Intelligent Indexing and Retrieval of Images
A Machine Learning Approach

Yixin Chen
Dept. of Computer Science and Engineering
The Pennsylvania State University
http://www.cse.psu.edu/~yixchen

Outline

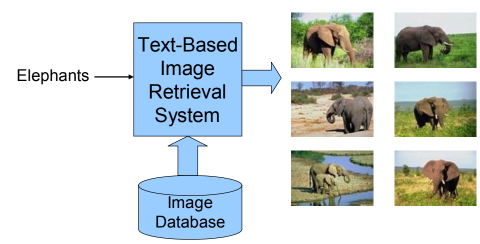
- Introduction
- Concept learning for image classification – learning sets of rules
- Cluster-based retrieval of images by unsupervised learning
- Conclusions and future work

Image Retrieval

- The driving forces
  - Internet
  - Storage devices
  - Computing power
- Two approaches
  - Text-based approach
  - Content-based approach

Text-Based Approach

- Input keywords descriptions

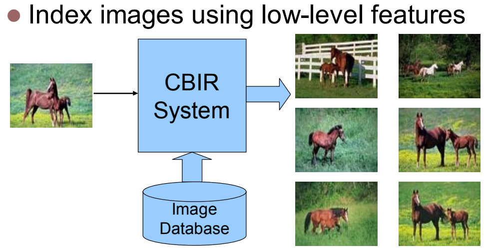


Text-Based Approach

- 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



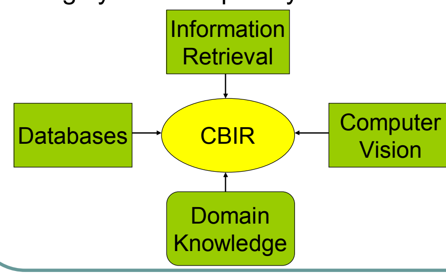
Content-based image retrieval (CBIR): search pictures as pictures

CBIR

- Applications
  - Commerce (fashion catalogue, ……)
  - Biomedicine (X-ray, CT, ……)
  - Crime prevention (security filtering, ……)
  - Cultural (art galleries, museums, ……)
  - Military (radar, aerial, ……)
  - Entertainment (personal album, ……)

CBIR

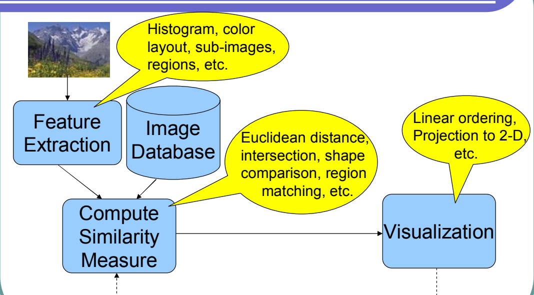
- A highly interdisciplinary research area



Previous Work on CBIR

- Starting from early 1990s
- General-purpose image search engines
  - IBM QBIC System and MIT Photobook System (two of the earliest systems)
  - VIRAGE System, Columbia VisualSEEK and WebSEEK Systems, UCSB NeTra System, UIUC MARS System, Standford SIMPLIcity System, NECI PicHunter System, Berkeley Blobworld System, etc.

A Data-Flow Diagram



Open Problem

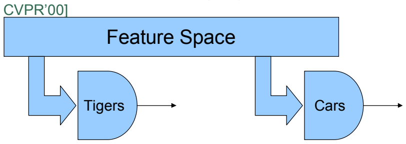
- Nature of digital images: arrays of numbers
- Descriptions of images: high-level concepts
  - Sunset, mountain, dogs, ……
- Semantic gap: discrepancy between low-level features and high-level concepts
  - High feature similarity may not always correspond to semantic similarity
  - Different users at different time may give different interpretations for the same image

Richness of User Semantics

- What is this image?
  - Sunset scene
  - River or lake
  - Camping site
  - People
  - Kayak
  - Picnic
  - “National park” by Google

Narrowing the Semantic Gap

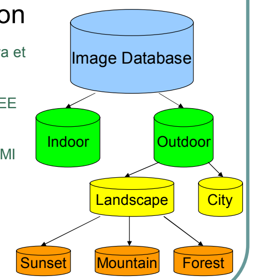
- Imagery features and similarity measure
  - Select effective imagery features [Tieu et al., IEEE CVPR'00]



  - Subjective experiments [Mojsilovic et al., IEEE Trans. IP 9(1)]

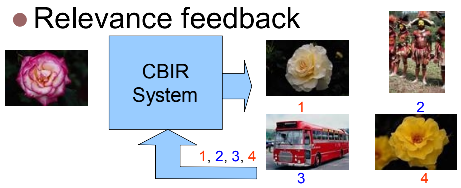
Narrowing the Semantic Gap

- Semantic Classification
  - Vacation images [Vailaya et al., IEEE Trans. IP 10(1)]
  - SIMPLIcity [Wang et al., IEEE Trans. PAMI 23(9)]
  - ALIP [Li et al., IEEE Trans. PAMI 2003]



Narrowing the Semantic Gap

- Relevance feedback



  - Adjusting similarity measure [Picard et al., IEEE ICIP'96], [Rui et al., IEEE CSVT 8(5)], [Cox et al., IEEE Trans. IP 9(1)]
  - Support vector machine [Tong et al. ACM MM'01]

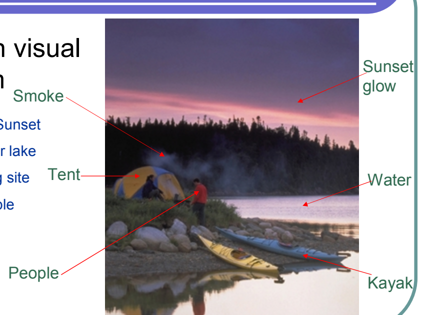
Outline

- Introduction
- Concept learning for image classification – learning sets of rules
- Cluster-based retrieval of images by unsupervised learning
- Conclusions and future work

Objects and Image Semantics

- Human visual system

Sunset glow ➔ Sunset
Water ➔ River or lake
Tent ➔ Camping site
Persons ➔ People
Kayak ➔ Kayak
Smoke ➔ Picnic



Objects and Image Semantics

- Image semantics may be related to objects in the image
- Semantically similar images may contain semantically similar objects
- Can a computer program learn semantic concepts about images based on objects?

Problem Formulation

- Binary classification

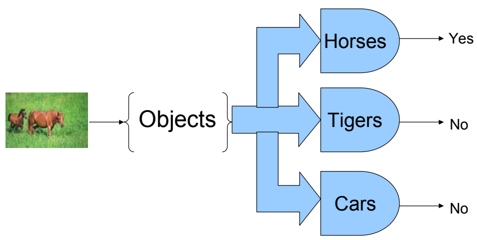
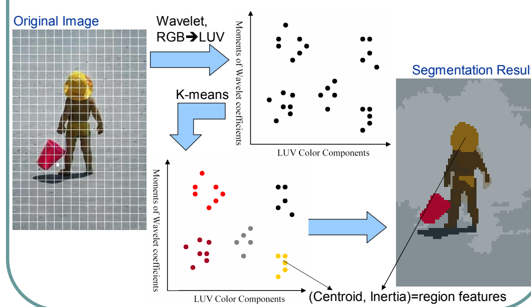


Image Segmentation



Learning

- Semantically similar images may contain semantically similar objects
  - Find similar objects (feature vectors) among “positive” images
  - At the same time they should be as distinct from all objects in “negative” images as possible
- Conceptual feature vector
  - Multiple-Instance Learning (MIL) using diverse density [Maron 1998], [Zhang et al., ICML’02]

Example

Example

- Three conceptual feature vectors
  - Water, rock, trees
- Rule description of a semantic concept

1: **IF** one of the regions is similar to water **AND** one of the regions is similar to rock **THEN** it is a waterfall image, **OR**,

2: **IF** one of the regions is similar to water **AND** one of the regions is similar to trees **THEN** it is a waterfall image

Learning

- Rule-based classification system and Support Vector Machine (SVM) [Chen et al., IEEE Trans. FS 2003]
- General learning scheme
  - Image segmentation
  - Learn a collection of conceptual feature vectors using MIL
  - Map images to a new feature space defined by conceptual feature vectors
  - Support vector learning

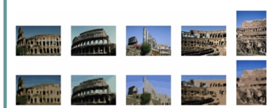
An Image Classification Example

- 10 image categories each containing 100 images
  - Africa, beach, buildings, buses, dinosaurs, elephants, flowers, horses, mountains, and food (labels are provided by COREL)


Africa

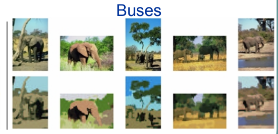

Beach

An Image Classification Example

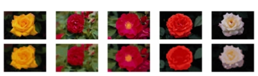

Buildings

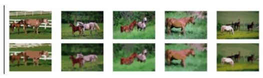

Buses


Dinosaurs

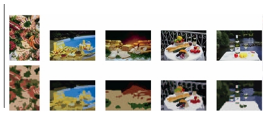

Elephants

An Image Classification Example


Flowers


Horses


Mountains


Food

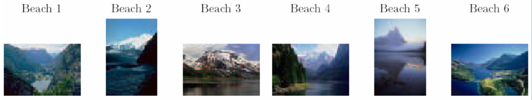
An Image Classification Example

- Confusion matrix

%	Africa	Beach	Buildings	Buses	Dinosaurs	Elephants	Flowers	Horses	Mountains	Food
Africa	72	2	4	2	0	12	2	0	0	6
Beach	0	54	8	6	0	2	2	0	26	2
Buildings	4	0	84	2	0	6	0	0	2	2
Buses	0	4	0	94	0	0	0	0	0	2
Dinosaurs	0	0	0	0	100	0	0	0	0	0
Elephants	14	0	6	0	0	74	0	4	0	2
Flowers	2	0	0	0	0	0	92	0	0	6
Horses	2	2	0	0	0	0	0	96	0	0
Mountains	0	8	12	4	0	4	0	0	72	0
Food	4	2	0	2	0	0	4	0	0	88

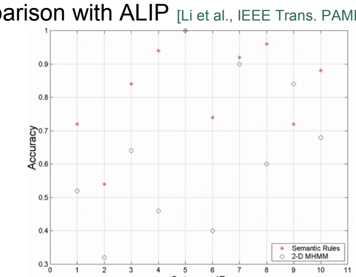
Beach vs. Mountains


Beach 1 Beach 2 Beach 3 Beach 4 Beach 5 Beach 6


Mountains 1 Mountains 2 Mountains 3 Mountains 4 Mountains 5 Mountains 6

An Image Classification Example

- Comparison with ALIP [Li et al., IEEE Trans. PAMI, 2003]



Outline

- Introduction
- Concept learning for image classification – learning sets of rules
- Cluster-based retrieval of images by unsupervised learning
- Conclusions and future work

CLUE: CLUsters-based rEtrieval of images by unsupervised learning

- Motivation

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

CLUE

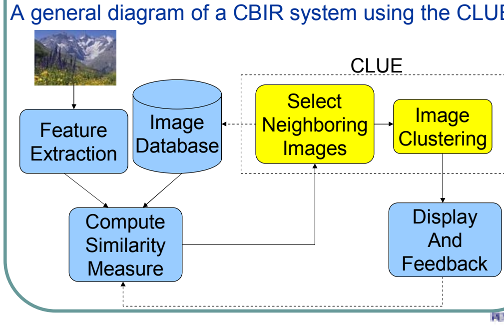
- Hypothesis

  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

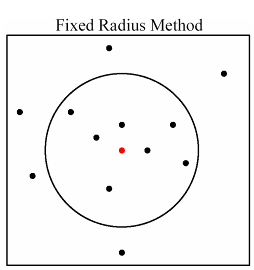
System Overview

A general diagram of a CBIR system using the CLUE



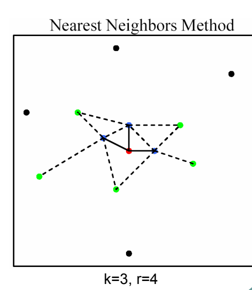
Neighboring Images Selection

- Fixed radius method
  - Take all target images within some fixed radius with respect to the query



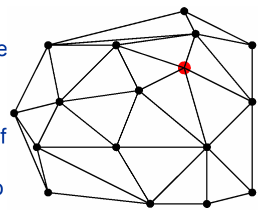
Neighboring Images Selection

- 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



Weighted Graph Representation

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



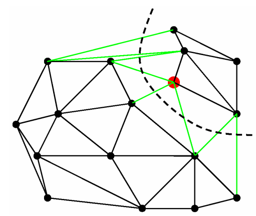
Clustering

- Graph partitioning and cut

$$cut(\mathbf{A},\mathbf{B}) = \sum_{i\in \mathbf{A}, j\in \mathbf{B}} w_{ij}$$

- Normalized cut (Ncut) [Shi et al., IEEE Trans. PAMI 22(8)]

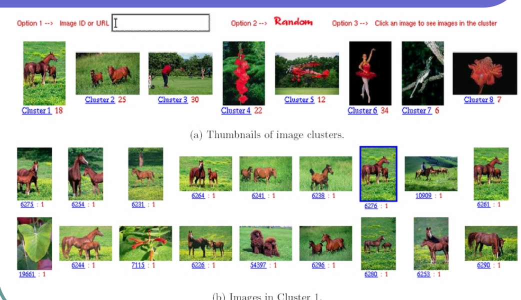
$$Ncut(\mathbf{A},\mathbf{B}) = \frac{cut(\mathbf{A},\mathbf{B})}{assoc(\mathbf{A},\mathbf{V})} + \frac{cut(\mathbf{A},\mathbf{B})}{assoc(\mathbf{B},\mathbf{V})}$$



An Experimental System

- Similarity measure
  - UFM [Chen et al. IEEE PAMI 24(9)]
- Database
  - COREL
  - 60,000
- Computer
  - Pentium III 700MHz PC, Linux

User Interface



Query Examples

- Query Examples from 60,000-image COREL Database

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

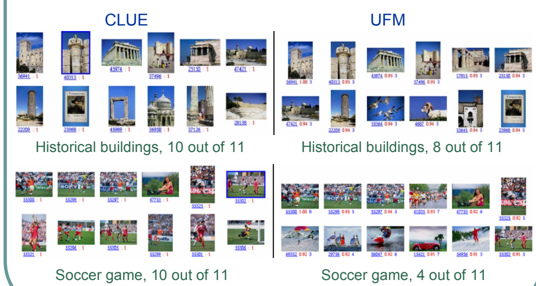
Bird, 6 out of 11 | Bird, 3 out of 11

Query Examples

Car, 8 out of 11 | Car, 4 out of 11

Food, 8 out of 11 | Food, 4 out of 11

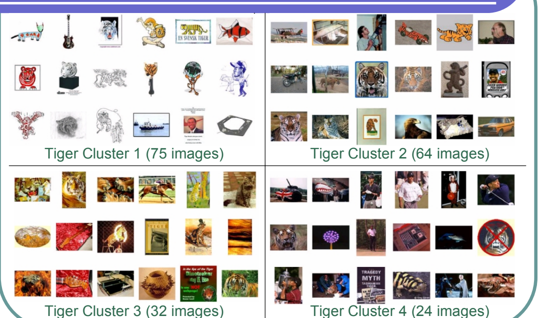
Query Examples



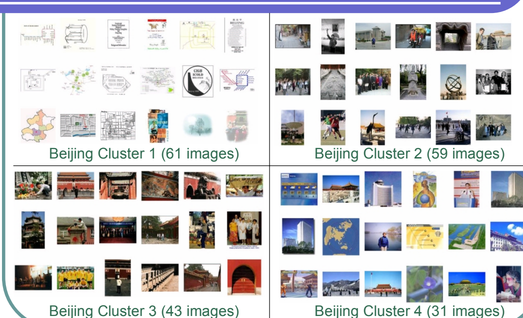
Clustering WWW Images

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

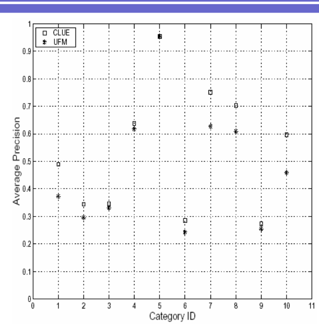
Clustering WWW Images



Clustering WWW Images



Retrieval Accuracy



Outline

- Introduction
- Concept learning for image classification – learning sets of rules
- Cluster-based retrieval of images by unsupervised learning
- Conclusions and future work

Conclusions

- Two approaches to tackle the "semantic gap" problem
  - Learning semantic concepts about images based on regions (objects)
  - Retrieving image clusters by unsupervised learning

  Tested using 60,000 images from COREL and images from WWW

Related Publications

- Journal Publications
  - [IEEE Trans. PAMI, 2002], computer vision & soft computing
  - [IEEE Trans. FS, 2003], machine learning theory & soft computing
- Refereed Conference Publications
  - [ACM MM'01], computer vision & soft computing
  - [FUZZ-IEEE'03], machine learning theory & soft computing
  - [FUZZ-IEEE'03, invited], computer vision & soft computing

Other Contributions

- Robotics and control (University of Wyoming)
  - Modeling and identification of payload (in collaboration with Prof. McInroy)
    - [IEEE Trans. AC, 2002], [ICRA'02], control
  - Decoupled control algorithms (in collaboration with Prof. McInroy)
    - [IEEE Trans. CST, 2003], [ICRA'00], control
  - Motion planning (in collaboration with Prof. McInroy)
    - [IEEE Trans. RA, 2003], [ACC'03], robotics
  - Fault-tolerant kinematics analysis (in collaboration with Prof. McInroy and Yong Yi)
    - [IEEE Trans. RA, 2003], [ISSSM'02], robotics

Future Work

- Continue to make contributions to the areas of
  - machine learning theories
  - robotics and automatic control
  - computer vision
  - soft computing

Future Work

- Apply theories to real world problems
  - Internet
  - Digital libraries
  - Robot vision systems
  - Sensor, imaging and video systems
  - Biomedicine
  - Autonomous intelligent agents
  - Homeland security

Potential Funding Sources

- NSF
  - Information Technology Research
  - SENSORS
  - Digital Libraries
  - Information and Data Management
  - Knowledge and Cognitive Systems
- NIH
- Office of Naval Research
- DARPA
- DOE
- Industrial funding

Support by

- The National Science Foundation
- The Pennsylvania State University
- The PNC Foundation
- SUN Microsystems
- NEC Research Institute

Acknowledgment

- Dissertation committee at Penn State
  - Professor James Z. Wang, IST and CSE
  - Professor Lee C. Giles, IST and CSE
  - Professor John Yen, IST and CSE
  - Professor Jia Li, STAT
  - Professor Donald Richards, STAT
- NEC Research Institute
  - Dr. Robert Krovetz

More Information

- Papers in PDF, 60,000-image DB, demonstrations, etc.

  http://wang.ist.psu.edu
  http://www.cse.psu.edu/~yixchen