# HCI / CprE / ComS 575 - Computational Perception

**Spring 2020**
**Tuesday and Thursday 2:10 - 3:30 p.m.**
**Howe Hall, Room 1252**
**Iowa State University**
**Ames, Iowa 50011**

**Instructor:** Vladimir Sukhoy

| | |
|---|---|
| Office: | Howe Hall, Room 2624E |
| Phone: | 630-445-1332 (email preferred) |
| Email: | sukhoy@iastate.edu |
| Web Page: | `http://sukhoy.public.iastate.edu/` |
| Office Hours: | Tuesday and Thursday 3:30-4:00pm (after class), or by appointment |

| | |
|---|---|
| **Teaching Assistants:** | Karter Krueger |
| TA Office: | TLA, Coover Hall |
| TA Office Hours: | Fridays noon–2:00pm |


**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. 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. Audio filtering. Auditory feature extraction. Sound classification and recognition. Hidden Markov models for activity recognition and speech recognition. Gesture recognition.

**Readings:** There are two **required** 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 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, 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, lack of sleep, and diarrhea. Talk to your instructor if this class is right for you.

**Students with Disabilities:** Iowa State University is committed to assuring that all educational activities are free from discrimination and harassment based on disability status. All students requesting accommodations are required to meet with staff in Student Disability Resources (SDR) to establish eligibility. A Student Academic Accommodation Request (SAAR) form will be provided to eligible students. The provision of reasonable accommodations in this course will be arranged after timely delivery of the SAAR form to the instructor. Students are encouraged to deliver completed SAAR forms as early in the semester as possible. SDR, a unit in the Dean of Students Office, is located in room 1076, Student Services Building or online at `www.dso.iastate.edu/dr/`. Contact SDR by e-mail at disabilityresources@iastate.edu or by phone at 515-294-7220 for additional information.

**Harassment and Discrimination:** Iowa State University strives to maintain our campus as a place of work and study for faculty, staff, and students that is free of all forms of prohibited discrimination and harassment based upon race, ethnicity, sex (including sexual assault), pregnancy, color, religion, national origin, physical or mental disability, age, marital status, sexual orientation, gender identity, genetic information, or status as a U.S. veteran. Any student who has concerns about such behavior should contact his/her instructor, Student Assistance (`http://www.dso.iastate.edu/sa/`) at 515-294-1020 or email dso-sas@iastate.edu, or the Office of Equal Opportunity and Compliance (`http://www.eoc.iastate.edu/`) at 515-294-7612.

**Religious Accommodation:** If an academic or work requirement conflicts with your religious practices and/or observances, you may request reasonable accommodations. Your request must be in writing, and your instructor or supervisor will review the request. You or your instructor may also seek assistance from the Dean of Students Office (`http://www.dso.iastate.edu/`) or the Office of Equal Opportunity and Compliance (`http://www.eoc.iastate.edu/`).

**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.

**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 12. 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.

**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. 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:

| | |
|---|---|
| Homework Assignments: | 55% ($5 \times 11\%$ each) |
| Final Project Proposal: | 10% |
| Final Project: | 30% |
| Class Participation: | 5% |

# 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.

## PRELIMINARY PROJECT PRESENTATIONS

## MATHEMATICS ON THE COMPLEX PLANE

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

## 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.

## 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 | 1/14 | Introduction | |
| | Thursday | 1/16 | Motivation and Inspiration | Homework 1 out |
| **2** | Tuesday | 1/21 | Matlab Tutorial | |
| | Thursday | 1/23 | Binary Image Processing | |
| **3** | Tuesday | 1/28 | Mathematical Morphology | Homework 2 out |
| | Thursday | 1/30 | OpenCV Tutorial | |
| **4** | Tuesday | 2/4 | Image Filtering | |
| | Thursday | 2/6 | Project Ideas/Updates | |
| **5** | Tuesday | 2/11 | Color and Movement Detection | Homework 3 out |
| | Thursday | 2/13 | Review of Complex Math | |
| **6** | Tuesday | 2/18 | Review of Complex Math | |
| | Thursday | 2/20 | Complex Sinusoids | |
| **7** | Tuesday | 2/25 | Preliminary Project Presentations | |
| | Thursday | 2/27 | Preliminary Project Presentations | |
| **8** | Tuesday | 3/3 | The Spectrum of a Signal | Homework 4 out |
| | Thursday | 3/5 | The FFT Algorithm | |
| **9** | Tuesday | 3/10 | IFFT and Matlab Examples | Project Proposals due |
| | Thursday | 3/12 | Audio Processing | |
| **10** | Tuesday | 3/17 | NO CLASS: Spring Break | |
| | Thursday | 3/19 | NO CLASS: Spring Break | |
| **11** | Tuesday | 3/24 | Extracting Audio Features | |
| | Thursday | 3/26 | Hidden Markov Models | |
| **12** | Tuesday | 3/31 | Hidden Markov Models | Homework 5 out |
| | Thursday | 4/2 | Speech Processing and Recognition | |
| **13** | Tuesday | 4/7 | TBD | |
| | Thursday | 4/9 | Face Recognition | |
| **14** | Tuesday | 4/14 | Face Detection | |
| | Thursday | 4/16 | Tracking Techniques | |
| **15** | Tuesday | 4/21 | Tracking Techniques | Project writeups due |
| | Thursday | 4/23 | TBD | |
| **16** | Tuesday | 4/28 | Project Presentations | |
| | Thursday | 4/30 | Project Presentations | |