An Unsupervised Learning Approach to Content-Based Image Retrieval

Yixin Chen & James Z. Wang
The Pennsylvania State University
Outline

- Introduction
- Cluster-based retrieval of images
- Experiments
- Conclusions and future work
Image Retrieval

- The driving forces
  - Internet
  - Storage devices
  - Computing power

- Two approaches
  - Text-based approach
  - Content-based approach
Text-Based Approach

- **Input keywords descriptions**

![Diagram showing a Text-Based Image Retrieval System with images of elephants as input and output.]
Text-Based Approach

- Index images using keywords (Google, Lycos, etc.)
  - Easy to implement
  - Fast retrieval
  - Web image search (surrounding text)
  - Manual annotation is not always available
  - A picture is worth a thousand words
  - Surrounding text may not describe the image
Content-Based Approach

- Index images using low-level features

Content-based image retrieval (CBIR): search pictures as pictures
A Data-Flow Diagram

Feature Extraction -> Image Database

- Histogram, color layout, sub-images, regions, etc.
- Euclidean distance, intersection, shape comparison, region matching, etc.

Compute Similarity Measure

- Linear ordering, Projection to 2-D, etc.

Visualization
Open Problem

- Nature of digital images: arrays of numbers
- Descriptions of images: high-level concepts
  - Sunset, mountain, dogs, ……
- Semantic gap
  - Discrepancy between low-level features and high-level concepts
  - High feature similarity may not always correspond to semantic similarity
Outline

- Introduction
- Cluster-based retrieval of images
- Experiments
- Conclusions and future work
Motivation

A query image and its top 29 matches returned by a CBIR system

Horses (11 out of 29), flowers (7 out of 29), golf player (4 out of 29)
Hypothesis

In the “vicinity” of a query image, images tend to be semantically clustered

CLUE attempts to capture high-level semantic concepts by learning the way that images of the same semantics are similar
A general diagram of a CBIR system using CLUE

- Feature Extraction
- Image Database
- Compute Similarity Measure
- Select Neighboring Images
- Image Clustering
- Display And Feedback

CLUE
Neighboring Images Selection

- Nearest neighbors method
  - Pick $k$ nearest neighbors of the query as seeds
  - Find $r$ nearest neighbors for each seed
  - Take all distinct images as neighboring images

Nearest Neighbors Method

$k=3$, $r=4$
Weighted Graph Representation

- Graph representation
  - Vertices denote image
  - Edges are formed between vertices
  - Nonnegative weight of an edge indicates the similarity between two vertices
Clustering

- Graph partitioning and cut
  \[ \text{cut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \]

- Normalized cut (Ncut) [Shi et al., IEEE Trans. PAMI 22(8)]
  \[ N\text{cut}(A, B) = \frac{\text{cut}(A, B)}{\text{assoc}(A, V)} + \frac{\text{cut}(A, B)}{\text{assoc}(B, V)} \]

- Recursive Ncut
Outline

- Introduction
- Cluster-based retrieval of images
- Experiments
- Conclusions and future work
An Experimental System

- Similarity measure
  - UFM [Chen et al. IEEE PAMI 24(9)]

- Database
  - COREL
  - 60,000
User Interface

(a) Thumbnails of image clusters.

(b) Images in Cluster 1.
Query Examples from 60,000-image COREL Database

Bird, car, food, historical buildings, and soccer game

CLUE

Bird, 6 out of 11

UFM

Bird, 3 out of 11
Query Examples

CLUE

<table>
<thead>
<tr>
<th>Car, 8 out of 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food, 8 out of 11</td>
</tr>
</tbody>
</table>

UFM

<table>
<thead>
<tr>
<th>Car, 4 out of 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food, 4 out of 11</td>
</tr>
</tbody>
</table>
Query Examples

CLUE

Historical buildings, 10 out of 11

Soccer game, 10 out of 11

UFM

Historical buildings, 8 out of 11

Soccer game, 4 out of 11
Clustering WWW Images

- Google Image Search
  - Keywords: tiger, Beijing
  - Top 200 returns
  - 4 largest clusters
  - Top 18 images within each cluster
Clustering WWW Images

Tiger Cluster 1 (75 images)

Tiger Cluster 2 (64 images)

Tiger Cluster 3 (32 images)

Tiger Cluster 4 (24 images)
Clustering WWW Images

Beijing Cluster 1 (61 images)

Beijing Cluster 2 (59 images)

Beijing Cluster 3 (43 images)

Beijing Cluster 4 (31 images)
Retrieve Accuracy
Outline

- Introduction
- Cluster-based retrieval of images
- Experiments
- Conclusions and future work
Conclusions

- Retrieving image clusters by unsupervised learning
- Tested using 60,000 images from COREL and images from WWW
Future Work

- Recursive Ncut
- Representative image
- Other graph theoretic clustering techniques
- Nonlinear dimensionality reduction
Thank You!