CS 598 – Visual Scene Understanding – Spring 2009

Instructor: **Derek Hoiem**
Room: **Siebel Center 1109**
Times: **Tuesday and Thursday 11:00am – 12:15pm**
Office Hours: Tuesday and Thursday 12:15-1pm; by appointment
Contact: d<last name>@uiuc.edu, Siebel 3312
Class website: http://www.cs.uiuc.edu/homes/dhoiem/courses/cs598_spring09

Visual scene understanding is the ability to infer general principles and current situations from imagery in a way that helps achieve goals. Goals are highly varied, depending to a large extent on the agent. In biological systems, vision is usually just one piece of the puzzle, but is often a particularly important one, as it allows us to make detailed inferences about distant surfaces and objects.

In this course, we will explore several aspects of visual scene understanding through tutorials, papers, and projects. This class is a group effort. To get the most out of it, everyone should participate in class discussions and collaborate extensively in projects. The goal is to provide each student with

- A broader and deeper understanding of computer vision research
- An improved ability to discern important contributions of papers and how they fit in
- Improved communication skills: presenting, discussing, and writing
- Research skills and experience: researching, designing, and implementing a project

Assignments and Grading

Students are responsible for presenting papers, attending and participating in class discussions, and performing a group project. When not presenting, students must still read the assigned material and post relevant comments to the bulletin board. Grades will be assigned based on the following criteria:

- Class Participation (25%): timely posting of comments, class attendance, and discussion
- Paper Presentation (25%): thoughtful and well-organized presentations on a paper or topic
- Group Project (50%)
    - Proposal: goals, related work, and planned experiments (15%)
    - Interim report: approach and initial results (15%)
    - Final report: written (50%), oral (20%)

I will give informal feedback on paper presentation afterwards and on class participation at midterm. The proposal and interim report will be graded as satisfactory or unsatisfactory; resubmission is required if unsatisfactory. I will assign a grade to the final report (both written and oral). Each team member will also send me a confidential short team evaluation, saying how each member contributed to the development of the key ideas, the background search, the data collection, coding and experiments, writing, and preparation for the oral presentation.

You *will get an A* if you always post to the discussion boards on time, attend classes, participate in class discussions, deliver thoughtful and organized presentations, meet project deadlines, and make a strong effort in projects. You *will not get an A* if you miss several classes, often do not post to discussion boards or discuss in class, make half-hearted attempts at presentation, or do not contribute substantially to your project.

Policies

Discussion and interaction is an important part of the class. For this reason, **attendance is required**. If you cannot be present (due to unavoidable travel or sickness), let me know before the class (at least one week beforehand in the case of travel). Each *unexcused* absence beyond the first will result in a half-grade reduction.

Bulletin board postings and all other assignments **must be turned in on time**.

There are no formal **prerequisites**. However, a basic understanding of statistics, probability, linear algebra, and image processing will be needed to understand many of the papers. Familiarity with machine learning algorithms and computer vision techniques will be very helpful. The best way to see if you have sufficient background is to try to read one or two of the papers. Contact me if you are concerned.

Academic dishonesty will not be tolerated. For presentations it is permissible to use slides from online material, but the sources should be credited. For the final paper, all of the text and figures should be your own. Copying sentences or paragraphs from other papers is dishonest, unless you are quoting them. If you use a figure from an outside source, you must cite the source in the figure (and obtain permission if you want to publish it).

To obtain **disability-related academic adjustments** and/or auxiliary aids, students with disabilities must contact the course instructor and the Disability Resources and Educational Services (DRES) as soon as possible. To contact DRES you may visit 1207 S. Oak St., Champaign, call 333-4603 (V/TDD), or e-mail a message to disability@uiuc.edu.

Bulletin Board

The bulletin board is to be used primarily for posting comments about assigned readings. To sign up, go to http://www.nicenet.org, click “Join a Class”, and enter the class key: 6NX987W46

Guidelines for Presenting/Reading

Instructions to the Presenter

The assigned papers are a starting point for exploring a particular topic. Your presentation should present the key ideas of the assigned works and explain important technical aspects. It should also tell how the work fits into the broader dialogue. The presentation should be about 40 minutes long and cover the following

- Motivation: applications and challenges
- Summary of approach
- Key ideas and results
- Strengths and weaknesses
- Ideas for improvement
- Questions or issues raised on bulletin board
- Related work (at least three related papers: key ideas and how they relate)

Ideally, these should be covered as part of a single, smoothly flowing presentation on the topic, rather than creating separate presentations for each paper. It can be assumed that everyone has read the papers in advance.

The presenter must go over a full draft of the presentation with me at least two days before the presentation, if it is his/her first presentation for the class. Be sure to practice at least once.

Everyone will present twice. When signing up for papers, if there are no more open slots, you can team up with someone who has already selected a paper. In this case, you must do a demonstration as part of the presentation. For the **demo**, you should run the software of the proposed paper (this is easier if you can download it) on many examples and use the results to illustrate the strengths and weaknesses of the method. If possible, compare to another approach or implement your own.

Instructions to Non-Presenters

Read the assigned papers sufficiently well to understand them. On the message board, post a comment about the paper. This could be a question about something that is unclear (and you can't easily figure out), a comment about a critical flaw or especially nice aspect of the paper, or an idea for how to build on this work. The objectives of these notes are 1) to make sure everyone reads the papers beforehand; and 2) to alert the presenter to confusing topics; 3) to provide insights that we can think about before class. *Posts to the message board should be made at least 24 hours before the papers are presented.*

Instructions to Note-Taker

Each class, one person will be assigned to write down

- Questions for speaker to resolve and discuss next class period
- Ideas for improvement
- Ideas for new projects
- Papers or topics of particular interest, to follow up on later

These notes should be posted to the class discussion board before the next class.

Group Project

The group projects are an opportunity to collaboratively explore a topic of interest. These projects should try to answer a question, rather than simply achieve good results on some dataset. These projects can be related to your main research, but should be compartmentalized. See separate document for some project ideas.

Project proposal (due Feb 12)

Describe goals and planned experiments. Answer the following questions:

- What are we trying to show?
- How does this relate to existing work? Compile a list of related work.
- How will we show it (datasets, experimental setup)? For this, design one experiment that can be done within 4 weeks to validate the main idea and others to more fully explore it.

Turn in a report of length ½-1 page, plus references.

Project interim report (due Mar 19)

Describe approach (strategy and method) and results of proposed proof-of-concept experiment. Turn in a report of roughly ½-1 page.

Final report (due May 5)

Turn in a NIPS format paper (6-8 pages). It need not be highly polished, but should provide the motivation, background, approach, and experimental results. Each group will also present their work in a 15 minute talk, plus 5 minutes for questions.

Schedule (initial and subject to change… see website for updates)

Date	Presenter	Topic and Papers	Authors
Jan 20 (Tues)	Derek Hoiem	Intro: Logistics and Overview	
Jan 22 (Thurs)	Derek Hoiem	Tools and Techniques: Features and Representations Perspectives in Vision: Art The Artists as Neuroscientists (2005)	 Cavanagh
Jan 27 (Tues)	Derek Hoiem	Tools and Techniques: Classification Perspectives in Vision: Semiotics Semiotics for Beginners: signs (web) Principles of Categorization (1978)	 Chandler Rosch
Part I. Getting Around: Spatial Inference and Navigation			
Jan 29 (Thurs)	Derek Hoiem	Introduction to Spatial Inference and Navigation Purple Peril: Notes on Perception of Slant (1966) Recovering Surface Layout from an Image (2007)	 Gibson Hoiem et al.
Feb 3 (Tues)	1	Scene Gist Modeling the Shape of the Scene (2001) Depth Estimation from Image Structure (2002)	 Oliva-Torralba Oliva-Torralba
Feb 5 (Thurs)	2	Monocular Depth Estimation Make3D: Depth from a Single Image (2008)	 Saxena et al.
Feb 10 (Tues)	3	Robot Navigation Vision-based mobile robot navigation and mapping (2001) High Speed Obstacle Avoidance using Monocular Vision (2005) Use of Vision to Push Back the Path-Planning Horizon (2006)	 Se et al. Michels et al. Nabbe et al.
Feb 12 (Thurs)		Project Discussion: Goals and Experiments	
Part II. Identifying and Observing Objects			
Feb 17 (Tues)	Derek Hoiem	Introduction to Objects	
Feb 19 (Thurs)	4	Instance recognition Object Recognition from Local Scale-Invariant Features (1999) Distinctive image features from scale-invariant keypoints (2004)	 Lowe Lowe
Feb 24 (Tues)	5	Sliding window methods for category recognition A Statistical Method for 3D Object Detection (2000) Rapid Object Detection using a Boosted Cascade of Simple Features (2001)	 Schneiderman-Kanade Viola-Jones

Feb 26 (Thurs)	6	Part-based Models Pictorial Structures for Object Recognition (2005) A Discriminatively Trained, Multiscale Deformable Part Model (2008)	Felzenszwalb-Huttenlocher Felzenszwalb et al.
Mar 3 (Tues)	7	Biological Perspectives on Recognition Biological Object Recognition: Scholarpedia Face recognition by humans: 20 results (2005) Object recognition with features inspired by visual cortex (2005)	Kreiman Serre et al. Sinha et al.
Mar 5 (Thurs)	8	Recognizing Stuff On Seeing Stuff (2002) Semantic Texton Forests (2008)	Adelson Shotton et al.
Mar 10 (Tues)	9	Shared Representations Sharing Features (2004) Learning an Alphabet of Shape and Appearance (2008)	Torralba et al. Opelt et al.
Part III. Observing and Performing Actions			
Mar 12 (Thurs)	Derek Hoiem	Introduction to Actions	
Mar 17 (Tues)	10	Robotic grasping Robotic Grasping of Novel Objects Using Vision (2008) STAIR: Hardware and Software Architecture (2007)	Saxena et al. Quigley et al.
Mar 19 (Thurs)		Project Discussion: Approach and Initial Results	
Mar 24 Mar 26		No Class -- Spring Break!	
Mar 31 (Tues)	11	Flow-Based Action Recognition The Representation and Recognition of Action Using Temporal Templates (2001) Recognizing Action at a Distance (2003)	Bobick-Davis Efros et al.
Apr 2 (Thurs)	12	Keypoint-Based Action Recognition Behavior Recognition via Sparse Spatio-Temporal Features (2005) Learning Realistic Human Actions from Movies (2008)	Dollár et al. Laptev et al.
Apr 7 (Tues)	13	Recognizing Actions from Still Images Recovering Human Body Configurations (2008) Pose Primitive Based Human Action Recognition (2008)	Mori et al. Thurau-Hlavac
Apr 9 (Thurs)	14	Social Interaction Recognizing Facial Expressions in Image Sequences using Parameterized Models of Image Motion (1997) Facial expression recognition and tracking for intelligent human-robot interaction (2008)	Black-Yacoob Yang et al.

Part IV. Integrative Scene Analysis: Putting Things Together			
Apr 14 (Tues)	Derek Hoiem	Tools and Techniques: Contextual Representations and Graphical Models Graphical Models in a Nutshell Introduction to Integrative Scene Analysis	Koller et al.
Apr 16 (Thurs)	15	Objects and Scenes Context-based vision system for place and object recognition (2003)	Torralba et al.
Apr 21 (Tues)	16	Object Identity and Boundaries Image parsing: unifying detection segmentation and recognition (2003) (journal version)	Tu et al.
Apr 23 (Thurs)	17	Integration via Feature Passing Auto-context and its Application to High-level Tasks (2008) Multi-class Segmentation with Relative Location Prior (2008)	Tu Gould et al.
Apr 28 (Tues)	18	TBD	
Apr 30 (Thurs)	19	TBD	
May 5 (Tues)		Class wrap-up	