HCI 575 X (ComS 575 X) - Computational Perception

Spring 2006
Monday and Wednesday 4:10 - 5:30 p.m.
Howe Hall, Room 1324
Iowa State University
Ames, Iowa 50011

Instructor: Alexander Stoytchev

Office:	Howe Hall, Room 1620F
Phone:	515-294-5904 (email preferred)
Email:	alex@cs.iastate.edu
Web Page:	http://www.cs.iastate.edu/~alex
Office Hours:	TBD, or by appointment

Teaching Assistant: Tyler Streeter

Office:	Howe Hall, room TBD
Email:	tylerstreeter@gmail.com
Web Page:	http://www.vrac.iastate.edu/~streeter
Office Hours:	TBD, or by appointment

Course Description: This class covers statistical and algorithmic methods for sensing, recognizing, and interpreting the activities of people by a computer. This semester we will focus on machine perception techniques that facilitate and augment human-computer interaction. The main goal of the class is to introduce computational perception on both theoretical and practical levels. You will work in small groups to design, implement, and evaluate a prototype of a human-computer interaction system that uses one or more of the techniques covered in the lectures.

At the end of this class you will have an understanding of the current state of the art in computational perception and will be able to conduct original research. In addition to that, you will have the skills to design novel human-machine interfaces that push the limits of current interfaces which, in general, are deaf and blind to the human user.

Topics to be Covered: The class will cover the following topics: Overview of computational perception. Tutorials on Matlab, open computer vision (openCV), and speech recognition packages. Basic image processing. Color and movement detection. Human activity recognition based on motion history images. Tracking techniques including Kalman filters and particle filters. Gaze tracking. Face detection and face recognition: eigenfaces, cascades, and neural network-based approaches. Brain-machine interfaces. Hidden Markov models for activity recognition and speech recognition. Gesture recognition. Handwriting recognition. Affective computing, i.e., computing that relates to, arises from, or deliberately influences human emotions.

Textbook & Readings: There is no required textbook for this class. The lectures will be based on a number of sources most of which are available for download from the Internet (links will be provided on the class web page). Reading material that is not available on-line will be placed on reserve in the library. A tentative list of readings to be covered in this class is provided at the end of this document.

Organization: This class will be taught as a seminar. The students will be expected to read the assigned papers for each lecture in advance and to actively participate in class discussions. In addition, each student will be expected to present one technical paper and to lead the class discussion on that paper.

Prerequisites: This is a joint graduate and advanced undergraduate class. Previous exposure to at least 2-3 of the following fields is highly recommended: statistics, linear algebra, computer vision, artificial intelligence, human-computer interaction. Programming skills will be required for the homework assignments and for the final project. The most important prerequisite of all, however, is your interest in the course, motivation, and commitment to learning. If you are not sure whether this class is for you, please talk to the instructor.

Advanced undergraduates are welcome.

Homework Assignments: There will be four homework assignments. You will have two weeks to complete each one of them. These assignments will be used to emphasize and clarify important concepts from the lectures.

Final Project: The final project must be a research or design project that is related to the topics covered in class. You may choose to work individually or in small groups (2-3 members each). Working in groups, however, is highly recommended.

You are encouraged to select a topic for your final project as soon as possible. A written project proposal (3-5 pages) will be due on March 9. The final project report (10-15 pages) will be due on April 20. Each team will be required to present the results of their final project during the last week of the semester.

Policy on Collaboration: You are encouraged to form study groups and discuss the reading materials assigned for this class. You are allowed to discuss the homework assignments with your colleagues. However, each student will be expected to write his own solutions/code. Sharing of code is not allowed.

Attendance: You are expected to attend every class and participate in the class discussions. If you miss a class, it is your responsibility to find out what we talked about, including any announcements that were made in class.

Grading: Your grade will be determined as follows:

Class Participation:	10%
Paper Presentation:	20%
Homework Assignments:	40% (4 × 10% each)
Final Project:	30%

Tentative Schedule and Reading List

INTRO (1 week)

    Overview of the class

    Intro to Computational Perception
- "2001: HAL's Legacy", PBS Show. The documentary was produced by David Kennard and Michael O'Connell (InCA Productions) and funded by the Alfred P. Sloan Foundation.
- Rosenfeld, A. (1997). "Eyes for Computers: How HAL could see?", Chapter 10 in "HAL's Legacy, 2001's Computer as Dream and Reality", Stork, D. (Editor), MIT Press.
- Irfan A. Essa (1999). "Computers Seeing People", AI Magazine 20(2): pp. 69-82.

TUTORIALS AND BACKGROUND MATERIAL (1 week)

    Matlab Tutorial

    OpenCV Tutorial

    Speech Recognition Packages Tutorial

    Review of Probability and Linear Algebra

BASIC IMAGE PROCESSING (2 weeks)

    Mathematical Morphology
- Jain, Kasturi, and Schunck (1995). Machine Vision, "Chapter 2: Binary Image Processing," McGraw-Hill, pp. 25-72.
- Haralick and Shapiro (1993). Computer and Robot Vision, "Chapter 5: Mathematical Morphology," Addison-Wesley.

    Image Filtering
- Jain, Kasturi, and Schunck (1995). Machine Vision, "Chapter 4: Image Filtering," McGraw-Hill, pp. 112-139.
- Burt and Adelson (1983). "The Laplacian Pyramid as a Compact Image Code," IEEE Transactions on Communications, vol. 31(4), pp. 532-540.

COLOR AND MOVEMENT (1 week)

    Color and Skin detection
- Yang, Lu, and Waibel (1997). "Skin-color modeling and adaptation", CMU-CS-97-146, May 1997.

    Motion Energy and Motion History
- A. F. Bobick and J.W. Davis. "An apearance-based representation of action". In Proceedings of IEEE International Conference on Pattern Recognition 1996, August 1996, pp. 307-312.
- Davis, J. and A. Bobick (1997). "The Representation and Recognition of Action Using Temporal Templates",

In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, June 1997, pp. 928-934.

Applications
- J. Yang, W. Lu, and A. Waibel (1998). "A real time face tracker". In Proceedings of Asian Conference on Computer Vision (ACCV), volume 2, pp. 687-694.
- A. Bobick, S. Intille, J. Davis, F. Baird, C. Pinhanez, L. Campbell, Y. Ivanov, A. Schutte, and A. Wilson (1999). "The Kidsroom: A Perceptually-Based Interactive and Immersive Story Environment", Presence: Teleoperators and Virtual Environments, Vol. 8, No. 4, 1999, pp. 367-391.
- J. Davis and A. Bobick (1998). "Virtual PAT: A Virtual Personal Aerobics Trainer", Workshop on Perceptual User Interfaces, November 1998, pp. 13-18.

TRACKING TECHNIQUES (1 week)

Kalman Filter
- Maybeck, Peter S. (1979). Chapter 1 in "Stochastic models, estimation, and control",Mathematics in Science and Engineering Series, Academic Press.
- Greg Welch and Gary Bishop (2001). SIGGRAPH 2001 Course: "An Introduction to the Kalman Filter".

Particle Filters
- Michael Isard and Andrew Blake (1998). "CONDENSATION – conditional density propagation for visual tracking", International Journal of Computer Vision, 29, 1, 5–28.
- Ioannis Rekleitis (2004). A Particle Filter Tutorial for Mobile Robot Localization. Technical Report TR-CIM-04-02, Centre for Intelligent Machines, McGill University, Montreal, Quebec, Canada.

GAZE TRACKING (1 week)

Guest lecture(s) by Derrick Parkhurst
- Derrick's papers go here.

FACE DETECTION AND RECOGNITION (1 week)

Eigenfaces
- M. Turk and A. Pentland (1991). "Eigenfaces for recognition". Journal of Cognitive Neuroscience, 3(1).
- Dana H. Ballard (1999). "An Introduction to Natural Computation (Complex Adaptive Systems)", Chapter 4, pp 70-94, MIT Press.

Neural Network-Based Approaches
- Henry A. Rowley, Shumeet Baluja and Takeo Kanade (1997). "Rotation Invariant Neural Network-Based Face Detection," Carnegie Mellon Technical Report, CMU-CS-97-201.

Cascades
- Paul Viola and Michael Jones (2001). "Robust Real-time Object Detection", Second International Workshop on Statistical and Computational Theories of Vision Modeling, Learning, Computing, and Sampling, Vancouver, Canada, July 13, 2001.

BRAIN-MACHINE INTERFACES (1 week)

Guest lectures by Byron Olson
- Byron Olson papers go here.

SPRING BREAK (1 week)

HIDDEN MARKOV MODELS (1 week)
- Rabiner, Lawrence, and Juang (1993). "Theory and Implementation of Hidden Markov Models", Chapter 6 in Fundamentals of Speech Recognition, Prentice-Hall, pp. 321-389.

GESTURE RECOGNITION (1 week)
- Stefan Waldherr, Roseli Romero, Sebastian Thrun (2000). "A Gesture Based Interface for Human-Robot Interaction", Autonomous Robots, Volume 9, Issue 2, September 2000, pp. 151 - 173.
- Thad Starner and Alex Pentland (1996) "Real-Time American Sign Language Recognition from Video Using Hidden Markov Models" PAMI July 1997.
- Tanawongsuwan, R., Stoytchev, A., and Essa, I. (1999). "Robust Tracking of People by a Mobile Robotic Agent", Technical Report GIT-GVU-99-19.

HANDWRITING RECOGNITION (1 week)
- Larry Yaeger, Brandyn Webb, and Richard Lyon (1998). "Combining Neural Networks and Context-Driven Search for On-Line, Printed Handwriting Recognition in the Newton", Spring 1998 issue of AAAI's AI Magazine.
- Larry Yaeger, Richard Lyon, and Brandyn Webb (1996). "Effective Training of a Neural Network Character Classifier for Word Recognition", NIPS 1996.
- MacKenzie and Zhang (1997). "The Immediate Usability of Graffiti", Graphics Interface 1997, pp. 29-137.

TOPIC TO BE DETERMINED (1 week)

AFFECTIVE COMPUTING (1 week)

Affective Computing
- Rosalind W. Picard (1997). "Affective Computing", MIT Press.
- Rosalind W. Picard (1995). "Affective Computing", MIT Media Lab TR-321, November 1995 (abbreviated version of the book).
- A. R. Demasio (1994). "Descartes' Error: Emotion, Reason and the Human Brain",New York: Gosset/Putnam Press (excerpt).

FINAL PROJECT PRESENTATIONS (1 week)

TOTAL: 16 weeks

Week	Day/Date		Topic	Assignment
1	Monday	1/9	Introduction	
	Wednesday	1/11	Overview of Computational Perception	
2	Monday	1/16	NO CLASS: MLK Day	
	Wednesday	1/18	Matlab Tutorial, OpenCV Tutorial	Homework 1 out.
3	Monday	1/23	Basic Image Processing	
	Wednesday	1/25	Basic Image Processing	Homework 1 due.
4	Monday	1/30	Image Filtering	Homework 2 out.
	Wednesday	2/1	Image Filtering	
5	Monday	2/6	Color and Movement Detection	
	Wednesday	2/8	Color and Movement Detection	
6	Monday	2/13	Tracking Techniques	Homework 2 due.
	Wednesday	2/15	Tracking Techniques	Homework 3 out.
7	Monday	2/20	Gaze Tracking	
	Wednesday	2/22	Gaze Tracking	
8	Monday	2/27	Face Detection and Recognition	
	Wednesday	3/1	Face Detection and Recognition	Homework 3 due.
9	Monday	3/6	Brain-Machine Interfaces	
	Wednesday	3/8	Brain-Machine Interfaces	Project Proposals due.
10	Monday	3/13	NO CLASS: Spring Break	
	Wednesday	3/15	NO CLASS: Spring Break	
11	Monday	3/20	Hidden Markov Models	
	Wednesday	3/22	Hidden Markov Models	Homework 4 out.
12	Monday	3/27	Gesture Recognition	
	Wednesday	3/29	Gesture Recognition	
13	Monday	4/3	Handwriting Recognition	
	Wednesday	4/5	Handwriting Recognition	Homework 4 due.
14	Monday	4/10	TBD	
	Wednesday	4/12	TBD	
15	Monday	4/17	Affective Computing	
	Wednesday	4/19	Affective Computing	Project writeups due.
16	Monday	4/24	Project Presentations	
	Wednesday	4/26	Project Presentations	

Recommended Books

Human-Computer Interaction
- Donald A. Norman (2002). "The Design of Everyday Things", Basic Books.
- Ben Shneiderman and Catherine Plaisant (2004). "Designing the User Interface : Strategies for Effective Human-Computer Interaction," 4th Edition, Addison Wesley.
- Alan Dix, Janet Finlay, Gregory Abowd, and Russell Beale (2004). "Human Computer Interaction," 3rd edition, Prentice Hall.

Computer Vision
- Jain, Kasturi, and Schunck (1995). Machine Vision, McGraw-Hill.
- Haralick and Shapiro (1993). Computer and Robot Vision, Addison-Wesley.
- David Stork (1998). "HAL's Legacy: 2001's computer as dream and reality", MIT Press.
- Rosalind W. Picard (1997). "Affective Computing", MIT Press.

Mathematical Background
- Richard O. Duda, Peter E. Hart, David G. Stork (2000)."Pattern Classification," 2nd Edition, Wiley-Interscience.
- William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling (1992). "Numerical Recipes in C : The Art of Scientific Computing," 2nd Edition, Cambridge University Press.
- Dana H. Ballard (1999). "An Introduction to Natural Computation (Complex Adaptive Systems)", MIT Press.
- Robert V. Hogg, Allen Craig, and Joseph W. McKean (2004). "Introduction to Mathematical Statistics," 6th Edition, Prentice Hall.
- Howard Anton, Chris Rorres (2004). "Elementary Linear Algebra with Applications," 9th edition, John Wiley and Sons.

Artificial Intelligence
- Stuart Russell and Peter Norvig (2002). "Artificial Intelligence: A Modern Approach", 2nd Edition, by
- Tom M. Mitchell (1997). "Machine Learning," McGraw-Hill.