Outline

- Background
- Objective of the class
- Prerequisite
- Lecture Plan
- Course Project
- Q&A

An image is worth a thousand words....

- What we observe are pixels....
- The story:
  - The train wreck at La Gare Montparnasse, 1895
- What computer can do these days:
  - Figure out the building
  - The train
  - People walking around
- Still long way to go to figure out the semantics
  - Train crashes
  - It is an abnormal event (context)

Advances in Image Sensors: pixels and voxels

- Hyperspectral Image Sensor
  - $I(x,y)$ in $\mathbb{R}^D$, $D=48$, e.g.
- 3D/Depth Sensor: LiDAR, Stereo Capture
  - $I(x,y,z)$ in $\mathbb{R}$
- Panoramic Video Cameras
  - $I(\alpha, \beta)$, $\alpha, \beta$ in $[0, 2\pi]$
- Lightfield Capture
  - Lenslet images
More than 20 years of Image Retrieval Research...

- IEEE Computer 1995 Special Issue on Content Based Image Retrieval (CBIR)

Content-Based Image Retrieval Systems

Z. Li: Image Analysis & Retrieval, Fall 2016. p.5

Dr. Raghavan, Vijay
Distinguished Professor
Center for Adv. Computer Studies
Univ of Louisiana at Lafayette

MPEG-7 Visual Features (Circa 2003)

- Color, Shape, Texture Features for Image Search

Color
- Texture
- Shape
- Motion

1. Histogram
   • Scalable Color
   • Color Structure
   • GOF/GOP
2. Dominant Color
3. Color Layout

- Texture Browsing
- Homogeneous texture
- Edge Histogram
- Contour Shape
- Region Shape
- Camera motion
- Motion Trajectory
- Parametric motion
- Motion Activity

ImageNet

- Tasks: Image Classification, Object Detection & Localization
  - 2012: Fisher Vector
  - 2013: Deep Learning ~ Conv Neural Networks (CNN)
  - 2016: (Very) Deep Learning ~ Residual Neural Networks (RNN)

Content-Based Image Retrieval at the End of the Early Years

Arnold W.M. Smeulders, Senior Member, IEEE, Marcel Worring, Simone Santini, Member, IEEE, Annamath Gupta, Member, IEEE, and Ramesh Jain, Fellow, IEEE

Abstract—The paper presents a review of 200 references in content-based image retrieval. The paper starts with discussing the working conditions of content-based retrieval: patterns of use, types of pictures, the role of semantics, and the sensory gap. Subsequent sections discuss computational steps for image retrieval systems. Step one of the review is image processing for retrieval sorted by color, texture, and local geometry. Features for retrieval are discussed next, sorted by: accumulative and global features, salient points, object and shape features, signs, and structural combinations thereof. Similarly, of pictures and objects in pictures is reviewed for each of the feature types, in close connection to the image and purpose of feedback: the user of the system is able to give by interaction. We briefly discuss aspects of system engineering: databases, system architecture, and evaluation. In the concluding section, we present our view on the driving forces of the field, the heritage from computer vision, the influence on computer vision, the role of similarity and of interaction, the need for databases, the problem of evaluation, and the role of the semantic gap.

Z. Li: Image Analysis & Retrieval, Fall 2016. p.6

Z. Li: Image Analysis & Retrieval, Fall 2016. p.8
**MPEG CDVS**

- **Compact Descriptor for Visual Search (CDVS)**
  - Object Re-Identification
  - Applications: Navigation, Query by Capture, AR/VR

- **Technology**
  - Key Point (SIFT) detection
  - Fisher Vector Aggregation and Hashing (for shortlisting)
  - SIFT compression

- **Performance**
  - Verification: 90+% precision on 1% recall
  - Retrieval: mAP in 80~90%

**Image Analysis Pipeline**

- **Holistic Image Analysis**
  - Direction Pixel Projection Subspace Models

\[
Y = AX
\]

- Convolutional Neural Networks

**About Myself**

- **Research Experiences**:
  - 10+ years experiences in leading industrial labs (Motorola, Samsung Research) and academia:
    - Associate Professor, Dept of CSEE, **UMKC**, 2015.08~
    - Sr. Staff Researcher/Sr. Manager, **Samsung Research America**, Dallas, 2012~15
    - Sr. Staff Researcher, **Huawei** Media Lab, 2010-12
    - Asst. Prof, Dept of Computing, **Hong Kong Polytechnic Univ.**, 2008-10,
    - Principal Staff Research Engineer, **Motorola** Labs, 2000-08.

- **Career Highlights**
  - Associate Editor, *IEEE Trans on Multimedia, IEEE Trans on Circuits & System for Video Tech (T-CSVT)*, 2 top journals in Multimedia area (google scholar)
  - TC Member: MMSP (SPS), MSA (CAS), MMTC (ComSoc)
  - 30+ issued or pending patents
  - 90+ publications in journal, conference proceedings and standards contributions.
  - Extensive project and leadership experiences.
Prerequisite & Textbook

- **Prerequisite**
  - For senior and graduate students in EE/CS
  - Good Matlab/C programming skills. Some Python is also desirable.
  - Taken Digital Signal Processing and/or Digital Image Processing, or consent of the instructor
  - Will have different expectation for MS and undergrad students

- **Textbook:**
  - None required (saving $), will distribute relevant chapters, papers, and notes.

- **Key References:**

- **Outline**
  - Background
  - Objective of the class
  - Prerequisite
  - Lecture Plan
  - Course Project
  - Q&A

- **Media Computing & Communication (MC²) Lab**
  - **Location:** FH 262
  - **Contact:** em: [lizhu@umkc.edu](mailto:lizhu@umkc.edu), ph: 816-235-2346

- **Projects**
  - Point Cloud Compression for Free Viewpoint Video
  - DNA Compression with Deep Learning
  - Next Gen Media Transport and Low Latency Video
  - Immersive Visual Communication
  - Device Based Deep Learning

- **Tentative Lecture Plan**
  - **Image Processing Basics**
    - Camera model and image formation
    - Image filtering
  - **Image Features for Retrieval**
    - Color Features
    - Texture and Shape Features
    - Basic Image Retrieval System and Metrics
  - **Object Identification in Image**
    - Key Point Detection
    - Key Point Feature Description
    - Fisher Vector Aggregation
    - MPEG Mobile Visual Search Technology and Standard
  - **Holistic Approach in Image Understanding**
    - Subspace methods for face recognition: Eigenface, Fisherface, Laplacianface.
    - Deep Learning in Image Understanding

- **Homework**
  - **Homework 1:** Image Filtering and Features
  - **Homework 2:** CalTech 101 Image Retrieval
  - **Homework 3:** Fisher Vector Aggregation
  - **Homework 4:** Laplacian Face Recognition
Potential Course/MS thesis Project

- Resources from last year:

- Potential projects with 25% bonus points
  - Boosting based key point feature detection via box filtering for Hyperspectral images
  - Compact deep learning models for embedded vision
  - Gesture recognition from point cloud capture
  - Image registration with point cloud
  - Rate-agnostic hash for video de-duplication in cache and networks.

Course Outcome

- Upon completion of the course you will be able to:
  - Understand the basic operations in image formation and filtering
  - Understand basic image features for retrieval: color, shape, texture
  - Understand key point features and aggregation in object identification
  - Understand the holistic appearance modeling approach in image understanding
  - Understand the latest image analysis and understanding techniques like deep learning.
  - Can apply the knowledge an algorithms to solve real world image understanding and retrieval problems
  - Well prepared for conducting advanced research and pursing career/PhD in this topic area.

Grading

- Homeworks (40%)
  - Image Filtering and Basic Features
  - Image Retrieval System and Performance Metrics
  - Key Point Feature and Fisher Vector Aggregation in Object Identification
  - Subspace Models in Image Understanding

- 2 Quizes (20%) : relax, quiz is actually on me, to see where you guys stand
  - Quiz-1: Part I and II
  - Quiz-2: Part III and IV

- Project (40%)
  - Original work leads to publication, discuss with me by the mid of October. (25% bonus point, worth 50 points)
  - Regular project: assign papers to read, implement certain aspect, and do a presentation.

Logistics

- Office Hour:
  - Mon, Wed: 2:30-4:00pm, 560E FH
  - Or by appointment

- TA:
  - TBD
  - Lab Sessions are planned to cover certain software tools aspects.
  - Office Hour: TBA

- Course Resources:
  - Will share a box.com folder with slides, references, data set, and software
  - Additional reference, software, and data set will be announced.