Non-Parameter Local Transformation of Low Frequency Wavelet Coefficients Applied in Aerial Texture Retrieval

Xubing Zhang [1]
[1]School of Computer, Wuhan Textile University, Wuhan, China
E-mail:zxbwf001@163.com

Bo Cai [2*]
[2]Computer School of Wuhan University, Wuhan, China
bo_cai@yeah.net

***Abstract*—Low frequency wavelet coefficients include much important vision information of the image, and are useful to image recognition and understanding. While at present, the applications of the low frequency wavelet coefficients are limited in the researches of image analysis. In this paper, the authors extracted the BFV (Binary Feature Vector, BFV) and TFV (Ternary Feature Vector, TFV) of low frequency wavelet coefficients based on non-parameter local transformation, which adopts the comparison results of the coefficient amplitudes in the neighborhood with the center coefficient to extract the feature rapidly. The TFV describes the texture more accuracy than BFV by adopting the two adaptive thresholds, and the by adjusting the parameter *f* TFV can adapt itself to the different texture data. The authors apply the BFV and TFV in aerial textures retrieval. In the experiments, our method is compared with the GLCM, Markov and Fractal algorithms, and the results prove that our methods behave well in the retrieval rate, especially the rapid processing speed.**

***Index Terms*—Non-parameter local transformation, low frequency, wavelet coefficient, texture retrieval**

I. INTRODUCTION

When the image is decomposed with the wavelet transform, the low frequency sub-band, which involves much important visual information of the image, is very useful to image recognition and understanding. Many researchers have used the low frequency wavelet data in various applications of image information processing, such as Fei Yuanyuan extracted the gray level concurrence matrix of the low frequency wavelet coefficients, and then combined with the high frequency feature to retrieve the texture images [1]. Cai used the low frequency coefficients of the Harr wavelet to detect the position of the license plate roughly, which can save the dynamic memory, decrease the complexity and improve the efficient of the detection, and further recognize the license plate accurately by the character recognition system [2]. He Jiazhong presented a face recognition method by combining the train samples with the reconstruct image of low frequency subband, whose performance is better than the (PC) ^{2}A and SPCA methods [3]. Li yan applied the difference of low frequency wavelet coefficients in the texture classification of remote sensing images [4].

Although the low frequency information is used in so many occasions, is usually used as an assistant measure, and the high frequency wavelet information or the other spatial domain method is combined with it to calculate the further processing results.

In this paper, the low frequency wavelet feature vectors are extracted base on non-parameter transformation and applied in the image retrieval as an independently approach. Except the discussions abovementioned, the follow researches would support our intention.

Firstly, the low frequency sub image can hold the main information of the images, and it is robust to the noise. For example, the same view's images which gained by different sensors, usually have the unlike high frequency information while the similar low frequency components [5]. Wang ronghui presented that the background of the image is mainly consist of the low frequency components, although whose shape is arbitrary [6]. Especially, the coherent noises are mainly distributed in the high frequency part, and the wavelet decomposition has the fine time-frequency characteristic, so that we can extract the image features by analyzing the low frequency information [7]. Moreover, Bow indicated that low frequency component can well reconstruct the face image [8].

Secondly, besides the important vision information, the low frequency can carry a certain amount of texture information when the image is decomposed by appropriate levels of wavelet transformation. Such as, He Jiazhong presented that, the low frequency sub image could well hold the features of the face, such as the contour of the hair and face, the positions of the eyes, nose and mouth [3]. Li yan indicated that, the low frequency component showed more self-similar features of the image, when extracting the texture feature of the remote sensing image based on the self-similar model of low frequency components [4].

* Contact author: bo_cai@yeah.net

doi:10.4304/jsw.6.7.1313-1320

Finally, Daidi. Zhong extracted the binary and ternary feature vector of the DC coefficients in image DCT (Discrete Cosine Transform) domain by non-parameter local transformation, and the results of the face recognition experiments proved that his method is efficient [9]. Furthermore, Au K. M. indicates that the characteristic of the DC coefficients is very similar to the low frequency wavelet coefficients [10]. So we can infer that the feature vector of low frequency wavelet coefficients could be extracted and used in some applications of image information processing independently.

The aerial images usually have the larger size, and the amount of the aerial image date is growing up rapidly. They are usually compressed before storage. Especially, a larger number of images are compressed based on the wavelet transformation because of the excellent multi-scale and time-frequency analysis performance. If we could extract the feature of the low frequency wavelet to retrieve the image in the massive data, which don't need to completely decode the compressed images, the speed and the efficient of the image retrieval will certainly be improved.

According to the above discussion, the authors apply the feature vector of low frequency wavelet domain in the aerial image retrieval. In the experiments, the Gray Level Concurrence Matrix (GLCM), Markov random fields (Markov), and fractal dimensions (Fractal) are compared with our method, and the experimental results demonstrated that this method is feasible and effective.

II. CHARACTRISTIC OF THE LOW FREQUENCY WAVELET COEFFICIENTS

The energies are distributed on the 3k+1 sub-bands when an image is decomposed by k level DWT (Discrete Wavelet Transformation), and among the 3k+1 sub-bands, the LL is low frequency sub-band and others are high frequency sub-bands. Table Ⅰ shows the energy percentage of LL sub-band when the standard test images are decomposed by the 3, 4, 5, 6 level CDF 9/7 biorthogonal wavelet and Daubechies 4 wavelet..

TABLE I. THE ENERGY PERCENTAGE OF WAVELET LL SUB-BAND

Wavelet	Daubechies 4				CDF97			
levle	3	4	5	6	3	4	5	6
lena	95.2	92.1	87.4	83.1	97.1	93.7	88.7	83.1
camera	96.2	94.4	93.0	93.7	97.4	96.0	94.6	93.7
peppers	96.7	94.1	89.5	86.7	98.0	95.4	92.2	86.7
bridge	95.4	93.7	92.0	91.8	96.5	95.1	93.0	91.8
boat	97.5	95.6	94.6	93.8	98.0	96.8	94.4	93.8
goldhill	95.5	93.5	91.5	92.9	97.2	95.7	94.0	92.9
barbara	95.3	92.8	89.6	87.8	96.4	94.1	92.0	87.8
baboo	94.7	93.5	91.5	89.8	95.1	94.3	93.2	89.8
Average	95.8	93.7	91.1	90.0	97.0	95.1	92.7	90.0
Standard deviation	0.90	1.1	2.24	3.85	0.96	1.07	1.90	3.85

We can draw the conclusion from table Ⅰ that, Most energies of the image are aggregated on the LL sub-band. There are still 90% percent energies of image in the low frequency coefficients, even the area of LL sub-band is 1/4096 of the image after decomposed by the 6 levels of DWT. So the LL coefficients are very important to the visual effect of the images

III. NON-PARAMETER LOCAL TRANSFORMATION

Non-parameter local transformation is kind of a region correspondence algorithm of computer vision, while it is different from the general region correspondence algorithms [11]. According to the usual algorithms, the statistics of the local regions pixels' gray values are used to match the two corresponding objects, while in terms of the non-parameter local transformation, the feature of the local region is described by comparing the every neighborhood pixel's gray value with the center pixel, then the results of comparison is used to analyze the corresponding relationship between the objects in two images.

Matching algorithms which usually based on standard statistical methods are suited to a single population in the images. Parametric measures, such as the mean or variance, do not behave well in the presence of distinct subpopulations, each with its own coherent parameters. This problem, which will usually refer to as factionalism, is a major issue in computer vision [11].

The fundamental ideal behind the non-parameter local transformation is to define a local image transform that tolerates factionalism. Non-parametric statistics is distinguished by the use of ordering information among data, rather than the data values themselves. There mainly two types of non-parameter transformation, rank transformation and census transformation.

A. Rank Transformation

Let P be a pixel, $I(P)$ its intensity (usually an 8-bit integer), and $N(P)$ the set of pixels in some square neighborhood of diameter d surrounding P. All nonparametric transforms depend upon the comparative intensities of P versus the pixels in the neighborhood $N(P)$. The transforms we will discuss only depend on the sign of the comparison. Define $\varepsilon(p,p')$ to be "1" if $I(p') < I(p)$ and "0" otherwise. The non-parametric local transforms depend solely on the set of pixel comparisons, which is the set of ordered pairs.

The rank transform is defined as the number of pixels in the local region whose intensity is less than the intensity of the center pixel. Formally, the rank transform $R(P)$ is defined as

$$R(P) = \sum_{P' \in W(P,d)} \varepsilon(P,P') \tag{1}$$

Where W is a square region whose center is P. P' is the any pixel in the W region except the pixel P, and d is the diameter of the region W. Note that $R(P)$ is not an intensity at all, but rather an integer in the range of $0 \sim ((2d+1)^2 - 1)$. This distinguishes the rank transform from other attempts to use non-parametric measures such as median filters, mode filters or rank

filters. For the example as the neighborhood region in (2), $R(121) = 6$.

$$\begin{matrix} 98 & 67 & 128 \\ 24 & 121 & 106 \\ 57 & 226 & 49 \end{matrix} \tag{2}$$

B. Census Transformation

The census transform $C(P)$ maps the local neighborhood surrounding a center pixel P to a bit string representing the set of neighboring pixels whose intensity is less than that of P. let $\otimes$ denote the concatenation of binary "0" or "1". The census transform can then be specified as follows,

$$C(P) = \bigotimes_{P' \in W(P,d)} \varepsilon(p, p') \tag{3}$$

For the example as (2), the result of census transformation is $C(121) = 11011101$.

Two pixels of census transformed images are compared for similarity using the Hamming distance, which means the number of bits that differ in the two bit strings, such as (4).

$$Ham(00101110, 01100110) = 2 \tag{4}$$

When matching images with the census transformation, a pixel's corresponding pixel image in the matching is the point which has the least hamming distance with it. Actually, the census transformation includes more information than rank transformation, we can know the comparison order of every neighborhood pixel with the center pixel from the census transformation results, while just the amount of the pixels which are smaller than the center from rank transformation. We also can calculate the rank transformation by analyzing the results of the census transformation, on the contrary, can not.

C. Charactristic of non-parameter transformation

These local transforms rely solely upon the set of comparisons, and are therefore invariant under changes in gain or bias. The tolerance of these transforms for factionalism also results from their reliance upon the comparisons. If a minority of pixels in a local neighborhood has a very different intensity distribution than the majority, only comparisons involving a member of the minority are affected. Such pixels do not make a contribution proportional to their intensity, but proportional to their number. This limited dependence on the minority's intensity values is a major distinction between non-parameter transformation and parametric measures. To illustrate the manner in which these transforms tolerate factionalism, consider a three-by-three region of an image whose intensities are

$$\begin{matrix} 127 & 127 & 129 \\ 126 & 128 & 129 \\ 127 & 131 & A \end{matrix} \tag{5}$$

For some value $0 \le A \le 256$, consider the effect on various parametric and non-parametric measures, computed at the center of this region, as A varies over its 256 possible values. The mean gray value of this region varies from 114 to 142, while the variance ranges from 2 to 1823. These parametric measures exhibit continuous variation over a substantial range as A changes.

Non-parametric transforms are more stable, however. All the elements of the comparison except one will remain fixed as A changes. The comparison result will be

$$\begin{matrix} 1 & 1 & 0 \\ 1 & & 0 \\ 1 & 0 & \alpha \end{matrix} \tag{6}$$

where α is 1 if A < 128, and otherwise 0. The result of census transform is as follows

$$C(128) = \{1,1,0,1,0,1,0,\alpha\} \tag{7}$$

The rank transform will give the result

$$R(128) = \begin{cases} 5 & A < 128 \\ 4 & A \ge 128 \end{cases} \tag{8}$$

This comparison shows the tolerance that non-parametric measures have for factionalism. A minority of pixels can have a very different value, but the effect on the rank and census transforms is limited by the size of the minority.

While there are some limitations about the non-parameter local transformation, such as, the comparison result is represented by "0" and "1", which can not express the detail degree of the intensity comparisons. Also, the results of the comparison are strongly depended on the intensity of the center pixel.

Ⅳ. Low Frequency wavelte Feature Vetor Extract Based on Non-parameter Local Transformation

Non-parameter local transformation is kind of a region correspondence algorithm of computer vision, which is different from the general region correspondence algorithms. Parametric measures, such as the mean or variance, do not behave well in the presence of distinct subpopulations, each with its own coherent parameters. This problem, which is referred to as factionalism, is a major issue in computer vision. The fundamental ideal behind the non-parameter local transformation is to define a local image transform that tolerates factionalism. Non-parametric statistics is distinguished by the use of ordering information among data, rather than the data values themselves.

Daidi. Zhong have extracted the feature vector of the DC coefficients in image DCT (Discrete Cosine Transform) domain based on non-parameter local transformation [9]. Au K. M. indicates that the characteristic of the DC coefficients is very similar to the low frequency wavelet coefficients [10]. By deriving expressions for the lowpass filters in both block-based

DCT and DWT, it is found that these two filters are very similar. Au K. M. analyzed the lowpass filters of DCT between with Haar, Daubechies 4, Biorthogonal (9, 3), and Biorthogonal (9, 7) wavelets kernels, and the similarities between them are respectively 1, 0.9831, 0.9522 and 0.9659. Both theoretical and experimental studies showed that features that are common to DCT and DWT low frequency domains can be extracted in Ref. [10].

Based on the foundation of the above researches we applied this measurement in low frequency wavelet domain, and the detail of the binary and ternary feature extraction methods are as follows.

A. Binary Feature Vector

The binary feature vector (Binary Feature Vector, BFV) is expressed by the string of "0" and "1". For each coefficient in the low frequency wavelet sub-band, there are eight coefficients surrounding it. Such a 3×3 coefficients region is used here to generate the BFV. By measuring and thresholding the magnitude of differences between the non-central wavelet coefficients and the central wavelet coefficient, a binary vector length 8 is formed. The amplitude of central wavelet coefficient is thus used as a threshold to binarize neighboring coefficients. According to the situation that the amplitude of the non-central coefficient equals to the amplitude of the central coefficient, two different cases are considered here, they are

Binary Feature Vector of Inclusive (BFVI)

$$\underset{C' \in W(C,1)}{BFVI(C',C)} = \begin{cases} 0 & |C'| \le |C| \\ 1 & |C'| > |C| \end{cases} \tag{9}$$

Binary Feature Vector of Exclusive (BFVE)

$$\underset{C' \in W(C,1)}{BFVE(C',C)} = \begin{cases} 0 & |C'| < |C| \\ 1 & |C'| \ge |C| \end{cases} \tag{10}$$

BFVI is inclusive case and BFVE is exclusive case. In (9) and (10), where W is the 8 pixels neighborhood with the center of coefficient C. For example, there is a 3×3 low frequency wavelet coefficients region "S" as follows,

$$\begin{array}{ccccc} 216 & \rightarrow & -179 & \rightarrow & -67 \\ & & & & \downarrow \\ 53 & & 118 & & 118 \\ \uparrow & & & & \downarrow \\ 207 & \leftarrow & -131 & \leftarrow & 98 \end{array} \tag{11}$$

We define the clockwise sequence of comparison which is show in (11). According to the feature extract method, the result of BFVI(s, 118) equals to the binary string (11000110), and the result of BFVE(s, 118) is (11010110). There is a little difference between the results of BFVI and BFVE, because the fourth non-central coefficient is equal to the central coefficient.

B. Ternary Feature Vector

Because BFV only takes the amplitude of the central coefficient as the threshold, it is strongly depended on the amplitude of the central coefficient. Actually, each BFV has the bit length of 8, which gives 256 binary vectors. BFV has a good ability to distinguish different visual features. However, a single threshold makes the variations unequally distributed. Most BFV are concentrated over limited amount of bins. This property is quite harmful for the image retrieval ability. To relieve this problem, a self-adaptive method is used to set the threshold [9]. There are two thresholds are adopted which can samples every comparison result to 3 octaves, and so this feature is called Ternary Feature Vector (Ternary Feature Vector, TFV).

Similar to BFV, TFV is also calculated from a 3×3 coefficient matrix. Within each matrix, assuming the maximal wavelet coefficient amplitude is "Max", the minimal wavelet coefficient amplitude is "Min", the mean amplitude of the coefficient region is "M", the two thresholds are calculated by

$$T_+ = \lfloor M + (Max - Min) \times f \rfloor \tag{12}$$

$$T_- = \lfloor M - (Max - Min) \times f \rfloor \tag{13}$$

Where T_+ and T_- are thresholds, and f is a real number in range of (0,1). The parameter f is adaptive, and we can change the value of f when TFV method is applied to different situations. Usually, f is trained to the optimal by processing the samples of the experiments before applied to the some fields. The same as the BFV method, two different cases are applied here in TFV.

Ternary Feature Vector of Inclusive (TFVI)

$$\underset{C' \in W(C,1)}{TFVI(C',C)} = \begin{cases} 0 & |C'| \le T_- \\ 1 & T_- < |C'| < T_+ \\ 2 & |C'| \ge T_+ \end{cases} \tag{14}$$

Ternary Feature Vector of Exclusive (TFVE)

$$\underset{C' \in W(C,1)}{TFVE(C',C)} = \begin{cases} 0 & |C'| < T_- \\ 1 & T_- \le |C'| \le T_+ \\ 2 & |C'| > T_+ \end{cases} \tag{15}$$

TFVI is inclusive case and TFVE is exclusive case. The ternary feature vector uses "0", "1", "2" to describe the comparison results between the wavelet coefficients and the two thresholds. The comparison sequence is also the clockwise as same as the BFV. For example, to the coefficient region S as (11) when f equals 0.4, T_+ equals to 197 and T_- equals to 66, then the result of TFVI(s, 118) and TFVE(s, 118) both equal to the ternary string (21111120). While for another coefficient region as showed in (16), the result of TFVI(s, 118) equals to (21121020), and TFVE(s, 118) still equals to (21111120). The results of TFVI and TFVE are not the same because

there are two coefficients whose amplitudes are respectively equal to the threshold T_+ and T_-.

$$\begin{array}{ccccc} 216 & \rightarrow & -179 & \rightarrow & -67 \\ & & & & \downarrow \\ 53 & & 118 & & 197 \\ \uparrow & & & & \downarrow \\ 207 & \leftarrow & 66 & \leftarrow & 98 \end{array} \tag{16}$$

V. EXPERIMENT RESULTS ANALYSIS

Our experiments data are selected from the aerial image database of some area in China, and the photograph scale of the aerial image is 1:60000. We select four kinds of textures from 6 aerial images with the width of 14646 pixels and the height of 14871 pixels, and they are residential area, naked land, woodland and river textures. Every kind of texture data consists of 50 pieces of 256×256 pixel texture images. The examples of the four kinds of texture images are showed in Fig. 1.

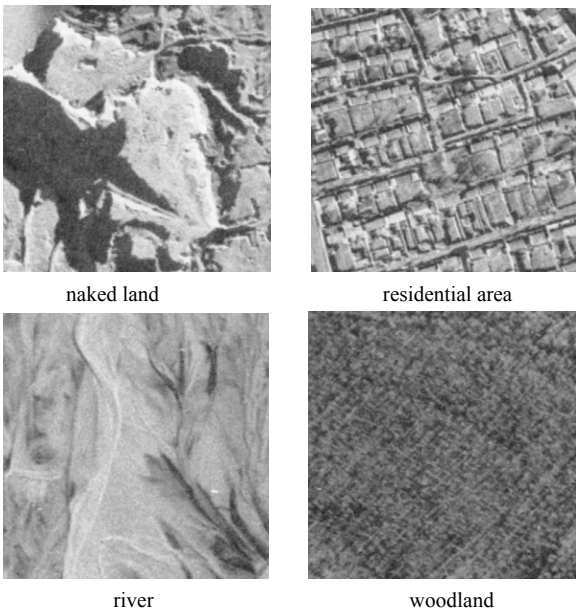

Figure 1. examples of the four kinds of texture images

Those texture feature extraction algorithms, such as gray level concurrence matrix (GLCM) [12], Markov random fields (Markov) [13] and fractal dimensions (Fractal) [14], are compared with BFV (BFVI and BFVE) and TFV (TFVI and TFVE) on the retrieval capability of the aerial texture images in our experiments. The experimental platform is P4 2.6G CPU, 1G DDR2 memory, Windows XP operation system and Matlab 7.0 software.

In this paper, the retrieval ratio are adopted to measure the retrieval capability. Given a queried texture image, we need search the images which have the similar texture with the queried image from the image database which consists of naked land, residential area, woodland and river textures. The amount of the texture images in the database is 200, and every kind of texture has 50 images in it. We define two kinds of the retrieval rates are based on the amount of the retrieval images as follows.

Definition of η_1 is as (17),

$$\eta_1 = n/T,\ T = N/2 \tag{17}$$

Definition of η_2 is as (18),

$$\eta_2 = n/T,\ T = N \tag{18}$$

Where T is the amount of the retrieval images, and N is the amount of a certain class of texture images of the image database. In this paper N equal to 50. n is the amount of the right retrieval images, whose textures belong to the same kind texture with the queried image.

According to the merits of D53 wavelet such as, it has the least computation and the fastest decomposed speed in all of the integer transformation wavelets which have much faster transformation speed than the first generation classical wavelet. The results of the 3 levels D53 wavelet decompose of the texture images are Fig. 2.

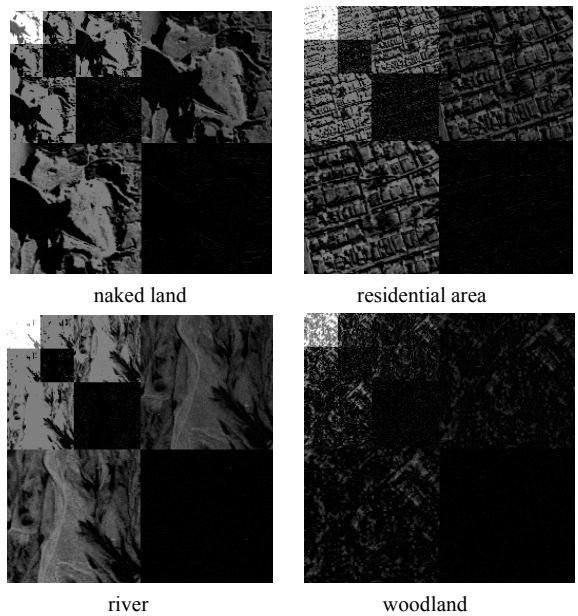

Figure 2. 3 levels of D53 wavelet decompose of the texture images

TABLE II. THE COMPARISON OF THE RETRIEVAL CAPABILITY OF THE EIGHT ALGORITHMS

Texture	Rate %	GLCM	Mar-kov	Fractal	BFVI	BFVE	TFVI	TFVE
Naked land	η_1	92.84	71.6	83.6	74.44	72.12	93.56	93.78
	η_2	83.17	52.8	80.72	64.67	61.45	83.4	83.22
Residen-tial	η_1	87.52	86.25	95.68	78.07	79.52	85.84	86.4
	η_2	78.84	69.03	93.36	67.28	67.96	74.24	74.6
River	η_1	91.68	72.08	65.6	88.64	90.4	87.12	87.36
	η_2	82.15	62.17	63.8	70.6	72.48	74.36	74.88
Wood-land	η_1	84.4	68.15	40.96	82.48	80.8	86.16	85.2
	η_2	68.92	45.26	33.12	71.6	68.28	73.12	72.44

Table II shows the comparison of the retrieval capability of the eight algorithms, and the Euclidean distance of feature vector histogram is adopted here as

the similarity measurement between the queried image and the retrieval images. In GLCM algorithm, Angular Second Moment, Contrast, Correlation and Entropy are used, and the aerial images' 256 levels gray is compressed to 32 levels gray. Markov is the texture feature algorithm based on Conditional Markov Model, Fractal is the feature extraction algorithm based on fractal dimension, and the 3×3 pixel texture analysis window is used in Markov and Fractal algorithms. The images are decomposed by three levels D53 wavelet, and the features of low frequency coefficients are extracted according to BFV and TFV algorithms. The results showed in the Table Ⅱ are the best results with the optimal values of *f* are adopted.

According to the results showed in Table Ⅱ, about the naked land texture, TFV is the best. The naked land retrieval result of GLCM (η_1) is 0.72% lower than the TFVI (η_1), and 0.94% lower than TFVE (η_1), the difference is very little, and the similar as GLCM (η_2). The BFV behaved not well, and is worse than GLCM and Fractal.

For residential area texture, the Fractal is the best, and the retrieval rates η_1 and η_2 of Fractal are both higher than 90%. TFV is close to Markov, and the BFV is the worst.

For river texture, the GLCM is the best, whose retrieval rates η_1 and η_2 are respectively 91.68% and 82.15%. BFV behaved well in the retrieval of river texture, especially for η_1, BFVI (η_1) and BFVE (η_1) are respectively 88.64% and 90.4%. TFV is a little worse than TBV. The results of Markov and Fractal are both unsatisfactory.

For woodland texture retrieval, TFV behaved better than the other methods. GLCM was not bad, for example GLCM (η_1) is just respectively 1.76% and 0.8% lower than TFVI and TFVE. BFV is a little difference with GLCM, and the BFVI (η_1) and BFVE (η_1) both exceed 80%. Fractal is the worst, and the Markov is similar with Fractal.

We can see that, when *f* is optimal, the mean retrieval rate of TFVI (η_1) is 88.17%, TFVI (η_2) is 76.28%, TFVE (η_1) is 88.185%, and TFVE (η_2) is 76.285%. The retrieval capability of TFV algorithm is better than that of Fractal, Markov, BFV algorithms, and is corresponding to GLCM algorithm. The BVF algorithm is better than the Markov and Fractal algorithms. The mean retrieval rate of Markov (η_1) is 74.52%, and Markov (η_2) is 57.315%, which has a poor retrieval capability. The Fractal algorithm is instable, whose retrieval rates are 95.68% (η_1) and 93.36% (η_2) on the residential area texture images but are 40.96% (η_1) and 33.12% (η_2) on the woodland texture images.

When adopting the BFV and TFV methods to index the aerial texture images, the difference of retrieval rates are not great on the different textures. The retrieval rate of river is highest than the other textures by means of BFV, and retrieval rate of naked land is highest than the other textures by means of TFV. The inclusive case and exclusive case are adopted for BFV and TFV methods in this paper, and there is some different retrieval capability between the two cases. Such as that BFVI is better than BFVE on naked land and woodland retrievals, while contrary when adopted in residential area and river textures. The capability of TFVE is little better than TFVI on retrievals of naked land, residential area and river textures. Actually, the difference between the inclusive and exclusive case is related with the characteristics of the texture be retrieved.

According to the retrieval results and the above discussion, we can see that, the BFV and TFV of low frequency wavelet coefficients can distinguish some different aerial image textures when the image is decomposed by 3 levels of wavelet transformation. Which demonstrates that although the low frequency sub-band only include less information, and the abundant texture details and edges are filtered after processing by lowpass, it can hold some useful texture characteristics for image retrieval.

On the control of the parameter *f*, the retrieval capability of TFV is better than BFV. By adopting the adaptive threshold, the TFV doesn't depend on the center coefficient so much as the BFV. TFV compares the coefficient amplitude with two thresholds T_+ and T_-, which can get 3 kinds of results and then is more accuracy quantified than BFV. So that the texture feature that TFV extracted is more reasonable than BFV, which enable the TFV more efficient than the BFV. Especially by adjusting the parameter *f*, the TFV can adapt itself to the different textures, and when the parameter *f* is optimal, the retrieval performance of TFV is fine.

Figure 3. Some retrieval images by used the TFVI algorithm on the river texture image retrieval (f = 0.4)

Some results of the river texture image retrieva1 with the TFVI algorithm are showed in Fig. 3. There are 16 pieces of retrieval images in the Fig. 1, and the first image River27 is the queried image. From the Fig. 3 we

can see that, there are three erroneous return images, Woodland30, Woodland06 and Woodland27. Except these three woodland images the other images are the correct river texture images and they are similar to the queried image.

TABLE III. THE COMPARISON OF THE TFVI RETRIEVAL CAPABILITY UNDER DIFFERENT ADJUSTIVE COEFFICIENT F

Texture		Naked land		Residential area		River		Woodland	
Retrieval rate %		η_1	η_2	η_1	η_2	η_1	η_2	η_1	η_2
f	0.25	79.56	66.0	78.2	67.04	74.72	57.6	84.56	72.84
	0.3	84.22	74.22	85.84	74.24	79.28	64.32	86.16	73.12
	0.35	84.44	72.78	81.84	72.32	82.96	68.64	83.36	71.28
	0.4	82.66	69.0	76.0	75.6	87.12	74.36	80.24	68.96
	0.45	79.56	63.78	65.92	56.84	85.28	73.28	66.96	59.96
	0.5	72.89	58.3	60.56	55.64	83.6	73.96	59.76	56.76
	0.55	83.34	69.67	57.2	50.96	78.32	70.32	58.88	53.64
	0.6	88.67	77.67	54.32	51.44	79.12	71.72	58.24	51.0
	0.65	91.34	84.12	54.96	48.28	80.48	71.88	54.96	51.32
	0.7	93.56	83.4	45.92	45.28	75.04	66.24	52.16	58.64

TABLE IV. THE COMPARISON OF THE TFVE RETRIEVAL CAPABILITY UNDER DIFFERENT ADJUSTIVE COEFFICIENT F

Texture		Naked land		Residential area		River		Woodland	
Retrieval rate %		η_1	η_2	η_1	η_2	η_1	η_2	η_1	η_2
f	0.25	80.0	65.67	78.2	67.04	74.72	57.6	84.56	72.84
	0.3	85.11	74.0	85.84	74.24	79.28	64.32	86.16	73.12
	0.35	84.0	72.67	81.76	72.56	83.2	68.28	83.2	70.92
	0.4	81.78	68.33	77.6	66.28	87.36	74.88	81.52	69.64
	0.45	79.78	63.78	66.24	56.88	85.04	73.32	67.12	60.0
	0.5	74.67	59.22	59.84	55.6	83.2	72.92	60.24	56.28
	0.55	88.89	78.67	56.88	51.04	78.24	70.36	58.96	53.6
	0.6	91.12	84.12	52.24	49.68	78.72	71.32	57.76	51.8
	0.65	93.78	83.22	54.56	48.08	80.8	71.84	54.72	50.96
	0.7	92.0	80.11	46.88	45.04	75.28	66.08	52.16	48.04

TABLE V. THE OPTIMAL COEFFICIENT F OF THE TFV RETRIEVAL ON THE TEXTURE IMAGES

f	Naked land	Residential area	River	Woodland
TFVI	0.7	0.3	0.4	0.3
TFVE	0.65	0.3	0.4	0.3

Table Ⅲ and Ⅳ respectively show the effect of parameter *f* on the retrieval of TFV. Table Ⅴ shows the optimal *f* of TFI and TFVE, that the retrieval results is the best in the 4 kinds of textures.

From the tables we can see that *f* has significant influence on the texture retrieval. The maximum difference of retrieval rate is about 20% on the naked land and river texture, and about 30% on the woodland and residential area texture images retrieval with different values of *f*. When *f* equals to 0.3, the TFV method get the best results in the retrieval of the woodland and residential textures. The optimal *f* is respectively 0.7 and 0.65 when TFVI and TFVE algorithms are used to retrieval the naked land. So that the optimal value of *f* is difference on the retrieval of different texture. We can improve the retrieval result of the TFV by adjusting the value of *f*. The usually approach that how to get the optimal *f* is to select a certain amount of samples to training the parameter *f* on the retrieval experiments with TFV method.

TABLE VI. THE COMPARISON OF THE RETRIEVAL TIME OF THE DIFFERENT ALGORITHMS

algorithm	GLCM	Markov	Fractal	BFV	TFV
Retrieval time (s)	6745.7	22916.5	3315.9	98.344	885.83

Table Ⅵ shows the comparison of the retrieval time on all the 200 texture images by means of the different algorithms. From this table we can see that the BFV and TFV algorithms have the significant advantage on the retrieval speed, especially the BFV algorithm, whose retrieval time is respectively 2.97%, 1.46%, and 0.43% of the retrieval time of Fractal, GLCM, and Markov.

We analyse the reason as follows. Firstly, the non-parameter local transform only need the simple compare operation, and no complex operations needed. Secondly, the BFV and TFV are the features of the low frequency wavelet coefficients. The data amount of LL sub-band is greatly less than that of the initial image, The LL sub-band is 1/64 of the initial image when the image is decomposed by 3 level wavelet transform, that is to say, the processing data amount of BFV and TFV algorithms are greatly less than that of the other algorithms.

Although the TFV exceed the BFV on the retrieval rate, while the retrieval speed of it is smaller than that of the BFV algorithm. Because that the operation of TFV is more complex than that of BFV, such as the threshold calculation. For example, when processing a 32×32 low frequency sub-band, the added computational cost of TFV algorithm is 12288 add and subtract, 3072 multiplication operations, while the BFV does not need these operations.

Ⅵ. CONCLUSION

The authors analyse the characteristics of the low frequency wavelet sub-band, such as it can hold the main vision information of the image, and also can carry a certain amount of texture information. Based on the non-parameter local transformation, the BFV and TFV of the low frequency wavelet coefficients are extracted and used in aerial images retrieval. In our experiments, the four kinds of texture, naked land, residential area, river, and woodland aerial images are retrieved, and the BFV and TFV algorithms are compared with the GLCM, Markov and Fractal algorithms. From the results of retrieval rate and the retrieval time, we can see that, the TFV algorithm has fine performance both on the retrieval rate and retrieval time, the BFV algorithm is the quickest while it's retrieval rate is general, and also the adaptive coefficient *f* has the significant effect on the retrieval capability of the TFV algorithm.

We can draw the conclusion that low frequency wavelet coefficients have a certain amount of texture differentiation capability, and the binary and ternary feature vectors are efficient methods of texture feature extraction of aerial images. Further, by adopting the adaptive thresholds and the parameter *f*, TFV can

describe the texture feature of the local wavelet coefficients more accurate than BFV, and then improve the retrieval capability of aerial textures. While the retrieval speed of TFV is worse than BFV, because of the computational cost brought by the calculation of the adaptive thresholds and more complex comparison operation than BFV. Finally there are some difference of retrieval capability between the inclusive case and the exclusive case, which may be decided by the texture data of the experiments.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of project "The research of ASIP modeling of infrared image simulation under the atmospheric attenuation environment 2008CDB332", and "The Fabric Defect Detection System Based on ASIP".

REFERENCES

[1] Fei. Yuanyuan, Sun Jinguang, and Tao Zhiyong, "Textre image retirval based on wavelet decomposition and gray level co-occurrence matrix," *Modern Computer*, no. 269, pp. 58–59, October 2007.

[2] Ming-xin Cai, "License-plate recognition using wavelet transform and neural network," *Chaoyan university of science and technology, Master thesis*, June 2004.

[3] He Jiazhong, Du Minghui, "Principal component analysis combined with wavelet low-frequency band," *Journal of South China University of Technology (Natural Science Edition),* vol. 35, no. 1, pp. 44–48, January 2007.

[4] Li Yan, Peng Jiaxiong, "Wavelet transform based multiscale hurst parameter texture features and its application," *Acta Electronica Sinica*, vol. 30, no. 7, pp. 1041–1043, July 2007.

[5] Guo Zhiqiang, "Wavelet transform image fusion based on regional features," *Journal of Wuhan University of Technology*, vol. 27, no. 2, pp. 65–71, February 2005.

[6] Wang Ronghui, Liu Gang, "Detecting of targets in natural texture background based on wavelet energy," *Journal of Changchun University of Science and Technology*, vol. 28, no. 3, pp. 70–72, September, 2005.

[7] Wu Yan, "The study on texture extration of SAR image in wavelet domain," *Xidian University, Master thesis*, June, 2007.

[8] Bow, S. T., *Pattern Recognition and Image Preprocessing*. New York, marcel dekker, 1992.

[9] Zhong, D. and I. Defee, "Study of image retrieval based on feature vectors in compressed domain," Signal Processing Symposium, NORSIG 2006, pp. 202–205, 2006.

[10] Au, K. M., L. N. F and S. W. C. "Unified feature analysis in different compressed domains," the 4th International Conference on Information, Communications and Signal Processing, 2003.

[11] Zabih, R. and W. Johns, "Non-parametric local transforms for computing visual correspondence," *Lecture Notes In Computer Science*, vol. 801, pp. 151–158, 1994.

[12] Haralick R M, Shamugam K, "Textural features for image classification", *IEEE Trans. on SMC*, vol. 6, pp. 610-612, March 1973.

[13] Zheng Zhaobao, *Markov Random Field Method of Image Analyses*. Wuhan, Wuhan technical university of surveying and mapping, 2000.

[14] Huang Guilan, Zheng Zhaobao, "The analysis and experiment on aerial image texture classification using three kinds of methods", WTUSM Bulletin of Science and Technology, no. 3, pp. 12-15, 1996.

Xubing Zhang was born in 1977. He received B.S. degree from Department of Automation Engineering, ordnance engineering college; and received M.S. degree from School of Computer, National University of defense technology; and received Ph.D. degree from School of Remote Sensing and Information Engineering, Wuhan University, China. He is major in remote sensing and medical image processing, pattern recognition, image feature extraction, registration and deformation measurement, artificial intelligence.

He had been an assistant professor and a lectuer in National University of Defense Technology, China, from 1999 to 2006. Now, he is an associate professor in School of Computer, Wuhan Textile University, China; and also a post doctoral researcher in Department of Science and Engineering, Ritsumeikan University, Japan. He has published more than 10 research papers in his research fields.

Bo Cai received his Ph.D. in computer application from the Wuhan University. His research interests are in the areas of image processing and video information processing. His research group develops novel analytical methods for video, such as the clustering and similarity algorithm of video shots, browse and retrieval method of video database and text region extraction algorithm in digital videos, vehicles and objects detection algorithm in digital videos.