HCI / CprE / ComS 575 - Computational Perception

Spring 2024
Tuesday and Thursday 2:10 - 3:30 p.m.
Howe Hall, Room 1242
Iowa State University
Ames, Iowa 50011

Instructor: Alexander Stoytchev

Office:	Durham Hall, Room 303
Phone:	515-294-5904 (email preferred)
Email:	alexs@iastate.edu
Web Page:	`https://www.ece.iastate.edu/~alexs/`
Office Hours:	Tuesday and Thursday 3:30-4:00pm (after class), or by appointment

Teaching Assistants:

Manish Manepalli	Email: mmanish@iastate.edu
TA Office Hours:	TBD

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 and open computer vision (openCV). Basic image processing. Image filtering. Color and movement detection. Tracking techniques, including Kalman filters and particle filters. Face detection and face recognition: eigenfaces, cascades, and neural network-based approaches. Audio processing and speech segmentation. DFT and FFT. IDFT and IFFT. The spectrum of a signal. Audio filtering. Auditory feature extraction. Sound classification and recognition. The sense of self. Hidden Markov models for activity recognition and speech recognition. Gesture recognition.

Readings: There are two **recommended** books for this class: 1) "Learning OpenCV: Computer Vision with the OpenCV Library" by Gary Bradski and Adrian Kaehler; 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 can be found in the library. A tentative 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, signal processing, 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, and lack of sleep. Talk to your instructor if this class is right for you.

Accessibility Statement: Iowa State University is committed to assuring that all educational activities are free from discrimination and harassment based on disability status. Students requesting accommodations for a documented disability are required to work directly with staff in Student Accessibility Services (SAS) to establish eligibility and learn about related processes before accommodations will be identified. After eligibility is established, SAS staff will create and issue a Notification Letter for each course listing approved reasonable accommodations. This document will be made available to the student and instructor either electronically or in hard-copy every semester. Students and instructors are encouraged to review contents of the Notification Letters as early in the semester as possible to identify a specific, timely plan to deliver/receive the indicated accommodations. Reasonable accommodations are not retroactive in nature and are not intended to be an unfair advantage. Additional information or assistance is available online at www.sas.dso.iastate.edu, by contacting SAS staff by email at accessibility@iastate.edu, or by calling 515-294-7220. Student Accessibility Services is a unit in the Dean of Students Office located at 1076 Student Services Building.

Harassment and Discrimination: Iowa State University does not discriminate on the basis of race, color, age, ethnicity, religion, national origin, pregnancy, sexual orientation, gender identity, genetic information, sex, marital status, disability, or status as a U.S. Veteran. Inquiries regarding non-discrimination policies may be directed to Office of Equal Opportunity, 3410 Beardshear Hall, 515 Morrill Road, Ames, Iowa 50011, Tel. 515-294-7612, Hotline 515-294-1222, email eooffice@iastate.edu

Free Expression: Iowa State University supports and upholds the First Amendment protection of freedom of speech and the principle of academic freedom in order to foster a learning environment where open inquiry and the vigorous debate of a diversity of ideas are encouraged. Students will not be penalized for the content or viewpoints of their speech as long as student expression in a class context is germane to the subject matter of the class and conveyed in an appropriate manner.

Religious Accommodation: Iowa State University welcomes diversity of religious beliefs and practices, recognizing the contributions differing experiences and viewpoints can bring to the community. There may be times when an academic requirement conflicts with religious observances and practices. If that happens,

students may request the reasonable accommodation for religious practices. In all cases, you must put your request in writing. The instructor will review the situation in an effort to provide a reasonable accommodation when possible to do so without fundamentally altering a course. For students, you should first discuss the conflict and your requested accommodation with your professor at the earliest possible time. You or your instructor may also seek assistance from the Dean of Students Office at 515-294-1020 or the Office of Equal Opportunity at 515-294-7612.

Homework Assignments: There will be five homework assignments. You will have approximately two weeks to complete each one of them. These assignments will be used to emphasize and clarify important concepts. Some homeworks will have extra credit questions.

Final Project Proposal: You are encouraged to select a topic for your final project as soon as possible. You will have to present your project idea in a short 5 minute presentation (or 5 minute video for online students) that is due on Feb 8. A more detailed, but still preliminary, project idea presentation will be due on Feb 27. This is a 10-15 minute presentation/video that you and your team will have to record and upload on Canvas. A written project proposal (10-15 pages) will be due on March 7. A template for the proposal document will be provided in both Word and LaTex format.

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. The final project report (25-30 pages) will be due on April 25. Each team will be required to present the results of their final project during the last week of the semester. You will have to record a video of the presentation that includes a demo of your system and upload it on Canvas. The presentation will count as 10% of the overall grade for this class; the written report will be 20%.

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 is expected to write his/her own solutions/code. Sharing of code is not allowed.

IMPORTANT: Cheating, plagiarism, and other academic misconducts will not be tolerated and will be handled according to the ISU's academic dishonesty procedures, which are posted here:
`http://catalog.iastate.edu/academic_conduct/#academicdishonestytext`

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

Additional standard **syllabus statements** for all classes at ISU are posted on Canvas.

Grading: Your grade will be determined as follows:

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

Tentative Reading List and Schedule

INTRO

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.

Matlab Tutorial

OpenCV Tutorial

Review of Probability and Linear Algebra

BASIC IMAGE PROCESSING

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.

PROJECT UPDATES/FORM GROUPS

COLOR AND MOVEMENT DETECTION

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.

COMPLEX NUMBERS

- Howard Anton and Chris Rorres (1991). "Elementary Linear Algebra: Applications Version," 6th ed., Upper Saddle River, NJ: Prentice-Hall, Section 10.4: Complex Vector Spaces. pp. 501–505.
- TBD

PRELIMINARY PROJECT PRESENTATIONS

FFT AND IFFT

- Cormen, T., Leiserson, C., and Rivest R. (1990). "Introduction to Algorithms," The MIT Press, Cambridge, MA. Chapter 32: Polynomials and the FFT.
- TBD

AUDIO PROCESSING

- Ian Vince McLoughlin (2016). "Speech and Audio Processing: A MATLAB-Based Approach," Cambridge University Press, Cambridge. Chapter 2: Basic Audio Processing, pp. 9-53.
- TBD

HIDDEN MARKOV MODELS

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.

FACE DETECTION AND RECOGNITION

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.

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.

THE SENSE OF SELF

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.

TRACKING TECHNIQUES

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.

FINAL PROJECT PRESENTATIONS

Week	Day/Date		Topic	Assignment
1	Tuesday	1/16	Introduction	
	Thursday	1/18	Motivation and Inspiration	Homework 1 out
2	Tuesday	1/23	Matlab Tutorial	
	Thursday	1/25	Binary Image Processing	
3	Tuesday	1/30	Mathematical Morphology	Homework 2 out
	Thursday	2/1	OpenCV Tutorial	
4	Tuesday	2/6	Image Filtering	
	Thursday	2/8	Project Ideas/Updates	
5	Tuesday	2/13	Color and Movement Detection	Homework 3 out
	Thursday	2/15	Complex Numbers	
6	Tuesday	2/20	Roots of Unity	
	Thursday	2/22	Complex Sinusoids	
7	Tuesday	2/27	Preliminary Project Presentations	
	Thursday	2/29	Preliminary Project Presentations	
8	Tuesday	3/5	The Spectrum of a Signal	Homework 4 out.
	Thursday	3/7	The FFT Algorithm	Project Proposals due.
9	Tuesday	3/12	NO CLASS: Spring Break	
	Thursday	3/14	NO CLASS: Spring Break	
10	Tuesday	3/19	FFT Shift	
	Thursday	3/21	The IFFT Algorithm	
11	Tuesday	3/26	Audio Processing Examples	
	Thursday	3/28	Hidden Markov Models	
12	Tuesday	4/2	Hidden Markov Models	Homework 5 out.
	Thursday	4/4	Hidden Markov Models	
13	Tuesday	4/9	Face Recognition	
	Thursday	4/11	Face Detection	
14	Tuesday	4/16	Theories of Vision	
	Thursday	4/18	The Sense of Self	
15	Tuesday	4/23	Kalman Filters	
	Thursday	4/25	Particle Filters	Project writeups due.
16	Tuesday	4/30	Project Presentations	
	Thursday	5/2	Project Presentations	