

Rotation and Scale Invariant Texture Analysis with Tunable Gabor Filter Banks

Xinqi Chu and Kap Luk Chan

School of Electrical and Electronics Engineering,
Nanyang Technological University,
Singapore 639798
{chux0001,ek1chan}@ntu.edu.sg

Abstract. In this paper, we propose a method that can be used for image texture recognition in the presence of concurrent rotation and scale changes with tunable directional bandpass Gabor filter banks. The method relies on the analysis of the frequency spectra of the image textures, and from which the rotation and scale changes are estimated using a new spectral shift measure. Tunable Gabor filter banks are designed based on the spectral shift measure. Spectral features obtained from applying the tuned Gabor filter bank are used in a novel search strategy to achieve texture recognition. The proposed method is compared with a non-tunable Gabor filter bank and the improvement in recognition performance is demonstrated through the experimental results on 112 Brodatz textures.

1 Introduction

Texture analysis is important in image processing and computer vision. Texture recognition/retrieval has been an active research topic for more than three decades. Applications of texture recognition can be found in medical image analysis, remote sensing, content-based image retrieval, etc. The process of texture recognition is to identify the type of image texture from a finite set of texture classes upon presentation of a newly captured image texture.

Numerous approaches have been reported on texture recognition/retrieval. Ma and Manjunath[13] has reported best retrieval rate on the entire Brodatz[1] and USC texture database using Gabor filters, which is widely accepted as the benchmark method for texture classification and retrieval. However, the vast majority of these approaches assume, either explicitly or implicitly, that the textures have identical scale and orientation. Such an assumption cause their methods to fail in many practical applications in which a texture maybe observed at different orientation and scale from the trained reference texture. Noticing the limit, many works started to address rotation and scale invariance issue. However, many of these methods can only achieve rotation invariance[12][17][21][6][16][8][11].

Scale and rotation invariance is also considered in [7], but their method cannot achieve concurrent rotation and scale invariance since if the spectrum is rotated and scaled, the sum along each scale level will not hold constant while the rotation of orientation also exists. The same happens while we sum along each orientation,

since the sum will not hold constant if the rotation and scale is taking place concurrently. However, most textures in real life are seen as rotated as well as scaled concurrently. Therefore, it is a very important issue that must be addressed.

Despite of its importance, work on simultaneous scale and rotation invariant texture recognition/retrieval is scarce. The major existing approaches include psycho-physical transformation by [9], multi-resolution simultaneous autoregressive(MRSAR) model[18], log-polar wavelet signatures[19], multichannel Gabor filtering[5] as well as using the Wold model for invariant texture analysis[20]. However, their methods were evaluated on a small selection of the image textures. In this paper, motivated by the success of multichannel Gabor filtering, we explore and develop a method by tuning the directional bandpass Gabor filters to achieve recognition of concurrently rotated and scaled textures, thus accomplishing invariant texture recognition with originally non-invariant directional bandpass filters. We carried out the evaluation of our method on the entire 112 Brodatz textures. We assess the performance of our method against using a non-invariant Gabor filter bank with due consideration to the perceptual properties of texture based on the Wold texture model[10].

In section 2, we present the theoretical basis of our method, the derivation of a tunable Gabor filter bank, the spectral shift measure and our search strategy for texture recognition. In section 3, we introduce the Wold texture model and the grouping of Brodatz texture perceptually based on the notion of periodicity, directionality and randomness. We present the experimental evaluation on the entire 112 Brodatz textures in section 4. Conclusion is given in section 5.

2 Achieving Rotation and Scale Invariance

2.1 Our Approach to the Problem of Invariance

According to Leung *et.al*[9], to a human being, matching scaled and rotated patterns is a daily encounter and the biological human visual system appears to be very good at this task. Experiments in human psychology show that the task of matching rotated and scaled shapes is done by mentally rotating and scaling one of the shapes into the orientation and size of the other one and then tested for a match[15][3][2][4].

Therefore, our method proposed here is to find out the rotation and scale factor before the search in the database. A texture spectrum is compared with the reference spectra to estimate the rotation and scale factor. Then the bank of Gabor filters is tuned according to the rotation and scale factor determined from the dominant frequency component by a spectral shift measure.

2.2 Conventional Gabor Filter Bank Design

A 2-D Gabor function $g(x, y)$ and its Fourier transform $G(u, v)$ can be expressed, respectively, as:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j W x \right] \quad (1)$$

and

$$G(u, v) = \exp \left\{ -\frac{1}{2} \left[\frac{(u - W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right] \right\}, \quad (2)$$

where $\sigma_u = \frac{1}{2\pi\sigma_x}$ and $\sigma_v = \frac{1}{2\pi\sigma_y}$. In (1) and (2), σ_x and σ_y characterize the spatial extent and frequency bandwidth of the Gabor filter, and W represents the center frequency of the filter. A set of Gabor functions $g_{m,n}(x, y)$ can be generated by rotating and scaling $g(x, y)$ to form an almost complete and non-orthogonal basis set:

$$g_{m,n}(x, y) = a^{-2m} g_{m,n}(x', y') \quad (3)$$

where

$$x' = x \cos(n\theta) + y \sin(n\theta), \quad (4)$$

$$y' = -x \sin(n\theta) + y \cos(n\theta), \quad (5)$$

where $m = 0, 1, \dots, S-1$, and $n = 0, 1, \dots, K-1$. Parameter S is the total number of scales, and parameter K is the total number of orientations.

The filter parameters are selected such that the adjacent half-peak magnitude contours of the filter responses in the frequency domain are tangent to each other. Follow the same derivation in [13],

$$a = \left(\frac{U_h}{U_l} \right)^{\frac{1}{S-1}} \quad (6)$$

$$\sigma_u = \frac{(a-1)u_H}{(a+1)\sqrt{2 \ln 2}}, \quad (7)$$

$$\sigma_v = \tan(\pi/2K) \sqrt{\frac{u_H^2}{\sqrt{2 \ln 2}} - \sigma_u^2}, \quad (8)$$

2.3 Tunable Gabor Filters Design

Here, we derive a Gabor filter that is tunable to scale and rotation changes. As it is well-known, rotating and scaling a 2-D signal would also rotate and dilate or shrink its spectrum. Such a spectral property gives the motivation of tuning filters. The DFT of a 2-D scaled spatial signal $f(x, y)$ satisfy the property:

$$f(\omega_1 x, \omega_2 y) \leftrightarrow F\left(\frac{u}{\omega_1}, \frac{v}{\omega_2}\right) \quad (9)$$

where $F(u, v)$ is the DFT of the signal $f(x, y)$. The DFT of a rotated 2-D spatial signal $f(x, y)$ is given by:

$$F_{\phi_0}(u, v) \leftrightarrow F(u', v') \quad (10)$$

where $(u', v') = (u \cos \phi_0 + v \sin \phi_0, -u \sin \phi_0 + v \cos \phi_0)$ is the rotated coordinates in the spatial-frequency domain.

Suppose an input texture to be classified is from the same texture class as the reference texture, but is spatially rotated and scaled. In order to locate the

corresponding channels of frequency components, which are just rotated and scaled in the spatial-frequency domain, the reference Gabor filter bank may be tuned by rotating and scaling it accordingly. Therefore, each tuned Gabor filter band will cover the corresponding shifted and/or rotated frequency components.

Suppose that the rotation and scale change $(\delta\phi_t, \delta\rho_t)$ can be estimated through the analysis of the spectrum, the original Gabor filter bank can be tuned by,

$$G_{mn}(x, y) = a^{-2m}G(u'', v''), \quad (11)$$

where

$$u'' = u \cos(n\theta + \delta\phi_t) + v \sin(n\theta + \delta\phi_t), \quad (12)$$

$$v'' = -u \sin(n\theta + \delta\phi_t) + v \cos(n\theta + \delta\phi_t), \quad (13)$$

and the frequency domain standard deviation becomes:

$$\sigma'_u = \sigma'_u / \delta\rho_t = \frac{(a-1)u_H}{\delta\rho_t(a+1)\sqrt{2\ln 2}}, \quad (14)$$

$$\sigma'_v = \sigma'_v / \delta\rho_t = \frac{\tan(\pi/2K)}{\delta\rho_t} \sqrt{\frac{u_H^2}{\sqrt{2\ln 2}} - \sigma_u^2}, \quad (15)$$

Hence, through above formulae, the tunable Gabor filter bank can be constructed and used to locate the corresponding frequency components.

2.4 The Paradox of Using the Tunable Gabor Filter Bank

The problem of applying a tunable Gabor filter is that, the class label of an image texture is not known unless the rotation and scale change is estimated. This means that, the parameters to tune the Gabor filters could not be obtained and the texture features cannot be extracted from the filtered output and thus the matching with reference texture feature can not be carried out. However, if we know about the rotation and scale change, then we know the class of that texture already, this is because the rotation and scale can only be determined correctly when we compare the texture with the corresponding reference texture. Hence, we have to solve the chicken and egg problem.

The solution to the above problem is that, we don't attempt to get the correct parameter in the first place. Instead, we tune the Gabor filter using an estimate based on the texture spectrum. The correct estimation will eventually emerge at the last step since it'll supposedly give us the smallest distance after the comparison of the feature vectors with respect to the reference feature vector for classification. The whole idea is explained in the next subsection.

2.5 Solution to the Paradox

The reference Gabor filter bank is used to extract texture features from all the training reference texture samples. Whereas for any input texture, which may have undergone a rotation and scale change, the bank of filters would be tuned by estimating the rotation and scale change with respect to all the n reference textures.

Rotation and Scale Estimation. Our method of estimation works in the polar coordinates of the frequency spectrum. We find the dominant frequency peak after removing the DC component. The peak will be shifted along the ρ axis if the texture is scaled and it will be rotated if the peak is shifted along the θ axis. A more stable peak estimation can be performed by picking the highest n peaks and then find the value that occur most often by a weighted mean shifting algorithm.

Gaussian Tapering. One crucial issue related to the parameter estimation is the image boundaries. Proper handling of the boundaries is especially important when the image has irregular boundaries or is not quite homogeneous[10]. The spreading out or “leakage” effect of the spectrum is due to energy being artificially generated by the discontinuity at the end points of the waveform, and the solution is to multiply by a window weighting function before the FFT is performed.

Therefore, before transforming the spectrum into polar coordinate, a Gaussian Tapering is applied to reduce this spectrum leakage. The Gaussian tapering window is defined as:

$$g_t(m, n) = e^{-\frac{(m-N/2)^2+(n-N/2)^2}{2\sigma^2}}, (m, n) \in D, \quad (16)$$

where the standard deviation σ is 0.375[10], normalized by the image half-size $N/2$.

The dominant peak of the spectrum will be more stable after the Gaussian tapering, hence the estimation can be performed by measuring the peak shift in the spectrum in polar coordinates. The polar coordinates of peak location of the reference texture in each class(scale=1, rotation 0°) is calculated and saved as: ρ_r and ϕ_r . Then, for input texture image, the location of peak is also estimated as ρ_t and ϕ_t . Therefore, the rotation and scale change is calculated by the following spectral shift measure:

$$\delta\phi_t = \phi_t - \phi_r \quad (17)$$

$$\delta\rho_t = \frac{\rho_t}{\rho_r} \quad (18)$$

Searching for the Correct Class. These rotation and scale change estimates, only the estimation from correct reference class make sense, the other rotation and scale estimation parameters all make no sense since we are comparing with respect to incorrect texture references. This correct estimation hide somewhere inside the estimations. These estimations of rotation and scale change are used to tune the Gabor filter banks for respective classes. After that, the tuned Gabor filter banks(among which, only one of them is correctly tuned) are then used to extract feature vectors from the input image. The feature vectors are stacked to form a matrix of n rows and the number of columns equals to the number of dimensions of the feature vector. Therefore, out of all feature vectors extracted, only one of them is the correct feature vector and will be located at the row whose index is the class index. Let A be such a matrix, and B be the matrix

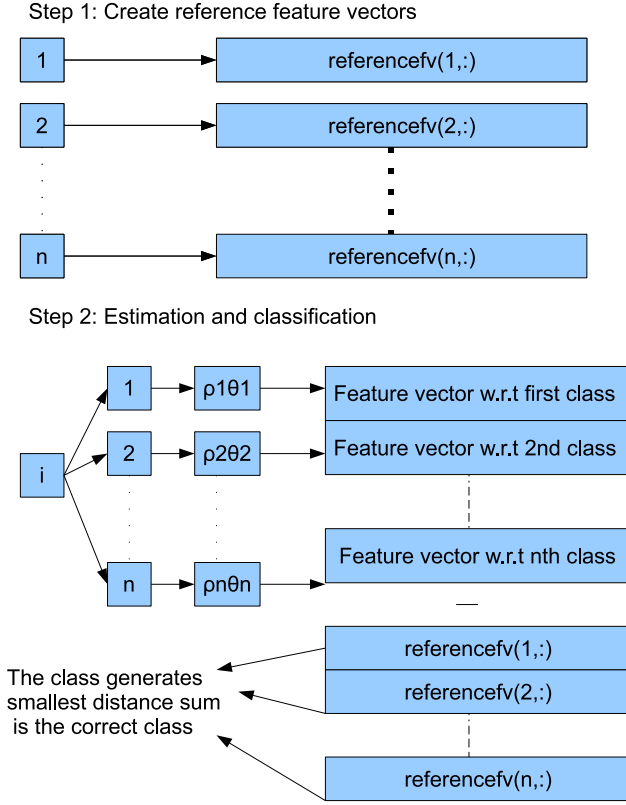


Fig. 1. Illustration of the principle of distance-based classification

formed in a similar way using reference textures(one for each class). The distance measure is defined by

$$d_i = \sum_{j=1}^l \left| \frac{A_{ij} - B_{ij}}{\sigma_r^l} \right| \tag{19}$$

where A_i and B_i are the i th row of the matrix A and B , l is the dimension of the feature vector and σ_r^l is the standard deviation of the l th dimension over all the n classes. Then we determine the class of input texture as:

$$c = \arg \min_i \{d_i\}, i = 1, 2, \dots, n. \tag{20}$$

This idea is illustrated in Figure 1.

All in all, we use the discriminative power of the Gabor filter bank combined with the accurate tuning method to achieve rotation and scale invariance.

3 Brodatz Textures According to the Wold Texture Model

The Wold texture model is a recent method used to model texture[24]. The model allows homogenous texture to be decomposed into three mutually orthogonal components. The 2-D Wold theory applied to the textures is based on some approximations that texture images are homogeneous random fields. A texture image can be considered by the following decomposition:

$$y(m, n) = w(m, n) + p(m, n) + g(m, n),$$

where $w(m, n)$ is the purely indeterministic component, $p(m, n)$ the half-plane deterministic component, and $g(m, n)$ the generalized evanescent component. A 2-D homogeneous regular random field can be represented as the sum of a harmonic components, a countable number of evanescent components and a purely indeterministic component. Brodatz textures have been grouped according to the wold texture model in [10]

Based on this decomposition of homogenous textures, plus the inhomogeneous classes from the Brodatz database, we perceptually classify all the 112 textures into 4 categories according to the visual appearance of the textures and their spectrums, namely:

- Inhomogeneous textures(Category1)
- Textures that exhibits periodicity(Category2)
- Textures that exhibits directionality(Category3)
- Textures that exhibits randomness(Category4)

We are unable to show the textures in each category due to page limit, but these are mostly visually similar to those in [10]. Among the homogenous textures, the textures in category 2 tend to exhibit strong spectral energy peaks. The directional textures may have the spectral energy peaks and ridges along some directions. The random textures show a isotropic Gaussian like spread of spectral energies. By using the above categories, we are able to study the behavior of our spectrum based texture analysis method and reveal some insight on our method and the Gabor-filtering texture features.

4 Experiments and Discussion

4.1 Database Construction

The Brodatz database is widely used as the benchmark for testing recognition results. We also use the Brodatz database to evaluate our method. The scaling range is limited because human itself will perceive a different texture pattern and while scaling is above a certain level, and Phil Brodatz himself regard 1x and 2x magnifications as different classes of textures, so we restrict the scaling range to $\frac{1}{\sqrt{2}} \sim 1 \times \sqrt{2}$. Figure 2 illustrates this point: each column represents

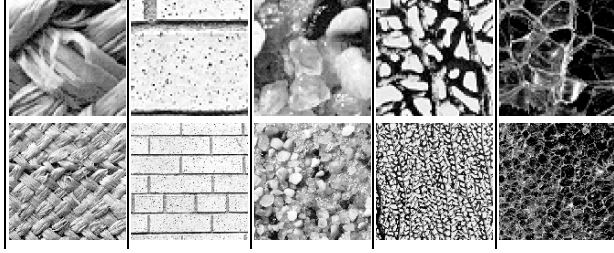


Fig. 2. Each column represents two images from the same texture (from left to right:D18,D26,D27,D87,D112) but of 4 times scale difference, and you can observe that the texture is entirely different though the upper row is just a 4 up-scaled version of the lower row

two different images of 4 times scale difference, and you can observe that the texture is entirely different though the upper row is just a 4 up-scaled version of the lower row. Hence, in this paper we use the entire Brodatz[1] database with each original texture image rotated by 10 different angles (0° to 180°) with 20° as interval, and scaled with 8 different scales (0.7 to 1.4 with 0.1 intervals). Therefore for each of the 112 texture classes, we create 8 scales ($0.7 \sim 1.4$), and for each scale, 10 rotations ($0^\circ \sim 180^\circ$). Hence we'll have 80 images for each class in which 63 images are concurrently rotated and scaled. Then the center portion of 128×128 is cropped for each image. Therefore we have $112 \times 8 \times 10 = 8960$ of 128×128 images created for this database.

4.2 Experimental Results

We carried out classification experiments using the database similar to [13] and the Gabor filter based on conventional design which is not invariant. We studied the effects of rotation and scale changes on the method using conventional Gabor filter banks and tunable Gabor filter bank using the large database containing rotated and scaled textures.

Conventional Gabor filter yields the well-known (and probably the best) overall recognition rate of 74.7 percent on the entire 112 texture classes of Brodatz database. However, because of its band directionality and fixed band coverage, it is not rotation and scale invariant in nature. That's the reason why the overall classification rate dropped to an average of around 19 percent when using a large database with rotated and scaled texture images. From Table 1, it is shown that, by using our method, the adaptive Gabor filter's overall classification rate is pushed back by 44.2 percent to show the effectiveness of our method even on our extensive database.

We can see from the result that, for inhomogeneous textures tested on conventional Gabor filters, the rate drop is the lowest when switched to a rotated and scaled dataset, this is because the spectrum of those textures are highly Gaussian and no spectral peak can be located while doing feature extraction. Therefore, the Gabor features between the original and the rotated/scaled are

Table 1. Comparison With Respect to Conventional Gabor Implementation(%) In-homo – Inhomogenous; Rand – Random

Methods	Datasets	Inhomo.	Periodic	Directional	Rand.	Overall
Conventional	no-rotation/scale	44.9	97.9	87.6	61.2	74.7
Conventional	rotated & scaled	19.7	19.5	19.1	19.7	19.4
Our method	rotated & scaled	50.2	79.7	71.7	56.7	63.7
Rotation/scale effects on conventional		-25.2	-78.4	-68.5	-40.5	-55.3
Rate increase due to tuning		+30.4	+60	+51.6	+37.0	+44.2

not as distinguishable as the ones with high spectral peaks. The corresponding Gabor filter band will still locate similar energy values though not exactly the same while the scale and rotation is small. Whereas for the ones with high spectral peaks, a small change in scale might cause the peak shift to another filter band and the each Gabor filter band could not locate the same amount of energy anymore. In this case, the Gabor filter tuning will be necessary. This is again justified by the column containing periodical textures. The classification rate experience a significant drop by 78.4 percent while the conventional Gabor filter is applied on a rotated and scaled dataset, however, we are able to push the result back by 60 percent to 79.7 on this category by Gabor filter tuning method. Our method improved the result for the directional category by 51.6 percent and random textures by 37 percent, which is in line with our expectation before the experiment, since they’ve got less significant peaks than the periodical ones.

5 Conclusion and Future Work

We address the problem of rotation and scale invariance in texture recognition and propose an tunable Gabor filter approach to achieve concurrent rotation and scale invariant recognition. We carried out experiments on the Brodatz texture with a database whose size is much larger than those presented in any of the existing methods for invariant texture recognition. The performance of our method was tested by a comprehensive experiments across the dataset of the entire brodatz database. The results were carefully analyzed by the grouping based on 2-D Wold-like decomposition. Interesting pattern appears while we compared the performance on the 4 groups side by side with the results of the conventional Gabor filter. Those patterns can be explained by the spectral energy distributions according to the Wold texture model. Our classification result on the entire database($112 \times 80 = 8960$ images) achieves concurrent rotation and scale invariance by 44 percent better accuracy than the well-known method by [13]. Further improvements can be done for estimating rotation and scale change for non-homogenous and random textures which might improve our method further.

Acknowledgements

This work was supported by project of reference SERC TSRP grant number 062 130 0055 of Agency for Science, Technology and Research (A*STAR), Singapore.

References

1. Brodatz, P.: Textures: A Photographic Album For Artists and Designers. Dover (1966)
2. Larsen, A., Bundesen, C.: Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance* 1, 214–220
3. Shepard, R.N., Cooper, L.A.: *Mental Images & Their Transformation*. MIT Press, Cambridge (1982)
4. Bundesen, L.A.C., Farrell, J.E.: Mental transformations of size and orientation. In: *Attention and Performance IX*, pp. 279–294. Lawrence Erlbaum, Hillsdale
5. Fountain, S.R., Tan, T.N.: Extraction of noise robust rotation invariant texture features via multichannel filtering. In: *Proc. International Conference on Image Processing*, October 26–29, vol. 3, pp. 197–200 (1997)
6. Hayley, G.M., Manjunath, B.M.: Rotation invariant texture classification using modified gabor filters. In: *Proc. of IEEE ICIP 1995*, pp. 262–265 (1994)
7. Ma Ju Han, K.-K.: Rotation-invariant and scale-invariant gabor features for texture image retrieval. *Image and Vision Computing* 25(9), 1474–1481 (2007)
8. Kashyap, R., Khotanad, A.: A model based method for rotation invariant texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(4), 786–804 (1986)
9. Leung, M.M., Peterson, A.M.: Scale and rotation invariant texture classification. In: *Conference Record of The Twenty-Sixth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 461–465 (1992)
10. Liu, F.: *Modeling Spatial and Temporal Textures*. PhD thesis, Massachusetts Institute of Technology (September 1997)
11. Madiraju, S.V.R., Liu, C.-C.: Rotation invariant texture classification using covariance. In: *Proc. ICIP 1994. IEEE International Conference Image Processing*, vol. 2, pp. 655–659 (1994)
12. Mahersia, H., Hamrouni, K.: New rotation invariant features for texture classification. In: *Proc. International Conference on Computer and Communication Engineering ICCCE 2008*, pp. 687–690 (2008)
13. Manjunath, B.S., Ma, W.Y.: Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18(8), 837–842 (1996)
14. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7), 971–987 (2002)
15. Shepard, R.: The role of transformation in spatial cognition. In: *Spatial Cognition, Brain Bases and Development*. Lawrence Erlbaum Associates, Mahwah (1988)
16. Greenspan, H., et al.: Rotation invariant texture recognition using a steerable pyramid. In: *Proc. of ICPR 1994*, pp. 162–167 (1994)
17. Zhang, L., Ma, J., Xu, X., Yuan, B.: Rotation invariant image classification based on mpeg-7 homogeneous texture descriptor. In: *Proc. Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing SNPDC 2007*, vol. 3, pp. 798–803 (2007)

18. Jain, A., Mao, J.: Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern Recognition* 25(2), 173–188 (1992)
19. Pun, C.-M., Lee, M.-C.: Log-polar wavelet energy signatures for rotation and scale invariant texture classification 25(5), 590–603 (2003)
20. Wu, Y., Yoshida, Y.: An efficient method for rotation and scaling invariant texture classification. In: *Proc. International Conference on Acoustics, Speech, and Signal Processing ICASSP 1995*, May 9–12, 1995, vol. 4, pp. 2519–2522 (1995)
21. Xu, Z., Pietikainen, M., Ojala, T.: Rotation-invariant texture classification using feature distributions. *Pattern Recognition* 33(2000), 43–52 (2000)
22. Ivry, R., Beck, J., Sutter, A.: Spatial frequency channels and perceptual grouping in texture segmentation. *Computer Vision, Graphics, Image Processing* 37, 299–325 (1987)
23. Jones, J.P., Palmer, L.A.: An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *The Journal of Neuroscience* 58(6) (1987)
24. Francos, J.M., Meiri, A.Z., Porat, B.: A Wold-Like Decomposition of Two-Dimensional Discrete Homogenous Random Fields. *The Annals of Applied Probability* 5(1) (1995)