AED	73.4961	71.958	48.8379	79.3945	78.4473	79.7754
Contourlet	ED	73.8867	73.75	44.1748	78.5303	74.4873	77.085

	AED	75.249	74.4775	53.0908	79.5508	80.7129	81.3867
NSCT	ED	75.3223	75.4492	59.4482	80.7129	79.2627	81.2598
	AED	76.8945	76.2012	63.2617	81.7383	83.8037	84.1113
NSST	ED	76.6260	76.8164	55.7129	81.9531	79.7461	81.6699
	AED	78.5010	77.9248	61.7578	83.0273	84.6094	85.0098

are retrieved.

To evaluate the performance of the proposed retrieval system, we use two different databases.

a. VisTex database. 640 texture images from 40 classes in Massachusetts Institute of Technology (MIT) Vision Texture database are constructed by dividing each 512×512 class image.

Table 3. Overall recognition rate (%) with 6 features and 2 distances based on different decomposition levels of NSST in Vistex database

Scale Direction	Distance			Feature			
		$\{\mu\}$	$\{\sigma\}$	$\{K\}$	$\{\mu, \sigma\}$	$\{\mu, K\}$	$\{\sigma, K\}$
1 [6]	ED	59.1602	58.8867	33.2715	66.0742	60.0879	62.7344
	AED	60.9082	61.2305	34.7559	68.4375	68.4082	69.2676
2 [6 6]	ED	70.8008	69.7754	39.1211	74.1992	66.3477	69.1504
	AED	73.9160	72.4316	43.0957	77.2266	76.679	78.2129
3 [6 6 10]	ED	76.2305	75.2051	41.1426	77.8418	70.3027	72.6758
	AED	79.0039	77.3730	46.9629	80.3613	80.8301	81.4355

Table 4. Overall recognition rate (%) with 6 features and 2 distances based on different decomposition levels of NSST in Brodatz database

Scale Direction	Distance			Feature			
		$\{\mu\}$	$\{\sigma\}$	$\{K\}$	$\{\mu, \sigma\}$	$\{\mu, K\}$	$\{\sigma, K\}$
1 [6]	ED	58.3594	60.3174	41.709	68.8770	67.8076	71.1035
	AED	60.7959	62.4316	43.3691	70.9570	72.5	74.5801
2 [6 6]	ED	71.5381	71.7188	54.0234	78.0322	76.665	79.3604
	AED	73.9160	73.3301	57.0068	79.7949	81.2354	82.2949
3 [6 6 10]	ED	76.6260	76.8164	55.7129	81.9531	79.7461	81.6699
	AED	78.5010	77.9248	61.7578	83.0273	84.6094	85.0098

into 16 non-overlapping 128×128 subimages. In retrieval experiments, the query image is taken randomly from 640 subimages, and relevant candidate images are the other 15 subimages from the same class.

b. Brodatz database. 1280 texture images from 80 different classes are taken from the Brodatz texture database. Each original 640×640 Brodatz texture image is divided into sixteen nonoverlapping 160×160 subimages. The total number of images from this database is 1280 (80×16). The query image is any one of the 1280 subimages; the other 15 subimages from the same class are relevant candidate images.

We show three sets of experiments and study the results. First, we compare the variance- kurtosis-based method across five transforms: DWT, Contourlet, NSCT, DT-CWT and NSST. Secondly, we change the decomposition levels of NSST to evaluate the retrieval rate with the method mentioned above. Thirdly, we show the image retrieval output interfaces derived from the proposed method. All the experiments are evaluated on Vistex database and Brodatz database. Meanwhile, each method is executed with two similarity measurements: Euclidean distance (ED) and Average Euclidean distance (AED).

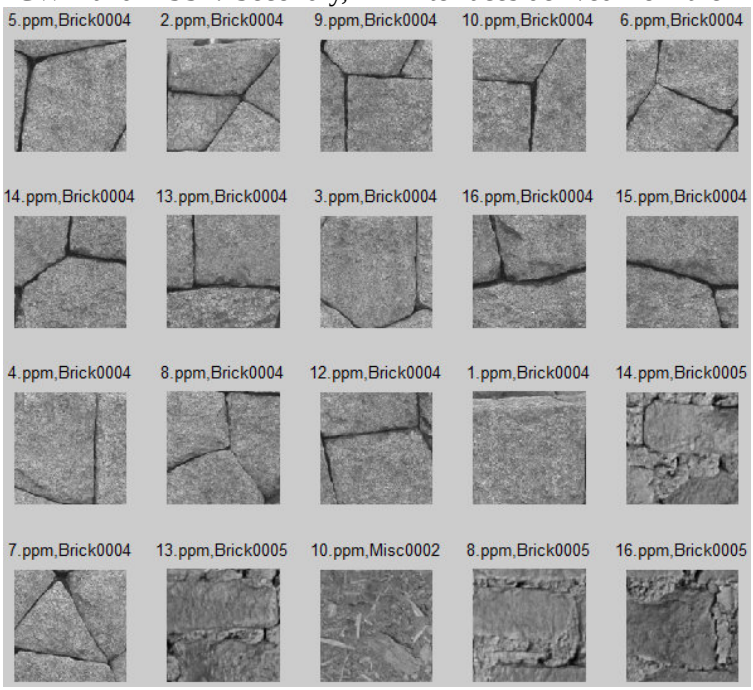

Figure 1. Retrieval result with proposed method. Vistex, Brick0004, 5.ppm, Precision = 0.75, Recall = 0.9375.

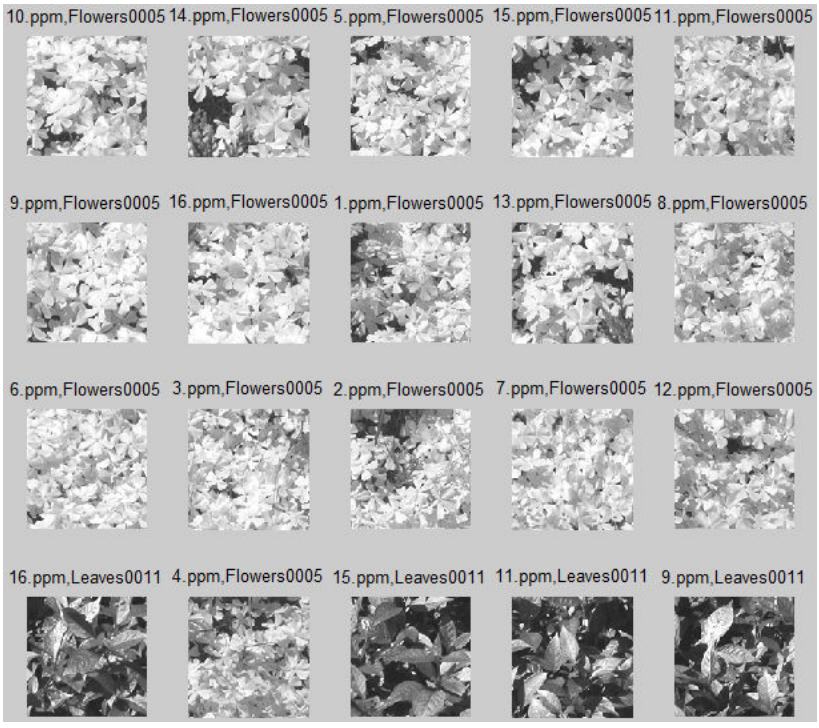

Figure 2. Retrieval result with proposed method. Vistex, Flowers0005, 10.ppm, Precision = 0.8, Recall = 1

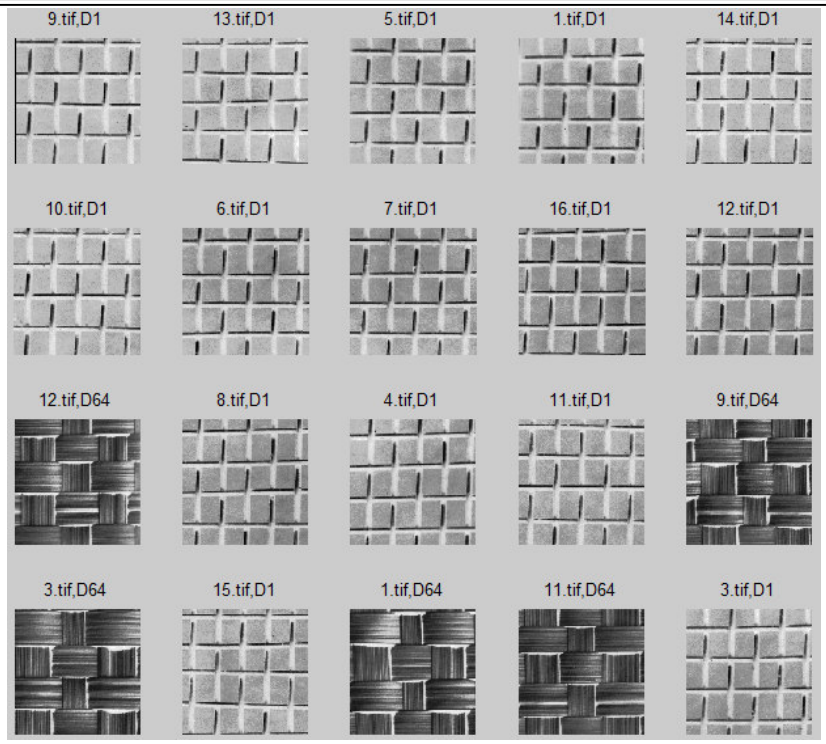

Figure 3. Retrieval result with proposed method. Brodatz, D1, 9.tif, Precision = 0.75, Recall = 0.9375

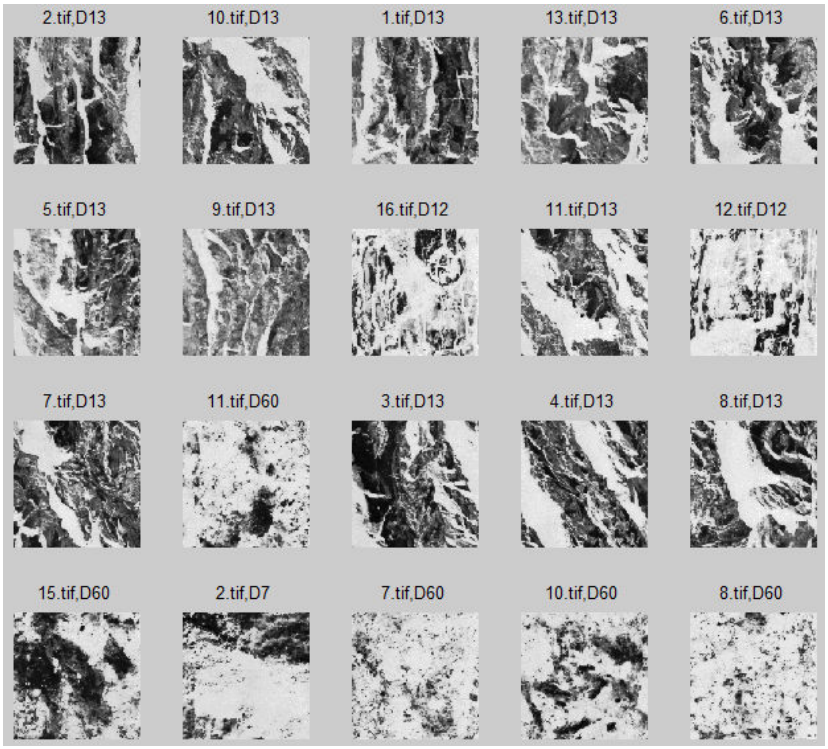

Figure 4. Retrieval result with proposed method. Brodatz, D13,2.tif, Precision = 0.6, Recall = 0.75

3.1. Results based on four transforms

In this experiment, the images are all decomposed into 3 levels by DWT, Contourlet, NSCT, DT-CWT and NSST. For the proposed method, 6, 6, 10 directions are decomposed in the scales from coarser to finer for NSST, and the pyramid filter of NSST is set as ‘maxflat’.

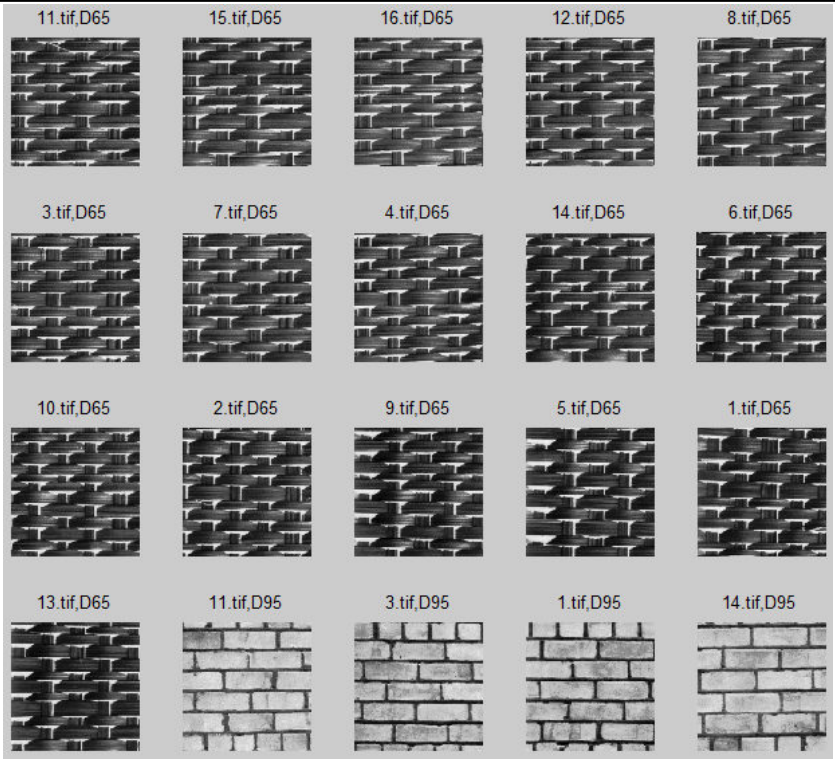

Figure 5. Retrieval result with proposed method. Brodatz, D65, 11.tif, Precision = 0.8, Recall = 1.0

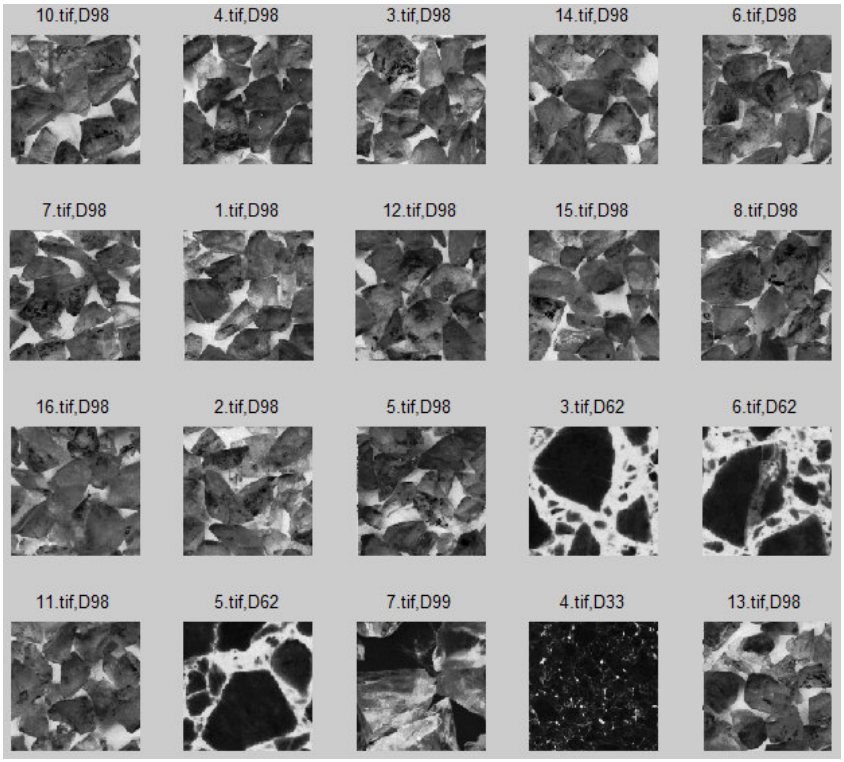

Figure 6. Retrieval result with proposed method. Brodatz, D98, 10.tif, Precision = 0.75, Recall = 0.9375

As shown in Table 1, we computed the overall recognition rate with three methods on five kinds of multi-scale transforms on Vitex database. Overall recognition rate denotes the *average recall* for all the images in the database. *recall* refers to the ratio of the number of relevant image retrieved to total number of relevant images in the database. Here, total number of relevant images in the database equals to 16, and the output is 16 retireve images. $\{\mu\}$, $\{\sigma\}$ and $\{K\}$ denote only the mean, standard deviation and kurtosis of each NSST subband are extracted respectively to measure the similarities by ED and AED. Table 2 shows the same methods executing on Brodatz database.

From Table 1 and Table 2, we can see that: Two feature methods are better than one feature methods, since more features bring more information of images, which improves the efficiency of retrieval system. And, AED measurement is better than ED. As shown on the last row of Table 1 and Table 2, methods based on NSST with AED show the best results comparing with all the other methods. That demonstrates the good performance of NSST with AED measurement in comparison with other transforms. And the proposed method could achieve the highest recognition rate.

3.2. Results from different decomposition levels of NSST

We also evaluate the performance of proposed method based on different decomposition levels of NSST. Table 3 and Table 4 show the overall recognition rate with three methods on NSST of one scale with 6 directions, 2 scales with 6 directions respectively and three scales with 6, 6, 10 directions on Vitex database and Brodatz database respectively. NSST of three decomposition levels with AED gives the best results than that of one and two decomposition levels. The proposed method still derives best results comparing with the other methods.

3.3. Image retrieval system outputs

Six retrieval system outputs of proposed method are shown in Figure 1 to Figure 6. Each output result is based on one retrieval. The retrieval results are computed from the texture images based on the VisTex and Brodatz database. In each output result, the query image is on the top left corner and all other images are ranked in the order of similarity with the query image from left to right, top to bottom. The query images are chosen randomly from following classes: Brick0004 (Vistex), Flowers0005 (Vistex), D1 (Brodatz), D13 (Brodatz) D65 (Brodatz) and D98 (Brodatz), respectively. To each query, *Precision* and *Recall* are calculated. *Precision* is the ratio of the number of relevant images retrieved to the total number of images retrieved, where the total number of images retrieved in this experiment is 20 (While the former experiments output 16 retrieved images). *Precision* and *Recall* of each query are quite different, that is due to the retrieval method, the selection of database and so on. In general, 68.75%~100% relevant images can be retrieved correctly with the proposed method.

4. Conclusions

In this paper, we present a new texture image retrieval method. The image firstly is decomposed by non-subsampled shearlet transform. The standard deviation and kurtosis of NSST subbands are facilitated to represent texture features. The Average Euclidean Distances is used to measure the similarity between the query image and the database images. The experiments based on six methods, four different kinds of transforms and two distances show the effectiveness of the proposed method, which can achieve higher overall recognition rate than the other methods.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities of China (56XAA14043); Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

1. Singhai N., Shandilya S.K. (2010) A survey on: Content based image retrieval systems. *International Journal of Computer Applications*, 7(2), p.p. 22-26.
2. Gonzalez D. I., Hormigos C.E.B., Maria F.D.D. (2014) A generative model for concurrent image retrieval and ROI segmentation. *IEEE Transactions on Multimedia*, 16(1), p.p. 169-183.
3. Dai D. X., Riemenschneider H., Gool L.V. (2014) The Synthesizability of Texture Examples. *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, p.p. 3027-3034.
4. Do M. N., Vetterli M. (2002) Wavelet–based texture retrieval using generalized gaussian density and kullback-leibler distance. *IEEE Transactions on Image Processing*, 11(2), p.p. 146-158.
5. Singha M., Hemachandran K., Paul A. (2012) Content–based image retrieval using the combination of the fast wavelet transformation and the colour histogram. *IET Image Processing*, 6(9), p.p. 1221-1226.
6. Atto A. M., Berthoumieu Y., Bolon P. (2013) 2D wavelet packet spectrum for texture analysis. *IEEE Transactions on Image Processing*, 22(22), p.p. 2495-2500.
7. Do M. N., Vetterli M. (2005) The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 14(12), p.p. 2091-2106.
8. Qu H., Peng Y., Sun W. (2007) Texture Image Retrieval Based on Contourlet Coefficient Modeling with Generalized Gaussian Distribution. *ISICA 2007*, Heidelberg, p.p. 493-502.
9. Cunha A. L., Zhou J. P. (2006) The nonsubsampled contourlet transform: theory, design and application. *IEEE Transactions on Image Processing*, 15(10), p.p. 3089-3101.
10. Labate G.K.D., Lim W. (2005) Sparse

multidimensional representation using shearlets. *SPIE Proc*, Bellingham, p.p. 254-262.
11. Easley W.L.G., Labate D. (2008) Sparse directional image representations using the discrete shearlet transform. *Applied and Computational Harmonic Analysis*, 25(1), p.p. 25-46.
12. Zhu Z., Zhao C., Hou Y. (2012) Research on Similarity Measurement for Texture Image Retrieval. *Plos One*, 7(9), p.p.1-14.