Journal of Electronic Imaging 17(4), 043026 (Oct–Dec 2008)

Efficient rotation- and scale-invariant texture classification method based on Gabor wavelets

Xudong Xie
Qionghai Dai
Tsinghua University
Department of Automation
Beijing, China
E-mail: xdxie@tsinghua.edu.cn

Kin-Man Lam
The Hong Kong Polytechnic University
Centre for Signal Processing
Department of Electronic and Information Engineering
Hong Kong

Hongya Zhao
City University of Hong Kong
Department of Electronic Engineering
Hong Kong

Abstract. *An efficient texture classification method is proposed that considers the effects of both the rotation and scale of texture images. In our method, the Gabor wavelets are adopted to extract local features of an image and the statistical properties of its gray-level intensities are used to represent the global features. Then, an adaptive, circular orientation normalization scheme is proposed to make the feature invariant to rotation, and an elastic cross-frequency searching mechanism is devised to reduce the effect of scaling. Our method is evaluated based on the Brodatz album and the Outex database, and the experimental results show that it outperforms the traditional algorithms. © 2008 SPIE and IS&T.* [DOI: 10.1117/1.3050071]

1 Introduction

Texture analysis performs fundamental and important roles in many image-based applications, such as remote sensing analysis, medical image interpretation, pattern recognition, and content-based image retrieval. Many methods have been proposed for texture analysis and can be divided into three categories: statistical methods, structural methods, and model-based methods.[1–3] Among these methods, multichannel analysis algorithms, such as the wavelet model[4,5] and the Gabor model,[6] have gained a lot of attention due to their ability to characterize features at different frequencies and orientations. With the wavelet-based method, only three directions (i.e., the horizontal, vertical and diagonal orientations) are considered, while the Gabor wavelets (GWs) can be used to extract features at any specific orientation.

Paper 08046R received Mar. 26, 2008; revised manuscript received Sep. 16, 2008; accepted for publication Oct. 31, 2008; published online Dec. 22, 2008.

1017-9909/2008/17(4)/043026/7/$25.00 © 2008 SPIE and IS&T.

In image retrieval applications, because it is difficult to ensure that a captured texture image has the same orientation and scale as the training images, invariant texture analysis is highly desirable from both the practical and theoretical viewpoints.[7] In Ref. 8, a circular shift technique (CST) is used for rotation normalization based on the extracted Gabor features, so that all images have the same dominant direction. In Ref. 9, rotation-invariant texture features are derived from the symmetrical Gabor filtered images of texture, and the feature used is a modified average absolute deviation from the mean. Gabor features are also used for rotation-invariant texture classification in Ref. 10, where rotation invariance is achieved by using the Fourier expansion of these features with respect to orientation. The above-mentioned methods consider only the effect of rotation. Cui *et al.*[7] proposed a rotation- and scale-invariant feature set based on the Radon transform and multiscale analysis. In Ref. 11, the rotational invariance is achieved by using the two kinds of wavelets with their directional properties, and the scale invariance is achieved by using a method that is an extension of the fractal dimension features. The first- and second-order statistical parameters and the entropy characterize the quality of the features extracted.

In this paper, an efficient texture classification method, which is invariant to rotation and scale, is proposed. Our algorithm uses Gabor wavelets to extract features at different frequencies and orientations. The mean and variance of the Gabor filtered image form the feature to represent the homogeneous texture image and can also be used for texture classification. An adaptive circular orientation normalization (ACON) technique is proposed to reduce the effect of rotation in the images. Compared to the CST in Ref. 8, two modifications are made in our algorithm. First, orien-

tation normalization is performed within each frequency; this can reduce the cross-frequency disturbance. Second, because the Gabor features can only represent the local characteristics of an image, we also consider the statistical properties of an image's gray-level intensity as a global feature. Then, an elastic cross-frequency searching (ECFS) mechanism is devised to search the scales of the features used in matching. Experimental results based on the Brodatz album[12] and the Outex database[13] show that our method can greatly improve the classification performance.

This paper is organized as follows. In Section 2, the fundamentals of GWs are described. The ACON technique, which is used for rotation normalization, and the ECFS mechanism, which can effectively reduce the effect of scaling, are presented in Section 3. Experimental results are given in Section 4, evaluating the different invariant texture image classification algorithms based on two different texture databases. Finally, conclusions are drawn in Section 5.

2 Gabor Feature Extraction

The GWs, whose kernels are similar to the response of the two-dimensional receptive field profiles of the mammalian simple cortical cell,[14] exhibit the desirable characteristics of capturing salient visual properties, such as spatial localization, orientation selectivity, and spatial frequency.[6] In the spatial domain, a GW is a complex exponential modulated by a Gaussian function, which is defined as follows:[14]

$$\psi_{\omega,\theta}(x,y) = \frac{1}{2\pi\sigma^2} e^{-[(x\cos\theta + y\sin\theta)^2 + (-x\sin\theta + y\cos\theta)^2/2\sigma^2]} \times [e^{i(\omega x\cos\theta + \omega y\sin\theta)} - e^{-(\omega^2\sigma^2/2)}], \quad (1)$$

where (x,y) denotes the pixel position in the spatial domain, ω is the radial center frequency, θ is the orientation of the GW, and σ is the standard deviation of the Gaussian function along the x- and y-axes, where $\sigma_x = \sigma_y = \sigma$ is assumed. The value of σ can be derived as follows:

$$\sigma = \kappa/\omega, \quad (2)$$

where $\kappa = \sqrt{2\ln 2}[(2^{\phi}+1)/(2^{\phi}-1)]$ and ϕ is the bandwidth in octaves. By selecting different center frequencies and orientations, we can obtain a family of Gabor kernels from (1) that can be used to represent an image. The Gabor filters with different center frequencies and orientations are shown in Fig. 1.

Given a gray-level image $f(x,y)$, the convolution of $f(x,y)$ and $\psi_{\omega,\theta}(x,y)$ is given as follows:

$$Y_{\omega,\theta}(x,y) = f(x,y) * \psi_{\omega,\theta}(x,y), \quad (3)$$

where the asterisk denotes the convolution operator. Concatenating the convolution outputs, we can obtain a one-dimensional Gabor representation of the input image,

$$\mathbf{Y}_{\omega,\theta} = [Y_{\omega,\theta}(0,0), Y_{\omega,\theta}(0,1), \ldots, Y_{\omega,\theta}(0,N_{\mathrm{H}}-1), Y_{\omega,\theta}(1,0), \ldots, Y_{\omega,\theta}(N_{\mathrm{W}}-1,N_{\mathrm{H}}-1)]^T, \quad (4)$$

where T represents the transpose operation, and N_{W} and N_{H} are the width and height of the image, respectively. In this paper, we consider only the magnitude of the Gabor representations, which can provide a measure of the local properties of an image (for convenience, we also denote it as $\mathbf{Y}_{\omega,\theta}$). Then these Gabor features, with different ω and θ, are concatenated to form a high-dimensional feature vector, as follows:

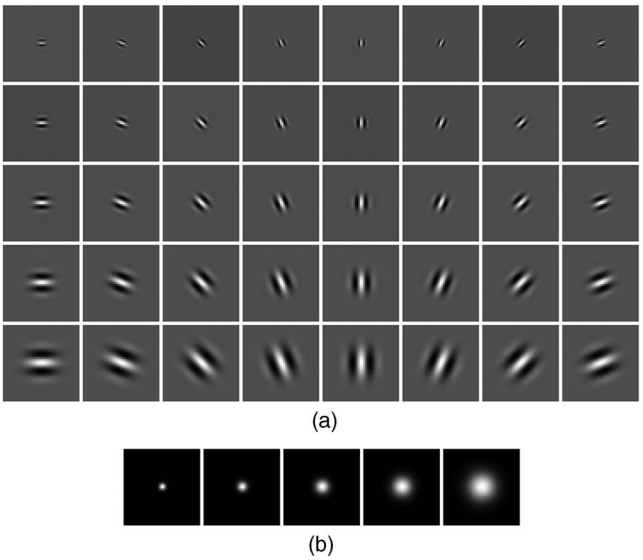

Fig. 1 Gabor filters: (a) Real parts of the Gabor filters at five different center frequencies and eight orientations. The frequencies are $\pi/2$, $\sqrt{2}\pi/4$, $\pi/4$, $\sqrt{2}\pi/8$, and $\pi/8$ from the top to the bottom row, respectively. The orientations are from 0 to $7\pi/8$ in steps of $\pi/8$, from the left to the right column, respectively. (b) Magnitudes of the Gabor filters at the corresponding five different center frequencies.

$$\mathbf{Y} = [\mathbf{Y}^T_{\omega_1,\theta_1}, \mathbf{Y}^T_{\omega_2,\theta_1}, \ldots, \mathbf{Y}^T_{\omega_m,\theta_n}]^T, \quad (5)$$

where m and n are the numbers of center frequencies and orientations used, respectively.

3 Rotation- and Scale-Invariant Feature Representation and Texture Classification

The mean $\mu_{\omega,\theta}$ and the standard deviation $\sigma_{\omega,\theta}$ of $\mathbf{Y}_{\omega,\theta}$ in (6) and (7), respectively, are used to form the features of a homogeneous texture image, and can be used for texture classification,[8]

$$\mu_{\omega,\theta} = \frac{\Sigma_y \Sigma_x Y_{\omega,\theta}(x,y)}{N_{\mathrm{W}} N_{\mathrm{H}}}, \quad (6)$$

$$\sigma_{\omega,\theta} = \sqrt{\frac{\Sigma_y \Sigma_x [Y_{\omega,\theta}(x,y) - \mu_{\omega,\theta}]^2}{N_{\mathrm{W}} N_{\mathrm{H}}}}. \quad (7)$$

We can obtain a feature vector, which includes all the means and standard deviations at different frequencies and orientations, i.e.,

$$\mathbf{P} = [\mu_{\omega_1,\theta_1}, \sigma_{\omega_1,\theta_1}, \mu_{\omega_2,\theta_1}, \sigma_{\omega_2,\theta_1}, \ldots, \mu_{\omega_m,\theta_n}, \sigma_{\omega_m,\theta_n}]^T. \quad (8)$$

This vector is used for texture classification. If the query image is rotated, then the order of the elements in $\mathbf{P}$ will be altered and, therefore, we cannot use $\mathbf{P}$ directly for classification. In this case, an orientation normalization process is necessary.

From (1), we can see that the kernel function of a Gabor wavelet is a periodic function, whose period is 2π. On the basis of this property, the simple circular shift technique is used in Ref. 8 to reduce the effect of rotation. In this technique, for each feature vector $\mathbf{P}$, the orientation with the highest energy (i.e., the largest $\mu_{\omega,\theta}$) is considered the dominant orientation. The corresponding element in the dominant orientation is shifted circularly to become the first element in the feature vector $\mathbf{P}$, and the other elements are also shifted accordingly. Assuming that the dominant orientation is θ_i, (8) is rewritten as

$$\mathbf{P} = [\mu_{\omega_1,\theta_i}, \sigma_{\omega_1,\theta_i}, \mu_{\omega_2,\theta_i}, \sigma_{\omega_2,\theta_i}, \ldots, \mu_{\omega_m,\theta_i}, \sigma_{\omega_m,\theta_i}, \mu_{\omega_1,\theta_{i+1}}, \sigma_{\omega_1,\theta_{i+1}}, \ldots, \mu_{\omega_m,\theta_{i-1}}, \sigma_{\omega_m,\theta_{i-1}}]^T. \quad (9)$$

This method is simple and efficient and can reduce the effect of rotation. However, it does not consider the effect of scaling. In fact, because $\mu_{\omega,\theta}$ represents the energy of an image at a specific frequency and at a specific orientation, its magnitude should be dependent on both the frequency and the orientation. In other words, for a scaled image, due to the effect of scaling, its largest mean $\mu_{\omega,\theta}$ may have a different orientation θ from that of the original image. In this case, performing the circular shift on all the elements of a feature extracted at different frequencies and orientations may result in mismatching. Therefore, in this paper, an ACON technique is proposed to reduce the effect of rotation and to eliminate any cross-frequency disturbance.

For each frequency ω_i, the corresponding extracted features are denoted as $\mathbf{P}_{\omega_i}$, where $\mathbf{P}_{\omega_i} = [\mu_{\omega_i,\theta_1}, \sigma_{\omega_i,\theta_1}, \mu_{\omega_i,\theta_2}, \sigma_{\omega_i,\theta_2}, \cdots, \mu_{\omega_i,\theta_n}, \sigma_{\omega_i,\theta_n}]^T$. For those elements in $\mathbf{P}_{\omega_i}$, if μ_{ω_i,θ_j} has the largest magnitude, then the corresponding θ_j will be considered the dominant orientation at the frequency ω_i. The feature element μ_{ω_i,θ_j} is moved to the first element in $\mathbf{P}_{\omega_i}$, and the other elements are circularly shifted accordingly. The new feature vector is denoted as $\mathbf{P}'_{\omega_i}$, where $\mathbf{P}'_{\omega_i} = [\mu_{\omega_i,\theta_j}, \sigma_{\omega_i,\theta_j}, \mu_{\omega_i,\theta_{j+1}}, \sigma_{\omega_i,\theta_{j+1}}, \ldots, \mu_{\omega_i,\theta_{j-1}}, \sigma_{\omega_i,\theta_{j-1}}]^T$. Finally, the $\mathbf{P}'_{\omega_i}$ from different frequencies are concatenated to build a high-dimensional feature vector, i.e.,

$$\mathbf{P} = [\mathbf{P}'^T_{\omega_1}, \mathbf{P}'^T_{\omega_2}, \ldots, \mathbf{P}'^T_{\omega_m}]^T. \quad (10)$$

As the orientation normalization is performed for each individual frequency, $\mathbf{P}'_{\omega_i}$ is only affected by the orientation and the disturbance from cross frequency is eliminated.

Because the Gabor features can only represent the local characteristics of an image, we also consider the statistical properties of its gray-level intensities, which can give the global information about the image. Denote the mean and the standard deviation of the gray-level intensities of a texture image as μ_0 and σ_0, respectively. Adding these two feature values to (10), we have the following feature vector to represent the texture image:

$$\mathbf{P} = [\mu_0, \sigma_0, \mathbf{P}'^T_{\omega_1}, \mathbf{P}'^T_{\omega_2}, \ldots, \mathbf{P}'^T_{\omega_m}]^T. \quad (11)$$

$\mathbf{P}$ is invariant to rotation and is therefore effective for texture classification. Compared to the CST, two improvements have been made in our algorithm: not only is the orientation normalization performed for each individual frequency, which can reduce the disturbance from cross frequency, but the global information about the texture image is also included.

After rotation normalization, an ECFS mechanism is proposed to reduce the effect of image scaling. As the extracted Gabor features based on different frequencies represent the characteristics of images at different scales,[14] we should consider these features separately. For two texture images to be compared, the distance metric is defined as

$$D = \sum_i D_i + D_0, \quad (12)$$

where

$$D_i = \min_k \sum_j \sqrt{(\mu^{(1)}_{\omega_i,\theta_j} - \mu^{(2)}_{\omega_k,\theta_j})^2 + (\sigma^{(1)}_{\omega_i,\theta_j} - \sigma^{(2)}_{\omega_k,\theta_j})^2}, \quad (13)$$

and

$$D_0 = \sqrt{(\mu^{(1)}_0 - \mu^{(2)}_0)^2 + (\sigma^{(1)}_0 - \sigma^{(2)}_0)^2}. \quad (14)$$

In (13), $\mu^{(1)}_{\omega_i,\theta_j}$ and $\sigma^{(1)}_{\omega_i,\theta_j}$ are the mean and the standard deviation of the first image with frequency ω_i and orientation θ_j, respectively; and $\mu^{(2)}_{\omega_k,\theta_j}$ and $\sigma^{(2)}_{\omega_k,\theta_j}$ denote the mean and the standard deviation of the second image with frequency ω_k and orientation θ_j. In this formulation, the feature vector of the second image at a frequency of, say ω_k, which results in a minimum distance to the first image at frequency ω_i, is considered. This minimum distance is denoted as D_i. In other words, the feature at frequency ω_k of the second image is the most similar or matched one to the feature at frequency ω_i of the first image. In (14), D_0 is a measure of the similarity between the global features of the two images. With this distance measure, the effect of scaling can be effectively reduced. Then, the nearest-neighbor rule is applied to find the most similar pairing between a query image and the images in an image database.

4 Experimental Results

In this section, we will evaluate the performance of our proposed algorithm for texture classification based on the Brodatz album and the Outex database, which contain 112 and 319 texture patterns, respectively. All the texture images are normalized to a size of 64×64. First, we will investigate the selection of the parameters for the Gabor wavelets to be used (i.e., the number of frequencies and orientations). Then, our method is compared to the CST, which mainly focuses on rotation-invariant texture classification. Furthermore, our method is also compared to other methods for rotation- and scale-invariant texture classification.

4.1 *Parameter Setting*

As described in Section 2, the Gabor features are extracted at different frequencies and orientations. The number of frequencies and orientations should be predetermined. To determine these numbers, the 10 images D1–D10 from the Brodatz album were selected. The testing set was generated by rotating and scaling these 10 images. Each of these images is scaled with 10 different scales (0.5–1.5, with 0.1

Table 1 Texture classification performance (in percent) based on Gabor wavelets with different numbers of frequencies and orientations.

m and *n*	2	3	4	5	6
4	87.2	85.4	90.6	93.0	91.9
6	87.9	90.9	94.1	96.9	96.8
8	87.4	90.9	92.9	97.2	97.1
10	86.4	89.9	93.2	98.1	97.9
12	86.4	89.8	91.2	96.4	96.3

intervals, 1.0 is excluded) and rotated by 36 different orientations (10–360 deg, with 10 deg intervals); therefore, a total of 360 testing images is generated. Then, the performance of our method in terms of its recognition rate is measured with different numbers of frequencies and orientations, as tabulated in Table 1. Here, the frequencies and orientations adopted are $\pi/(\sqrt{2})^{p+1}$, where $p=1,2,\ldots,m$, and $2q\pi/n$, where $q=1,2,\ldots,n-1$, respectively. From Table 1, we can see that the best performance can be achieved when five frequencies and ten orientations are employed. This result coincides with the analysis in Ref. 8. Therefore, in our following experiments, *m* and *n* are set as 5 and 10, respectively. To demonstrate a better visualization of the classification performance, the corresponding confusion matrix for this optimal number of frequencies and orientations is shown in Table 2, and some incorrectly classified samples are illustrated in Fig. 2. We can see that, sometimes, one texture may look similar to another texture after the rotation and scaling operations.

4.2 *Experimental Results Compared to CST*

In this experiment, all 112 images in the Brodatz database are used for training, and three different sets of testing images are produced that are used to evaluate the relative performances of the different methods. In set I, all 112 images in the Brodatz database are rotated in increments of 10 deg from 10–360 deg; thus, 4032 rotated texture images ($112\times36=4032$) are generated. In set II, the images are scaled by scaling factors from 0.5 to 1.5 with 0.1 intervals (the image with scaling factor 1.0 is excluded). Thus, this image set has 1120 ($112\times10=1120$) images. For set III, the images are both rotated and scaled as in the previous section [i.e., 36 orientations (10–360 deg, with 10 deg intervals) and 10 scales (0.5–1.5, with 0.1 intervals; 1.0 is excluded). In other words, a total of 40,320 ($112\times36\times10=40,320$) images are generated. From this procedure for producing the various testing sets, we can see that there is no overlap between the images in the training set and

Table 2 The confusion matrix of the result when *m* and *n* are set as 5 and 10, respectively.

Actual targets	Classified Targets									
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	355	0	0	0	0	0	0	0	0	5
D2	0	353	0	0	0	0	0	0	7	0
D3	0	0	360	0	0	0	0	0	0	0
D4	0	0	0	357	0	0	0	3	0	0
D5	0	0	0	0	360	0	0	0	0	0
D6	0	0	0	0	0	360	0	0	0	0
D7	0	0	0	0	12	0	348	0	0	0
D8	0	0	0	0	0	0	0	319	41	0
D9	0	0	0	0	0	0	0	0	360	0
D10	0	0	0	0	0	0	0	0	0	360

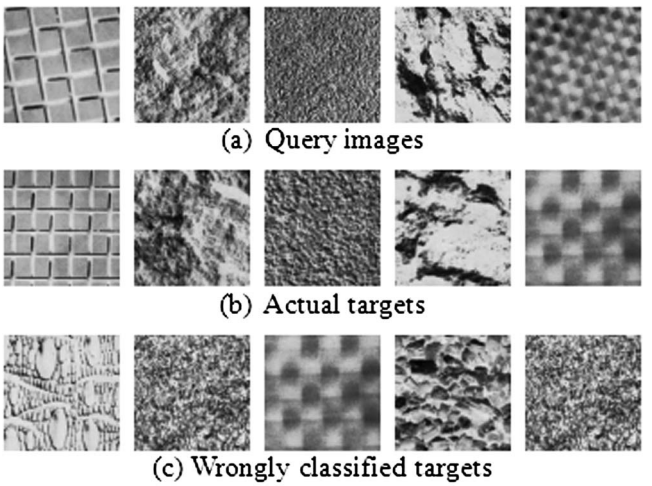

Fig. 2 Some incorrectly classified samples. For the query images in (a), the scaling factors used are 1.3, 0.6, 0.6, 0.6, and 0.5, respectively, while the corresponding rotation angles are 100, 150, 130, 120, and 110 deg, respectively, from the left to the right column.

those in the testing set. Figure 3 shows the texture image D26 and some of its rotated and scaled images.

The performance of our proposed algorithm is evaluated and compared to the circular shift technique. The experimental results are shown in Table 3. The "Gabor Wavelets" method refers to the use of the feature in (8), and the Euclidean distance measure is used for classification. The "Intensity Values" method employs the global statistical properties of the image only; in other words, (14) instead of (12) is used for classification. From Table 3, we have the following observations:

1. For those rotated testing images, the GWs method achieves a poor recognition rate, even lower than that based on the intensity values method. This is because Gabor features are sensitive to the orientation of an image. Consequently, if a query image is rotated, it cannot give a satisfactory result.
2. With the CST, the recognition rate is greatly increased; however, this method does not perform well when the images are scaled. Furthermore, with the use of a more robust and flexible rotation normalization mechanism, our proposed method outperforms the CST when the images are under rotation only.
3. For set I, the ACON technology can achieve the best performance; whereas while for set II, our proposed ECFS mechanism outperforms the other methods. If the images are both rotated and scaled, then combining these two techniques can obtain the best performance. For both sets I and II, the combined method has a performance level slightly lower than the respective best performance; however, it still significantly outperforms the other methods.
4. The statistical properties of the gray-level intensities can provide additional information to the Gabor features, which represent the local characteristics of an image. Combining these two different kinds of features together, the best recognition performance can be achieved.

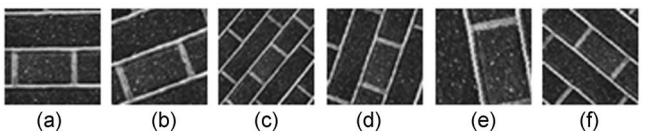

Fig. 3 Sample texture images: (a) original texture image, and (b)–(f) the rotated and scaled images, where the rotation angles are 20, 50, 70, 100, and 140 deg, respectively, and the scaling factors are 1.1, 0.5, 0.7, 1.4, and 0.8, respectively.

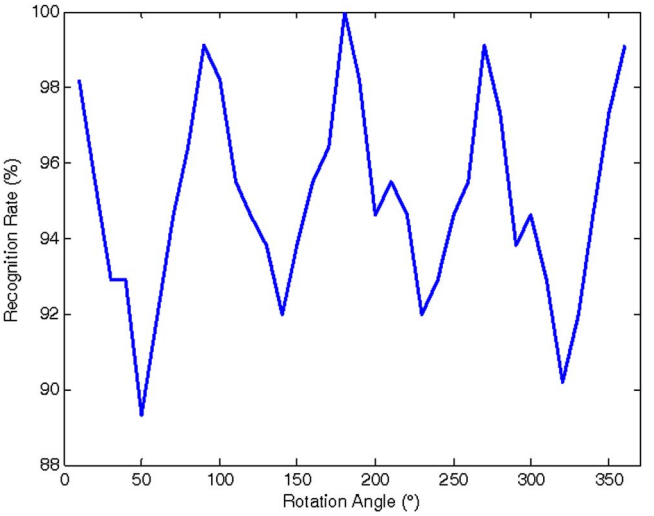

Fig. 4 Texture classification for images with different rotated angles.

In order to investigate the effect of rotating and scaling texture images on the classification performance, we consider sets I and II. The images in sets I and II are divided into 36 and 10 subcategories according to their angles of rotation and scaling factors, respectively. The respective recognition rates are shown in Figs. 4 and 5. From Fig. 4, we can see that the peak values locate at the angles of 10, 90, 180, 270, and 360 deg. This is because when an image is rotated, some information in the edge areas, which cannot be handled by the rotation normalization, is lost. The distorted information between these rotated images and the

Table 3 Texture classification based on the Brodatz album.

Recognition rate (%)	Set I	Set II	Set III
Gabor wavelets	44.7	71.3	26.1
GW+CST	90.3	59.3	52.3
GW+ACON	94.1	60.8	53.5
Intensity values	76.6	51.6	51.0
GW+ACON+intensity values	95.7	66.4	61.5
GW+intensity values+ECFS	41.7	89.6	33.4
GW+ACON+ECFS+ Intensity values	95.0	87.1	80.1

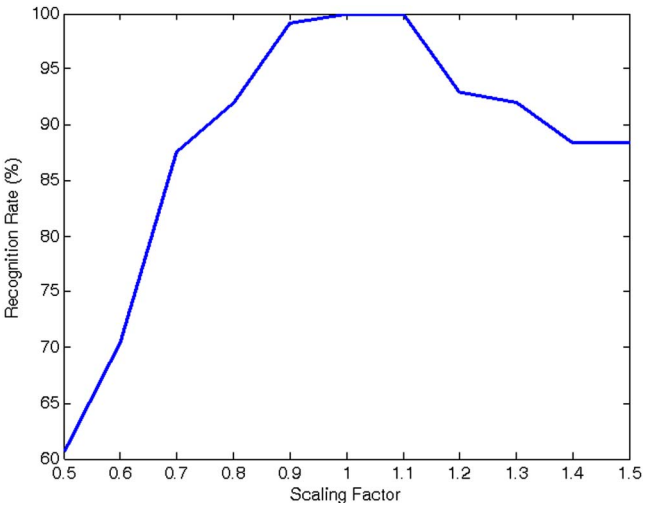

Fig. 5 Texture classification for images with different scaling factors.

original image is smallest at these angles. From Fig. 5, the performance is optimal when the scaling factor is close to 1.0. Furthermore, we can see that the proposed method can achieve a better performance for enlarged images than for the reduced images. This is because more detailed texture information will be lost when the resolution is reduced.

4.3 *Experimental Results Based on the Outex Database*

The experiments in the previous section were based on the Brodatz album, where the testing images are produced by artificially scaling and rotating the original images. In order to make the results independent of artificially introduced artifacts, we also consider another texture database (i.e., the Outex database),[13] where the textures have been rotated and scaled physically before being imaged by a camera. In our experiment, 319 feature patterns are considered (i.e., the training set includes 319 images). For each texture, only one image is adopted for training, whose resolution is 360 dpi. Three testing sets are built according to the characteristics of the images. Set I considers only those images with rotation, where the rotated angles are 5, 10, 15, 30, 45, 60, 75, and 90 deg. Set II includes images with different resolutions: 300, 500, and 600 dpi. In Set III, both the scaling and rotation effects are considered. The numbers of images in the three testing sets are 2552 (319×8), 957 (319×3), and 8613 ($319\times9\times3$), respectively. The experimental results based on the different methods are shown in Table 4. We can see that, similar to the conclusions in Section 4.2, ACON can reduce the effects of rotation, ECFS is the most effective method for feature classification under scaling variations, and when both rotation and scaling effects exist, the combined algorithm can achieve the best performance.

4.4 *Experimental Results Compared to Other Methods*

In this section, we compare the performance of our method to some other texture classification algorithms that are invariant to rotation and scale. The simulation was performed according to the descriptions in Ref. 7, where 25 classes of natural texture images from the Brodatz database were selected and three testing image sets were built. The images in Set I are both rotated and scaled, with 24 orientations (0–345 deg, with 15 deg intervals) and 5 scaling factors (0.6–1.4, with 0.2 intervals). Therefore, this set has 3000 (25×120) images. The images in set II are rotated with 72 orientations only (0–355 deg, with 5 deg intervals). In this way, this image set has 1800 (25×72) images. For set III, each image is scaled by different factors to form 18 samples (0.48–1.84, with 0.08 intervals); thus, this set has 450 (25×18) images. In the experiment, 20 samples from each class in set I (i.e., a total of 500 images) are used for training and the other images in all the image sets are used for testing. The average recognition rates for the different methods are shown in Table 5. We see that our proposed method can achieve the best performance.

We have also evaluated our method for the case when only one image from each class in set I (i.e., 25 images only) is selected for training, and the other images in all image sets are used for testing. The results are shown in the last row of Table 6. We can see that although the number of training images is greatly reduced, our proposed method can still achieve almost the same recognition rate as the methods in Refs. 7 and 15.

Table 4 Texture classification based on the Outex database.

Recognition rate (%)	Set I	Set II	Set III
Gabor wavelets	46.6	37.8	19.5
GW+CST	76.1	31.2	26.6
GW+ACON	74.5	30.8	25.9
Intensity values	78.1	63.7	59.4
GW+ACON+intensity values	93.1	61.9	58.9
GW+intensity values+ECFS	75.2	80.1	61.2
GW+ACON+ECFS+ Intensity values	92.9	76.5	72.8

5 Conclusion

In this paper, we have proposed an efficient rotation- and scale-invariant texture classification method. In our method, GWs are used for extracting an image's local fea-

Table 5 Texture classification based on different methods.

	Radon transform[7]	Log-polar wavelet energy signature[15]	Standard wavelet packet energy signature[16]	Our method
Average recognition rate (%)	92.2[a]	92.1[a]	83.5[a]	99.2

[a]Data are from Ref. 7.

Table 6 Texture classification based on different training sets.

(%)	Set I	Set II	Set III	Average
500 images for training	99.4	100.0	95.3	99.2
25 images for training	88.7	98.7	90.4	92.3

tures, and then the means and standard deviations of these features at different frequencies and orientations are calculated. After an ACON operation, an ECFS mechanism is proposed for classification that can effectively reduce the effect of scaling. Considering the fact that the Gabor features represent only the local features of an image, the statistical properties of the gray-level intensities of an image are also used in our algorithm. The experiments were conducted using the Brodatz album and the Outex database. Experimental results show that our proposed method can achieve the best performance when compared to other algorithms.

Acknowledgments

This work was supported, in part, by The Distinguished Young Scholars of NSFC (Grant No. 60525111), The General Program of NSFC (Grant No. 60872085), and in part, by the 863 Program (Grant No. 2007AA01Z332).

References

1. T. R. Reed and J. M. H. Du Buf, "A review of recent texture segmentation and feature extraction techniques," *CVGIP: Image Understand.* **57**(3), 359–372 (1993).
2. T. Ojala, M. Pietikinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recogn.* **29**(1), 51–59 (1996).
3. J. G. Zhang and T. N. Tan, "Brief review of invariant texture analysis methods," *Pattern Recogn.* **35**(2), 735–747 (2002).
4. R. Manthalkar, P. K. Biswas, and B. N. Chatterji, "Rotation and scale invariant texture features using discrete wavelet packet transform," *Pattern Recogn. Lett.* **24**(14), 2455–2462 (2003).
5. S. Arivazhagan and L. Ganesan, "Texture classification using wavelet transform," *Pattern Recogn. Lett.* **24**(9–10), 1513–1521 (2003).
6. C. Liu and H. Wechsler, "Independent component analysis of Gabor features for face recognition," *IEEE Trans. Neural Netw.* **14**(4), 919–928 (2003).
7. P. Cui, J. Li, Q. Pan, and H. Zhang, "Rotation and scaling invariant texture classification based on radon transform and multiscale analysis," *Pattern Recogn. Lett.* **27**(5), 408–413 (2006).
8. S. Arivazhagan, L. Ganesan, and S. Padam Priyal, "Texture classification using Gabor wavelets based rotation invariant features," *Pattern Recogn. Lett.* **27**(16), 1976–1982 (2006).
9. R. Manthalkar, P. K. Biswas, and B. N. Chatterji, "Rotation invariant texture classification using even symmetric Gabor filters," *Pattern Recogn. Lett.* **24**(12), 2061–2068 (2003).
10. F. Lahajnar and S. Kovačič, "Rotation-invariant texture classification," *Pattern Recogn. Lett.* **24**(9–10), 1151–1161, (2003).
11. K. Muneeswaran, L. Ganesan, S. Arumugam, and K. Ruba Soundar, "Texture classification with combined rotation and scale invariant wavelet features," *Pattern Recogn.* **38**(10), 1495–1506 (2005).
12. P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover, Toronto (1966).
13. T. Ojala, T. Maenpaa, M. Pietikainen, J. Viertola, J. Kyllonen, and S. Huovinen, "Outex—New framework for empirical evaluation of texture analysis algorithms," in *Proc. of Int. Conf. on Pattern Recognition*, vol. 1, pp. 701–706, IEEE, Piscataway, NJ (2002); http://www.outex.oulu.fi/outex.php.
14. C. K. Chui, *An Introduction to Wavelets*, Academic Press, New York (1992).
15. C. M. Pun and M. C. Lee, "Log-polar wavelet energy signatures for rotation and scale invariant texture classification," *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(5), 590–603 (2003).
16. A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(11), 1186–1191 (1993).

Xudong Xie received his BEng in electronic engineering and MSc in signal and information processing from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 1999 and 2002, respectively. In 2006, he received his PhD from the Department of Electronic and Information Engineering, at The Hong Kong Polytechnic University, Hong Kong. From March 2008 to October 2008, he was a research associate at the Department of Electronic Engineering, at The City University of Hong Kong, Hong Kong. After being a postdoctoral research fellow at the Department of Electronic Engineering, Tsinghua University, Beijing, he joined the Broadband Network and Digital Media Center at the Department of Automation, School of Information Science and Technology, Tsinghua University, Beijing, China, where he is currently an assistant professor. His research interests include image analysis, pattern recognition, and computer vision.

Qionghai Dai received his BS in mathematics from Shanxi Normal University, China, in 1987, and ME and PhD in computer engineering and automatic control from Northeastern University, China, in 1994 and 1996, respectively. After being a postdoctoral research affiliate at the Department of Automation, Tsinghua University, Beijing, he is currently a professor in that department. His research interests include digital video processing and communication, computer vision, and computational photography.

Kin-Man Lam received his associateship in electronic engineering with distinction from The Hong Kong Polytechnic University (formerly called Hong Kong Polytechnic) in 1986. He received his MSc in communication engineering from the Department of Electrical Engineering, at the Imperial College of Science, Technology and Medicine, England, in 1987, and his PhD from the Department of Electrical Engineering at the University of Sydney, Australia, in 1996. Currently, he is an associate professor in the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University. He is the chairman of the IEEE Hong Kong Chapter of Signal Processing and an associate editor of the *EURASIP Journal of Image and Video Processing*. He has published more than 100 technical articles in international journals and conferences. His current research interests include facial image analysis, image and video processing, and computer vision.

Hongya Zhao received her MSE degree from Northeastern University, Shenyang, China, in 2002, and her PhD degree from the Department of Mathematics of Hong Kong Baptist University, Hong Kong, in 2006. She is currently a research associate in the Department of Electronic Engineering, City University of Hong Kong, Hong Kong. Her current research interests include bioinformatics, statistics, and pattern recognition.