Rotation and Scaling Invariant Texture Classification Based on Gabor Wavelets

Xudong Xie*, Jie Gong*, Qionghai Dai*, Feng Xu*

*Department of Automation, Tsinghua University, Beijing, China
xdxie@mail.tsinghua.edu.cn
gongjie99@mails.tsinghua.edu.cn
qhdai@mail.tsinghua.edu.cn
xufeng07@mails.tsinghua.edu.cn

Keywords: Texture classification; Gabor wavelets; rotation invariant classification; scaling invariant classification

Abstract

In this paper, an efficient texture classification method is proposed, which not only considers the effect of rotation, but also is scaling invariant. In our method, the Gabor wavelets are adopted to extract local features of an image, and the statistical property of the intensity values is used to represent the global feature. Then, an adaptive circular orientation normalization scheme is proposed for rotation normalization; and an elastic inter-frequency searching mechanism is used to reduce the effect of scaling. Our method is evaluated based on the Brodatz album, and the experimental results show that it outperforms the traditional algorithms.

1 Introduction

Texture analysis acts 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 which can be divided into three categories: statistical methods, structural methods and model-based methods [1-3]. Within these methods, multi-channel analysis algorithms, such as wavelet model [4, 5] and Gabor model [6], have gained a lot of attentions for their abilities to characterize features at different frequencies and orientations. For the wavelet-based method, only three directions, i.e., horizontal, vertical and diagonal orientations, are considers in wavelet transform, while the Gabor wavelets (GW) can extract features at a specific orientation.

In real applications, because it is difficult to ensure that the captured texture image has the same rotation angle and scaling factor with the training images, invariant texture analysis is highly desirable from both the practical and the theoretical viewpoint [7]. In this paper, an efficient rotation and scaling invariant texture classification method is proposed. In our algorithm, Gabor wavelets are used to extract features at different frequencies and orientations. Then, an adaptive circular orientation normalization technique is used for reducing the effect of rotation; and an elastic inter-frequency searching mechanism is proposed to tackle the problem of scaling. Because the Gabor features can only represent the local characteristics of an image, we also consider the statistical property of the intensity values. The experimental results based on the Brodatz album [8] show that our method can greatly improve the classification performance.

This work was supported in part by The Distinguished Young Scholars of NSFC, No.60525111 and in part by the 863 Program, 2007AA01Z332 .The authors are with Tsinghua University, Beijing 100084, China (e-mail: xdxie@mail.tsinghua.edu.cn).

2 Gabor Feature Extraction

The Gabor wavelets, whose kernels are similar to the response of the two-dimensional receptive field profiles of the mammalian simple cortical cell [9], exhibit the desirable characteristics of capturing salient visual properties such as spatial localization, orientation selectivity, and spatial frequency [10]. In the spatial domain, a Gabor wavelet is a complex exponential modulated by a Gaussian function, which is defined as follows [9]:

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

where x, y denote the pixel position in the spatial domain, ω is the radial center frequency, θ is the orientation of the Gabor wavelet, 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:

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

where $\kappa=\sqrt{2\ln 2}\left(\left(2^{\phi}+1\right)/\left(2^{\phi}-1\right)\right)$, and ϕ is the bandwidth in octaves. By selecting different center frequencies and orientations, we can obtain a family of Gabor kernels from (1), which can be used for representing an image. The Gabor filters with different center frequencies and orientations are shown in Figure 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:

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

where * denotes the convolution operator. The convolution can be computed efficiently by performing the fast Fourier transform (FFT), then point-by-point multiplications, and finally the inverse fast Fourier transform (IFFT). Concatenating the convolution outputs, we can obtain a one-dimensional Gabor representation of the input image,

$$\mathbf{Y}_{\omega,\theta} = \left[Y_{\omega,\theta}(0,0), Y_{\omega,\theta}(0,1), \cdots, Y_{\omega,\theta}(0, N_H - 1), Y_{\omega,\theta}(1,0), \cdots, Y_{\omega,\theta}(N_W - 1, N_H - 1)\right]^T, \quad (4)$$

where N_W and N_H are the width and height of the image, and T represents the transpose operation. In this paper, we only consider 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 vector as (5),

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

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

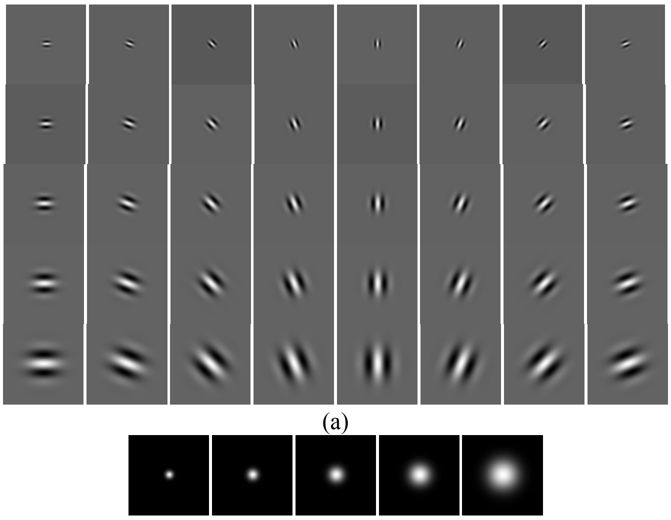


Fig. 1: Gabor filters. (a) Real part of the Gabor filters at five different center frequencies and eight orientations. The frequencies are $\pi/2$, $\sqrt{2}\pi/4$, $\pi/4$, $\sqrt{2}\pi/8$ and $\pi/8$ from the top row to the bottom, respectively. The orientations are from 0 to $7\pi/8$ in steps of $\pi/8$,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

3 Rotation and Scaling Invariant Feature Representation and Texture Classification

For $\mathbf{Y}_{\omega,\theta}$, we consider the mean $\mu_{\omega,\theta}$ and standard deviation $\sigma_{\omega,\theta}$, which represent the feature of the homogeneous texture image and can be used for texture classification [11].

$$\mu_{\omega,\theta} = \frac{\sum_y \sum_x Y_{\omega,\theta}(x,y)}{N_W \cdot N_H}, \quad (6)$$

$$\sigma_{\omega,\theta} = \sqrt{\frac{\sum_y \sum_x \left(Y_{\omega,\theta}(x,y) - \mu_{\omega,\theta}\right)^2}{N_W \cdot N_H}}. \quad (7)$$

We can obtain a features vector, which includes all the mean values and standard deviation values at different frequencies and orientations, i.e.

$$P = \left[\mu_{\omega_1,\theta_1}, \sigma_{\omega_1,\theta_1}, \mu_{\omega_2,\theta_1}, \sigma_{\omega_2,\theta_1}, \cdots, \mu_{\omega_m,\theta_n}, \sigma_{\omega_m,\theta_n}\right]^T, \quad (8)$$

and then this vector can be used for texture classification. If the query image is rotated, the order of the parameters in P is altered, and therefore we cannot directly use P for classification. In this case, orientation normalization technology is necessary.

From (1), we can see that the kernel function of a Gabor wavelet is a periodic function, where the period is 2π. Based on this property, a simple circular shift technique is used in [11] to reduce the effect of rotation. In [11], for each feature vector P, the orientation with the highest energy, i.e. the largest value of $\mu_{\omega,\theta}$, is considered the dominant orientation. The feature element in dominant orientation is moved as first element in the feature vector P. Then, the other elements are circularly shifted. Assume that the dominant orientation is θ_i, (8) is rewritten as

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

This method is simple and efficient, which 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 value should be determined by not only the frequency but also the orientation. In other words, if two images have different scaling factors, the largest value of $\mu_{\omega,\theta}$ from different frequency ω may not have the same orientation θ. In this case, if perform circular shift based on all parameters from different frequencies and orientations, some mismatches may occur. Therefore, in this paper, an adaptive circular orientation normalization technique is proposed for reducing the effect of rotation.

For each frequency ω_i, the extracted features are denoted as P_{ω_i}, where $P_{\omega_i} = \left[\mu_{\omega_i,\theta_1}, \sigma_{\omega_i,\theta_1}, \mu_{\omega_i,\theta_2}, \sigma_{\omega_i,\theta_2}, \cdots, \mu_{\omega_i,\theta_n}, \sigma_{\omega_i,\theta_n}\right]^T$. Within these parameters in P_{ω_i}, the largest value of μ_{ω_i,θ_j} is considered from the dominant orientation θ_j at the frequency ω_i. Then the feature element μ_{ω_i,θ_j} is moved as first element in P_{ω_i} and the other elements are circularly shifted. The new feature vector is denoted as P'_{ω_i}, where $P'_{\omega_i} = \left[\mu_{\omega_i,\theta_j}, \sigma_{\omega_i,\theta_j}, \mu_{\omega_i,\theta_{j+1}}, \sigma_{\omega_i,\theta_{j+1}}, \cdots, \mu_{\omega_i,\theta_{j-1}}, \sigma_{\omega_i,\theta_{j-1}}\right]^T$. Then P'_{ω_i} from different frequencies are concatenated to build a high-dimensional feature vector, i.e.

$$P = \left[P'^{T}_{\omega_1}, P'^{T}_{\omega_2}, \cdots, P'^{T}_{\omega_m}\right]^T . \tag{10}$$

In (10), because the orientation normalization is performed within each frequency, P'_{ω_i} is only affected by the orientation and the disturbance from inter-frequency is eliminated.

Because the Gabor features can only represent the local characteristics of an image, we also consider the statistical property of the intensity values, which can give the global information of an image. Define the mean value and standard deviation of the intensity values of the texture image are μ_0 and σ_0, respectively. From (10), we can obtain a new feature vector, i.e.

$$P = \left[\mu_0, \sigma_0, P'^{T}_{\omega_1}, P'^{T}_{\omega_2}, \cdots, P'^{T}_{\omega_m}\right]^T . \tag{11}$$

We can see that P is a rotation invariant feature vector and can be used for texture classification.

After rotation normalization, an elastic inter-frequency searching mechanism is proposed to reduce the effect of image scaling. Because the extracted Gabor features based on different frequencies represent the characteristics of image at different scales, we should consider these features separately. For two texture images compared, the distance metric is defined as

$$D = \sum_i D_i + D_0 , \tag{12}$$

where

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

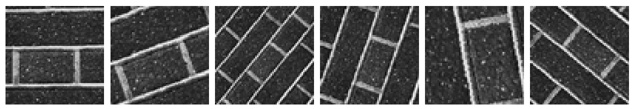
and

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

In (13), $\mu^{(1)}_{\omega_i,\theta_j}$ and $\sigma^{(1)}_{\omega_i,\theta_j}$ are the mean and standard deviation of the first image with the 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 standard deviation of the second image with the frequency ω_k and orientation θ_j. From (13) we can see that for the extracted feature of the first image at frequency ω_i, we should find the minimum distance between it and the feature of the second image at different frequency. In other words, we should find the most matched frequency ω_k in the second image with the frequency ω_i in the first image. In (14), D_0 is used to measure the similarity between the global features of the two images. In this way, the effect of scaling can be effectively reduced. Then, the nearest neighbor rule is adopted to find the most similar pair between a target image and the images within a training set.

4 Experimental Results

In this section, we will evaluate the performance of our proposed algorithm for texture classification based on the Brodatz album [8]. The database used includes 112 different texture patterns. All textures from the Brodatz album (D1–D112) are used for training and three testing image sets are produced for simulation. In Set I, all textures are rotated in steps of 10° up to 360°, and 4032 rotated texture images (112×36 = 4032) are included. In Set II, images are with scaling transform only, where the scaling factor is from 0.5 to 1.5 with 0.1 intervals (the images with scaling factor 1.0 are excluded). In this way, an image set with 1120 (112×10 = 1120) images is created. For images in Set III, which are with joint rotation and scaling transforms, 36 orientations (10° to 360° with 10° intervals) and 10 scaling (0.5 to 1.5 with 0.1 intervals, 1.0 is excluded) are considered. In other words, total 40320 (112×36×10 = 40320) images are created. All images are normalized to a size of 64×64. In Figure 2, the texture image D26 and its several rotated and scaled images, which are used for testing, are shown. For Gabor feature extraction, we select five center frequencies, which are $\pi/2$, $\sqrt{2}\pi/4$, $\pi/4$, $\sqrt{2}\pi/8$ and $\pi/8$, respectively, and ten orientations from 0 to $9\pi/10$ in increments of $\pi/10$.



(a) (b) (c) (d) (e) (f)

Fig.2: 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°, respectively, and the scaling factors are 1.1, 0.5, 0.7, 1.4, and 0.8, respectively.

The performance of our proposed algorithm is evaluated and compared with the circular shift technique [11]. The experimental results are shown in Table 1. In Table 1, the method "Gabor Wavelets" uses the Gabor wavelets to extract features, and (8) is adopted for classification, where the Euclidean distance measure is used. The method "Intensity Values" only considers the global statistical property of the image; in other words, (14) instead of (12) is used for discrimination.

From Table 1, we can see that for the rotated testing images, if we directly use the Gabor wavelets to extract features, and adopt (8) for classification, the recognition rate is poor, even lower than the result based on the intensity values. This is because Gabor features are sensitive to the orientation of an image. If the query image is rotated, it can not give satisfied results. For the method given in [11], where a circular shift technique is applied based on the Gabor features for orientation normalization, the recognition rate can be greatly increased; however, this method is not suitable for images with scaling. For Set I, the method based on the adaptive circular orientation normalization technology can achieve the best performance, and for Set II, the method based on the elastic inter-frequency searching mechanism outperforms others. If the images have both rotation and scaling transforms, combining these two techniques can obtain the best results. Also, we can see that the statistical property of the intensity values can provide additional information and the best recognition performance can be achieved.

Recognition Rate (%)	Set I	Set II	Set III
Gabor Wavelets	44.7	71.3	26.1
GW+ Circular Shift [11]	90.3	59.3	52.3
GW+ Adaptive Circular Orientation Normalization	94.1	60.8	53.5
Intensity Values	76.6	51.6	51.0
GW+ Adaptive Circular Orientation Normalization + Intensity Values	**95.7**	66.4	61.5
GW+ Intensity Values+Elastic Inter-Frequency Searching	41.7	**89.6**	33.4
GW+ Adaptive Circular Orientation Normalization+Elastic Inter-Frequency Searching + Intensity Values	95.0	87.1	**80.1**

Table 1: Texture classification based on different methods.

We also compare our method with the algorithm based on Radon transform [7], which is also rotation and scaling invariant. The simulation is performed according to the descriptions in [7], and the experimental results are shown in Table 2. We can see that our proposed method can still achieve the best result.

	Radon transform [7]	Lgo-polar wavelet energy signature [12]	Standard wavelet packet energy signature [13]	Our method
Recognition Rate (%)	92.2*	92.1*	83.5*	**99.2**

Table 2: Texture classification based on different methods.

* The data is from the Ref. [7].

5 Conclusion

In this paper, we propose an efficient rotation and scaling invariant texture classification method. In our method, the Gabor wavelets are used for extracting image local features, and then the mean values and standard deviations of these features at different frequencies and orientations are calculated. After an adaptive circular orientation normalization operation, an elastic inter-frequency searching mechanism is proposed for classification, which can effectively reduce the effect of scaling. Considering the Gabor features only represent the local features of an image, the statistical property of the intensity values of an image is also used. The experimental results based on the Brodatz album show that our proposed method can achieve the best performance comparing with other algorithms.

References

[1] T.R. Reed, J.M.H Du Buf, "A review of recent texture segmentation and feature extraction techniques", *CVGIP: Image Understanding*, vol.**57**, no.3, pp.359-372, (1993).

[2] T. Ojala, M. Pietikinen, D. Harwood, "A comparative study of texture measures with classification based on feature distributions", *Pattern Recognition*, vol.**29**, no.1, pp.51-59, (1996).

[3] J.G. Zhang, T.N. Tan, "Brief review of invariant texture analysis methods", *Pattern Recognition*, vol.**35**, no.2, pp.735-747, (2002).

[4] R. Manthalkar, P.K. Biswas, B.N. Chatterji, "Rotation and scale invariant texture features using discrete wavelet packet transform", *Pattern Recognition Letters*, vol.**24**, no.14, pp.2455-2462, (2003).

[5] S. Arivazhagan, L. Ganesan, "Texture classification using wavelet transform", *Pattern Recognition Letters*, vol.**24**, no.9-10, pp.1513-1521, (2003).

[6] C. Liu, H. Wechsler, "Independent component analysis of Gabor features for face recognition", *IEEE Trans. Neural Networks*, vol.**14**, no.4, pp.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 Recognition Letters*, vol.**27**, no.5, pp.408-413, (2006).

[8] P. Brodatz, "Textures: A photographic album for artists and designers", Dover, Toronto (1966).

[9] C.K. Chui, "An introduction to wavelets", Boston: Academic Press (1992).

[10] C. Liu, H. Wechsler, "Independent component analysis of Gabor features for face recognition", *IEEE Trans. Neural Networks*, vol.**14**, no.4, pp.919-928, (2003).

[11] S. Arivazhagan, L. Ganesan, S. Padam Priyal, "Texture classification using Gabor wavelets based Rotation invariant features", *Pattern Recognition Letters*, vol.**27**, pp.1976-1982, (2006).

[12] C.M. Pun, M.C. Lee, "Log-Polar wavelet energy signatures for rotation and scale invariant texture classification", *IEEE Trans. Pattern Anal. Machine Intell.*, vol.**25**, no.5, pp.590-603, (2003).

[13] A. Laine, J. Fan, "Texture classification by wavelet packet signatures", *IEEE Trans. Pattern Anal. Machine Intell.*, vol.**15**, no.11, pp.1186-1191, (1993).