Advanced Multimedia

CSE/ISE 364
Tamara Berg
Course Info

Instructor: Tamara Berg  tlberg@cs.sunysb.edu
Office: 1411 Computer Science

Lectures: Tues/Thurs 3:50-5:10pm, Rm 2205 CS
Office Hours: Tues/Thurs 1:00-2:00pm and by appt

Course Webpage:
http://tamaraberg.com/teaching/Spring_12/cse364
About me

– PhD from UC Berkeley 2007.
– Assistant Professor at SBU since 2008.

• 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
Automatically Created Face Dictionary

Browse the news according to the individuals present

Face Dictionary

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)
Baby Talk: Understanding and Generating Simple Image Descriptions

Girish Kulkarni, Visruth Premraj, Sagnik Dhar, Siming Li, Yejin Choi, Alexander C Berg, Tamara L Berg

CVPR 2011
This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair, and near the potted plant.
“This picture shows one person,
“This picture shows one person, one grass,
“This picture shows one person, one grass, one chair,
“This picture shows one person, one grass, one chair, and one potted plant."
“This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass,
“This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair."
“This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair,
“This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair, and near the potted plant.”
“This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair, and near the potted plant.”
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 should be proficient in programming.
- 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 focus on multimedia as accessed via the web. We will look at algorithms for storing, organizing, retrieving, and manipulating various forms of digital media.
Topics will include:

- Text, Sound, Images, and Video
- Retrieval
- Morphing
- Tagging & Annotation
- Social Media
- Location Information
- Interaction with Xbox Kinect
- Recommendation systems

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 talk about 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: Spam Filter

- Input: email
- Output: spam/ham
- Setup:
  - Get a large collection of example emails, each labeled “spam” or “ham”
  - Note: someone has to hand label all this data!
  - Want to learn to predict labels of new, future emails
- Features: The attributes used to make the ham/spam decision
  - Words: FREE!
  - Text Patterns: $dd, CAPS
  - Non-text: SenderInContacts
  - ...

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 signal basics

- Filtering
- Manipulation and analysis
- Create your own composition through wave manipulation
Sound

• Sound Applications
  – Beat detection
  – Genre classification
  – Retrieval
Sound Guest Lecture

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

Multimedia is not just about computer science!
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
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 Manipulation

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

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

Women In Art

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
Interaction
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

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

Multi-modal Interaction
Vision, sound, and text interaction using Xbox Kinect
Other kinds of Digital Media

Recommendation Systems – Predict what you will like

* 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

New For You

Personal

Item

Social

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

Assignments:

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

All assignments should be submitted by email to:

cseise364@gmail.com
Workload

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

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

Projects can be completed in the programming language of your choice.
Workload

Projects Continued:

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

Past projects include:

“Epic Bunny Dash: Character Control via Motion Tracking”: interactive web-cam based tracking game
“What do images sound like”: translating between images and sound
“Text2Sound”: converting between text and sound
“Texture Synthesis”: synthesizing large textures from small texture patches
“Virtual Drum Kit”: interactive virtual drum set with Kinect
Workload

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

Book:
“Digital Multimedia” by Nigel Chapman and Jenny Chapman

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.

You will be allowed 5 free homework late days of your choice over the semester. After those are used late assignments will be accepted with a 10% reduction in value per day late.
We will be implementing the assignments in Matlab - a software package with useful built in functionality for working with digital media.

I expect most (all) of you have never used Matlab. Don’t worry! We will do in-class labs for each media type over the semester.

Matlab is installed in the Multimedia lab, and department 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/](http://www.mathworks.com/academia/student_version/)
Course Info

Instructor: Tamara Berg
tlberg@cs.sunysb.edu
Office: 1411 Computer Science
Lectures: Tues/Thurs 3:50-5:10pm,
  Multimedia Lab - Rm 2205 CS
Office Hours: Tues/Thurs 1:00-2:00pm and by appt.

Course Webpage:
http://tamaraberg.com/teaching/Spring_12/CSE364
For Next Time

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