ABSTRACTS – 2016 AIC

Wyatt Bair

University of Washington

A binocular model for motion integration in MT neurons

Visual neuroscience is a field of great specialization, where submodalities such as motion, depth, form and color are often studied in isolation. One disadvantage of this isolation is that results from each subfield are not integrated to constrain common underlying neural circuitry. Yet, to deeply understand the cortical computations underlying visual perception, it is important to unify our fragmentary models to reach a critical mass of constraints so that robust and powerful circuits can emerge. I will discuss our efforts to unify models of direction selectivity, binocular disparity and motion-in-depth (MID, or 3D motion) to reveal circuits and to understand computations from V1 to area MT and beyond. In spite of the great attention given to area MT in terms of its role in motion perception, past efforts to model this area have largely overlooked the binocular integration of motion signals. Recent electrophysiological studies have tested binocular integration in MT and found surprisingly that (i) MT neurons lose their hallmark pattern motion selectivity when stimuli are presented dichoptically and (ii) many MT neurons are selective for motion-in-depth (MID), contrary to the prevailing view. By unifying these novel observations with insights from monocular, frontoparallel motion studies concurrently in a binocular MT motion model, we have generated clear, testable predictions about the circuitry and mechanisms underlying visual motion processing. We built binocular models in which signals from left- and right-eye streams could be integrated at various stages from V1 to MT and attempted to create the simplest plausible circuits that accounted for the physiological data. Our successful models make predictions about the existence and the order of critical operations along the pathway leading to MT, and hint at unexpected relationships between pattern- and 3D-motion processing. I will present our results and insights and discuss the challenges of trying to unify constraints across visual sub-modalities to build more robust and comprehensive models of early- and mid-level visual processing.

Dirk Bernhardt-Walther

University of Toronto

Contour junctions underlie neural representations of scene categories in human v

Authors: Dirk Bernhardt-Walther and Heeyoung Choo

Humans can quickly recognize a novel real-world environment according to its basic-level scene category. Which visual features facilitate such efficient categorical scene perception in humans? To answer this question we combine multi-voxel pattern analysis of neural representations of scene categories throughout visual cortex with computational models of scene categorization. Participants viewed line drawings of six scene categories (beaches, forests, mountains, city streets, highways, and offices) while in an MRI scanner. We decoded scene categories from locally distributed neural activity patterns. We compared decoding error patterns to the error patterns from five computational models of scenes, each relying on the statistics of only one kind of contour property: orientation, length, curvature, junction types, or junction angles. We found that statistics of junction properties exhibited the largest contribution to the neural representations of scene categories in high-level scene-selective brain regions (parahippocampal place areas, PPA; occipital place area, OPA; lateral occipital complex, LOC), followed by orientation statistics (see figure). To assess the causal involvement of these visual properties in neurally representing scene categories, we manipulated the images in such a way that junctions or orientations were disrupted. In early visual cortex, scene categories could be successfully decoded under both manipulations. In the PPA, OPA and LOC, on the other hand, disruption of junction statistics severely reduced the extent of category-specific neural activity. When orientations were disrupted, scene categories could still be decoded successfully in the PPA and LOC. Based on these results we suggest a causal role for contour junctions, which provide cues for the 3D arrangement of surfaces in a scene, in the neural mechanisms of scene categorization.

Vincent Billock

The Ohio State University

Visual Amplification Via Sensory Integration in Rattlesnakes, Cats and Humans

Sensory integration and sensory binding are similar problems separated by a vast methodological gulf. The dominant paradigm of binding theory is neural synchronization, while sensory integration is built on observations of bimodal neurons. These cells show large increases in firing rates for bimodal presentation of weak stimuli, but little improvement for strong stimuli, a finding known as the Principle of Inverse Enhancement. It would be useful to link these two fields so that methods from each could be used by the other. The best case for such a bridge is the rattlesnake, which has two dissimilar visual systems, one for light and one for heat. Although this sounds like a binding problem, the rattlesnake has been studied using the methods of sensory integration. Many cells in rattlesnake optic tectum are sensitive only to light but can be strongly modulated by heat stimuli, or vice versa. I simulated these cells by assuming that they are members of synchronized pairs of excitatory-coupled rate-coded neurons. The same synchronized neuron model, without any parameter changes, accounts for a population of cells in cat visual cortex whose firing rates are enhanced by auditory stimuli consistent with the Principle of Inverse Enhancement. It is intriguing that the most important principle in sensory integration can be derived from binding theory. The same mechanism can be used to model within-vision nonlinear perceptual amplifications, such as those seen in binocular vision and color vision.

Alyssa Brewer

University of California Irvine

Rod scotoma fMRI elucidates cortical rod pathways and lesion measurement concerns

Are silencing, ectopic shifts, and receptive field scaling in cortical scotoma projection zones (SPZs) the result of long-term reorganization (plasticity) or short-term adaptation? Electrophysiological studies of SPZs following retinal lesions in animal models remain controversial, because they are unable to conclusively answer this question due to limitations of the methodology. Here we used functional MRI visual field mapping via population receptive field (pRF) modeling with moving bar stimuli under photopic and scotopic conditions to measure the effects of the rod scotoma in human early visual cortex. As a naturally-occurring central scotoma, it has a large cortical representation, is free of traumatic lesion complications, is completely reversible, and has not reorganized under normal conditions (but can, as seen in rod monochromats). We found that the pRFs overlapping the SPZ in V1, V2, V3, hV4, and VO-1 generally (1) reduced their BOLD signal coherence, and (2) shift ed their pRFs more eccentric, but (3) scaled their pRF sizes in variable ways. Thus, silencing, ectopic shifts, and pRF scaling in SPZs are not unique identifiers of cortical reorganization; rather, they can be the expected result of short-term adaptation. However, are there differences between rod and cone signals in V1, V2, V3, hV4, and VO-1? We did not find differences for all five maps in more peripheral eccentricities, outside of rod scotoma influence, neither in coherence, in eccentricity representation, nor pRF size. Thus, rod and cone signals appear to be processed similarly in cortex.

==

Bruce Brown

Brigham Young University

Whole-wave EEG analysis in the identification of neuropsychiatric illnesses
Authors: Bruce Brown, Dawson Hedges, Jack Silcox

An eigenvector-based approach is used to extract unique ideographic cognitive components for individual subjects from ERP data gathered during a cognitive task. These cognitive components appear to have impressive personal identification and diagnostic capabilities. For example, females are found to have a substantially different Sternberg memory search process than males ($F[1,33] = 579.84$, $p<.0001$, Rsquare=.946). In similar studies, the discrimination between clinically depressed males and healthy controls is strong ($F[1,318]= 1802.31$, $p<.0001$, Rsquare=.850), as is the discrimination between OCD males and healthy controls ($F[1,318]= 1939.76$, $p<.0001$, Rsquare=.916). In the

neurological area cognitive components discriminated well between mild AD subjects and healthy controls (Wilks"lambda = .4297, p<.0001, Rsquare=.5703), even when using a weak cognitive task (auditory oddball). Additional studies are in progress.
==

Patricia Cheng

University of California, Los Angeles

Analytic knowledge of causal invariance in rats
Authors: Julia Schroeder, Aaron Blaisdell, Rui He, Jeffrey Bye, Patricia Cheng

Abstract:

none yet

==

Lawrence Cormack

The University of Texas at Austin

The perception of depth vs. frontoparallel motion probed with motor tracking.

From traditional psychophysics, we know that stereoscopic acuity is exquisite, yet under some conditions depth motion is more difficult to see than frontoparallel motion, this despite the ecological importance of detecting approaching objects. Being fond of more naturalistic tasks, we wondered how observers would fare if asked to dynamically track objects moving in three dimensions. In the main experiment, subjects pointed with their index finger to track the center of a stereoscopic target as it moved in a 3-dimensional Brownian random walk. We measured tracking performance by calculating the cross-correlograms for each of the cardinal motion axes. Tracking performance was selectively impaired for depth motion compared to horizontal and vertical (frontoparallel) motion, by which we mean that the peak correlation between stimulus and response was both lower and occurred after a longer time delay. Moreover, this impairment was greater than was expected from the small relative size of the retinal signals arising from depth motion. Further experiments ruled out motor explanations (e.g., the notion that moving the hand along a depth axis is more difficult than moving it within a frontoparallel plane). In another experiment, observers tracked a disparity-defined target in a dynamic random element stereogram as it moved in a 3-dimensional Brownian random walk. When frontoparallel tracking required a disparity computation, it showed the same response delay observed in the original 3D tracking experiment – a fact we share while remaining agnostic about the order (in the Lu and Sperling sense) of stereomotion processing. Thus, despite the crucial importance of egocentric depth motion, its perception and the motor responses thereto are actually impaired relative to that of frontoparallel motion.
==

Sven Dickinson

University of Toronto

Model-Based Perceptual Grouping and Shape Abstraction
Authors: Sven Dickinson and Pablo Sala

For many object classes, shape is the most generic feature for object categorization. However, when a strong shape prior, i.e., a target object, is not available, domain independent, mid-level shape priors must play a critical role in not only grouping causally related features, but regularizing or abstracting them to yield higher-order shape features that support object categorization. In this talk, I will present a framework in which mid-level shape priors take the form of a vocabulary of simple, user-defined 2-D part models. From the vocabulary, we learn to not only group oversegmented regions into parts, but to abstract the shapes of the region groups, yielding a set of abstract part hypotheses. However, the process of shape abstraction can be thought of as a form of "controlled hallucination",

which comes at the cost of many competing 2-D part hypotheses. To improve part hypothesis precision, we assume that the 2-D parts represent the component faces of aspects that model a vocabulary of 3-D part models. We then exploit the relational structure (spatial context) of the faces encoded in the aspects, and again formulate hypothesis selection in a graph-theoretic, probabilistic framework. Finally, we introduce a technique that is able to recover the pose and shape of a volumetric part from a recovered aspect, yielding a framework that revisits the classical problem of recovering a set of qualitative 3-D volumetric parts from a single 2-D image.

Barbara Dosher

University of California, Irvine

Perceptual Learning

==

Frank Durgin

Swarthmore College

The possible role of angular expansion in the misperception of large-scale space
Authors: Frank Durgin and Zhi Li

It has long been observed that perceptions of distances and slant in large-scale spaces (e.g., those affording locomotion) are typically distorted: Hills look much steeper than they are -- even when looking downhill; distances along the ground are foreshortened. Based on a series of studies of perceived visual direction in pitch and yaw, on parametric examination of slant estimates for large and small surfaces at near and far distances, and on a variety of perceptual measures of perceived distance and direction in large-scale space, we propose that the exaggeration of perceived angular deviations from horizontal/straight-ahead (typically with a gain of 1.5 in pitch) can parsimoniously account for a great deal of new and existing data on perceptual bias - including the large-scale horizontal-vertical illusion. We speculate that these angular exaggerations must have a functional basis and propose that orientation biases may reflect a coding scheme for retaining precision at the cost of accuracy. Such a trade-off could aid in calibrated action control, because the calibration of action control is limited by the precision of feedback guiding calibration processes rather than by the accuracy of estimation.

==

James Elder

York University

The Southampton-York Natural Scenes (SYNS) Dataset
Authors: James Elder, Wendy Adams, Erich Graf, Alex Muryy, Arthur Lugitigheid

The inference of 3D structure from 2D images is a central function of biological and machine vision systems. Since the problem is ill-posed, optimal inference depends upon knowing the joint statistics of 3D surfaces and 2D images. We have developed a new public dataset (syns.soton.ac.uk) that can be used to ground theories of human visual processing and machine vision algorithms in the ecological statistics underlying the problem.

To fairly represent the diverse visual environments we experience, we randomly sampled scenes from 19 outdoor and 6 indoor categories across Hampshire, UK. Outdoor categories, identified by the UK Land Use dataset, include cropland, coastal dunes, woodlands, industrial estates, wetlands, residential areas, farms and orchards. Indoor categories include residential, theatres, cafes and offices. Each scene is represented by three types of co-registered data: (i) 3D point clouds spanning 360° x 135° captured by a laser rangefinder (LiDAR), (ii) High dynamic range images spanning 360° x 180° captured by a spherical camera and (iii) 18 Stereo image pairs, each spanning 35° x 24° and tiling a 360° horizontal panorama, captured by a custom-built high-resolution stereo rig, with camera separation matched to average human interpupillary distance.

LiDAR data were analysed to determine the distribution of egocentric surface attitudes in outdoor environments. Surface normals were computed at each LiDAR point, using an adaptive scale selection method. Overall, the

distribution is dominated by the ground plane. To relate these natural scene statistics to human perception, we conducted psychophysical experiments to measure both discrimination and absolute judgments of the 3D orientation of real textured planar surface; together these allowed estimation of the observer's underlying prior. We found that priors varied substantially across observers, with an overall bias toward surface normals lying in the vertical meridian. We discuss potential reasons for the variability across observers and divergence between ecological statistics and human perceptual judgments.
==

Gregory Francis

Purdue University

A small part of the Human Brain Project: Neural dynamics of visual segmentation

The Human Brain Project is a European funded effort to build a "scaffold" model and simulation of the human brain by 2023. In its current phase the HBP is developing simulation and database technologies that will be used to achieve the long-term goal. Some of these technologies may be of interest to cognitive scientists, especially if they develop neural models of human behavior. I will briefly describe some of the relevant HBP technologies and share how they contributed to a neural network model of visual segmentation. I will also describe how the model explains complex effects of perceptual grouping for empirical data on visual crowding.

Wilson Geisler

University of Texas at Austin

Measurements and models of detection in natural scenes
Authors: Steve Sebastian, Jared Abrams, Wilson S. Geisler

An ultimate goal of vision science is to measure and predict performance in visual tasks under natural conditions. Perhaps the most fundamental task is to detect target objects in the natural backgrounds that surround us. It is known from experiments with simple stimuli that the specific properties of a background (luminance, contrast, similarity to the target) have a strong influence on detectability. It is also know from experiments with simple stimuli that the uncertainty created by randomizing the amplitude and/or location of the target ("target uncertainty"), and randomizing properties of the background ("background uncertainty") are additional factors influencing detectability. What is relatively unknown are how these known factors individually affect detection in natural scenes, and how these factors combine in affecting detection in natural scenes. We address these two questions using a direct experimental approach that is quite efficient and could be used to address similar questions for other natural tasks. A large collection of calibrated natural images is divided into millions of background patches that are sorted into narrow bins along dimensions of interest. In the present study, each bin corresponds to a particular narrow range of luminance, contrast, and target similarity. Detection performance is then measured in a sparse subset of bins spanning the entire space, with and without target and background uncertainty. We find that detection thresholds in natural backgrounds vary linearly along all three dimensions and that humans are remarkably unaffected by simultaneous background and amplitude uncertainty. We show that these results are predicted by a Bayesian signal-detection model (a generalized matched template model) derived from first principles.
==

Joseph Houpt

Wright State University

Revisiting Stereoscopic Disparity as a Feature in Visual Search
Authors: Joseph W. Houpt, Leslie M. Blaha, Megan B. Morris, John P. McIntire

With the recent explosion in the number of commercially available stereoscopic 3D displays, it has become financially feasible to use 3D in many operational environments. Almost certainly there will be a wrong way to exploit the technology in the sense that it leads to worse performance and outcomes than when employing standard 2D displays, but there is still a lot of potential for the technology to enhance capabilities. The focus of this talk will be on the use of stereo-3D displays for visual search tasks and, in particular, on the use of stereoscopic disparity as a target feature. I will begin with the standard demonstration of a pop-out effect for 3D targets and an extended analysis using distributional level measures. Next, I will discuss a follow-up study, in which we explored the perceptual processing of targets based on a combination of shape and stereoscopic disparity. Using Systems Factorial Technology, we found that participants are faster at identifying combined shape depth targets than would be predicted by independent, parallel search and that most participants employed a coactive strategy (i.e., they searched for the target based on pooled shape depth information rather than treating each source of information separately).

==

Xiaoping Hu

Emory University

The effect of prenatal alcohol exposure on brain connectivity in adolescents
Authors, Zhihao Lik Bing Ji, Claire Coles, Mary Elln Lynch, Xiaoping Hu

We performed resting state connectivity analysis and structural connectivity analysis in data from a sample of 72 prenatally alcohol exposed (defined as 13.3 oz absolute alcohol/week) individuals (age: 13±3; 37 male and 35 female) and 72 matched controls (age: 13.5±3; 40 male an 32 female). Our analysis successfully identified 7 functional networks (default mode, left frontal-parietal, right frontal-parietal, primary motor, primary visual, extrastriate visual, and salient) using independent analysis, and identified significant ($p<0.05$ corrected) reduction of functional connectivity in the exposed cohort all networks except the salient network. In addition, structural connectivities in the primary and extrastriate visual regions were examined. The results showed significantly decreased connectivity in the connections between the left and right primary visual cortices ($p<0.03$, Tukey HSD corrected), as well as between the left primary visual cortex and left extrastriate area ($p<0.019$, Tukey HSD corrected) in the PAE group. These results indicate that prenatal alcohol exposure leads to reduction in both functional and structural connectivities.

==

Alan Johnston

University of Nottingham

The harmonic vector average: a new approach to the aperture problem

A central problem for the visual system is how to compute the speed of objects as they move in the world. The motion processing system in monkey and man is arranged in a hierarchy of anatomically distinct brain areas containing neurones whose receptive fields increase in size, response selectivity and computational complexity with increase in rank. The evident variation across the hierarchy in the spatial range of motion analysis, as indicated by receptive field size, leads to some unavoidable computational problems. Initial local analysis is limited by the aperture problem, a result of which is that neurones typically signal motion orthogonal to contours. These essentially independent estimates need to be brought together through some computation that can deliver the global motion of an object as a whole.

We can study motion integration experimentally using an array of Gabor elements (Gaussian windowed moving sine gratings). The local speeds of object contours vary systematically with the cosine of the angle between the component of the local velocity orthogonal to the contour and the global object motion direction. A spatial distributed array of Gabor elements whose speed depends on local spatial orientation in accordance with this pattern can appear to cohere and move as a single surface. A number of models have been proposed to explain how the visual system might achieve this result. If we assume a single rigid translation, the global direction of motion can be found from at least two elements with different orientations, using a strategy known as the intersection of constraints (IOC). This strategy has usually been contrasted with the vector average. However, the vector average over local vectors that vary in direction always provides an underestimate of the true global speed, and if we have a biased set of local motions with respect to the global motion, the global percept is shifted towards the average direction, which is inconsistent with the IOC strategy. We need to look for an approach that will resolve these problems.

If we plot the normal component motion vectors in a velocity space they lie on a circle through the origin. This circle when inverted in the unit circle maps to a line, allowing a least square estimate of the IOC and an average

inverse velocity for a set of normal components. This average, inverted once more in the unit circle, is the harmonic vector average (HVA). The harmonic vector average provides the correct global speed and direction for an unbiased sample of local velocities with respect to the global motion direction. The HVA over biased samples provides an aggregate velocity estimate that can still be combined through an IOC computation to give an accurate estimate of the global velocity, which is not true of the vector average. Psychophysical results for a biased distribution of Gabor arrays show perceived direction and speed falls close to the intersection of constraints direction for Gabor arrays having a wide range of orientations but the IOC prediction fails as the mean orientation shifts away from the global motion direction and the orientation range narrows. In this case perceived velocity generally defaults to the harmonic vector average. Neither the IOC nor the HVA can account for human global motion perception in biased arrays, however the perceived direction of motion appears to be bounded by the IOC and the HVA.
==

Philip Kellman

University of California, Los Angeles

Title: Spatiotemporal Boundary Formation
Authors: Philip Kellman and Gennady Erlikhman

Spatiotemporal boundary formation (SBF) is the perception of illusory contours, global form, and global motion from spatially and temporally sparse transformations of texture elements. Because it produces complete contours from elements lying along as little as 3% of an object's boundaries, SBF may be the "most from the least" in perceptual organization.
In this talk, I consider recent progress in understanding how SBF works. Evidence suggests that local oriented edge fragments are somehow generated from discrete element changes. These fragments then connect to form continuous object boundaries through well-known contour interpolation processes. The mystery is how local oriented fragments are generated. Following formal proofs that local orientation could theoretically be derived from triplets of non-collinear element changes, we developed a paradigm for examining minimal conditions in SBF. We found that same display can appear as a single moving element along a sawtooth arrangement of dots or a larger oriented edge moving through the array, depending on timing. Experimental results indicating precise temporal constraints on SBF and the operation of edge formation in parallel across the visual field suggests that the local edge formation stage in SBF may depend on known spatiotemporal filter mechanisms (e.g., Adelson & Bergen, 1985; van Santen & Sperling, 1984). These “motion energy" filters are ordinarily studied with spatial orientation given unambiguously by luminance contrast, but our results suggest a duality whereby, when orientation is not specified by static information these motion filters are also spatiotemporal orientation detectors. A combination of known contour interpolation processes in middle vision and outputs of early spatiotemporal filters may explain the processes that produce shape and continuous boundaries in SBF.

Michael Landy

New York University

Criterion learning in static and dynamic environments
Authors: Michael S. Landy, Elyse H. Norton, Stephen M. Fleming, Nathaniel D. Daw

Humans often make decisions based on uncertain sensory information. Signal detection theory describes detection and discrimination decisions as a comparison of stimulus strength to a fixed decision criterion. How is the criterion set? Here, we examine how observers learn to set a decision criterion in an orientation-discrimination task under both static and dynamic conditions. To investigate mechanisms underlying trial-by-trial criterion placement we compared covert and overt discrimination tasks. In each task, stimuli were ellipses with principle orientations drawn from two categories: Gaussian distributions with different means and equal variance. In the covert-criterion task, observers categorized a displayed ellipse. In the overt-criterion task, on every trial observers adjusted the orientation of a line that served as the discrimination criterion for a subsequently presented ellipse. We compared performance to the ideal Bayesian model and several suboptimal models that varied in both computational and memory demands. Under static and dynamic conditions, we found that, in both tasks, observers used suboptimal learning rules. A model in which the recent history of past samples determines a belief about category means fit the data best for most

observers and on average. Our results reveal dynamic adjustment of discrimination criterion, even after prolonged training.
==

Zhong-Lin Lu

The Ohio State University

qPR: An adaptive partial report procedure based on Bayesian Inference
Authors: Zhong-Lin Lu, Jongsoo Baek, Luis Lesmes

Iconic memory is best assessed with the partial report procedure, in which an array of letters appears briefly on the screen and a post-stimulus cue directs the observer to report the identity of the cued letter(s). Typically 6-8 cue delays or 600-800 trials are tested to measure the sensory memory decay function. Here we develop a quick partial report or qPR procedure based on a Bayesian adaptive framework to directly estimate the parameters of the sensory memory decay function with much reduced testing time. The exponential decay function is characterized by three parameters with a joint probability distribution. Starting with a prior of the parameters, the method selects the stimulus to maximize the expected information gain in the next test trial. It then updates the probability distribution of the parameters based on the observer's response using Bayesian inference. The procedure is re-iterated until either the total number of trials or the precision of the parameter estimates reaches a certain criterion. Simulation studies showed that only 100 trials were necessary to reach an accuracy of ±2.5% correct and precision of 7.5%. A psychophysical validation experiment showed that estimates of the sensory memory decay function obtained with 100 qPR trials exhibited good precision (the half width of the 67% credible interval = 5.1%) and excellent agreement with those obtained with 1600 trials of the conventional procedure (mean RMSE = 5.7%). qPR relieves the data collection burden in characterizing sensory memory and makes it possible to assess sensory memory in clinical populations.

Laurence Maloney

New York University

Representing and distorting probability and probability density
Authors: Laurence T Maloney and Hang Zhang

The movement we execute is not always the movement we plan; what we see is not always what is there; events in the world may turn out other than expected. Bayesian decision theory prescribes how to act so as to maximize expected value despite uncertainty concerning the outcomes of our actions. While human performance is impressive in many decision tasks, it is not optimal (Maloney, 2002). Small, systematic deviations from optimality are potentially a valuable source of information concerning how humans distort probability and probability density (Maloney & Mamassian, Visual Neurosci., 2009; Ting et al, J. Neurosci., 2015; Wu et al, PNAS, 2009, J. Neurosci., 2011; Zhang et al, PLoS Comp. Biol., 2010; Zhang & Maloney, 2012; Zhang et al, Frontiers in Neurosci., 2015). Based on this recent experimental work, I'll outline an alternative to Bayesian decision theory based on more accurate characterizations of the human representation of probability and probability density and discuss why we distort probability and probability density as we do.
==

Jeff Mulligan

NASA Ames Research Center

Measuring and Modeling Shared Visual Attention
Authors: Jeff Mulligan and Patrick Gontar

Multi-person teams are sometimes responsible for critical tasks, such as flying an airliner. Here we present a method using gaze tracking data to assess shared visual attention, a term we use to describe the situation where team members are attending to a common set of elements in the environment. Gaze data are quantized with respect to a set of N areas of interest (AOIs); these are then used to construct a time series of N dimensional vectors, with each

vector component representing one of the AOIs, all set to 0 except for the component corresponding to the currently fixated AOI, which is set to 1. The resulting sequence of vectors can be averaged in time, with the result that each vector component represents the proportion of time that the corresponding AOI was fixated within the given time interval. We present two methods for comparing sequences of this sort, one based on computing the time- varying correlation of the averaged vectors, and another based on a chi-square test testing the hypothesis that the observed gaze proportions are drawn from identical probability distributions.

We have evaluated the method using synthetic data sets, in which the behavior was modeled as a series of activities, each of which was modeled as a first-order Markov process. By tabulating distributions for pairs of identical and disparate activities, we are able to perform a receiver operating characteristic (ROC) analysis, allowing us to choose appropriate criteria and estimate error rates. Using these criteria, we have applied the methods to data from airline crews, collected in a high-fidelity flight simulator (Gontar & Hoermann, 2014). We conclude by considering the problem of automatic (blind) discovery of activities, using methods developed for text analysis.

Anitha Pasupathy

University of Washington

Visual shape representation in the intermediate stages of the primate brain

Decades of research have yielded detailed models of visual form processing in the primary visual cortex (V1) of the primate. Beyond V1, however, our understanding is quite limited. Past studies have shown that in V4, an intermediate stage in the ventral visual pathway, neurons are sensitive to both the curvature of the bounding contour and to the luminance contrast of the stimulus surface relative to the background. These shape selective neurons also display position and size invariance within the confines of the receptive field, but we currently have no model of V4 neurons that can simultaneously achieve all of these stimulus preferences. To attain this elusive goal, we are pursuing two strategies. First, we are conducting paired neurophysiology and modelling experiments aimed at improving the most promising biologically motivated model of V4 form selectivity (the HMAX model), which emphasizes boundary orientation, ignores surface contrast and exhibits limited invariance. For example, we are comparing model units to V4 neurons in terms of whether their shape selectivity is maintained for stimuli defined by an outline alone versus stimuli defined by an outline and surface contrast. To achieve more realistic invariance properties, we are manipulating attributes of the model including normalization equations, design of low-level convolutional features and receptive field density. Second, in parallel to developing this explicitly biologically plausible model, we are also exploring the ability of high-performing artificial object recognition networks to achieve the selectivities and invariances observed in V4. In particular, we have identified and are analyzing the properties of V4-like units in a deep convolutional neural network (AlexNet) trained to recognize objects in NATURAL scenes. We hope to understand what architectures underlie V4-like selectivity, and ultimately what training regimes may be responsible for the emergence of the relevant representations. In my talk, I will present results that reveal how our current best V4 model matches up with V4 physiology, and I will describe plans for future experiments to address the challenges that lie ahead.

==

Misha Pavel

Northeastern University

Decomposing Liquid Intelligence

We will discuss issues arising when attempting to decompose tasks that are thought to require fluid intelligence into components such as update and inhibit. The notion of these basic components is challenged these executive function tasks are embedded in more real-life-like scenarios such as computer games. We will illustrate these issues using a subset of data from a large study designed to investigate whether fluid intelligence, as measured by tasks such as Ravens Progressive Matrices, can be improved by training the component cognitive functions and their combinations. We plan to discuss both theoretical and practical implications of these results and insights.

==

Zygmunt Pizlo

Purdue University

The role of 3D symmetry in figure-ground organization of real scenes.
Zygmunt Pizlo, Aaron Michaux and Vijai Jayadevan, Purdue University.

The first step in visual perception is determining the presence, the number and the location of objects in front of the observer. In human vision, this step is traditionally called "Figure-Ground Organization (FGO). According to common wisdom, FGO is solved by grouping (clustering) operations on the basis of the similarity of nearby retinal regions such as their similarity with respect to color, motion, texture, etc. The main challenge to this common wisdom derives from the fact that the retinal image always confounds information about the permanent characteristics of the 3D physical objects within a scene with the constantly varying viewing conditions, including illumination, distance, and occlusions. This confound has been an unsurmountable obstacle in formulating a theory of FGO that could produce, even in principle, anything like the level of performance we all achieve in our everyday life. I will describe our recent attempt to develop such a theory. In this theory, the symmetry of objects is the key concept. Symmetry can do the job because (i) *all natura*l *objects* are characterized by one or more types of 3D symmetry, and (ii) a *3D configuration of unrelated objects* is, itself, almost never symmetrical. It follows that detecting 3D symmetries in a real 3D scene will lead to nearly perfect FGO.

Nicholas Port

Indiana University School of Optometry

Ocular-motor Performance of IU Athletes 2.0: A Nefarious Slope
Authors: Nicholas Port,Steve Hitzeman, Kacie Monroe, Melissa Elrod-Schmidt
Tina

It has been suggested that sports ability relates to ocular motor performance. We set out to directly test this hypothesis among ~1400 Indiana University athletes across all 24 sports. Baseline ocular motor data was collected on the first day of training camp just prior to the beginning of the freshman year. Ocular motor tasks included two smooth pursuit tasks (sinusoidal and step-ramp), one self-paced saccade task, and one fixation task (with and without a whole-field motion distractor). Over the ensuing 4 to 7 years of their collegiate athletic careers, we then collected longitudinal data relating to each subject's athletic performance. Seven or more years was sometimes needed in order to obtain a sufficient sample size (e.g., 20 subjects) for sports with small team sizes (e.g., golf and tennis). Large differences were found between sports, with an overall trend for ball sport athletes to have ocular motor performance profiles that differ from non-ball sport athletes e.g. swimming and cross-country). We also found some ocular-motor variables correlate with athletic performance in some sports (e.g. football). Our results, therefore, support the idea that athletic performance is correlated with some smooth pursuit and saccadic ocular motor variables. Additional research is needed to ascertain any causal connections, however.

Jenny Read

Newcastle University

Mantis stereopsis in complex scenes
Authors: Vivek Nityananda, Ghaith Tarawneh, Jenny Read

Praying mantises are the only non-vertebrate known to possess stereoscopic 3D vision. Yet, very little is known about the capabilities of insect 3D. To date, they have been tested only in very simple scenes containing one or two target objects. Last year at AIC, I presented our lab's technique for displaying 3D stimuli to mantids using blue/green anaglyph glasses. This enables us to probe mantids' response to arbitrarily complex 3D scenes. We have shown that (i) mantids can successfully identify target disparity even in a complex scene containing many moving objects; (ii) mantids can identify disparity of a target object even when that object is perfectly camouflaged on its background in any given frame, and revealed only by its motion. However, we have not yet found evidence that mantids can use disparity to break camouflage when an object is camouflaged even in the motion domain and

revealed only by its disparity. Investigating this in further detail will be crucial to determining the differences between insect and vertebrate stereopsis.
==

Bas Rokers

University of Wisconsin

The structural consequences of abnormal binocular input in child- and adulthood

Introduction
Complete blindness in one eye is associated with reduced volume of optic radiation and increased diffusivity in visual white matter regions of the human brain (Park et al., 2009). The effects of abnormal, rather than absent, visual input are less established. In amblyopia, which is thought to involve suppression of visually-evoked responses, at least one study suggests underdeveloped optic radiation (Ming-xia et al., 2007). In this study, we were interested in the consequences of abnormal visual input during development as well as during adulthood on the white matter integrity of thalamo-cortical projections.

Methods
We obtained diffusion-weighted MRI images (2 mm isotropic resolution; 32 diffusion directions; b0 = 1000) in 10 normally-sighted participants, 10 amblyopes, and 10 patients with monocular glaucoma. We used constrained spherical deconvolution and probabilistic tractography (Tournier et al., 2012) to estimate visual pathways from the optic chiasm, to extra-striate cortex.

Results
Participants with amblyopia as well as participants with glaucoma exhibited significantly reduced white matter integrity in the optic radiation. We subsequently related the structural deficits to the extent of behavioral impairment.

Discussion
Our results suggest that abnormal visual input during childhood, as well as adulthood, affects the white matter integrity of thalamo-cortical pathways. These results help provide valuable insight into the development and plasticity of the thalamo-cortical projections in the visual system.
==

Constantin Rothkopf

Technical University Darmstadt

What is the cost of a glance?

Perceptual tasks involve inherent costs and benefits that are usually manipulated explicitly by the experimenter. Even in tasks, in which we ask subjects to judge, which of two stimuli was brighter, to classify a sequence of stimuli, or to search for a target stimulus among distractors, such costs underlie subjects" behavior. Considering active gaze, what is the cost of a glance?

In the first set of experiments subjects had to detect an event at one of two spatial locations in a classic laboratory task. The observed behavior is not well explained by an ideal observer model without involving internal costs for switching gaze. Instead, a Bayesian ideal learner model that takes into account the time of processing and the specific uncertainty that subjects have about the elapsed time can capture key aspects of the observed behavior. Crucially, this model allows quantifying, how much detection probability subjects were willing to give up, to save one additional glance.

In the second set of experiments subjects carried out an extended sequential visuomotor task, navigating a sidewalk. In such naturalistic every day tasks it is not clear, what the costs imposed by the experimenter are. We show that it is possible to infer these inherent costs underlying subjects" behavior from their walking trajectories using inverse optimal control on a subject by subject and trial to trial basis. This methodology allows recovering the costs that best

explain the sequence of actions observed in navigation. Relating these inferred costs to subjects" eye movements during navigation allows calculating how much a single glance at an object relates to the inherent costs and benefits.
==

Christopher Rozell

Georgia Institute of Technology

Learning manifold transport operators of 3D transformations from 2D imagery
Authors: Christopher Rozell and Marissa Norko

An important task of the visual system is to use two-dimensional retinal images to make inferences about the three-dimensional world that are useful for interacting with the environment. It has been a topic of vigorous debate whether or not the visual pathway uses an explicit internal 3D model. While there is some support for the existence of a 3D model in visual perception, there have been few attempts at creating a computational model that can explain the ability to learn and perform inference about the 3d world from 2D imagery. In this talk I will describe our recent work building such a model using machine learning tools based on a generative model of transformations with a manifold structure. The central idea is that object transformations (e.g., rotation) in three dimensions form a manifold in the original 3D space and the projected 2D observation space contains an embedding of that manifold. Using this model, we have been able to learn the structure of transformation manifolds in 3D object space from unsupervised 2D imagery (i.e., the model has no knowledge of the specific transformation being observed). This learning produces a transport operator that replicates the action of the transformation in 3D (i.e., movement along a manifold), and can be used to perform inference or apply that transform to novel data. This work represents the foundation for a possible model of the 3D world that explicitly captures the action of identity preserving transformations and is learnable from 2D imagery.
==

Michele Rucci

Boston University

Beyond the sensory-motor loop: the inseparable link between vision and action

The visual system is normally thought to operate within a sensory-motor loop: visual processes guide eye movements, which in turn determine the next stimulus on the retina. In this talk, I will argue that this common framework is in many regards misleading, as the distinction between acquisition and processing of visual information is not clear cut. I will focus on fixational eye movements, the small eye movements that continually occur in the periods between "voluntary" relocations of gaze, and argue that this incessant motion of the eye is a critical information processing stage: a computational element of an active sensorimotor strategy by which the visual system processes spatial information in the temporal domain. I will review recent experimental and theoretical findings to address three main questions: (1) How is spatial information encoded in the modulations of luminance resulting from eye movements? (2) How is this information extracted and interpreted? (3) Can this stage of processing be tuned to the task by controlling fixational eye movements? The proposal that the visual system actively represents space through time implies that eye movements contribute to fundamental properties of spatial vision that are, at present, solely attributed to neural mechanisms.
==

Michael Rudd

University of Washington School of Medicine

Cortical computational model explains lightness contrast and assimilation

In a series of recent papers, I have developed an object-centered model of cortical lightness computation (Rudd, 2010, 2013, 2014) in which local directed steps in log luminance are first encoded by oriented spatial filters in early visual cortex, then the filter outputs are appropriately integrated along image paths directed towards the target at a subsequent cortical stage. A contrast gain control mechanism adjusts each filter's gain on the basis of the outputs of

other nearby filters. Here, I present computer simulations of this model to account for the Phantom Illusion (Galmonte, Soranzo, Rudd, & Agostini, submitted), a new luminance gradient illusion in which either an incremental or decremental target surrounded by a shallow gradient can be made to appear as an increment or decrement, depending on the gradient width. For wide gradients, incremental targets appear as increments and ecremental targets appear as decrements. For narrow gradients, the reverse is true. I discuss these results in the context of earlier modeling results that explain similar contrast and assimilation in other types of displays. To explain the entire set of results with a unitary theory highly constrains the model parameters. The required constraints imply that the lightness computations must be object-centered and thus midlevel. This conclusion is consistent with the model hypotheses that long-range contrast integration occurs in or beyond cortical area V4, following midlevel cortical computations related to image segmentation (completion, border ownership) in V2.
===

Keith Schneider

York University

Revisiting the discrete perception hypothesis

Early in the 20th century, when motion pictures were new, scientists and philosophers noted that, when people watched a movie, it was not subjectively apparent that the movie was actually a series of static snapshots. They wondered whether perhaps our perception always worked like this, with our impressions reconstructed from discrete samples of the world. Although there was no success in connecting such discrete sampling with any physiological mechanism, such as the alpha rhythm, numerous psychophysical experiments in the 1950â€™s and 60â€™s supported the idea of discrete perception. Later, in the late 1960s and early 70s, several experiments contradicted a narrow conceptualization of discrete perception, and the hypothesis largely vanished from the literature until the discrete wagon wheel illusion was identified in 1996. Here, I will briefly review the successes and failures of the discrete perception hypothesis, examine the range of parameters that are permitted by the experimental evidence, and test a simple model with flash-lag effect and temporal order judgment data. I will show that discrete perception with a variable frame rate, with motion as a primary rather than strictly inferred property, and occurring at the level of object recognition, remains a viable hypothesis.

Thomas Serre

Brown University

Towards a unified model of classical and extra-classical receptive field computations across visual modalities.

Authors: D.A. Mély, J.-K. Kim, J. Zhang and T. Serre

One of the major goals in visual neuroscience is to understand how the cortex processes visual information. A substantial effort has thus gone into characterizing input-output relationships across areas of the visual cortex, which has yielded an array of computational models. These models have, however, typically focused on one or very few visual areas (e.g., V1, MT), functions (e.g., object recognition, boundary detection, action recognition, etc.) or modules (form, motion, disparity or color). An integrated framework that would explain the computational mechanisms underlying visual processing beyond any specific area, module or function, while being at least consistent with the known anatomy and physiology of the visual cortex is still lacking.

Here, we present an integrated computational model of early vision that comprehensively describes neural responses in the primary visual cortex across modalities (form, motion, disparity and color) using a small set of elementary operations (linear-nonlinear cortical filter model, push-pull layout of excitation and inhibition, selective pooling, divisive normalization, etc.). We describe a basic circuit that combines “untuned” interactions in the classical receptive field (cRF) and “tuned” interactions in the extra-cRF. The circuit displays a range of interesting behaviors including digital selection and analog gain control and offers a characterization of disparate contextual phenomena across modalities as general induction phenomena. We show that the resulting circuit seems sufficient to capture the extent of psychophysical data on color constancy – offering a possible computational-level justification for the observed center-surround interactions.

===

Steven Shevell

University of Chicago

Color-Motion Feature Binding Errors Without Color: An AIC Story
Authors: Steven Shevell and Natalie Stepien

Peripheral and central moving objects of the same color may be perceived to move in the same direction even though peripheral objects have a different true direction of motion (Wu Kanai, & Shimojo, 2004). The perceived, illusory direction of peripheral motion is a color-motion feature binding error. Experiments considered whether color is essential to elicit these motion-binding errors, and tested two hypotheses that attempt to explain them. One hypothesis holds that binding errors occur because peripheral and central objects become linked if they have a combination of features in common. Another hypothesis is that binding errors depend on the overall feature correspondence among central and peripheral features represented at a preconjunctive level. Results showed that color (1) is not required to elicit motion binding errors, as David Knill posited at AIC 4 years ago, and (2) can, under some circumstances, inhibit rather than facilitate feature binding errors.

George Sperling

University of California, Irvine

Abstract: TBA

Alan Stocker

University of Pennsylvania

Perception: when Efficient coding meets Bayesian decoding

Two theoretical ideas have been instrumental in describing the process of perception. On one hand, "Efficient coding" argues that neural resource limitations lead to efficient sensory representations that are optimized with regard to the specific stimulus statistics of the environment. On the other hand, “Bayesian decoding” proposes that perceptual judgments are the result of interpreting the sensory information in the context of the observer's prior beliefs about the environment. While these theories have each played major roles in explaining properties of sensory neural responses and perceptual behavior, they have not been considered within a single model framework - until now. I present a new model description for perception that unifies the two theoretical ideas. Specifically, I introduce a more holistic Bayesian observer model that is constrained by assuming an efficient representation of the sensory input. The model makes a series of new and counter-intuitive ("anti-Bayesian") predictions. I will discuss these predictions in the light of existing psychophysical data and previous model descriptions, and will explore the model's implications for future experimental and theoretical investigations.

Bosco Tijan

University of Southern California

A computational account for the development of a preferred retinal locus
Authors: Bosco Tjan, Helga Mazyar

In normal vision, saccades bring the fovea to a visual target of interest. Central field loss (CFL) caused by macular degeneration means that such foveating saccades would be counterproductive. Patients with CFL often adopt a preferred retinal locus (PRL) in the peripheral retina for saccades and fixation. Factors such as acuity, eccentricity or contrast sensitivity cannot predict PRL formation. The optimality of adopting a PRL has also been questioned. Here we show that a conceptually simple computational model, which optimizes the expected target acuity for each saccade while at the same time estimating the unknowns needed for computing the expected acuity, can account for the formation of a PRL and its idiosyncrasies. We assume that the visual system always intends to bring the target of interest, with a saccade, to a retinal locus with the highest expected acuity. The expected post-saccade acuity of a retinal locus is a function of the physiological acuity at each retinal location, the positional uncertainty associated with each retinal location intrinsic to the visual system, and the expected oculomotor error. We assume that oculomotor error does not improve, since for each PRL-referred saccade there is a fovea-referred saccade with identical oculomotor movement at asymptotic accuracy. We assume that positional uncertainty associated each retinal location is optimally re-estimated after each saccade given the observed landing error using Bayesian update. Since the landing error is assessed with respect to the saccade target, the reassessment of positional certainty is local. We further assume a global forgetting function to prevent position uncertainty from vanishing. Simulations showed that immediately after CFL, the utilized retinal loci are close to the edge of the scotoma on the side nearest to saccade targets. After each saccade, positional uncertainties associated with two retinal locations -- the pre-saccade target location and the intended retinal locus for the saccade -- decrease. Decrease in positional uncertainty increases the expected post-saccade acuity of a retinal locus. Thus, a previously selected retinal locus is more likely to be selected for a future saccade because it has a higher expected post-saccade acuity, leading to further reductions of the positional uncertainty associated with it. A preference for a particular retinal locus hence emerges, even though this retinal locus may not be the closest to the target or with the highest physiological acuity. Idiosyncrasies at the early stages of CFL strongly influence PRL formation.

Laurie Wilcox

York University

Depth magnitude judgments for physical targets

Authors: Laurie M. Wilcox, L. M. Deas, B. Hartle

Depth percepts from stereoscopic targets are consistently and dramatically reduced when they form parts of a closed object. In a series of experiments we have shown that this reduction in perceived depth is contingent on both 2D (collinear L-junctions) and 3D (good disparity continuation) properties. We have proposed that this phenomenon is due to object-based disparity smoothing operations that serve to reduce perceived depth for object parts, but enhance object cohesion and detectability (Deas and Wilcox, 2014, 2015). These results are surprising, particularly given the well-known precision of the stereoscopic system in psychophysical studies (McKee, 1983). It is possible that the impact of figural grouping on stereoscopic depth perception we have demonstrated is limited to situations in which stereopsis is studied in isolation. In this presentation I will discuss a set of experiments in which we assess depth magnitude percepts using physical targets presented using a custom built physical stereoscopic display system. This apparatus uses precision motorized stages to position physical targets and records observers' magnitude judgments using a computerized sensor strip. In a series of condition, we evaluate the impact of additional depth information (e.g. shading, binocular disparity, and motion parallax) on suprathreshold depth estimates. Our results speak to the impact of figural grouping on depth perception in natural multi-cue environments, and to the impact of these additional sources of depth information on observers' reliance on binocular disparity in general.

Ali Winkler

University of Nevada, Reno
Alissa D. Winkler, John Erik Vanston, Nikolai Oh, Michael A. Crognale, Michael A. Webster

Psychophysical Asymmetries and Higher-Order Representations of Color Explored Using Unipolar VEPs

Previous research has shown that there is greater individual variation in achromatic settings, and a tendency for weaker sensitivity in many perceptual tasks along bluish-yellowish axes than along reddish-greenish axes that are chosen to have equivalent modulations along the LM and S cardinal axes. In addition to this "blue-yellow" (BY) asymmetry, we recently demonstrated an additional "blue vs. yellow" (BvsY) asymmetry in color naming, such that desaturated blues are more likely to be categorized as achromatic than equivalent yellows. The absence of a BvsY asymmetry in standard threshold and suprathreshold saturation tasks suggests that the color naming bias reflects a higher-order inference about color. To begin to explore whether BvsY biases can be evidenced in early cortical processing, we recorded visual evoked potentials (VEP)s for participants from site Oz while they fixated and passively viewed equiluminant grating fields in a 200ms onset, 800ms offset design. Chromatic contrast was modulated along axes in a cone opponent space, including the cardinal (0º-180º and 90º-270º), nominal blue-yellow (135º-315º), and its orthogonal (45º-225º) axis. Contrast thresholds at which along each axis the peak-to-trough response amplitudes meet a 2 x noise criterion threshold, form an iso-response contour that is elevated along the blueish-yellowish axes relative to the reddish-greenish axes. Other recent work using fMRI has suggested the presence of a non-cardinal bias in early cortical color coding, potentially reflecting a signature of cortical adaptation to the stronger blue-yellow variations in natural scenes and lighting. Further, within blueish-yellowish axes yellowish directions have higher thresholds than their bluish counterparts.

Jessica Witt

Colorado State University

Action Influences Perception but Only When Action is Reliable.

A person's ability to act influences spatial perception. Hills looks steeper and distances look farther to perceivers who would have to exert more effort to traverse the terrain. Objects that are easier to hit (e.g. baseballs, golf holes) look bigger and slower compared with objects that are difficult to hit. The optical information is constant, yet spatial perception varies as a function of action. In the current experiments, the reliability of the action was manipulated to determine whether perception only takes action into account when the action is reliable. To manipulate reliability, participants were given full control over their actions, and then control was taken away. Participants could anticipate the loss of control, and were theorized to be less likely to anticipate the outcomes when they were going to lose control. Action influenced perception much more when participants retained control than when control was lost. However, even when control was lost, action had a small influence on perception. Thus, perception incorporates many different sources of information and weights these sources according to the reliability of each.

Deyue Yu

Ohio State University College of Optometry

Letter recognition: crowding and perceptual learning

In everyday reading, letters in text are almost always closely flanked by other letters. The increased difficulty in identifying target letters due to the proximity of adjacent letters is referred to as crowding which is especially problematic in the periphery. It has been shown that crowded letter recognition is the major factor limiting reading. A better understanding of letter crowding has significance for both basic and clinical vision science. In this talk, I will discuss how crowding changes letter recognition and how perceptual learning changes the impact of crowding.
