Schedule:

May 8 (Wednesday)

9:00-9:30	Qasim Zaidi, Elias Cohen and Anshul Jain Model for the salience of mirror symmetry in natural patterns.
9:30-10:00	Ingo Fründ and James Elder Human selectivity for statistical properties of natural shapes.
10:00-10:30	Break
10:30-11:00	Vicky Froyen, Jacob Feldman and Manish Singh Perceptual grouping as Bayesian estimation of mixture models.
11:00-11:30	Gregory Francis Publication bias, modeling, and theorizing in vision science.
11:30-12:00	Discussion
12:00-2:00	Lunch
2:00-2:30	Stephen Grossberg How do object reference frames and motion vector decomposition emerge in laminar cortical circuits?
2:30-3:00	Bo Cao and Arash Yazdanbakhsh A neural model of MSTd neurons for eye movement compensation.
3:00-3:30	Break
3:30-4:00	Jihyun Kim A biophysical model of retinal circuitry: spatiotemporal dynamics.
4:00-4:30	Oliver Layton and Arash Yazdanbakhsh A multi-scale model of figure-ground segregation in the primate visual system.
4:30-5:00	Discussion

May 9 (Thursday)

9:00-9:30	Michele Rucci, Jonathan Victor and Murat Aytekin Seeing space through time during natural fixation.
9:30-10:00	Emre Akbas and Miguel Eckstein Object detection with a foveated visual system.
10:00-10:30	Break
10:30-11:00	Michael E. Rudd Perceptual weighting of local contrast information in an edge integration model of neural lightness computation.
11:00-11:30	Ruth Rosenholtz What is our baseline model of visual processing?
11:30-12:00	Discussion
12:00-2:00	Lunch
2:00-2:30	Holly Gerhard, Felix Wichmann and Matthias Bethge How Sensitive Is the Human Visual System to the Local Statistics of Natural Images?
2:30-3:00	Christopher Tyler Modeling ‘Binoculomotor’ Dynamics.
3:00-3:30	Break
3:30-4:00	Harald Ruda and Ennio Mingolla A model of depth ordering from motion occlusion and disocclusion in textured scenes.
4:00-4:30	Inna Tsirlin, Laurie Wilcox and Robert Allison A biologically-plausible computational model of depth from monocular occlusions and binocular disparity.
4:30-5:00	Discussion
5:00-5:30	Business meeting

Poster:

Ram P. Sharma, Donna Strickland, and Melanie C. W. Campbell
Nonlinear Laser pulse induced plasma effects on the performance of crystalline lens.

May 10 (Friday)

9:00-9:30	Jian Ding, Stanley Klein and Dennis Levi Modeling studies on binocular combination.
9:30-10:00	Stanley Klein, Thom Carney, Austin Roorda, Lawrence Sincich, Claudio Privitera, Ram Sabesan and Gene Switkes Modeling color perception from single and multiple cone stimulation.
10:00-10:30	Break
10:30-11:00	Tadamasa Sawada A geodesic on a non-differentiable part of a 3D surface and its role in visual perception.
11:00-11:30	Ilan Kadar and Ohad Ben-Shahar From Perceptual Relations to Scene Gist Recognition.
11:30-12:00	Conclusion

Abstracts:

Qasim Zaidi[1], Elias Cohen[2] and Anshul Jain[1]
[1]Graduate Center for Vision Research, SUNY Optometry, [2]Vanderbilt University

Model for the salience of mirror symmetry in natural patterns

Mirror symmetry detection is considered extremely rapid, based on experiments with random noise. Symmetry detection in natural settings, however, is accomplished against structured backgrounds. We measured temporal thresholds for detecting the mirror symmetry axis in patterns assembled from 101 natural textures. Thresholds ranged from 28 to 568 msec indicating a wide range of salience (1/Threshold). For estimating symmetric energy, we used pairs of mirror-symmetric cortex-like filters, connected with AND junctions, in a filter-prune-filter-select model. The model easily identified the axis of symmetry for all patterns, but symmetry magnitude quantified at this axis correlated weakly with salience. To examine context effects, we used the model to estimate magnitudes of approximate symmetry at flanking and orthogonal axes. These showed strong negative correlations with salience, revealing context interference with symmetry detection. Expanded regression modeling including the context-based measures explained the salience results, and yielded understanding of why perceptual symmetry differs from mathematical characterizations. We also apply the model to discriminating the rotation direction of non-rigid symmetric shapes when perceived rotation could correspond only to rotation of global symmetry, and tracking of shape contours or local features was uninformative.

Ingo Fründ and James Elder
York University, Toronto

Human selectivity for statistical properties of natural shapes

We study statistical regularities in the bounding contours of natural shapes, and how humans exploit these regularities for visual perception. We approximated naturalistic shapes by equilateral polygons and extracted a set of low-order statistical features based on the turning angles of these polygons.
To measure human selectivity for these features, we developed a method for generating contour metamers that match the natural contours on selected subsets of these features, but which lack all other statistical regularities found in the natural shapes. We measured the length of the contour fragment required for threshold performance in discriminating between fragments from naturalistic shapes and fragments from metamers.

Vicky Froyen, Jacob Feldman and Manish Singh
Dept. of Psychology, Center for Cognitive Science, Rutgers University, New Brunswick

Perceptual grouping as Bayesian estimation of mixture models

We propose a Bayesian approach to perceptual grouping in which the goal of the computation is to estimate the organization that best explains an observed configuration of image elements. We formalize the problem as a mixture estimation problem, where it is assumed that the configuration of elements is generated by a set of distinct components ("objects" or "clusters"), whose underlying parameters we seek to estimate (including location and "ownership" of image elements). Among other parameters, we can estimate the number of components in the image, given a set of assumptions about the underlying generative model. We illustrate our approach, and compare it to human perception, in the context of one such generative class: Gaussian dot-clusters. We find that numerical estimates derived from our model

closely match subjects' perceptual estimates of number of clusters. Thus our Bayesian approach to perceptual grouping, as one side-effect, effectively models the perception of cluster numerosity.

Gregory Francis
Purdue University

Publication bias, modeling, and theorizing in vision science

Vision science is not immune to the issues that are affecting areas of psychology, and an important part of the problems involves misunderstandings about data and theory. A common practice is to "let the data define the theory," but among modelers it is well known that quantitative models can over fit experimental data. Such practices promote developing a model that tries to explain random noise, a task that is essentially hopeless. In the other direction, researchers often claim that a theory predicts an experimental outcome; but without a quantitative model such predictions are unjustified. To help address the crisis in psychology, modelers should take a leading role in clarifying the important relationships between data and theory.

Stephen Grossberg
Boston University

How do object reference frames and motion vector decomposition emerge in laminar cortical circuits?

How do spatially disjoint and ambiguous local motion signals in multiple directions generate coherent and unambiguous representations of object motion? Various motion percepts, including those of Duncker and Johansson, obey a vector decomposition rule, whereby object parts are seen moving relative to a reference frame, whose direction is subtracted from the directions of the parts. A neural model predicts how vector decomposition results from interactions within and between the form and motion processing streams. These mechanisms solve the aperture problem, carry out figure-ground separation, group spatially disjoint moving object parts via illusory contours, and capture object motion direction signals on real and illusory contours. Inter-depth directional inhibition causes a vector decomposition whereby motion directions of a moving frame at a nearer depth suppress these directions at a farther depth and cause a *peak shift* in the perceived directions of object parts.

Bo Cao and Arash Yazdanbakhsh
Boston University

A neural model of MSTd neurons for eye movement compensation

Neurons in the dorsal part of monkey medial superior temporal area (MSTd) can respond to pursuit eye movement as well as visual motion patterns (Komatsu and Wurtz, 1988). Some of these neurons in MSTd are tuned to the on-screen speed and direction of moving stimuli, even when there is eye movement (Bradley et al., 1996; Inaba et al., 2010). It seems that these MSTd neurons are able to compensate the corresponding visual effect caused by the eye movement, so that the actual visual motion can be represented in brain. However, how these neurons integrate the eye movement and the visual motion information to achieve this compensation is still controversial. In this study, we develop and discuss a dynamical neural model to account for this compensation. This result indicates one possible principle of neurons' computation in integrating the signals from different channels.

Jihyun Kim
Purdue University

A biophysical model of retinal circuitry: spatiotemporal dynamics

Retina is comprised of several feedforward cell layers (photoreceptors, bipolar and ganglion cells) that are interleaved by inhibitory cell layers (horiozontal and amacrine cells). While neurophysiological evidence suggests that a complex synaptic structure underlies in this network, its functional consequence has not been evaluated. Here, a biophysical retinal model is designed to investigate the spatiotemporal dynamics of the retinal processing. The model simulations demonstrate that kinetic (response time course and gain) and spatial (receptive-field) properties of horizontal cells and M and P retinal ganglion cells interact to differently modulate M and P cell response gains over low-to-high spatial-frequency range. This produces not only different spatial-frequency sensitivity characteristics between M and P cells, but also spatial-frequency dependent time course of information processing. The model predicts psychophysical data on visual persistence and afterimage, each demonstrating short-term and long-term effect of retinal processing on perception.

Oliver Layton and Arash Yazdanbakhsh
Boston University

A multi-scale model of figure-ground segregation in the primate visual system

Local analysis of shape curvature is ambiguous with respect to figure-ground assignment because contours may constitute convex parts of the figure or concave parts of the background or vice versa. Determining whether a local region within a visual scene belongs to the interior or exterior of a shape requires analysis across multiple spatial scales. Neurons in area V4 selectively respond to shapes defined by convex and concave contours, despite their size and local contour curvature and acuteness (Pasupathy & Connor, 1999). We present a dynamical model of primate V4 with convex receptive field units that elicit activity to the interior of figures (e.g. those with convexities, concavities, and holes) at the appropriate size, but not to their exterior. Model V4 cells perform a nonlinear multi-scale integration of subunits that vary in retinotopic location and receptive field size, and feedback to units that are selective to figures of a single size.

Michele Rucci, Jonathan Victor and Murat Aytekin
Boston University

Seeing space through time during natural fixation

We have recently shown (Kuang et al., 2012) that continually occurring microscopic eye movements remove predictable correlations in natural scenes, a transformation, previously attributed to center-surround receptive field organization. Here we extend this work by examining the characteristics of the retinal input during natural head-free fixation. We show that head and eye movements combine to transform spatial patterns of luminance into temporal modulations in a way that counterbalances the spectral distribution of natural images and whitens the retinal image over a wide range of spatial frequencies. These results further support the proposal that fixational behavior is a critical component of the encoding of visual information.

Emre Akbas and Miguel Eckstein
University of California, Santa Barbara

Object detection with a foveated visual system

We present a foveated object detector (FOD) as a biologically-inspired alternative to the sliding window (SW) approach, the dominant paradigm of visual search in computer vision. Similar to the human visual system, the FOD has higher resolution at the fovea and lower resolution at the periphery. Consequently, more computational resources are allocated at the fovea and relatively fewer at the periphery. FOD makes saccades to align its fovea with regions of interest in the input image and integrates observations through multiple fixations. Our approach combines state of the art object detectors from computer vision [Felzenszwalb et al.,TPAMI 2010] with a recent model of peripheral pooling regions found at the V1 layer of the human visual system [Freeman & Simoncelli, Nature Neuroscience 2011]. We assessed various eye movement strategies and show that the FOD gives similar detection accuracy to the SW approach while bringing significant computation cost savings.

Michael E. Rudd
University of Washington

Perceptual weighting of local contrast information in an edge integration model of neural lightness computation

Edge integration theory posits that lightness is computed by a neural process that first extracts information about local oriented contrast (generalized "edges") then spatially integrates this information to construct a cortical representation of surface reflectance. The full process entails several steps occurring in the following order: edge classification -> weight setting –> edge integration –> anchoring –> lightness. I have previously proposed that in human vision edge weights are influenced by attention (Rudd, J Vision, 2010). Here I show how lightness judgments in disk-annulus, Gilchrist dome, and Gelb displays are together explained by an edge integration model that incorporates three additional principles governing perceptual edge weightings. First, weights decline with distance from the surface whose lightness is computed. Second, larger weights are associated with edges whose dark sides point towards the surface. Third, edge integration is carried out only along paths leading from a common "ground" region to the surface.

Ruth Rosenholtz
MIT

What is our baseline model of visual processing?

Here is a composite, cartoon view of vision: Our visual systems process one thing at a time. Resources focus on, and are local to, that thing. The visual system switches modes to process textures or scenes instead of “things”. Difficulties arise when multiple items compete for resources. Call this view, “one thing, right here, right now”. Recent evidence indicates: (1) The visual system performs “texture processing” in peripheral vision, even during an object recognition task. (2) Some of the logic behind “one thing, right here, right now” relied on faulty assumptions. (3) A wide range of stimuli can be well represented by a rich set of local summary statistics. Such a representation obviates the need for mode switching in human vision. (4) For many tasks, it is not optimal to limit processing to the object location. These arguments suggest a different baseline model of vision: “stuff, everywhere, all the time.”

Holly Gerhard, Felix Wichmann and Matthias Bethge
Werner Reichardt Centre for Integrative Neuroscience, Bernstein Center for Computational Neuroscience, Tuebingen

How Sensitive Is the Human Visual System to the Local Statistics of Natural Images?

Several physiological links between natural image regularities and visual representation have been made using probabilistic natural image models. However, such models had not yet been linked directly to perceptual sensitivity. Here we present results from a new test of model efficacy based on perceptual discrimination. Observers viewed two sets of image samples on every trial: one set of natural images, the other set matched in joint probability under a natural image model (generated by shuffling the natural set's content subject to model assumptions). Task: which set contains true natural images? We tested 8 models from one capturing only 2nd-order correlations to one among the current state-of-the-art in capturing higher-order correlations. Discrimination performance was accurately predicted by model likelihood, an information theoretic measure of model efficacy, and was overall quite high indicating that the visual system's sensitivity to higher-order regularities in natural images far surpasses that of any current image model.

Christopher Tyler
Smith-Kettlewell Eye Research Institute

Modeling ‘Binoculomotor’ Dynamics

The dynamics of the binocular oculomotor (‘binoculomotor’) coordination of saccades are an important index of the underlying oculomotor physiology and its disorders. Saccadic dynamics may be characterized by the four parameters of onset latency, duration, peak velocity and temporal asymmetry. Across a non-academic population of 100 participants, each parameter distribution was well-fitted by a Gaussian distribution function, with a coefficient of variation of ~0.5 (e.g., 150-350 ms for onset latency. The individual variabilities, on the other hand, were not well fit by Gaussians but required asymmetric gamma distribution functions. The interocular difference distributions were again Gaussian, but much tighter, of the order of 1% of the width of the population distributions (e.g., from 0.5-2 ms in onset latency sigmas), implying highly characteristic concatenations of binocular coordination dynamics for each participant, at ~1 billionth of the volume of the population variability parameter space.

Harald Ruda and Ennio Mingolla
CVL, Northeastern University

A model of depth ordering from motion occlusion and disocclusion in textured scenes

The Form-Motion-Occlusion (FMO) model of Barnes & Mingolla (2013) describes how the primate visual system can determine figure-ground relationships between static and moving surfaces. The model was tested on black-and-white textures where one surface moves in front of or behind another. We are extending the FMO model to work with more general scenes, and describe the model and modifications necessary for working with grayscale imagery. We have chosen the Michotte rabbit-hole slit phenomenon as a key stimulus. This effect is complex and dynamic, since the perceived depth order of regions of a scene changes during a brief (e.g., 500 msec) sequence. In addition, we have developed a purely texture-based version of this phenomenon (http://neu.edu/cvl/demos/movies/michotte-texture.gif), which while

equally salient, is a challenging stimulus for model development. We demonstrate results using this stimulus.

Inna Tsirlin, Laurie Wilcox and Robert Allison
York University, Toronto

A biologically-plausible computational model of depth from monocular occlusions and binocular disparity.

Occlusion of one object by another gives rise to monocular occlusions - areas visible only in one eye. Although binocular disparities cannot be computed for these areas, they can be precisely localized in depth and can induce the perception of illusory occluding surfaces. Psychophysical experiments suggest that the visual system assigns depth from monocular occlusions using the constraints imposed by occlusion geometry. However, none of the existing models of stereopsis use it as the primary mechanism for depth computation in occluded areas and none were shown to recover depth in illusory occluders. We propose a model of depth perception from disparity and monocular occlusions in which monocular areas are detected explicitly and depth from occlusions is calculated based on occlusion geometry. It successfully reconstructs depth in a large range of stimuli demonstrating that a dedicated set of occlusion detectors combined with classical disparity detectors can underpin many stereoscopic percepts.

Ram P. Sharma[1], Donna Strickland[2], and Melanie C. W. Campbell[2]
[1]Indian Institute of Technology, [2]University of Waterloo

Nonlinear Laser pulse induced plasma effects on the performance of crystalline lens.

Age-related loss of the ability of the human crystalline lens and eye to focus clearly on objects close to the eye is known as presbyopia. Restoring the elasticity of the human crystalline lens is one of the most challenging problem in vision research. Use of moderate power femtosecond laser micro-machining has now been successfully implemented as a surgical tool for cutting a corneal flap for laser-assisted in situ keratomileusis (LASIK) surgery. Ophthalmologists are quite hopeful for micro-machining the crystalline lens of the eye. Microscopic caviation bubbles are expected to be formed due to the laser pulse and lens interaction. These cavitation bubbles result in an increase of elasticity of the lens and can help in curing the presbyopia. But nobody properly understands this complex nonlinear process. Crystalline lens is having gradient index (GRIN) structure. Due to laser pulse, nonlinear effects are introduced in the lens on account of nonlinear properties of lens material and plasma generation. Therfore, the laser pulse now propagates in a nonlinear media and can breakup into small filamentary structures. This may result the breaking of the bonds of the molecules of the proteins by which the lens is made of. To date, the underlying physical mechanism of cavitation bubble formation is not fully understood. In the present paper we have proposed a theoretical and computational model based on the GRIN and plasma nonlinearity. Intensity and profile of the filaments have been calculated and relevance to the cavitation bubble formation pointed out.

Jian Ding, Stanley Klein and Dennis Levi
University of California, Berkeley

Modeling studies on binocular combination

An elaborated model (the DSKL model, Ding, Klein and Levi, JOV 2013) modified from Ding and Sperling (PNAS, 2006), was proposed to account for binocular combination when the two eyes' inputs varied over a large range of contrast and phase differences. The perceived phase of a cyclopean sinewave with information about the interocular contrast ratio just before the binocular combination site provides strong constraints on the model, and the perceived contrast plus contrast discrimination provides further constraints. Both the DSKL and Ding-Sperling models predict the linear summation at low contrast and the constant-contrast perception at high contrast in binocular combination. However, by adding interocular gain-enhancement to the Ding-Sperling model that only includes interocular gain-control, the DSKL model significantly improved the data fitting for both normal and abnormal visual systems. A motor/sensory fusion mechanism is also included in the DSKL model to account for the contrast-dependent phase-dependence of binocular contrast combination.

Stanley Klein[1], Thom Carney[1], Austin Roorda[1], Lawrence Sincich[2], Claudio Privitera[1], Ram Sabesan[1] and Gene Switkes[3]
[1]University of California, Berkeley, [2]University of Alabama at Birmingham, [3]University of California, Santa Cruz

Modeling color perception from single and multiple cone stimulation

Hofer, Singer & Williams (2005) reported a wide diversity of color percepts upon stimulating single same pigmented cones. A new Adaptive Optics Scanning Laser Ophthalmoscope (AOSLO) with stabilization capability enables one to stimulate the same individual cones month after month (Sincich, Zhang, Tiruveedhula, Horton, Roorda, 2009). We are using this new technology to test several color-naming hypotheses: 1) Centers of single, near foveal ON parvo LGN cells get input from multiple cones (Sincich et al. 2009). 2) Local surround effect. 3) Bayesian inference based on natural scene statistics (Brainard, Williams, Hofer, 2008). Disambiguating these hypotheses requires multicone stimulation experiments plus serious computational modeling of cone-cone nonlinearities. We have gathered color naming data to demonstrate that repeatable single cone color naming is feasible, but we have several hurdles before the hypotheses are testable. Other modeling opportunities made available by AOSLO will be discussed.

Tadamasa Sawada
Ohio State University

A geodesic on a non-differentiable part of a 3D surface and its role in visual perception

A geodesic on a 3D surface is defined as a path which is locally the shortest. Any pair of neighboring points on a surface can be connected by a geodesic with the shortest length. It was pointed out that geodesic contours on a 3D surface induce 3D perception of the surface. Note that the role of geodesics in visual perception has been studied with 3D surfaces which are twice-differentiable everywhere. However, there are surfaces with sharp non-differentiable parts (e.g. angular edges and pointed apices) and we perceive them to be sharp as they are. In this study, I generalize the concept of geodesics to non-differentiable parts of 3D surfaces and discuss the relevance of geodesics in visual perception.

Ilan Kadar and Ohad Ben-Shahar
Ben-Gurion University of the Negev

From Perceptual Relations to Scene Gist Recognition

The ability to recognize visual scenes rapidly and accurately is highly constructive for both biological and machine vision. Following the seminal demonstrations of the ability of humans to recognize scenes in a fraction of a second much research has been devoted to understanding its underlying visual process, as well as its computational modeling. In this work we focus on one aspect of the scene recognition process and investigate whether prior knowledge about the perceptual relations between the different scene categories may help facilitate better scene recognition. We first describe a psychophysical paradigm that probes human scene recognition, extracts perceptual relations between scene categories, and suggests that these relations do not always conform the semantic structure between categories. We then incorporate the obtained perceptual relations into a computational classification scheme, which takes inter-class relationships into account. Finally, we demonstrate how our approach provides significant improvements on various scene recognition operations.