COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Lecture 10 (4.14.07)

Image Representation and Description

Shahram Ebadollahi

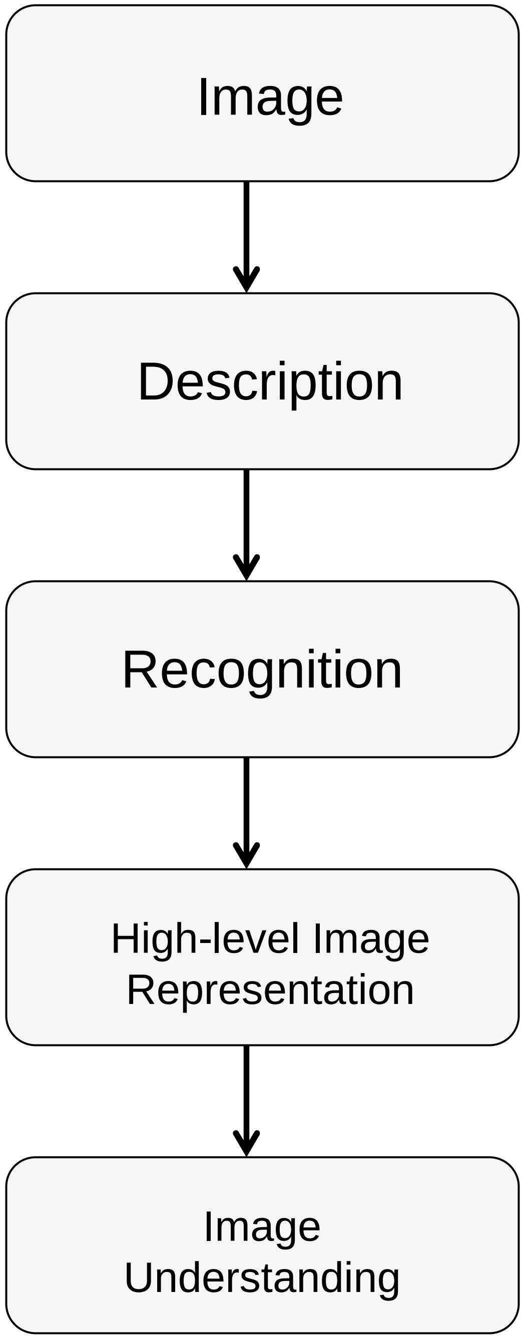


Image
Description
Recognition
High-level Image Representation
Image Understanding

Lecture Outline

- Image Description
  - Shape Descriptors
  - Texture & Texture Descriptors
  - SIFT
  - Motion Descriptors
  - Color Descriptors

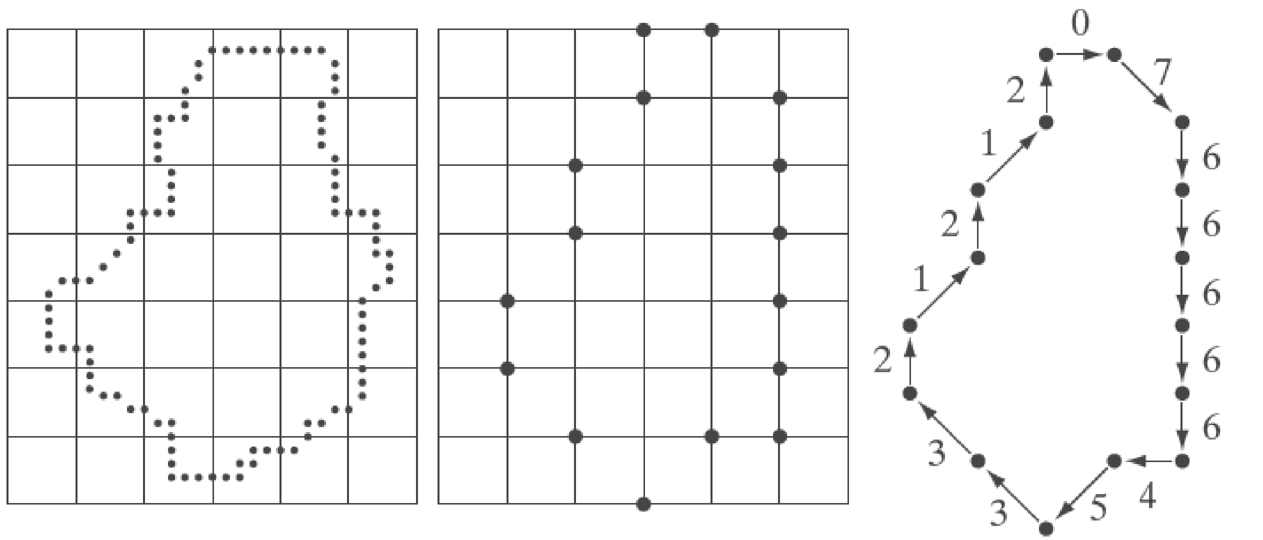
Shape Description

- Shape Represented by its Boundary
  - Shape Numbers,
  - Fourier Descriptors,
  - Statistical Moments

- Shape Represented by its Interior
  - Topological Descriptors
  - Moment Invariants

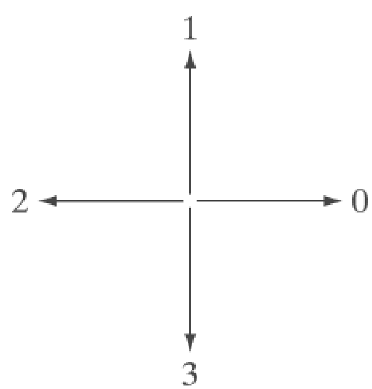
Boundary Representation: (Freeman) Chain Code

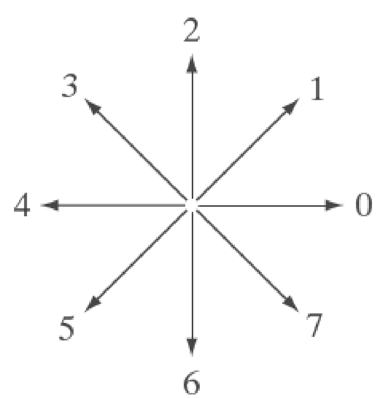
Boundary representation = 0766666453321212



Original boundary   Sub-sampled boundary   Chain code of boundary

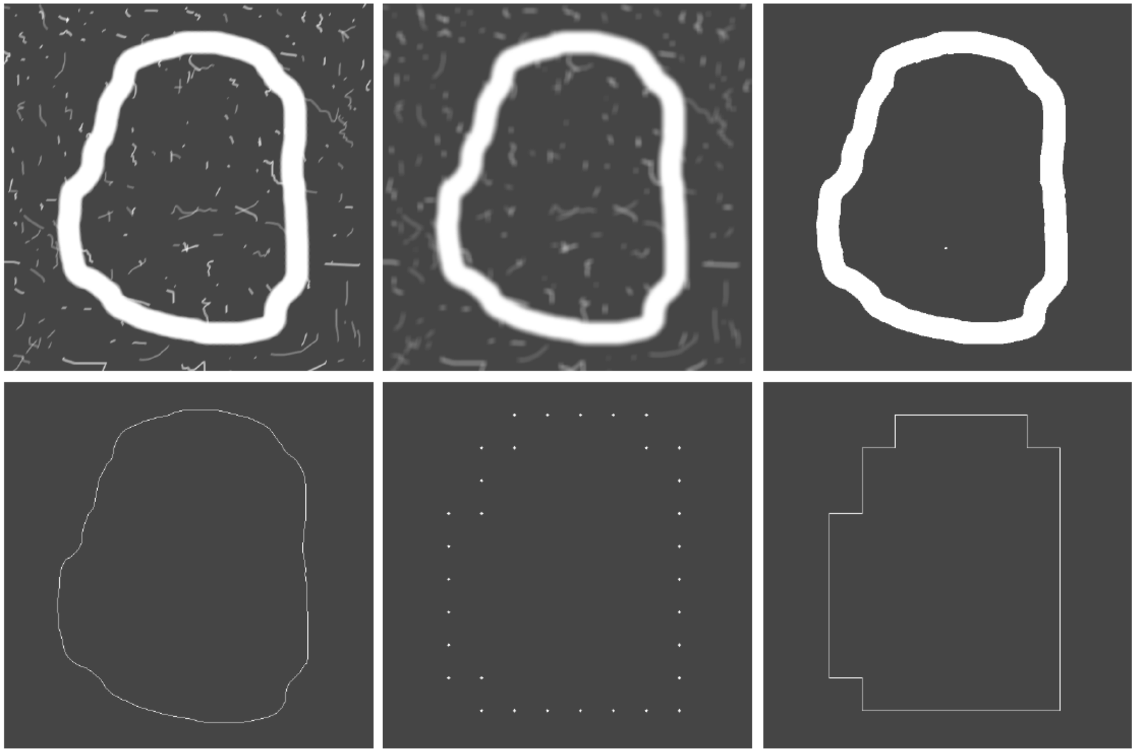
Chain code for 4-neighborhood





Chain code for 8-neighborhood

Chain Code: example



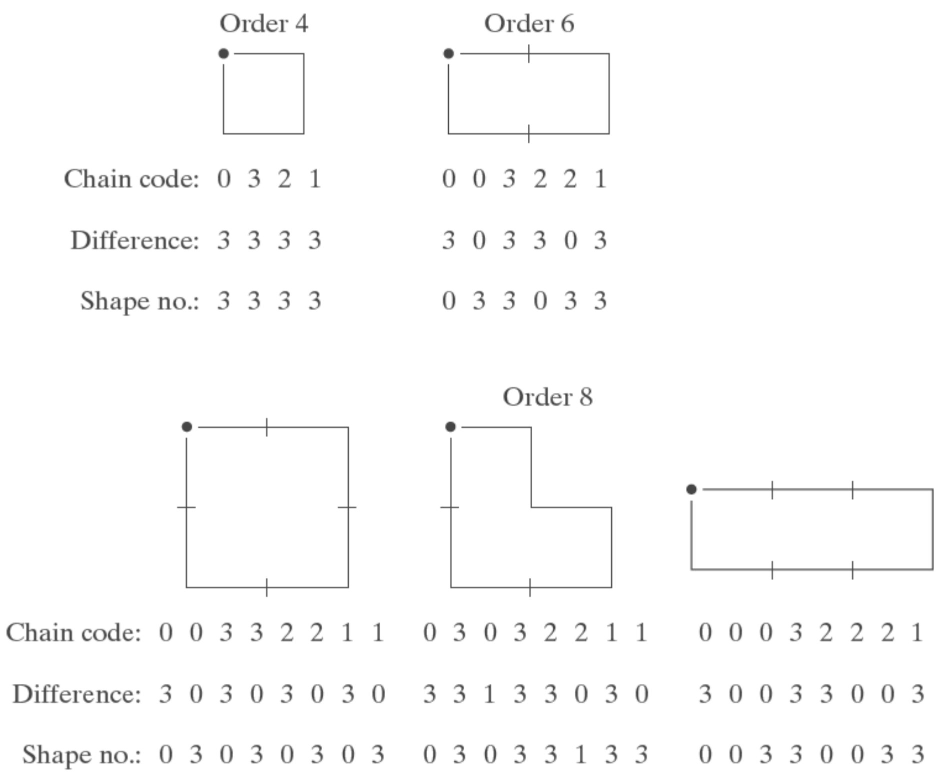
8-directional chain code → 0000606666666644444242222202202

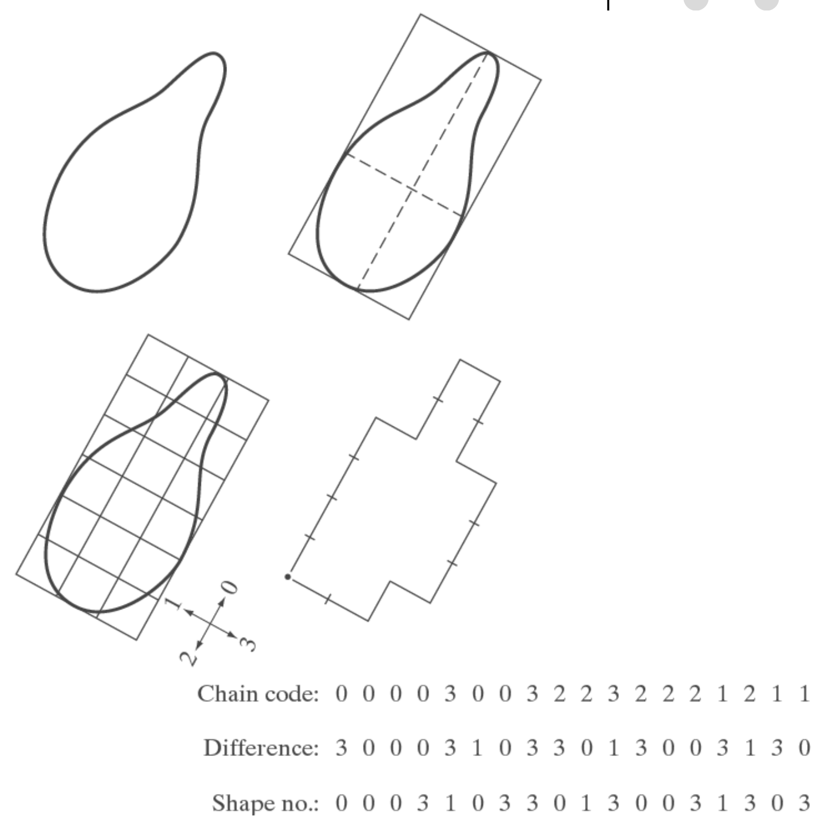
Starting point normalized chain code → 0000606666666644444242222202202

Rotation normalized chain code → 0006200000006000006260000620626

First difference of chain code

Shape Number –
A boundary descriptor



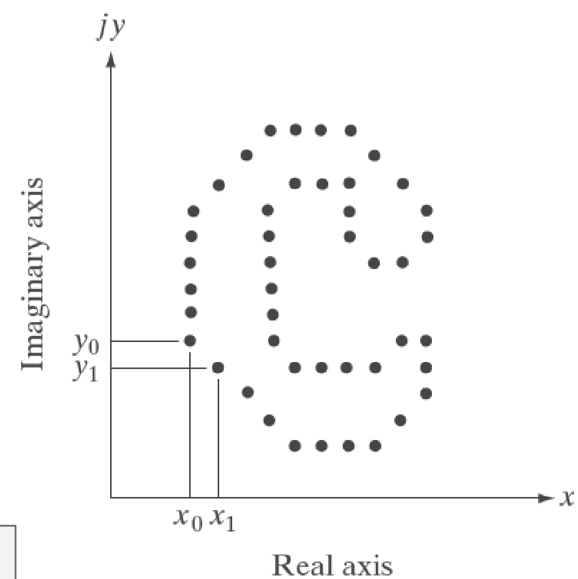


Boundary descriptor – Fourier

$$s(k) = x(k) + jy(k) \qquad k = 0,1,2,\cdots,K-1$$

$$a(u) = \sum_{k=0}^{K-1} s(k) e^{-j2\pi uk/K} \qquad u = 0,1,2,\cdots,K-1$$ ← Fourier Descriptor

$$s(k) = \frac{1}{K}\sum_{u=0}^{K-1} a(u) e^{j2\pi uk/K} \qquad k = 0,1,2,\cdots,K-1$$

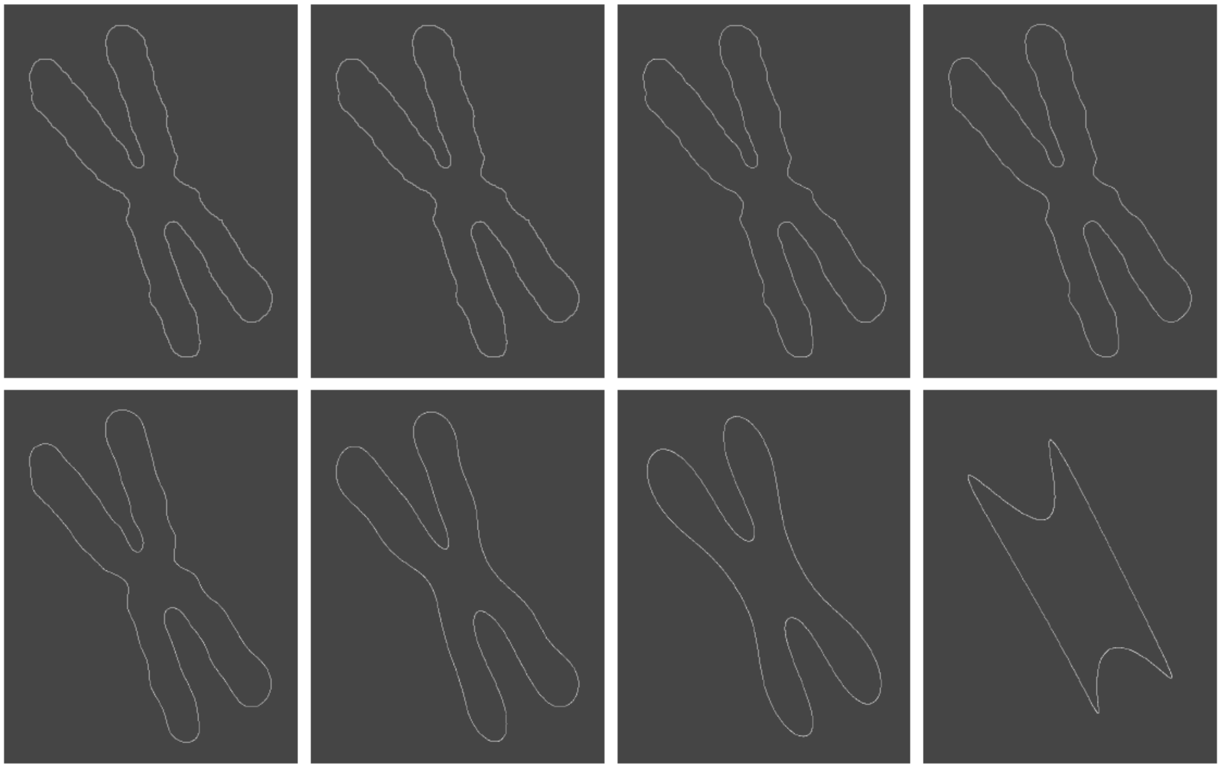


Transformation	Boundary	Fourier Descriptor
Identity	$s(k)$	$a(u)$
Rotation	$s_r(k) = s(k)e^{j\theta}$	$a_r(u) = a(u)e^{j\theta}$
Translation	$s_t(k) = s(k) + \Delta_{xy}$	$a_t(u) = a(u) + \Delta_{xy}\delta(u)$
Scaling	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a(u)$
Starting point	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u)e^{-j2\pi k_0 u/K}$

Boundary Reconstruction using Fourier Descriptors

100% 50% 10% 5%

2868 descriptors

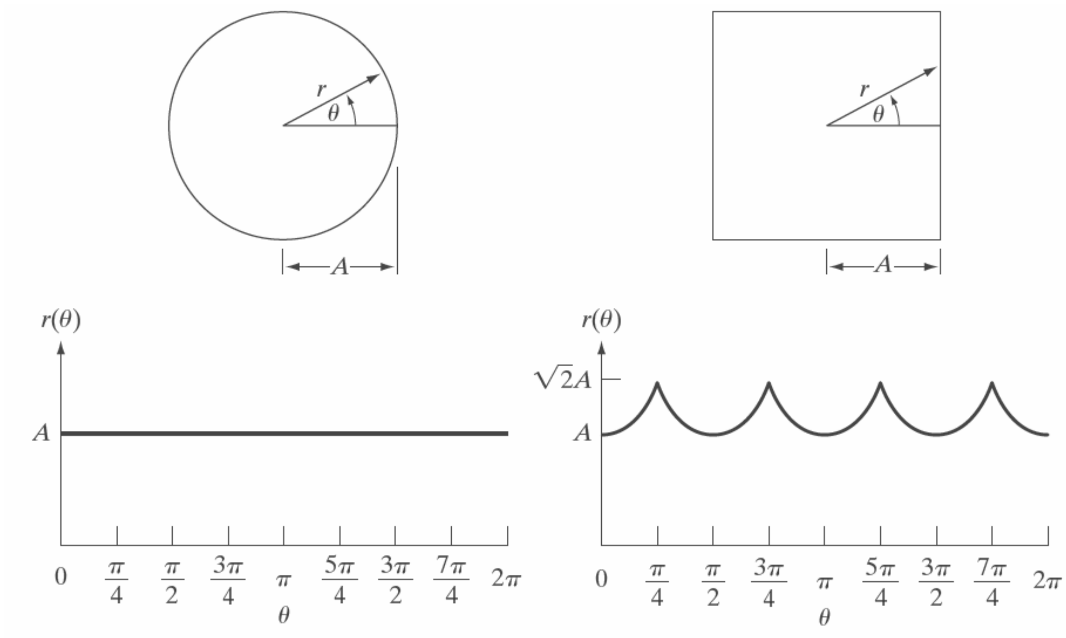


Only 8 descriptors

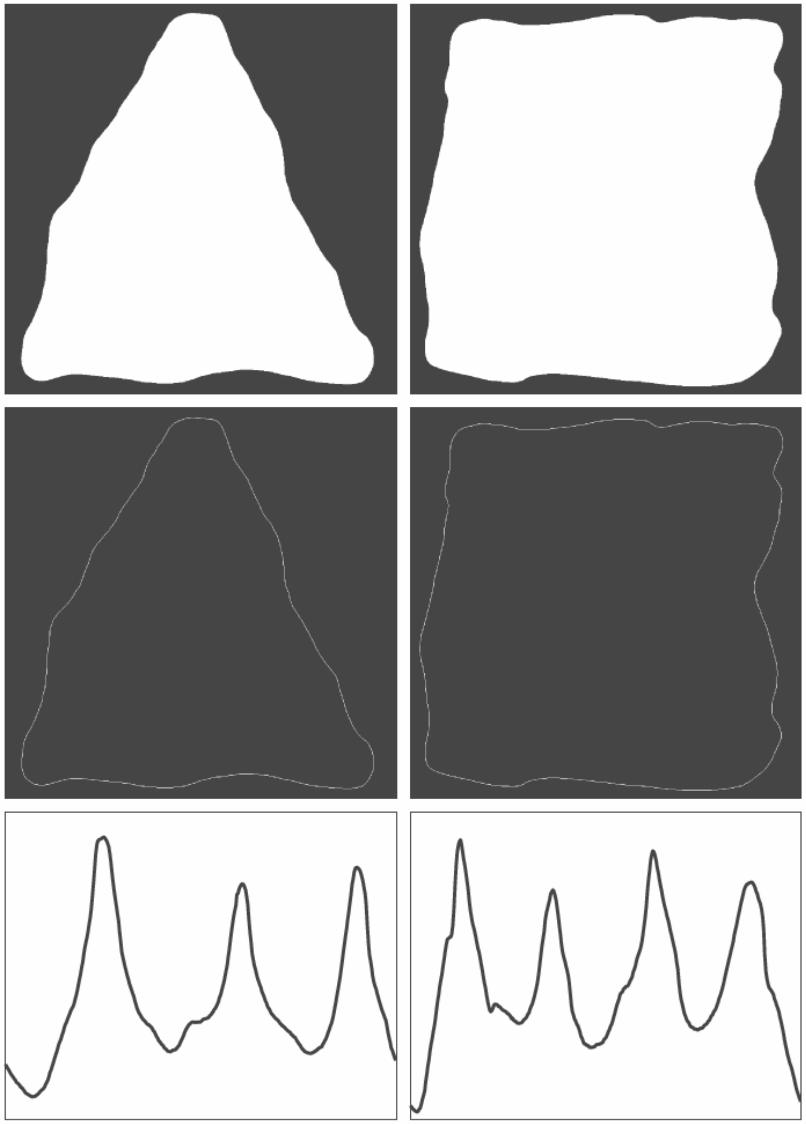
2.5% 1.25% 0.63% 0.28%

Boundary Representation: Signatures

- Represent 2-D boundary shape using 1-D signature signal



Boundary Representation: Signatures

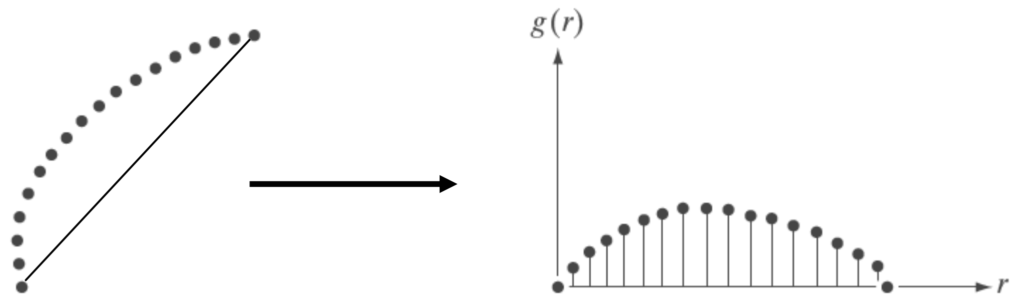


Boundary Description using Statistical Moments

$$\mu_n(v) = \sum_{i=0}^{A-1} (v_i - m)^n p(v_i)$$

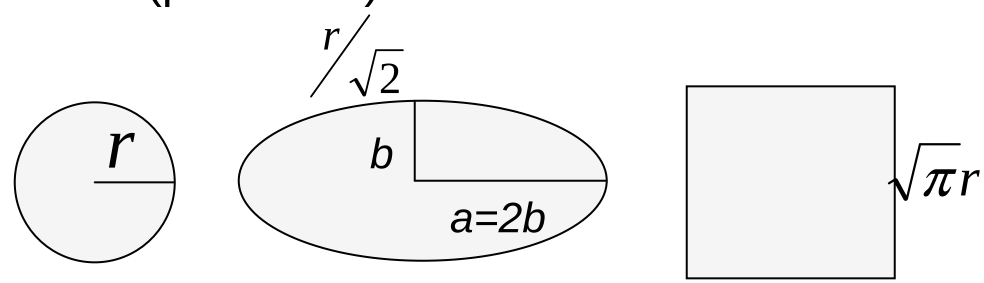
n-th moment of v

$$m = \sum_{i=1}^{A-1} v_i p(v_i)$$



Region Descriptors - Simple

- Area
- Perimeter
- Compactness → (perimeter)^2/Area
- Circularity Ratio
- Mean/Median intensity
- Max/Min intensity
- Normalized area

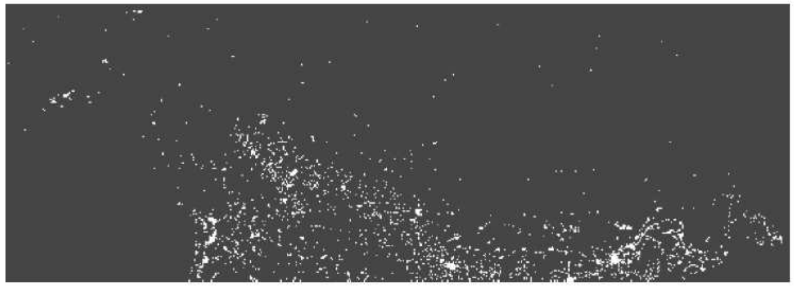


$C: \quad 4\pi \qquad 5\pi \qquad 16$

$R_c: \quad 1 \qquad 4/5 \approx 0.8 \qquad \pi/4 \approx 0.78$

$$R_c = \frac{A}{P^2 / 4\pi}$$

Area of circle with same perimeter as the shape



Region no. (from top)	Ratio of lights per region to total lights
1	0.204
2	0.640
3	0.049
4	0.107

Topological Region Descriptors

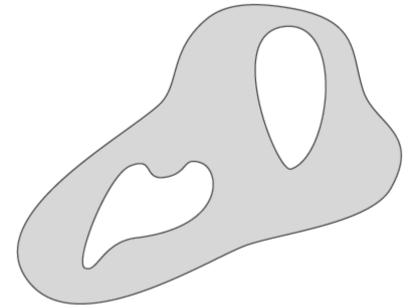
• <u>Topological properties:</u> Properties of image preserved under rubber-sheet distortions

H: # holes in the image

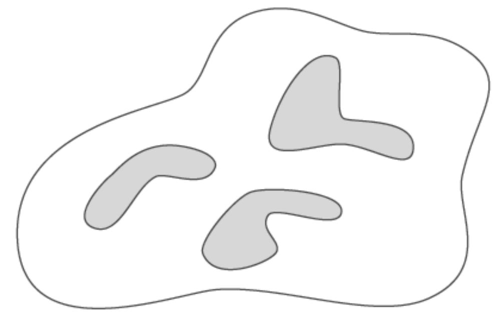
C: # connected components

E = C-H: Euler Number

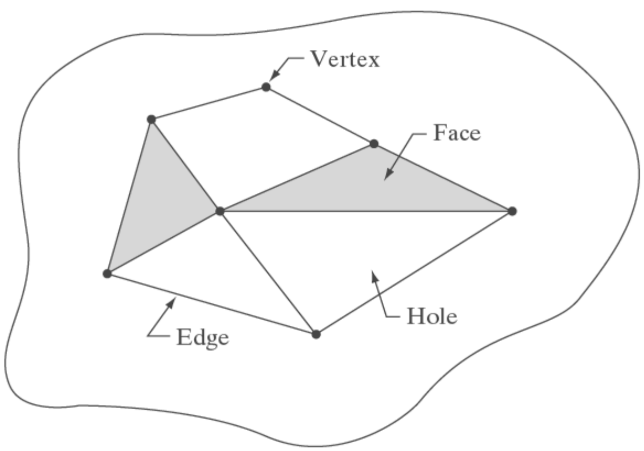
H=2, C=1, E=-1



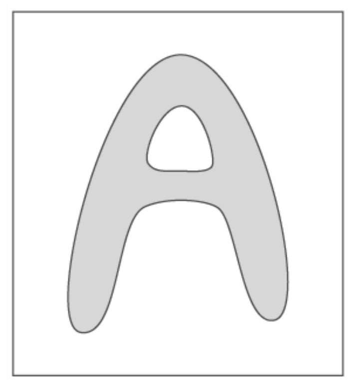
H=0, C=3, E=3



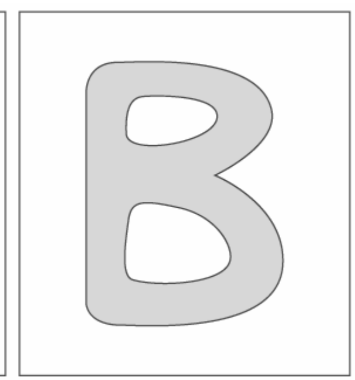
$$V - Q + F = C - H = E$$



H=1, C=1, E=0



H=2, C=1, E=-1



Geometric Moment Invariants

$$m_{pq} = \int\int x^p y^q f(x,y)\,dxdy$$

(p+q)-th 2D geometric moment

$$m_{pq} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} x^p y^q f(x,y)$$

Projection of f(x,y) onto monomial $x^p y^q$

- Why use moments?
  - Geometric moments of different orders represent spatial characteristics of the image intensity distribution

m_{00} Total intensity of image. For binary image → area

$x_0 = m_{10} / m_{00}$ Intensity centroid

$y_0 = m_{01} / m_{00}$ binary image → geometrical center

Central Moments

$$\mu_{pq} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}(x-x_0)^p (y-y_0)^q f(x,y)$$

[Translation invariance]

$\mu_{00} = m_{00}$

$\mu_{10} = \mu_{01} = 0$

μ_{02}, μ_{20} Variance about the centroid

μ_{11} covariance

Scaled Central Moment

$$\lambda_{pq} = \mu'_{pq} / (\mu'_{00})^{(p+q+2)/2}$$

Scale and translation invariant

$$\mu'_{pq} = \frac{\mu_{pq}}{\alpha^{p+q+2}}$$

Normalized Un-Scaled Central Moment

$$\eta_{pq} = \mu_{pq} / (\mu_{00})^{(p+q+2)/2}$$

Moment Invariants (translation, scale, mirroring, rotation)

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

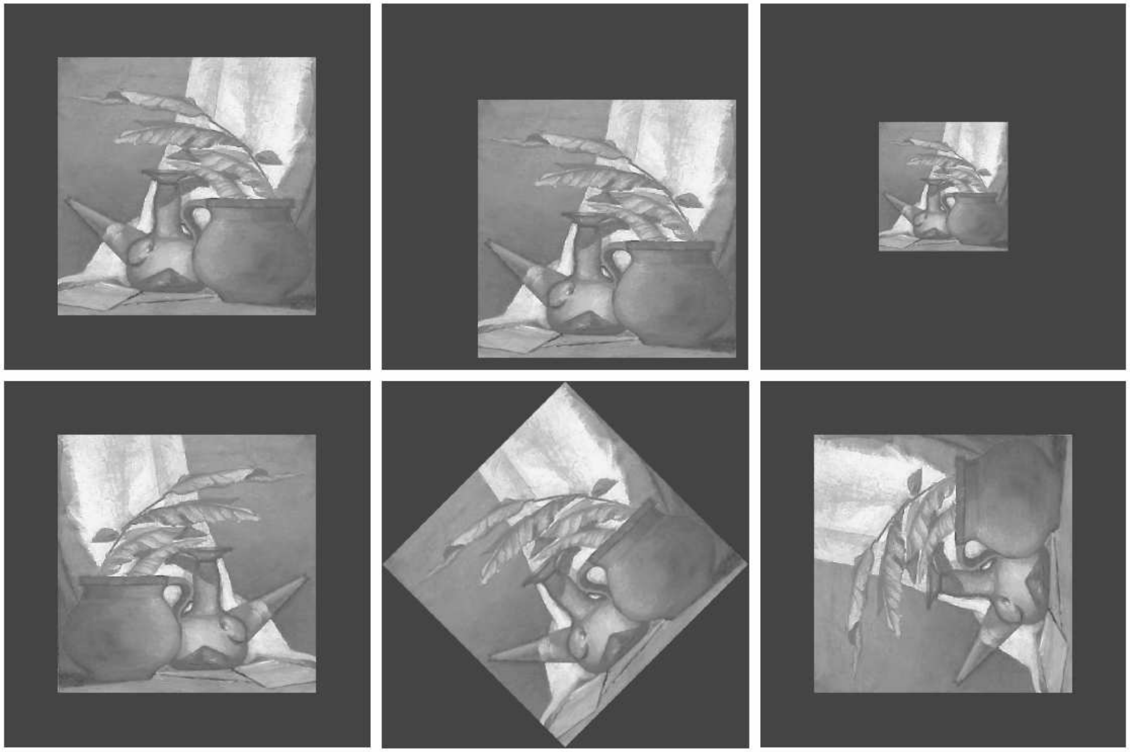
$$\phi_3 = (\eta_{30} - \eta_{12})^2 + (\eta_{21} - \eta_{03})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\phi_5 = \cdots$$

$$\phi_6 = \cdots$$

$$\phi_7 = \cdots$$



Moment Invariant	Original Image	Translated	Half Size	Mirrored	Rotated 45°	Rotated 90°
ϕ_1	2.8662	2.8662	2.8664	2.8662	2.8661	2.8662
ϕ_2	7.1265	7.1265	7.1257	7.1265	7.1266	7.1265
ϕ_3	10.4109	10.4109	10.4047	10.4109	10.4115	10.4109
ϕ_4	10.3742	10.3742	10.3719	10.3742	10.3742	10.3742
ϕ_5	21.3674	21.3674	21.3924	21.3674	21.3663	21.3674
ϕ_6	13.9417	13.9417	13.9383	13.9417	13.9417	13.9417
ϕ_7	−20.7809	−20.7809	−20.7724	20.7809	−20.7813	−20.7809

Affine Transform & Affine Moment Invariants

$$x' = T_x(x, y)$$
$$y' = T_y(x, y)$$

$\longrightarrow$

$$x' = \sum_{r=0}^{m} \sum_{k=0}^{m-r} a_{rk} x^r y^k$$
$$y' = \sum_{r=0}^{m} \sum_{k=0}^{m-r} b_{rk} x^r y^k$$

In practice: bilinear transform

$$x' = a_0 + a_1 x + a_2 y + a_3 xy$$
$$y' = b_0 + b_1 x + b_2 y + b_3 xy$$

4 pairs of corresponding points needed to find coefficients

In practice: affine transform

$$x' = a_0 + a_1 x + a_2 y$$
$$y' = b_0 + b_1 x + b_2 y$$

3 pairs of corresponding points needed to find coefficients

Rotation:

$$x' = x\cos\phi + y\sin\phi$$
$$y' = -x\sin\phi + y\cos\phi$$

Scale:

$$x' = ax$$
$$y' = by$$

Skew:

$$x' = x + y\tan\phi$$
$$y' = y$$

Elliptical Shape Descriptors

Principal moment of inertia:

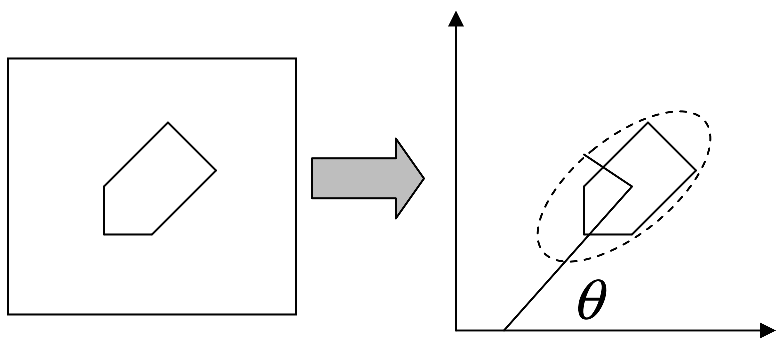
$$I_1 = \frac{(\mu_{20} + \mu_{02}) + [(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2]^{1/2}}{2}$$

$$I_2 = \frac{(\mu_{20} + \mu_{02}) - [(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2]^{1/2}}{2}$$

$(I_1 + I_2) / m_{00}^2$ spreadness

$(I_2 - I_1)/(I_1 + I_2)$ elongation

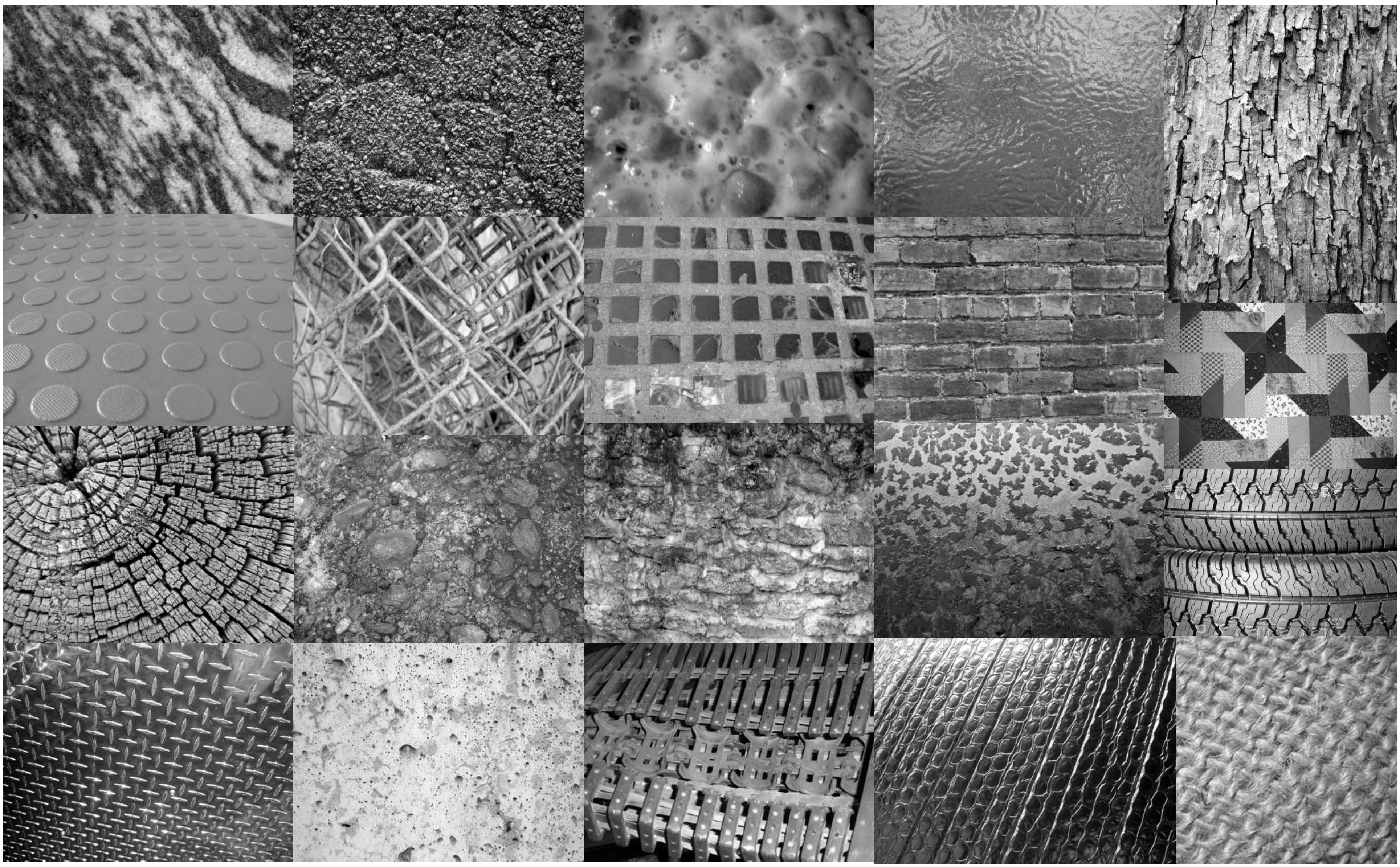
Image ellipse characterizes fundamental shape features and also 2D position and orientation



$$\theta = 0.5 \tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right)$$

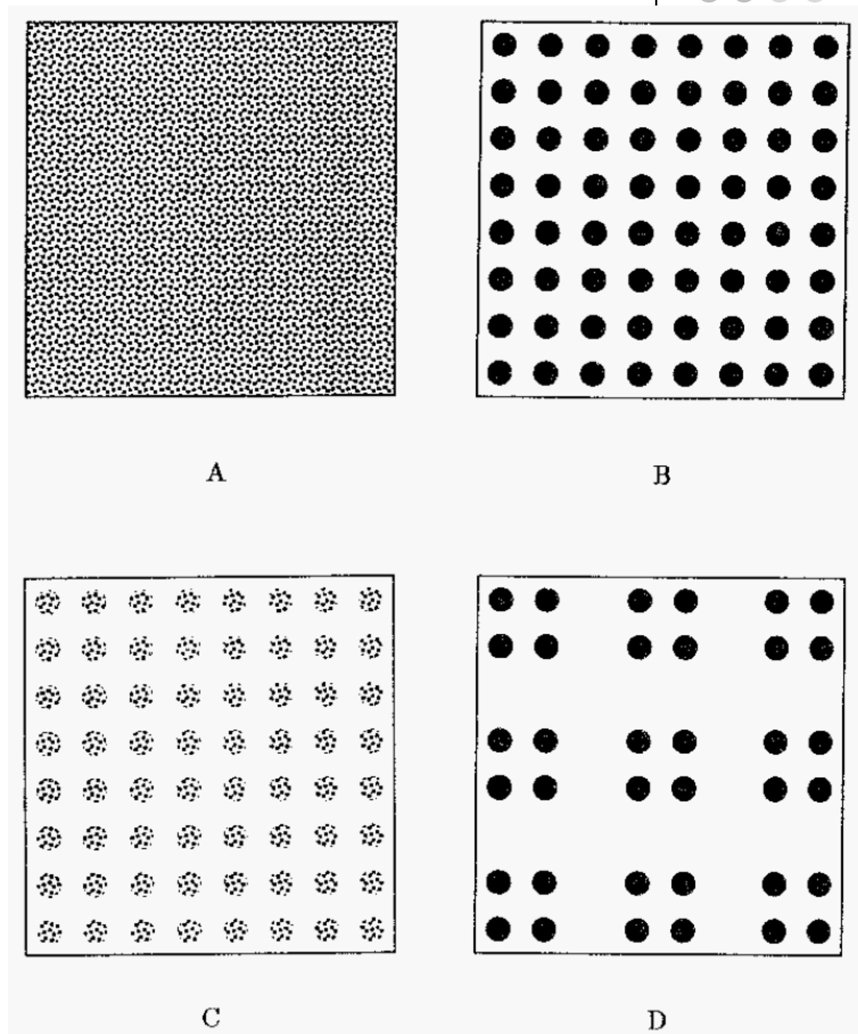
$$a = 2(I_1 / \mu_{00})^{1/2} \qquad b = 2(I_2 / \mu_{00})^{1/2}$$

Texture - Definition



Texture – Quantification Methods

- Statistical: compute local features at each point in image and derive a set of statistics from the distribution of local features
    - 1st, 2nd, and higher-order statistics based on how many points are used to define local features
- Structural: texture is considered to be composed of “texture elements”. Properties of the “texture elements” and their spatial placement rules characterizes the texture
    - Original texture can be reconstructed from its structural description



Statistical Texture Analysis
1st order statistics

image $f \rightarrow h_f$ histogram

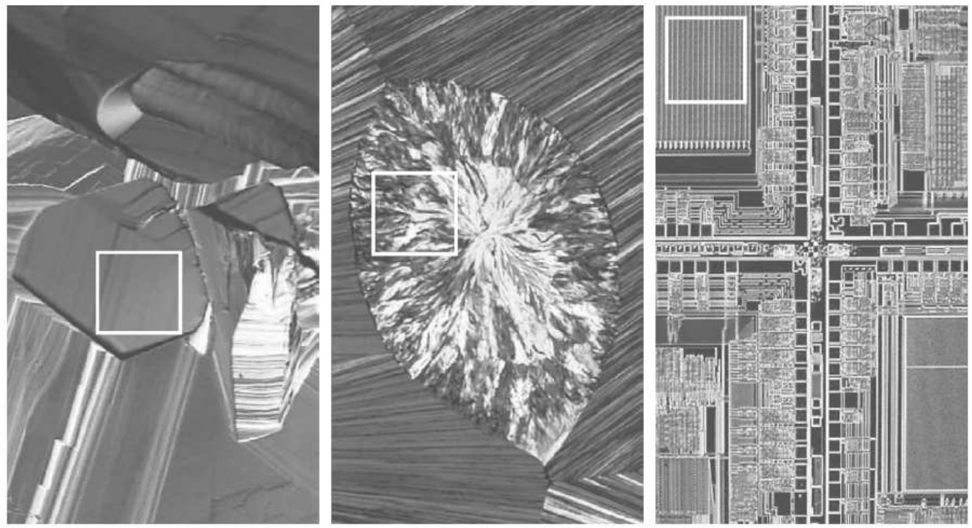
- Obtain statistics of the histogram:

Mean: $\sum_{i=0}^{L-1} ih(i)$: average intensity

Variance: $\sum_{i=0}^{L-1} (i-\mu)^2 h(i)$: measure of intensity contrast

Skewness: $\sum_{i=0}^{L-1} (i-\mu)^3 h(i)$

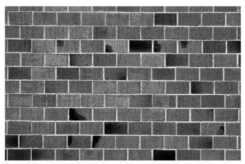
Entropy: $-\sum_{i=0}^{L-1} h(i) \log h(i)$ Measure of variability of intensity

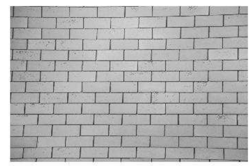


Texture	Mean	Standard deviation	R (normalized)	Third moment	Uniformity	Entropy
Smooth	82.64	11.79	0.002	−0.105	0.026	5.434
Coarse	143.56	74.63	0.079	−0.151	0.005	7.783
Regular	99.72	33.73	0.017	0.750	0.013	6.674

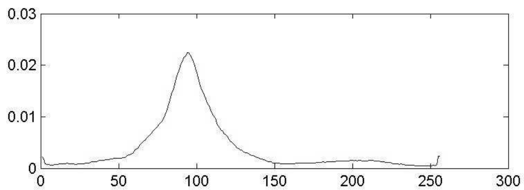
Statistical Texture Analysis
1^{st} order statistics

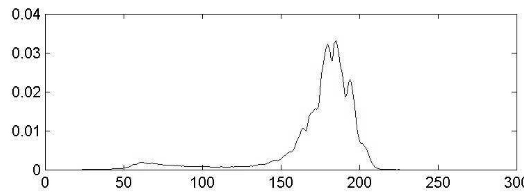
image

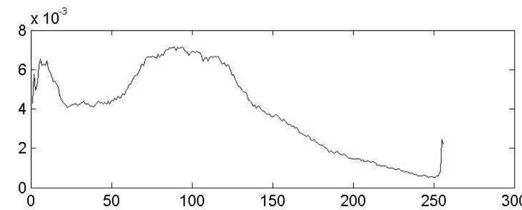




histogram





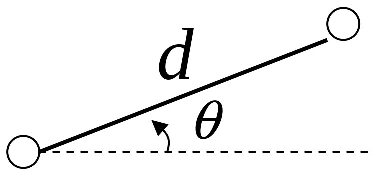


statistics

Skewness = 2.08	Skewness = 2.44	Skewness = -0.092
Entropy = 0.88	Entropy = 0.77	Entropy = 0.97

Statistical Texture Analysis
2^{nd} order statistics: Co-occurrence

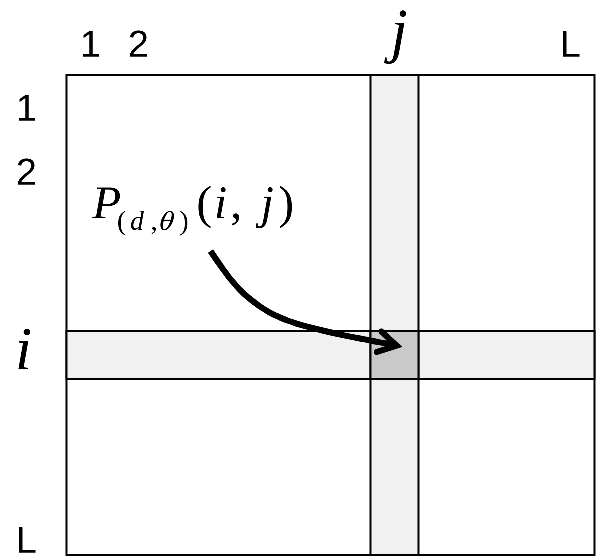
$f(m_2, n_2) = j$



$f(m_1, n_1) = i$

$$P_{(d,\theta)}(i, j) \approx \Pr[\, f(m_1, n_1) = i,\ f(m_2, n_2) = j\,]$$

- Joint gray-level histogram of pairs of pixels
  - 2D histogram



1	1	7	5	3	2
5	1	6	1	2	5
8	8	6	8	1	2
4	3	4	5	5	1
8	7	8	7	6	2
7	8	6	2	6	2

Image f

	1	2	3	4	5	6	7	8
1	1	2	0	0	0	1	1	0
2	0	0	0	0	1	1	0	0
3	0	1	0	1	0	0	0	0
4	0	0	1	0	1	0	0	0
5	2	0	1	0	1	0	0	0
6	1	3	0	0	0	0	0	1
7	0	0	0	0	1	1	0	2
8	1	0	0	0	0	2	2	1

Co-occurrence matrix **G**

Statistical Texture Analysis

2^{nd} order statistics: Co-occurrence

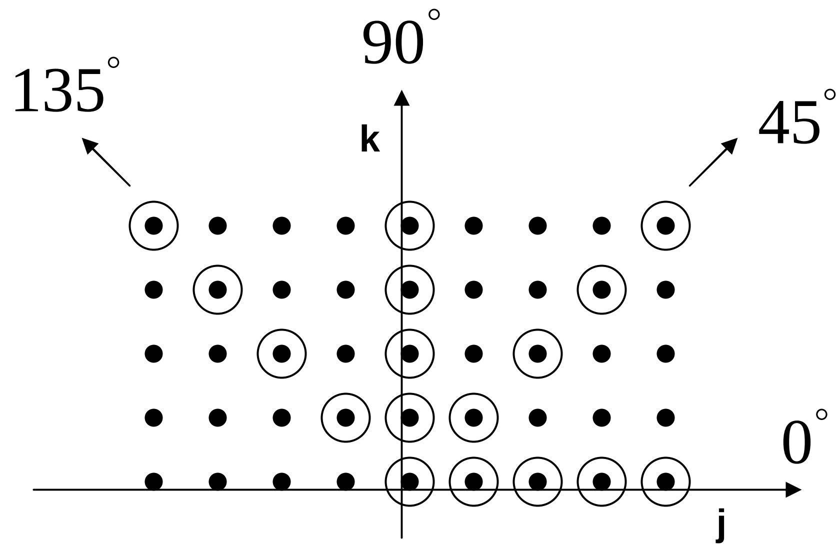
$$P_{(d,\theta=0^\circ)}(i,j) = |\{((k,l),(m,n)) \in (M \times N) \times (M \times N):$$
$$k-m=0, |l-n| = d,\ f(k,l)=i,\ f(m,n)=j\}|$$

$$P_{(d,\theta=45^\circ)}(i,j) = |\{((k,l),(m,n)) \in (M \times N) \times (M \times N):$$
$$(k-m=d, l-n=-d) \vee (k=m=-d, l-n=d),\ f(k,l)=i,\ f(m,n)=j\}|$$

$$P_{(d,\theta=90^\circ)}(i,j)$$

$$P_{(d,\theta=135^\circ)}(i,j)$$

$\left|\{\cdots\}\right|$ is set cardinality



Statistical Texture Analysis

2nd order statistics: Co-occurrence (example)

image

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Co-occurrence matrix

Grey Tone		Grey Tone 0	1	2	3
	0	#(0,0)	#(0,1)	#(0,2)	#(0,3)
	1	#(1,0)	#(1,1)	#(1,2)	#(1,3)
	2	#(2,0)	#(2,1)	#(2,2)	#(2,3)
	3	#(3,0)	#(3,1)	#(3,2)	#(3,3)

0°

$$P_H = \begin{pmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

90°

$$P_V = \begin{pmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

135°

$$P_{LD} = \begin{pmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

45°

$$P_{RD} = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Statistical Texture Analysis

2nd order statistics: Co-occurrence (statistics)

Angular 2nd moment (energy):
(measure of image homogeneity)

$$\sum_{i=1}^{L}\sum_{j=1}^{L} P^2_{(d,\theta)}(i,j)$$

Maximum Probability:

$$\max_{i,j} P_{(d,\theta)}(i,j)$$

Entropy:

$$-\sum_{i=1}^{L}\sum_{j=1}^{L} P_{(d,\theta)}(i,j)\log P_{(d,\theta)}(i,j)$$

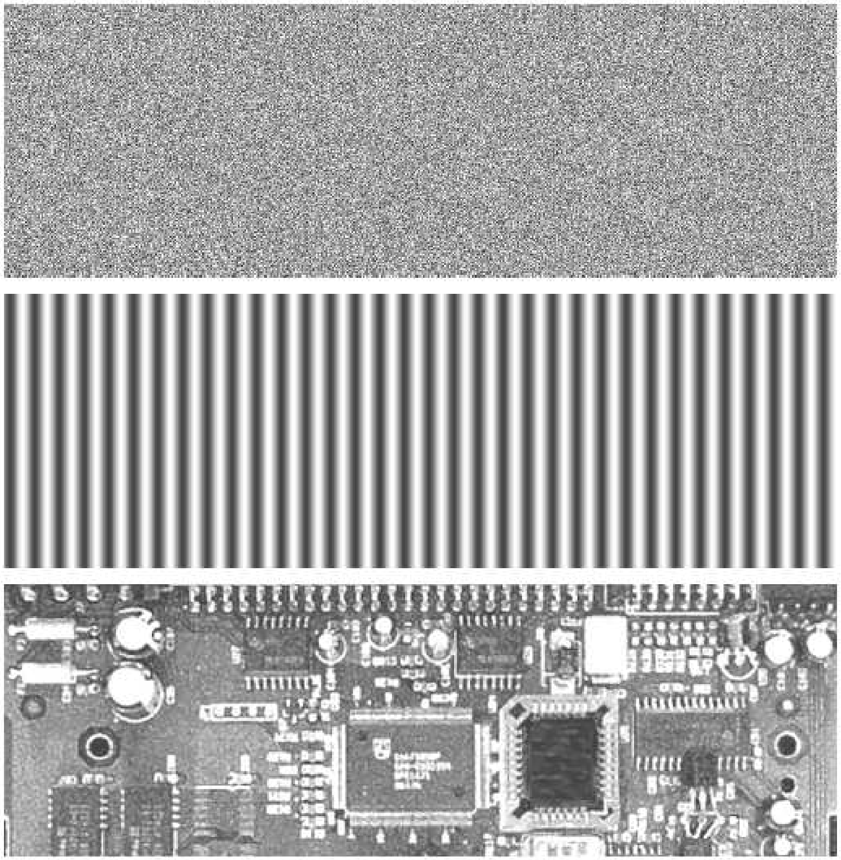
Contrast:
(measure of local variations)

$$\sum_{i=1}^{L}\sum_{j=1}^{L} |i-j|^{\kappa} P^{\lambda}_{(d,\theta)}(i,j)$$

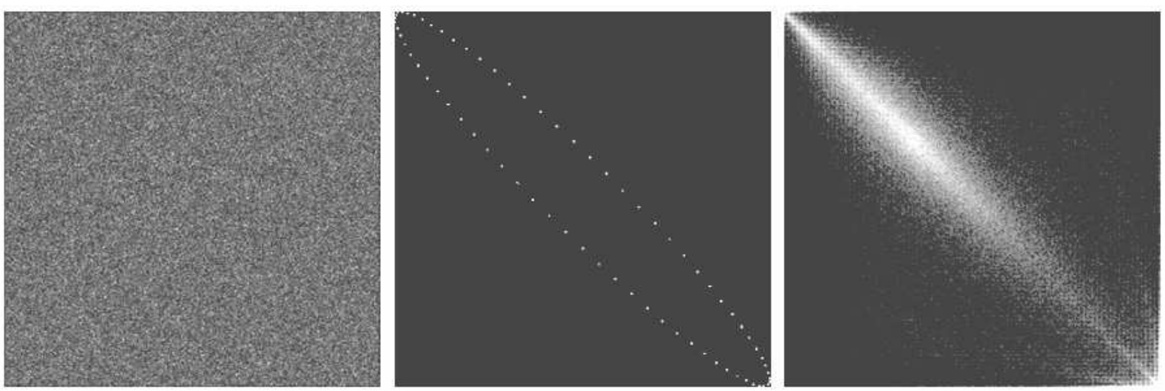
Correlation:
(measure of image linearity)

$$\frac{\sum_{i=1}^{L}\sum_{j=1}^{L}[ijP_{(d,\theta)}(i,j)]-\mu_x\mu_y}{\sigma_x\sigma_y}$$

$$\mu_x = \sum_{i=1}^{L} i \sum_{j=1}^{L} P_{(d,\theta)}(i,j) \qquad \sigma_x = \sum_{i=1}^{L}(i-\mu_x)^2\sum_{j=1}^{L} P_{(d,\theta)}(i,j)$$



Normalized Co-occurrence Matrix	Descriptor					
	Max Probability	Correlation	Contrast	Uniformity	Homogeneity	Entropy
$\mathbf{G}_1/n_1$	0.00006	−0.0005	10838	0.00002	0.0366	15.75
$\mathbf{G}_2/n_2$	0.01500	0.9650	570	0.01230	0.0824	6.43
$\mathbf{G}_3/n_3$	0.06860	0.8798	1356	0.00480	0.2048	13.58



Statistical Texture Analysis

2^{nd} order statistics: Difference Statistics

$$P_{(d,\theta)}(k) = \sum_{\substack{i,j \in \{1,\cdots,L\} \\ |i-j|=k}} P_{(d,\theta)}(i,j)$$ is a subset of co-occurrence matrix

Angular 2^{nd} moment (energy):

$$\sum_{k=0}^{L-1} P_{(d,\theta)}^{2}(k)$$

Mean:

$$\sum_{k=0}^{L-1} k P_{(d,\theta)}(k)$$

Entropy:

$$-\sum_{k=0}^{L-1} P_{(d,\theta)}(k) \log P_{(d,\theta)}(k)$$

Contrast:

$$\sum_{k=0}^{L-1} k^{2} P_{(d,\theta)}(k)$$

Statistical Texture Analysis
2nd order statistics: Autocorrelation

$$C_{ff}(p,q) = \frac{MN}{(M-p)(N-q)} \frac{\sum_{k=1}^{M-p}\sum_{l=1}^{N-q} f(k,l)\, f(k+p,l+q)}{\sum_{k=1}^{M}\sum_{l=1}^{N} f^2(k,l)}$$

Large texture elements ➔ autoccorrelation decreases slowly with increasing distance

Small texture elements ➔ autoccorrelation decreases rapidly with increasing distance

Periodic texture elements ➔ periodic increase & decrease in autocorrelation value

Statistical Texture Analysis

2nd order statistics: Fourier Power Spectrum

$$f(x,y) \leftrightarrow F(u,v)$$

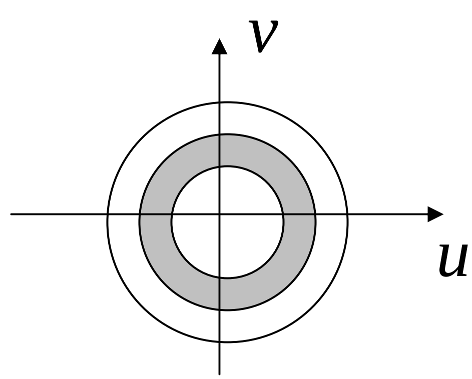
Power Spectrum

$$P(u,v) = |F(u,v)|^2$$

Note:

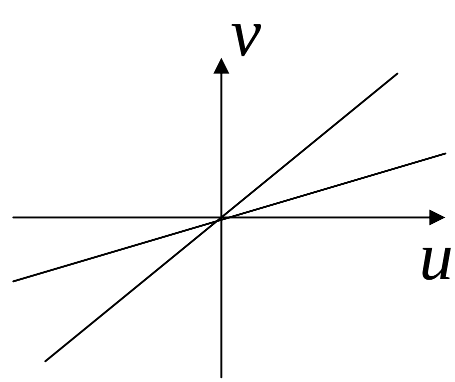
$$C_{ff} = F^{-1}\{|F(u,v)|^2\}$$

$$P(r) = 2\sum_{\theta=0}^{\pi} P(r,\theta)$$



Indicator for size of dominant texture element or texture coarseness

$$P(\theta) = \sum_{r=0}^{L/2} P(r,\theta)$$



Indicator for the directionality of the texture

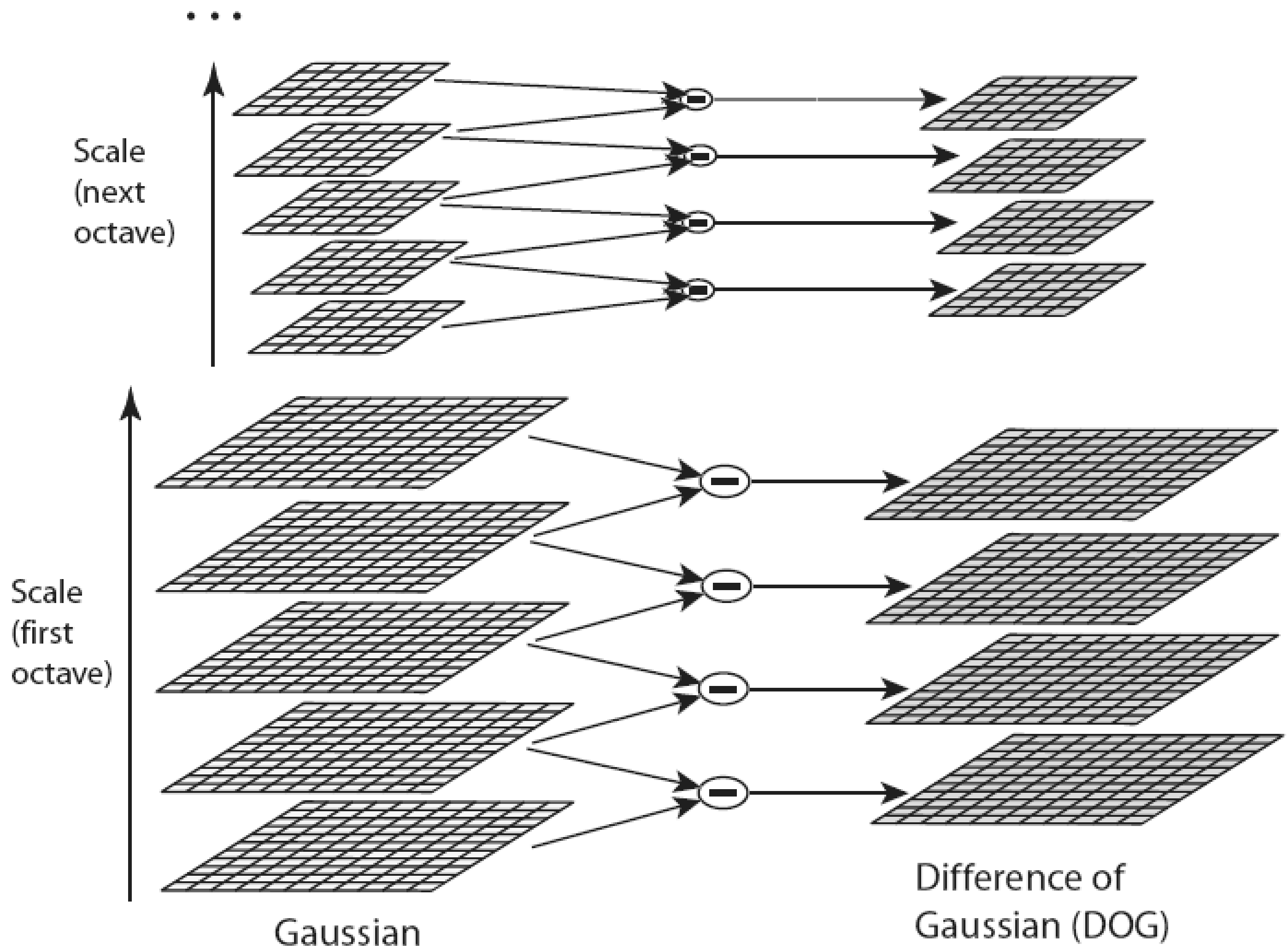
Law's Texture Energy Measures

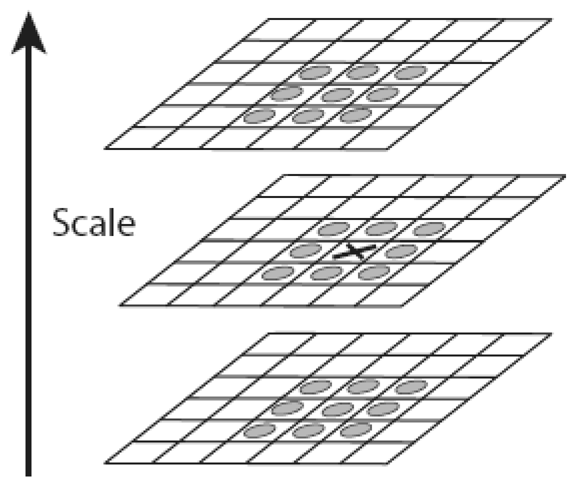
$L_3 = [1,2,1]$ $E_3 = [-1,0,1]$ $S_3 = [-1,2,-1]$

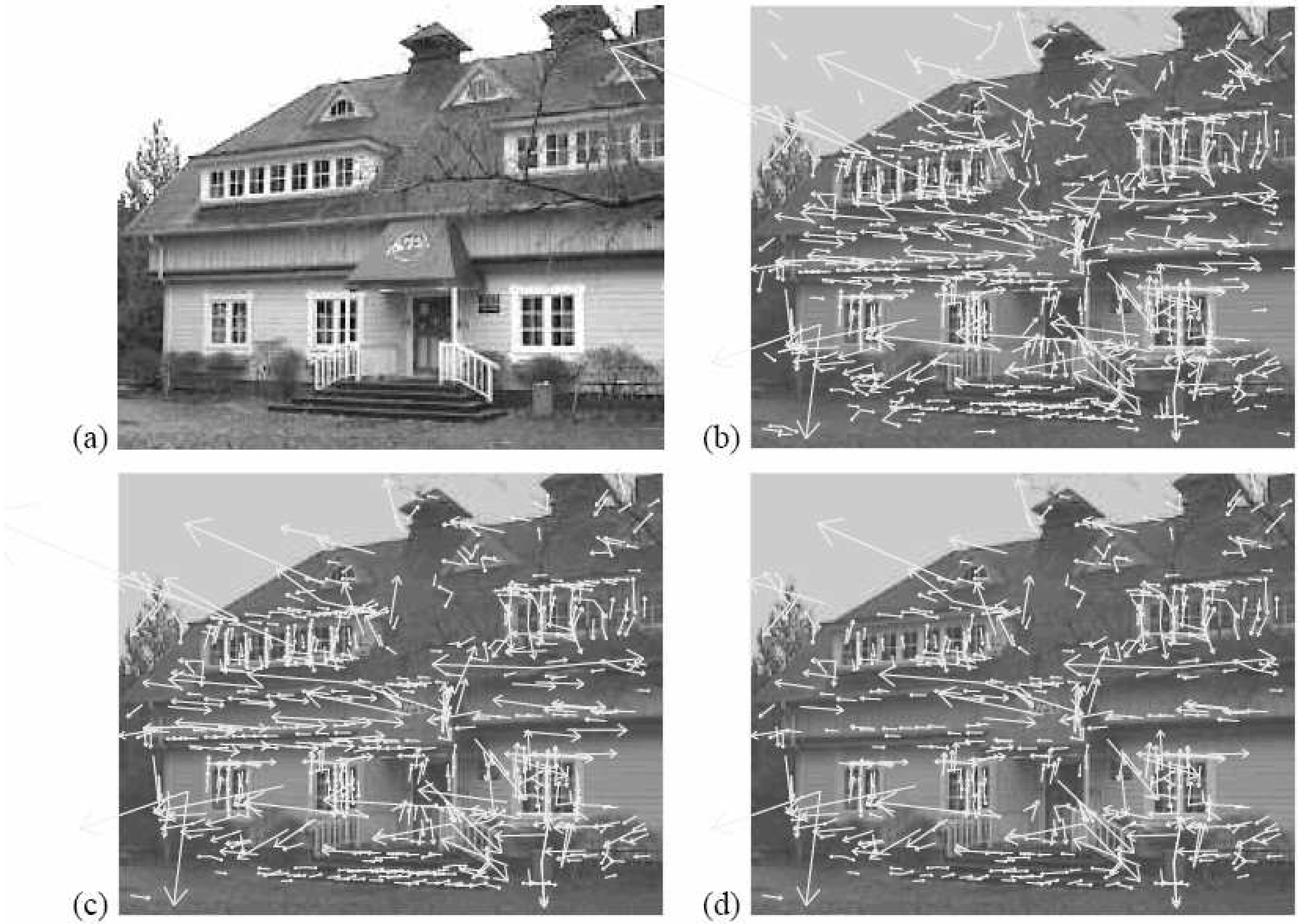
$$L_5 = [1,4,6,4,1]$$
$$E_5 = [-1,-2,0,2,1]$$
$$S_5 = [-1,0,2,0,-1]$$
$$R_5 = [1,-4,6,-4,1]$$
$$W_5 = [-1,2,0,-2,-1]$$

$$L_5^T \times S_5 = \begin{bmatrix} -1 & 0 & 2 & 0 & -1 \\ -4 & 0 & 8 & 0 & -4 \\ -6 & 0 & 12 & 0 & -6 \\ -4 & 0 & 8 & 0 & -4 \\ -1 & 0 & 2 & 0 & -1 \end{bmatrix}$$

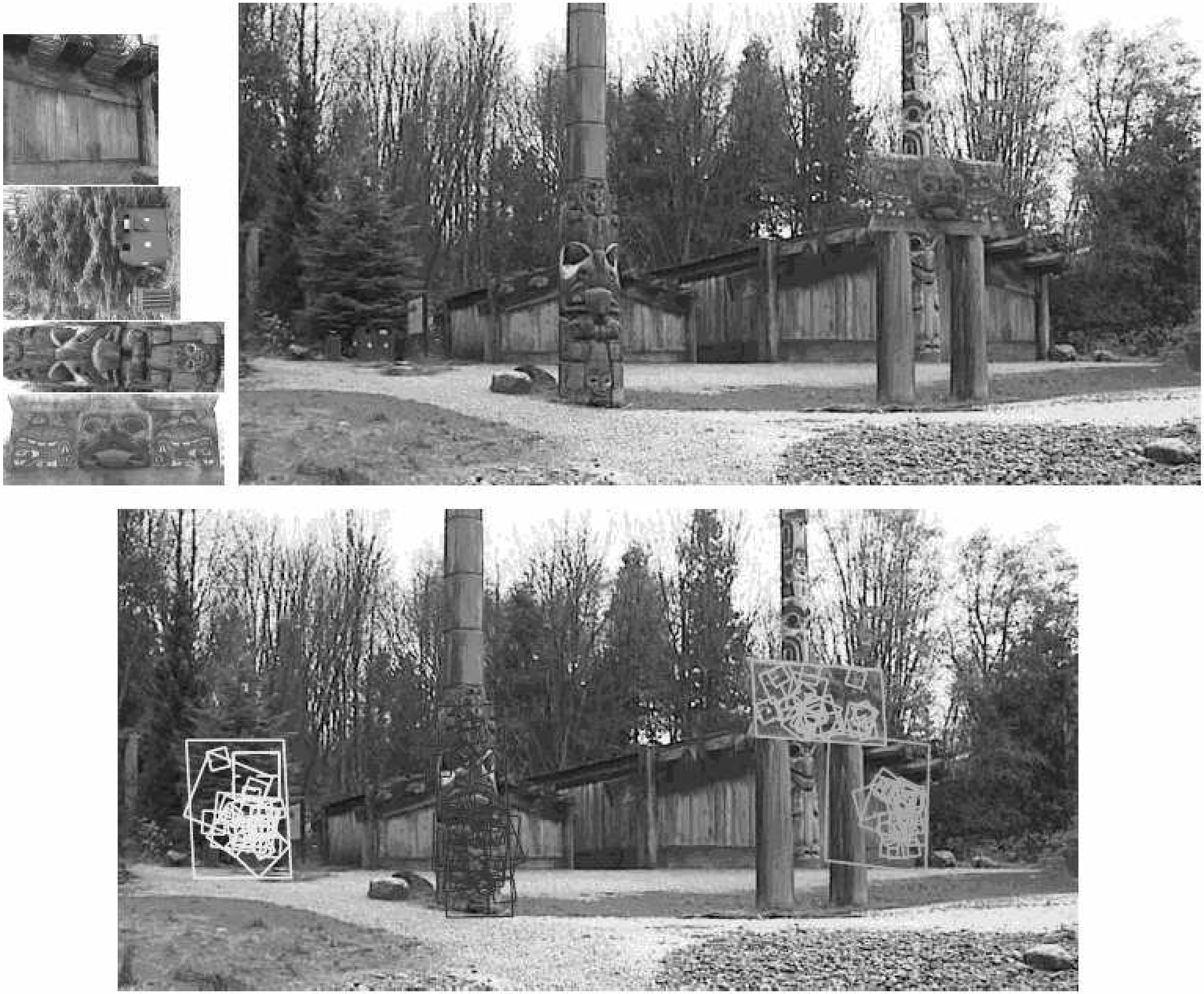
- Convolute different Law's masks with image
- Compute energy statistics


. . .
Scale
(next
octave)
Scale
(first
octave)
Gaussian
Difference of
Gaussian (DOG)


Scale


(a)
(b)
(c)
(d)

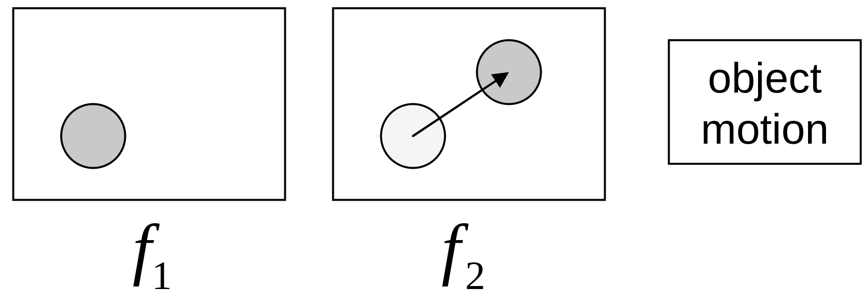

10
PACK
maxell
CD-R



Motion – object

Difference image:

$$d(i,j) = \begin{cases} 0 & \text{if} \quad |f_1(i,j) - f_2(i,j)| \le \varepsilon \\ 1 & \text{otherwise} \end{cases}$$



No motion direction information !

$$d_{cum}(i,j) = \sum_{k=1}^{n} a_k \, | f_1(i,j) - f_k(i,j) |$$

Tells us how often the image gray level was different from gray-level of reference image

Cumulative difference image

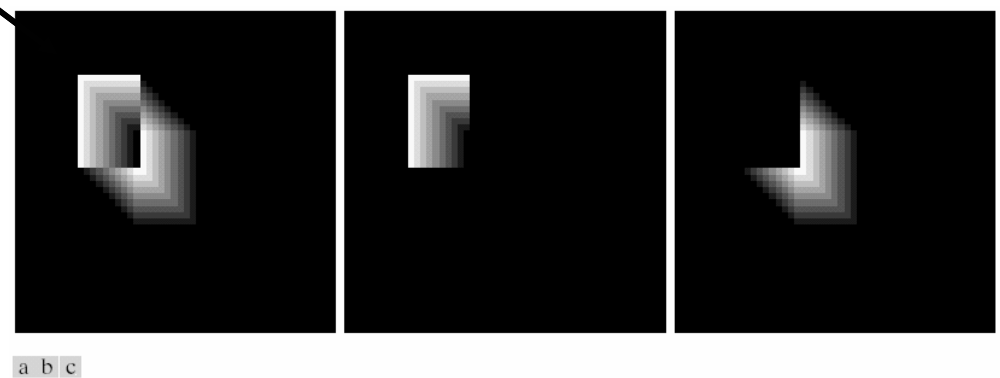
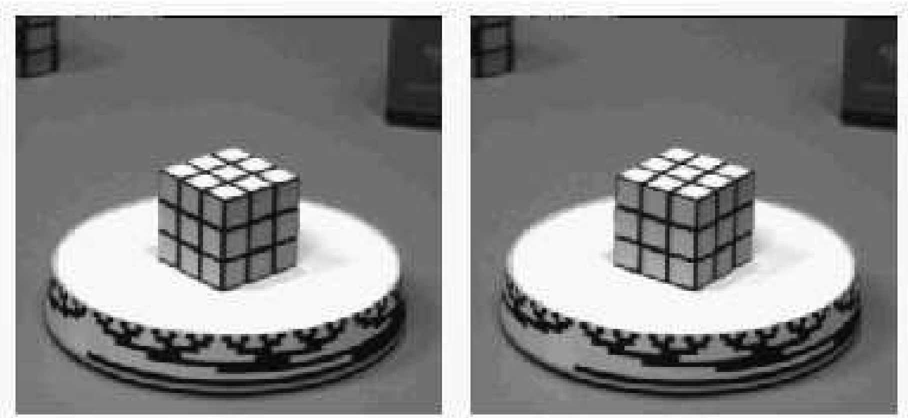
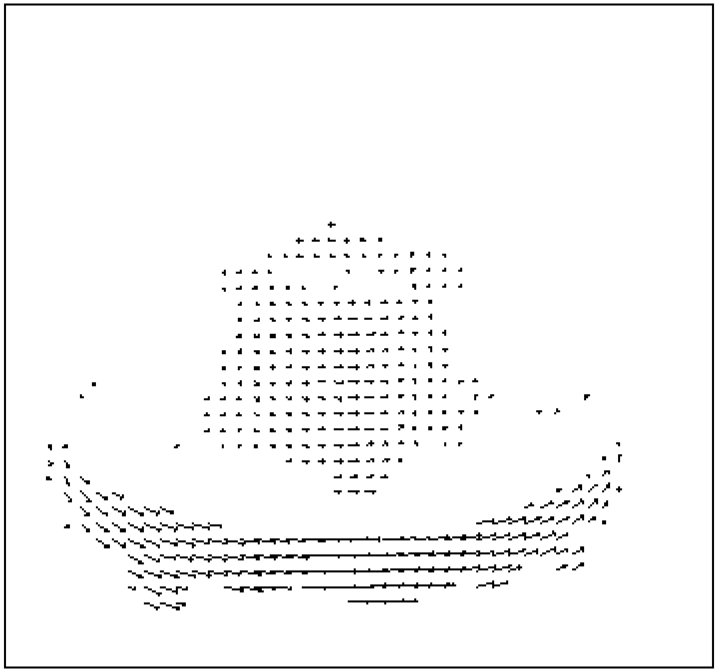


FIGURE 10.49 ADIs of a rectangular object moving in a southeasterly direction. (a) Absolute ADI. (b) Positive ADI. (c) Negative ADI.

Motion Field

- A velocity vector is assigned to each pixel in the image
- Velocities due to relative motion between camera and the 3D scene
- Image change due to motion during a time interval dt
- Velocity field that represents 3-dimensional motion of object points across 2-dimensional image

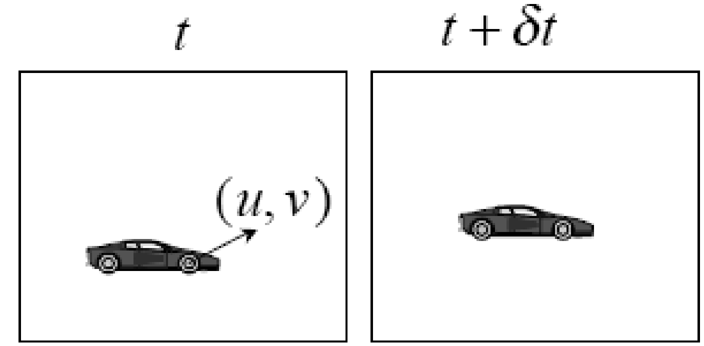




Motion field

Optical flow

- Motion of brightness patterns in image sequence
- Assumptions for computing optical flow:
  - Observed brightness of any object point is constant over time
  - Nearby points in the image plane move in a similar manner



$$f(x+dx, y+dy, t+dt) = f(x,y,t) + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial t}dt + O(\partial^2)$$

$$= f(x,y,t) + f_x dx + f_y dy + f_t dt + O(\partial^2)$$

$$f(x+dx, y+dy, t+dt) = f(x,y,t) \Rightarrow -f_t = f_x \frac{dx}{dt} + f_y \frac{dy}{dt}$$

$$c = (\frac{dx}{dt}, \frac{dy}{dt}) = (u,v)$$

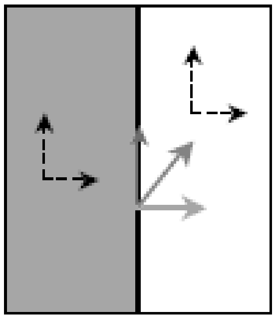
Gray-level difference at same location over time is equivalent to product of spatial gray-level difference and velocity

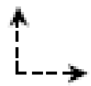
$$-f_t = f_x u + f_y v = \nabla f . c$$

known unknown

Optical Flow Constraints

$$-f_t = f_x u + f_y v$$

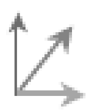




• no spatial change in brightness, induce no temporal change in brightness → no discernible motion

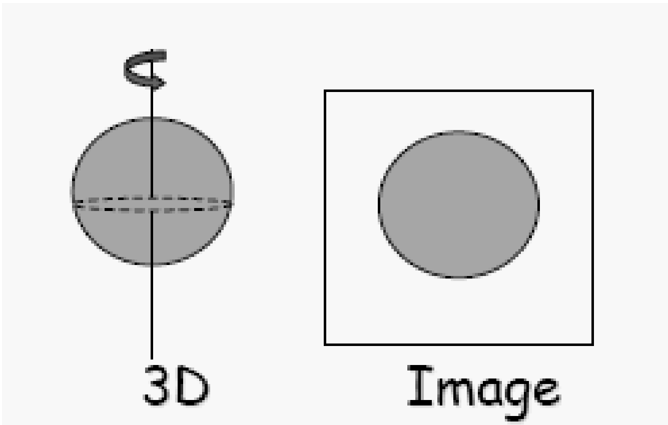
• motion perpendicular to local gradient induce no temporal change in brightness → no discernible motion

• motion in direction of local gradient, induce temporal change in brightness → discernible motion



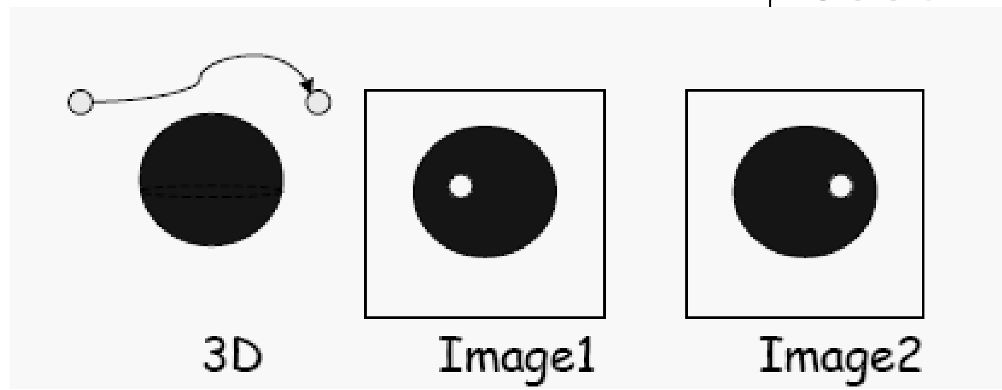
• only motion in direction of local gradient induces temporal change in brightness and discernible motion

Optical flow != Motion Field



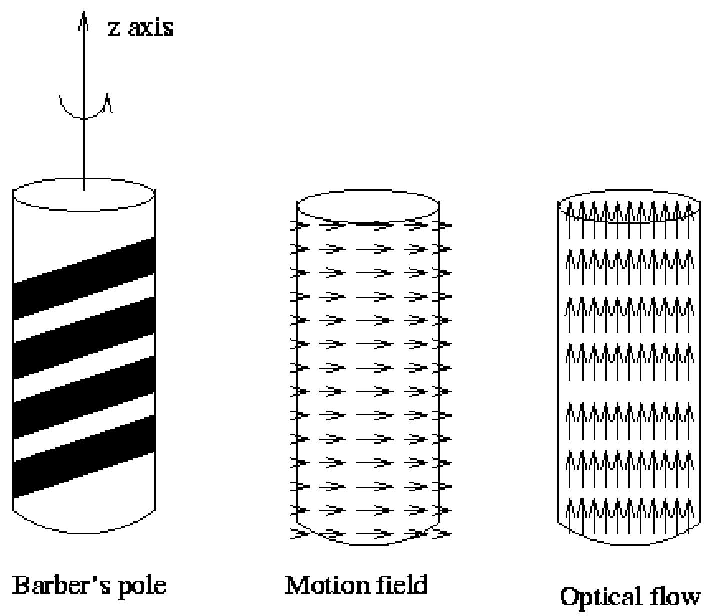
$$MF \neq 0$$

$$OF = 0$$



$$MF = 0$$

$$OF \neq 0$$



Which descriptors? Image Feature Evaluation

1. Prototype Performance
   - Classify (Segment) image using different features
   - Evaluate which feature is optimal (minimum classification error)

2. Figure of Merit
   - Establish functional distance measurements between set of image features (large distance → low classification error)
   - Bhattacharyya distance