Multimedia Information Systems

Samson Cheung

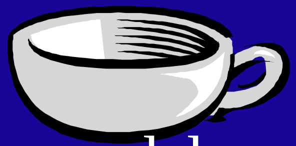
EE 639, Fall 2004

Lecture 6: Content-based Image Retrieval

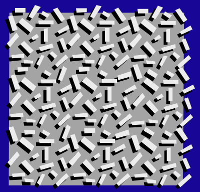
Basic + Color Feature

How to search an image collection?

- Text based
  - Subject navigation
  - Keyword based search
- Visual
  - Models
  - Features



model

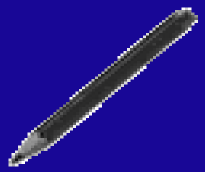


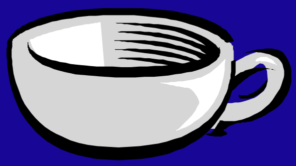
texture

Optimal choice depends upon the task

Content based retrieval: origins

- **CBR has its origins in computer vision – geometric shape hashing**
- **Computer Vision: model driven**
  - **Create a model for each object or class**
  - **geometric hashing**
- **Problems:**
  - **Too many models!**
  - **3D models are computationally intensive**

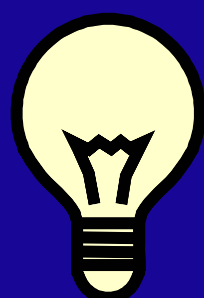


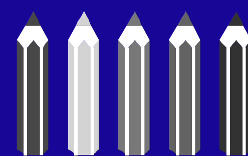


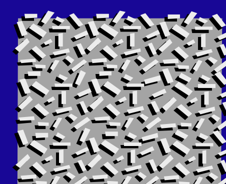
Q: Can we do away with models?

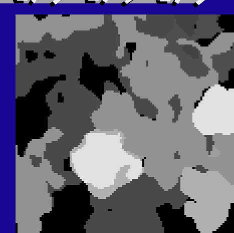
Query By Information Content

- Worked on images directly: No models
- Used features derived from the images:
  - Color
  - Texture
  - Shape
  - Composition







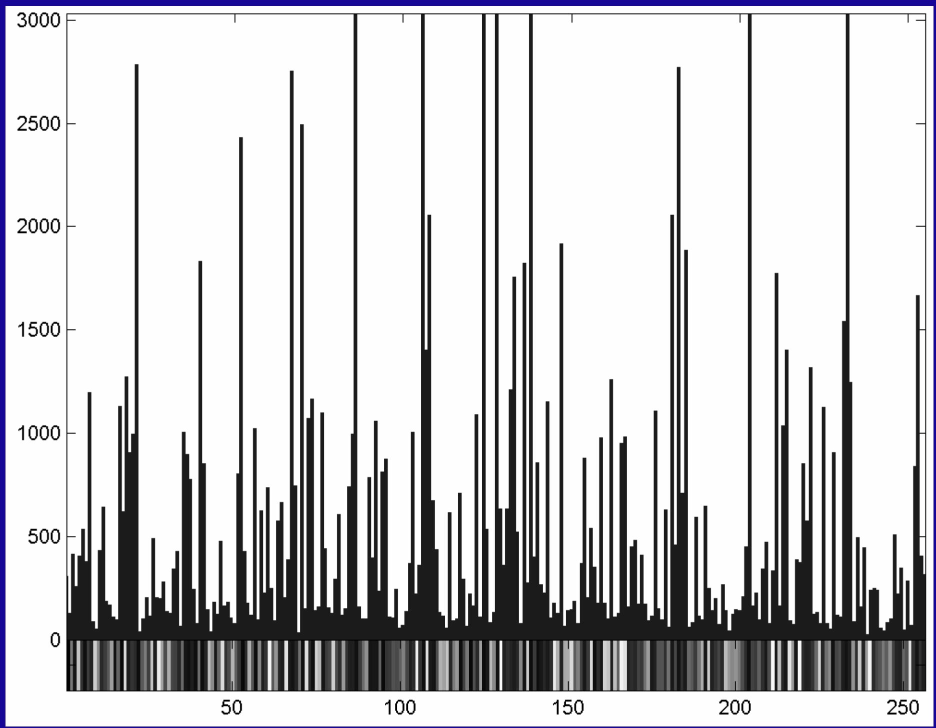


A feature example: histograms

- A histogram is count of the “amount” of a particular color in an image
- Indexed vs. non indexed images
  - indexed images have a colormap (e.g. gif) – the index functions as a look up table
  - non-indexed images (e.g. jpeg) need to be converted into an index
- The conversion involves quantization of the color space of the image
  - Uniform quantization
  - Non-uniform quantization

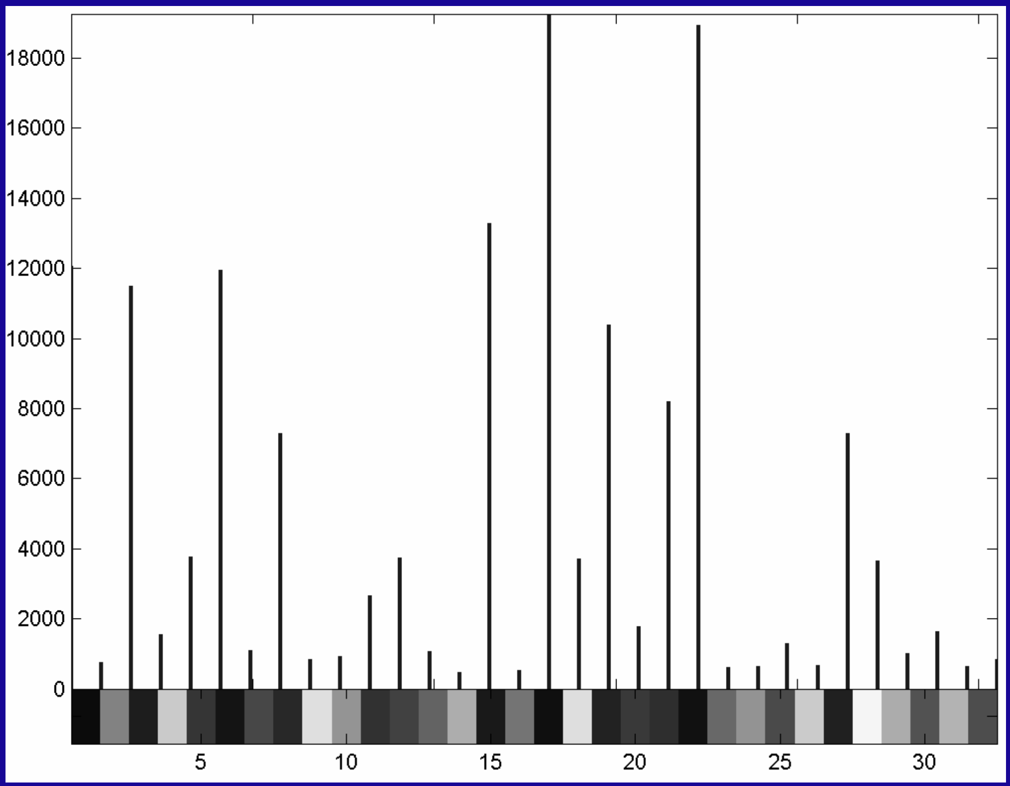
Histograms #1

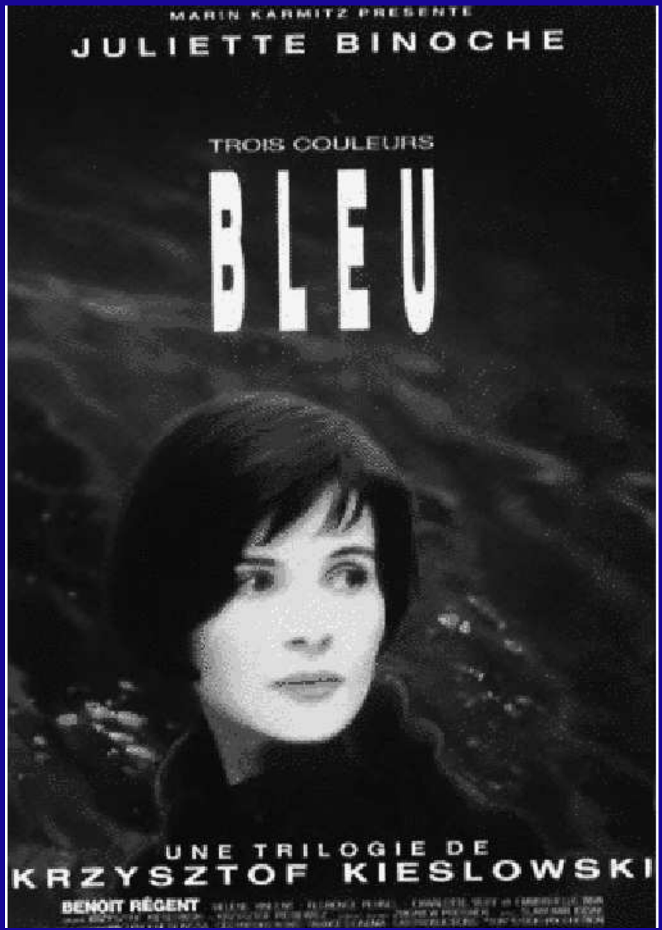
- **256 bin quantized image**



histograms #2

- 32 bin quantized image





An Example

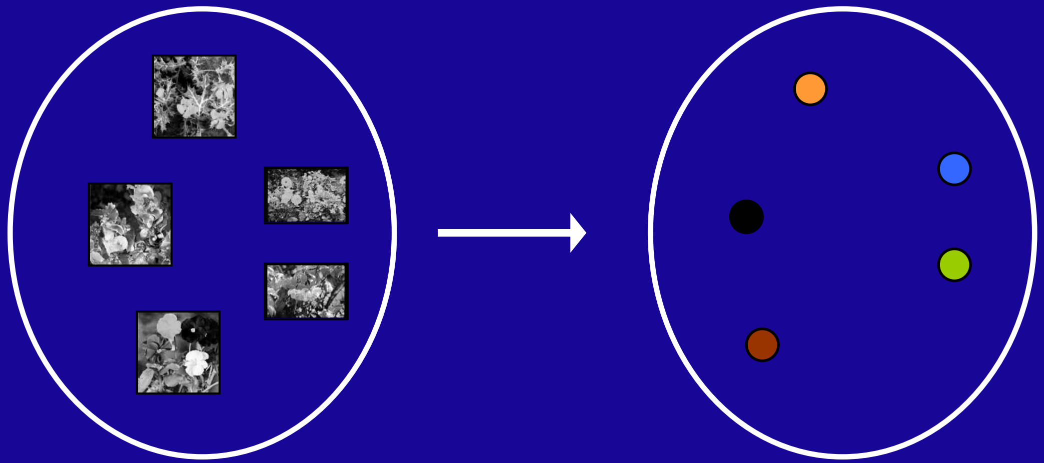
Query:

Results:

Color Histograms

How?

- Images are represented as high-dimensional feature vectors (e.g. color histogram) some space.
- So, given a query:
  - find nearest neighbors
  - these are the “closest” matches



Key ideas in CBR

- What features?
- How to measure similarity?
  - Distance (discriminative) vs. Clustering (generative)
- fast indexing
  - Sequential is too slow
- interfaces
  - query formulation
  - result visualization
- feedback
  - query refinement vs. adaptive metrics
  - As few iteration as possible
- evaluation

Common characteristics of visual features

- Local or Global
- Complexity
  - Extraction – raw/compressed domain
  - Dimensionality – for indexing
- Invariance
  - rotation + translation + scaling + occlusion + illumination
- Relevance
  - expert / machine-learning
- Any reasonable metric?

Color

- Perceptual color spaces:
  - Munsell system (~ HSV)
  - Hunter (LAB) space
  - Invariant color ratios
- Pros:
  - easy to understand; samples available
  - easy to use and compare colors
  - number, and sampling can be adapted to application
- Cons:
  - too many color systems! ☹ cannot easily translate
  - color under a particular illuminant
  - difficult to apply to monitors (self-luminous displays)

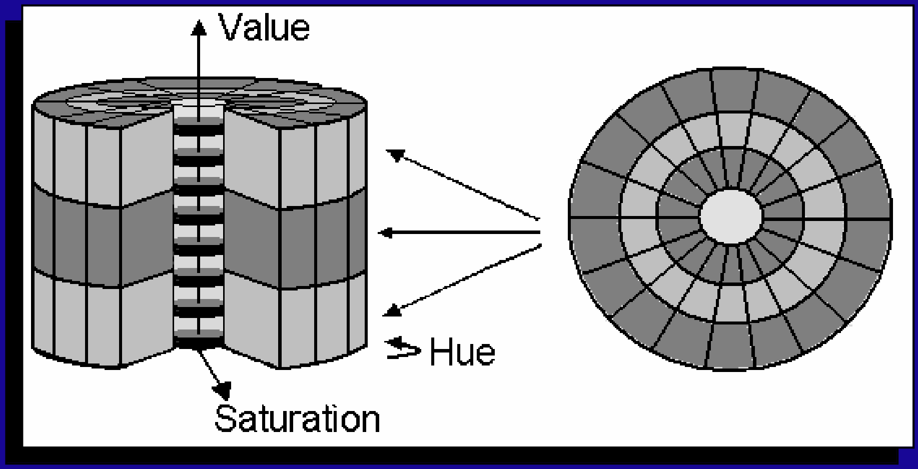
Color Invariance (Gevers & Smeulders 99)

- Insensitive to the varying image conditions
  - Illumination
  - Object pose – shadows, shading, highlighting
- Assumption : narrow-band sensors
- Any one of these models
  - $N(C^x, C^y) = (C^x - C^y) / (C^x + C^y)$
  - $F(C^x, C^y) = C^x / C^y$
  - $M(A^x, B^x, A^y, B^y) = A^x B^y / A^y B^x$

where $A, B, C \in \{R,G,B\}$; x and y denote neighboring pixels

Color Space quantization

- QBIC (IBM)
  - 16m (rgb) → 4096 rgb → 64 munsell colors
- VisualSeek (columbia)
  - 16m (rgb) → 166 HSV (18h, 3s, 3v + 4 grays)
- Independent vs. Joint quantization

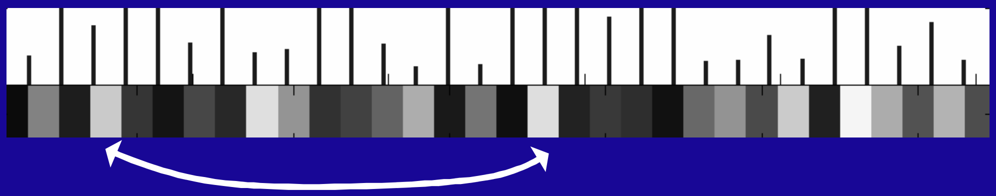


Histograms (again!)

- computes the distribution of colors amongst the bins (the dominant colors) in the image
  - the histogram is the feature vector
- pros:
  - compact
  - global distribution
  - simple Euclidean metrics
- cons:
  - no spatial information
  - high-dimensional
  - What about IDF?

histogram metrics

- l_1 distance: $$d_1(x,y)=\sum_n \left|h_x(n)-h_y(n)\right|$$
- l_2 distance: $$d_2(x,y)=\left(\sum_n \left|h_x(n)-h_y(n)\right|^2\right)^{1/2}$$
- Intersection: $$d_I(x,y)=1-\sum_n \min\left[h_x(n),h_y(n)\right]$$
  - $=0.5*l_1$
- Quadratic $$\sum_{n=1}^{N}\sum_{m=1}^{N}(h_x(n)-h_z(m))(1-\alpha_{n,m})(h_x(n)-h_z(m))$$
  - crosstalk



α could represent the perceptual distance amongst the colors

Incorporating spatial relationship

- Color Coherence Vector

- Color Correlogram

Color Coherence Vectors

(Pass, Zabith 99)

$$\begin{bmatrix} 2 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 2 & 1 & 1 \\ & & & & & \end{bmatrix} \longrightarrow$$

Color Quantization
Region Segmentation
Labeling

$$\rightarrow \begin{bmatrix} B & C & B & B & A & A \\ B & B & C & B & A & A \\ B & C & D & B & A & A \\ B & B & B & A & E & E \\ B & B & A & A & E & E \\ B & B & A & A & E & E \end{bmatrix}$$

Region Labels

	A	B	C	D	E
Color	1	2	1	3	1
size	12	15	3	1	5

Color Histogram

Color	1	2	3
α	17	0	0
β	3	0	1

$$G_I = \langle (\alpha_1, \beta_1) \cdots (\alpha_n, \beta_n) \rangle$$

$$G_{I'} = \langle (\alpha_1', \beta_1') \cdot (\alpha_n', \beta_n') \rangle$$

$$\Delta_G \triangleq \sum_{i=1}^{n} |\alpha_i - \alpha_i'| + |\beta_i - \beta_i'|$$

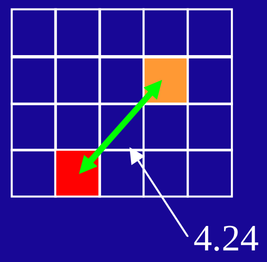
$$\Delta_H \triangleq \sum |(\alpha_i - \alpha_i') + (\beta_i - \beta_i')|$$

$\Delta_G > \Delta_H$ by triangular inequality

Color Correlogram (Huang et al 97)

- **3-dimensional table :**

  P (Red, Orange, 4.24) = Probability of



- **Hugh table :**
  - **Limit Range**
    - use small distances only (local)
  - **Dimension reduction**
    - only look at same color (color autocorrelogram)