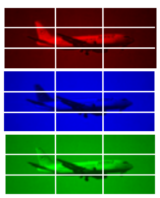
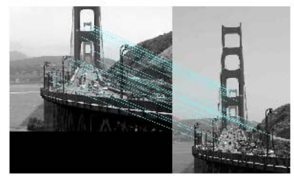
EE 6882
Visual Search Engine

Prof. Shih-Fu Chang, Jan. 30, 2012

Lecture #2

- Visual Features: Global features and matching
- Evaluation metrics





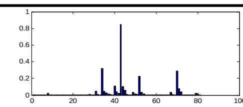
(Many slides from A. Efors, W. Freeman, C. Kambhamettu, L. Xie, and likely others)
(Slides preparation assisted by Rong-Rong Ji)

Course Format

- Lectures + two hands-on homeworks (due 2/13, 2/27)
- Mid-term project
  - Review and implement topics of interest, 2 students each team
  - Proposal due 3/5, narrated slides due 3/26
  - Selected projects presented and discussed in class (3/26-4/9)
- Final project
  - Extension of mid-term projects encouraged, 2 students each team
  - Proposal due 4/2, narrated slides due 4/30
  - Selected projects presented and discussed in class (4/30-5/7)
- Grading:
  - Class participation (20%), homework (20%), mid-term (20%), final (40%)
  - Late policy: a total “budget” of 4 days for late submissions. No other delays accepted.

2

Image Features



merl.com

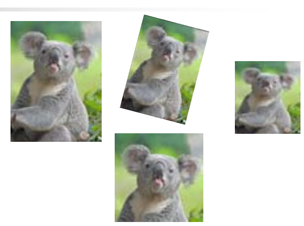
photoguides.net

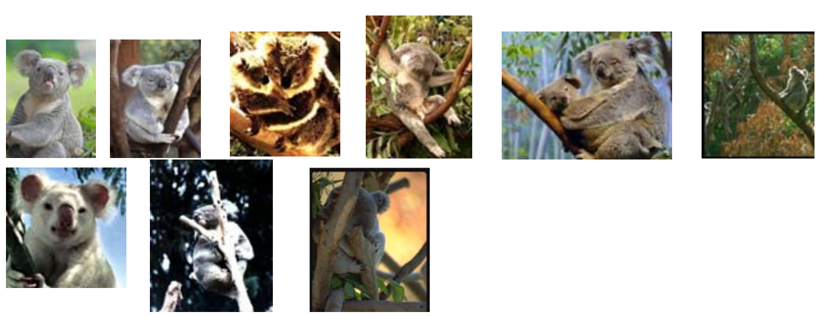
- Why features are needed?
  - Finding similar images in database
  - Classifying images to categories
  - Tracking objects in video
  - Creating panorama
  - Stereo matching -> 3D
- Desired properties
  - Compact (~100 – 1000 dimensions)
  - Easy to compute (30 fps for video)
  - Robust (invariant to photometric, geometric, content variations)

3

Desired Properties of Visual Features

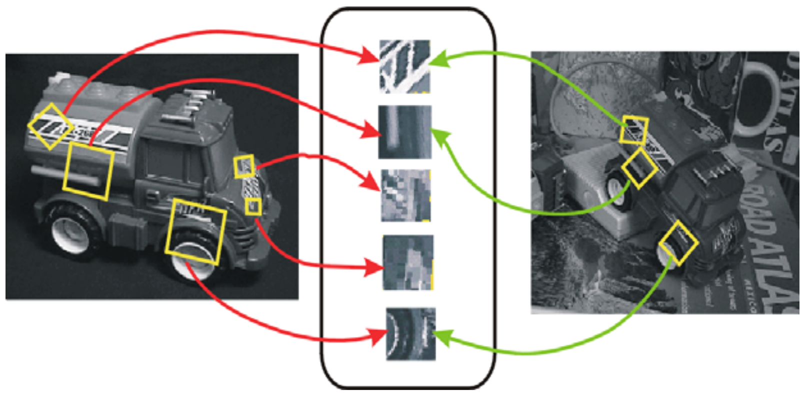
- Invariance:
  - Rotation, scaling, cropping, shift, etc.
  - illumination, pose, clutter, occlusion, viewpoint





Invariant Local Features

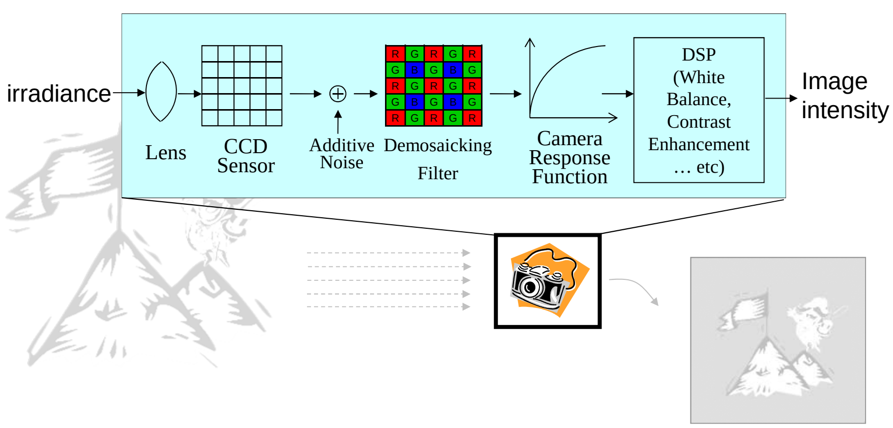
Image content is transformed into local feature coordinates that are invariant to translation, rotation, scale, and other imaging parameters



Features Descriptors

(Slide of A. Efros)

(review) Imaging Formation

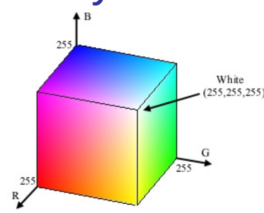


6

Color Spaces and Color Order Systems

- Color Spaces
  - RGB – cube in Euclidean space

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B}$$

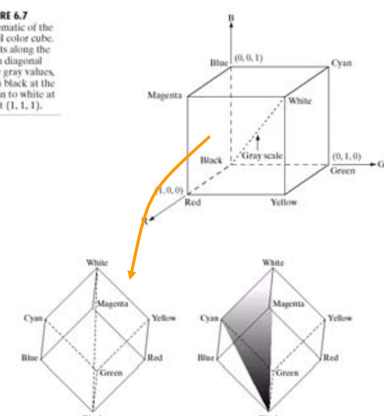


  - Standard representation used in color displays
  - Drawbacks
    - RGB basis not related to human color judgments
    - Intensity should be one of the dimensions of color
    - Important perceptual components of color are hue, saturation, and brightness
  - Perceptual color spaces: HIS, HSV

Understanding HSI from RGB

- Turn the RGB cube so that Black-White axis is vertical
- Each plane containing the B-W axis and a color point contains all the colors of the same hue
- <u>Hue</u> represented as angle between the plane and a reference plane (e.g. Red)
- <u>Saturation</u>: distance to axis, less saturated by mixing more grey colors
- <u>Intensity</u> measured by intersection with the B-W axis.
  - Cross section shape: triangle – hexagon - triangle

FIGURE 6.7 Schematic of the RGB color cube. Points along the main diagonal have gray values, from black at the origin to white at point (1, 1, 1).



a b
FIGURE 6.12 Conceptual relationships between the RGB and HSI color models.

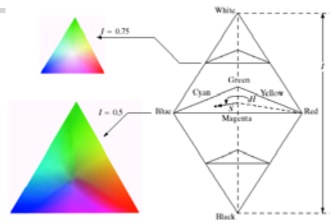
Images from Gonzalez and Woods

8

Colors on the HSI color cone

- Cross section approximated by triangle or circle
- HSI values computed by various geometrical models, e.g.,

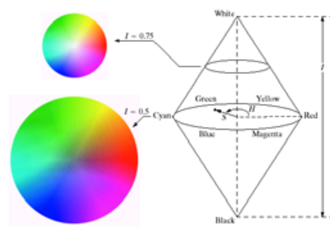
$$\begin{bmatrix} I \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{6} & -1/\sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad H = \tan^{-1}\left(\frac{V_2}{V_1}\right) \qquad Chroma = (V_1{}^2 + V_2{}^2)^{1/2}$$



- More suitable for measuring perceptual distance

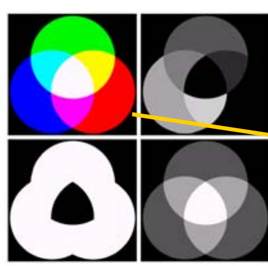
$$D(C_1, C_2) = \omega_1 D_{hue} + \omega_1 D_{sat} + \omega_1 D_{int}$$

- Can be quantized unevenly, e.g., Columbia VisualSEEK System: 16M colors (in RGB) quantized to 166 HSV colors (18 Hue, 3 Sat, 3 Val, 4 Gray)



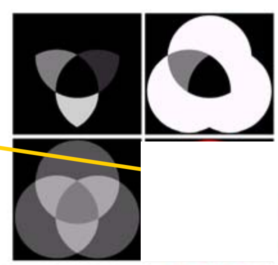
9

Manipulations in the HSI space



a b
c d

FIGURE 6.16 (a) RGB image and the components of its corresponding HSI image: (b) hue, (c) saturation, and (d) intensity.



a b

- HSI values of primary/secondary colors
- HSI allows independent manipulations of colors

- Hue of Green & Blue set to 0.
- Saturation of Cyan reduced by half.
- Intensity of White reduced by half.

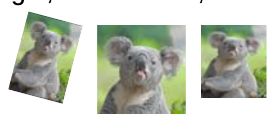
10

Color Histogram

- Feature extraction from color images
  - Choose GOOD color space
  - Quantize color space to reduce number of colors

$$h_{RGB}[r,g,b] = \sum_m \sum_n \begin{cases} 1 & \text{if } I_R[m,n]=r, I_G[m,n]=g, I_B[m,n]=b \\ 0 & \text{otherwise} \end{cases}$$

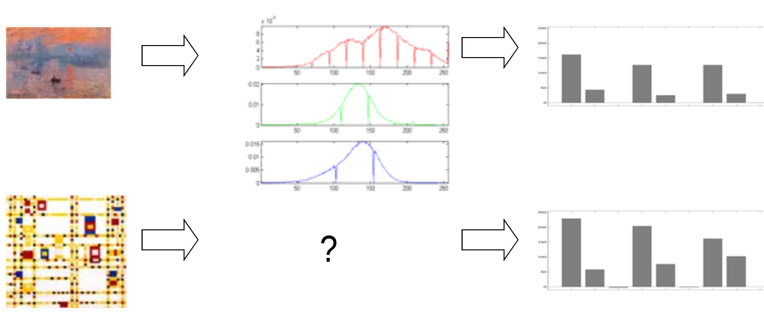
- Invariance?
  - Scale, shift, rotation, crop, view angle, illumination, clutter, occlusion



- Advantages
  - Easy to compute and compare
- Cons
  - Lack spatial information, dimension may be high

Color Moments

- Is there a more compact representation than color histogram?
- Compute moment statistics in each color channel.



Histograms of Partitioned Image

Divide image up into rectangles.
Compute separate histogram for each partition.

Rectangles can overlap.

http://www.ai.mit.edu/courses/6.801/Fall2002/

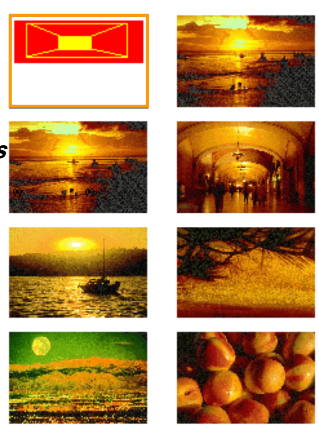
Color Layout Search

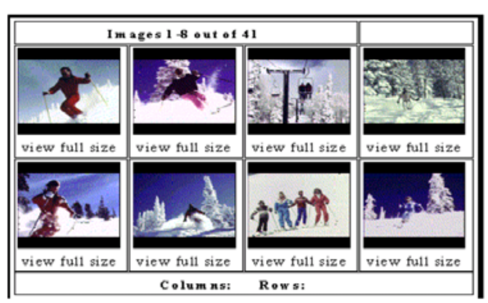
Columbia VisualSEEk (Smith & Chang, '96)

IBM QBIC (Flickner et al '95)

Query

results





Indexing with Color Correlograms

[Zabih, et al.]

Problem: Pictures taken from slightly different view positions can look substantially different with a color histogram similarity measure.

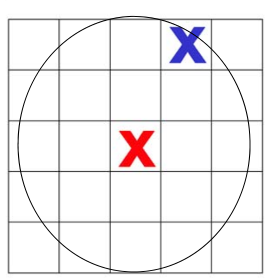
Proposed solution: Compute color co-occurance statistics [Haralick 1979].

http://www.ai.mit.edu/courses/6.801/Fall2002/

Color Correlogram

[Zabih, et al.]

For each image, estimate the probability that a pixel of some color lies within a particular distance of pixel of another color.



http://www.ai.mit.edu/courses/6.801/Fall2002/

Estimating Color Correlogram

Consider set of distances of interest $[d]=\{1,2,\ldots,d\}$
Measure pixel distance with L_∞ norm.
Consider m possible colors $c_i \in \{c_1, c_2, \ldots, c_m\}$.

```
Offline, for each image:
    Construct a correlogram that has m×m×d bins, initialize=0.
    For each pixel p_i in the image, find it's color c_i
        for each distance k∈{1,2,…,d}
            for each pixel at distance k from p_i
                increment bin (i,j,k)
            end
        end
    end
    Normalize correlogram by a scale factor (see Zabih, et al.)
```

http://www.ai.mit.edu/courses/6.801/Fall2002/

Color Coherence Vector (CCV) (Pass *et al*, 1997)

$$\begin{matrix} 2 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 2 & 1 & 1 \\ 2 & 1 & 3 & 2 & 1 & 1 \\ 2 & 2 & 2 & 1 & 1 & 3 \\ 2 & 2 & 1 & 1 & 3 & 3 \\ 2 & 2 & 1 & 1 & 3 & 3 \end{matrix} \xrightarrow{\text{Region segmentation}} \begin{matrix} B & C & B & B & A & A \\ B & B & C & B & A & A \\ B & C & D & B & A & A \\ B & B & B & A & A & E \\ B & B & A & A & E & E \\ B & B & A & A & E & E \end{matrix}$$

Not just count of colors, also check adjacency

regions	*A*	*B*	*C*	*D*	*E*
color	1	2	1	3	3
size	12	15	3	1	5

Coherent!
Size threshold: 3

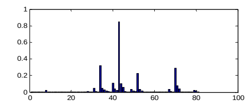
CCV	*color*	1	2	3
	α	12	15	5
	β	3	0	1

$$G_I = \langle (\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n) \rangle \quad G'_I = \langle (\alpha'_1, \beta'_1), \ldots, (\alpha'_n, \beta'_n) \rangle$$

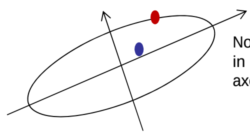
$$\Delta_C = \sum_{i=1}^{n} |\alpha_i - \alpha'_i| + |\beta_i - \beta'_i| \quad \Delta_H = \sum_{i=1}^{n} |(\alpha_i - \alpha'_i) + (\beta_i - \beta'_i)|$$

$\Delta_G > \Delta_H$ *by triangular inequality*

Distance Metrics between Feature Vectors



- Lp distance $$D_p = (\sum_i |x_1(i) - x_2(i)|^p)^{1/p}$$
- Quadratic distance $$D_q = (\sum_j \sum_i |x_1(i) - x_2(i)| C(i, j) |x_1(j) - x_2(j)|)$$ $$= (\vec{x}_1 - \vec{x}_2)^T C (\vec{x}_1 - \vec{x}_2)$$ *C(i,j):* color distance
- Histogram Intersection $$D_I = 1 - \frac{\sum_i \min\{x_1(i), x_2(i)\}}{\min\{\sum_i x_1(i), \sum_i x_2(i)\}}$$
- Mohalanobis distance $$D^2_{mah} = (x_1 - x_2)^T C_x^{-1} (x_1 - x_2)$$



Normalize distance in the major/minor axes

where C_x is the covariance matrix

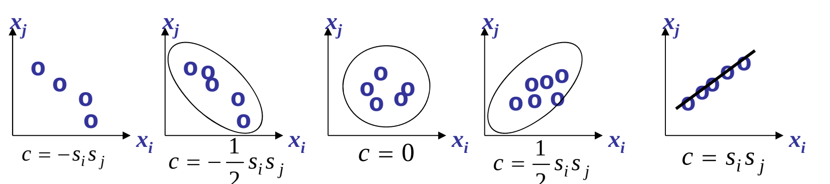
Mohalanobis Metric

$$D_{mah}{}^2 = (x_1 - x_2)^T C_x^{-1} (x_1 - x_2)$$

$$\text{covariance matrix } C_x = \begin{bmatrix} c(1,1) & c(1,2) & \dots & c(1,d) \\ \dots & \dots & \dots & \dots \\ c(d,1) & c(d,2) & \dots & c(d,d) \end{bmatrix}$$

d: dimension of features

$$c(i,j) = \sum_{k=1}^{N} [x_k(i) - m(i)]\,[x_k(j) - m(j)] \;/\; N-1, \qquad N: \text{number of samples}$$



s_i, s_j: std. deviation

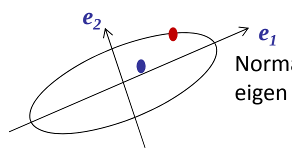
Mohalanobis Metric

$$D^2_{mah} = (x_1 - x_2)^T C_x^{-1} (x_1 - x_2)$$

where C_x is the covariance matrix

$$C_x = \left[\vec{e}_1 \,|\, \vec{e}_2 \ldots |\, \vec{e}_d\right] \; diag(\lambda_1, \lambda_2, \ldots, \lambda_d) \left[\vec{e}_1 \,|\, \vec{e}_2 \ldots |\, \vec{e}_d\right]^T$$

$$C_x^{-1} = \left[\vec{e}_1 \,|\, \vec{e}_2 \ldots |\, \vec{e}_d\right] \; (diag(\lambda_1, \lambda_2, \ldots, \lambda_d))^{-1} \left[\vec{e}_1 \,|\, \vec{e}_2 \ldots |\, \vec{e}_d\right]^T$$



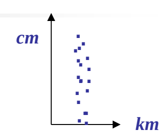
Normalize distance in the eigen vector axes

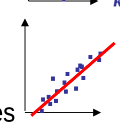
Project data to the eigen vectors, divide with the sd of each eigen dimension, and compute Euclidian distance

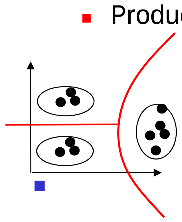
Mohalanobis Metric (cont.)

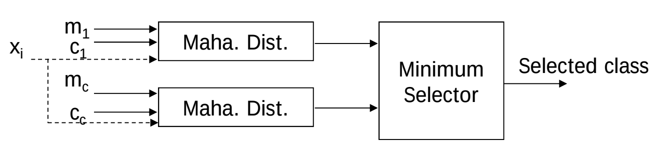
- Advantages of Mahalanobis metric
  - Account for scaling of coordinate axes
  - Invariant under linear transformation

    $If \; y = Ax \Rightarrow C_y = AC_xA^T \qquad , D_y^2 = D_x^2$
  - Correct for correlation
  - Produce curved as well as linear decision boundaries





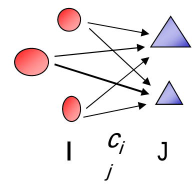




- Potential issue
  - Need enough training data to estimate Cov. Matrix

Earth Mover's Distance (EMD)

- Rubner, Tomasi, Guibas '98
- Mallow's distance in statistics in 1950's
- Transportation Problem [Dantzig'51]
  - I: set of suppliers
  - J: set of consumers
  - c_{ij} : cost of shipping a unit of supply from i to j



- Problem: find the optimal flows f_{ij}

$$minimize \sum_{i \in I} \sum_{i \in I} c_{ij} f_{ij} \qquad s.t.$$

$$f_{ij} \geq 0, i \in I, j \in J \qquad (No\ reverse\ shipping)$$

$$\sum_{i \in I} f_{ij} = y_j, j \in J \qquad (satisfy\ each\ consumer\ need\ /cacacity)$$

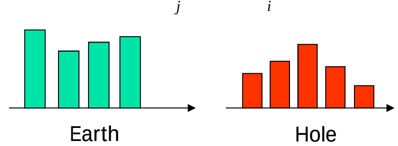
$$\sum_{j \in J} f_{ij} \leq x_i, i \in I \qquad (bounded\ by\ each\ supplier's\ limit)$$

$$\sum_{j \in J} y_j \leq \sum_{i \in I} x_i \qquad (feasibility)$$

EMD of Color Histogram

$$h = \left[h(1), h(2), ..., h(M)\right], g = \left[g(1), g(2), ..., h(N)\right], assume \sum_j g(j) \leq \sum_i h(i)$$

$$EMD(h,g) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} C_{ij} f_{ij}}{\sum_{i=1}^{M} \sum_{j=1}^{N} f_{ij}}$$



$$= \sum_{i=1}^{M} \sum_{j=1}^{N} C_{ij} f_{ij} / \sum_{j=1}^{N} g_j \quad Fill\ up\ each\ hole$$

C_{ij}: *distance between color i in color space h and color j in color space g*

f_{ij}: *move* f_{ij} *units of mass from color i in h to color j in g*

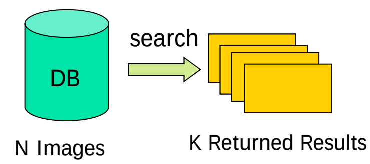
- Normalization by the denominator term
  - Avoid bias toward low mass distributions (i.e., small images)
  - what's the difference if both h and g are normalized first?

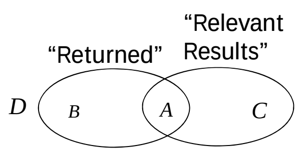
Evaluation

<u>Ground truth</u>

$V_n = 1$ "Relevant"

$\quad 0$ "Irrelevant" $\quad n = 0 \cdots N-1$





- Detection $\quad A = \sum_{n=0}^{K-1} V_n$
- False Alarms $\quad B = \sum_{n=0}^{K-1} (1 - V_n)$
- Misses $\quad C = (\sum_{n=0}^{N-1} V_n) - A$
- Correct Dismissals $\quad D = (\sum_{n=0}^{N-1} (1 - V_n)) - B$

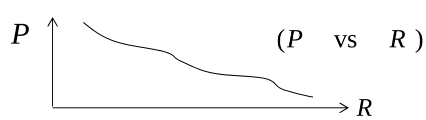
- Recall $\quad R = A/(A+C)$
- Precision $\quad P = A/(A+B)$
- Fallout $\quad F = B/(B+D)$
- Combined $\quad F_1 = \dfrac{P \cdot R}{(P+R)/2}$

Evaluation Measures

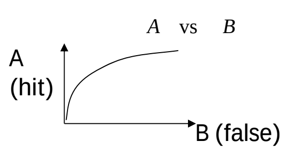
Precision at depth K

$$\boxed{P_k = (\textstyle\sum_{n=0}^{k-1} V_n) / K}$$

Precision Recall Curve



Receiver Operating Characteristic (ROC Curve)



Evaluation Metric: Average Precision

AP approximates areas under PR curve

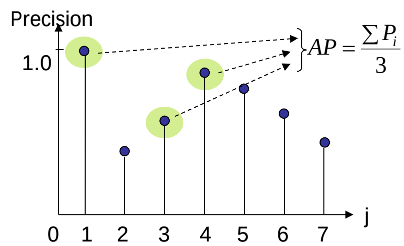
$$AP = \frac{1}{\min(K, R)} \sum_{j=1}^{K} [P_j \cdot I(D_j \text{ is correct})]$$

R : *total* # of relevant data, I : *indicator function*

Example:

Ranked list of data in response to a query

	D_{15}	D_8	D_{63}	D_{21}	...	...	D_s
Ground truth	*1*	*0*	*1*	*1*	*0*	*0*	*0*
Precision	*1/1*	*1/2*	*2/3*	*3/4*	*3/5*	*3/6*	*3/7*



Evaluation Metric: Average Precision

- Observations (AP)
  - AP depends on the rankings of relevant data and the size of the relevant data set. E.g., R=10

Case I: + + + + + + + + + + - - - - - - -

Pre: 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 AP=1

Case II: - + - + - + - + - + - + - + - + - + - +

Pre: 1/2 1/2 1/2 1/2 1/2 1/2 1/2 1/2 1/2 1/2 AP=1/2

Case II: - - - - - - - - - - + + + + + + + + +

Pre: 1/11 2/12 ... ... 10/20 AP~0.3

Homework #1

- Given a small image database and a few queries
- Implement codes to extract color histogram
- Implement codes to measure L_2 image similarity
- Use image object labels to measure precision/recall curves
- Bonus:
  - Add new color features or similarity metrics to improve performance
  - Design GUI for result browsing

Reading List

- Rui, Y., T.S. Huang, and S.-F. Chang, *Image retrieval: current techniques, promising directions and open issues. Journal of Visual Communication and Image Representation, 1999. 10(4): p. 39-62.*
- Smith, J.R. and S.-F. Chang. *VisualSEEk: a Fully Automated Content-Based Image Query System. in ACM International Conference on Multimedia. 1996. Boston, MA.*
- David G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision, 60(2), 2004, pp91-110.
- Randen, T. and J. Husoy, *Filtering for texture classification: A comparative study. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2002. 21(4): p. 291-310.*
- Mikolajczyk, K. and C. Schmid, *A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005: p. 1615-1630.*
- Brown, M., R. Szeliski, and S. Winder. *Multi-image matching using multi-scale oriented patches. in IEEE CVPR. 2005.*