Virtual Eye: a spatio-temporal bottom-up eye sensitivity model

Todd Goodall, Facebook Reality Labs

Introduction. Video quality and compression models use the spatial contrast sensitivity function (CSF) [5], which is solved based on a linear system approximation. This function measures the eye's sensitivity to sinusoid gratings, ignoring the subtle connectivity and inhomogeniety of cell density across the visual field. Non-linear aspects of the eye, such as the change in frequency sensitivity with changing illumination, are not captured by this simple approximation. We propose Virtual Eye, a bottom-up approach that models the spatio-temporal dynamics of the eye across the visual field. Each functional retinal cell layer in the eye is modeled using non-uniform spatial cell responses, which can be easily extended to incorporate complex retinal nonlinearities. Given any grayscale signal input, Virtual Eye produces a dense output that describes the total retinal energy transmitted to the brain for each point in the visual field.

Model. The T frame, $M \cdot N$ resolution signal V is first treated with a lens distortion function to form the retinal projection, V_R. This projection is then processed by a continuous inhomogenous 'field' of photoreceptors R_{cones}, which transmit pointwise information along center and surround pathways. Lastly, midget and parasol cell fields difference these paths over time to produce dense spatio-temporal outputs R_{midget} and $R_{parasol}$.

Each cell layer is computed using a spatially continous Gaussian cell approximation. The local width of Gaussian per pixel coordinate is computed using physiological cell density measurements [2], display distance, and display resolution to ensure output spatial units of cells per pixel. The spatial connectivity and receptive fields of cones, midgets, and parasols are all captured by Virtual Eye.

After computing R_{midget}, the responses are normalized for sensitivity analysis using

$$\frac{1}{TMN} \sum_{(t,m,n)} \left| R_{midget}(t,m,n) + \tau_m \right|^{\alpha} D^2_{midget}(m,n) \quad (1)$$

where t, m, and n index the temporal axis, the vertical spatial axis, and the horizontal axis respectively. The parameters τ_m and α model cell firing rate and output non-linearity. $D^2_{midget}(m,n)$ refers to the squared density of the midget cells. Dividing by TMN normalizes for resolution and frame rate. Finally, sensitivity can be directly predicted by adding (1) to the complementary computation involving $R_{parasol}$.

Validation. We calibrated Virtual Eye simultaneously to multiple datasets. Each dataset was chosen to cover different aspects of retinal sensitivity. Figure 1 depicts the spatiotemporal fit for flickering sinusoids presented at the fovea [3]. Figure 2 depicts sensitivity falloff matching across the visual field for select spatial frequencies [4]. We also calibrated with the Modelfest dataset [1]. For each dataset, we observing Spearman's correlation greater than 0.95. For visualizing the calibrated system response, we plot the resulting Virtual Eye midget cell output energy field given the image of a spatial CSF stimuli in Figure 3, demonstrating the emergent property of the csf from the retinal cell populations.

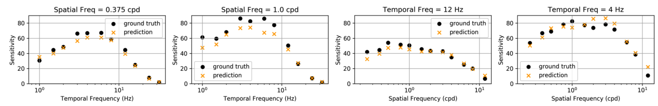
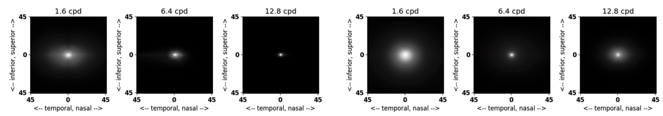


Fig. 1: Virtual Eye fits to spatio-temporal data.



(a) Measurements (b) Virtual Eye

Fig. 2: Sensitivity vs spatial frequency and eccentricity.

Conclusion. We consider this early validation of an exciting new approach, which has the power to model traditionally more challenging phenomena such as eye movement, cortical magnification, and luminance adaptation through emergent properties of cell populations.

References

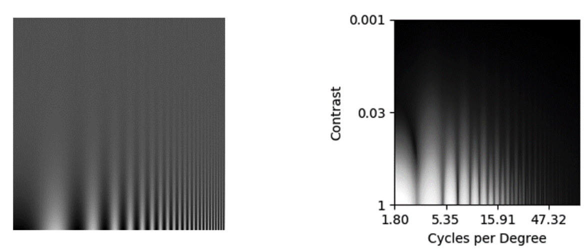
[1] Thom Carney, Christopher W Tyler, Andrew B Watson, Walter Makous, Brent Beutter, Chien-Chung Chen, Anthony M Norcia, and Stanley A Klein. Modelfest: Year one results and plans for future years. In *Human vision and electronic imaging V*, volume 3959, pages 140–152. International Society for Optics and Photonics, 2000.

[2] Dennis M Dacey and Michael R Petersen. Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proceedings of the National Academy of sciences*, 89(20):9666–9670, 1992.

[3] DH Kelly. Motion and vision. ii. stabilized spatio-temporal threshold surface. *Josa*, 69(10):1340–1349, 1979.

[4] JS Pointer and RF Hess. The contrast sensitivity gradient across the human visual field: With emphasis on the low spatial frequency range. *Vision research*, 29(9):1133–1151, 1989.

[5] J. G. Robson. Spatial and temporal contrast-sensitivity functions of the visual system. *JOSA A*, 56(8):1141–1142, Aug 1966.



(a) Spatial CSF input (b) Virtual Eye midget energies

Fig. 3: Sensitivity of local foveal midget cells to CSF.