

Research Statement

Computational Perception Overview

My research in Computational Perception is motivated by the view that it is essential to understand the information-processing principles operating in biological sensory systems in order to better understand human perception and to develop effective machine sensory systems. Central to this goal is elucidating the neural representation of sensory stimuli which allows the brain to extract useful information from a complex and noisy world. Since sensory systems evolved to represent the natural environment, a working hypothesis which guides my research is that sensory representations should be optimized to efficiently represent natural stimuli. Understanding the details of how natural stimuli are efficiently represented at multiple hierarchical layers of sensory processing is therefore the main goal of my work. The three main approaches that I take towards this goal are: (1) Collection and analysis of natural stimulus (image and sound) databases, (2) Development of biologically-inspired statistical models of natural stimulus coding and (3) Psychophysical experiments which examine the relationship between perception, neural coding, and the statistics of natural stimuli. In addition to these main interests, I am also interested in analysis of publically available sensory neurophysiology data, as well as pursuing applications of my work to practical problems in machine vision or hearing. Specific projects that I am currently interested in are described in the second section of this statement.

In terms of traditional disciplines, my research program in Computational Perception can be thought of as a three-legged stool comprised of computation, psychophysics and neuroscience. Therefore, I offer a rare flexibility in the sense that I can supervise and mentor students and from computational and biological backgrounds alike who share my deep interest in sensory systems. Ideally I would like to have a joint appointment in both a biological and computational department so that I can contribute fully to the intellectual life of the university with cross-listed interdisciplinary courses and seminars, as well as the ability to mentor a variety of students. Hopefully by teaching and mentoring students from a variety of departments, I will be able to partially bridge the unfortunate gap that often exists between the biological and quantitative sciences at many schools.

My research requires no expensive equipment, requires no external funding if I do not have graduate students, and can be done in a small room. All that I need apart from a salary and a room are computers to analyze natural stimulus databases and run simulations, and monitors for running visual psychophysics experiments. Additional equipment which I might be interested in acquiring as funding becomes available are an infrared eye tracker and a 3-D object scanner, but each of these items can be acquired for less than five thousand dollars. Therefore, a modest startup package could outfit my laboratory for the near future. Furthermore, my main research program is completely independent, and unlike many computational neuroscience colleagues I do not require data from other laboratories. Nevertheless, despite the independence of my research program, I am fortunate to have access to a rich network of colleagues at top universities with whom I can potentially collaborate, including my postdoctoral advisor Mike Lewicki at Case Western and my graduate advisor Kechen Zhang and other mentors like Ed Connor, Ernst Niebur, Eric Young and Xiaoqin Wang at Johns Hopkins.

A large part of my research involves computer programming and scientific computing, and I am very interested in utilizing the latest tools, including parallel cluster computing as and cloud supercomputing. I plan to have an 'open source' laboratory, where all code, models, and psychophysical data will be made available for download by the scientific community in the spirit of reproducible research. Programs will be fully documented and version tracked as is done by professional software developers.

Specific Project Examples

Project 1: Representation of texture edges by statistical image models and human observers

One important set of cues for determining depth relations in visual scenes are occlusion edges, where one object blocks the view of another. In natural images, multiple cues may define these boundaries, including luminance differences, texture differences and color differences. However, it is unclear what neural information-processing strategies can efficiently represent natural occlusion boundaries defined by various cues, and how they predict the performance of human observers in a boundary detection task. Preliminary work by myself and others in our laboratory has shown that a hierarchical distribution coding model of natural scenes (Karklin & Lewicki 2009) which models the conditional Gaussian distribution most likely to generate a given image patch (and learns complex cell receptive fields) far outperforms linear sparse representations (which learn simple cell receptive fields) at representing textures and texture-defined edges. Using a large database of 100 images whose occlusion boundaries have been hand-labeled by 6 human subjects, I will be performing psychophysical experiments on human observers and comparing their occlusion edge detection abilities to those of Bayesian classifiers defined using the sparse coding and distribution coding model representations. By analysis of the models and how they perform on the task, as well as comparison with human performance, I hope to gain insights into potential neural coding strategies the brain may employ to detect occlusion edges as well insights into how machine vision systems can solve these problems.

Project 2: Effects of global context on local edge perception by humans and statistical models

Quite often the local information available in an image patch is insufficient to determine whether or not an occlusion edge is present. Therefore, the visual system relies upon integration of global information from nearby image regions in order to accurately perceive contours. It is well known that contextual information affects both neuronal and behavioral responses to local edges. However, it is unclear to what extent these modulations of neuronal and behavioral responses are related to the statistical structure of natural scenes. Although previous work has shown that human performance on a contour integration task can be modeled by a Bayesian ideal observer applying a simple association rule based on two-point correlations measured from natural images (Geisler et. al. 2001), it is unclear how contextual information affects the perception of arbitrary image patches, since this requires an accurate conditional probabilistic model of large-scale image regions. In order to address this, I will extend existing hierarchical statistical models of natural image patches to integrate information across large images in a Bayes-optimal manner, and compare the ability of these models to accurately predict the nature of context-dependent observer biases in psychophysical tasks like local orientation discrimination and detection of low contrast local edges. This study will also make use of the image database.

Project 3: Efficient representation of three dimensional structure in natural images

Classical work has demonstrated that learning a sparse code for two-dimensional images yields a set of filters which resemble simple cell receptive fields (Olshausen & Field 1996). However, recent studies in higher visual areas have demonstrated neurons tuned for structure which defines three-dimensional objects (Yamane et. al. 2008). It is unknown how the response properties of these cells may be related to efficient coding strategies for natural three dimensional objects. This project will involve scanning natural 3-D objects, parameterizing them as topological deformations of a sphere, and finding a sparse efficient code for this parameterization using ICA, distribution coding or other methods for efficient representation. Hopefully, this will end up learning filters which resemble the 3-D receptive field tuning found in IT cortex and provide insight into efficient neural representation of 3-D structure.