

Information Retrieval

Multimedia Retrieval

Open Video example

Overview

- Generations of multimedia systems
- Feature-based representations
- Multidimensional indexing
- Features in 2D still images
- OpenVideo Project example

Multimedia Data Types

- Audio
 - speech, music, environmental sounds
- Images
 - 2d still [+ text + audio]
 - Video [+ text + audio]
 - 3D images [+text + audio]
- Other sensory signals
 - Haptic, olfactory; changes in vision

Who needs it?

- Associative reasoning, recognition
 - Journalism; Security & Law Enforcement; Data Mining
- Classification
 - historical site retrieval, archeology;
 - chemistry (structure of molecules), etc.
- Design
 - architecture, decor, industrial design
- Future? [Why not your own life?!]

Multimedia databases

- Art galleries, museums
- Photo archives
- Textiles
- Medical images
- Facial detection db
- Engineering design db
- Geographic
- Satellite image databases

Key Issues in Multimedia IR

- Raw data representation [limited use; last time]
- Query generation [how to generate a query for images, sounds?]
- Large number datasets of each object
 - and redundancy
- Same content, same representation, but different only in scale

Raw Data Representation

- Image file: a list of pixels [compare pixels to characters or bits]
- Video: a list of images
- Audio: list of samples

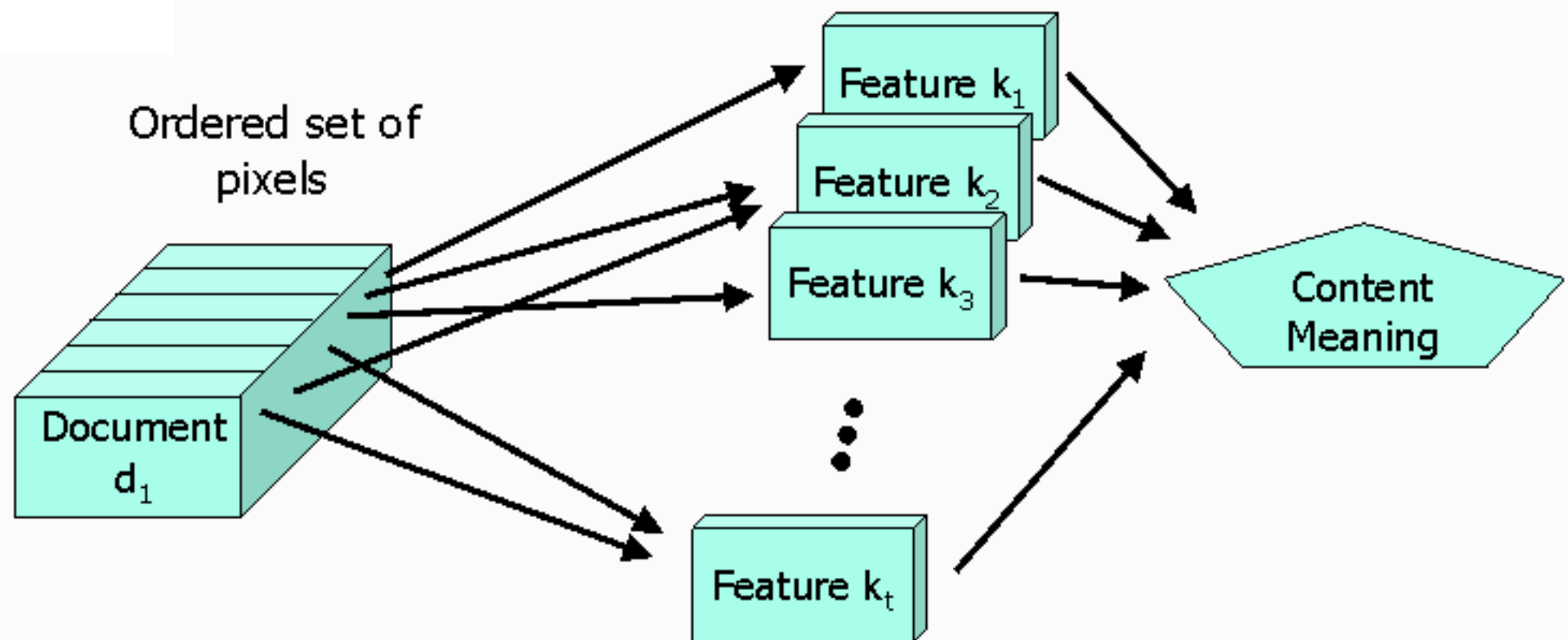
First Generation Image Retrieval

- Images & video are described by metadata
 - figure captions, keywords, associated paragraphs
- Retrieval based on text
 - keywords; text features

Retrieval of 2D still images

- Goal: retrieve based on the document (image) content:
 - physical and perceptual properties
 - semantic properties (e.g., objects)
 - abstract properties (events, emotions, intent)
- Level of abstraction

Set of features



Vector Space Representation

- $d = \{f_1, f_2, \dots, f_n\}$
- Dimensions represent the “amount of features” in the multimedia document
- Query is represented as a vector
- Retrieval is governed by similarity in vector space (cosine of the angle or Euclidian distance)

Features (for retrieval)

- Feature: real number representing the presence or absence of a “feature”
- ID feature: each document represented by a single number (term, color, face, ...)
- More than one feature → multidimensional representation
- A generalization of term vectors, each element corresponds to a feature!

Indexing multidimensional data [optional]

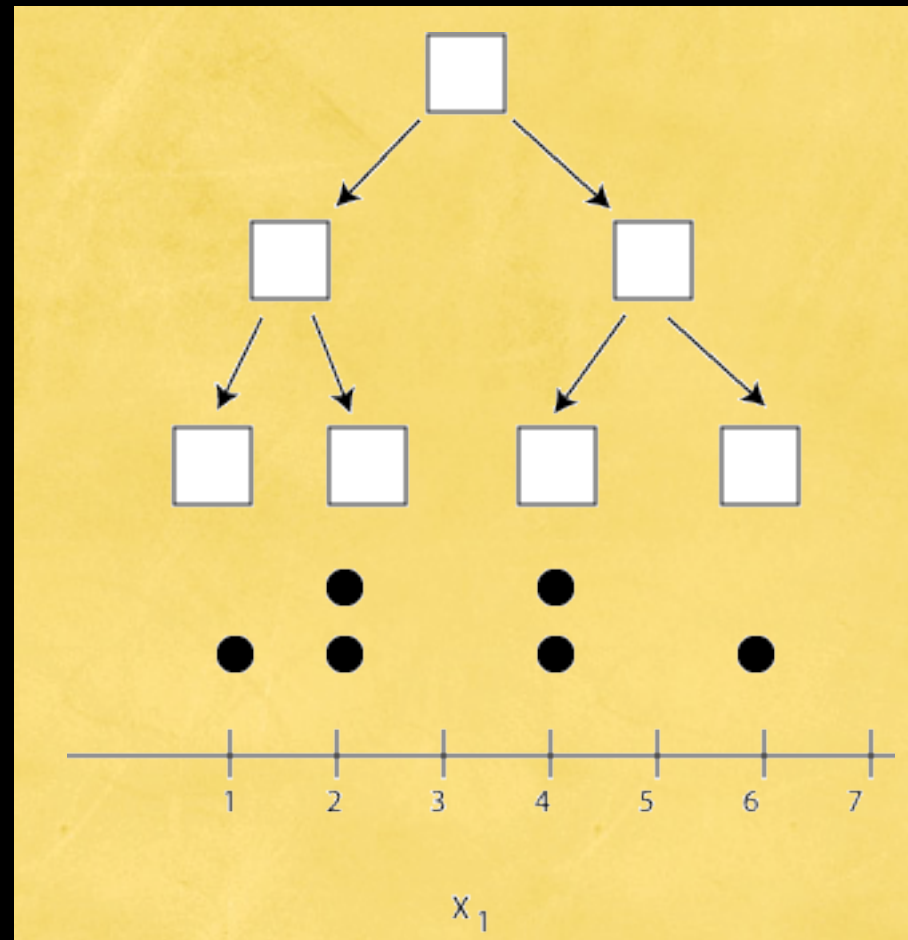
- Implementing indexing analog to inverted indexing for terms:
- K-d terms
- R-trees
- R* trees
- R+ trees
- SS-trees (weighted Euclidian distance)

Example: Retrieval in Vector Space

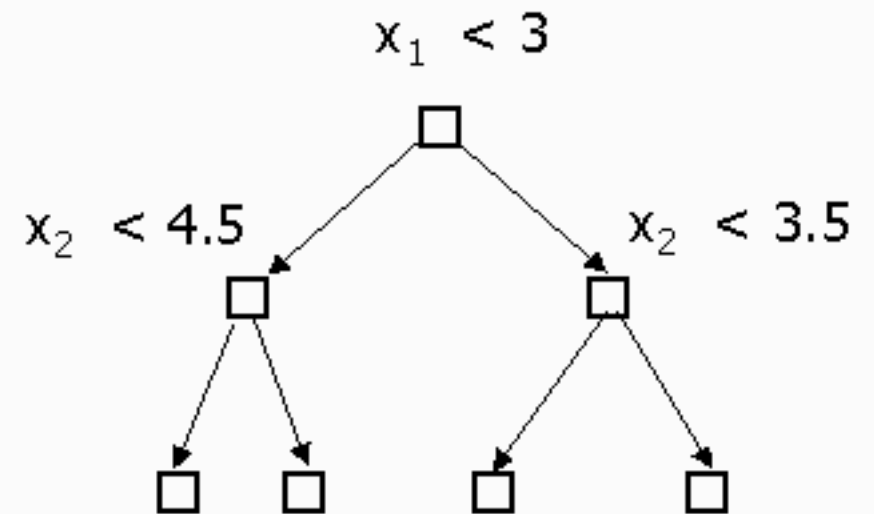
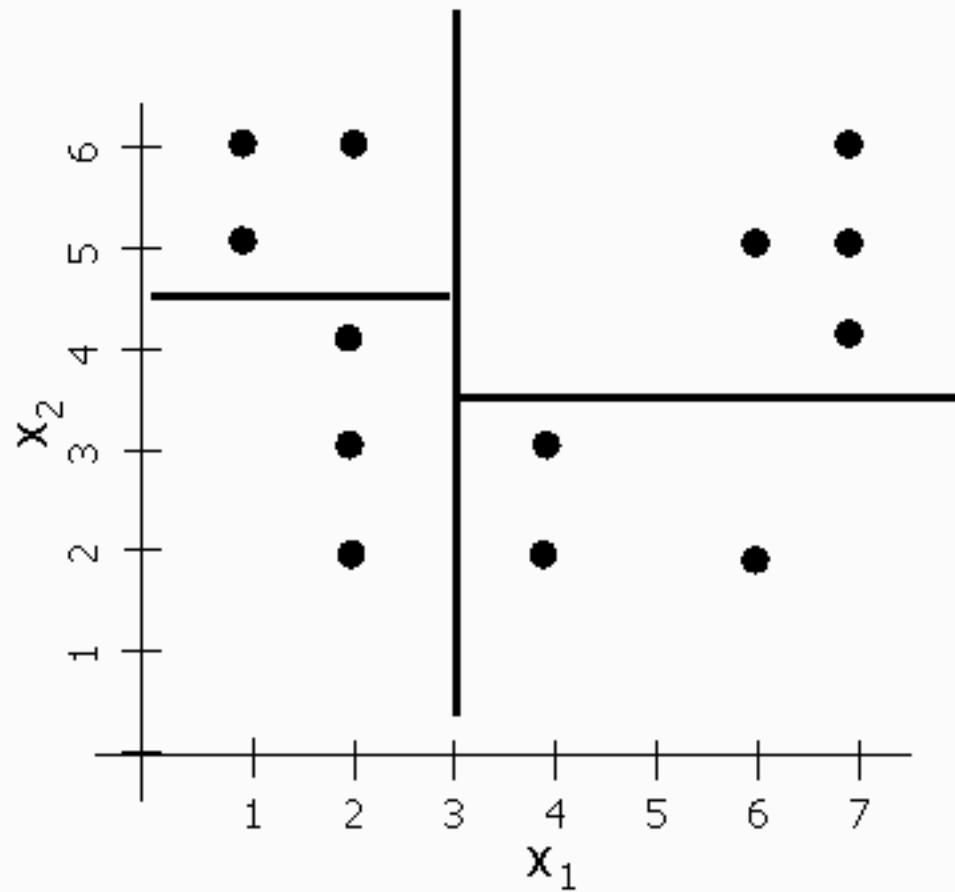
- Imagine the data reflect features...
- Data: (1,6)(1,5)(2,6)(2,4)(2,3)(2,2) ...
- Query: (3,5)
- Range query: $(3,5) \pm (1,1)$ [(2,4) - (4,6)]

I-dimensional indexing

- Similar to a binary search: keep dividing [whatever you're given] 'til you can't go further:



2D K-d Tree Example



2D R-Tree example

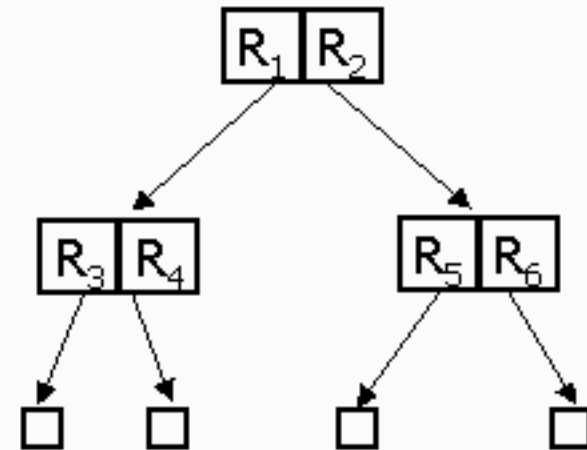
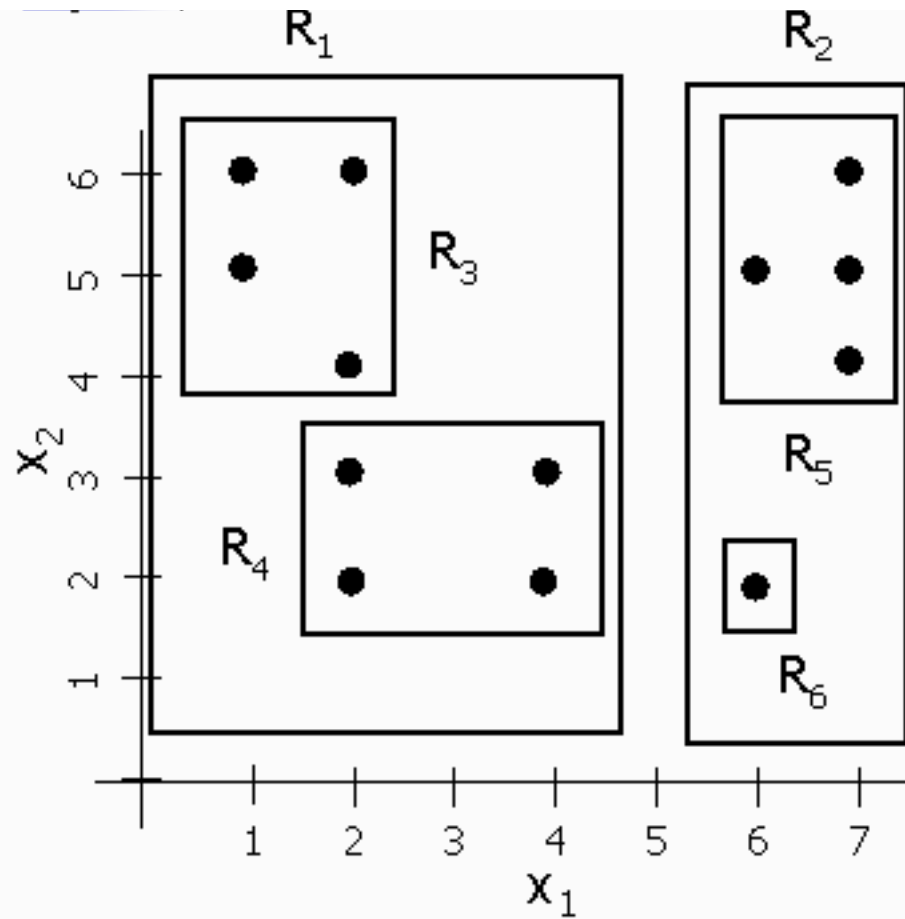


Image & Video Features

Image & Video Features

- Color of the pixels
 - Line segments and edges (remember Blobworld)
 - Texture
 - Shape
 - Motion ...
-
- But what would make a feature useful for IR?

Potentially good features: invariant representations

Potentially good features: invariant representations

- Invariance with respect to geometric transformations:
 - translation, rotation, scale, projection

Potentially good features: invariant representations

- Invariance with respect to geometric transformations:
 - translation, rotation, scale, projection
- Invariance with respect to illumination:
 - intensity (of the color)
 - Limited color variation

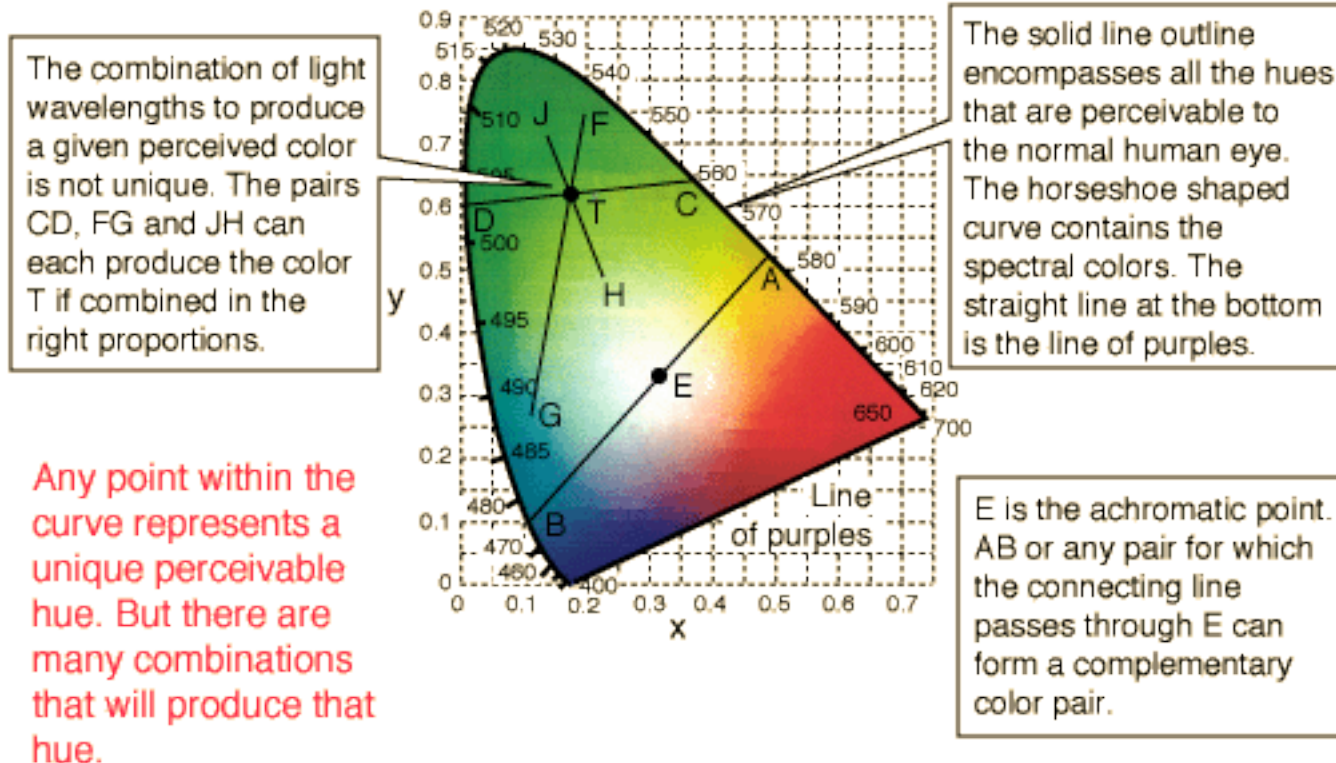
Potentially good features: invariant representations

- Invariance with respect to geometric transformations:
 - translation, rotation, scale, projection
- Invariance with respect to illumination:
 - intensity (of the color)
 - Limited color variation
- The idea is that if the feature is good for IR, it is likely a good feature for human recognition

Representations of Color

- What *is* color?
- 3D color spaces
 - perception-based models
 - CIE chromaticity (x,y,z)
 - Opponent color model: Luv
 - Hue, saturation, values or brightness
- Hardware-oriented models; RGB, CYMK

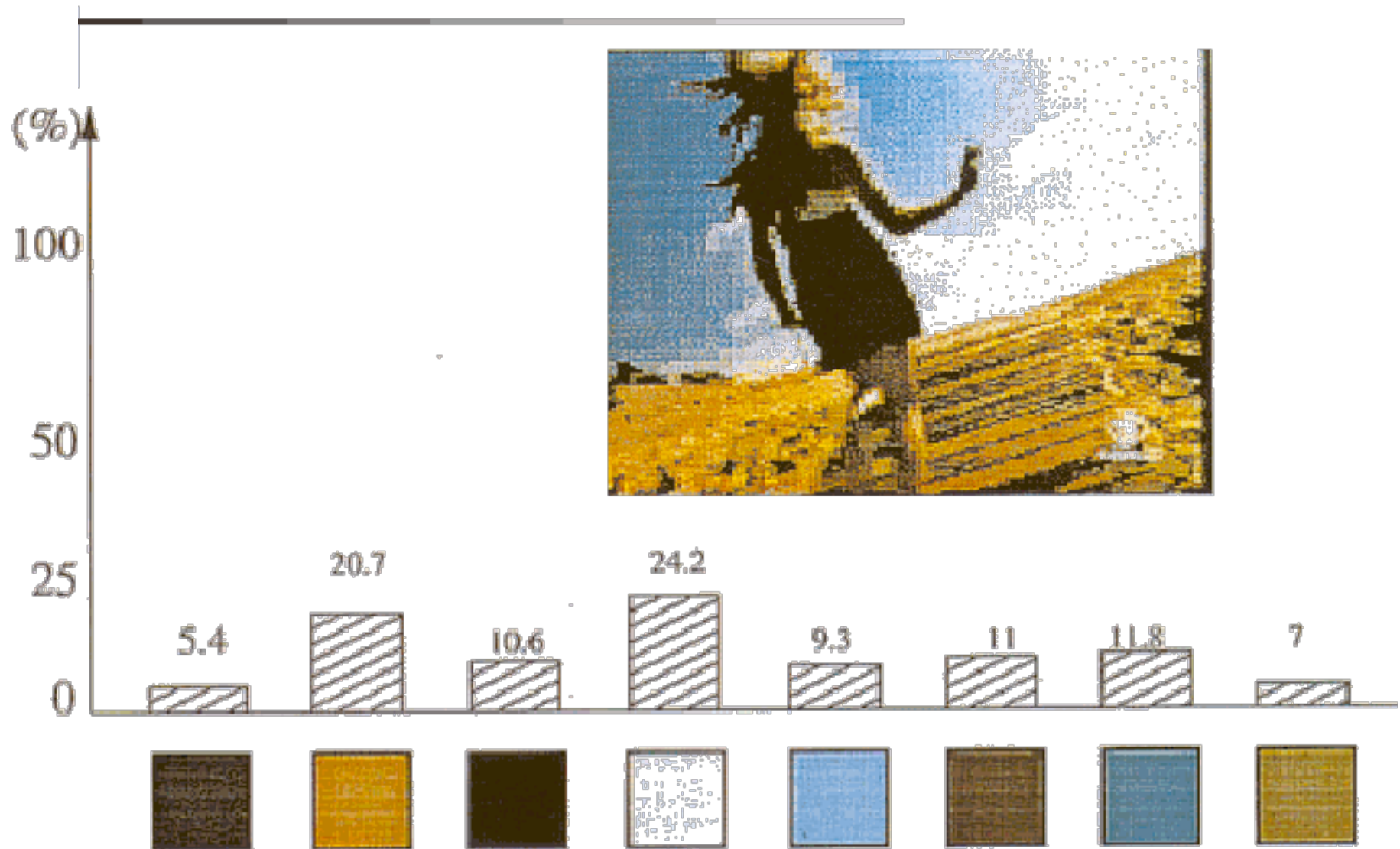
Color



Representation of images with color histograms

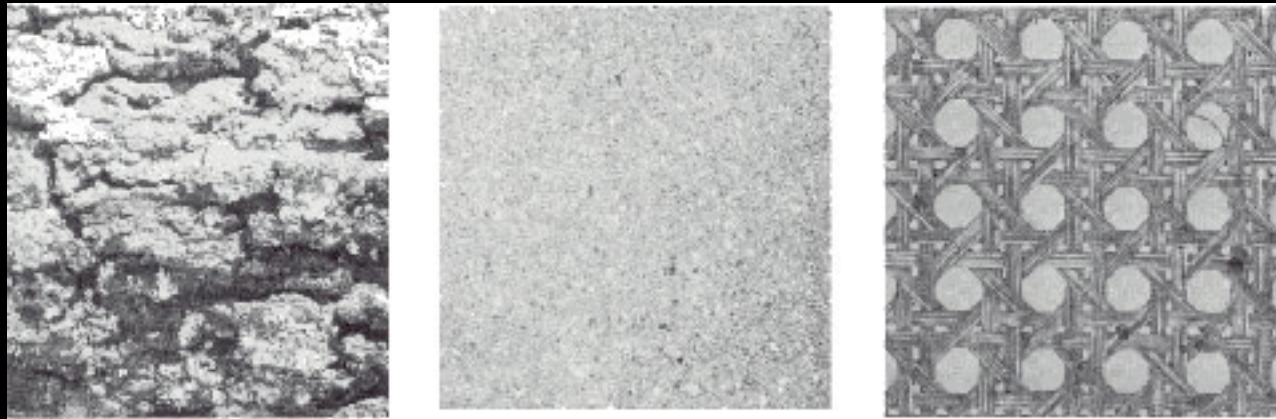
- Relative frequency distribution of each color dimension [compare to term distribution]
- Compute similarity between corresponding histograms of each color dimension [as if a query term has synonyms; search for all of them, including the query]

Example: Color Histogram



Texture representation

- average pattern properties
 - Pixels: co-occurrence matrix, Markov, others
 - Patterns: contrast, orientation, directionality of edges and line segments, principal components analysis

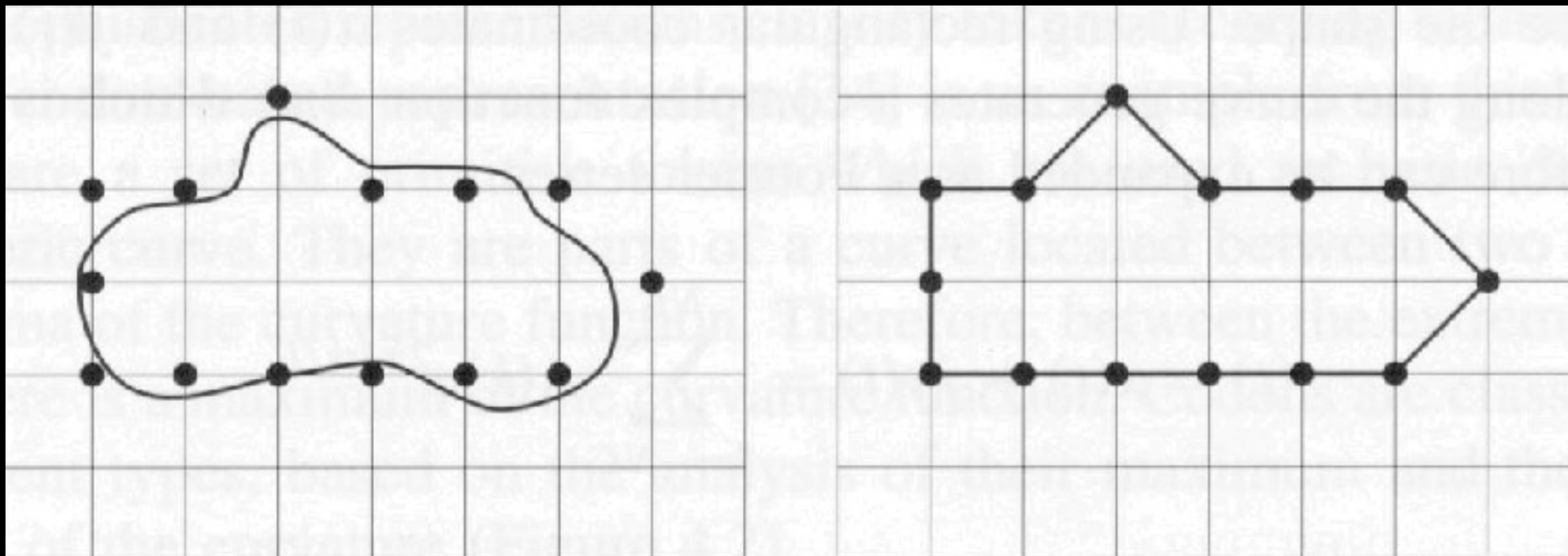


What to look for...

- Shade representations [to find boundaries]
- Edge detection [corners, edges, links, regions]



Chain code representation



Video Features

Video Features

- Segments, scenes, and basic frames
- Transitions (differences between 2 frames)
- Motion
- Compression standards:
 - MPEG2 - region coding and motion compensation
 - MPEG4 - content-based compression and synthetic data representation
 - MPEG7 - standardization of structures and arbitrary description schemes

Multimedia User Interface

- How to input (query) images and sound?
- Content visualization, browsing, and non-linear editing
- It's believed if you compress the file, you remove redundancies ... improves IR
- QBIC <http://www.qbic.almaden.ibm.com/>
- Virage
- Excalibur <http://www.excalib.com/>

Example

- Open Video Project
 - <http://www.open-video.org/>