

# VC 17/18 – TP15

## Visual Feature Extraction

Mestrado em Ciência de Computadores  
Mestrado Integrado em Engenharia de Redes e  
Sistemas Informáticos

***Miguel Tavares Coimbra***

# Outline

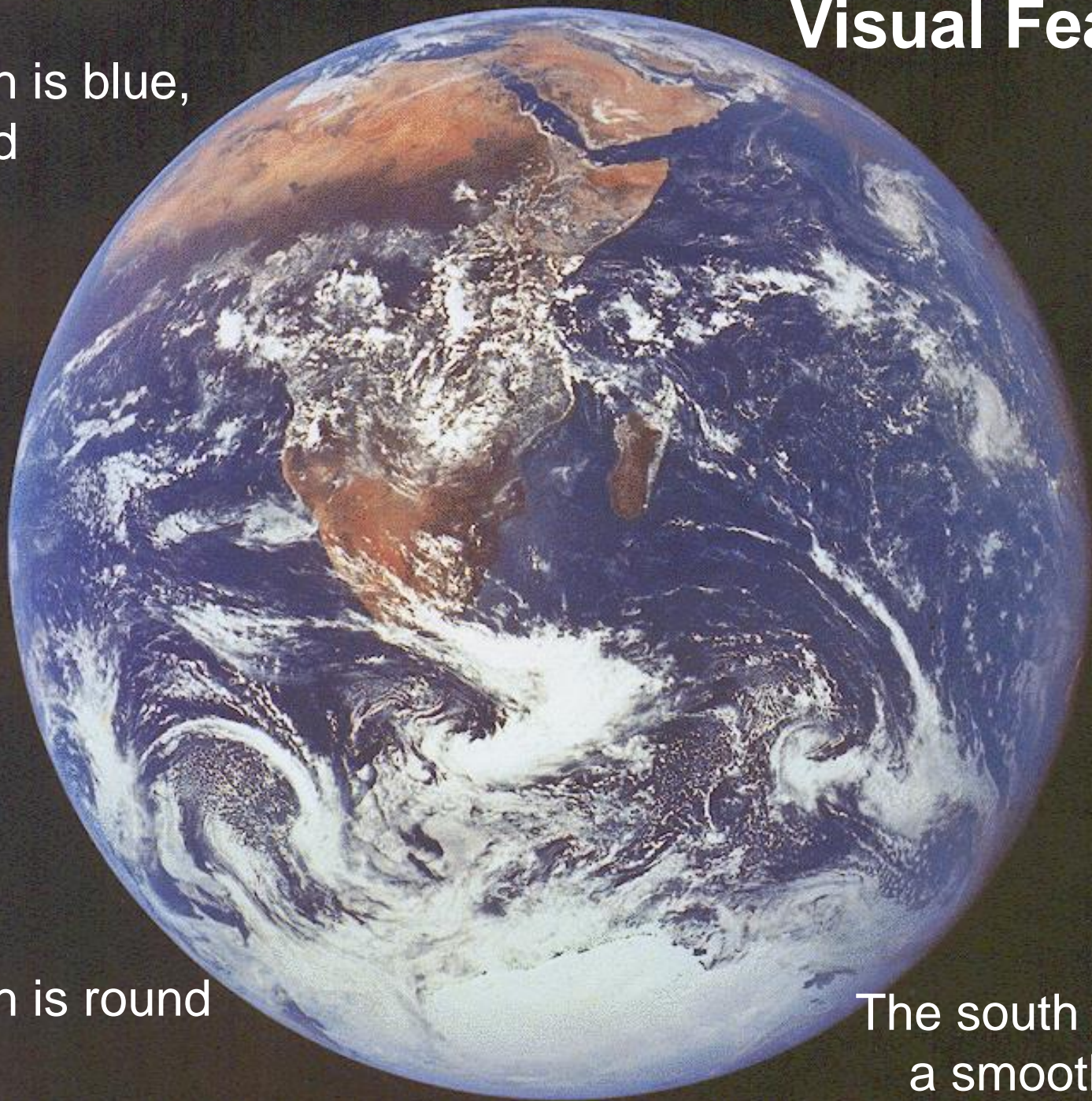
- Feature Vectors
- Colour
- Texture
- Shape

# Topic: Feature Vectors

- Feature Vectors
- Colour
- Texture
- Shape

# Visual Features

The earth is blue,  
white and  
brown



The earth is round

The south pole has  
a smooth texture

# Visual Features

- **Features**
  - Measure specific characteristics.
  - Numerical values.
  - May have multiple values.
- **Visual Features**
  - Quantify visual characteristics of an image.
  - Popular features.
    - Colour, Texture, Shape

# Feature vector

- Feature  $F_i$       $F_i = [f_i]$

- Feature  $F_i$  with  $N$  values.

$$F_i = [f_{i1}, f_{i2}, \dots, f_{iN}]$$

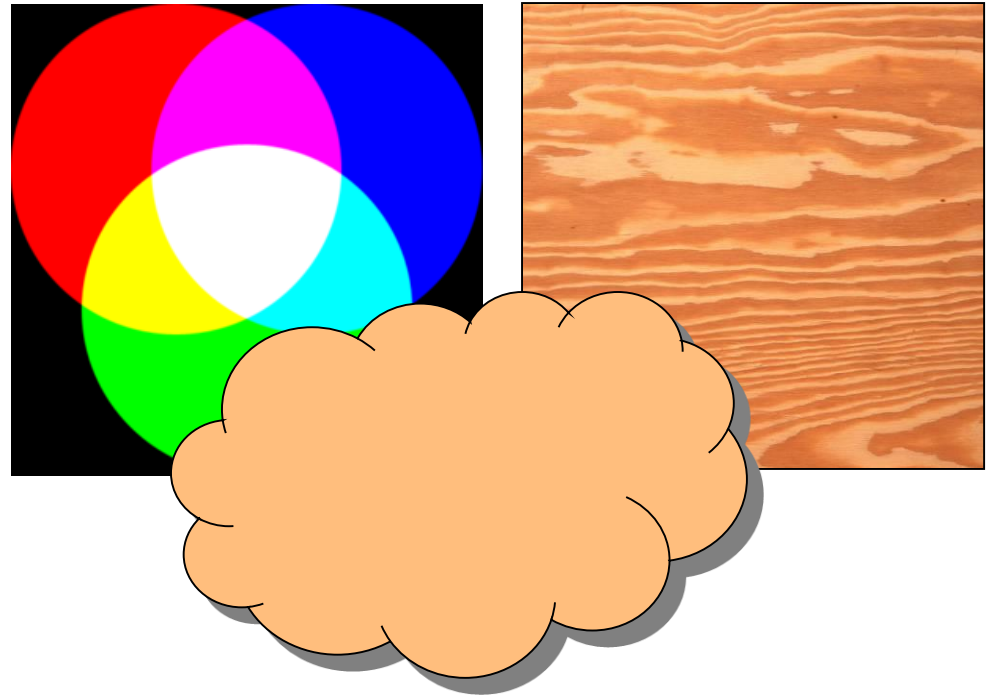
- Feature vector  $F$  with  $M$  features.

$$F = [F_1 | F_2 | \dots | F_M]$$

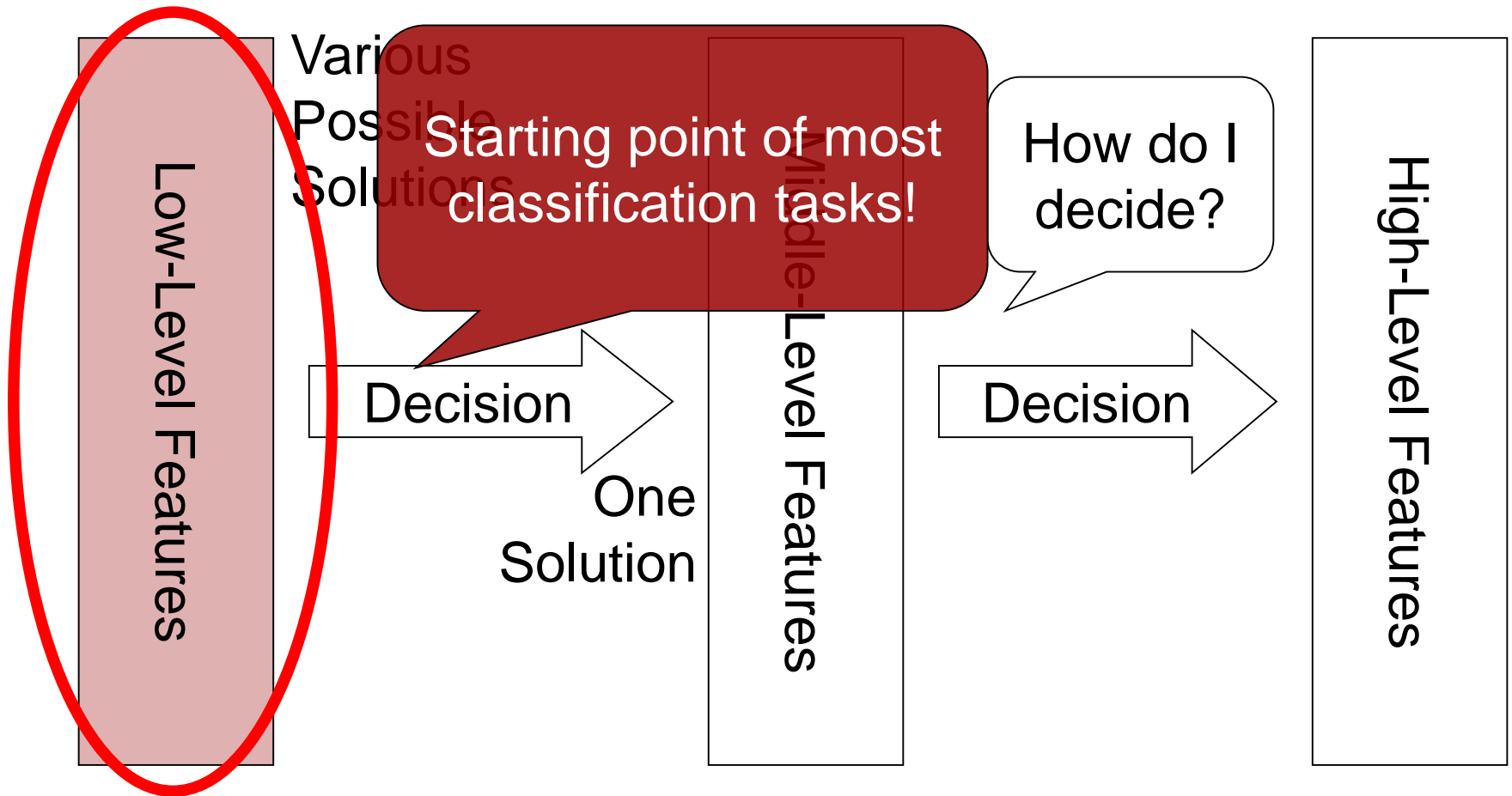
- Naming conventions for this module:
  - Elements of a **feature vector** are called **coefficients**.
  - **Features** may have one or more **coefficients**.
  - **Feature vectors** may have one or more **features**.

# Low-level visual features

- Objective
- Directly reflect specific visual features.
  - Colour
  - Texture
  - Shape
  - Etc.



# Features & Decisions





# How to quantify visual features?

- Many possibilities!
- We need a standard.
- **MPEG-7 Standard**
  - Developed by the Moving Pictures Expert Group.
  - “is a standard for describing the multimedia content data that supports some degree of interpretation of the information meaning, which can be passed onto, or accessed by, a device or a computer code”

*[MPEG-7 Overview (version 10),  
ISO/IEC JTC1/SC29/WG11N6828]*

# The MPEG-7 standard

- Provides a rich set of standardized tools to describe multimedia content.
  - Computer annotation.
  - Human annotation.
- **Audiovisual Description Tools**
  - Descriptors
  - Descriptor Schemes
- **Target functionality:**
  - Efficient search, filtering and browsing of multimedia content.



Feature  
Vectors

# MPEG-7 Links

- MPEG website  
<http://www.chiariglione.org/mpeg>
- MPEG-7 Industry Forum website  
<http://www.mpegif.com>
- MPEG-7 Consortium website  
<http://mpeg7.nist.gov>
- MPEG-7 Overview (version 10)  
<http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm#E9E3>

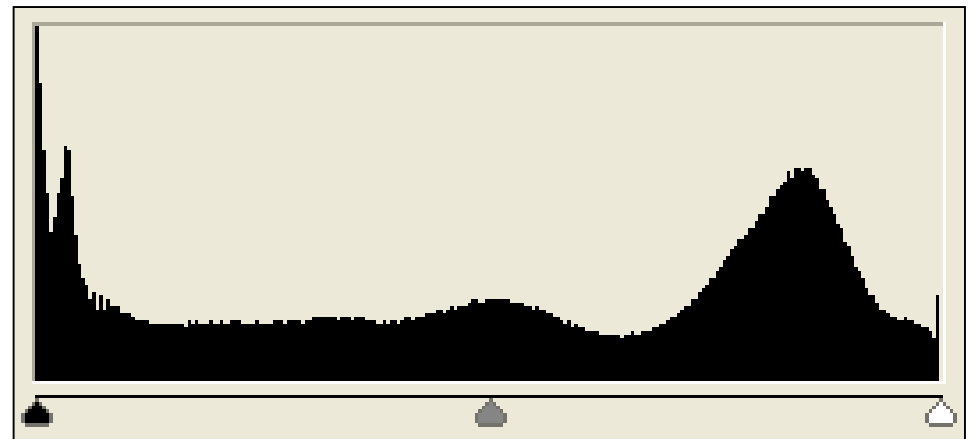
# Topic: Colour

- Feature Vectors
- **Colour**
- Texture
- Shape

# Gray-Level Histogram

- Intensity distribution (HSI).
- We can define the number of histogram bins.
- Histogram bins = Feature coefficients.

$$F = [f_0, \dots, f_{255}]$$



# Colour Histogram

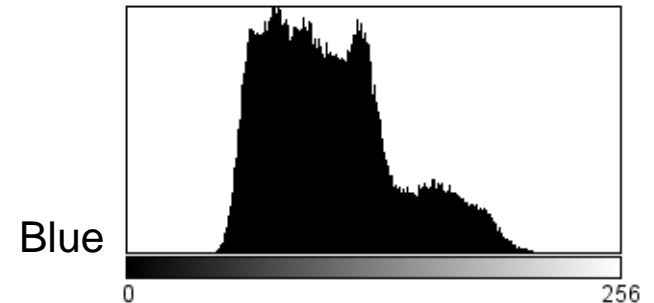
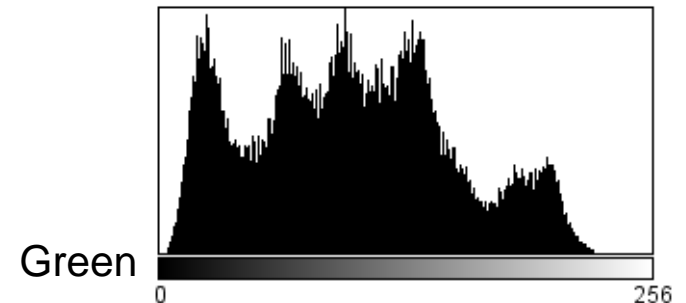
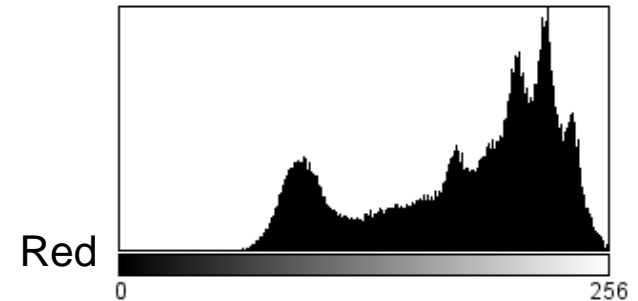
- We typically have three histograms

Ex: RGB Colour space

- Red Histogram
- Green Histogram
- Blue Histogram

- How do we build a feature vector?

- Concatenate vectors.
- Multi-dimensional quantization of colour space.



# RGB Histogram

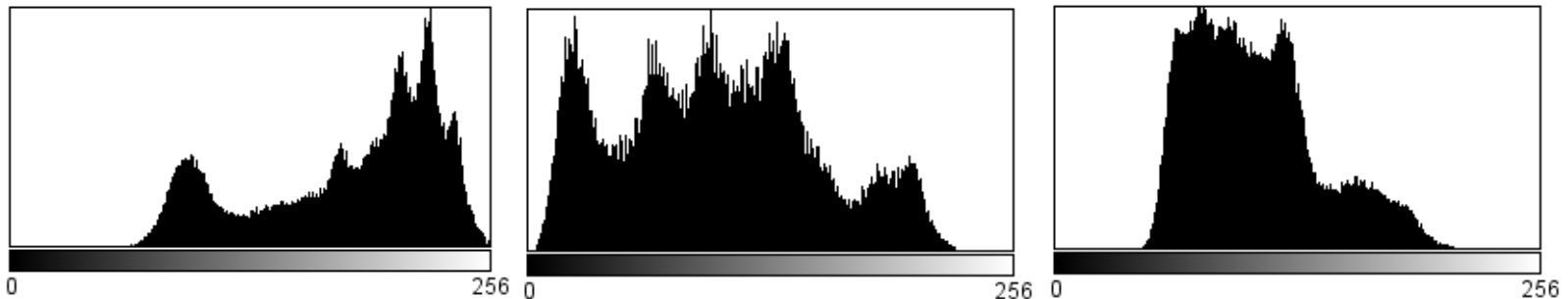
- Simply concatenate vectors.
- Not very smart. (why?)

$$F_R = [f_{R0}, \dots, f_{R255}]$$

$$F_G = [f_{G0}, \dots, f_{G255}]$$

$$F_B = [f_{B0}, \dots, f_{B255}]$$

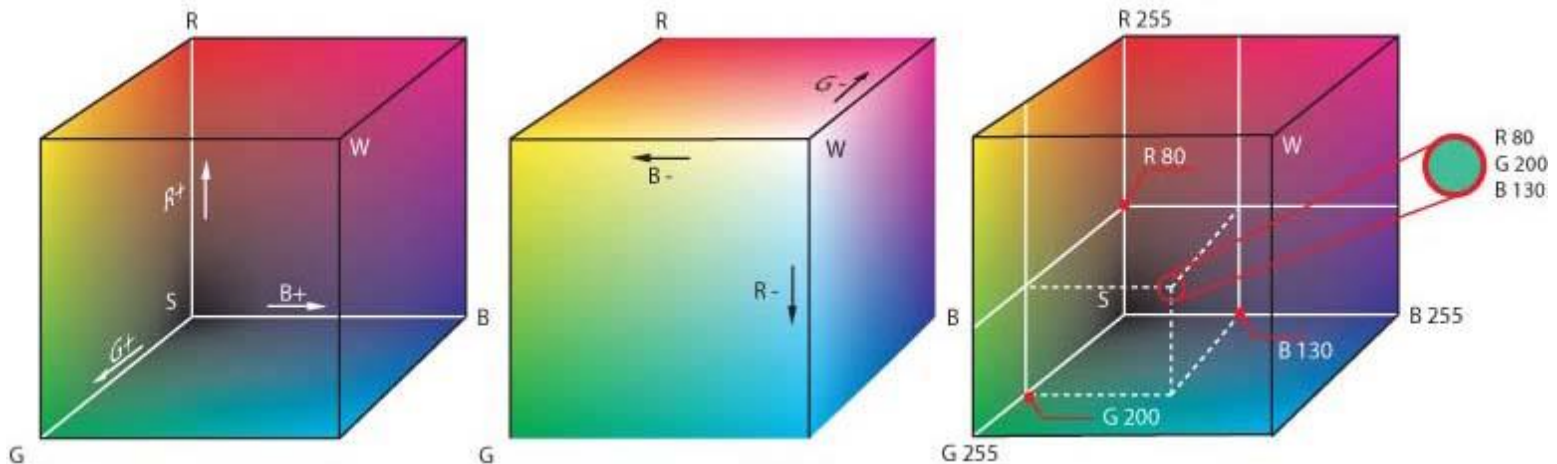
$$F_{RGB} = [F_R \mid F_G \mid F_B]$$



# Combined Histogram

- Quantize multi-dimensional colour space.
- RGB
  - Each coefficient is a small ‘cube’ inside the RGB cube.

$$F = [f_0, \dots, f_N]$$



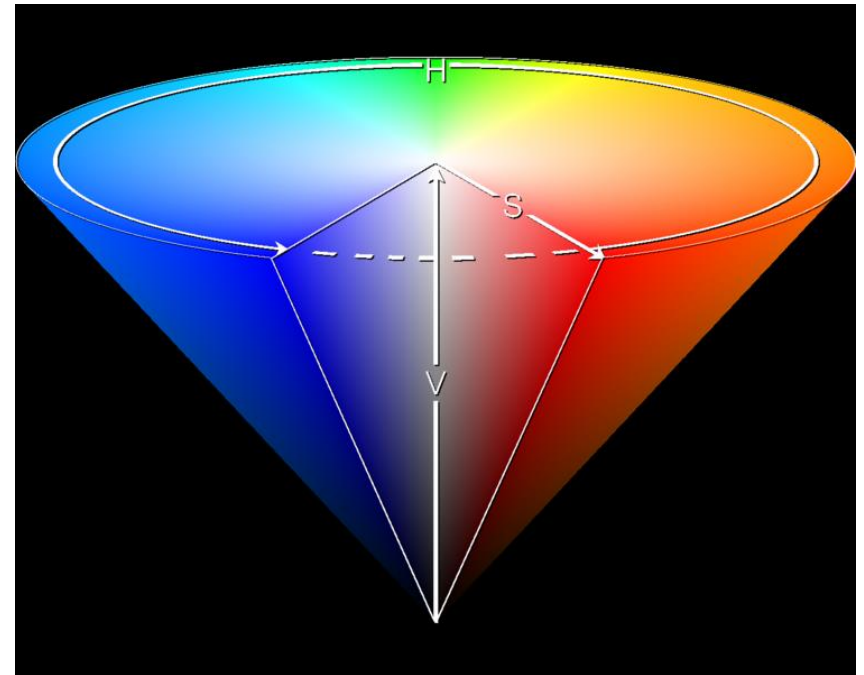


# HSI Histogram

- Quantize HSI space.
  - Define number of bins  $N$ .
  - Feature vector

$$F_{HSI} = [f_0, \dots, f_N]$$

- Typically better for object description



# MPEG-7 Dominant Colour

- Clusters colors into a small number of representative colors (salient colors)
- **$F = \{ \{c_i, p_i, v_i\}, s \}$** 
  - $c_i$  : Representative colors
  - $p_i$  : Their percentages in the region
  - $v_i$  : Color variances
  - $s$  : Spatial coherency

# MPEG-7 Scalable Colour

- HSI Histogram
- Typical quantization: 256 bins.
  - 16 levels in H
  - 4 levels in S
  - 4 levels in I
- Very popular for CBIR (Content-Based Image Retrieval).

$$F_{SC} = [f_0, \dots, f_{255}]$$

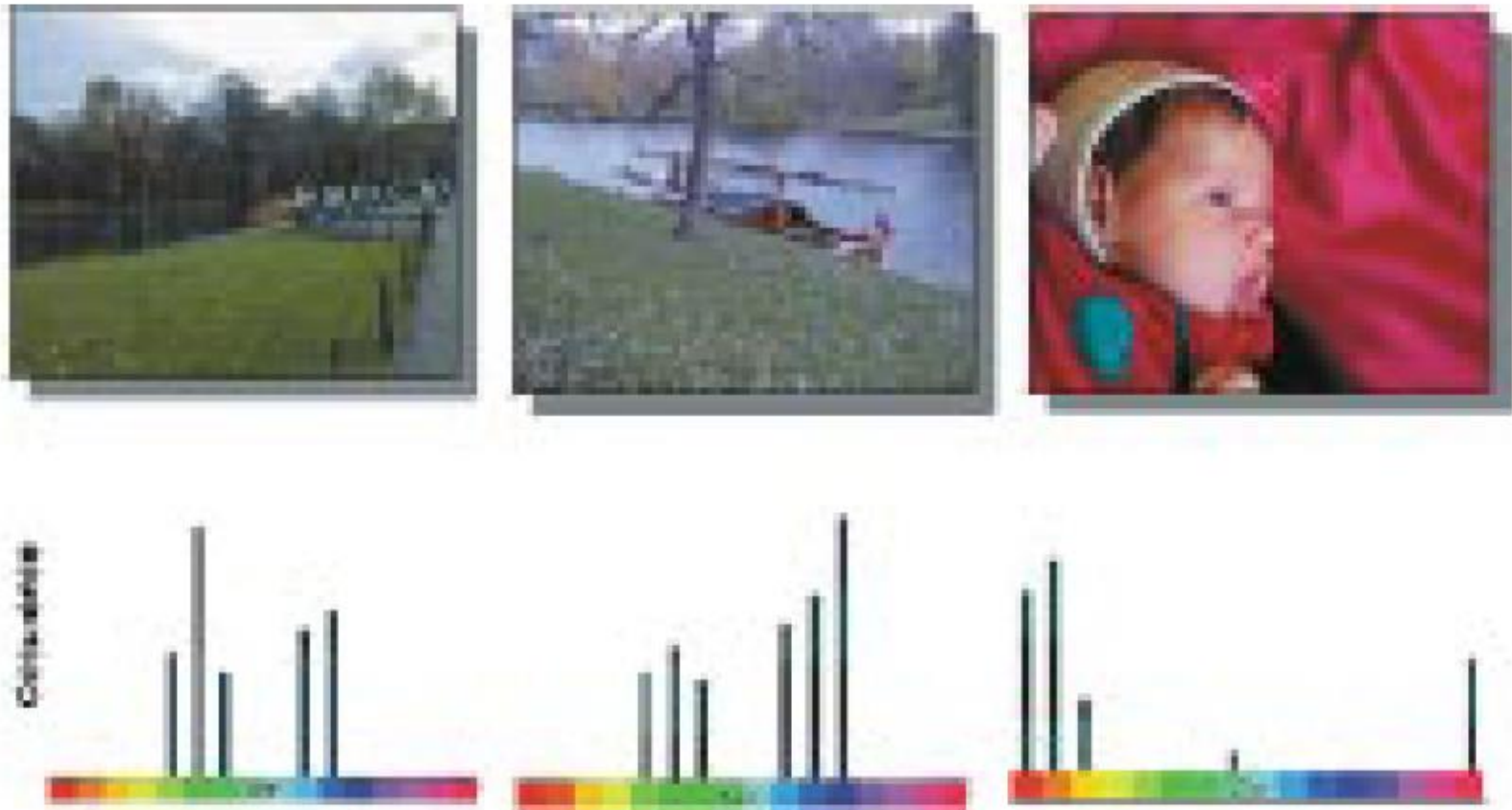
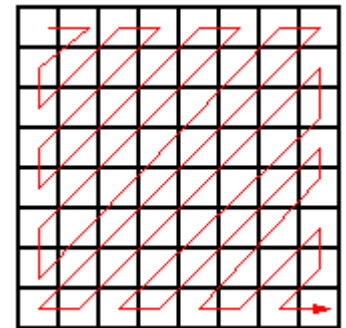


Fig. 2. Three color images and their MPEG-7 histogram color distribution, depicted using a simplified color histogram. Based on the color distribution, the two left images would be recognized as more similar compared to the one on the right.

[Sikora 2001]

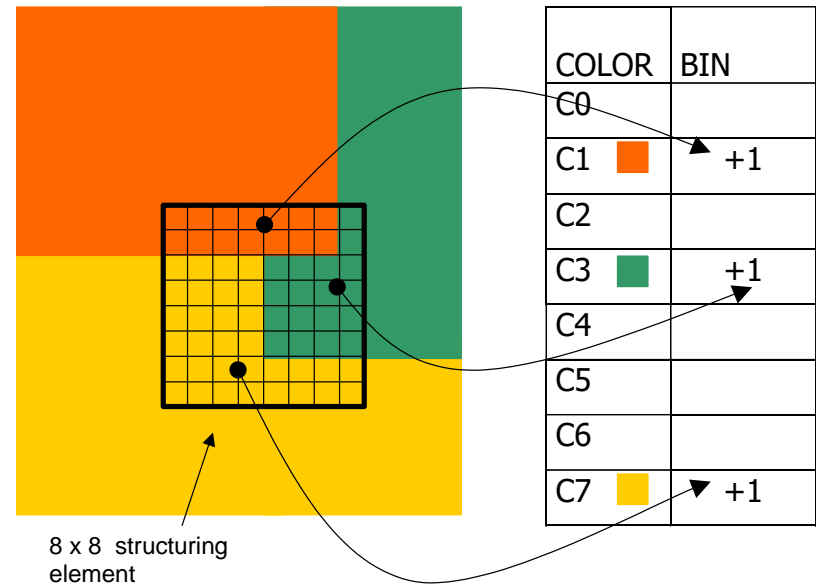
# MPEG-7 Colour Layout

- Clusters the image into 64 (8x8) blocks
- Derives the average color of each block (or using DCD)
- Applies (8x8)DCT and encoding
- Efficient for
  - Sketch-based image retrieval
  - Content Filtering using image indexing



# MPEG-7 Colour Structure

- Scans the image by an 8x8 struct. element
- Counts the number of blocks containing each color
- Generates a color histogram (HMMD/4CSQ operating points)



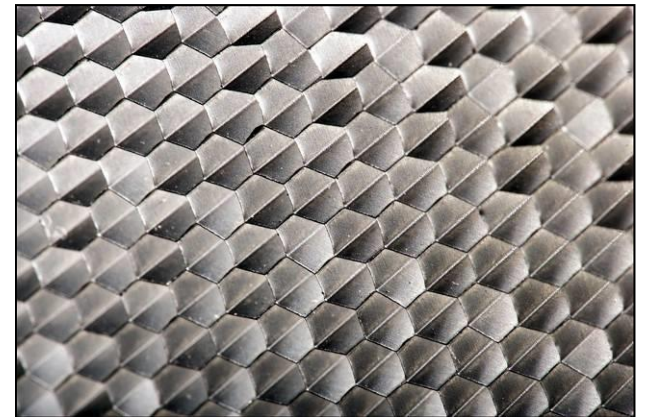
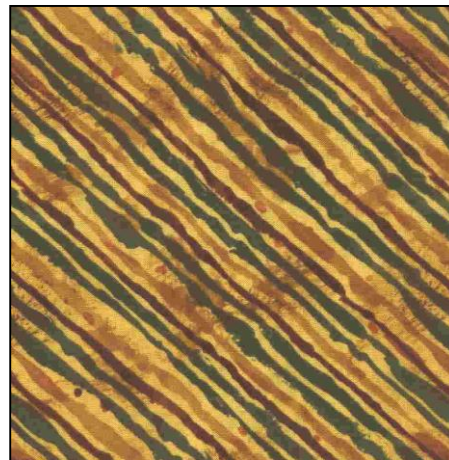
# Topic: Texture

- Feature Vectors
- Colour
- **Texture**
- Shape

# What is texture?

“Texture gives us information about the spatial arrangement of the colours or intensities in an image”.

[L. Shapiro]





# Two approaches to texture

- **Structural approach**
  - Texture is a set of primitive *texels* in some regular or repeated relationship.
  - Good for regular, ‘man-made’ textures.
- **Statistical approach**
  - Texture is a quantitative measure of the arrangement of intensities in a region.
  - More general and easier to compute.

# Statistical approaches

- Grey level of central pixels
- Average of grey levels in window
- Median
- Standard deviation of grey levels
- Difference of maximum and minimum grey levels
- Difference between average grey level in small and large windows
- Sobel feature
- Kirsch feature
- Derivative in x window
- Derivative in y window
- Diagonal derivatives
- Combine features



How do I pick one??

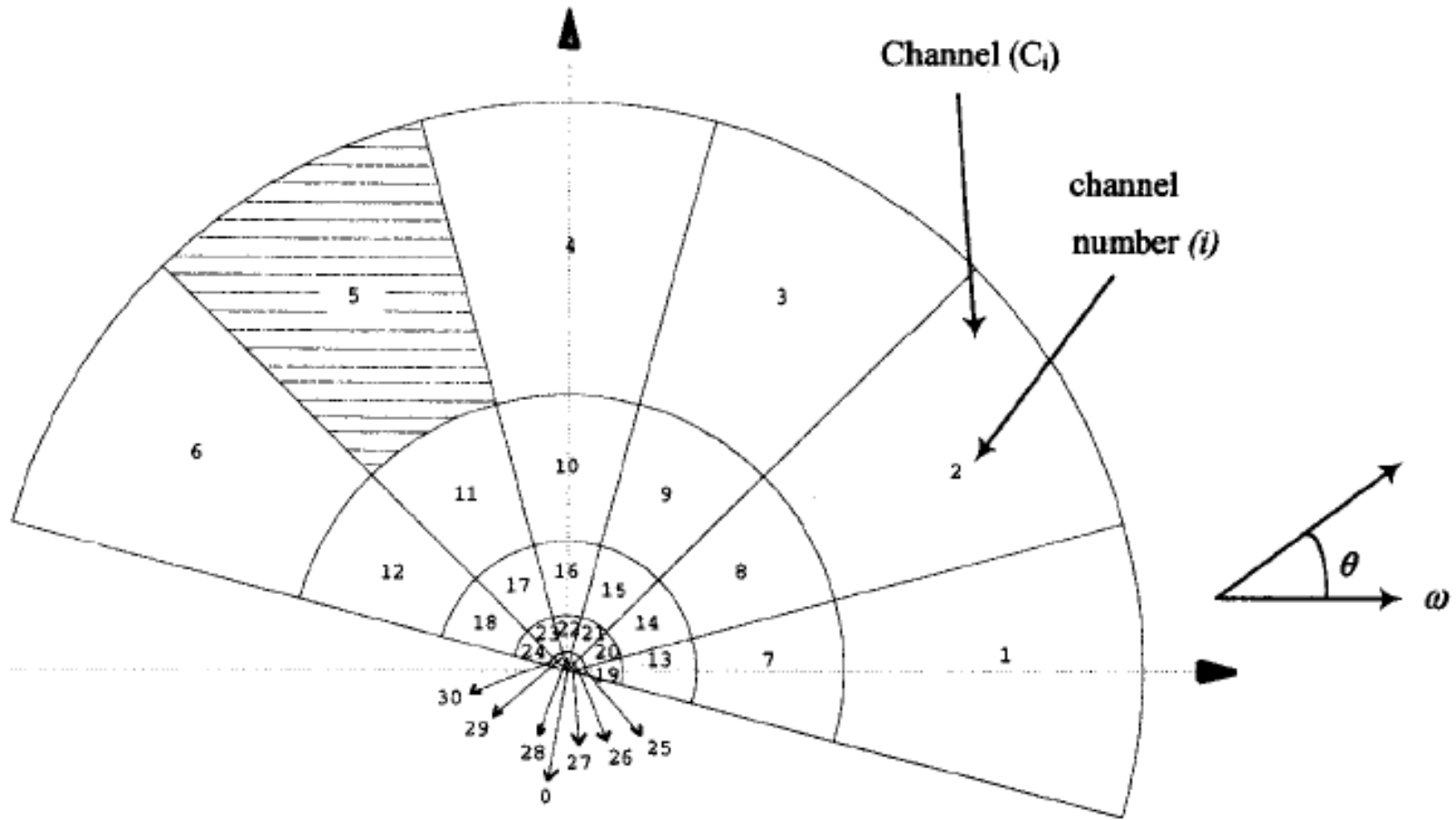
# MPEG-7 Homogenous Texture

- Filters the image with a set of orientation and scale sensitive filters.
- Computes mean and standard deviation of response.
- 30 channels
  - 6 in angular direction, 5 in radial direction.

$$F_{HT} = [f_{DC}, f_{SC}, e_1, e_2, \dots, e_{30}, d_1, d_2, \dots, d_{30}]$$

$f_{DC}$ ,  $f_{SC}$  are the mean intensity and the standard deviation of image texture), where  $e_x$  and  $d_x$  are the logarithmically scaled texture energy and texture energy deviation coefficients.

# HT Channels



# MPEG-7 Local Edge Histogram

- Image divided into 4x4 sub-regions.
- Edge histogram computer for each sub-region.
- Five bins:
  - Vertical, horizontal, 45 diagonal, 135 diagonal, and isotropic.
- 80 total bins.

$$F_{LEH} = [f_0, \dots, f_{79}]$$

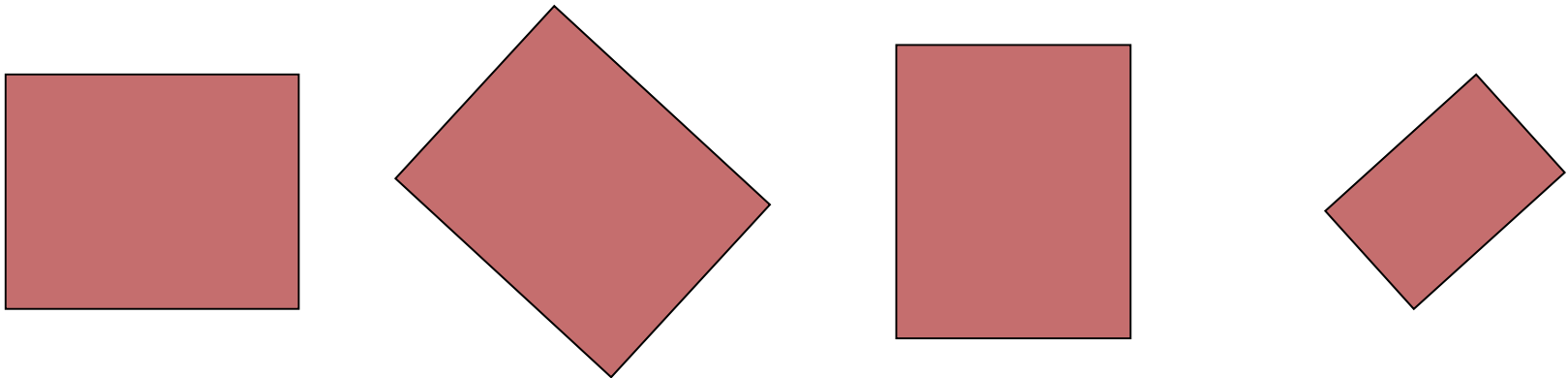
# Topic: Shape

- Feature Vectors
- Colour
- Texture
- **Shape**

# Definitions

- **Geometric definition**

Two sets have the same shape if one can be transformed into another by a combination of translations, rotations and uniform scaling operations.



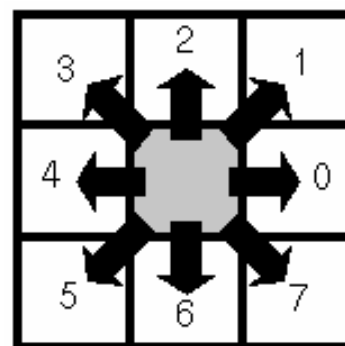
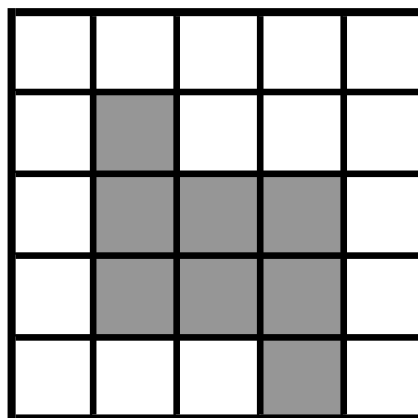
# Shape and Segmentation

- **Shape implies a segmentation step.**
  - Segmentation has multiple solutions (middle-level feature).
  - But the shape feature itself has a single solution!
- **How do we describe shapes?**
  - Chain-codes
  - Statistical descriptors.



# Freeman Chain Code

- Chains represent the borders of objects.
- Coding with *chain codes*.
  - Relative.
  - Assume an initial starting point for each object.
- How do we build a feature vector?



Freeman Chain Code

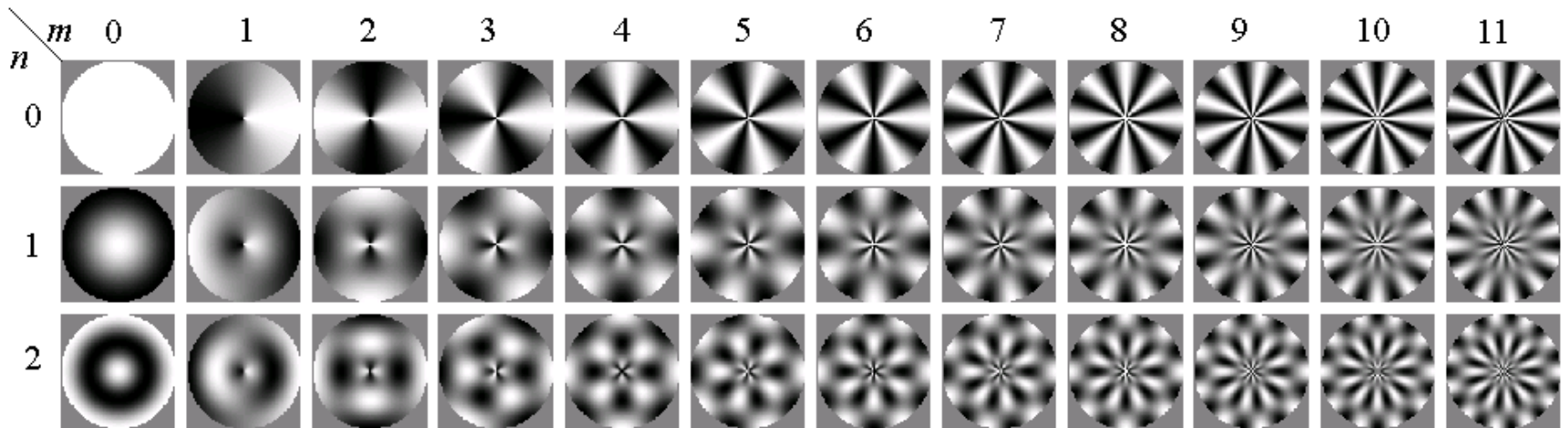
Using a Freeman Chain Code and considering the top-left pixel as the starting point:  
70663422

# MPEG-7 – Region-Based Shape

- Uses a set of separable ART (angular radial transformation) functions.
- Classifies shape along various angular and radial directions.
- Totals 35 coefficients.

$$F_{RBS} = [f_0, \dots, f_{34}]$$

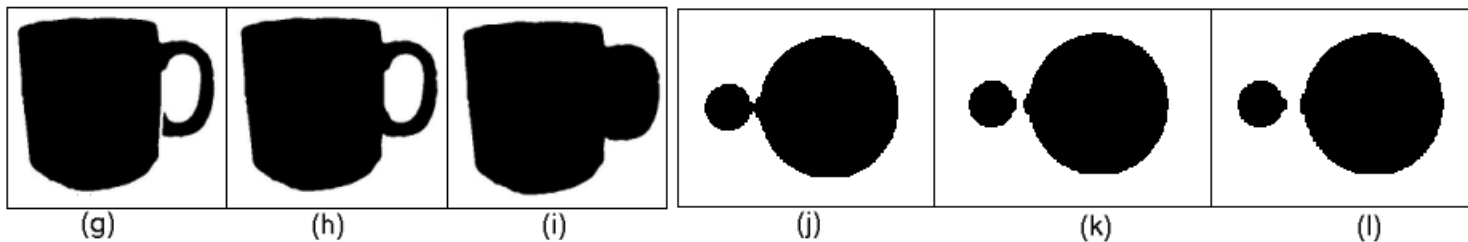
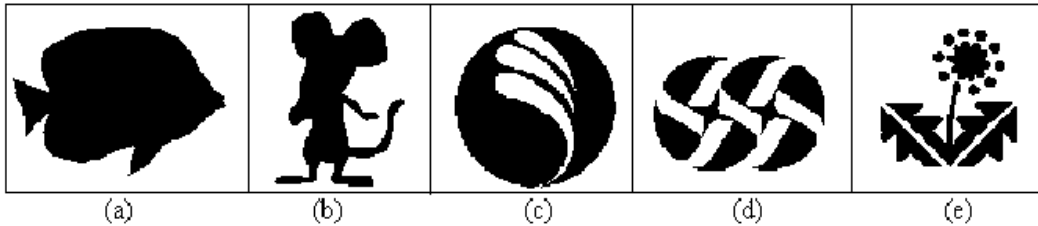
# ART Basis Functions



•Applicable to figures (a) – (e)

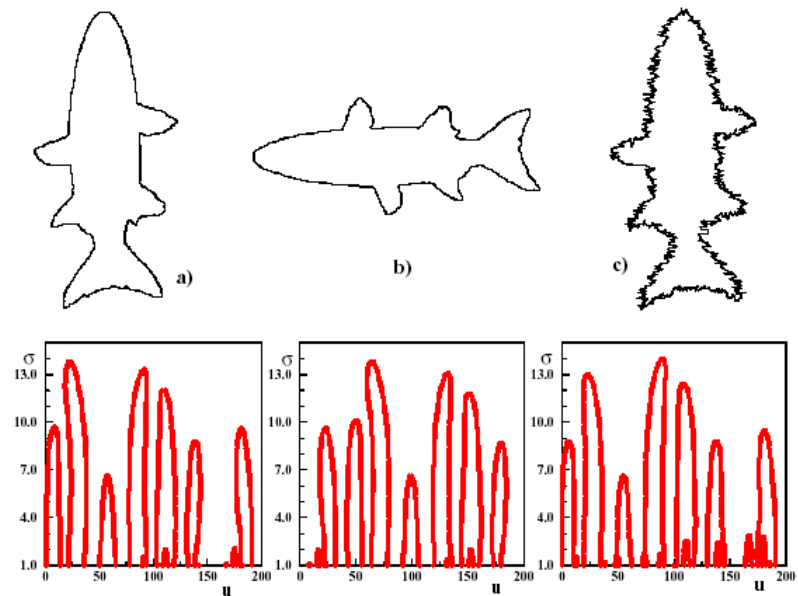
•Distinguishes (i) from (g) and (h)

•(j), (k), and (l) are similar



# MPEG-7 – Contour-Based Descriptor

- Finds curvature zero crossing points of the shape's contour (key points)
- Reduces the number of key points step by step, by applying Gaussian smoothing
- The position of key points are expressed relative to the length of the contour curve



- Applicable to (a)
- Distinguishes differences in (b)
- Find similarities in (c) - (e)



(a)



(b)



(c)



(d)

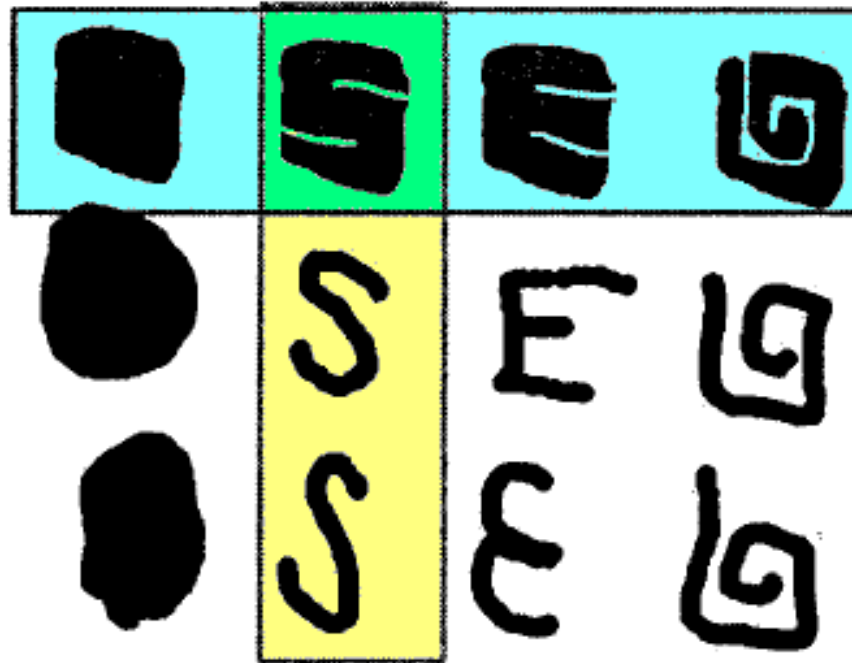


(e)

### Advantages:

- Captures the shape very well
- Robust to the noise, scale, and orientation
- It is fast and compact

# Comparison



- Blue: Similar shapes by Region-Based
- Yellow: Similar shapes by Contour-Based

# Resources

- L. Shapiro, Chapters 6 and 7
- T. Sikora, “MPEG-7 Visual Standard for Content Description—An Overview”,  
<http://ieeexplore.ieee.org/iel5/76/20050/00927422.pdf?arnumber=927422>