Advanced Multimedia

CSE/ISE 364
Tamara Berg
Instructor: Tamara Berg
tlberg@cs.sunysb.edu
Office: 1411 Computer Science
Lectures: T/R 3:50-5:10pm,
    Multimedia Lab - Rm 2205 CS
Office Hours: T/R 5:10-6:10pm and by appointment.

Course Webpage: http://tamaraberg.com/teaching/Spring_10/
About me

• New Professor last year
  – PhD from UC Berkeley 2007.
  – 2007-2008 Yahoo! Research

• Research in computer vision and natural language processing - combining information from multiple forms of digital media for applications like image search and recognition.
President George W. Bush makes a statement in the Rose Garden while Secretary of Defense Donald Rumsfeld looks on, July 23, 2003. Rumsfeld said the United States would release graphic photographs of the dead sons of Saddam Hussein to prove they were killed by American troops. Photo by Larry Downing/Reuters
Intuition

George Bush
Browse the news according to the individuals present

Face Dictionary

Actor: Duddy Eisen
Actor: Charlton Heston
Actor: Cheech Marin
Actor: Christopher Reeve
Actor: Christopher Walken
Actor: Chris Eastwood
Actor: Eliah Hawke
Actor: Harrison Ford
Actor: Jackie Chan
Actor: Jake Gyllenhaal

Actor: James Gandolfini
Actor: Jason Priestley
Actor: Johnny Depp
Actor: Kevin Spacey
Actor: Kirk Douglas
Actor: Leonard Nimoy
Actor: Liam Neeson
Actor: Matt Dillon
Actor: Mel Gibson
Actor: Michael Caine

Actor: Nick Nolte
Actor: Paul Reubens
Actor: Pierce Brosnan
Actor: Ray Romano
Actor: Richard gere
Actor: Rob Lowe
Actor: Robert Blake
Actor: Robert Duvall
Actor: Robert Moore
Actor: Russell Crowe

Actor: Salma Hayek
Actor: Steven Seagal
Actor: Tom Cruise
Actor: Tom Sizemore
Actor: William Shatner
Actor: Angela Bassett
Actor: Cameron Diaz
Actor: Diane Lane
Actor: Doris Roberts

Actor: Gwyneth Paltrow
Actor: Halle Berry
Actor: Hilary Duff
Actor: Holly Hunter
Actor: Jennifer Aniston
Actor: Jennifer Connolly
Actor: Jennifer Garner
Actor: Jennifer Lopez
Actor: Julianne Moore
Actor: Kate Winslet

Actress: Kirsten Dunst
Actress: Kristen Davis
Actress: Laura Elena Harring
Actress: Laura Linney
Actress: Meg Ryan
Actress: Meryl Streep
Actress: Naomi Watts
Actress: Nia Vardalos
Actress: Nicole Kidman
Actress: Queen Latifah

Actress: Jennifer Lopez was nominated for a Golden Raspberry or Razzie award as "the year’s worst actress" for "Enough" and "Maid in Manhattan" on February 10, 2003. Lopez is shown at the premiere of "Maid in Manhattan" on Dec. 8 and is joined by Madonna, Britney Spears, Winona Ryder and Angelina Jolie for the dubious honor. (Jeff Christensen/Reuters)
Animals on the Web

Goal: Classify images depicting semantic categories of animals in a wide range of appearances
Animals on the Web Outline:

Harvest pictures of animals from the web using Google Search.
Animals on the Web Outline:

Harvest pictures of animals from the web using Google Search.

Select some example images using text based information most image search engines only use text info!
Animals on the Web Outline:

Harvest pictures of animals from the web using Google Search.

Select some example images using text based information

Use visual + textual content analysis to extend to similar images.
Retrieval Comparison

Words

Present range of the American Bullfrog in British Columbia

Present range of the Toad Frog in British Columbia

Present range of the American Bullfrog in British Columbia

Words + Pictures

Fredro's Frog Gallery
Consortium for Digital Arts, Culture & Technology (cDACT)

http://emedia.art.sunysb.edu/cDACT/index.html
About you?
Pre-Requisites

• You do not need to have taken CSE/ISE 334.
• You are expected to be proficient in programming and the basics of digital media, but I will provide most necessary background for the course as we go.
• Come talk to me if you have any questions!
What is Multimedia?
What is Multimedia?

• From Wikipedia: “Multimedia includes a combination of text, audio, still images, animation, video, and interactivity content forms. Multimedia is usually recorded and played, displayed or accessed by information content processing devices, such as computerized and electronic devices, but can also be part of a live performance.”

• In this class we will mainly consider multimedia as accessed via the web. We will look at algorithms for storing, organizing, retrieving, and manipulating various forms of digital media.
Topics

• Text
• Sound
• Images/Video
• Other types of Digital Media including location information, social media, collaborative filtering etc.

We will discuss the fundamentals of each type of digital media plus a selection of special topics such as retrieval, modeling, and manipulation.
Text Basics – words and more, intro to NLP

Lots of text on the web. How can we access it effectively?

Web document retrieval (we’ll even implement part of the algorithm that first made Google a success!)
Text modeling and classification

Classify document by topic, or as spam/ham etc
Basic probability and machine learning
Useful features/attributes to measure

Example from Dan Klein
Clustering

- Basic idea: group together similar instances
- Example: 2D point patterns

Document clustering
Group results returned by web search in some meaningful way
Need:
  A representation for documents
  A way to measure similarity between documents
  A way to find the groups

Example from Dan Klein
Sound

- Sound signal basics, manipulation and analysis
  - Filtering
  - Beat detection
  - Retrieval
  - Digital Music
Images/Video

- Image formation & the camera
- Popular file formats & compression
- Consumer photo sharing
Images/Video

Color

Edges

Texture

Shape

Image Content analysis
Features for describing content – color, edges, texture, shape

Color & edge examples - Lana Lazebnik
Images/Video

Recognition – What’s in this image?

Search by text query – Find me images related to “horse”

Search by sketch or image query – retrievr – Find me images that look like “this”

Image recognition & retrieval
Image Manipulation

1. Extract Sprites (e.g. using *Intelligent Scissors* in Photoshop)

2. Blend them into the composite (in the right order)

Image blending and compositing
Image Manipulation

Image warping & morphing
Combined Media

Various types of media often appear together
Web pages with text & images/video
  image search can be implemented as text search
  of nearby words, or as text analysis + image content
  analysis
Movies/games – combination of images, video, sound
Education – can use a combination of media for
  effective teaching
Medicine – virtual surgery for training and telesurgery
Other kinds of Digital Media

Social Media
Tagging & Annotation
Location Information
Speech
Multimedia for Mobile Devices
Recommendation Systems
Other kinds of Digital Media

Flickr “Interestingness” – completely based on social data surrounding pic

Photo labeling as a game

Social Media

Sharing photos/videos – flickr, youtube, picasa, kodak, fotki, photobucket, ...
Social networking – facebook, myspace, friendster, linkedin, ...
Harnessing the power of people – wikipedia, esp game, collabio, mechanical turk
Other kinds of Digital Media

Tags – banana, monkey, primate, homeless, cardboard, hairy, brown, zoo, funny photo, animal

Photo by Nicholaus Haskins

Tagging & Annotation

Why tag? How do people tag? What can we do with the tags?
Other kinds of Digital Media

Flickr World Map

Google Street View

Location Information

Associate photos with a location – drag & drop, gps, cell phone tower info

Associate videos with locations – Street View
Other kinds of Digital Media

Text to speech

Speech to text

Speech

Automatic translation of speech to text, text to speech
Other kinds of Digital Media

Adaptive web page layout for mobile devices

Multimedia for Mobile Devices
Other kinds of Digital Media

**Personal**

*Personalized recommendation – recommend based on the individual's past behavior*

*Social recommendation – recommend based on the past behavior of similar users*

*Item recommendation – recommend based on the thing itself*

*A combination of the three approaches above*

**Item**

Netflix Prize - $1 million to improve their recommendation system by 10%

**Social**

*Recommender Systems – Predict what you will like*

*Neulix Prize*
Guest Lectures

Margaret Schedel – composer and cellist specializing in the creation and performance of ferociously interactive media.

Christa Erickson – interdisciplinary artist who weaves together combinations of video, tactile materials, programming, physical cinematic devices, and live data in interactive installations.

Multimedia is not just about computer science!
Workload

Assignments:

There will be about 4 assignments, covering the core types of digital media we will be studying (text, sound, images, social).

All assignments should be submitted by email to: cseise364@gmail.com
Assignments (subject to change)
Assignment 1

• Ever wonder how Google search works?
• Here’s your chance to find out.
• Implement the original Google PageRank algorithm
• Use it to re-rank pages from the Stony Brook cs website.
Assignment 2

• Compose a piece of music using matlab by creating and manipulating sound waves and audio files.

• Your piece should include musical notes, mixtures of musical notes, recorded sounds, imported samples of audio files.

• Flair and creativity are a must and the best pieces (as voted on by the class) will receive prizes tbd.
Assignment 3

• Fun with Frames
• In this assignment you will experiment with inserting transformed images into images and video.
Assignment 4

• TBD
• Perhaps something related to analyzing twitter data in various ways.
Workload

Projects:
Each of you will be responsible for defining and developing a project over the course of the semester. Projects may be done alone or in groups.

This will include a project proposal presentation, Project update presentation, and final project presentation. A project write-up will serve as your final exam.

I will provide some project ideas, but you are also welcome to come up with your own.
Workload

Readings:
Readings will be assigned from the textbook as well as external sources that I will provide.


Other useful reference books
Foundations of Statistical Natural Language Processing, Christopher D. Manning, and Hinrich Schutze.
Grading

This will be a very hands-on course and project focused. Therefore, there will be no formal exams.

Grading will be based on:
55% assignments, 35% project, 10% participation

For each assignment and project milestone you should turn in a simple web page showing your results and displaying any related media. How many of you know basic html?

For assignments each day late will subtract 10% from the assignment’s grade.
Programming

We will be implementing the assignments in Matlab - a software package with useful built-in functionality for working with digital media.

If you haven’t used Matlab before don’t worry. It might seem daunting, but we will do in-class tutorials on Feb 4 and I will show some useful techniques on Feb 9.

Matlab is installed in the first 3 rows of the Multimedia lab, and computing labs (2126, 2114, 2129).

I highly recommend getting it for your laptop or home computer. Can be purchased from Mathworks for home use for $99: http://www.mathworks.com/academia/student_version/
For Next Time

• Read “The Anatomy of a Large-Scale Hypertextual Web Search Engine,” Sergey Brin and Lawrence Page – linked from course web page