Content Based Image Retrieval

R. Venkatesh Babu
Outline

• What is CBIR
• Approaches
• Features for content-based image retrieval
  – Global
  – Local
  – Hybrid
• Similarity measure
Traditional Image Retrieval

• Traditional text-based image search engines
  – Manual annotation of images
  – Use text-based retrieval methods

• E.g.

Water lilies

Flowers in a pond

<Its biological name>
Limitations of text-based approach

• Problem of image annotation
  – Large volumes of databases
  – Valid only for one language – with image retrieval this limitation should not exist

• Problem of human perception
  – Subjectivity of human perception
  – Too much responsibility on the end-user

• Problem of deeper (abstract) needs
  – Queries that cannot be described at all, but tap into the visual features of images.
What is CBIR?

• Images have rich content that can be extracted as various content features:
  – Mean color, Color Histogram etc…
• Take the responsibility of forming the query away from the user.
• Each image will now be described by its own features.
CBIR

• How are the images represented?
  – Features
    • Classification based approach
      – Features are numerical values computed from each image
    • Text retrieval based approach
      – Features are image properties that are present or absent
CBIR – Sample Query

• User wants to search for, say, many rose images
  – He submits an existing rose picture as query.
  – He submits his own sketch of rose as query.
• The system will extract image features for this query.
• It will compare these features with that of other images in a database.
• Relevant results will be displayed to the user.
Sample Query
Sample CBIR architecture
IR Example: Query by Text
IR Example: Query by Example
Why Image Retrieval?

• More and more multimedia information available
• Difficult to access
• Specific Image Retrieval, such as medical image retrieval for diagnosis
Pixels to Picture
Pixels to Picture
How do we perceive?

Courtesy: Arun, CNS, IISc
How do we perceive?

Courtesy: Arun, CNS, IISc
Eye Fixation Map

Courtesy: Arun, CNS, IISc
Image Representation

- Low Level Feature
  - Color
  - Texture
  - HoG
  - LBP
  - Haar
  - Shape
- Interest Point based Feature
  - SIFT (image)
  - STIP (Video)
  - SURF
- Segmentation and shape
- ...

- How to represent image should be dependent on the application
Image Representation: Global to Local

• Global Descriptors
  – Colour histograms
  – Texture Features
  – Shape Features

• Local Descriptors
  – Direct approach
  – Patch-histograms / bag-of-visual words
  – SIFT features
Color histograms

- Statistics of pixel distribution
- R,G,B space;
- H,S,V space
- Like term frequency In text
Colour Histograms: Example

RGB colour space

HSV colour space
Texture Features

- “Texture refers to the properties held and sensations caused by the external surface of objects received through the sense of touch”
Tamura Texture Features
[1978]

- Coarseness – coarse vs. fine
- Contrast – high vs. low
- Directionality – directional vs. non-directional
- Line-likeness – line-like vs. non-line-like
- Regularity – regular vs. irregular
- Roughness – rough vs. smooth
Texture Feature : GLCM

• Gray Level Co-occurrence Matrix
  – Consider a pair of two pixels with fixed relative position
  – Like bigram in text

\[ C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^{n} \sum_{q=1}^{m} \begin{cases} 
1, & \text{if } I(p, q) = i \text{ and } I(p + \Delta x, q + \Delta y) = j \\
0, & \text{otherwise}
\end{cases} \]
Gabor Features

• Obtain several values per pixel de-noising spatial frequencies and directions
Gabor Filters: Combining Gaussian and (co)sinusoids

Cosinusoid × Gaussian

Sinusoid × Gaussian

Mask value

Pixel

Mask value

Pixel

Courtesy: Rishi, MBU, IISc
Tiling the 2D frequency plane with Gabor filters


Courtesy : Rishi, MBU, IISc
Application to image processing

- Cosine part $\rightarrow$ Directional blur operator
- Sine part $\rightarrow$ Directional Edge detector
Pixel Values as Features

- Most straightforward
- Scale all images to a common size
- Compare pixels using e.g. Euclidean distance

- Multi-scale Representations:
Shape: GIST descriptor

• Describe the shapes occurring in an image with one descriptor
  – Subdivide image in 4×4 sub images
  – Calculate Gabor responses in each of these
  – Create histograms of Gabor responses in each sub image

• Has been shown to be helpful to distinguish
  – Naturalness
  – Openness
  – Roughness
  – Expansion
  – Ruggedness
Shape: GIST descriptor

GIST descriptor
Oliva and Torralba, IJCV 2001

Slide by James Hays and Alexei Efros
Local Descriptors

• Features extracted from local regions from the image
• E.g. patches, SIFT features, local color histograms, ...
• Extraction position determined by interest points
• Known to achieve good results in many tasks

• Active field of research in object recognition, detection, scene classification, image annotation, and image retrieval
Local Descriptors: Interest Points
Local Descriptors: SIFT

- Store a histogram of gradients in local areas
- SIFT = Scale Invariant Feature Transform

- Leading to 128-dimensional feature vectors
- Have been shown to perform well in many tasks
Combining Features

- Manually tuned
  - Have an `expert' find a proper set of parameters
- Heuristic to capture different image properties
- Combination to reflect human perception
- Combination to obtain optimal performance (given a set of training queries)
Histogram of Local Descriptors
Other Popular Features: HoG
• LBP–Local Binary Pattern
  – Used for face recognition, tracking, etc.

• Haar
  – Most Important feature for face detection on your PC
CBIR Summary

Types of signatures
- Feature vector
- Region-based signature
- Summary of local feature vectors

Mathematical formulation
- Single vector
- Sets of vectors
- Codebooks
- Prob. density

Distances
- Euclidean
- Geodesic
- Weighted sum of vector distances
- Select regions
- Text-based
- Kullback-Leibler

Techniques
- Manifold embedding
- Optimal matching
- Linear programming
- User input, optimization
- Vector quantization
- Text retrieval
- Density estimation
CBIR Systems
IBM’s QBIC

- QBIC – Query by Image Content
- First commercial CBIR system (1995).
- Model system – influenced many others.
- Uses color, texture, shape features
- Text-based search can also be combined.
- Uses R*-trees for indexing
QBIC – Search by color

** Images courtesy : Yong Rao
QBIC – Search by shape

Features:
- Area
- Circularity
- Eccentricity
- Major Axis Direction
- Moments ...

** Images courtesy : Yong Rao
QBIC – Query by sketch

Matching done by template matching

** Images courtesy : Yong Rao
Virage

• Developed by Virage inc.
• Like QBIC, supports queries based on color, layout, texture
• Supports arbitrary combinations of these features with weights attached to each
• This gives users more control over the search process
VisualSEEk

• Research prototype – University of Columbia
• Mainly different because it considers spatial relationships between objects.
• Global features like mean color, color histogram can give many false positives
• Matching spatial relationships between objects and visual features together result in a powerful search.