HCI / CprE / ComS 575 - Computational Perception

Spring 2010
Monday and Wednesday 2:10 - 3:30 p.m.
Howe Hall, Room 1242
Iowa State University
Ames, Iowa 50011

Instructor: Alexander Stoytchev

Office:	Howe Hall, Room 1620F
Phone:	515-294-5904 (email preferred)
Email:	alexs@iastate.edu
Web Page:	http://www.ece.iastate.edu/alexs
Office Hours:	Monday and Wednesday 3:30-4:00pm (after class), or by appointment

Teaching Assistants:	Ritika Sahai (ritika@iastate.edu)
	Connor Schenck (cschenck@iastate.edu)
	Taylor Bergquist (knexer@iastate.edu)
Office:	Howe Hall, downstairs VRAC lab (aka. room 1638)
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. The sense of self. Sensory substitution. Audio processing and speech segmentation. Sound classification and recognition. Tracking techniques including Kalman filters and particle filters. Face detection and face recognition: eigenfaces, cascades, and neural network-based approaches. Hidden Markov models for activity recognition and speech recognition. Gesture recognition. Theories of vision and intelligence. Affective computing, i.e., computing that relates to, arises from, or deliberately influences human emotions.

Readings: There are two **required** books for this class: 1) "Learning OpenCV: Computer Vision with the OpenCV Library" by Gary Bradski and Adrian Kaehler; and 2) "Getting Started with MATLAB 7: A Quick Introduction for Scientists and Engineers" by Rudra Pratap. The lectures will be based on a number of sources most of which are available for download from the Internet. Reading material that is not available on-line will be placed on reserve in the library. The reading list 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.

* The instructor reserves the right to change any and all aspects of this class for whatever reason or no reason at all (a.k.a., academic freedom).

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.

For best results take two lectures weekly. Common side effects may include sweatiness, nervousness, lack of sleep, and diarrhea. Talk to your instructor if this class is right for you.

Students with Disabilities: Iowa State University complies with the Americans with Disabilities Act and Section 504 of the Rehabilitation Act. Any student who may require an accommodation under such provisions should contact the instructor as soon as possible and no later than the end of the first week of classes or as soon as you become aware. No retroactive accommodations will be provided in this class.

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.

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 (10-15 pages) will be due on March 10. The final project report (25-30 pages) will be due on April 21. 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.

Grading: Your grade will be determined as follows:

Class Participation:	10%
Homework Assignments:	50% (5 $\times$ 10% each)
Final Project Proposal:	10%
Final Project:	30%

Tentative Schedule and Reading List

INTRO (1.5 weeks)

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.

Matlab Tutorial

OpenCV Tutorial

Review of Probability and Linear Algebra

BASIC IMAGE PROCESSING (1 week)

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 (1 week)

- 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.

PROJECT UPDATES/FORM GROUPS (0.5 weeks)

COLOR AND MOVEMENT DETECTION (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.

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.

THE SENSE OF SELF (1 week)

Phantoms in the Brain

- Ramachandran, V.S. and S. Blakeslee (1998). "Phantoms in the Brain: Probing the Mysteries of the Human Mind", William Morrow, New York. pp. 1-62.
- Melzack, R. (1992). "Phantom Limbs", Scientific American, 266, April, pp. 120-126.

Sensory Substitution

- New Scientist (2005). Cover story: "Why you have at least 21 senses", January 29, pp. 33-43.
- Andy Clark, (2003). "Who are we?", Ch. 5 in Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence, Oxford University Press.
- P. Bach-y-Rita, C. C. Collins, F. Sauders, B. White, and L. Scadden, (1969), "Vision substitution by tactile image projection". Nature, 221, pp. 963-964.
- Paul Bach-y-Rita and Stephen W. Kercel (2003). "Sensory substitution and the human-machine interface", Trends Cogn Sci, Dec;7(12):541-6.

PRELIMINARY PROJECT PRESENTATIONS (1 week)

TOPIC TO BE DETERMINED (1 week)

SPRING BREAK (1 week)

TRACKING TECHNIQUES (1.5 weeks)

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.

HIDDEN MARKOV MODELS (1 week)

Theory

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

Applications

- 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.
- 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.

WHAT IS INTELLIGENCE? (1.5 weeks)

Theories of Vision

- J. K. O'Regan and A. Noe, (2001). "A sensorimotor account of vision and visual consciousness", Behavioral and Brain Sciences, 24(5), 939- 1011.

What is Intelligence?

- Jeff Hawkins and Sandra Blakeslee, "On Intelligence: How a New Understanding of the Brain Will Lead to the Creation of Truly Intelligent Machines", Henry Holt, pp. 138-235, 2004.

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	1/11	Introduction	
	Wednesday	1/13	Matlab Tutorial	Homework 1 out
2	Monday	1/18	NO CLASS: MLK Day	
	Wednesday	1/20	Binary Image Processing	
3	Monday	1/25	Mathematical Morphology	Homework 2 out
	Wednesday	1/27	Image Filtering	
4	Monday	2/1	Color and Movement Detection	
	Wednesday	2/3	Project Ideas/Updates	
5	Monday	2/8	TBD	Homework 3 out
	Wednesday	2/10	Tracking Techniques	
6	Monday	2/15	Tracking Techniques	
	Wednesday	2/17	Tracking Techniques	
7	Monday	2/22	Preliminary Project Presentations	Homework 4 out
	Wednesday	2/24	Preliminary Project Presentations	
8	Monday	3/1	Audio Processing	
	Wednesday	3/3	Audio Processing	
9	Monday	3/8	Face Recognition	
	Wednesday	3/10	Face Detection	Project Proposals due.
10	Monday	3/15	NO CLASS: Spring Break	
	Wednesday	3/17	NO CLASS: Spring Break	
11	Monday	3/22	TBD	
	Wednesday	3/24	TBD	Homework 5 out.
12	Monday	3/29	The Sense of Self	
	Wednesday	3/31	Hidden Markov Models	
13	Monday	4/5	Hidden Markov Models	
	Wednesday	4/7	What is Intelligence?	
14	Monday	4/12	What is Intelligence?	
	Wednesday	4/14	Theories of Vision	
15	Monday	4/19	Affective Computing	
	Wednesday	4/21	Affective Computing	Project writeups due.
16	Monday	4/26	Project Presentations	
	Wednesday	4/28	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.
- Gonzalez and Woods (2008). "Digital Image Processing", Prentice Hall.

Mathematical Background

- 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.
- William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling (1992). "Numerical Recipes: The Art of Scientific Computing," 3nd Edition, Cambridge University Press.
- Christopher M. Bishop (2006). "Pattern Recognition and Machine Learning", Springer.
- Christopher M. Bishop (1995). "Neural Networks for Pattern Recognition, Claredon Press, Oxford.
- Ian T. Nabney (2004) "Netlab: Algorithms for Pattern Recognition", Springer.
- Richard O. Duda, Peter E. Hart, David G. Stork (2000)."Pattern Classification," 2nd Edition, Wiley-Interscience.
- 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.
- Tom M. Mitchell (1997). "Machine Learning," McGraw-Hill.