CBIR
Content-based Image Retrieval

J. M.
Department of Computer Science
Aristotle University

Goals of this Tutorial

- Introduce media retrieval

  - motivation
  - image dataset
  - feature extraction
  - image similarity
  - query by sample
  - classification

Motivation

- Motivation for development of efficient media retrieval systems

Motivation

- Motivation for development of efficient media retrieval systems

  - Explosion in the volume of media data over the Internet and wireless networks

Motivation

- Motivation for development of efficient media retrieval systems
  - Explosion in the volume of media data over the Internet and wireless networks
  - Increasing popularity of imaging devices such as digital camera and increasing proliferation of image data over communications networks

Motivation

- Motivation for development of efficient media retrieval systems
    - Explosion in the volume of media data over the Internet and wireless networks
    - Increasing popularity of imaging devices such as digital camera and increasing proliferation of image data over communications networks
    - Emergence of new consumerism where media technologies meet consumers' needs
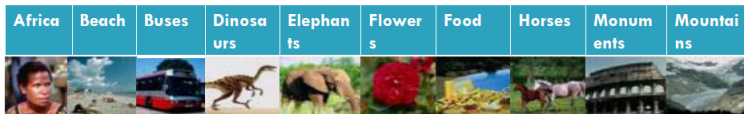
Content-based Image Retrieval (CBIR)

Schematic Overview of the System
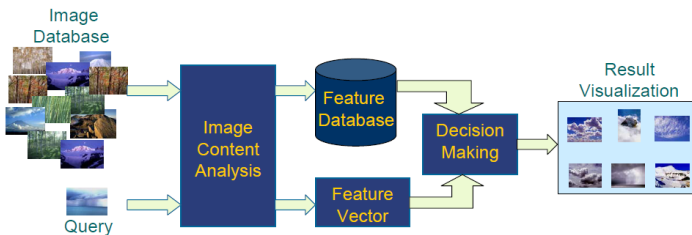
Image Dataset

The image dataset includes

- 1000 images from 10 different categories where 100 images correspond to each category

Feature Extraction

- Concatenate the features to form 190-dimensional feature vectors

| Features	Description	Dimensions
Color histogram	HSV space is chosen, each H, S, V component is uniformly quantized into 8, 2 and 2 bins respectively	32
Color auto-correlogram	The image is quantized into 4x4x4 = 64 colors in the RGB space	64
Color moments	The first two moments (mean and standard deviation) from the R, G, B color channels are extracted	6
Gabor wavelet	Gabor wavelet filters spanning four scales: 0.05, 0.1, 0.2, 0.4 and six orientations: $\theta_0 = 0$, $\theta_{n+1} = \theta_n + \frac{6}{\pi}$ are applied to the image. The mean and standard deviation of the Gabor wavelet coefficients are used to form the feature vector	48
Wavelet moments	Applying the wavelet transform to the image with a 3-level decomposition, the mean and the standard deviation of the transform coefficients are used to form the feature vector	40

Image Similarity

- There are many ways to define similarity. Similarity regarding *color distribution*, *shapes*, *textures* etc.
- Since the dataset is constructed from a combination of these features we need to define some similarity metrics to take advantage of it, such as:

Relative ℓ_1	$\sum\limits_{i=1}^{N} \frac{|X_i - Y_i|}{(1 + X_i + Y_i)}$
ℓ_2	$\sqrt{\sum\limits_{i=1}^{N} (X_i - Y_i)^2}$
$\vdots$	$\vdots$
etc.	etc.

Query by Sample

- Using the previously defined similarity metrics a user can submit queries and the system responds retrieving results similar to the query image using the *knn* method
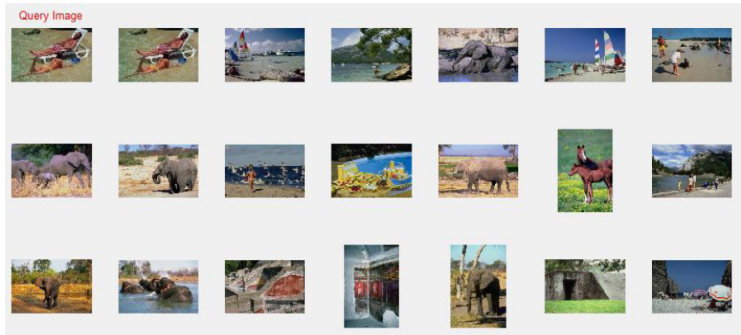
Classification

- In order to increase accuracy and retrieve better results compared to *knn* methods, classification techniques such as *svm* have been utilized

- For that reason $\frac{n!}{(n-k)!\,k!}$ binary classifiers have been trained for that particular task

- The final matrix has dimensions 1000x192 where the 191st column corresponds to the name of the image file and the 192nd column correponds to the true class label

- During the training 'hold out' cross validation is performed to increase accuracy where 50% of the samples is used for training and the other 50% for testing

Classification

Efficiency

To measure the efficiency of the classification the
contingency table is constructed as follows:

- Contingency matrix M is defined as

- $M(x, y) = \begin{cases} \forall\, x = y \Rightarrow \textit{true positive} \\ \forall\, x \neq y \Rightarrow \textit{false positive} \end{cases}$

THANK YOU! ☺