Texture for CBIR

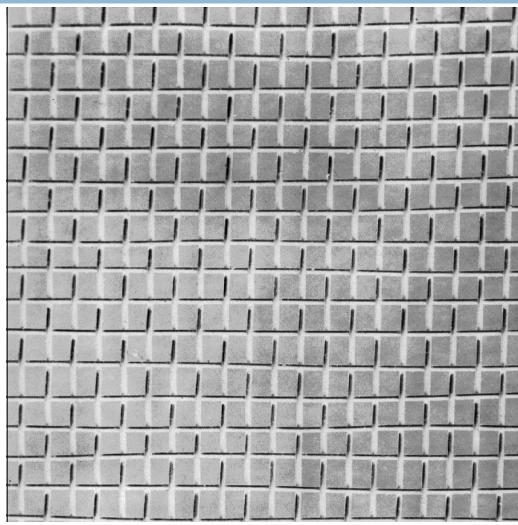
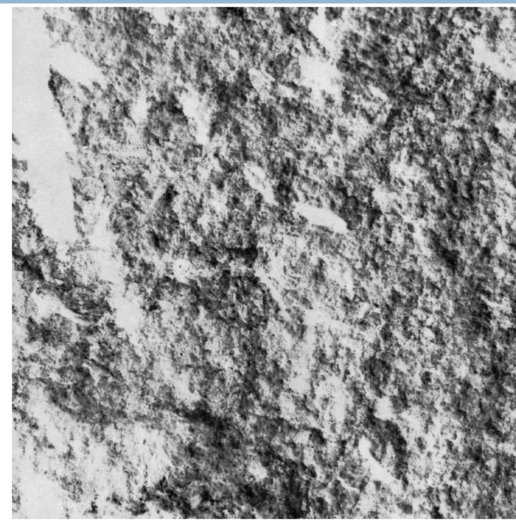
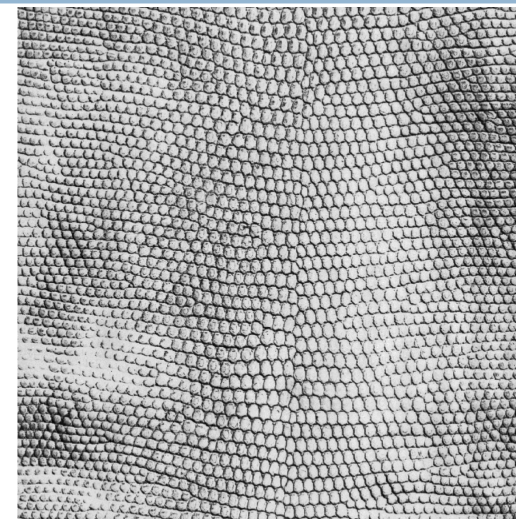
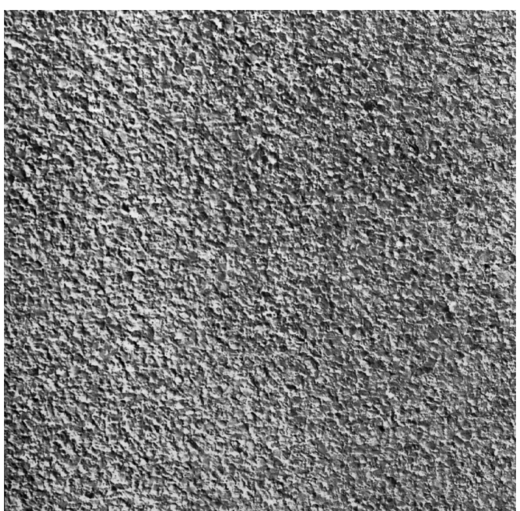
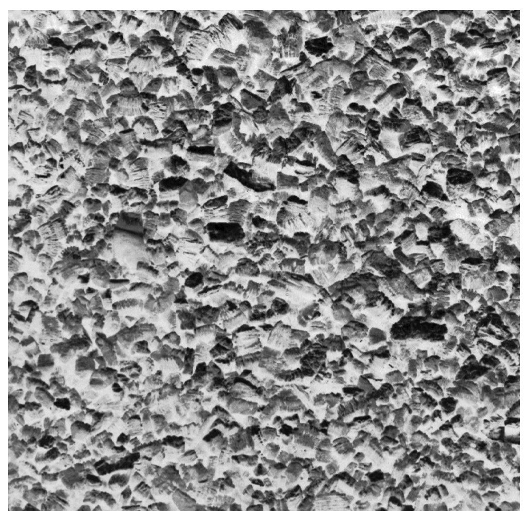
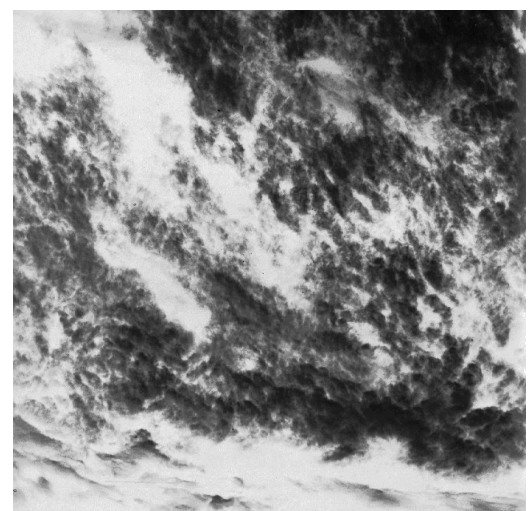
Wei-Ta Chu

Tamura, et al. “Textural feature corresponding to visual perception,” IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-8, no. 6, pp. 460-473, 1978.

Introduction

- Many natural and man-made objects are distinguished by their texture.
- Man-made textures
  - Brick walls, handwoven rugs, …
- Natural textures
  - Water, clouds, sand, grass, lizard skin, …
- Texture can be viewed as some basic primitives (micropatterns) that have specific spatial distributions.
- The spatial distributions can be regular or random.

Examples

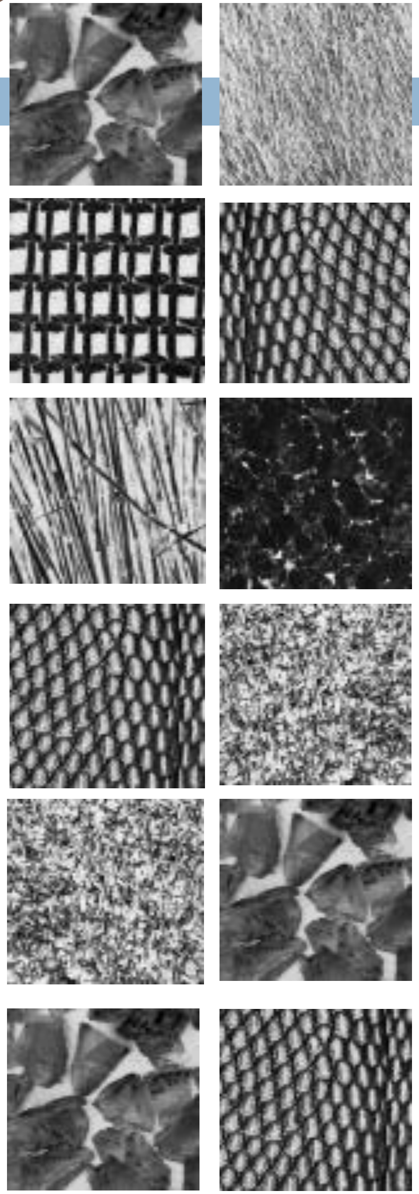
http://www.ux.uis.no/~tranden/brodatz.html

Texture Features

- Structural
  - Describe arrangement of texture elements
  - E.g., texton model, texel model
- Statistical
  - Characterize texture in terms of statistical features
  - E.g., co-occurrence matrix, Tamura features, multiresolution simultaneous autoregressive model, Markov random field
- Spectral
  - Based on analysis in spatial-frequency domain
  - E.g., Fourier transform, Gabor filter, Gaussian derivatives, wavelets

Tamura – Textual Properties

- Coarseness – coarse vs. fine
- Contrast – high contrast vs. low contrast
- Directionality – directional vs. nondirectional
- Line-likeness – line-like vs. blob-like
- Regularity – regular vs. irregular
- Roughness – rough vs. smooth



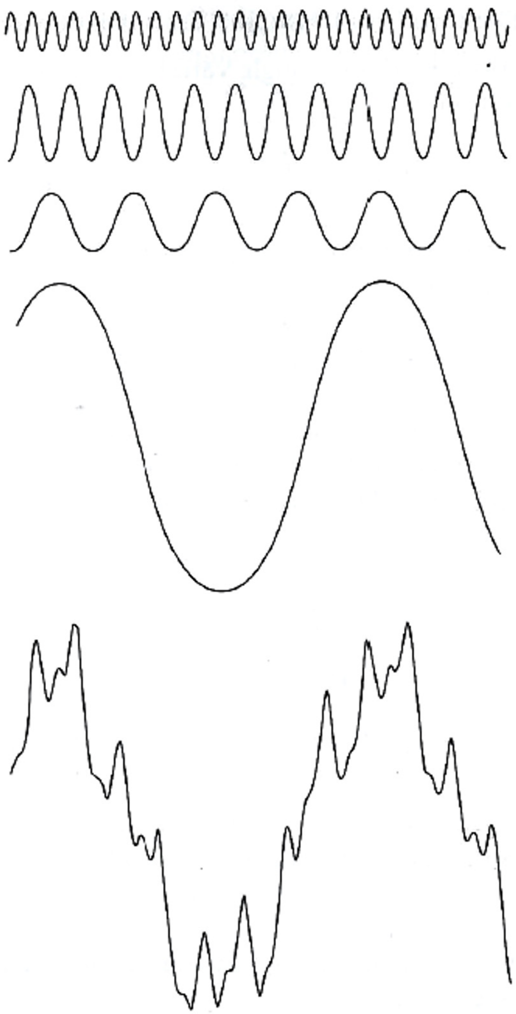
Tamura, et al. "Textural feature corresponding to visual perception,"IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-8, no. 6, pp. 460-473, 1978.

Spectral Texture Features

Wei-Ta Chu

Introduction

- Any function that is periodically repeats can be expressed as the sum of sines and/or cosines of different frequencies, each multiplied by a different coefficient – Fourier series.
- Even functions that are not periodic can be expressed as the integral of sines and/or cosines multiplied by a weighting function. The formulation is the Fourier transform.



Definition of the Fourier Transform

Forward Continuous-Time Fourier Transform

$$X(j\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt$$

Inverse Continuous-Time Fourier Transform

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega) e^{j\omega t} d\omega$$

- The forward transform is an *analysis* integral because it extracts spectrum information
- The inverse transform is a *synthesis* integral because it is used to create the time-domain signal from its spectral information.

Definition of the Fourier Transform

Example: Rectangular Pulse Signals

- Consider the rectangular pulse

$$x(t) = \begin{cases} 1 & -\frac{1}{2}T < t < \frac{1}{2}T \\ 0 & \text{otherwise} \end{cases}$$

- The Fourier transform is

$$X(j\omega) = \int_{-T/2}^{T/2} e^{-j\omega t} dt$$

$$= -\frac{e^{-j\omega t}}{-j\omega}\Bigg|_{-T/2}^{T/2} = -\frac{e^{-j\omega T/2} - e^{j\omega_0 T/2}}{-j\omega}$$

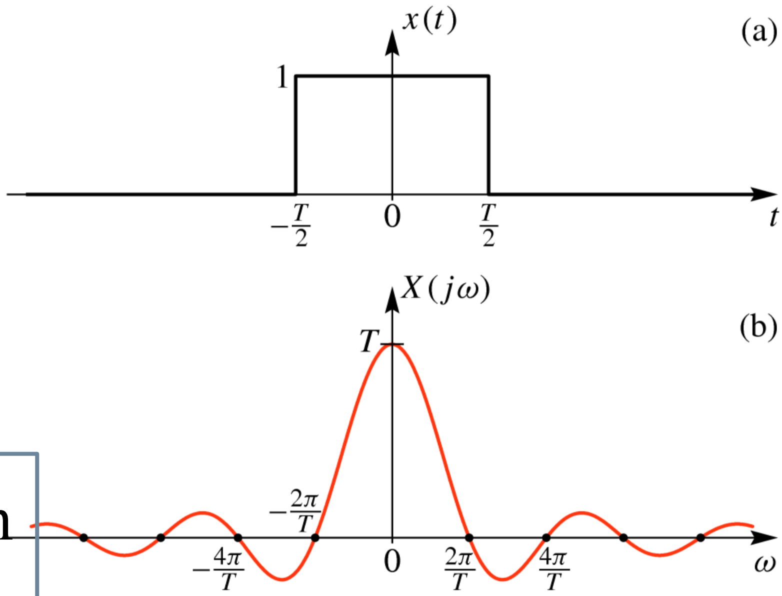
$$= \frac{\sin(\omega T/2)}{\omega/2}$$

Time-Domain		Frequency-Domain
$[u(t+\frac{1}{2}T) - u(t-\frac{1}{2}T)]$	$\overset{\mathcal{F}}{\longleftrightarrow}$	$\frac{\sin(\omega T/2)}{\omega/2}$

Rectangular Pulse Signals

- The Fourier transform of the rectangular pulse signal is called a *sinc function*.
- The formal definition of a sinc function is

$$sinc(\theta) = \frac{\sin(\pi\theta)}{\pi\theta}$$



Time-Domain		Frequency-Domain
$[u(t+\frac{1}{2}T) - u(t-\frac{1}{2}T)]$	$\overset{\mathcal{F}}{\longleftrightarrow}$	$\frac{\sin(\omega T/2)}{\omega/2}$

Discrete Fourier Transform

- One-dimensional DFT

$$F(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(x)e^{-j2\pi ux/M}$$ for $u = 0, 1, 2, \ldots, M\text{-}1$

$$f(x) = \sum_{u=0}^{M-1} F(u)e^{j2\pi ux/M}$$ for $x = 0, 1, 2, \ldots, M\text{-}1$

In order to compute $F(u)$, we start by substituting $u = 0$ in the exponential term and then summing for all values of x. We then substitute $u = 1$ …
Like $f(x)$, the transform is a discrete quantity, and it has the same number of components as $f(x)$.

Discrete Fourier Transform

$$F(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(x)e^{-j2\pi ux/M}$$

Euler's formula: $e^{j\theta} = \cos\theta + j\sin\theta$

$$F(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(x)[\cos 2\pi ux/M - j\sin 2\pi ux/M]$$

Each term of the Fourier transform (the value of $F(u)$) is composed of the sum of all values of the function $f(x)$.
The domain (values of u) over which the values of $F(u)$ range is called the frequency domain, because u determines the frequency of the components of the transform. Each of the M terms of $F(u)$ is called a frequency component of the transform.

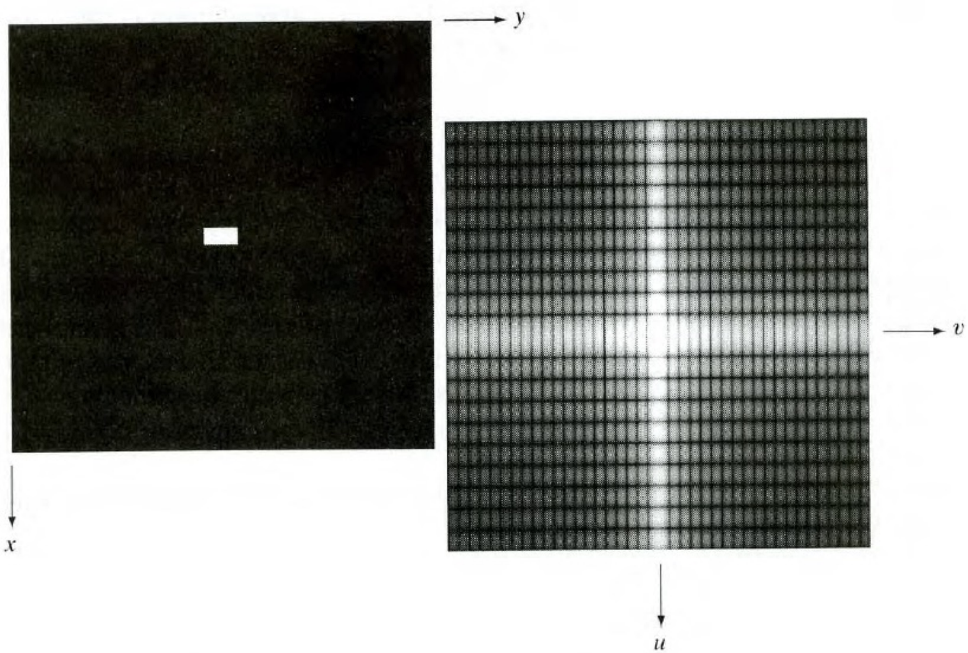
Discrete Fourier Transform

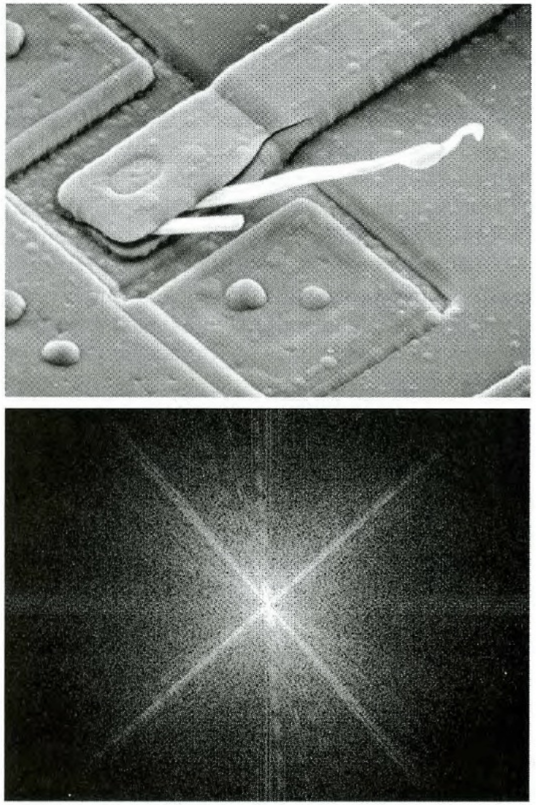
- Two-dimensional DFT

$$F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(ux/M+vy/N)}$$

$$f(x,y) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} F(u,v)e^{j2\pi(ux/M+vy/N)}$$

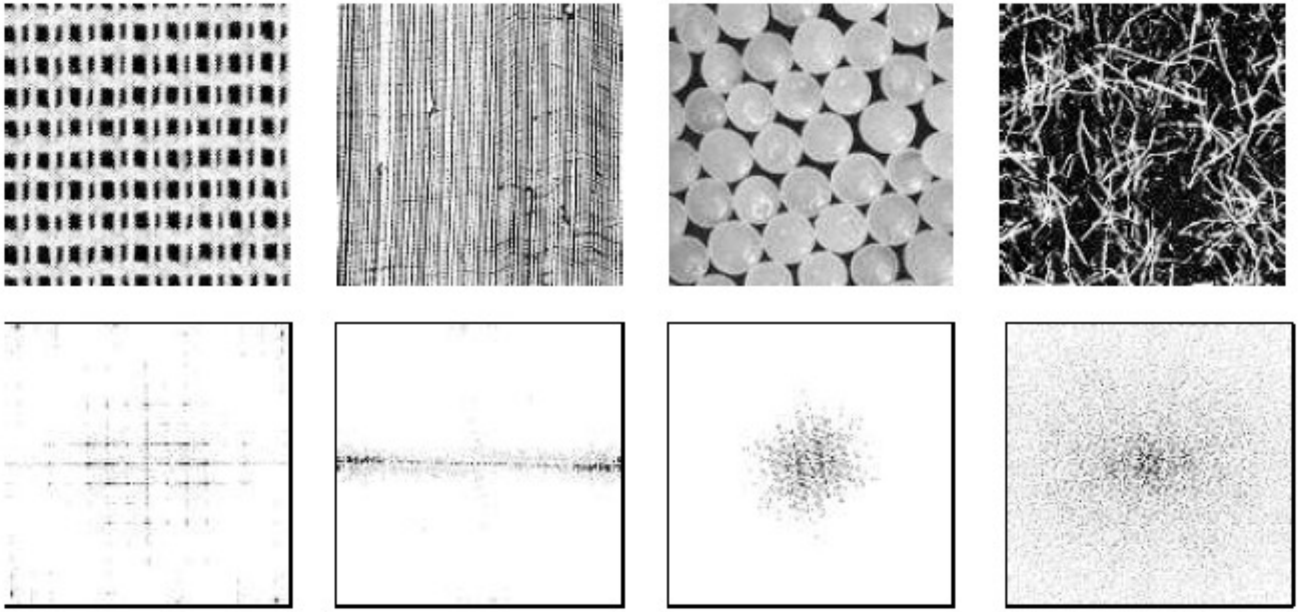
Images in Frequency Domain





Gonzalez and Woods, Chapter 4 of Digital Image Processing, Prentice-Hall, 2001.

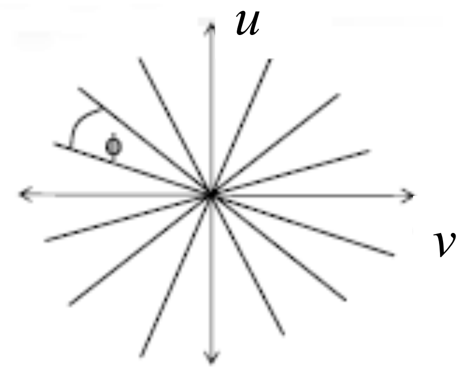
Images and Their FT



(a) structured (b) oriented (c) granular (d) random

Frequency Domain Features

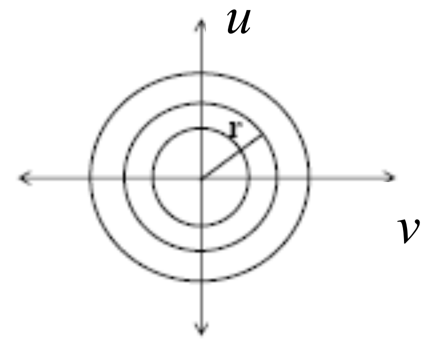
- Fourier domain energy distribution
- Angular features (directionality)



$$V_{\theta_1\theta_2} = \int\int |F(u,v)|^2 dudv$$

$$\theta_1 \leq tan^{-1}\frac{v}{u} \leq \theta_2$$

- Radial features (coarseness)



$$V_{r_1r_2} = \int\int |F(u,v)|^2 dudv$$

$$r_1 \leq u^2 + v^2 < r_2$$

- Uniform division may not be the best

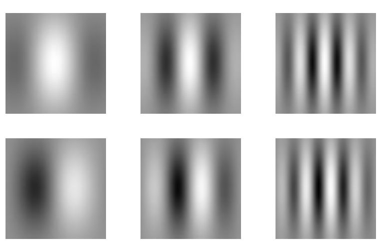
Gabor Texture

- The Gabor representation has been shown to be optimal in the sense of minimizing the joint two-dimensional uncertainty in space and frequency.
- These filters can be considered as orientation and scale tunable edge and line (bar) detectors.
- The statistics of these microfeatures in a given region are often used to characterize the underlying texture information.

B.S. Manjunath and W.Y. Ma, “Texture features for browsing and retrieval of image data,” IEEE Trans. on PAMI, vol. 18, no. 8, 1996, pp. 837-842.

Gabor Texture

- Fourier coefficients depend on the entire image (Global) → we lose spatial information
- Objective: local spatial frequency analysis
- Gabor kernels: looks like Fourier basis multiplied by a Gaussian
  - Gabor filters come in pairs: symmetric and anti-symmetric
- We need to apply a number of Gabor filters at different scales, orientations, and spatial frequencies



Symmetric kernel

Anti-symmetric kernel

Gabor Texture

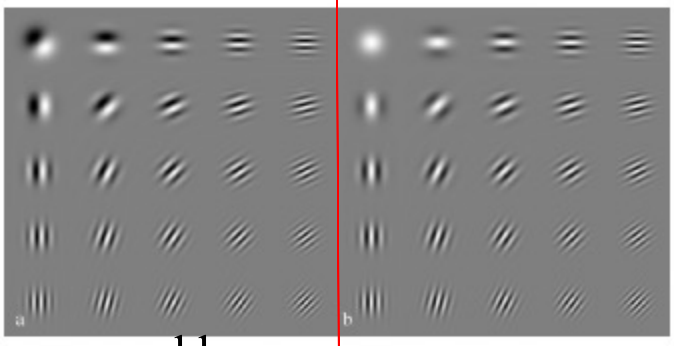
- Image $I(x,y)$ convoluted with Gabor filters h_{mn} (totally M x N)

$$W_{mn}(x,y) = \int I(x_1, y_1) h_{mn}(x - x_1, y - y_1) dx_1 dy_1$$

- Using first and 2nd moments for each scale and orientations

$$\mu_{mn} = \int \int |W_{mn}(x,y)| dx dy$$

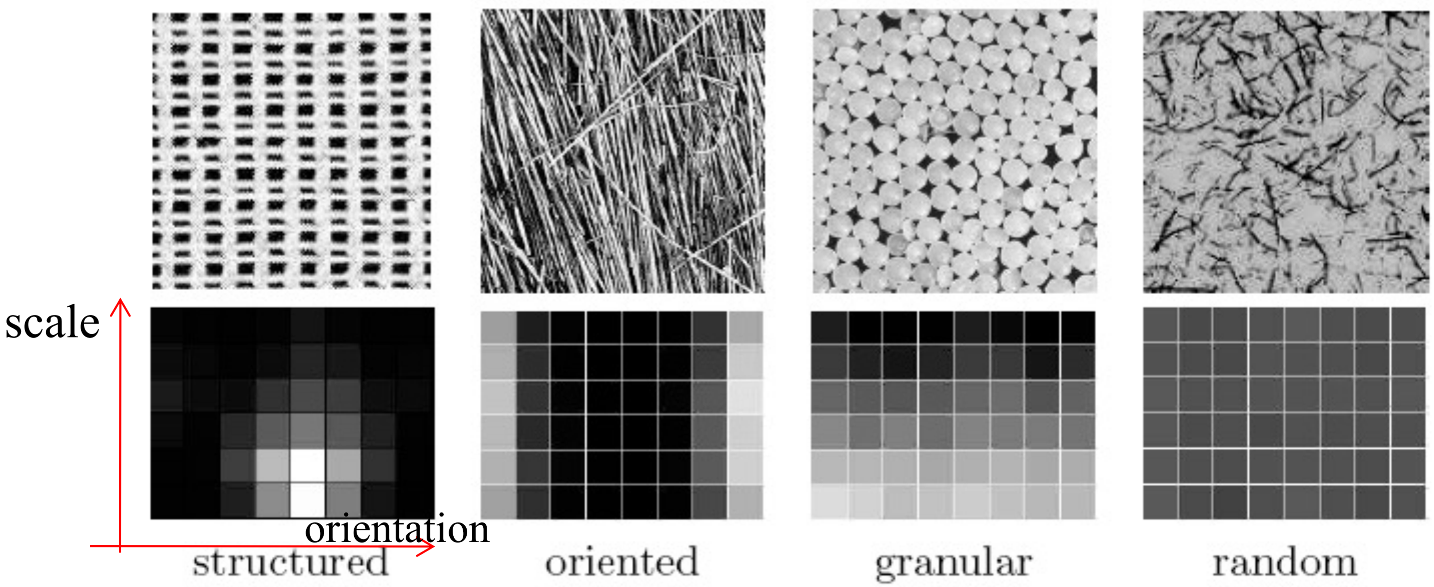
$$\sigma_{mn} = \sqrt{\int \int (|W_{mn}(x,y)| - \mu_{mn})^2 dx dy}$$



- Features: e.g., 4 scales, 6 orientations → 48 dimensions

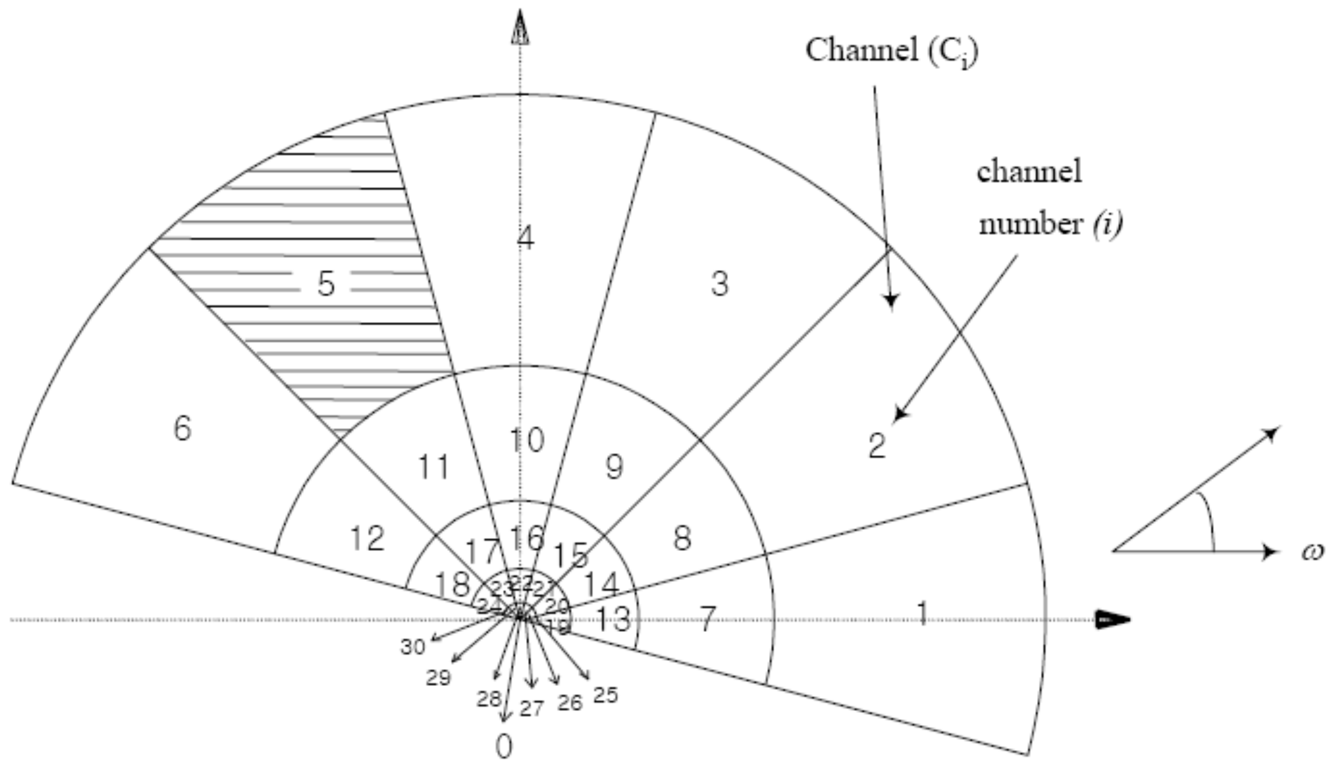
$$\bar{v} = [\mu_{00}, \sigma_{00}, \mu_{01}, \dots, \mu_{35}, \sigma_{35}]$$

Gabor Texture



- Arranging the mean energy in a 2D form
  - structured: localized pattern
  - oriented (or directional): column pattern
  - granular: row pattern
  - random: random pattern

Homogeneous Texture Descriptor



Frequency plane partition is uniform along the angular direction (30º), non-uniform along the radial direction (on an octave scale)

B.S. Manjunath and W.Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Trans. on PAMI, vol. 18, no. 8, 1996, pp. 837-842.

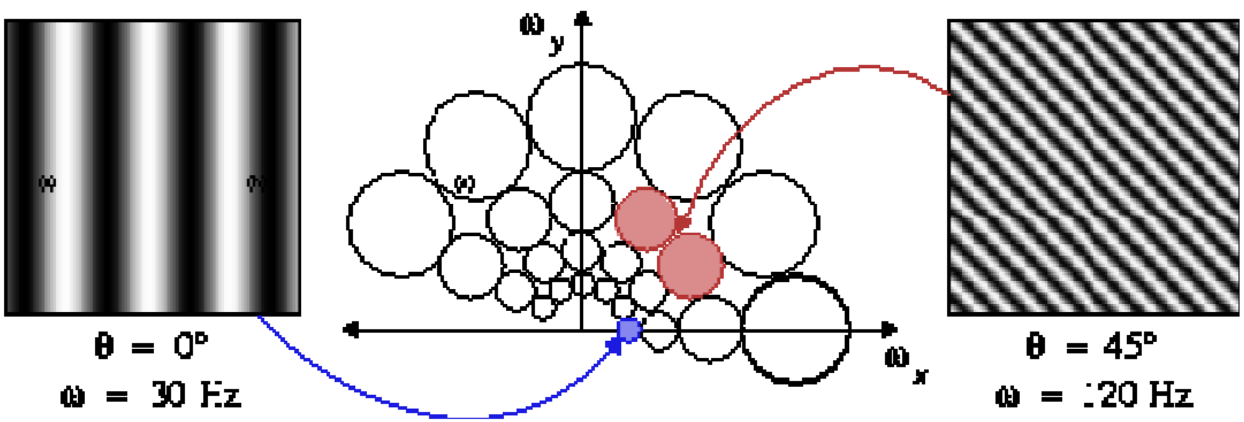
Gabor Function

- On the top of the feature channel, the following 2D Gabor function (modulated Gaussian) is applied to each individual channels.

$$G_{s,r}(\omega, \theta) = exp\left[\frac{-(\omega-\omega_s)^2}{2\sigma_s^2}\right] exp\left[\frac{-(\theta-\theta_r)^2}{2\tau_r^2}\right]$$

  - Equivalent to weighting the Fourier transform coefficients of the image with a Gaussian centered at the frequency channels as defined above

- Each channel filters a specific type of texture



Homogeneous Texture Descriptor

- Partition the frequency domain into 30 channels (modeled by a 2D Gabor function)
- Computing the energy and energy deviation for each channel
- Computing the mean and standard deviation of frequency coefficients
- HTD = $\{f_{DC}, f_{SD}, e_1, e_2, \ldots, e_{30}, d_1, d_2, \ldots, d_{30}\}$

f_{DC} and f_{SD} are the mean and standard deviation of the image
e_i and d_i are the mean energy and energy deviation of the corresponding ith channel

Distance Measure

Consider two image patterns i and j, and let $\vec{f}^{(i)}$ and $\vec{f}^{(j)}$ represent the corresponding feature vectors. Then the distance between the two patterns in the feature space is defined to be

$$d(i,j) = \sum_m \sum_n d_{mn}(i,j),$$

where

$$d_{mn}(i,j) = \left| \frac{\mu_{mn}^{(i)} - \mu_{mn}^{(j)}}{\alpha(\mu_{mn})} \right| + \left| \frac{\sigma_{mn}^{(i)} - \sigma_{mn}^{(j)}}{\alpha(\sigma_{mn})} \right|. \qquad (8)$$

$\alpha(\mu_{mn})$ and $\alpha(\sigma_{mn})$ are the standard deviations of the respective features over the entire database, and are used to normalize the individual feature components.

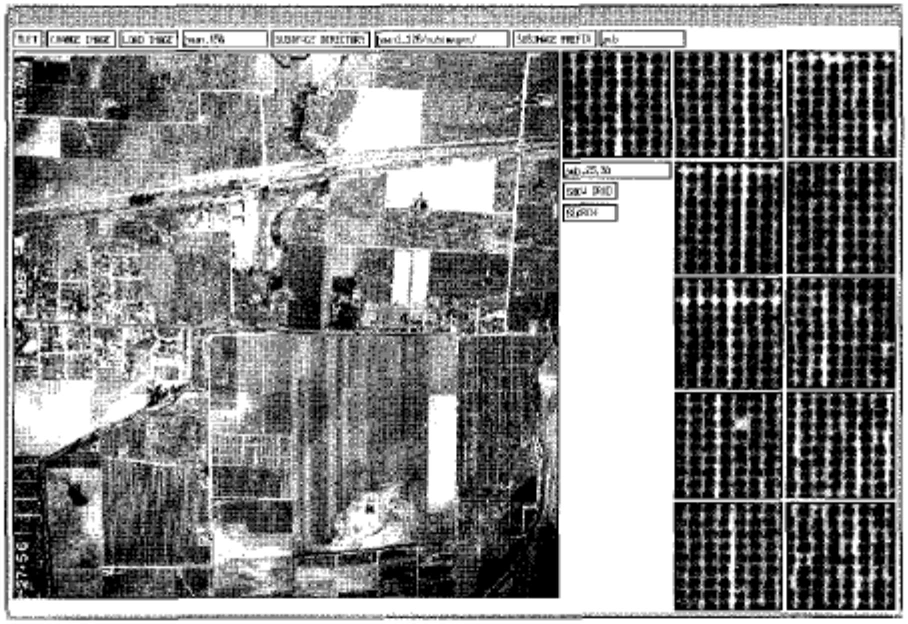
B.S. Manjunath and W.Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Trans. on PAMI, vol. 18, no. 8, 1996, pp. 837-842.

Resources: http://vision.ece.ucsb.edu/texture/feature.html

On-line demo: http://vision.ece.ucsb.edu/texture/mpeg7/index.html

Example: Browsing Satellite Images

- Find a vegetation patch that looks like this region



B.S. Manjunath and W.Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Trans. on PAMI, vol. 18, no. 8, 1996, pp. 837-842.

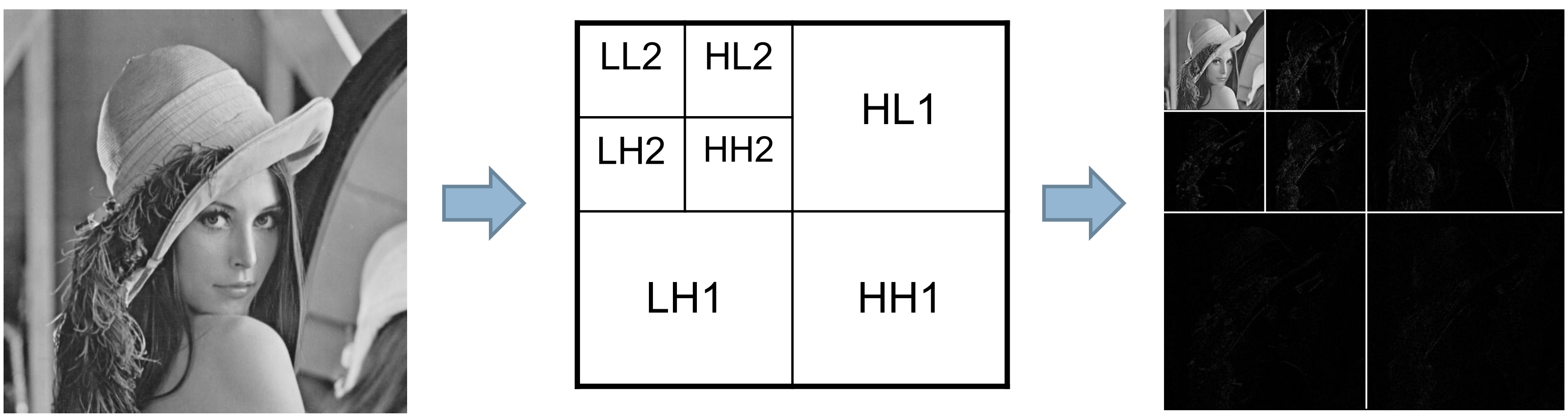
Example: Browsing Satellite Images

- (b) parts of highway
- (c) region containing some buildings (center of the image toward the left)
- (d) a number marked on the image (lower left corner)

(b) (c) (d)

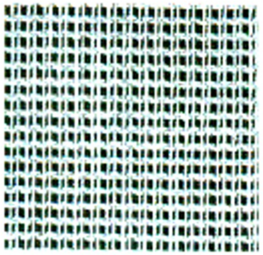
Wavelet Features

- Wavelet transforms refer to the decomposition of a signal with a family of basis functions with recursive filtering and subsampling
- At each level, it decomposes a 2D signal into four subbands, which are often referred to as LL, LH, HL, HH (L=low, H=high)

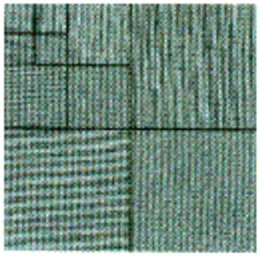


Wavelet Features

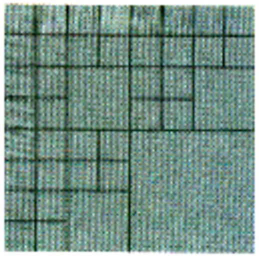
- Using the mean and standard deviation of the energy distribution in each subband at each level.
- PWT (Pyramid-structured wavelet transform)
  - Recursively decompose the LL band
  - Results in 30-dimensional feature vector (3x3x2+2=30)
- TWT (Tree-structured wavelet transform)
  - Some information appears in the middle frequency channels – decomposition is not restricted to the LL band
  - Results in 40x2 = 80 dimensional feature vector



Original image



PWT



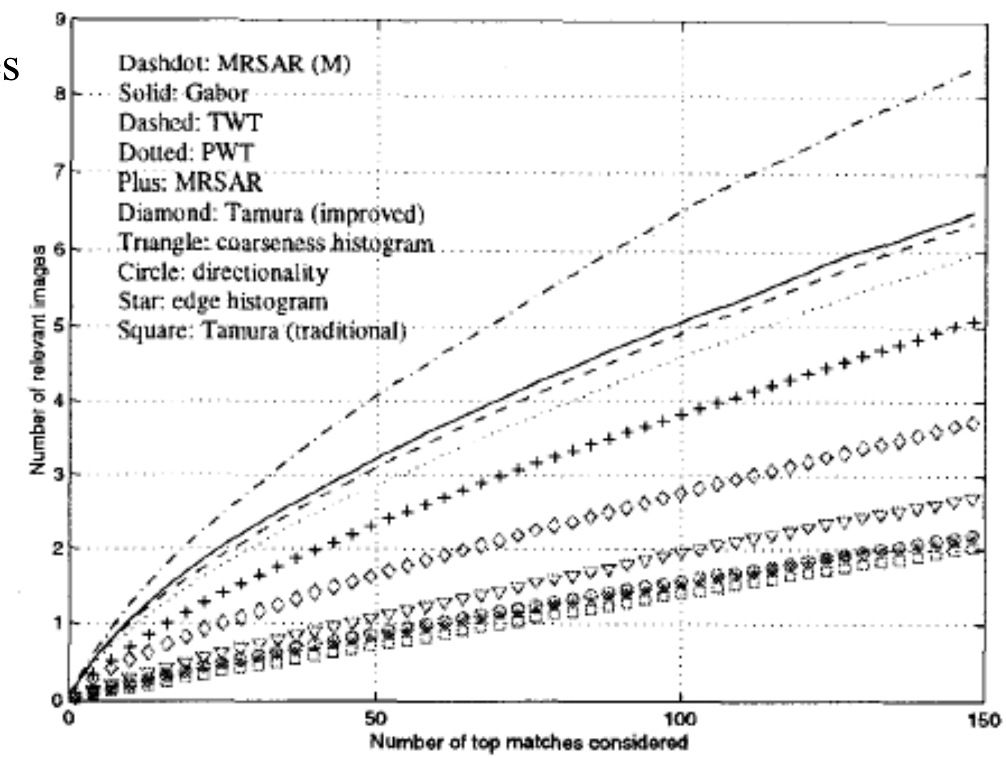
TWT

T. Chang and C.C.J. Kuo, "Texture analysis and classification with tree-structure wavelet transform," IEEE Trans. On Image Processing, vol. 2, no. 4, 1993, pp. 429-441.

Performance Comparison

- Retrieval performance of different texture features for the Corel photo databases.
- L_1 distance is used to computing the dissimilarity between images.
- For the MRSAR, Mahalanobis distance is used.

#relevant images



MRSAR (M)

Gabor
TWT
PWT

MRSAR

Tamura (improved)

Coarseness histogram
Directionality
Edge histogram
Tamura (traditional)

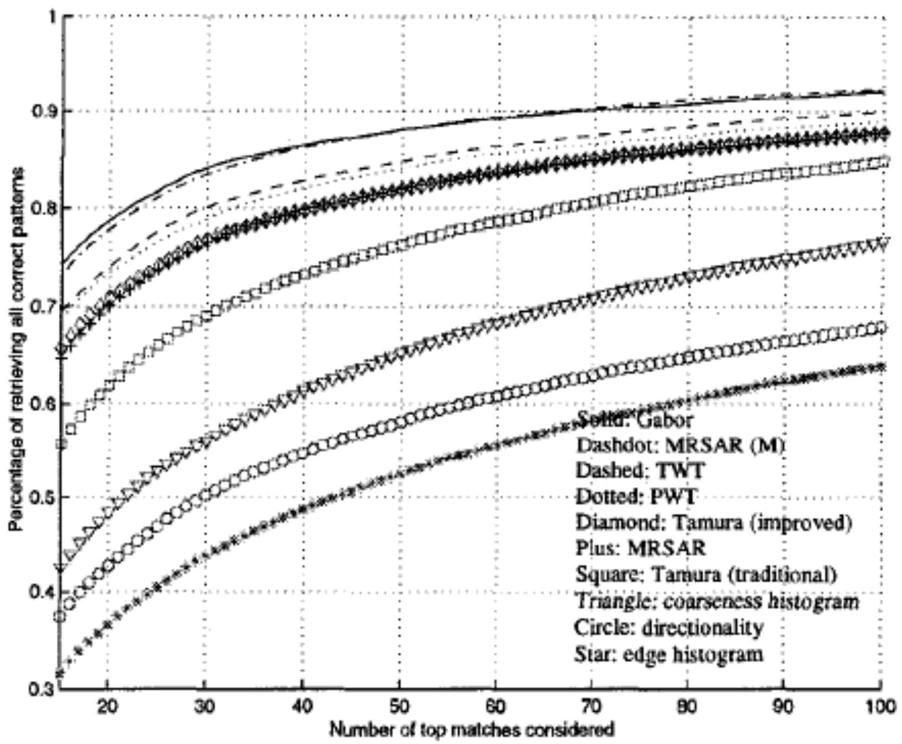
#top matches considered

Manjunath and Ma, Chapter 12 of Image Database: Search and Retrieval of Digital Imagery, edited by V. Castelli and L.D. Bergman, John Wiley & Sons, 2002.

Performance Comparison

- Retrieval performance of different texture features for the Brodatz texture image set.

Percentage of retrieving all correct patterns



#top matches considered

Gabor
MRSAR (M)
TWT
PWT
Tamura (improved)
MRSAR
Tamura (traditional)
Coarseness histogram
Directionality
Edge histogram