

Journal of  
**Applied Remote Sensing**

RemoteSensing.SPIEDigitalLibrary.org

## **Improved color texture descriptors for remote sensing image retrieval**

Zhenfeng Shao  
Weixun Zhou  
Lei Zhang  
Jihu Hou

# Improved color texture descriptors for remote sensing image retrieval

Zhenfeng Shao, Weixun Zhou,\* Lei Zhang, and Jihu Hou

Wuhan University, State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan 430079, China

**Abstract.** Texture features are widely used in image retrieval literature. However, conventional texture features are extracted from grayscale images without taking color information into consideration. We present two improved texture descriptors, named color Gabor wavelet texture (CGWT) and color Gabor opponent texture (CGOT), respectively, for the purpose of remote sensing image retrieval. The former consists of unichrome features computed from color channels independently and opponent features computed across different color channels at different scales, while the latter consists of Gabor texture features and opponent features mentioned above. The two representations incorporate discriminative information among color bands, thus describing well the remote sensing images that have multiple objects. Experimental results demonstrate that CGWT yields better performance compared to other state-of-the-art texture features, and CGOT not only improves the retrieval results of some image classes that have unsatisfactory performance using CGWT representation, but also increases the average precision of all queried images further. In addition, a similarity measure function for proposed representation CGOT has been defined to give a convincing evaluation. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JRS.8.083584](https://doi.org/10.1117/1.JRS.8.083584)]

**Keywords:** remote sensing image retrieval; texture feature; color texture; grayscale texture; Gabor wavelets.

Paper 14035 received Jan. 16, 2014; revised manuscript received May 26, 2014; accepted for publication Jun. 27, 2014; published online Jul. 24, 2014.

## 1 Introduction

Texture features have shown significant advantages in the field of image classification,<sup>1</sup> image segmentation,<sup>2</sup> and content-based image retrieval (CBIR),<sup>3,4</sup> etc. Particularly, texture features are one kind of low-level features which have been widely used in CBIR community because of characteristics of independence of image color and intensity.

Some popular textural descriptors, such as gray level co-occurrence matrix (GLCM),<sup>5</sup> Gabor filter,<sup>6</sup> wavelet transform,<sup>7</sup> and local binary pattern (LBP),<sup>8</sup> have been extensively used in CBIR community. Unfortunately, these conventional texture features mentioned above are extracted from grayscale images directly and leave the discriminative information derived from different color channels, which can be regarded as complementary information for different texture patterns, out of consideration. With the intention of fully exploiting the discriminative information to improve the retrieval results of remote sensing images, many studies have been conducted on this topic.

Strategies of such research can be roughly divided into two categories: (1) combination of color and texture features and (2) texture features integrating opponent process theory. Some works based on the former strategy are illustrated as follows. Lin et al.<sup>9</sup> proposed a smart CBIR system based on color and texture features. Chun et al.<sup>10</sup> presented a CBIR method based on a combination of color and texture features extracted in multiresolution wavelet domain. Liapis and Tzirita<sup>11</sup> illustrated a new image retrieval mechanism based on a combination of

---

\*Address all correspondence to: Weixun Zhou, E-mail: [weixunzhou1990@whu.edu.cn](mailto:weixunzhou1990@whu.edu.cn)

texture and color features using discrete wavelet frames analysis and one-dimensional histograms of CIELab chromaticity coordinates, respectively. This strategy has also been accepted as one of the important retrieval mechanisms in some famous image retrieval systems, such as query by image content (QBIC).<sup>12,13</sup> Some other similar works<sup>14–17</sup> could be found in this research. Although these works simultaneously take discriminative information and texture features into consideration, problems such as computational complexity and definition of weight parameters with combinational features are still an open question. In the 1800s, Hurvich and Jameson<sup>18</sup> proposed an opponent process theory of human color vision, and thus texture features integrating opponent process theory have increasingly drawn substantial attention in recent years. Jain and Healey<sup>19</sup> proposed a multiscale representation based on the opponent process theory for texture recognition, and later this method was applied to hyperspectral image texture recognition.<sup>20</sup> In one recent work by Choi et al.,<sup>21</sup> two features, namely color local Gabor wavelets and color LBP, are proposed for the purpose of face recognition, which share similar principles and can be treated as an extensive application of the theory proposed in Ref. 19. The opponent process theory provides complementary information among color channels and generates a simple but effective feature representation.

Motivated by the aforementioned applications of opponent process theory, in this study, we propose one descriptor named color Gabor wavelet texture (CGWT) for remote sensing image retrieval. Meanwhile, color Gabor opponent texture (CGOT) descriptor based on Gabor wavelets has also been presented so as to improve the retrieval results of certain image classes which have inferior precision using CGWT representation.

The rest of this study is organized as follows. Section 2 shows the framework of remote sensing image retrieval based on the proposed descriptors and illustrates the details of the proposed features, parameters used, and similarity measure defined for CGOT descriptor. In Sec. 3, comparative experimental results and discussions are presented. Conclusions and future work constitute Sec. 4.

## 2 Improved Color Texture Descriptors

### 2.1 Framework of Remote Sensing Image Retrieval Based on the Proposed Descriptors

Generally an image retrieval system contains image database, feature database, and some important functional modules, such as feature extraction, indexing mechanism, and similarity measure. Figure 1 illustrates the framework of improved color texture descriptors for remote sensing image retrieval in this study, which mainly contains two parts: feature extraction and image retrieval.

Feature extraction part indicates the extraction procedure of the proposed features. Given one RGB remote sensing image, three dependent color channel images, R, G, and B are obtained first. Then unichrome features corresponding to each color channel are extracted based on Gabor filter with orientation and scale  $(u, v)$ . Finally, R unichrome feature, G unichrome feature, and B unichrome feature are combined together to form unichrome feature. For opponent feature, two Gabor filters with orientation and scale  $(u, v)$  and  $(u, v')$  are used for two color channel images, respectively. As with unichrome feature, RG opponent feature, RB opponent feature, and GB opponent feature are combined together to form opponent features.

Image retrieval part illustrates a simple procedure of remote sensing image retrieval. All images and features are stored in image database and feature database, respectively. Meanwhile, images are associated with the corresponding features through an indexing mechanism. Given one query image, distances between query image and images in database are calculated using a predefined similarity measure, and then the first  $k$  most similar images are returned in ascending or descending order of similarity.

Feature extraction is an important and indispensable part in one image retrieval system. In Sec. 2.2, the details of extraction of proposed representations are illustrated. In addition, as the most important procedure of image retrieval part, similarity measure methods used in this study are discussed in Sec. 2.3 as well.

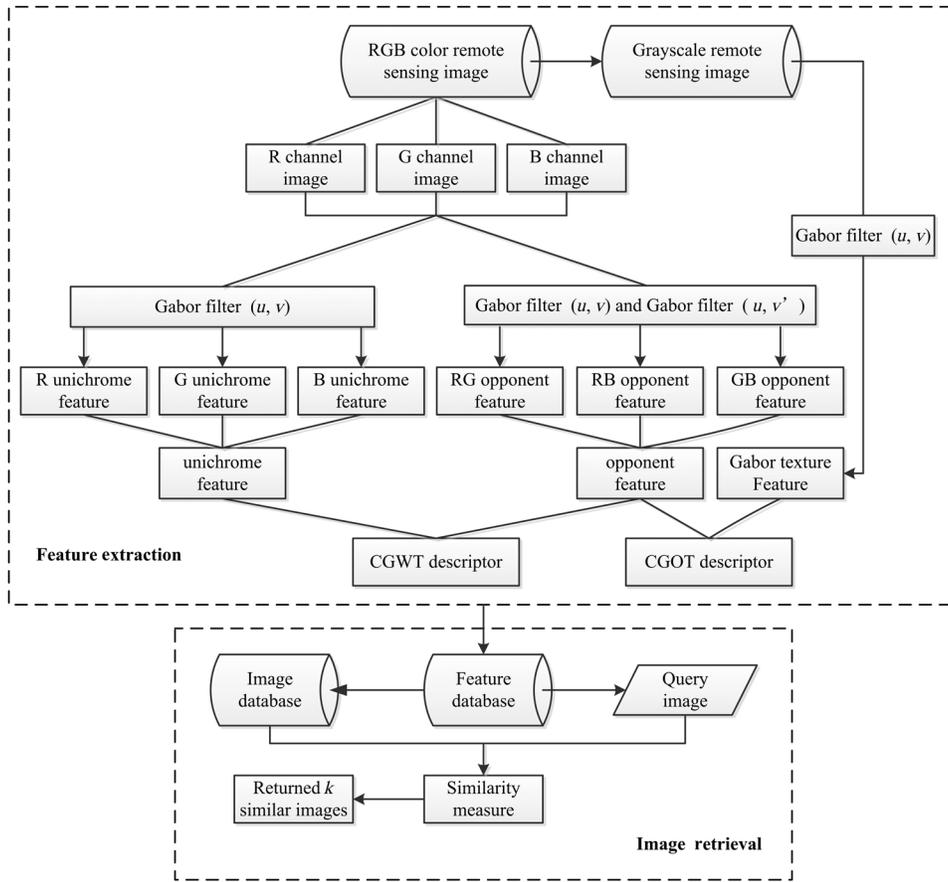


Fig. 1 Architecture of remote sensing image retrieval based on proposed descriptors.

## 2.2 Feature Extraction

In our methodology, all images are represented in RGB color space for convenience. Both CGWT and CGOT features are based on Gabor filter, illustrated as follows:

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{(-\|k_{u,v}\|^2 \|z\|^2 / 2\sigma^2)} [e^{ik_{u,v}z} - e^{-\sigma^2/2}], \quad (1)$$

where  $u$  and  $v$  are mean orientation and scale parameters of Gabor kernels, respectively,  $z = (x, y)$ ,  $\|\bullet\|$  means the norm operator, and  $k_{u,v}$  is defined as follows:

$$k_{u,v} = k_v e^{i\phi_u}, \quad (2)$$

where  $k_v = k_{\max}/f^v$  and  $\phi_u = \pi u/8$ .  $k_{\max}$  is the maximum frequency,  $f$  is the spacing factor, and  $\sigma$  is the standard deviation.

Note that the Gabor filter may have many formula forms, and Eq. (1) in Ref. 22 is chosen because of its conciseness and convenience for setting parameters, such as direction and scale, in our algorithm.

### 2.2.1 Extraction of CGWT descriptor

As illustrated in Fig. 1, CGWT representation consists of two parts, unichrome feature and opponent feature. The terms “unichrome feature” and “opponent feature” follow the definition in Ref. 19, where you can find detailed information about the two features. Let R, G, and B be the three grayscale images of corresponding color channels of an RGB image, respectively. The convolution results of three grayscale images and Gabor kernel  $\psi_{u,v}$  are denoted as follows:

$$\begin{cases} R_{u,v}(z) = R(z) * \psi_{u,v}(z) \\ G_{u,v}(z) = G(z) * \psi_{u,v}(z) \\ B_{u,v}(z) = B(z) * \psi_{u,v}(z) \end{cases} \quad (3)$$

where  $z$  has the same meaning as Eq. (1) and  $*$  means convolution operator.  $R_{u,v}(z)$ ,  $G_{u,v}(z)$ , and  $B_{u,v}(z)$  are the convolution results of three grayscale images with orientation  $u$  and scale  $v$ . Then, unichrome feature representation of one color image is represented by

$$\text{uni} = \left[ \sqrt{\sum R_{u,v}^2(z)}, \sqrt{\sum G_{u,v}^2(z)}, \sqrt{\sum B_{u,v}^2(z)} \right], \quad (4)$$

where the three components of unichrome feature are R unichrome feature, G unichrome feature, and B unichrome feature, respectively. It is clear that unichrome features are defined as values extracted from a single image band.

Then, the difference of normalized  $R_{u,v}(z)$ ,  $G_{u,v}(z)$ ,  $B_{u,v}(z)$  is defined by

$$\begin{cases} RG_{uvv'} = \frac{R_{u,v}(z)}{\sqrt{\sum R_{u,v}^2(z)}} - \frac{G_{u,v'}(z)}{\sqrt{\sum R_{u,v'}^2(z)}} \\ RB_{uvv'} = \frac{R_{u,v}(z)}{\sqrt{\sum R_{u,v}^2(z)}} - \frac{B_{u,v'}(z)}{\sqrt{\sum B_{u,v'}^2(z)}} \\ GB_{uvv'} = \frac{G_{u,v}(z)}{\sqrt{\sum G_{u,v}^2(z)}} - \frac{B_{u,v'}(z)}{\sqrt{\sum B_{u,v'}^2(z)}} \end{cases} \quad (5)$$

where  $u$  and  $\nu$ ,  $v'$  denote the orientation and scales of Gabor filters used, respectively. Note that according to Gabor kernels in Eq. (1), we choose  $\nu$  and  $v'$  as adjacent scales, which means they should meet restriction condition  $|v - v'| \leq 1$ . Then, opponent features can be defined by

$$\text{opp} = \left[ \sqrt{\sum RG_{u,v,v'}^2(z)}, \sqrt{\sum RB_{u,v,v'}^2(z)}, \sqrt{\sum GB_{u,v,v'}^2(z)} \right], \quad (6)$$

where the three components of opponent features are RG opponent feature, RB opponent feature, and GB opponent feature, respectively. It is clear that opponent features are defined as values extracted from the difference of two image bands.

According to Eqs. (5) and (6), we can obtain three equations in Eq. (7). During the feature extraction procedure, dimension and efficiency are the two factors needed to be considered, while in the work by Jain and Healey<sup>19</sup> and Choi et al.,<sup>21</sup> the above factors are not taken into consideration. In our study, we just choose three of them in Eq. (7) to constitute opponent feature so as to decrease feature dimension and increase efficiency. Finally, the CGWT representation of an image is denoted by Eq. (8)

$$\begin{cases} \sqrt{\sum RG_{u,v,v'}^2(z)} = \sqrt{\sum GR_{u,v,v'}^2(z)} \\ \sqrt{\sum RB_{u,v,v'}^2(z)} = \sqrt{\sum BR_{u,v,v'}^2(z)} \\ \sqrt{\sum GB_{u,v,v'}^2(z)} = \sqrt{\sum BG_{u,v,v'}^2(z)} \end{cases} \quad (7)$$

$$\text{CGWT} = \{\text{uni}, \text{opp}\}. \quad (8)$$

### 2.2.2 Extraction of CGOT descriptor

CGOT representation combines Gabor texture<sup>6</sup> and opponent feature together, which substantially decreases the feature dimension compared with CGWT representation. Given one grayscale image  $I$ , then the convolution of  $I$  and Gabor kernels  $\psi_{u,v}$  with orientation  $u$  and scale  $v$  is given by

$$g_{u,v}(z) = I(z) * \psi_{u,v}(z). \tag{9}$$

The mean  $\mu_{uv}$  and standard deviation  $\sigma_{uv}$  of the transform coefficients are defined by

$$\begin{cases} \mu_{u,v} = \iint |g_{u,v}(x, y)| dx dy \\ \sigma_{u,v} = \sqrt{\iint (|g_{u,v}(x, y)| - \mu_{u,v})^2 dx dy} \end{cases} \tag{10}$$

Gabor texture feature composed of  $\mu_{uv}$  and  $\sigma_{uv}$  is denoted using  $T = \{\mu_{u,v}, \sigma_{u,v}\}$ . Then, CGOT representation of an image is denoted by

$$\text{CGOT} = \{\mu_{uv}, \sigma_{uv}, \text{opp}\}. \tag{11}$$

### 2.2.3 Extraction of comparative texture features

Some widely used traditional texture features, such as wavelet texture, LBP, and GLCM, are introduced as comparative methods to give a quantitative analysis. Before extraction of these features, the color images are converted into intensity images using the equation  $\text{gray} = 0.299 * r + 0.587 * g + 0.114 * b$ , where  $r$ ,  $g$ , and  $b$  mean red, green, and blue channels, respectively. Details about these comparative methods are in the following.

Wavelet transform makes a great difference in the field of texture analysis. Let  $I$  be an original image, and then the extraction procedure is described as follows. First, ‘‘haar’’ wavelet is used to construct two decomposition filters, one low-pass filter, and one high-pass filter. Then, 2-level two-dimensional wavelet decomposition is applied to  $I$  by means of above constructed decomposition filters, and six subband images are obtained. Note that the decomposition level is an important parameter and the size of the smallest subimage should not be less than  $16 \times 16$ .<sup>7</sup> Finally, the energy of each subband image is calculated using the following equation:

$$E = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N |x(m, n)|, \tag{12}$$

where  $x(m, n)$  are subband images,  $M \times N$  is the size of original image,  $1 \leq m \leq M$ ,  $1 \leq n \leq N$ . Also, wavelet texture feature of image  $I$  is defined by mean and standard deviation of energy of each subband image using  $T = \{\mu_1, \sigma_1, \mu_2, \sigma_2, \dots, \mu_6, \sigma_6\}$ .

LBP describes the local structure of image texture through calculating the differences between each image pixel and its neighboring pixels. Ojala et al.<sup>8</sup> improved an original LBP operator and developed a generalized grayscale and rotation invariant operator  $\text{LBP}_{P,R}^{\text{riu}2}$  which can detect ‘‘uniform’’ patterns and is denoted by

$$\text{LBP}_{P,R}^{\text{riu}2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & U(\text{LBP}_{P,R}) \leq 2 \\ P + 1 & U(\text{LBP}_{P,R}) > 2 \end{cases} \tag{13}$$

where  $U(\text{LBP}_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|$ , and  $s(x)$  is defined by

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{14}$$

$R$  is the radius of the circularly symmetric neighbor set and  $P$  is the number of equally spaced pixels on the circle.  $g_c$  is the center pixel of circular neighbor and  $g_p (p = 0, 1, \dots, P - 1)$  are the neighbor pixels on the circle.  $U$  is a uniformity measure corresponding to the number of spatial transitions in the ‘‘pattern’’ and  $\text{riu}2$  stands for rotation invariant ‘‘uniform’’ patterns having  $U$  value of at most 2.

In our study, 8 pixels circular neighbor of radius 1, i.e.,  $\text{LBP}_{8,1}^{\text{riu}2}$  operator is used, and a total of 59 grayscale and rotation invariant LBP histogram is accepted.

GLCM is one widely used texture analysis method that considers spatial dependencies of gray levels from the perspective of mathematics. In the work by Haralick et al.,<sup>5</sup> 14 statistical

measures extracted from GLCM are introduced. Nevertheless, many of them are strongly correlated with each other and there are no definitive conclusions about which features are more important and discriminative than others. How to choose appropriate features for texture analysis from 14 statistical measures is still studied by some researchers. Haralick et al. selected four features, energy, entropy, correlation, and contrast, as texture features and conducted classification experiments using a satellite imagery data set, and good classification results are obtained.<sup>5</sup> Considering the good performance of the above four features on remote sensing images, energy, entropy, correlation, and contrast are used in our study. They are defined by

$$\begin{cases} f_1 = \sum_i \sum_j p_{d,\theta}^2(i, j) \\ f_2 = -\sum_i \sum_j p_{d,\theta}(i, j) \log p_{d,\theta}(i, j) \\ f_3 = \frac{\sum_i \sum_j ij p_{d,\theta}(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \\ f_4 = \sum_i \sum_j (i - j)^2 p_{d,\theta}(i, j) \end{cases}, \quad (15)$$

where  $f_1$ ,  $f_2$ ,  $f_3$ , and  $f_4$  stand for energy, entropy, correlation, and contrast, respectively.  $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$ , and  $\sigma_y$  are defined by  $\mu_x = \sum_i i \sum_j p_{d,\theta}(i, j)$ ,  $\mu_y = \sum_i j \sum_j p_{d,\theta}(i, j)$ ,  $\sigma_x^2 = \sum_i (i - \mu_x)^2 \sum_j p_{d,\theta}(i, j)$ , and  $\sigma_y^2 = \sum_j (j - \mu_y)^2 \sum_i p_{d,\theta}(i, j)$ .  $(i, j) \in [1, N_g]$  is the entry in GLCMs, where  $N_g$  means the number of distinct gray levels in the quantized image, and  $p_{d,\theta}(i, j)$  is the normalized GLCM of pixel distance  $d$  and direction  $\theta$ . In our study, pixel distance is set as 1 and four directions  $\theta \in \{0, 45, 90, 135 \text{ deg}\}$  are chosen. In addition, because the used images have 256 gray levels and excessive gray levels will increase the workload of calculating GLCMs drastically, we scale the images to eight gray levels, which means GLCM is one  $8 \times 8$  symmetric matrix and  $N_g = 8$ . Consequently, a total of eight texture features composed of mean and standard deviation of four features in Eq. (15) is obtained.

### 2.2.4 Parameters setting

How to choose optimal parameters for Gabor wavelets is still studied by some researchers because different parameters may result in different experimental results even for the same question. With respect to parameters used in this study, we choose default parameters used in Ref. 22, and the details are described as follows. Gabor wavelets of five scales  $v \in \{0, 1, 2, 3, 4\}$  and eight orientations  $u \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ , which have been used in most cases, are accepted because they can extract texture features from more scales and orientations. For the rest of the parameters,  $\sigma = 2\pi$ ,  $k_{\max} = \pi/2$ , and  $f = \sqrt{2}$  are accepted, which can be regarded as empirical values. In addition, the size of Gabor window is also an important parameter, and it is set as  $128 \times 128$  in this study. Then, a total of 80 Gabor texture features is obtained.

According to Eq. (5) and restriction condition  $|v - v'| \leq 1$ , we can obtain 13 scale groups of  $(v, v') \in \{(0,0), (1,1), (2,2), (3,3), (4,4), (0,1), (1,0), (1,2), (2,1), (2,3), (3,2), (3,4), (4,3)\}$  and eight orientations of  $u$ . Thus, CGWT and CGOT representations are a total of 432 ( $120 + 312$ ) and 392 ( $80 + 312$ ) feature vectors, respectively.

### 2.3 Similarity Measure

Similarity measure is an indispensable and important step in image retrieval systems, and different methods may result in great difference even for identical query images. Some widely used similarity measure methods, such as Minkowski distance, histogram intersection, K-L distance, and Jeffrey divergence, etc. tend to have their own scope of application. In such cases, specific similarity measure methods are defined for certain features in this study.

Given two images  $I_i$  and  $I_j$  with corresponding CGWT representations  $f_i^{\text{CGWT}}$  and  $f_j^{\text{CGWT}}$ , the distance measure of CGWT is defined as in Ref. 19

$$d_{ij}^{\text{CGWT}} = \sum \left( \frac{f_i^{\text{CGWT}} - f_j^{\text{CGWT}}}{\sigma^{\text{CGWT}}} \right)^2, \quad (16)$$

where  $\sigma^{\text{CGWT}}$  is the standard deviation of CGWT representation over the entire image database. This distance measure has been used to classify the textural images and ideal classification results have been achieved.<sup>19</sup>

For CGOT representation, considering it is the combination of Gabor texture feature and opponent feature, we integrate distance measure Eq. (16) and distance measure for Gabor texture features in Ref. 6 and define one much simpler distance measure by

$$d_{ij}^{\text{CGOT}} = \sum \left| \frac{f_i^{\text{CGOT}} - f_j^{\text{CGOT}}}{\sigma^{\text{CGOT}}} \right|, \quad (17)$$

where  $\sigma^{\text{CGOT}}$  is the standard deviation of CGOT representation over the entire image database,  $f_i^{\text{CGOT}}$  and  $f_j^{\text{CGOT}}$  are the corresponding CGOT representation of image  $I_i$  and  $I_j$ , respectively.

Note similarity measure Eq. (17) has similar form but different meanings as similarity measure Eq. (16). Since Gabor texture and opponent feature constitute CGOT representation, distance measure Eq. (17) taking both of them into consideration is appropriate. In this similarity measure, CGOT representation is regarded as a unitary feature, which means it is unnecessary to pay attention to each component of the feature when calculating standard deviation  $\sigma^{\text{CGOT}}$ .

### 3 Experiments and Discussions

#### 3.1 Data Set

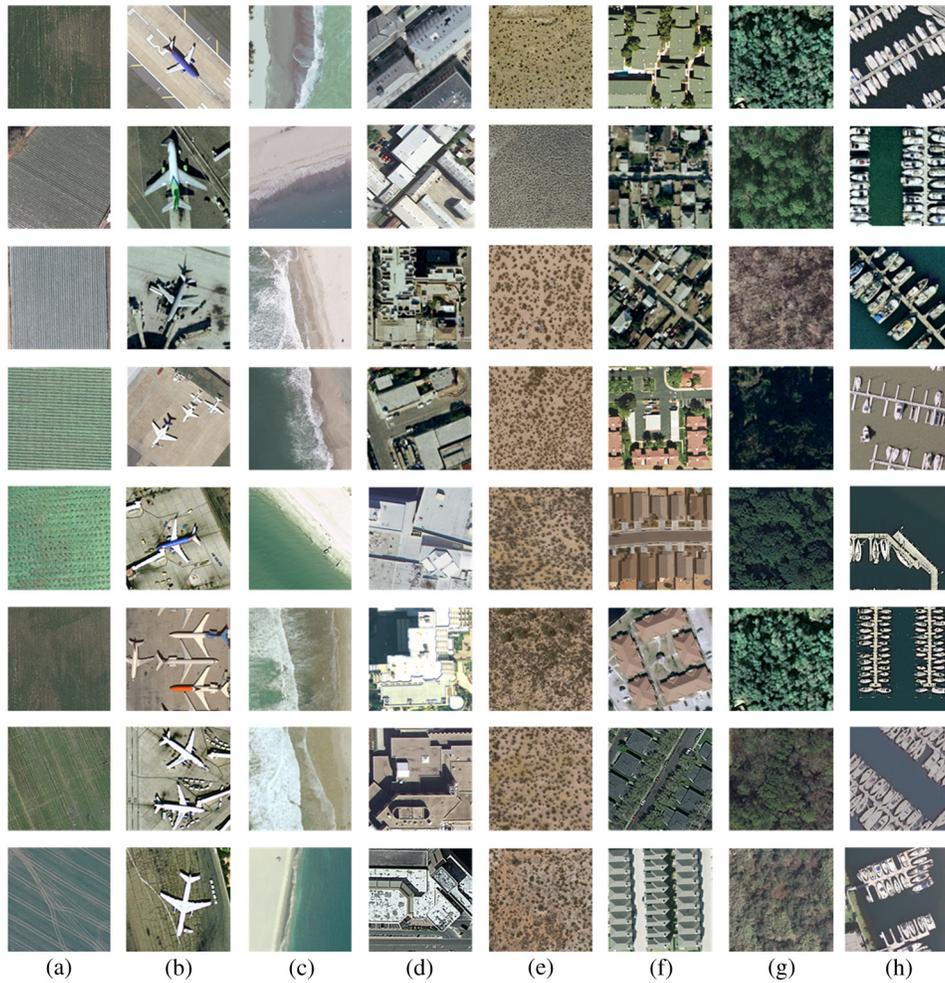
To evaluate the performance of proposed descriptors, eight land-use/land-cover (LULC) classes from UC Merced LULC data set are chosen as retrieval image database. Original LULC is one manually constructed data set consisting of 21 image classes, and the 100 images in each class are tiles with the size of  $256 \times 256$  from large aerial images with the spatial resolution of 30 cm of some US regions.<sup>23</sup> LULC data set has been used in many similar studies<sup>24,25</sup> and made publicly available to other researchers. Some image patches of eight LULC classes used in our experiments are shown in Fig. 2. From left to right, they are agricultural, airplane, beach, buildings, chaparral, residential, forest, and harbor, respectively.

#### 3.2 Performance of Proposed Descriptors

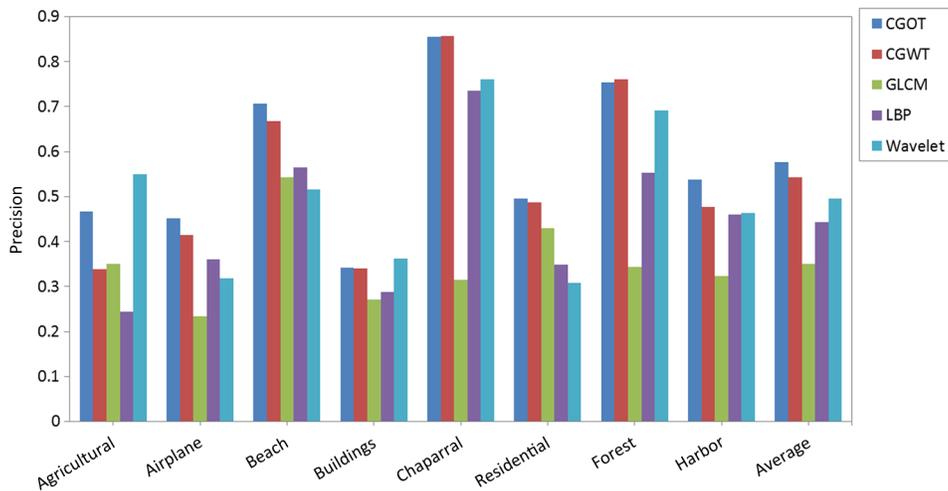
Accurate and objective evaluation criteria have also been a hot topic in the CBIR community. Precision, recall, precision-recall curves, and ANMRR are publicly accepted as evaluation criteria. However, due to the existence of semantic gap, evaluation of CBIR is not effortless. In addition, it is possible to get different performances with different evaluation methods even if the same data set is used.<sup>26</sup> In order to avoid such problems, precision and precision-recall curves are chosen as evaluation methods in this study, because they can be treated as similar evaluations from a different perspective. Precision is the fraction of correct retrievals and recall is the fraction of ground truth items retrieved for a given result set.<sup>23</sup>

Figure 3 shows the performance of proposed features and conventional texture features. The last bin of the histogram with the label “average” gives the average precision of corresponding features. The chart indicates that CGOT and CGWT representations perform better on five classes, i.e., airplane, beach, chaparral, residential, forest, and harbor, and less than perfect on the other two classes, i.e., agricultural and buildings, compared with wavelet texture. Nevertheless, the two proposed features achieve highest average precision on the whole image classes. Meanwhile, we can see CGOT feature increases the average precision of agricultural, airplane, beach, buildings, residential, and harbor by CGWT feature further, which is particularly obvious with respect to agricultural and harbor due to abundant texture information on these image classes.

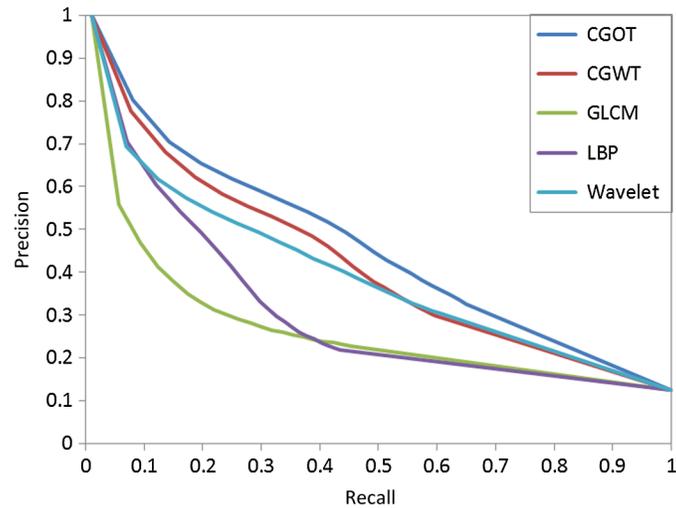
In order to demonstrate the superiority of the proposed representation, precision-recall curves for different features are presented in Fig. 4 through setting different number of returned images. With the increase of returned images, precision by conventional texture features



**Fig. 2** Image patches of eight land-use/land-cover classes used in our experiments, from left to right: (a) agricultural, (b) airplane, (c) beach, (d) buildings, (e) chaparral, (f) residential, (g) forest, and (h) harbor.



**Fig. 3** Average precision of each image class with the proposed features and other texture features.



**Fig. 4** Precision-recall curves for different features.

decrease rapidly, particularly GLCM and LBP. With regard to three rest features, it is evident that CGOT results in the best performance. For CGWT representation and wavelet texture, recall 0.5 can be treated as marginal value. When the value of recall is less than 0.5, CGWT representation performs better, and they have same performance with recall bigger than 0.5. Experimental results, here, are in accordance with the results in Fig. 3, and both of them have validated the effectiveness and good performance of the proposed color texture descriptors.

### 3.3 Comparisons of Used Similarity Measures

As aforementioned, appropriate similarity measure method is necessary in CBIR. For conventional texture features, i.e., GLCM, LBP, and wavelet texture, we choose  $L_2$  distance as similarity measure. For CGWT representation, distance measure presented in Ref. 19 is used. Also, for CGOT representation, characteristics of existed distance measure for Gabor texture, unichrome and opponent features are considered and a simpler distance measure for CGOT representation is defined. Table 1 compares the performance of CGOT representation using proposed similarity measure in Eq. (17) with some other similarity measures, such as  $L_1$  distance,  $L_2$  distance, Jeffrey divergence,<sup>27</sup> and distance measure in Ref. 19.

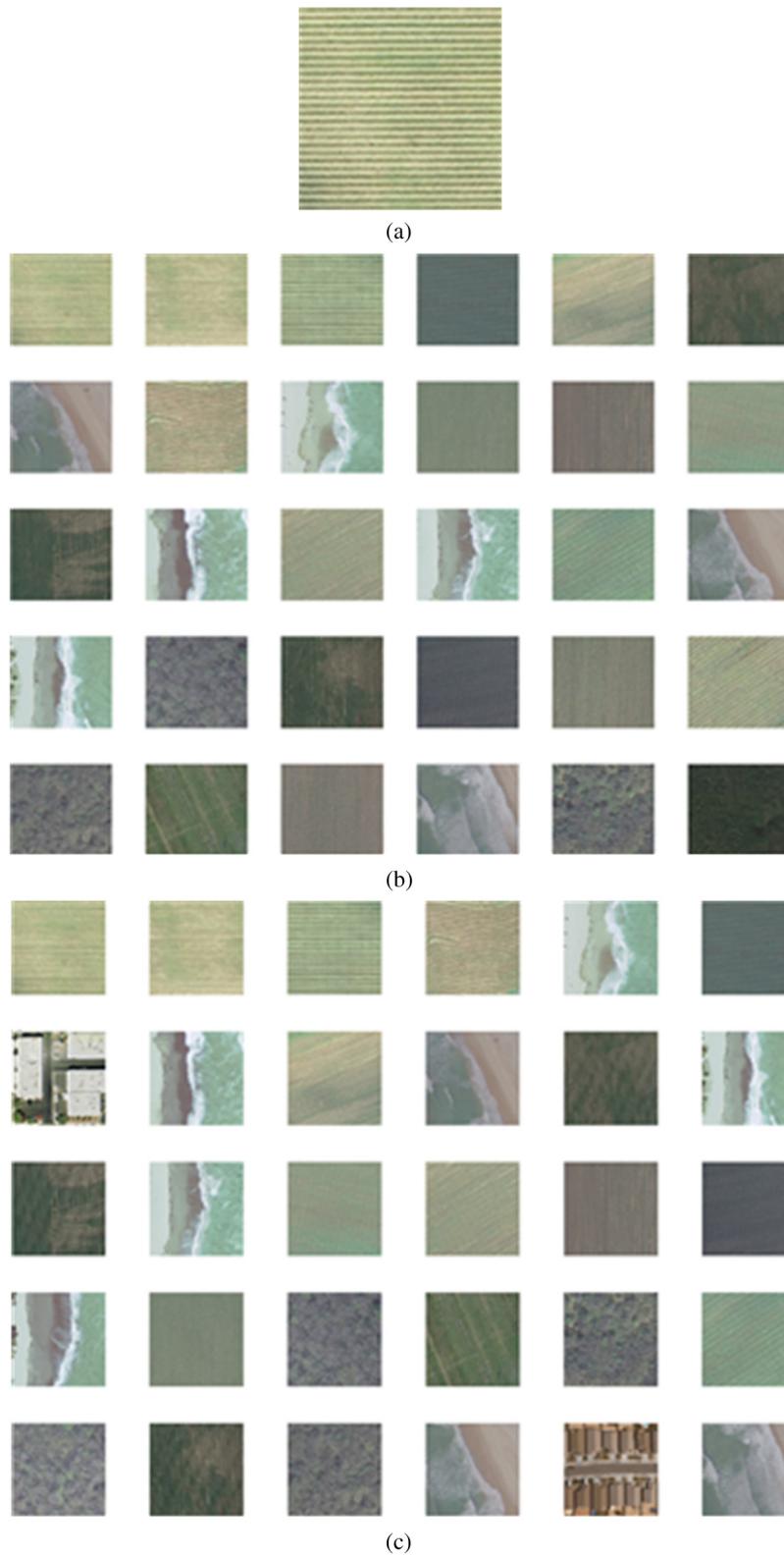
For each group of returned images, the proposed similarity measure achieves highest precision and the average performance is best as well. Table 1 demonstrates that proposed distance measure is an appropriate and effective similarity measure method.

### 3.4 Examples of Remote Sensing Image Retrieval

Figure 5 shows one remote sensing image retrieval example using two proposed descriptors. Figure 5(a) is the query image from agricultural class, and Figs. 5(b) and 5(c) are the first

**Table 1** Comparisons of CGOT using different distance measures.

Distance measure	Precision with various returned images										Average
	10	20	30	40	50	60	70	80	90	100	
Proposed	0.80	0.70	0.65	0.62	0.59	0.56	0.54	0.52	0.49	0.47	0.60
$L_2$ distance	0.76	0.66	0.61	0.57	0.54	0.51	0.48	0.46	0.44	0.41	0.54
$L_1$ distance	0.79	0.69	0.64	0.60	0.57	0.54	0.52	0.50	0.47	0.45	0.58
Jeffrey divergence	0.80	0.70	0.63	0.59	0.56	0.53	0.50	0.48	0.46	0.43	0.57
Distance in Ref. 19	0.78	0.68	0.65	0.59	0.56	0.53	0.51	0.49	0.46	0.44	0.57



**Fig. 5** One retrieval example of agricultural image: (a) query image, (b) performance of color Gabor opponency texture, (c) performance of color Gabor wavelet texture.

30 retrieved images of CGOT and CGWT, respectively. Note that these images are returned in the order of descending similarity, which means images ranking front are more similar to the query image.

According to the retrieval results of two descriptors, CGOT retrieves more similar images than CGWT. In addition, among the first 12 retrieved images, CGOT returns two irrelevant images, while CGWT returns five irrelevant images, which also indicates the better performance of CGOT descriptor.

### 3.5 Discussion

From the previous experiments of remote sensing image retrieval, some interesting points have been concluded.

1. Proposed color texture descriptors, CGWT and CGOT, describe the content of remote sensing images well and achieve a good performance compared with wavelet texture, LBP texture, and GLCM texture. The reason is that they have taken the discriminative information among color bands into consideration.
2. As shown in Fig. 3, CGOT improves the performance of CGWT and achieves highest average precision over the entire image database, and similar performance is obtained in Fig. 4. These results indicate that Gabor texture has better descriptive power than unichrome feature in terms of image texture.
3. The similarity measure defined for CGOT is appropriate. It reveals that the characteristics of one feature should be taken into consideration when defining a similarity measure, because it plays an important role in improving the performance of the proposed representations.

In this study, all experiments are conducted using aerial remote sensing images from one public image database. However, not all the selected images have regular texture structure, which will have an effect on the performance of proposed descriptors. In addition, proposed descriptors are likely to be suitable for hyperspectral image retrieval because they have high spectral resolution and more discriminative information can be extracted from image bands.

## 4 Conclusion

With the rapid development of remote sensing technology, the amount of accessible remote sensing data has been increasing at an incredible rate, which not only provides researchers more choices for various applications, but also brings more challenges. Under the circumstances, CBIR is a better choice for effective organization and management of massive remote sensing data.

Traditionally, low-level features, particular texture features, are widely used in CBIR community for their special characteristics. Nevertheless, conventional texture features tend to be extracted from grayscale images directly and ignore the complementary information that is of great importance between color bands.

To exploit the complementary information and perform remote sensing image retrieval, CGWT and CGOT representations have been proposed based on Gabor filter and opponent process theory. The filtered images by Gabor filter with five scales and eight orientations are obtained first and then unichrome features, opponent features, and Gabor texture features are extracted. Finally, CGWT and CGOT representations are constituted and used in remote sensing image retrieval.

Considering the existence of semantic gap and some other difficulties, two similar evaluations, i.e., precision and precision-recall curves are chosen to evaluate the performance of all texture features. Results demonstrate that CGWT and CGOT perform better than GLCM, LBP, and wavelet texture, and CGOT not only improves the performance of some image classes using CGWT but also increases overall precision of all queried remote sensing images. In addition, a similarity measure for CGOT based on two existed distance measures has been defined. Compared with some widely used distance measures, the proposed similarity measure shows better performance.

In the future, the fusion mechanism of unichrome features and opponent features, Gabor texture and opponent features, as well as the influence of color space on proposed descriptors will be considered.

## Acknowledgments

The author would like to thank Shawn Newsam for his LULC data set and the anonymous reviewers for their comments and corrections. This work was supported in part by National Science and Technology Specific Projects under Grant No. 2012YQ16018505 and National Natural Science Foundation of China under Grant No. 61172174.

## References

1. T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.* **29**(1), 51–59 (1996).
2. J. H. du Buf, M. Kardan, and M. Spann, "Texture feature performance for image segmentation," *Pattern Recognit.* **23**(3), 291–309 (1990).
3. A. W. Smeulders et al., "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(12), 1349–1380 (2000).
4. R. Datta et al., "Image retrieval: ideas, influences, and trends of the new age," *ACM Comput. Surv.* **40**(2), 5:1–5:60 (2008).
5. R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cybern.* **SMC-3**(6), 610–621 (1973).
6. B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(8), 837–842 (1996).
7. T. Chang and C. C. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Trans. Image Process.* **2**(4), 429–441 (1993).
8. T. Ojala, M. Pietikäinen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002).
9. C. H. Lin, R. T. Chen, and Y. K. Chan, "A smart content-based image retrieval system based on color and texture feature," *Image Vision Comput.* **27**(6), 658–665 (2009).
10. Y. D. Chun, N. C. Kim, and I. H. Jang, "Content-based image retrieval using multiresolution color and texture features," *IEEE Trans. Multimedia* **10**(6), 1073–1084 (2008).
11. S. Liapis and G. Tziritas, "Color and texture image retrieval using chromaticity histograms and wavelet frames," *IEEE Trans. Multimedia* **6**(5), 676–686 (2004).
12. M. Flickner et al., "Query by image and video content: the QBIC system," *Computer* **28**(9), 23–32 (1995).
13. C. W. Niblack et al., "QBIC project: querying images by content, using color, texture, and shape," *Proc. SPIE* **1908**, 173–187 (1993).
14. J. Yue et al., "Content-based image retrieval using color and texture fused features," *Math. Comput. Modell.* **54**(3), 1121–1127 (2011).
15. M. Singha and K. Hemachandran, "Content based image retrieval using color and texture," *SIPIJ* **3**(1), 39–57 (2012).
16. M. H. Tsai et al., "Color-texture-based image retrieval system using Gaussian Markov random field model," *Math. Prob. Eng.* **2009**, 1–17 (2010).
17. T. Mäenpää and M. Pietikäinen, "Classification with color and texture: jointly or separately," *Pattern Recognit.* **37**(8), 1629–1640 (2004).
18. L. M. Hurvich and D. Jameson, "An opponent-process theory of color vision," *Psychol. Rev.* **64**(6), 384–404 (1957).
19. A. Jain and G. Healey, "A multiscale representation including opponent color features for texture recognition," *IEEE Trans. Image Process.* **7**(1), 124–128 (1998).
20. M. Shi and G. Healey, "Hyperspectral texture recognition using a multiscale opponent representation," *IEEE Trans. Geosci. Remote Sens.* **41**(5), 1090–1095 (2003).
21. J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Color local texture features for color face recognition," *IEEE Trans. Image Process.* **21**(3), 1366–1380 (2012).

22. C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Trans. Image Process.* **11**(4), 467–476 (2002).
23. Y. Yang and S. Newsam, "Geographic image retrieval using local invariant features," *IEEE Trans. Geosci. Remote Sens.* **51**(2), 818–832 (2013).
24. V. Risojević and Z. Babić, "Fusion of global and local descriptors for remote sensing image classification," *IEEE Geosci. Remote Sens. Lett.* **10**(4), 836–840 (2013).
25. E. Aptoula, "Remote sensing image retrieval with global morphological texture descriptors," *IEEE Trans. Geosci. Remote Sens.* **52**(5), 3023–3034 (2013).
26. H. Müller, S. Marchand-Maillet, and T. Pun, "The truth about corel-evaluation in image retrieval," in *Image and Video Retrieval*, pp. 38–49, Springer, Berlin Heidelberg (2002).
27. Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *Int. J. Comput. Vision* **40**(2), 99–121 (2000).

**Zhenfeng Shao** received his PhD degree from Wuhan University, China, in 2004. He is now a professor of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China. His research interests are image retrieval, image fusion, and urban remote sensing application.

**Weixun Zhou** received his BS degree from Anhui University of Science and Technology, Anhui, China, in 2012. He is now working toward his master's degree at the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China. His research interests are remote sensing image retrieval and image processing.

**Lei Zhang** received her BS degree from Xinyang Normal University, Henan, China, in 2011. She is currently working toward the PhD degree in State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China. Her research interests include dimensionality reduction, hyperspectral classification, sparse representation, and pattern recognition in remote sensing images.

**Jihu Hou** received his BS degree from Hubei University, China, in 2012. He is now working toward his master's degree at the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China. His research interests are remote sensing image retrieval, image processing, and GIS applications.