Computational Vision

courses.kersten.org **Psychology Department , University of Minnesota Psy 5036W, Fall 2017, 3 credits #34359**
08:15 A.M. - 9:30 A.M. Mondays and Wednesdays
Elliott Hall N668

Instructor: Daniel Kersten. *Office*: S212 Elliott Hall. *Phone*: 612 625-2589 *email*: kersten@umn.edu
Office hours: Mondays 9:30-10:30 am or by appointment.

The visual perception of what is in the world is accomplished continually, instantaneously, and usually without conscious thought. The very effortlessness of perception disguises the underlying richness of the problem. We can gain insight into the processes and functions of human vision by studying the relationship between neural mechanisms and visual behavior through computer analysis and simulation. Students will learn about the anatomy and neurophysiology of vision and how they relate to the phenomona of perception. An underlying theme will be to treat vision as a process of statistical inference. There will be in-class programming exercises using the language Mathematica. No prior programming experience is required; however, some familiarity with probability, vector calculus and linear algebra is helpful.

Readings¶

Main¶

- Lecture notes, Main Readings & Supplementary Material are online.

Additional readings¶

Math and vision¶

- (**EV**) Early Vision. Yuille and Kersten. In *From Neuron to Cognition via Computational Neuroscience*, M.A. Arbib, James J. Bonaiuto Editors, Cambridge MA: The MIT Press.
- Understanding Vision: Theory, Models, and Data. Li Zhaoping. 2014.(publisher page) (author's web outline)

Functional human vision¶

- (**FV**) *Foundations of Vision*. Wandell (web)

Neurophysiology¶

- (**NVN**) *The New Visual Neurosciences*. John S. Werner and Leo M. Chalupa, edts. 2014.

Software¶

Mathematica¶

Mathematica is the primary programming environment for this course. Students who have registered for the course will have Google Docs access through the Psychology Department's site license.

Alternatives: Mathematica is available in several labs on campus, go to http://www.oit.umn.edu/computer-labs/software/index.htm
You may wish to purchase *Mathematica for Students* see http://www.wolfram.com/products/student/mathforstudents/index.html.
You can also access ***Mathematica*** on the CLA servers:

mac (Note: you may have to change the forward slash to a back slash)
windows

If you never programmed before go here. If you have programming experience, go here.

For user help on using Mathematica, see: http://mathematica.stackexchange.com

Python/IPython¶

http://ipython.org
http://jupyter-notebook-beginner-guide.readthedocs.org/en/latest/index.html
http://www.scipy.org

For an online course in using Python and PsychoPy for research in human vision see:
http://nbviewer.ipython.org/github/gestaltrevision/python_for_visres/blob/master/index.ipynb

Writing¶

- Gopen, G. D., & Swan, J. A., 1990. The Science of Scientific Writing. American Scientist, 78, 550-558.

- **Supplementary:**
    - The Sense of Style: The Thinking Person's Guide to Writing in the 21st Century (2014), Pinker, Steven. (amazon link)
    - Penrose, A. M., & Katz, S. B. (1998). Writing in the Sciences: Exploring Conventions of Scientific Discourse. New York: St. Martin's Press, Inc.
    - American Psychological Association. (2009). Publication manual of the American Psychological Association (6th ed.). Washington, DC: American Psychological Association

- **Writing assistance.** THE CENTER FOR WRITING offers free one-to-one writing assistance to undergraduate and graduate students, with appointments up to 45 minutes. Nonnative speaker specialists are available. For more information, see http://writing.umn.edu.

- Psychology department resources: http://writing.psych.umn.edu/student-resources

**Grade Requirements

**There will be programming assignments and a final project.

The grade weights are:

- Exercise/programming assignments: 55%

- Final project presentations: 5 %

- Final project : 40% (four parts: 2%+5%+5%+28%)

The programming assignments will use the *Mathematica* programming environment. No prior experience with *Mathematica* is necessary.

Assignment due By the midnight on the day due. _**Late Policy: Assignments turned in within 24 hours following the due date will have 15% deducted from the assignment score. Assignments turned in between 24 and 48 hours following the due date will have 30% deducted from the score. Assignments more than 48 hours late will receive a score of zero.**_

Lectures¶

Check this section before each class for recent additions and revisions.

(5036W Course material from 2015)

Lecture notes are in Mathematica Notebook and pdf format. You can download the Mathematica notebook files below to view with Mathematica or Wolfram CDF Player (which is free).

[University Calendar](http://onestop.umn.edu/onestop/calendar.html)	**Date**	**Lecture**	_**Main Readings**_
I. Introduction	Sep 6	1\. Introduction to Computational Vision	_[1.IntroToComputationalVision.nb](Lectures/1_MultidisciplinaryStudy/1.IntroToComputationalVision.nb) ([pdf](Lectures/1_MultidisciplinaryStudy/1.IntroToComputationalVision.nb.pdf))_ Olshausen, B. A. (2013). Perception as an Inference Problem. M. Gazzaniga (Ed.), The New Cognitive Neurosciences, 5th Edition (pp. 1â€“22). MIT Press. (pp. 1â€“18 MIT Press. ([pdf](../../coursepapers/Olshausen2013Perception_as_an_Inference_Problem.pdf))
	Sep 11	2.Limits to Vision	_[2.LimitsToVision.nb](Lectures/2_Limits to Vision/2_LimitsToVision.nb) [(pdf)](Lectures/2_Limits to Vision/2_LimitsToVision.nb.pdf)__ Hecht, S., Shlaer, S., & Pirenne, M. H. (1942). Energy, quanta, and vision. Journal of General Physiology, 25, 819-840\. ([pdf](../../coursepapers/hecht+al_42.pdf))_

	Sep 13	3\. The Ideal Observer	_[3.TheIdealObserver.nb](Lectures/3_TheIdealObserver/3_TheIdealObserver.nb) [(pdf)](Lectures/3_TheIdealObserver/3_TheIdealObserver.nb.pdf)_
	Sep 18	4\. Ideal observer analysis: Humans vs. ideals. Neurons vs. ideals	_[4.IdealObserverAnalysis.nb](Lectures/4_Ideal Observer Analysis/4_IdealObserverAnalysis.nb)_ ([pdf (Lectures/4_Ideal Observer Analysis/4_IdealObserverAnalysis.nb.pdf))
II. Image formation, pattern synthesis	Sep 20	5.Psychophysics: tools & techniques	_[5.Psychophysics.nb](Lectures/5_PsychophysicsSKEobserver/5_Psychophysics.nb)[](Lectures/5_PsychophysicsSKEobserver/5_Psychophysics.nb) [(pdf)](Lectures/5_PsychophysicsSKEobserver/5_Psychophysics.nb.pdf) _
	Sep 25	6\. Bayesian decision theory & perception	_[6.BayesDecisionTheory.nb](Lectures/6_BayesDecisionTheory/6_BayesDecisionTheory.nb.pdf) [(pdf) (Lectures/6_BayesDecisionTheory/6_BayesDecisionTheory.nb.pdf)_ _Geisler, W. S., & Kersten, D. (200 Illusions, perception and Bayes. Nat Neurosci, 5(6), 508-510\. ([pdf](http://gandalf.psych.umn.edu/%7Ekersten/kersten-lab/papers/GeislerKerstennn0602-508.pdf))_
	Sep 27	7\. Limits to spatial resolution, image modeling, introduction to linear systems	_[7.ImageModelLinearSystems.nb](Lectures/7.ImageModelingLinearSystems/7.ImageModelLinearSystems.nb) [(pdf)](Lectures/7.ImageModelingLinearSystems/7.ImageModelLinearSystems.nb.pdf)_ _Campbell, F. W., & Green, D. (1965). Optical and retinal factors affecting visual resolution. Journal of Physiology (Lond.), 181, 576-593\. ([pdf](../../coursepapers/CampbellGreen_JP1965.pdf))_
III. Early visual coding	Oct 2	8\. Linear systems analysis	_[8.LinearSystemsOptics.nb](Lectures/8\. Spatial filters/8.LinearSystemsOptics.nb) [(pdf)](Lectures/8\ Spatial filters/8.LinearSystemsOptics.nb.pdf) _
	Oct 4	9\. Features and filters. Spatial filter models of early human vision	_[9.NeuralSpatialFiltering.nb](Lectures/9\. Multi-scale analysis/9.NeuralSpatialFiltering.nb) [(pdf)](Lectures/9\. Multi-scale analysis/9.NeuralSpatialFiltering.nb.pdf)_
	Oct 9	10\. Features and filters. Local processing & image analysis	_[10.ImageProcessing.nb](Lectures/10.ImageManipulations/10.ImageProcessing.nb) [(pdf)](Lectures/10.ImageManipulations/10.ImageProcessing.nb.pdf)_ Gollisch, T., & Meister, M. (2010). Eye Smarter t Scientists Believed: Neural Computations in Circuits of the Retina. Neuron, 65(2), 150–164\. ([pdf](../.. coursepapers/Gollisch2010Eye_Smarter_than_Scientists_Believed_Neural_Computations_in_Circuits_of_the_Retina.

	Oct 11	11\. Coding efficiency: Retina	_[11.CodingEfficiency.nb](Lectures/11\. Efficient coding/11.CodingEfficiency.nb) [(pdf)](Lectures/11\. Efficient coding/11.CodingEfficiency.nb.pdf)_ Geisler, W. S. (2008). Visual perception and the statistica properties of natural scenes. Annu Rev Psychol, 59, 167-192\. [(pdf)](../../coursepapers/Geisler2008annurev.psych.58.110405.085632.pdf)
	Oct 16	12\. Coding efficiency: Cortex	_[12.SpatialCodingEfficiency.nb](Lectures/12\. Efficent coding Spatial/12.SpatialCodingEfficiency.nb) [(pdf)](Lectures/12\. Efficent coding Spatial/12.SpatialCodingEfficiency.nb.pdf)_ _Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. Annu Rev Neurosci, 24, 1193-1216.[(pdf)](../../coursepapers/SimoncelliOlshausenAnnRev1999.pdf))_
IV. Intermediate-level vision, integration, grouping	Oct 18	13\. Edge detection	[13.EdgeDetection.nb](Lectures/13\. Edges/13.EdgeDetection.nb) [(pdf)](Lectures/13\. Edges/13.EdgeDetection.nb.pdf)
	Oct 23	14. Objects and scenes from images. The visual cortical pathways and hierarchy.	_[14.ScenesfromImages.nb](Lectures/14.Scenes from images/14.ScenesfromImages.nb) [(pdf)](Lectu 14.Scenes from images/14.ScenesfromImages.nb.pdf)__ von der Heydt R (2003) Image parsing mechanisms of the visual cortex. In: The Visual Neurosciences (Werner JS, Chalupa LM, eds.), pp 1139-1150\. Cambridge, Mass.: MIT press.[(pdf)](../../coursepapers/VonderHeydt_ImageParsing_neuroscience.pdf))_ Kersten, D. J., & Yuille, A. L. (2014). Inferential Model the Visual Cortical Hierarchy. In M. S. Gazzaniga & G. R. Mangun (Eds.), The New Cognitive Neurosciences, 5th Edition (pp. 1â€"22). MIT Press. [(pdf)](http://vision.psych.umn.edu/users/kersten kersten-lab/coursepapers/KerstenYuilleNewCogNeuro2014.pdf)
	Oct 25	15\. Scene-based generative models	_[15.SurfaceGeometryDepth.nb](Lectures/15\. Surfacegeometrydepth/15.SurfaceGeometryDepth.nb [(pdf)](Lectures/15\. Surfacegeometrydepth/15.SurfaceGeometryDepth.nb.pdf)_ _Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian Inference. Annual Review of Psychology, 55, 271-304\. ([pdf](../../coursepapers/KerstenMamassianYuille_2004_annurev.psych.55.090902.142005.pdf))_
	Oct 30	16\. Shape-from-X	_[16.ShapeFromX.nb](Lectures/16\. Shape-from-X/16.ShapeFromX.nb) [(pdf)](Lectures/16\. Shape-fro X/16.ShapeFromX.nb.pdf)_

Nov 1	17\. Shape from shading Overview of python/ipython for computational vision[ ](https://wakari.io/sharing/bundle/kersten/Lect_19Intro_Python)	_[17. Shape from shading.nb](Lectures/17.Shape from shading/17.ShapeFromShading.nb)_ [(pdf)](Lectures/17.Shape from shading/17.ShapeFromShading.nb.pdf) [Lect_17Intro_Python.ipynb](Lecture 17_PythonForVision/Lect_17Intro_Python.ipynb) (source) ([pdf](Lectures/17_PythonForVision/Lect_17Intro_Python.pdf)) _[17\. IPython notebook](http://nbviewer.jupyter.org/url/gandalf.psych.umn.edu/users/kersten/kersten-lab/courses/Psy5036W2017/Lectures/17_PythonForVision/Lect_17Intro_Python.ipynb)_ [Demos](Lectures/17_PythonForVision/Demos/index.html) by Weichao Qiu and Dan Kersten, supplement to _**Early Vision**._ Yuille and Kersten. A chapter in _From Neuron to Cognition via Computational Neuroscience_, M.A. Arbib, James J. Bonaiut Editors, Cambridge MA: The MIT Press, in 2016
Nov 6	18\. Motion: optic flow	_[18.MotionOpticFlow.nb](Lectures/18.MotionOpticFlow/18.MotionOpticFlow.nb) [(pdf)](Lectures/18.MotionOpticFlow/18.MotionOpticFlow.nb.pdf)_ _OpenCV python demo:_ [OpticFlowSparse.ipynb](http://nbviewer.ipython.org/url/gandalf.psych.umn.edu/users/kersten/kersten-lab/courses/Psy5036W2015/Lectures/18.MotionOpticFlow/OpticFlowSparse.ipynb) needs: [648aa10.avi](Lectures/17_PythonForVision/648aa10.avi)
Nov 8	19\. Motion: biological, human perception	_[19.MotionHumanPerception.nb](Lectures/19.MotionIllusionsBayes/19.MotionHumanPerception.nb [(pdf)](Lectures/19.MotionIllusionsBayes/19.MotionHumanPerception.nb.pdf)_ _Weiss, Y., Simoncelli, P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. Nat Neurosci, 5(6), 598-604. ([pdf](../coursepapers/WeissSimonAdelNatNeu2002.pdf))_
Nov 13	20\. Material perception	_[20.SurfaceMaterial.nb](Lectures/20\. Surface material/20.SurfaceMaterial.nb) [(pdf)](Lectures/20\. Surface material/20.SurfaceMaterial.nb.pdf)_ V1 and lightness ([pdf](Lectures/20.%20Surface%20material/V1Lightness.pdf)) Doerschner, K., Fleming, R. W., Yilmaz, O., Schrater, P. R Hartung, B., & Kersten, D. (2011). Visual motion and the perception of surface material. Current Biolo 21(23), 2010–2016\. ([pdf](../../coursepapers/Doerschner2011Doerschner2011CURRENT-BIOLOGY-S-11-00953-3.pdf))
Nov 15	21\. Texture.	_[21.Texture.nb](Lectures/21\. Texture/21.Texture.nb) [(pdf)](Lectures/21\. Texture/21.Texture.nb.pd Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. Nature Publishing Group, 14(9 1195-1201\. http://doi.org/10.1038/nn.2889 ([pdf](../../coursepapers/Freeman2011Metamers_of_the_ventral_stream.pdf))
Nov 20	22.Science writing (Thanksgiving week)	_[22.ScienceWriting.nb](Lectures/22\. Science Writing/22.ScienceWriting.nb) [(pdf)](Lectures/22\. Scie Writing/22.ScienceWriting.nb.pdf)_

	Nov 22	23.Perceptual integration	_[23.PerceptualIntegration.nb](Lectures/23.PerceptualIntegration/23.PerceptualIntegration.nb) [(pdf) (Lectures/23.PerceptualIntegration/23.PerceptualIntegration.nb.pdf) _
V. High-level vision	Nov 27	24. Object recognition I	_[24.ObjectRecognition.nb](Lectures/24\. ObjectRecognition_I/24.ObjectRecognition.nb) [(pdf)](Lectu 24\. ObjectRecognition_I/24.ObjectRecognition.nb.pdf)_ _DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012 How does the brain solve visual object recognition? Neuron, 73(3), 415–434\._ ([pdf](../../coursepaper DiCarlo2012How_does_the_brain_solve_visual_object_recognition-2.pdf))
	Nov 29	25\. Object recognition II feeforward architectures	_25_Bidirectional_I.key.pdf ([pdf](Lectures/25\. ObjectRecognition_II/25_Bidirectional_I.pdf)) __ Ullman Vidal-Naquet, M., & Sali, E. (2002). Visual features of intermediate complexity and their use in classification. Nat Neurosci, 5(7), 682-687\. ([pdf](../../coursepapers/UllmanNatureNeuro2002.pdf)) _
	Dec 4	26. Object recognition III feedback architectures	_26_BidirectionalFeedback.key.pdf_ [(pdf](Lectures/26\. ObjectRecognition_III/ 26_BidirectionalFeedback.pdf))
	Dec 6	27. Empirical evidence for bidirectional computations	27.EmpiricalEvidenceBidirectionalProcessing([pdf](Lectures/27\. EmpiricalEvidenceBidirectionalProcessing/27\. EmpiricalEvidenceBidirectionalProcessing.pdf))
	Dec 11	28\. Vision for action, spatial layout, heading. Homegeneous coordinates.	_[28.SpatialLayoutScenes.nb](Lectures/28\. SpatialLayoutScenes/28.SpatialLayoutScenes.nb) [(pdf] (Lectures/28\. SpatialLayoutScenes/28.SpatialLayoutScenes.nb.pdf))_ _Kalman filter notes ([pdf](http: vision.psych.umn.edu/users/kersten/kersten-lab/courses/Psy5036W2015/Lectures/27\. SpatialLayoutScenes/kalman.pdf))_
	Dec 13 (Last day of class)		_In Class Project Presentations_
	Dec 21		

Final Project Assignment.

Goal: This course integrates the behavioral, neural and computational principles of perception. Students often find the interdisciplinary integration to be the most challenging aspect of the course. Through writing, you will learn to synthesize results from diverse and typically isolated disciplines. By writing about your project work, you will learn to think through the broader implications of your project, and to effectively communicate the rationale and results of your contribution in words. You will do a final page research report in which you will describe, in the form of a scientific paper, the results of an original computer program on a topic in computational vision.

Your final project will involve: 1) a computer program and; 2) a 2000-3000 word final paper describing your project. For your computer project, you will do one of the following: 1) Write a program to simulate a model from the computer vision literature ; 2) Design and program a method for solving some problem in perception. 3) Design and program a psychophysical experiment to study an aspect of human visual perception. The results of your final project should be written up in the form of a short scientific paper or Mathematica Notebook, describing the motivation, methods, results, and interpretation.

If you choose to write your program in Mathematica, your paper and program can be combined can be formated as a Mathematica notebook. See: Books and Tutorials on Notebooks. If you do your final project using Python, you can turn your paper in as a Jupyter notebook.

Your paper will be critiqued and returned for you to revise and resubmit in final form. You should write for an audience consisting of your class peers.

Completing the final paper involves 4 steps. Each step requires that you email a document to the teaching assistant.

1. **Outline (2% of grade)****.** You will submit a working title and paragraph outline by the deadline noted in the syllabus. These outlines will be critiqued in order to help you find an appropriate focus for your papers. (Consult with the instructor or TA for ideas well ahead of time).
2. **Complete draft (5% of grade).** A double-spaced, complete draft of the paper must be turned in by the deadline noted in the syllabus. Papers should be between **2000** and **3000** words. In addition to the title, author and date lines, papers must include the following sections: Abstract, Introduction, Methods, Results, Discussion, and Bibliography. Use citations to motivate your problem and to justify your claims. Cite authors by name and date, e.g. (Marr & Poggio, 1979). *Citations should be original sources, not wikipedia.* Use a standard citation format, such as APA. (The UM library has information on style guides, and in particular APA style.) Papers must be typed, with a page number on each page. Figures should be numbered and have figure captions. This draft will be reviewed by your instructor and one of you class peers. **The point break down for the total 5% is: 2 pts for completing Introduction, 2 pts for completing Methods, 1 pt for completing Discussion)**
3. **Peer commentary (5% of grade)**. You will submit a written commentary **(200** to **500** words) on a complete draft of one of your class peers. The project drafts and commentaries will be anonymous. The commentary should provide feedback to improve the quality and clarity of the writing.
4. **Final draft (20% of grade) and "Cover letter" (8% of grade).** The final draft must be turned in by the date noted on the syllabus. The "Cover letter" should describe how your revision addressed comments from your peer evaluator and from your instructor. It should itemize key criticisms together with a brief description of the changes you made to your draft manuscript.
5. **Some Resources:**
6. Student Writing Support: Center for Writing, 306b Lind Hall and satellite locations (612.625.1893) http://writing.umn.edu.

*NOTE: Plagiarism, a form of scholastic dishonesty and a disciplinary offense, is described by the Regents as follows: Scholastic dishonesty means plagiarizing; cheating on assignments or examinations; engaging in unauthorized collaboration on academic work; taking, acquiring, or using test materials without faculty permission; submitting false or incomplete records of academic achievement; acting alone or in cooperation with another to falsify records or to obtain dishonestly grades, honors, awards, or professional endorsement; altering, forging, or misusing a University academic record; or fabricating or falsifying data, research procedures, or data analysis. http://www1.umn.edu/regents/policies/academic/Code_of_Conduct.html. See too: http://writing.umn.edu/tww/plagiarism/ and*http://writing.umn.edu/tww/plagiarism/definitions.html

Privacy Statement