ROTATION-INVARIANT TEXTURE CLASSIFICATION USING MODIFIED GABOR FILTERS

George M. Haley

California Microwave, Incorporated
6022 Variel Ave.
Woodland Hills, CA 91367 USA

B. S. Manjunath

Department of Electrical and Computer Engineering, University of California
Santa Barbara, CA 93106 USA

ABSTRACT

A method of rotation invariant texture classification based on a joint space-frequency model is introduced. Multiresolution filters, based on a truly analytic form of a polar 2-D Gabor wavelet, are used to compute spatial frequency-specific but spatially localized *microfeatures*. These microfeatures constitute an approximate basis set for the representation of the texture sample. The essential characteristics of a texture sample, its *macrofeatures*, are derived from the statistics of its microfeatures. A texture is modeled as a multivariate Gaussian distribution of macrofeatures. Classification is based on a rotation invariant subset of macrofeatures.

1. INTRODUCTION

Several methods of rotation invariant texture classification have been proposed. Of spatial domain techniques those based on Markov Random Field (MRF) models predominate. While MRF models are inherently dependent on rotation, they have proven quite effective for many non-rotation invariant applications, and several methods have been introduced to obtain rotation invariance [1], [2], [3].

In other techniques, features are extracted that allow the formulation of a rotation invariant model. Leung and Peterson [4] present two approaches, one that transforms a Gabor-filtered image into rotation invariant features and the other of which rotates the image before filtering; however, neither utilizes the spatial resolving capabilities of the Gabor filter. Porat and Zeevi [5] use sample statistics based upon three spatially localized features, two of which (dominant spatial frequency and orientation of dominant spatial frequency) are derived from a Gabor-filtered image. You & Cohen [6] use filters that are tuned over a training set to provide high discrimination among its constituent textures.

The approach defined herein utilizes a Gabor wavelet transform to create a multiresolution space-frequency representation of the texture sample. A polar, analytic form of the 2-D Gabor function is used. This representation is transformed into a set of *microfeature* vectors that characterize local amplitude, frequency and directionality. *Macrofeatures* are derived from the sample statistics of the microfeatures and characterize a texture sample. The texture class model is a multivariate Gaussian distribution, the parameters of which are estimated from the statistics of the macrofeatures over several sample textures. A rotation invariant subset of macrofeatures is selected for classification, which is based on the MAP criteria.

2. THE GABOR FUNCTION

While many functions may be used for multiresolution space-frequency analysis, Gabor functions are particularly well suited since they achieve the theoretical minimum space-frequency bandwidth product [7],[8]. As filters, they provide optimal spatial resolution for a given bandwidth. Furthermore, they don't exhibit sidelobes. Finally, Gabor functions appear to share many properties with the human visual system [3].

One-dimensional analytic Gabor function

A Gabor function is the product of a Gaussian function and a complex sinusoid. Its general 1-D form is

$$g(x,\Omega,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(\frac{-x^2}{2\sigma^2}\right) \exp(j\Omega x) \qquad (1)$$

$$G(\omega,\Omega,\sigma) = \exp\left[-\frac{\sigma^2(\omega-\Omega)^2}{2}\right] \qquad (2)$$

$G(\omega)$ exhibits a relatively small but, depending on the bandwidth, potentially significant response at $\omega = 0$ (as shown in figure 1) and at very low frequencies. This manifests itself as an undesirable response to inter- and intra-image variations in contrast and intensity due to factors unrelated to the texture itself, such as lighting and shading.

0-8186-7310-9/95 $4.00

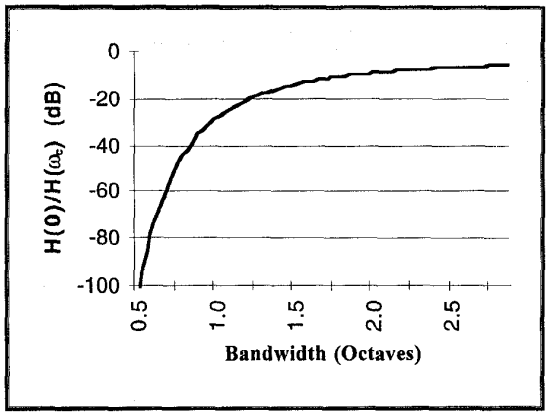

Figure 1. Response of Gabor function at $\omega = 0$ vs. bandwidth.

While point logarithmic [9] or local normalization preprocessing can reduce or eliminate these problems, a straightforward and effective approach is to force the Gabor function to be truly analytic by defining it to be

$$H(\omega) = \begin{cases} 2 \cdot G_A(\omega), & \omega \geq 0 \\ 0, & \omega \leq 0 \end{cases} \quad (3)$$

where $G_A(\omega) = [G(\omega) + G(-\omega)] / 2$ is the DFT of the imaginary part of $g(x)$, i.e., the conjugate antisymmetric part of $G(\omega)$. Thus, the real and imaginary components of $g(x)$ become a Hilbert transform pair. This modification eliminates most deleterious effects at very low frequencies, even when the octave bandwidth of $H(\omega)$ is relatively high.

Two-dimensional Gabor functions

Gabor functions may be extended into two dimensions. The Cartesian form is the product of two orthogonal one-dimensional Gabor functions:

$$h(x', y', \Omega_{x'}, \Omega_{y'}, \sigma_{x'}, \sigma_{y'}) = g(x', \Omega_{x'}, \sigma_{x'}) \cdot g(y', \Omega_{y'}, \sigma_{y'}) \quad (4)$$

$$H(\omega_{x'}, \omega_{y'}, \Omega_{x'}, \Omega_{y'}, \sigma_{x'}, \sigma_{y'}) = G(\omega_{x'} - \Omega_{x'}, \sigma_{x'}) \cdot G(\omega_{y'} - \Omega_{y'}, \sigma_{y'}) \quad (5)$$

The polar form is, in the spatial domain, the product of a Gabor function in the radial direction and a Gaussian function of orientation angle:

$$h(x, y, \Omega, \Theta, \sigma_\rho, \sigma_\theta) = g(\sqrt{x^2 + y^2}, \Omega, \sigma_\rho) \cdot \frac{1}{\sqrt{2\pi}\sigma_\theta} \cdot \exp\left[\frac{-\left(\tan^{-1}(y/x) - \Theta\right)^2}{2\sigma_\theta^2} \right] \quad (6)$$

3. TEXTURE REPRESENTATION

Representation in Gabor space

It has been shown [10], [11] that a set of Gabor functions of the form (4) can be used as a complete, albeit non-orthogonal, basis set for representing a 2-D signal $t(x,y)$. Similarly, $t(x,y)$ can be represented as

$$t(x, y) = \sum_{n_x=-\infty}^{\infty} \sum_{n_y=-\infty}^{\infty} \sum_{p=0}^{P-1} \sum_{r=0}^{R-1} \beta_{n_x, n_y, p, r} \cdot h_{n_x, n_y, p, r}(x, y) \quad (7)$$

where

$$h_{n_x, n_y, p, r}(x, y) = h(x - n_x D, y - n_y D, \Omega, \Theta, \sigma_\rho, \sigma_\theta) \quad (8)$$

is the polar form (6) of Gabor functions sampled appropriately in space, frequency and orientation, and P, R, D, Ω, Θ, σ_ρ and σ_θ are constants. While methods exist to compute exact coefficients for Gabor expansion [10], [11], a good approximation may be obtained by using

$$b_{p,r}(n_x, n_y) \approx \beta_{n_x, n_y, p, r} = h_{n_x, n_y, p, r}(x, y) * t(x, y) \quad (9)$$

from [12], provided that σ_ρ and σ_θ are chosen to provide bandwidths appropriate to the sampling intervals in the space and frequency domains.

Following Bovik, et. al. [13], $b_{p,r}(n_x,n_y)$ may be described as a "band" or "channel" of $t(x,y)$ tuned to the "carrier" frequency $p\Omega$ at the orientation $r\Theta$ and sampled in the x and y dimensions at intervals of D. And since $h_{n_x,n_y,p,r}(x, y)$ is a narrowband, analytic function, $b_{p,r}(n_x,n_y)$ is also narrowband and analytic.

Multiresolution representation

A 2-D polar Gabor function may be expressed as the basic wavelet

$$h_{n_x, n_y, s, r}(x, y) = h(x - n_x D, y - n_y D, \Omega_s, \Theta_r, \kappa, \sigma_\theta) \quad (10)$$

where $\Omega_s = \Omega_0 \cdot 2^{\gamma s}$, $\Theta_r = 360° r/R$, and D, S, R, γ, σ_θ and $\kappa = 1/(\Omega_s \sigma_s)$ are constants. By substituting the coefficient approximations from (9) into (7), the signal may be fully, although approximately, represented with the coefficients $b_{s,r}(n_x,n_y)$.

Transformation into microfeatures

Since $b_{s,r}(n_x,n_y)$ is analytic, it may be decomposed into its amplitude, $a_{s,r}(n_x,n_y) = |b_{s,r}(n_x,n_y)|$, and phase, $\psi_{s,r}(n_x,n_y) = \arg[b_{s,r}(n_x,n_y)]$, components. The gradient of the phase, $\boldsymbol{u}_{s,r}(n_x,n_y) = \nabla[\psi_{s,r}(n_x,n_y)]$[1], contains the frequency characteristics of $t(x,y)$, while $a_{s,r}(n_x,n_y)$ contains its amplitude characteristics.

From these components, microfeatures are defined as

$$f_{\mathrm{A}\,s,q}(n_x,n_y) = \left| \sum_{r=0}^{R/2-1} a_{s,r}(n_x,n_y) \cdot \exp\left(-\frac{2\pi \mathrm{j} rq}{R/2}\right) \right| \quad (11)$$

$$f_{\mathrm{F}\,s,q}(n_x,n_y) = \left| \sum_{r=0}^{R/2-1} \boldsymbol{u}_{s,r}(n_x,n_y) \cdot \exp\left(-\frac{2\pi \mathrm{j} rq}{R/2}\right) \right| \quad (12)$$

$$f_{\mathrm{DA}\,s,q}(n_x,n_y) = \arg\left[\sum_{r=0}^{R/2-1} a_{s,r}(n_x,n_y) \cdot \exp\left(-\frac{2\pi \mathrm{j} rq}{R/2}\right) \right] \quad (13)$$

$$f_{\mathrm{DF}\,s,q}(n_x,n_y) = \arg\left[\sum_{r=0}^{R/2-1} \boldsymbol{u}_{s,r}(n_x,n_y) \cdot \exp\left(-\frac{2\pi \mathrm{j} rq}{R/2}\right) \right] \quad (14)$$

$a_{s,r}(n_x,n_y)$ and $\boldsymbol{u}_{s,r}(n_x,n_y)$ are periodic in r with a period of R/2, and rotating $t(x,y)$ by ϕ produces circular r-shifts of $2R\phi/360°$ in $a_{s,r}(n_x,n_y)$ and $u_{s,r}(n_x,n_y)$. Thus, the DFTs in (11-14) map rotationally induced shifts into the phase angles $f_{\mathrm{DA}}(n_x,n_y)$ and $f_{\mathrm{DF}}(n_x,n_y)$[2], while the amplitude components $f_{\mathrm{A}}(n_x,n_y)$ and $f_{\mathrm{F}}(n_x,n_y)$ are unaffected .

$f_{\mathrm{A}}(n_x,n_y)$ and $f_{\mathrm{F}}(n_x,n_y)$ describe the local amplitude and frequency characteristics, respectively, of $t(x,y)$ and are rotation invariant. Both $f_{\mathrm{DA}}(n_x,n_y)$ and $f_{\mathrm{DF}}(n_x,n_y)$ contain only local directionality information and as such, they may be combined into $f_{\mathrm{D}}(n_x,n_y)$ to reduce the dimensionality of the feature set.

$f_{\mathrm{A}}(n_x,n_y)$, $f_{\mathrm{F}}(n_x,n_y)$ and $f_{\mathrm{D}}(n_x,n_y)$ provide a complete, though approximate, representation of $t(x,y)$. Since all of the transformations described in this section are either exactly or approximately invertible[3], little information has been lost in the representation of a texture sample as microfeatures.

4. THE TEXTURE MODEL

Assuming an underlying second order distribution, a sample may be accurately characterized by its macrofeatures

$$\boldsymbol{F}_{\mathrm{CA}} = \frac{1}{N_x N_y} \sum_{n_y=1}^{N_y} \sum_{n_x=1}^{N_x} f_{\mathrm{A}}(n_x,n_y) \quad (15)$$

$$\boldsymbol{F}_{\mathrm{CF}} = \frac{1}{N_x N_y} \sum_{n_y=1}^{N_y} \sum_{n_x=1}^{N_x} f_{\mathrm{F}}(n_x,n_y) \quad (16)$$

$$\boldsymbol{F}_{\mathrm{CD}} = \frac{1}{N_x N_y} \sum_{n_y=1}^{N_y} \sum_{n_x=1}^{N_x} f_{\mathrm{D}}(n_x,n_y) \quad (17)$$

$$\boldsymbol{F}_{\mathrm{AM}} = \sqrt{\frac{1}{N_x N_y} \sum_{n_y=1}^{N_y} \sum_{n_x=1}^{N_x} \left[f_{\mathrm{A}}(n_x,n_y) - \boldsymbol{F}_{\mathrm{CA}} \right]^2} \quad (18)$$

$$\boldsymbol{F}_{\mathrm{FM}} = \sqrt{\frac{1}{N_x N_y} \sum_{n_y=1}^{N_y} \sum_{n_x=1}^{N_x} \left[f_{\mathrm{F}}(n_x,n_y) - \boldsymbol{F}_{\mathrm{CF}} \right]^2} \quad (19)$$

$$\boldsymbol{F}_{\mathrm{DM}} = \sqrt{\frac{1}{N_x N_y} \sum_{n_y=1}^{N_y} \sum_{n_x=1}^{N_x} \left[f_{\mathrm{D}}(n_x,n_y) - \boldsymbol{F}_{\mathrm{CD}} \right]^2} \quad (20)$$

$\boldsymbol{F}_{\mathrm{CA}}$, $\boldsymbol{F}_{\mathrm{CF}}$ and $\boldsymbol{F}_{\mathrm{CD}}$ describe the amplitude, frequency and directional characteristics, respectively, of the "carrier". $\boldsymbol{F}_{\mathrm{AM}}$, $\boldsymbol{F}_{\mathrm{FM}}$ and $\boldsymbol{F}_{\mathrm{DM}}$ describe the amplitude modulation, frequency modulation and directional modulation characteristics, respectively. All, except $\boldsymbol{F}_{\mathrm{CD}}$, are rotation invariant.

Defining $\boldsymbol{F} = [\boldsymbol{F}_{\mathrm{CA}}\ \boldsymbol{F}_{\mathrm{CF}}\ \boldsymbol{F}_{\mathrm{AM}}\ \boldsymbol{F}_{\mathrm{FM}}\ \boldsymbol{F}_{\mathrm{DM}}]$, and assuming that the random vector $\mathbf{F}$ has a Gaussian distribution, the texture type c is modeled as:

$$p(\mathbf{F} \mid c) = \frac{1}{\sqrt{(2\pi)^M |\boldsymbol{C}_c|}} \exp\left(\frac{-(\mathbf{F}-\boldsymbol{F}_c)\boldsymbol{C}_c^{-1}(\mathbf{F}-\boldsymbol{F}_c)^{\mathrm{t}}}{2} \right) \quad (21)$$

where $\boldsymbol{F}_c = E\{\mathbf{F} \mid c\}$, $\boldsymbol{C}_c = E\{\mathbf{F}\cdot\mathbf{F}^{\mathrm{t}} \mid c\} - E\{\mathbf{F} \mid c\} \cdot E\{\mathbf{F}^{\mathrm{t}} \mid c\}$ and $M = (5 \times \mathrm{R}/2 \times \mathrm{S})$ is the number of macrofeatures.

5. EXPERIMENTAL RESULTS

Experiments were performed on 13 texture images from the Brodatz album [14] and other sources. Each texture was digitized at rotations of 0°, 30°, 60°, 90°, 120°, 150° and 200° as 512×512 images, each of which was then subdivided into 16 128×128 subimages. Half of the subimages were used to define the class texture model and the other half used as test samples. The Gabor transform parameter values were $S = 5$, $R = 16$, $\gamma = 1$, $\kappa = .283$ and $\sigma_\theta = 1.0$. The MAP criterion was used for classification decisions. Classification performance is presented in table 1. Overall, 99.0% of the texture samples were correctly classified.

[1] $\boldsymbol{u}_{s,r}(n_x,n_y)$ will be treated as complex, with the x and y components as real and imaginary, respectively.

[2] Vector notation (bold) will be used henceforth to indicate all components indexed by s and r at site (n_x,n_y).

[3] The gradient is exactly invertible if boundary conditions are available.

Texture class type	Texture sample type												
	bark	brick	bubbles	grass	leather	pigskin	raffia	sand	straw	water	weave	wood	wool
bark	56	0	0	0	0	0	0	0	0	0	0	0	0
brick	0	56	0	0	0	0	0	0	0	0	0	0	0
bubbles	0	0	56	0	0	0	0	0	0	0	0	0	0
grass	0	0	0	52	1	0	0	0	0	0	0	0	0
leather	0	0	0	4	55	0	0	0	0	0	0	0	0
pigskin	0	0	0	0	0	56	0	1	0	0	0	0	1
raffia	0	0	0	0	0	0	56	0	0	0	0	0	0
sand	0	0	0	0	0	0	0	55	0	0	0	0	0
straw	0	0	0	0	0	0	0	0	56	0	0	0	0
water	0	0	0	0	0	0	0	0	0	56	0	0	0
weave	0	0	0	0	0	0	0	0	0	0	56	0	0
wood	0	0	0	0	0	0	0	0	0	0	0	56	0
wool	0	0	0	0	0	0	0	0	0	0	0	0	55
%correct	100	100	100	93	98	100	100	98	100	100	100	100	98

Table 1. Number of texture samples classified as the texture type indicated in the leftmost column.

6. CONCLUSION

In conclusion, the approach of utilizing an information-conserving model based on Gabor features has been demonstrated to be highly effective in rotation invariant texture classification. The algorithm structure is well suited for parallel and/or pipeline implementations. This approach should be extensible to both texture segmentation and scale invariant classification.

Acknowledgments

The research was partially supported by NSS/IRI-9411330 and by NASA NAGW-3951.

7. REFERENCES

[1] R. L. Kashyap and A. Khotanzad, "A Model-Based Method for Rotation Invariant Texture Classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, pp. 472-481, July 1986.

[2] F. S. Cohen, Z. Fan and M. A. Patel, "Classification of Rotated and Scaled Textured Image Using Gaussian Markov Random Field Models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 2, pp. 192-202, Feb. 1991.

[3] R. Chellappa, R. L. Kashyap and B.S. Manjunath, "Model-Based Texture Segmentation and Classification," *Handbook of Pattern Recognition and Computer Vision.* (ed. C. H. Chen, L. F. Pau and P. F. P. Wang), World Scientific Publishing, 1992.

[4] M. M. Leung and A. M. Peterson, "Scale and Rotation Invariant Texture Classification," *Proc. 26th Asilomar Conference on Signals, Systems and Computers* (Pacific Grove, CA), Oct. 1992.

[5] M. Porat and Y. Zeevi, "Localized Texture Processing in Vision: Analysis and Synthesis in the Gaborian Space," *IEEE Trans. Biomedical Engineering*, vol. 36, no. 1, pp. 115-129, Jan. 1989.

[6] J. You and H. A. Cohen, "Classification and Segmentation of Rotated and Scaled Textured Images Using 'Tuned' Masks," *Pattern Recognition*, vol. 26, no. 2, pp. 245-258, 1993.

[7] D. Gabor, "Theory of Communication," *J. Inst. Elect. Eng.*, vol. 93, pp. 429-457, 1946.

[8] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer.*, vol. 2, pp. 1160-1169, 1985.

[9] A. C. Bovik, "Analysis of Multichannel Narrow-Band Filters for Image Texture Segmentation," *IEEE Trans.Signal Proc.*, vol. 39, no. 9, pp. 2025-2042, Sept. 1991.

[10] M. J. Bastiaans, "A sampling theorem for the complex spectrogram, and Gabor's expansion of a signal in Gaussian elementary signals ," *Optical Engineering*, vol. 20, pp. 594-598, 1981.

[11] M. J. Bastiaans, "Gabor's signal expansion and degrees of freedom of a signal," *Opt. Acta*, vol. 29, pp. 1223-1229, 1982.

[12] S. Marcelja, "Mathematical description of the responses of simple cortical cells," *J. Opt. Soc. Amer.*, vol. 70, pp. 1297-1300, 1980.

[13] A. C. Bovik, "Analysis of Multichannel Narrow-Band Filters for Image Texture Segmentation," *IEEE Trans. on Signal Processing*, vol. 39, no. 9, pp. 2025-2043, Sept. 1991.

[14] P. Brodatz, *Textures: A Photographic album for artists and designers.* New York: Dover, 1966.