Christopher DiMattina's Research Program in Computational Perception

Each item is 1 full-length journal paper. These projects will be completed during my post-doc or as an assistant professor.

Main projects in order

1) **Detection of occlusion boundaries in natural images by human observers and statistical analysis of natural occlusion boundaries**

   Occlusion boundaries are an important cue for determining ordinal depth relations from the two dimensional image, and are important for machine vision efforts to recover three dimensional structure and to segment an image into regions. Occlusion boundaries in natural images are often accompanied by the presence of multiple cues, including luminance differences, textural differences and color differences. In this study, we use a large database of color images whose occlusion boundaries have been hand-labeled by multiple human observers. We quantify the statistical differences between boundary regions and non-boundary textures which could potentially be exploited by an organism or computer program to locate potential occlusion boundaries. In order to carefully and quantitatively determine the local cues employed by human observers for discriminating boundary regions from non-boundary texture regions, we perform psychophysical experiments on human observers whose task is to indicate whether a given image patch contains an occlusion boundary or not. We compare the results of the experiment to ideal observers defined on the statistics measured from the database. We find that simple luminance differences and color contrast is not sufficient to predict behavior in detecting natural occlusion boundaries, but combining these cues with texture does a far better job of predicting human performance. Therefore, we see that texture representation, even at relatively local scales, is an important for detecting occlusion boundaries.

2) **Representation of natural textures by a distribution coding model of natural images**

   One important task for the visual system is segmenting a visual scene into regions. One useful set of cues for performing this segmentation are occlusion boundaries, which are formed when one object blocks the view of another. In natural images, scene regions and occlusion boundaries which divide them are often defined by textural differences as well as luminance and color differences. In this study, we seek to understand how the representations formed by a distribution coding statistical image model solve the problem of segregating different texture regions and identifying occlusion boundaries defined by textural cues. We demonstrate that the representations learned by this model when trained on natural images are well suited to discriminate textures and can also generalize within texture regions as well as generalizing along occlusion boundaries. We compare the performance of this generative model to that of a standard feed-forward model of texture segregation as well as simpler linear models on a natural texture discrimination task. We analyze individual model units in the distribution coding model and show that units sensitive to second-order edges similar to those seen in V2 are capable of reliably indicating the presence of natural texture-defined occlusion boundaries. We also consider the representation of edges by a recently developed complex phase model, and study empirically the correlated phase structure that exists along edges taken from a natural database. We develop a novel hierarchical model which combines phase and amplitude information and study the represenation of natural edge structure by this model.

3) Training distribution coding model with color image patches and extension of texture definition to include color

A previous study in our laboratory has demonstrated that a distribution coding model trained on natural images learns a representation which replicates many well-known nonlinear tuning properties of neurons in the early visual system. However, this study only considered training the models with gray-scale images and therefore does not take into account the distributions of natural color images. In this study, we train the distribution coding model with color image patches and analyze the resulting receptive fields. We find that similar to results with studies applying ICA to color images that the model learns first-order receptive fields (b vectors) which exhibit R-G and B-Y opponency as well as non-color tuned units which are more sensitive to spatial features. However, we also find that second-order receptive fields emerge which integrate color opponency information across various regions of the receptive field. Our model suggests the possibility of neurons in area V2 which may exhibit sensitivity to second-order color edges. We discuss the possibility of the combination of color and texture and basis elements of color 'textons' or 'colons'.

4) Definition of image patch saliency using statistical models of natural images and comparison to human eye movement data (Psychophysics)

Many previous studies have investigated models for computing the 'saliency' of image regions, and then compared the outputs of these models to the locations of human fixations free-viewing images. These models are feed-forwards and are based on prior assumptions about what sorts of visual features are likely to be salient, and are not trained using natural image data. In this study, we propose an alternative approach whereby we define the saliency of image patches using statistical models of natural images. Since a statistical model of natural image patches infers a set of sparse latent variables (or causes) which were most likely given the observed image patch, we can sensibly define the 'salience' of an image patch as the Shannon surprisal $S = -\ln p(s)$, where $p(s)$ is the prior probability evaluated on the inferred latent variables in the vector s. This measure allows us to say how 'unlikely' or 'surprising' a given image patch is under the generative scheme of the model. We compare this measure for two models of natural image patches: A standard 'sparse coding' model and a 'distribution coding' model. We hypothesize that fixations will be preferentially located at regions of the image which are considered to be less likely or more improbable.

5) **A hierarchical model of contextual interaction between neighboring image patches and predictions for role of context in perception**

In the brain, local processing of sensory information taking place in a particular cortical column is subject to contextual influences coming from adjacent cortical columns at the same level of processing, as well as feedback influences emanating from higher areas. In this study, we consider a model which incorporates lateral interactions between different image patches in order to model the joint distribution of latent variables in multiple neighboring image patches. We show that when trained with natural images, this model predicts long-range interactions between neurons with similar orientations like that seen in area V1. We find that these interactions improve local inference and de-noising for natural images by exploiting the long-range correlational structure seen in natural scenes. We compare the predictions of this model to several well-known neurophysiological and psychophysical effects of contextual influence on sensory processing, observing qualitatively similar results.

6) **Contextual effects of surrounds on local perception: Effects of target contrast and the predictions of a Bayesian ideal observer model (Psychophysics)**

It is well known from previous psychophysical studies and neurophysiological studies that global contextual information can have a strong influence on local perception. One example of this is the effects which a surrounding field of co-linear Gabor functions can have on the both the neuronal tuning of a V1 cell as well as the perception of its orientation. In particular, for a vertical bar and a field of vertical surrounding Gabor functions, it has been shown that observers perceptions of the central target orientation is biased towards the vertical, as are the neuronal rate responses. One can give these results a Bayesian interpretation and state that the visual system is exploiting prior information about the likely orientation of the central Gabor target given the context suggested by the surrounding field. If this is the case, then it would follow that as the central information becomes more weak or ambiguous then there would be a greater reliance on the contextual information, thus resulting in a larger bias. In order to investigate this possibility, we perform psychophysical experiments on human observers where we systematically vary the visibility of the central target relative to the surround.

7) **Statistical analysis of a 3-D database of scanned natural objects and predictions for neural coding**

Recent work recording from neurons in the infero-temporal cortex of the macaque monkey has demonstrated neurons tuned for the three-dimensional shapes. It is of great interest to relate the tuning properties of these neurons to the statistical properties of natural 3-D shapes. However, there is not presently any sort of large database of three-dimensional natural shapes which can be used to make the necessary measurements. Using commercially available scanners, we will scan in a large number of natural objects (rocks, sticks, fruit) and plastic figurines of animals. We will measure the statistical properties of these scanned images and consider how they may be represented efficiently by a neural code of units tuned for curvature at various locations. We will see to what extent the predicted code matches the tuning properties of neurons seen in real data (we will get real tuning curves from Ed Connor). We will also apply ICA and other methods to find sparse representations of 3-D structure.

8) Review: Statistical Learning Approaches to Sensory Neurophysiology

This review will consider recent quantitative approaches to sensory neurophysiology as a machine learning problem. It will be based on my thesis work and co-written with Kechen Zhang. We will broadly consider the system identification or model-based neurophysiology approach where the goal of experiments is to estimate the parameters of a sensory processing model. We will consider work on adaptive stimulus design, especially the work of Paninski and discuss some fundamental difficulties with identifying nonlinear models (DiMattina & Zhang 2010) and how they can potentially be overcome with active data collection. We will discuss the general approach proposed by DiMattina & Zhang in their most recent paper which utilizes a two-phase experiment where multiple competing models are estimated and then compared using critical stimuli. We will suggest future directions for research in this field and how one may integrate prior knowledge about receptive field properties and neural responses into these sensory processing models.

9) Review: Hierarchical Models of Sensory Processing

This review will consider recent models of higher-order structure in natural images and their implications for understanding visual processing in the brain. This model will function as a tutorial, with selected examples being implemented in MATLAB code which will be available for download. We will consider recent work by Lewicki, Olshausen, Simoncelli and collegues, as well as the Numenta HTM developed by Hawkins and George, and work by Hinton and colleagues on the Restricted Boltzman machine and related models. We will discuss mathematical relationships between the different models, and suggest directions for future research which may shed light on hierarchical neural representation in the primate visual system.