Image feature extraction and segmentation

Inel 5046
Pattern Recognition

Prof. Vidya Manian

Introduction

- Detection of discontinuities
- Edge linking and boundary detection
- Thresholding
- Region based segmentation
- Texture feature extraction
- Image Classification

Detection of discontinuities

- Detection of discontinuities using masks
- The response of the mask at any point in the image is $R = w_1z_1+w_2z_2+...+w_9z_9$

$$= \sum_{i-1}^{9} w_i z_i$$

- A general 3x3 mask

$\mathbf{w_1}$	$\mathbf{w_2}$	$\mathbf{w_3}$
$\mathbf{w_4}$	$\mathbf{w_5}$	$\mathbf{w_6}$
$\mathbf{w_7}$	$\mathbf{w_8}$	$\mathbf{w_9}$

Detection of Discontinuities

- A mask used for detecting isolated points different from a constant background

-1	**-1**	**-1**
-1	**8**	**-1**
-1	**-1**	**-1**

- Using the above mask a point is detected at the location on which the mask is centered if |R|>T where T is a nonnegative threshold, and

$$R = \sum_{i-1}^{9} w_i z_i$$

Line detection

Line masks

-1	**-1**	**-1**
2	**2**	**2**
-1	**-1**	**-1**

Horizontal

-1	**-1**	**2**
-1	**2**	**-1**
2	**-1**	**-1**

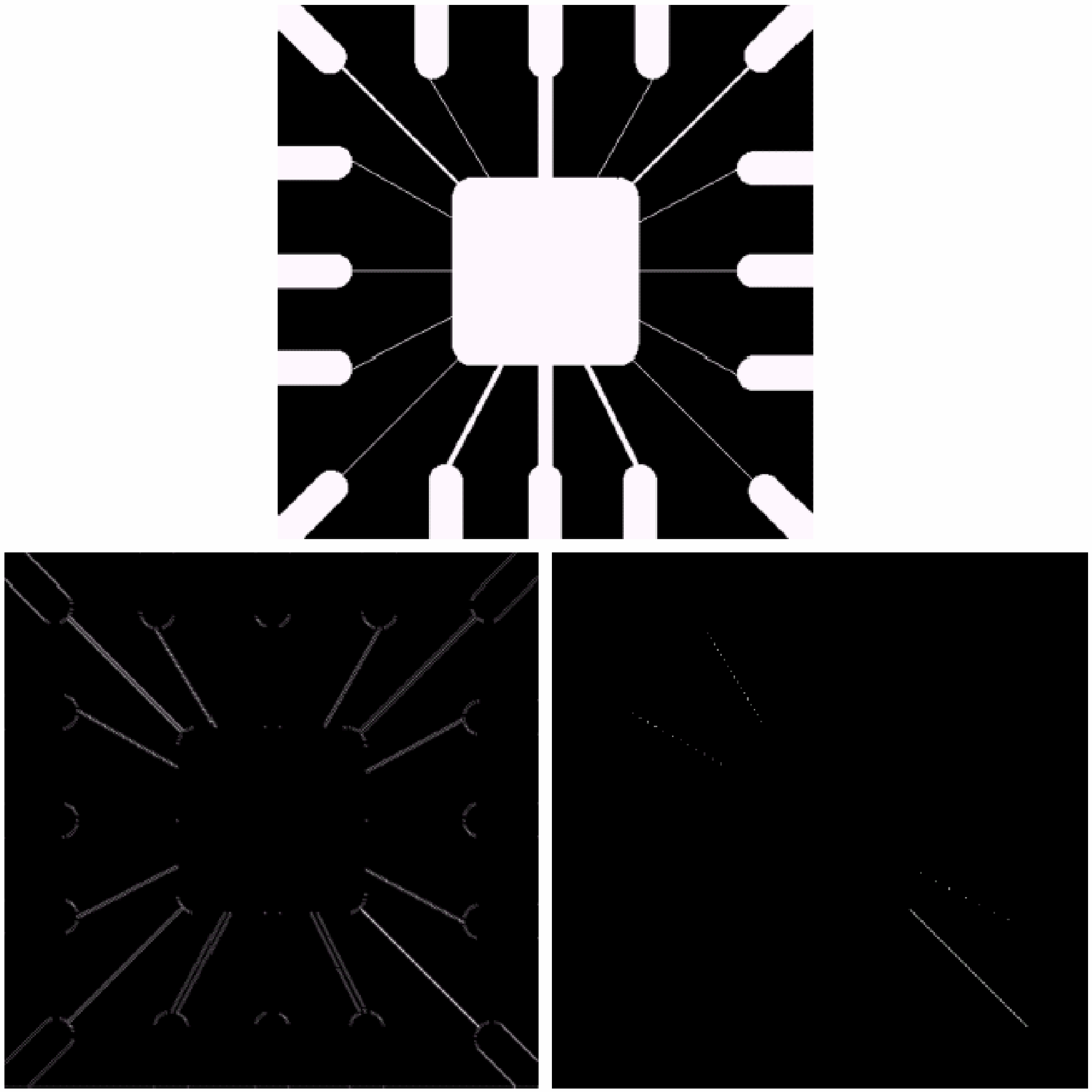
+45 degrees

-1	**2**	**-1**
-1	**2**	**-1**
-1	**-2**	**-1**

Vertical

2	**-1**	**-1**
-1	**2**	**-1**
-1	**-1**	**2**

-45 degrees



a
b c

Illustration of line detection.
(a) Binary wire-bond mask.
(b) Absolute value of result after processing with −45° line detector.
(c) Result of thresholding image (b).

Edge Detection

- Gradient operators - the gradient is used for image differentiation
- The gradient of an image f(x,y) at location (x,y) is the vector

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

- The magnitude of the gradient is $\nabla f = mag(\nabla f) = [G_x^2 + G_y^2]^{1/2}$
- The direction of the gradient is $\alpha(x, y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$

Edge Detection

- The Sobel operator provides both a differencing and a smoothing effect
- The smoothing effect is a particularly attractive feature as derivatives enhance noise
- $G_x=(z_7+2z_8+z_9)-(z_1+2z_2+z_3)$
- $G_y=(z_3+2z_6+z_9)-(z_1+2z_4+z_7)$ where z's are the gray levels of the pixels

Mask to compute G_x

-1	**-1**	**-1**
0	**0**	**0**
1	**2**	**1**

Mask to compute G_y

-1	**0**	**1**
-2	**0**	**2**
-1	**0**	**1**

	a
b	c
d	e
f	g

A 3 × 3 region of an image (the z's are gray-level values) and various masks used to compute the gradient at point labeled z_5.

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

−1	0
0	1

0	−1
1	0

Roberts

−1	−1	−1
0	0	0
1	1	1

−1	0	1
−1	0	1
−1	0	1

Prewitt

−1	−2	−1
0	0	0
1	2	1

−1	0	1
−2	0	2
−1	0	1

Sobel

0	1	1
−1	0	1
−1	−1	0

−1	−1	0
−1	0	1
0	1	1

Prewitt

0	1	2
−1	0	1
−2	−1	0

−2	−1	0
−1	0	1
0	1	2

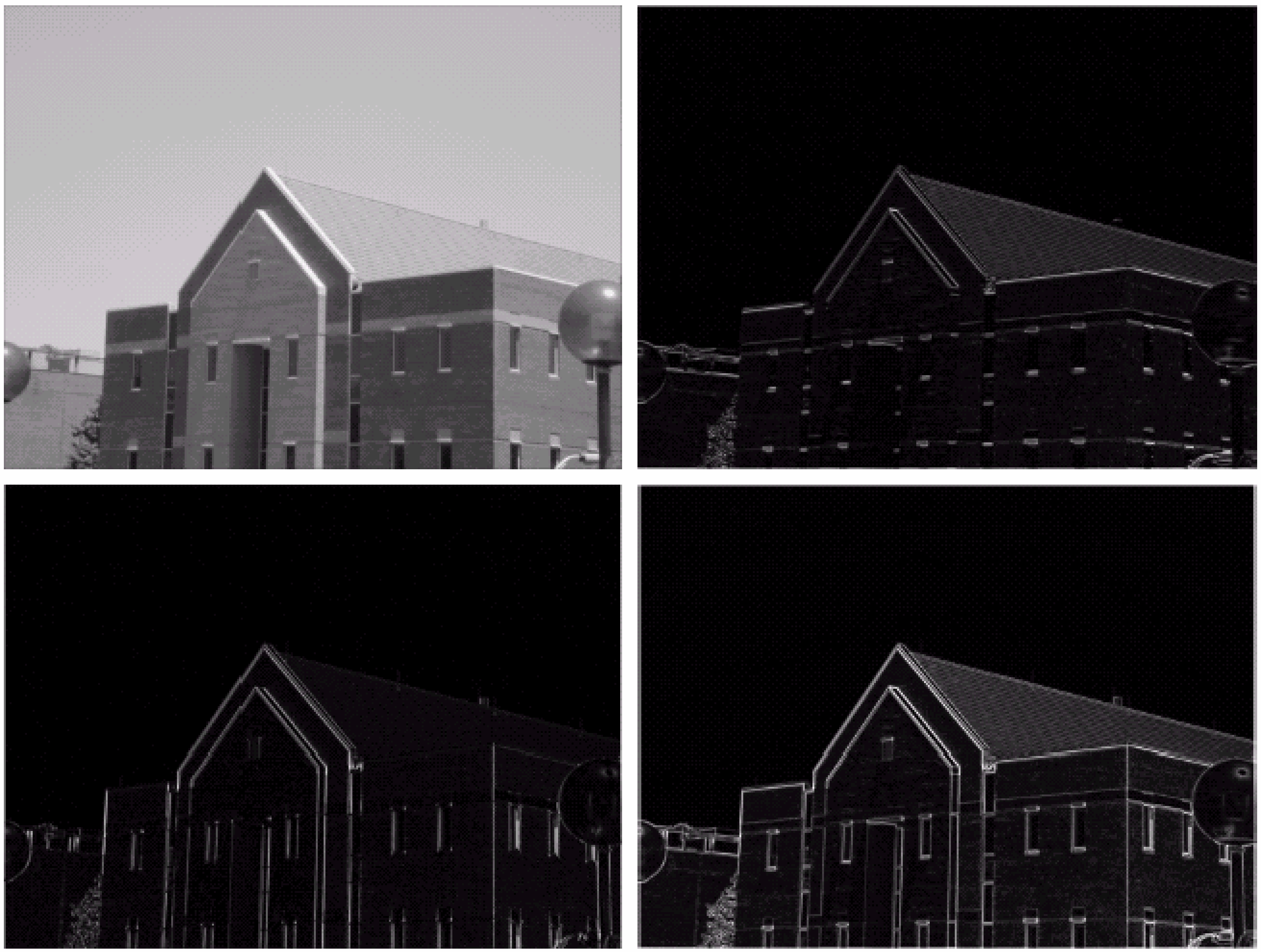
Sobel

a b
c d

Prewitt and Sobel masks for detecting diagonal edges.

a b
c d

(a) Original image. (b) $|G_x|$, component of the gradient in the x-direction. (c) $|G_y|$, component in the y-direction. (d) Gradient image, $|G_x| + |G_y|$.



a b
c d

Same sequence as in Fig. 10.10, but with the original image smoothed with a 5 × 5 averaging filter.

a b

Diagonal edge detection. (a) Result of using the mask in Fig. 10.9(c). (b) Result of using the mask in Fig. 10.9(d). The input in both cases was Fig. 10.11(a).

Edge Detection

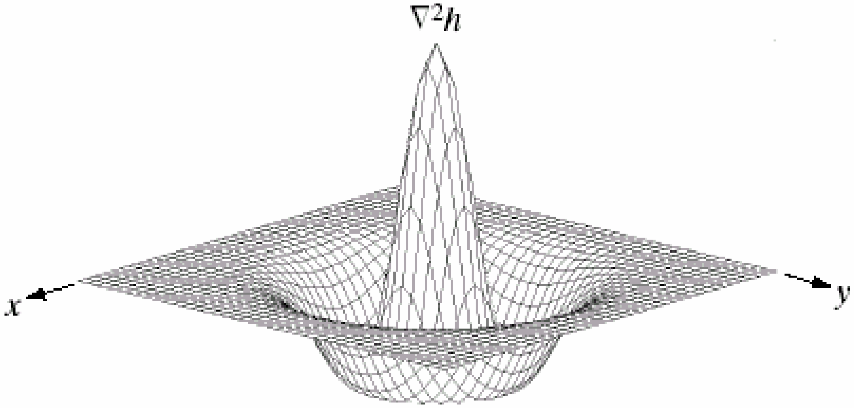
- The Laplacian of a 2-D function f(x,y) is a second-order derivative defined as

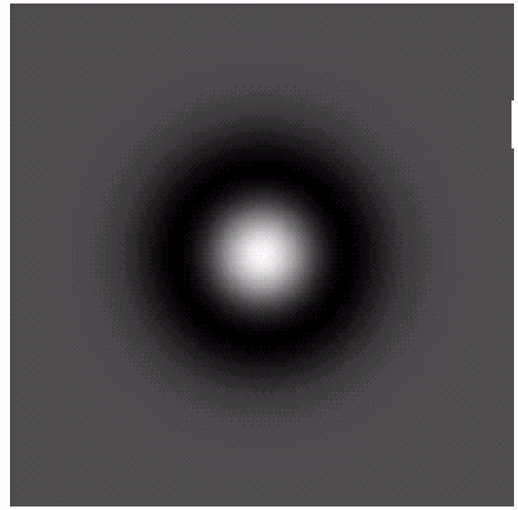
$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

- It is implemented in digital form as a 3x3 region, and is given by $\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$

<u>Mask used to compute Laplacian</u>

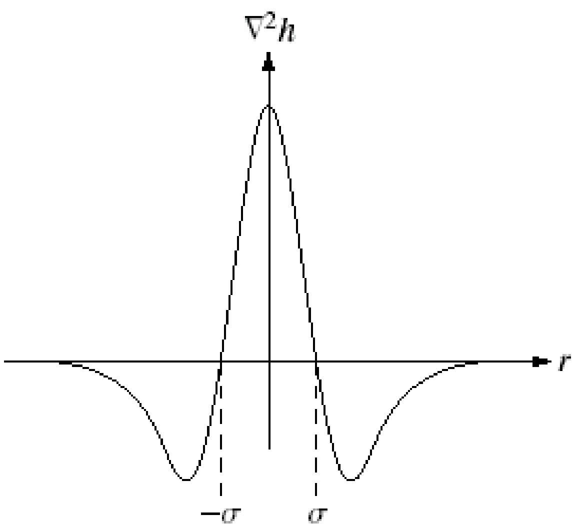
0	**-1**	**0**
-1	**4**	**-1**
0	**-1**	**0**





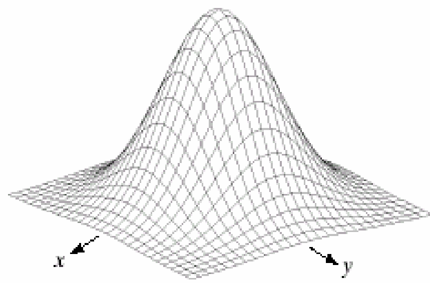
a	b
c	d

Laplacian of a Gaussian (LoG).
(a) 3-D plot.
(b) Image (black is negative, gray is the zero plane, and white is positive).
(c) Cross section showing zero crossings.
(d) 5×5 mask approximation to the shape of (a).



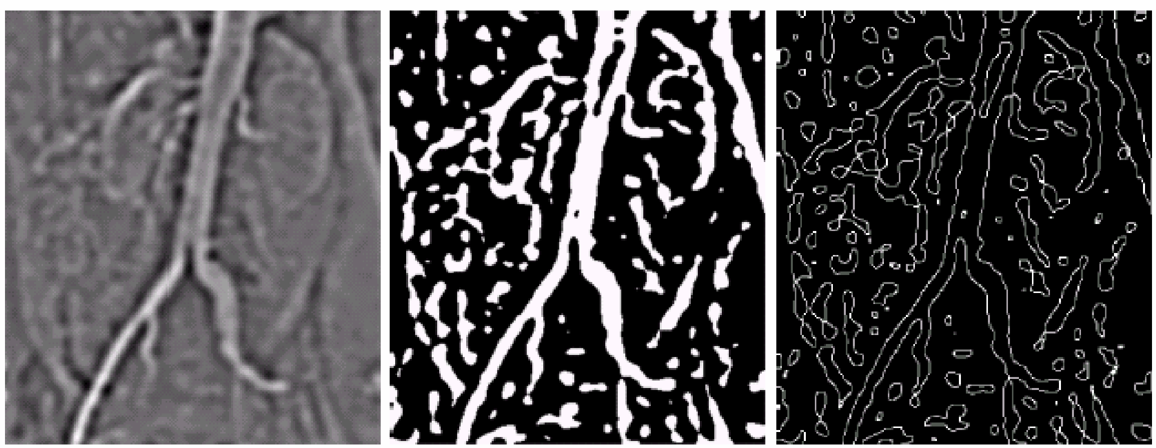
0	0	−1	0	0
0	−1	−2	−1	0
−1	−2	16	−2	−1
0	−1	−2	−1	0
0	0	−1	0	0

Sobel gradient



−1	−1	−1
−1	8	−1
−1	−1	−1

Spatial Gaussian
Smoothing function,
Laplacian mask



LoG
Thresholded LoG
Zero crossings

Orthogonal Masks

- Orthogonal masks are used for formulating multimasks

Basis of edge subspace

1	$\sqrt{2}$	-1
0	0	0
-1	$-\sqrt{2}$	-1

1	0	-1
$\sqrt{2}$	0	$-\sqrt{2}$
1	0	-1

0	-1	$\sqrt{2}$
1	0	-1
$-\sqrt{2}$	1	0

$\sqrt{2}$	-1	0
-1	0	1
0	1	$-\sqrt{2}$

Orthogonal Masks

Basis of line subspace

0	1	0
-1	0	-1
0	-1	0

-1	0	1
0	0	0
1	0	-1

1	-2	1
-2	4	-2
1	-2	1

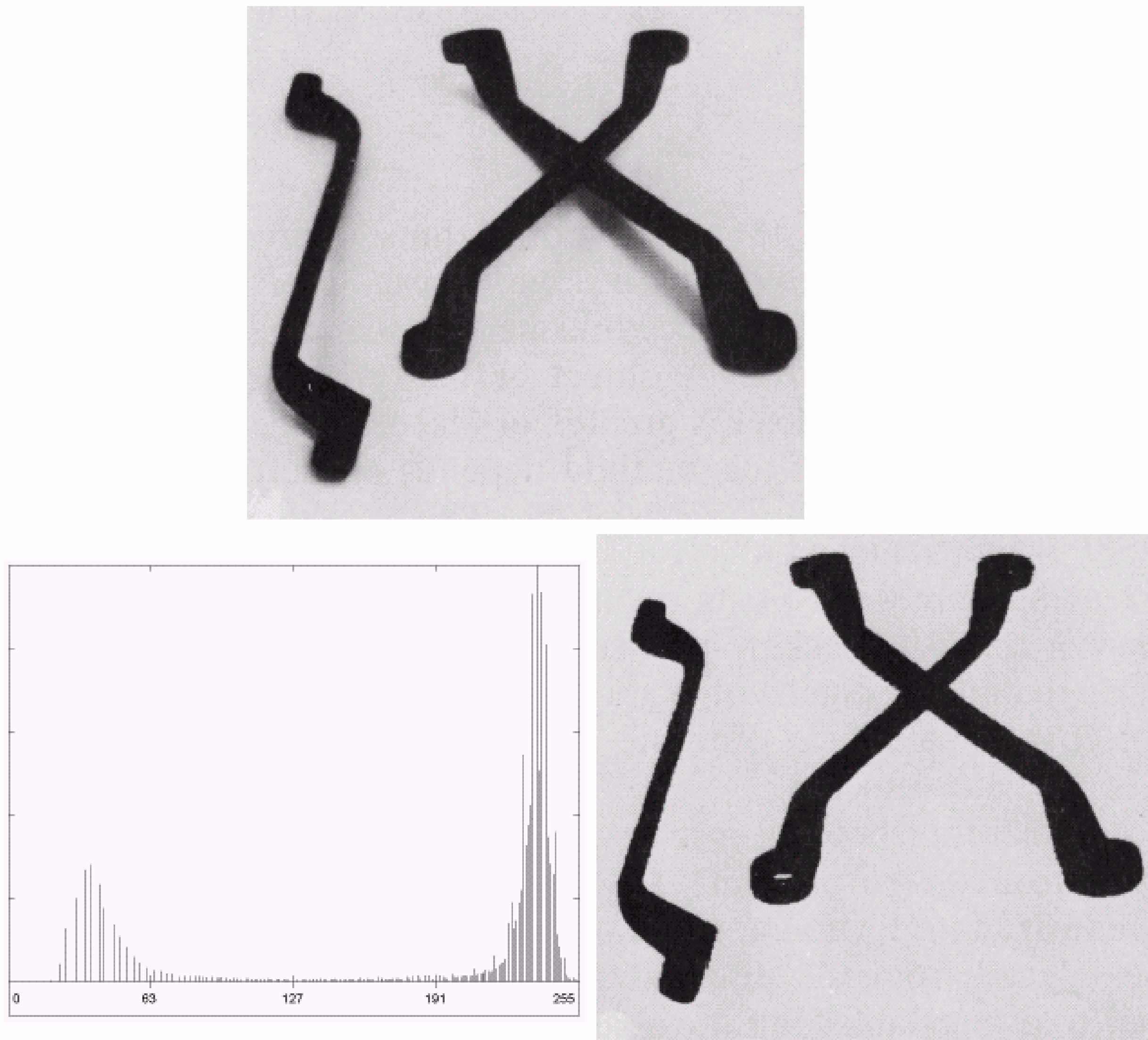
-2	1	-2
1	4	1
-2	1	-2

"Average" subspace

1	1	1
1	1	1
1	1	1

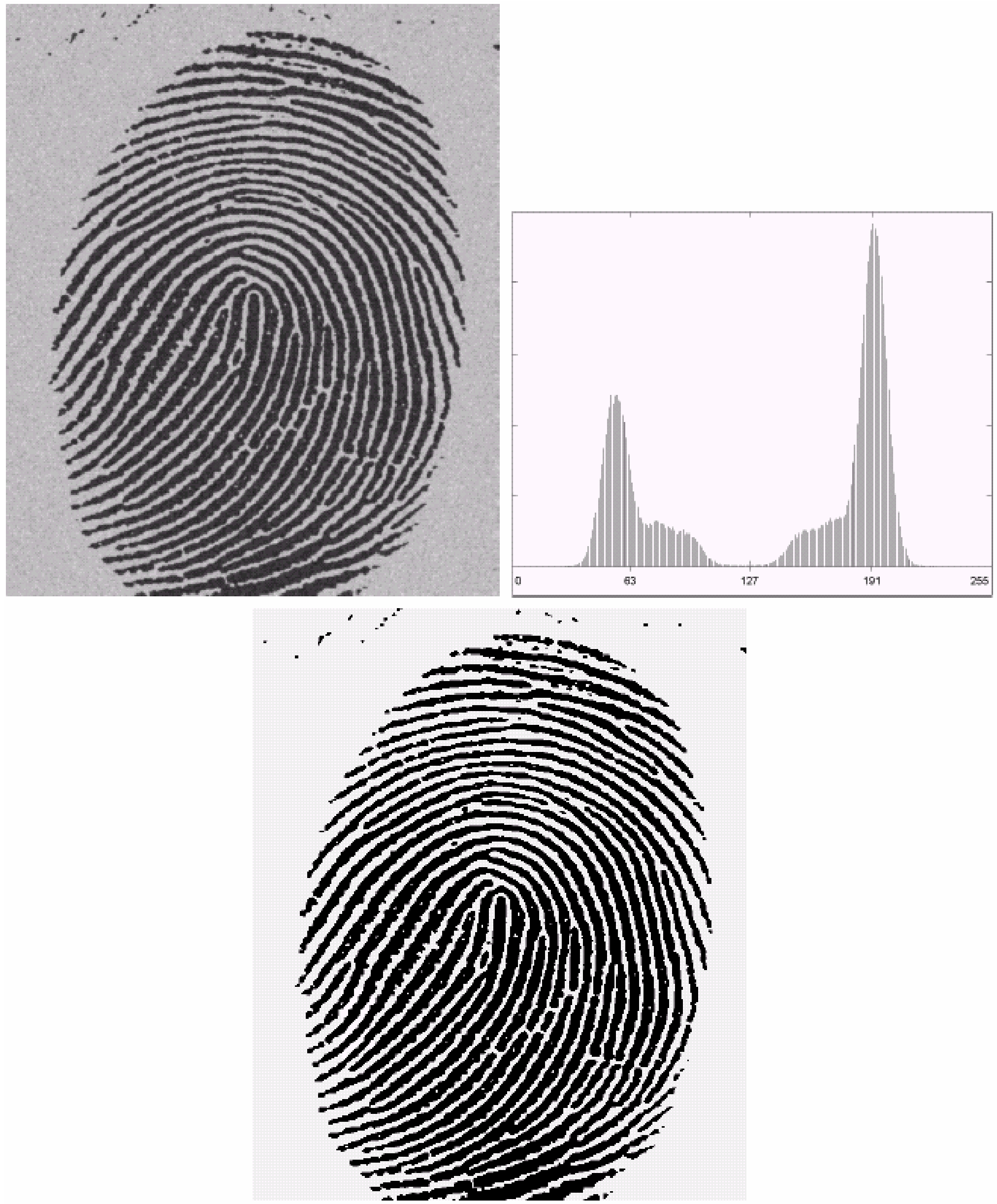
Thresholding

- Object and background pixels
- A threshold T is selected that separates the two modes
- Any point (x,y) for which $f(x,y)>T$ is called an object point; otherwise the point is the background point
- Use of histogram
- If the histogram has three dominant modes two thresholds T1 and T2 are selected
- Eg., two types of light objects on a dark background
- A point (x,y) belongs to: one object class if
- $T1< f(x,y) \leq T2$, to the other object class if $f(x,y) >T2$, to the background if $f(x,y) \leq T1$



a
b c

(a) Original image. (b) Image histogram. (c) Result of global thresholding with T midway between the maximum and minimum gray levels.

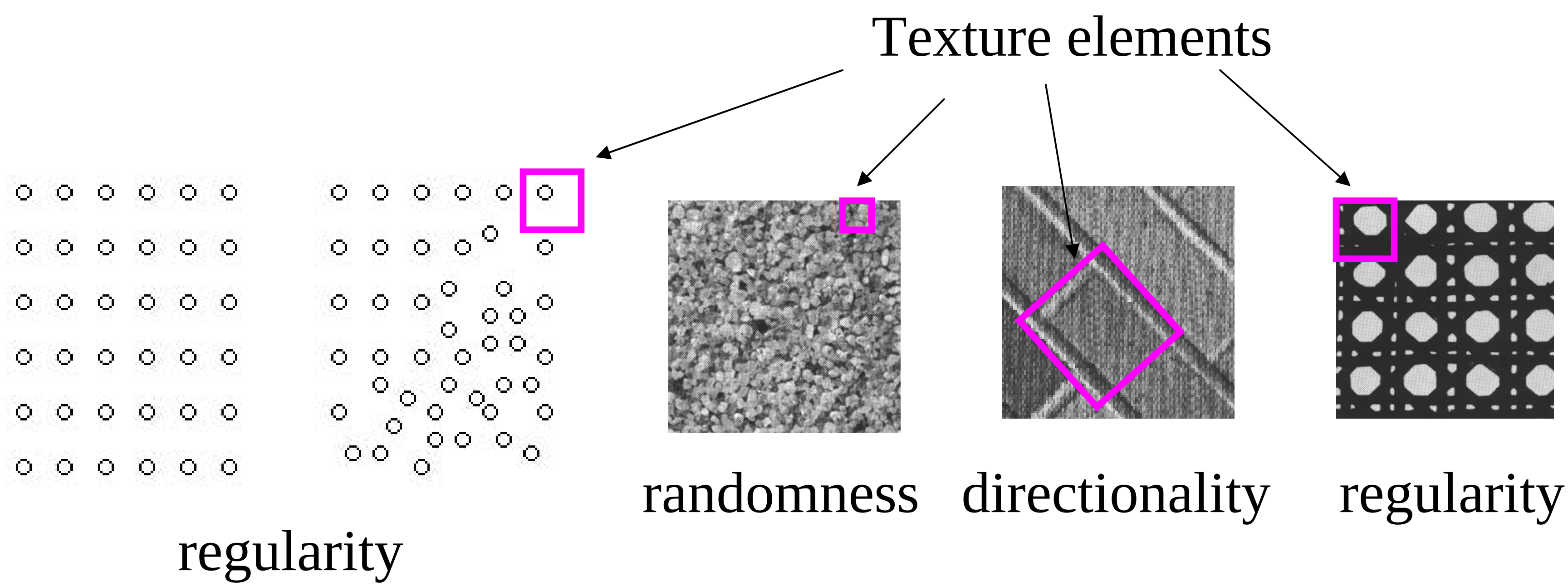


a b
c

(a) Original image. (b) Image histogram. (c) Result of segmentation with the threshold estimated by iteration. (Original courtesy of the National Institute of Standards and Technology.)

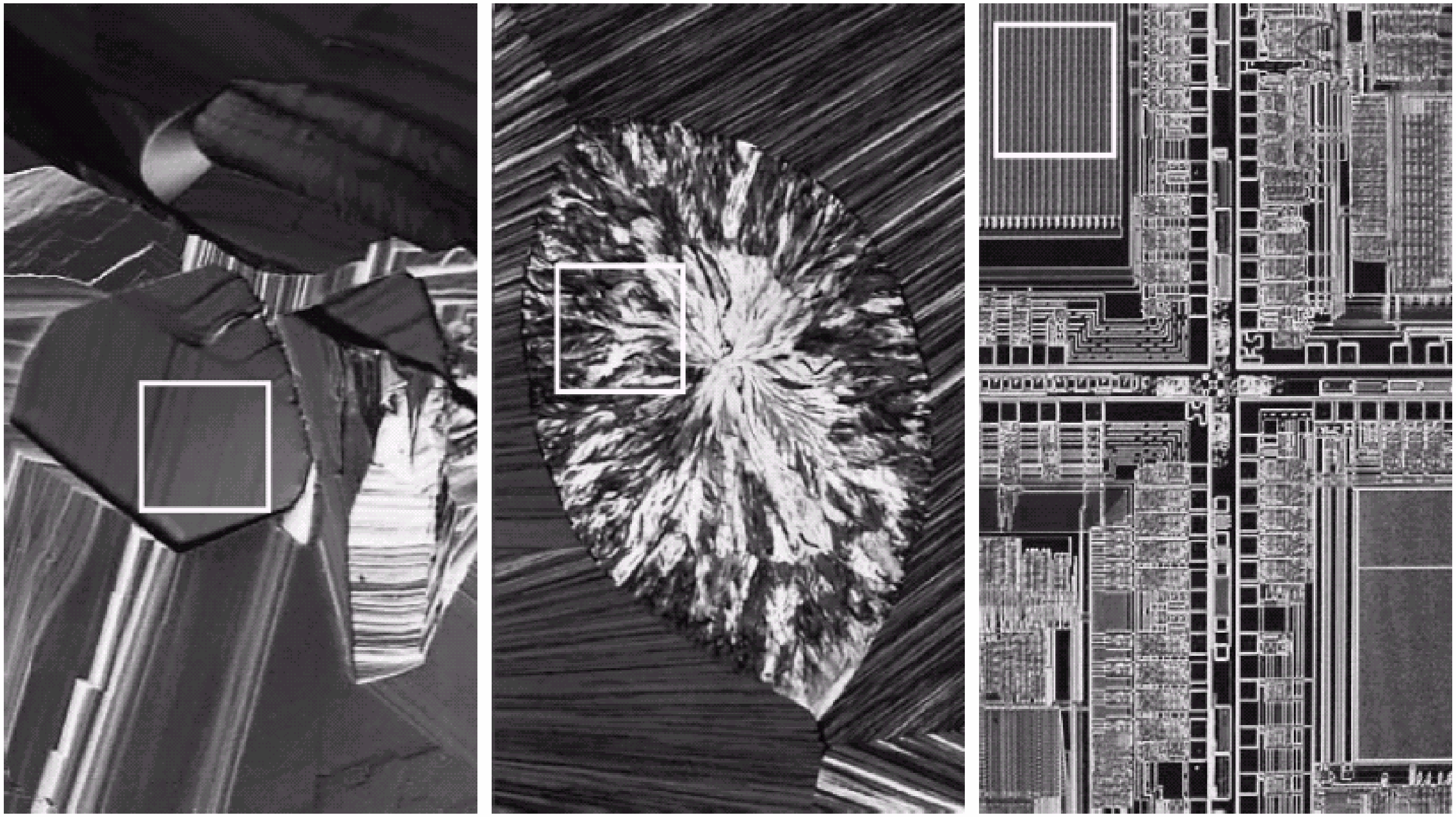
Texture

- Defined as the structural pattern of surfaces *which is homogeneous in spite of fluctuations in brightness and color*
- Most important visual cue in identifying regions
- Main categories: regularity, randomness and directionality



Texture features

- Statistical approach
  - From first order statistics (histogram)
  - From second order statistics (grey-level cooccurrence matrix)
- Spectral approach -Fourier power spectrum
- Gabor filter
- Multiresolution wavelet



a b c

The white squares mark, from left to right, smooth, coarse, and regular textures. These are optical microscope images of a superconductor, human cholesterol, and a microprocessor. (Courtesy of Dr. Michael W. Davidson, Florida State University.)

Texture measures for the subimages

Texture	Mean	Standard deviation	R (normalized)	Third moment	Uniformity	Entropy
Smooth	82.64	11.79	0.002	−0.105	0.026	5.434
Coarse	143.56	74.63	0.079	−0.151	0.005	7.783
Regular	99.72	33.73	0.017	0.750	0.013	6.674

Gray-Level Co-occurrence Matrix

- The gray-level co-occurrence matrix P[i,j] is defined by specifying a displacement vector d=(dx,dy) and counting all pairs of pixels separated by d having gray levels i and j

2	1	2	0	1
0	2	1	1	2
0	1	2	2	0
1	2	2	0	1
2	0	1	0	1

5x5 image with 3 gray levels 0, 1, and 2

i | **j**

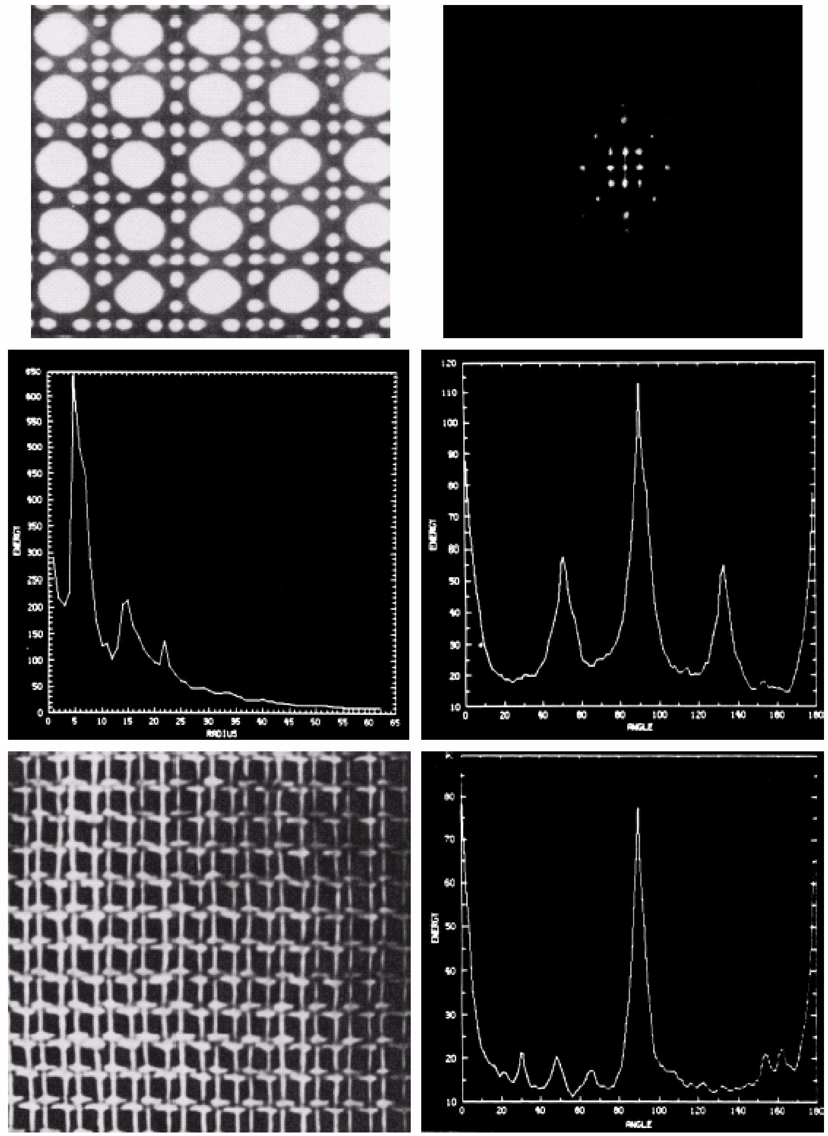
P[i,j]

i \ *j*	**0**	**1**	**2**
0	0	4	1
1	1	1	4
2	4	2	1

The co-occurrence matrix for d=(1,1)

Co-occurrence Features

- The co-occurrence matrices can be computed for 0, 45, 90 and 135 degress and at distances 1,2,3,...
- Statistical features are computed from the co-occurrence matrices
- *Entropy* is a feature which measures the randomness of gray-level distribution
- The *angular second-moment* is a measure of homogeneity of the image
- The *contrast* is a measure of the amount of local variations present in the image
- The *correlation* is a measure of linearity in the image



a	b
c	d
e	f

(a) Image showing periodic texture. (b) Spectrum. (c) Plot of $S(r)$. (d) Plot of $S(\theta)$. (e) Another image with a different type of periodic texture. (f) Plot of $S(\theta)$. (Courtesy of Dr. Dragana Brzakovic, University of Tennessee.)

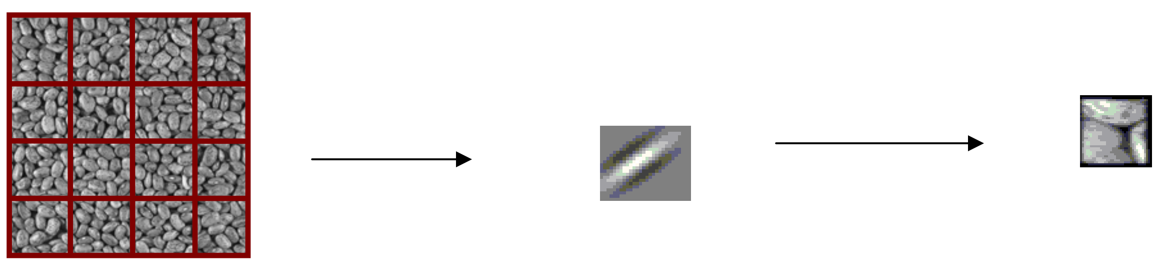
Gabor filter feature extraction

- Bank of Gabor filters in the spatial domain

$$f_{a,b}(x,y) = \frac{1}{2\pi\sigma_a^2}\exp\left\{-\frac{x^2+y^2}{2\sigma_a^2}\right\}\cos\left(2\pi\left(\omega_a x\cos\theta_b + \omega_a y\sin\theta_b\right)\right)$$

- Response from convolving Image sample with filter

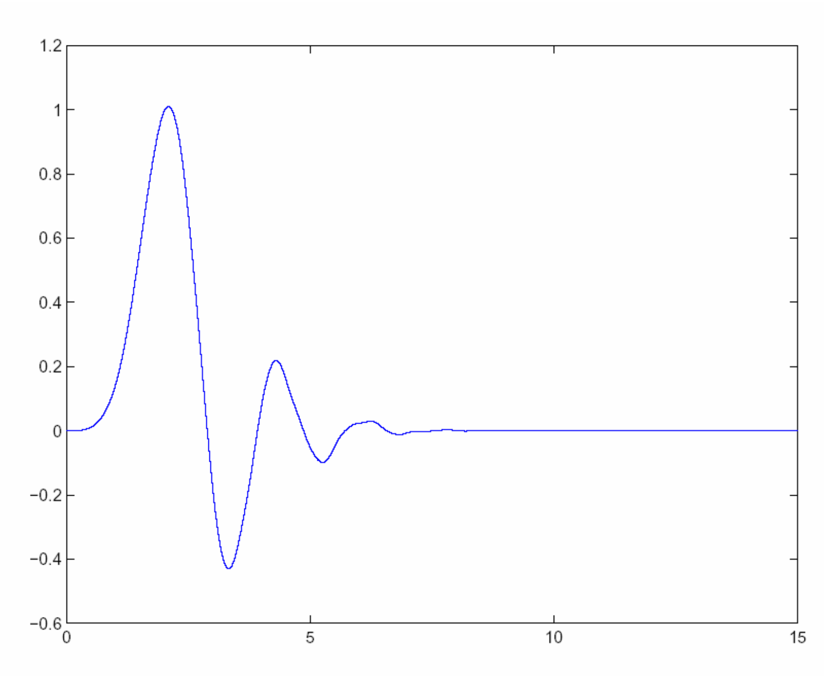
$$G_{dab}(x,y) = X_d(x,y) * f_{ab}(x,y)$$



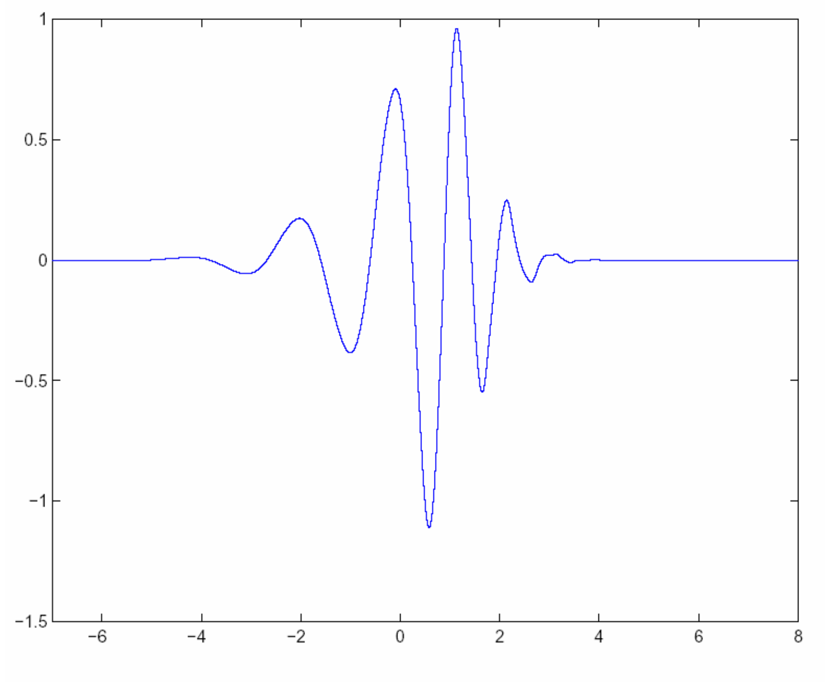
Sequential samples Gabor filter Response

Daubechies wavelet

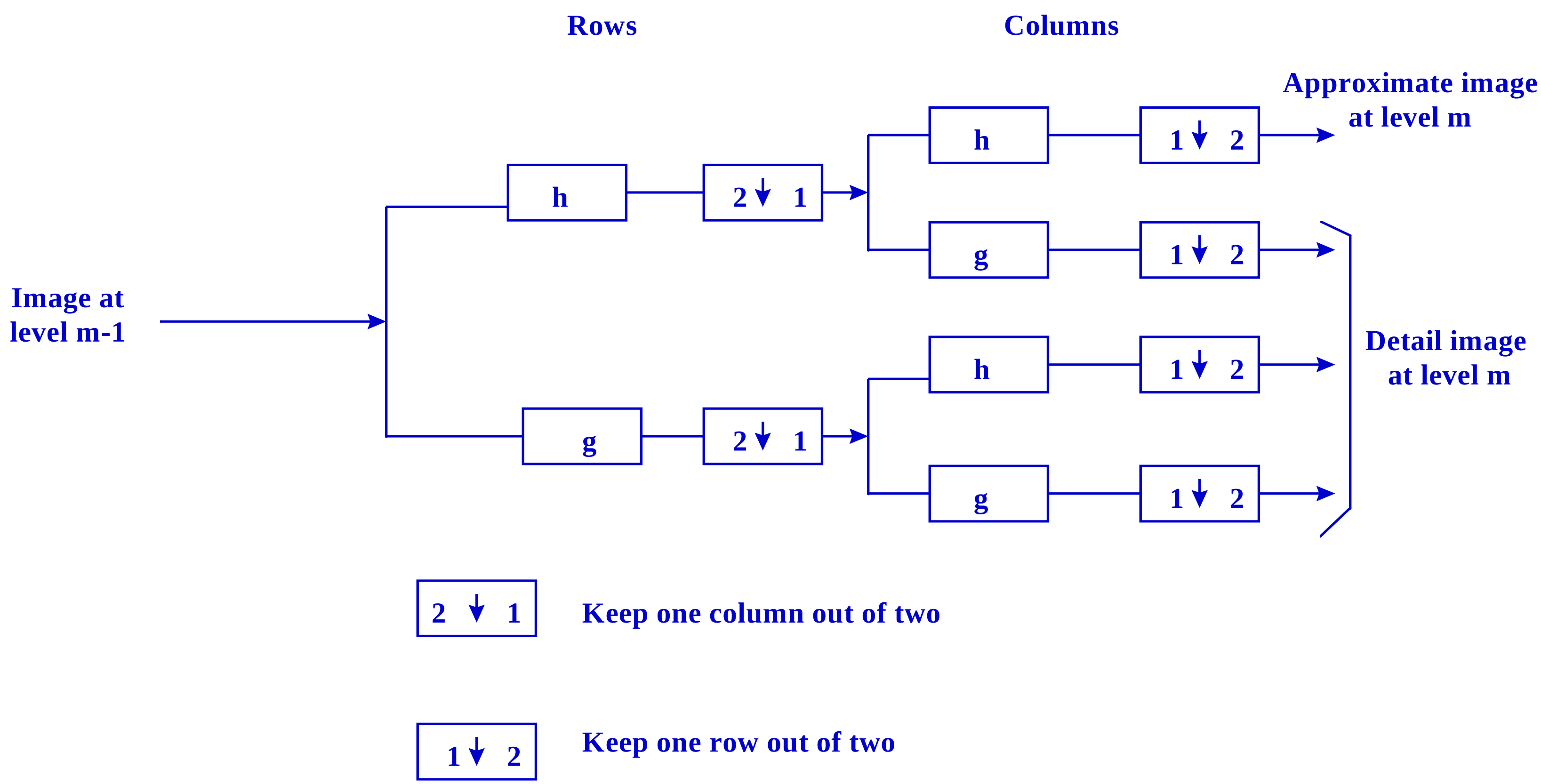
- Scaling function (h)



- Wavelet function (g)



Wavelet decomposition of an image



- Approximate subimages are further decomposed
- Detail images constitute the wavelet coefficients (C_{ij})

Multi-resolution wavelet features

- A 2-level wavelet transform of the image is constructed using the orthonormal Daubechies filter

- Features

1. Wavelet energy $$f_7 = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} | C_{ij} |$$

2. Variance $$f_8 = \left[\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} | C_{ij} - M |^2 \right]^{1/2}$$

3. Residual energy $$f_9 = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \left| C_{ij} - M \right|$$

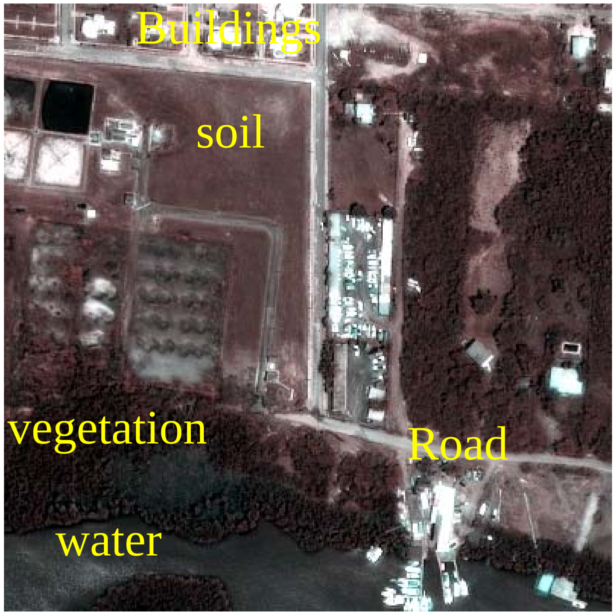
Classifier distance metric equation

$$D_{ij} = \sum_{r=1}^{q} \left(\frac{f_r^{\,i} - f_r^{\,j}}{\alpha(f_r)} \right)^2$$

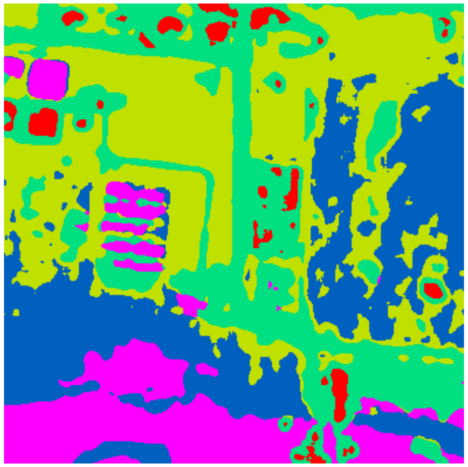
where

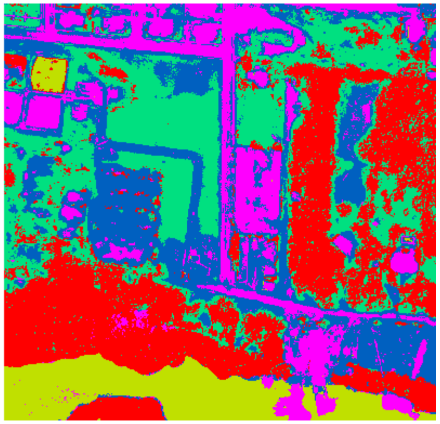
- $\left(f_1^i, f_2^i, \ldots, f_q^i\right)$ is the feature vector for texture class i
- $\left(f_1^j, f_2^j, \ldots, f_q^j\right)$ is the feature vector for texture class j
- q is the number of features
- $\alpha(fr)$ is the std. deviation of the feature fr over the C texture classes.

Classifications of IKONOS image (La Parguera)



4 bands
– R, Near IR, G, B

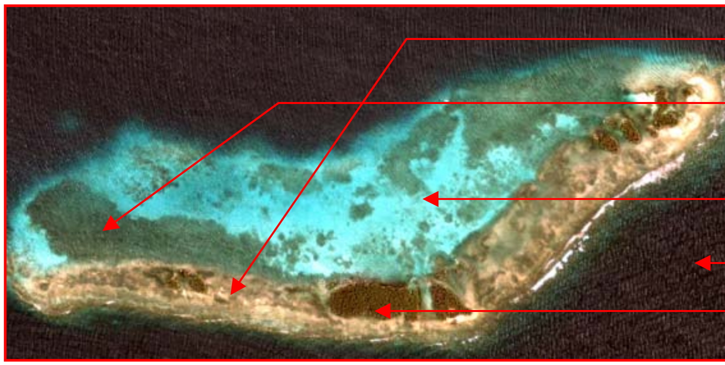




Total features= 4
(Band R spectral value
and 3
Statistical features)
8x8 regions
Accuracy (testing) =99%

Maximum likelihood classification using 4 bands of spectral values

Benthic habitat classification IKONOS image



4 bands – R,NIR, G and B
Resolution -1mt.

Class map – using spectral and gray level statistical features -3x3 regions
Total features = 5(NIR band and 4 statistical features)
Accuracy (testing samples) = 95.0745%