ROTATION INVARIANT TEXTURE CLASIFICATION USING GABOR WAVELETS

Ramchandra Manthalkar and P. K. Biswas
Department of Electronics and Electrical Communication Engineering , Indian Institute of Technology, Kharagpur PIN 721 302, INDIA
{rrm,pkb}@ece.iitkgp.ernet.in

ABSTRACT

A method of rotation invariant texture classification based on spatial frequency model is developed. Features are derived from the multi-channel Gabor filtering method. The classification performance is first tested on a set 1440 samples of 15 Brodatz textures rotated in 12 directions (0 to 165 in steps of 15 degrees). For the 13-class problem reported in [13] we got better classification with our features. The total Brodatz album is tested using the same features. 10752 samples from Brodatz album are classified (each texture rotated in 12 orientations). The percentage correct classification is 84.92 for Brodatz album.

Keywords: rotation invariance, Gabor wavelets, and texture classification

1. Introduction

Texture classification is very important in image analysis. Content based image retrieval, inspection of surfaces, object recognition by texture, OCR are few examples where texture classification plays a major role. Multi-channel filtering method [1-9] offers computational advantages over other methods for texture classification and segmentation. Most of these algorithms make an implicit assumption that all images are captured under the same orientation. In many practical applications this assumption is not valid. Therefore rotation invariant texture classification becomes necessary in such applications. For feature-based approaches, using anisotropic features rotation invariance can be achieved. Porat and Zeevi [10] use first and second order statistics based upon three spatially localized features, two of which are derived from Gabor filtered image i.e. dominant spatial frequency and orientation of dominant spatial frequency. Greenspan et al. [12], Haley and Manjunath [13], T. N. Tan [14] use rotation invariant features obtained via multi-resolution Gabor filtering. In our scheme almost complete spatial-frequency plane is covered nearly uniformly, and rotation invariance is achieved by transforming Gabor features into rotation invariant features using DFT magnitudes. The motivation for covering total spatial frequency plane nearly uniformly is for conserving spatial, spectral and directional information. Gabor wavelets are used because of their following properties:

- Gabor Elementary Function (GEF) achieves the minimum space-frequency bandwidth product.
- GEFs form a non-orthogonal basis for exact signal reconstruction.
- A narrowband GEF approximates an analytic signal.
- The magnitude of GEF in frequency domain has no sidelobes.
- Gabor decomposition only represents the lowest level of processing in the visual system. It mimics the image coding from the input (cornea and retina) to primary visual cortex, which is seen as the input stage for further and more complex cortical processing [17].

Section 2 explains Gabor wavelets. In section 3 rotation invariant texture features are explained. Section 4 describes the experiments performed. Lastly conclusion and future extensions of this work are discussed. The 15 Brodatz textures used for the first experiments are: crocodile skin (D10), bark of tree (D12), straw (D15), herringbone weave (D17), reptile skin (D22), pressed calf leather (D24), netting (D34), water (D37), oriental straw cloth (D53), straw matting (D56), handwoven oriental rattan (D65), wood grain (D68), oriental straw cloth (D80), oriental straw cloth (D82), raffia (D84). Figure 1 shows the samples of these Brodatz textures. This is a mixed

set of artificial and natural textures and used by tan [14] in his experiment. It is necessary to test oriented textures for rotation invariance and D15, D37, D68 are included in set for that purpose. Haley and Manjunath [14] use another set of 13 textures and they report 96.4 percent correct classification for 13-class problem. This texture set is D12, D29, D92, D112, D9, D24, D19, D84, D17, D37, D68, D15, and D95. Six textures are common to first set and remaining 7 textures are shown in figure 2. Our features give 100 percent classification for the set used in [13]. In [13] phase information is added to magnitude information. We experienced that phase information plays less role in texture discrimination; sometimes for difficult textures the classification performances degrades if phase information is used.

2. Gabor wavelets

Gabor Elementary Functions are Gaussians modulated by complex sinusoids. In two dimensions they are represented by [16]

$$G(x,y) = G_1(x,y)\exp(2pjWx) \qquad (1)$$

where

$$G_1(x,y) = \left(\frac{1}{2ps_x s_y}\right)\exp\left(-\frac{1}{2}\left(\frac{x^2}{s_x^2}+\frac{y^2}{s_y^2}\right)\right) \qquad (2)$$

The Fourier transform of $G(x,y)$ is

$$H(u,v) = \exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{s_u^2}+\frac{v^2}{s_v^2}\right]\right\} \qquad (2)$$

where $s_u = 1/2ps_x$ and $s_v = 1/2ps_y$.

For localized frequency analysis it is desirable to have a Gaussian envelope whose width adjusts with the frequency of the complex sinusoids. We have considered a class of self-similar functions, referred to as Gabor wavelets. Let $G(x,y)$ be the mother Gabor wavelet, then this self-similar filter set can be obtained by appropriate dilations and rotations of mother wavelet through the generating function:

$$G_{mn}(x,y) = a^{-m}G(x',y'), a>1, \qquad (3)$$

m, n = integer

$$x' = a^{-m}(x.\cos q + y.\sin q),$$

$$y' = a^{-m}(-x.\sin q + y.\cos q),$$

where $q = np/K$ and K is the total number of orientations. The scale factor a^{-m} in equation (3) ensures that the energy is independent of m.

$$E_{mn} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|G_{mn}(x,y)|^2 dxdy, \qquad (4)$$

This ensures that all filters in the set have the same energy. The non-orthogonality of Gabor wavelets implies that there is redundant information in the filtered images, and the following strategy is used to reduce this redundancy. Let U_l, U_h denote the lower and upper center frequencies of interest respectively. Let *K* be the number of orientations and *S* be the number of scales in the multi-resolution decomposition. Then the design strategy is to ensure that the half peak magnitude cross-sections of the filter responses in the frequency spectrum touch each other. This results in the following formulas for computing the filter parameters σ_u and σ_v (and thus σ_x and σ_y).

$$a = \left(U_h/U_l\right)^{\frac{1}{S-1}} \qquad (5)$$

$$s_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}} \qquad (6)$$

$$s_v = \tan\left(\frac{p}{2k}\right)\left[U_h - 2\ln 2\left(\frac{s_u^2}{U_h}\right)\right]\cdot \left[2\ln 2 - \frac{(2\ln 2)^2 s_u^2}{U_h^2}\right]^{-\frac{1}{2}} \qquad (7)$$

where

$$W = U_h, q = p/K \qquad m = 0,1,\ldots S-1$$

Here *m* is scale. In order to eliminate sensitivity of the filter response to absolute intensity values, the real (even) components of the 2-D Gabor filters are biased by adding a constant to make them zero mean. This can also be accomplished by setting $H(0,0) = 0$. Imaginary components of Gabor filter are zero mean and do not need this correction. Figure 3 shows the spatial frequency plane coverage of the chosen filter set.

3.Rotation invariant texture features

Gabor Elementary Functions give significant response at zero and very low frequencies. This results in undesirable response to interimage and intraimage variations in contrast and intensity and may cause misclassifications. The response to a zero frequency input relative to the response to an input of equal magnitude Gabor center frequency can be computed as a function of octave bandwidth. The response at zero frequency depends on bandwidth and not on Gabor center frequency. To take care of this aspect the octave bandwidth is kept one where the zero frequency response is 30 db below center frequency response. Each channel is formed by a pair of real Gabor filters. Let the output of each channel is given by

$$C_{ev}(x,y;U,q) = G_1(x,y).\cos(2pUx')* i(x,y)$$
$$C_{odd}(x,y;U,q) = G_1(x,y).\sin(2pUx')* i(x,y) \quad (8)$$

where $G_1(x,y)$ is 2D Gaussian and * denotes 2-D linear convolution.. The channel output $C(x,y)$ is computed as

$$C(x,y;U,q) = \sqrt{C^2{}_{ev} + C^2{}_{odd}} \quad (9)$$

Practical implementation is done in frequency domain for better computational efficiency. The mean value $M(U,q)$ of a channel output $C(x,y;U,q)$ is computed by

$$M(U,q) = \frac{1}{N_1 N_2}\sum C(x,y;U,q) \quad (10)$$

where $N_1 N_2$ is the area of $C(x,y;U,q)$. This value depends on the filter center frequency U and orientation q. The mean values provide powerful features for texture classification. These features are rotation dependant since $M(U,q_i) \neq M(U,q_j)$ for $i \neq j$. $M(U,q_i)$ is computed from $i(x,y)$ and $M(U,q_j)$ from $i(x,y)$ rotated by an angle Δq. $M(U,q_i) = M(U,q_j)$ if $q_j - q_i = \Delta q$. A rotation of input image $i(x,y)$ by Δq is equivalent to a translation of the average channel output $M(U,q)$ by the same amount Δq along the orientation axis. Conjugate symmetry gives $M(U,q_i) = M(U,p + q_i)$. This implies $M(U,q_i)$ is a periodic function of θ with a period of π. $M(U,q)$ contains information about the amplitude and amplitude modulation characteristics of texture's periodic properties within the band. Presently phase is not taken into account. Since rotation of input image $i(x,y)$ corresponds to the translation of $M(U,q)$, DFT of $M(U,q)$ would be rotation invariant feature. Equation (11) shows this operation mathematically

$$f_k = \sum_{k=1}^{8} M(U,q)e^{-2kqpj/K} \quad (11)$$

Total number of features required can be calculated easily. Let there be n orientations and s scales. The features are calculated by taking DFT over all rotations for each scale. Using the DFT property $|f_k| = |f_{-k}|$, n/2 number of features are obtained for each scale. Thus the total number of features will be (sn)/2. Practically the number of orientations for each scale is determined by keeping in view the computational complexity. Texture feature vector *F* is formed by taking the DFT at all scales. Thus in this case for 5 scales we get a 20-element feature vector. A texture may be modeled as a vector valued random field of features *F*. It is assumed that *F* is stationary and has a multivariate Gaussian distribution. For classification purpose, a texture *t* is modeled as a vector valued Gaussian random vector *F* with the conditional probability density function

$$p(\hat{F}/t) = \frac{1}{\sqrt{(2p)^N |C_{Ft}|}} \cdot \exp\left(\frac{-(\hat{F} - m_{Ft})C^{-1}{}_{Ft}(\hat{F} - m_{Ft})^T}{2}\right) \quad (12)$$

where

$$m_{Ft} = E\{F/t\} \quad (13)$$
$$C_{Ft} = E\{F.F^T/t\} - E\{F/t\}\cdot E\{F^T/t\} \quad (14)$$

are mean and covariance of *F*, respectively, *N* is the number of features and $\hat{F}$ is an estimate of *F* based on a sample of texture *t*. The parameters m_{Ft} and C_{Ft} are estimated from statistics over *M* samples for each texture *t*.

$$\hat{m}_{Ft} = \frac{1}{M}\sum_{m=1}^{M}\hat{F}_m \qquad (15)$$

and

$$\hat{C}_{Ft} = \frac{1}{M-1}\sum_{m=1}^{M}\left(\hat{F}_m - \hat{m}_{Ft}\right)\left(\hat{F}_m - \hat{m}_{Ft}\right)^T \qquad (16)$$

where $\hat{F}_m$ is estimate of F based on sample m of texture t.

4. Experimental Results

The number of scales chosen is 5 and orientations are 8. The center frequencies for the Gabor filters are 0.3536, 0.1768, 0.0884, 0.0442, 0.0221 and frequency bandwidth is one octave. The 8 orientations are 0, 22.5, 45, 67.5, 90, 112.5, 135, 157.5 degrees. Thus 40 Gabor filters are used in the experiments, which give 20 rotation invariant features for texture classification. The features are classified using the pattern recognition toolbox PRTools version 3.0 developed by R. P. W. Duin [15]. The four classifiers used are normal densities based quadratic (multi-class) classifier, normal densities based linear classifier (Bayes' rule), uncorrelated normal densities based quadratic classifier and k-nearest neighbour classifier (find k, build classifier). For first experiment only 4 classifiers are used and qdc gives the best performance amongst all. For all other experiments qdc classifier is used. All texture images are rotated in steps of 15 degrees from 0 to 165 degrees to form the training and test images. An 256×256 image is divided in 16 subimages of 64×64 size, half the samples are used for training and other half for testing the performance of the classifier. Thus each image has equal images for training and test phase. The number of images for testing can calculated as 8 images per rotations per texture. The images per texture for training and test phase are 96 each. For a 15-class problem this number becomes 1440. Table 1 gives the results of the 15-texture classification without feature reduction. Table 2 gives the comparison of performance with 13-texture set of [13]. Table 3 gives the results for total Brodatz album using normal densities based quadratic (multi-class) classifier. Using ldc classifier 13-texture classification accuracy is 98.72%. If number of scales is reduced to 16 (i.e. 4 scales as in [13]) % correct classification for ldc and qdc are 99.56 and 99.68 respectively. But for complete Brodatz album ldc gives 78.39% correct classification.

5. Conclusion

The proposed rotation invariant features give very good performance for the Brodatz textures. The textures for which classification error is more than 50 percent are really very hard to classify, there are 7 such textures. Phase information is not used in this classification scheme. Overall classification rate for Brodatz album is 84.92 percent. Figure 4 shows the textures from Brodatz album, for which misclassification arte is more than 50 percent. Figure 5 shows the textures from Brodatz album, which are classified correctly all the times, 4 samples of such images are given.

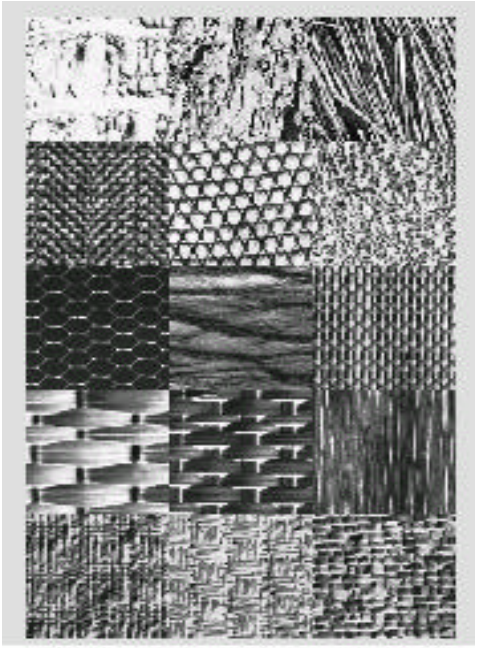

Figure 1: Samples of Brodatz textures used in first experiment. Row1: D10, D12, D15; row 2: D17, D22, D24; row 3: D34, D37, D53; row 4: D56, d65, D68; row 5: D80, D82, D84

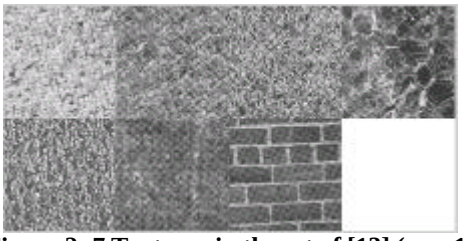

Figure 2: 7 Textures in the set of [13] (row 1: D29, D92, D9, D112; row 2: D24, D19, D95) other 6 are in figure 1.

Figure 3: Spatial frequency coverage of filter set

Short forms for classifiers in table.

qdc : normal densities based quadratic (multi-class) classifier
ldc: normal densities based linear classifier (Bayes' rule)
udc: uncorrelated normal densities based quadratic classifier
knnc: k-nearest neighbour classifier (find k, build classifier)

Table 1: Results of 15-texture classification

Sr. No.	Classifier	Percent correct classification
1	qdc	99.86
2	ldc	99.38
3	udc	99.10
4	knnc	98.68

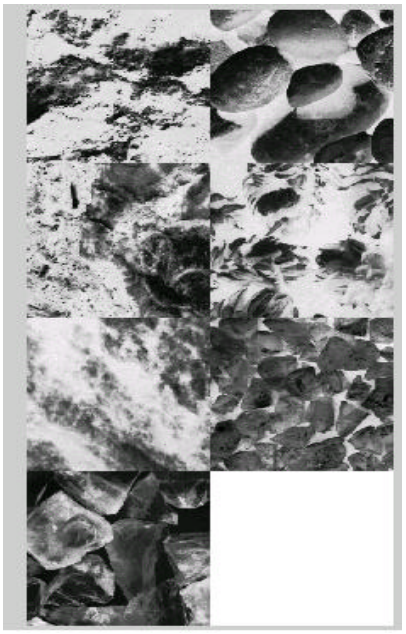

Figure 4: Textures having below 50% correct classification rate: D7, D31, D58, D89, D90, D98, D99 (right to left and top to bottom).

Table 2: Comparison with 13-class classification of [13]

Texture	% correct classification for proposed features	% correct classification reported in [13]
Bark (D12)	100	87.5
Sand (D29)	100	97.9
Pigskin (D92)	100	95.8
Bubbles (D112)	100	100
Grass (D9)	100	95.8
Leather (D24)	100	93.8
Wool (D19)	100	91.7
Raffia (D84)	100	100
Weave (D17)	100	100
Water (D37)	100	97.9
Wood (D68)	100	97.9
Straw (D15)	100	100
Brick (D95)	100	100
	100	96.8

Table 3: Brodatz Classification using qdc

Texture	% classi.	Texture	% classi.	Texture	% classi.	Texture	% classi.
D1	100	D29	97.92	D57	95.83	D85	100
D2	77.08	D30	67.71	D58	27.08	D86	73.96
D3	100	D31	45.83	D59	58.33	D87	100
D4	91.67	D32	92.71	D60	56.25	D88	55.21
D5	70.83	D33	96.88	D61	63.54	D89	44.79
D6	100	D34	100	D62	72.92	D90	33.33
D7	41.67	D35	100	D63	58.33	D91	53.13
D8	96.88	D36	100	D64	98.96	D92	87.5
D9	82.29	D37	92.71	D65	100	D93	97.92
D10	89.58	D38	98.96	D66	91.67	D94	96.88
D11	97.92	D39	84.38	D67	100	D95	100
D12	100	D40	67.71	D68	98.96	D96	88.54
D13	54.17	D41	67.71	D69	82.29	D97	86.46
D14	100	D42	94.79	D70	95.83	D98	42.71
D15	100	D43	72.92	D71	95.83	D99	33.33
D16	95.83	D44	91.67	D72	82.29	D100	79.17
D17	100	D45	81.25	D73	70.83	D101	86.46
D18	90.63	D46	96.88	D74	97.92	D102	92.71
D19	92.71	D47	98.96	D75	89.58	D103	70.83
D20	100	D48	72.92	D76	100	D104	62.5
D21	100	D49	100	D77	100	D105	92.71
D22	100	D50	100	D78	100	D106	85.42
D23	57.29	D51	100	D79	100	D107	69.79
D24	93.75	D52	98.96	D80	97.92	D108	83.33
D25	86.46	D53	100	D81	98.96	D109	76.04
D26	96.88	D54	89.58	D82	100	D110	83.33
D27	50	D55	98.96	D83	100	D111	85.42
D28	81.25	D56	100	D84	98.96	D112	61.46

Overall % Correct Classification rate = 84.92

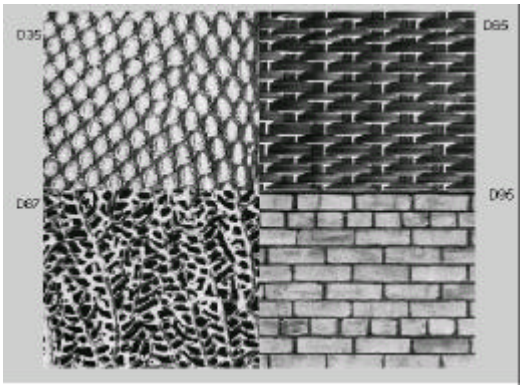

Figure 5: Four textures with 100% classification accuracy

REFERENCES

[1] I. Fogel, D. Sagi, " Gabor filters as texture discriminator," Biological cybernetics, Vol. 61, pp. 103-113, 1989.

[2] D. Dunn, W.E. Higgins and J. Wakeley, " Texture segmentation using 2-D Gabor elementary functions," IEEE trans. on Pattern Analysis and Machine Intelligence, Vol. 16, pp. 130-149, 1994.

[3] Phillippe P. Ohanian, Richard C. Dubes. " Performance evaluation for four classes of textural features," Pattern Recognition, Vol. 25, pp. 819-833, 1992.

[4] J. Malik, P. Perona, "Preattentive texture discrimination with early vision mechanisms," J. Opt. Soc. Am. A, Vol. 7, pp 923-932, 1990.

[5] Trygve Randen, Hakon Husoy, "Filtering for texture classification: a comparative study," IEEE trans. On pattern analysis and machine intelligence, Vol. 21, no. 4, pp. 291-310, 1999.

[6] A. C. Bovik, Analysis of multichannel narrow-band filters for image texture segmentation," IEEE Transactions on signal processing, vol. 39, pp. 2025-2042, September 1991.

[7] A. K. Jain, Farshid Farrokhnia, "Unsupervised texture segmentation using Gabor filters," Pattern Recognition, Vol. 4, no. 12, pp. 1167-1186, 1991.

[8] A. C. Bovik, M. Clark, W. S. Geisler, "Multichannel texture analysis using localized spatial filters," IEEE trans. On pattern analysis and machine intelligence, Vol. 21, no. 1, pp. 55-73, 1990.

[9] M. R. Turner, "Texture discrimination by Gabor functions, "Biological Cybernetics," Vol. 55, pp. 71-82, 1986.

[10] M. Porat and Y. Y. Zeevi, " The generalized scheme of image representation in biological and machine vision," IEEE Transanctions on pattern analysis and machine intelligence, vol. 36, pp. 115-129, January 1989.

[11] F. S. Cohen, Z. Fan and M. A. Patel, " Classification of rotated and scaled textured image using Gaussian Markov random field models," IEEE Transanctions on pattern analysis and machine intelligence, vol. 13, pp. 192-202, February 1991.

[12] H. Greenspan, S. Belongie, R. Goodman. And P. Perona, " Rotation invariant texture recognition using a steerable pyramid," Proc. IEEE international conference on image processing, Jerusalem, Israel, October 1994.

[13] George M. Haley and B. S. Manjunath, "Rotation-invariant texture classification using a complete space-frequency model," IEEE transactions on Image Processing, vol. 8, no. 2, pp. 255-269, February 1999.

[14] T. N. Tan, "Rotation invariant texture features and their use in automatic script identification," IEEE Transactions on pattern analysis and machine intelligence, vol. 20, pp. 751-756, July 1998.

[15] R. P. W. Duin, " PRTools version 3.0 A Matlab toolbox for pattern recognition," January 2000. (http://www.ph.tn.tudelft.nl/prtools.)

[16] Wei-Ying Ma, "Netra: A toolbox for navigating large image databases," Ph. D. dissertation, Electrical and Computer Engineering department, Uni. Of California Santa Barbara, June 1997.

[17] S. Marcelja, "Mathematical description of the responses of simple cortical cells, " J. Opt. Soc. Am., Vol. 70, pp. 1297-1300, April 1980.

[18] L. Van Gool et al, " Texture analysis anno 1983," Computer Vision, Graphics, and Image Processing, Vol. 29, pp. 336-357, 1985.

[19] Todd R. Reed and J. M. H. Du Buf, "A review of recent texture segmentation and feature extraction techniques," Computer Vision, Graphics, and Image Processing, Vol. 57, no. 3, pp. 359-372, 1993.

[20] P. Broadtz, "Textures: A photographic album for artists and designers," Dover, New York, 1966.