Automated depth perception as a behaviorally-relevant metric for the innate learning paradigm

Provost fellowship research proposal
Gordon Kratz (BS in Computer Science)

Objective

Implement an algorithm for depth perception using binocular linear filters – the mathematical description used for primary visual cortex receptive fields in visual neuroscience. Use this algorithm to score the efficacy of model visual systems that have been developed using various patterns of spontaneous neural activity.

Background and significance

Highly organized patterns of spontaneous activity have been demonstrated in the developing visual systems of a wide variety of species. These neural activity patterns are bounded, amorphous patterns that form spontaneously, move over time, and dissipate. They show much more organized structure than would be expected from random neural firing alone.

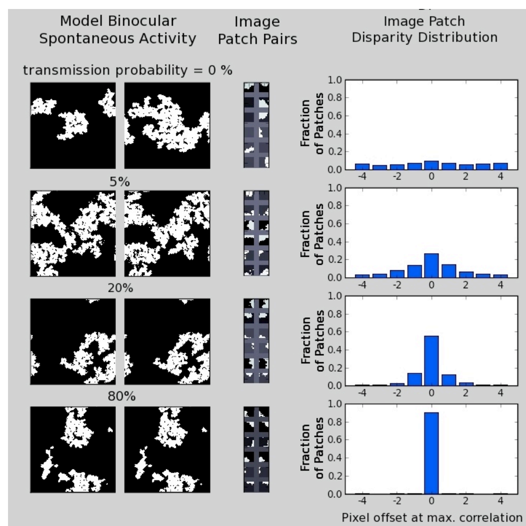

Fig. 1. **One current method of evaluating the binocular model.** Versions of the generated binocular activity patterns (left) with six pairs of image patches per version (middle) demonstrating the amount of similarity between eye layers. While the image patch disparity distribution (right) is one way of judging potential activity models - with the middle two being closer to typical binocular experience - the extent to which each version permits depth perception would be more behaviorally relevant and accessible to the general public (reproduced with permission from Dr. Albert)

What is the purpose of such structured spontaneous activity? Previous work has shown that such patterns contain the same low-level statistical structure as is present in natural visual scenes (Albert et al 2008). Because of this, any learning algorithm that relies on this structure can be used to adapt to both natural scenes as well as these early patterns of neural activity. The early visual systems of most animals adapt to changes in their visual experience early in life (e.g. more vertically-oriented cells when more vertical stimuli are present). The innate learning hypothesis presupposes that such learning is occurring even before the eyes open, on these patterns of neural activity.

The efficient coding approach is a way to understand the early visual system as a computational objective. Previous work has demonstrated that one of the goals of the early visual system is to produce a sparse (Olshausen & Field 1996) or statistically independent code (Bell & Sejnowski 1997). These coding strategies have been successfully applied to activity patterns resembling retinal waves (Albert et al 2008). The results showed this abstract, physiological spontaneous activity model is capable of refining primary visual cortex receptive fields.

Aims and Approach

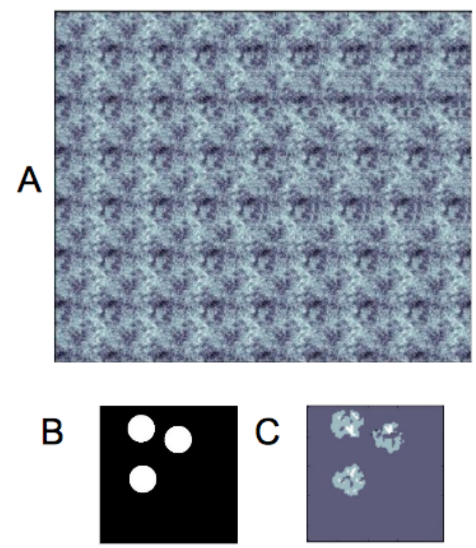

Fig. 2. **Model evaluation by depth perception – "magic eye" example.** A) an autostereogram (e.g. "magic eye") which was produced using the depth map represented in B. An estimated depth map can be generated, as in C, and by scoring the difference between the true depth map B and the derived depth map C, the quality of the binocular model can be evaluated (reproduced with permission from Dr. Albert)

The goal of my work will be to improve the way these models are evaluated. The current technique for evaluating computational models in visual neuroscience is to compare statistical properties of their receptive fields. Although properties like orientation bandwidth, phase, and spatial frequency bandwidth can be related between physiological and computational models, it can be difficult to determine the impact of deviations from these properties observed in nature.

My work will be in creating a clearer, more behaviorally-relevant metric for evaluating a developing visual system – the ability to perceive depth.

There will be four stages to my approach. First, I will use a set of idealized filters (ones which resemble the filters found in the visual system and are known to perform well in depth perception algorithms) and a straightforward, but crude strategy for estimating depth (aim 1). Once that prototype is complete, I will implement more advanced algorithms for optimally combining the output of a set of binocular filters (e.g. Qian 1994) to create a perceived depth estimate from binocular images (aim 2).

- **Aim 1:** Develop a prototype system using idealized filters
- **Aim 2:** Refine depth perception system using current literature
- **Aim 3:** Real-time demo
- **Aim 4:** Automatically score resulting depth images

While improving the depth perception algorithm, I will concurrently create a real-time system for demonstrating the capabilities of the software (aim 3). This will allow presentations of the "innate learning" approach to be understood intuitively by an audience, whether scientists or the general public. Throughout development, a method of scoring the estimated depth maps to the known depth will be applied (aim 4). This automated scoring will be critical to finding the right innate learning patterns to train the model visual system.

After completion of the project, I will present my work at a local or regional computer science conference in addition to the Loyola spring research symposium. My presentation will have a live demo, developed over the summer, for use during an oral or poster presentation.

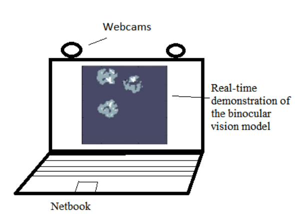

Fig. 3. A mock-up of the live demo

Conclusion

In studying the development of the animal visual system, an efficient coding approach has proven useful as a tool for modeling the early stages of visual processing. Metrics used to judge various models remain inaccessible to those not familiar with the field. A visual, behaviorally-relevant metric would better represent the underlying principles while lowering the barriers to understand and evaluate this type of research. I propose to implement such a metric on the binocular vision models of the Theoretical Neuroscience Lab at Loyola, open-sourcing and generalizing them for use by anyone who could benefit from the research.

Deliverables/Timeline

May 26th	Create idealized binocular linear filters (gabor wavelets) Create autostereogram test instance
June 2nd	Select database of coregistered binocular and range images Purchase netbook and dual webcam setup
June 9th	Complete the initial, crude depth inference algorithm (aim 1)
June 30th	Optimize depth perception algorithm for coregistered images (aim 2)
July 14th	Create depth perception real-time demo (aim 3)
July 28th	Automate scoring of depth images (aim 4)
August 11th	Streamline for easy application to the innate learning biological models

Roles and responsibilities

- Gordon Kratz, undergrad in Computer Science, will be focused on the implementation of a depth perception algorithm based solely on disparity using binocular linear filters. Later, these inferred depth patterns will be automatically scored to “close the loop” in optimizing the spontaneous patterns.

- Dr. Mark V. Albert, professor of Computer Science, heads the Theoretical Neuroscience lab. He will provide guidance on integrating this project with the lab’s current work in visual development.

- Neil (Abishek) Rao, undergrad in Biology, will be responsible for continued work on the binocular biological model of spontaneous pattern generation.