Brightness Illusions as Optimal Percepts

Santiago Jaramillo
sjara@ieee.org

Barak A. Pearlmutter
barak@cs.may.ie

Hamilton Institute. National University of Ireland Maynooth
Maynooth, Co. Kildare, Ireland.

Abstract

We show that Mach bands and a number of other low-level brightness illusions can be accounted for by assuming that the perceptual system performs simple Bayesian inference using a Gaussian image prior with noisy retinal gangion cells. This theory accounts for phenomena which have proven problematic for simple energy-based and lateral-interaction models while avoiding the complexities of mid-level vision theories that involve the estimation of structure and albedo.

1 Introduction

Hartline shared the 1967 Nobel Prize in Physiology and Medicine for his discovery of lateral inhibition in the retina (Hartline et al., 1956). This physiological phenomenon was interpreted as performing "sharpening of contrast so that forms stand out more clearly" (Bernhard, 1967), and accounting for low-level brightness illusions such as Mach bands (Ratliff, 1965). Ernst Mach himself proposed a Laplacian derivative operator as an explanation of these phenomena, following Ewald Hering. More recent theories of lateral inhibition in the retina eschew "contrast enhancement" as an explanation and are instead based on coding theory under the normative assumption that the retina attempts to encode images so as to maximize fidelity while transmitting them through the optic nerve (Atick and Redlich, 1990b,a). This theory and its elaborations (Atick et al., 1992; Linsker, 1993; Li, 1996) have enjoyed enormous success at accounting for receptive field properties. They do however assume that the rest of the brain knows the point-spread function of the retina, and therefore, in contrast to simple retinal lateral inhibition models, do not predict that percepts will correspond to simply applying the retinal transfer function to the image.

For this and other reasons, lateral inhibition in the retina is no longer taken seriously as an explanation of brightness illusions by experts in that area, although it does continue to dominate survey courses and textbooks. Modern brightness illusion theories fall into three classes, none of which ascribes a direct role to retinal effects. The first,

so-called energy-based models, are the philosophical descendent of the retinal lateral inhibition theory. These assume that a post-retinal relaxation process constructed to minimize a conflict (or energy) measure leads to brightness percepts. This relaxation process in the end results in (cortical) lateral inhibition. Such energy-based models have elaborate architectures carefully tuned to the phenomena they attempt to explain.

The second class, constancy models, are cast at a higher level of abstraction and hold that brightness and color illusions result from mechanisms the visual system uses to obtain stable veridical percepts in the face of varying illumination (Adelson, 2000, 1993; Sun and Perona, 1996a,b). Constancy models have excellent explanatory power for a broad class of images, particularly those with apparent transparency or geometric structure, but when faced with low-level brightness illusions such as Mach bands and simultaneous contrast effects they resort to special pleading concerning non-veridical geometric structure and albedo supposedly inferred by mid-level visual processes.

The third, physiological models, start with measurements of physiological activity in visual areas during presentation of an ensemble of stimuli that includes stimuli known to produce illusions and show that the measured neuronal activity, interpreted using a posited representation, are consistent with known percepts (Albright, 1994, 1995; Rossi et al., 1996; Rossi and Paradiso, 1996; Newsome et al., 1989). Although such work elucidates mechanism and representation, it is predictive only to the extent that the neurophysiology is fully understood.

The study of motion perception has revealed a rich set of seemingly counterintuitive non-veridical percepts. These motion illusions are well explained by a recent theory which posits that the visual system performs optimal perceptual inference in estimating motion, and that this processing necessarily takes into account both signal and noise (Weiss et al., 2002). Here we introduce a theory philosophically similar to the constancy and optimal motion estimation theories. The theory proposed here extends the optimal perception hypothesis to encompass low-level brightness illusions without invoking complex priors that involve three-dimensional structure, grouping, illumination, transparency, or albedo.

2   Methods

Van Beers et al. (2002) posit that *"neural signals are corrupted by noise and this places limits on information processing"* in the context of sensorimotor control. In this section we calculate such a limit in a perceptual system by finding the optimal Bayesian reconstruction of an image presented to the retina under a set of assumptions concerning noise between transduction and transmission along the optic nerve, and using a simple generic prior.

The input to the model presented here is a one-dimensional discrete signal (corresponding, for example, to one row of an image) which represents the lightness of the stimulus. The intensity range of the pixels is centered at zero (representing mid-level gray). Negative values correspond to darker tones, and positive values to lighter tones.

The transformation of the stimulus $\mathbf{x}$ at our one-dimensional annular retina is modeled as
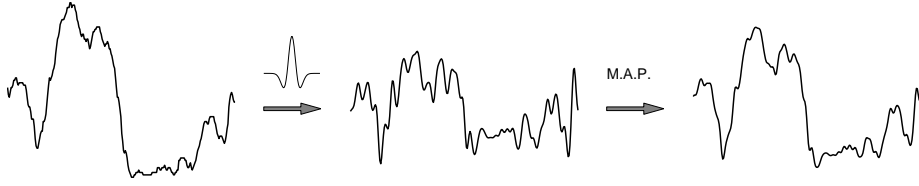
$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \eta \tag{1}$$

Figure 1: Transformation and inference process. The stimulus (left) is convolved with the *Mexican hat* vector to produced a new signal (middle). The brain has to guess the original pattern (here using a *maximum a-posteriori* estimation approach) from a noisy version of the transformed signal.

where $\mathbf{\Phi}$ is a circulant matrix and η is the noise due to transduction and transmission. Each row of $\mathbf{\Phi}$ corresponds to a shifted version of a *Mexican hat*-shaped vector[1] defined by

$$\phi(k) = a(b - k^2)e^{-k^2/2} \tag{2}$$

where a determines the power of the signal and b defines the area ratio between the positive and negative lobes. This transformation thus corresponds to convolving the input with a Mexican hat function. We assume that the noise η is white, Gaussian, zero-mean, and uncorrelated with the input signal. The covariance matrix is simply $\mathbf{C}_\eta = \sigma_\eta^2 \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

Given $\mathbf{y}$ as the signal received from the retina, we assume that the brain knows how to estimate the stimulus $\mathbf{x}$ that maximizes the posterior probability

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \tag{3}$$

We assume a zero-mean white Gaussian prior on stimuli, with diagonal covariance matrix $\mathbf{C}_\mathbf{x} = \sigma_\mathbf{x}^2 \mathbf{I}$

$$p(\mathbf{x}) \propto e^{-\mathbf{x}^T \mathbf{C}_\mathbf{x}^{-1} \mathbf{x}/2} \tag{4}$$

Given our simple assumptions the likelihood can be written as

$$p(\mathbf{y}|\mathbf{x}) \propto e^{-(\mathbf{y}-\mathbf{\Phi}\mathbf{x})^T \mathbf{C}_\eta^{-1}(\mathbf{y}-\mathbf{\Phi}\mathbf{x})/2} \tag{5}$$

The estimated stimulus $\mathbf{x}$ which maximizes the consequent log-posterior

$$\log p(\mathbf{x}|\mathbf{y}) = -\frac{1}{2\sigma_\eta^2}(\mathbf{y} - \mathbf{\Phi}\mathbf{x})^T(\mathbf{y} - \mathbf{\Phi}\mathbf{x}) - \frac{1}{2\sigma_\mathbf{x}^2}\mathbf{x}^T\mathbf{x} - \text{const.} \tag{6}$$

must then satisfy the linear equation

$$(\sigma_\eta^{-2}\mathbf{\Phi}^T\mathbf{\Phi} + \sigma_\mathbf{x}^{-2}\mathbf{I})\mathbf{x} = \sigma_\eta^{-2}\mathbf{\Phi}^T\mathbf{y} \tag{7}$$

---

[1]Expression (2) defines a wave centered at zero, with effective support $(-5, 5)$. The vector used in the model consists of a discrete shifted version of this wave, with sampling step depending on the desired effective width.
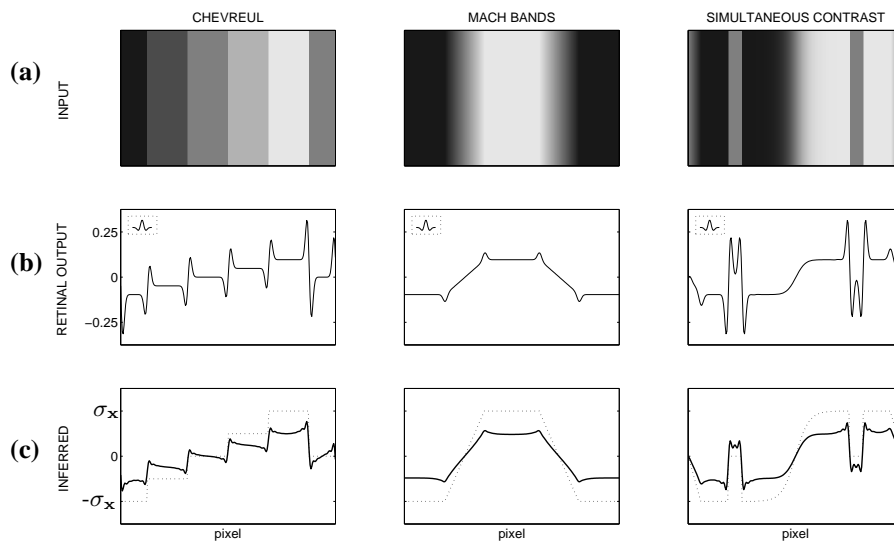
Figure 2: Optimal perceptual inference for 3 different patterns. The top row (a) presents the stimuli as images. The middle row (b) shows the transformed pattern after being convolved with the Mexican hat function. The actual width of the Mexican hat is shown on the top-left corner of each plot. The bottom row (c) presents the inferred pattern (solid) for each stimulus (dotted).

which is equivalent to the Wiener filter for the deconvolution of a noisy signal. Figure 1 shows an example of the transformation and inference process for a strip of a natural image.

For the simulations presented in the next section, we used the parameter values $a = 0.15$ and $b = 1.1$ for the Mexican hat transformation (see eq. 2). This value of b makes the total area of the two negative side lobes equal to 82% of the area of the positive central lobe. The Gaussian prior had standard deviation $\sigma_{\mathbf{x}} = 1$ and the standard deviation of the ganglion cell noise was $\sigma_{\eta} = 0.1$.

The stimuli consist of 256 pixels, with the maximum lightness set to $\sigma_{\mathbf{x}}$. The effective support (non-zero region) of the Mexican hat signal used for these simulations is approximately one tenth of the length of the input pattern (see figure 2, top-left corner of each plot in the middle row).

3   Results

We applied this simple cartoon model of the retina to the three most popular one-dimensional brightness illusion stimuli. None involve overt mid-level visual elements like perceived transparency or obvious geometric structure. Figure 2 shows the stimuli along with the inferred pattern following the retinal transfer function (without any noise

actually being injected) and Bayesian reconstruction.

Changes in the simulation parameters result in the following qualitative behavior. If the variance of the noise increases (approaching the value of the variance of the prior) the inferred signal is pulled toward zero. In contrast, if the variance of the noise is very small, the reconstructed pattern becomes closer to the original stimulus.

4 Discussion

The results from Figure 2 are consistent with reported percepts for each of these patterns. In the Chevreul/Staircase pattern the bars with homogeneous lightness are perceived as ramps. Mach Bands (peaks of brightness) appear at the edges between a ramp and a plateau. And, for the simultaneous contrast stimulus, we perceive the gray bar on the light background as darker than the other one, even though they have the same actual intensity.

Due to the simple linear imaging model and Gaussian prior on stimuli, the forward process and the optimal estimation process are both linear. This means that just as the Mexican hat function can be calculated as the impulse response function of the forward process, we can calculate the impulse response of the full model consisting of the composition of the forward process and the optimal estimation process. The result is shown in Figure 3.

Changing the magnitude of the noise or the prior modifies the inference process in the intuitively expected fashion. As the noise becomes larger the prior is weighted more heavily, which in our case indicates that pixels at the mean lightness become more probable. On the other hand, if the noise is small enough the inference consist of inverting the retinal transfer function, obtaining a pattern nearly identical to the actual stimulus. It is important to note that spike rates in retinal ganglion cells are not high, and the visual system must estimate image properties rapidly. Even if the retina itself is nearly deterministic, shot noise introduces considerable uncertainty concerning the activity level of a ganglion cell as estimated by a brief observation of its spike train.

This theory challenges the traditional edge-enhancement view of retinal lateral inhibition. According to the currently accepted theories, lateral inhibition in the retina is optimized to encode Gaussian signals of limited frequency content, using an optimal lossy encoder. Sharp edges or sudden gradient changes are exceedingly unlikely under the Gaussian image hypothesis, and an optimal lossy encoder will introduce greater distortion into low probability inputs. Hence one might expect sharp edges to cause local distortion of the estimated brightness. Far from "enhancing edges" as the traditional theory holds, lateral inhibition in the retina produces a code that performs well on Gaussian images but poorly on images containing sharp brightness or gradient discontinuities.

This account of low-level brightness illusions is an optimal perception theory in that it posits optimal interpretation of the signals that reach the brain through the optic nerve. This differs from the optimal perception account of motion illusions (Weiss et al., 2002), which result from optimal interpretation of the visual stimuli themselves.
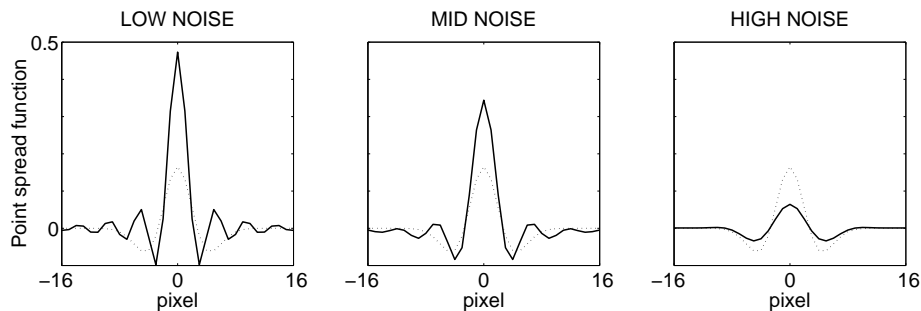
Figure 3: The retinal processing impulse response is shown (dotted line) along with the impulse response function of the composition of the retinal processing and the optimal reconstructor (solid line). In all three panels $\sigma_{\mathbf{x}} = 1$, while moving from left to right $\sigma_{\eta} = 0.01, 0.1, 1.0$. The center panel corresponds to the simulations of Figure 2.

5 Conclusion

We have seen that a number of low-level brightness illusions can be accounted for by assuming that the perceptual system performs simple Bayesian inference using a Gaussian image prior, and positing measurement noise on the outputs of the retinal gangion cells. This theory accounts for phenomena which have proven problematic for simple energy-based and lateral-interaction models while avoiding the complexities of mid-level vision theories that involve the estimation of structure and albedo.

As in other sensory domains, a general Bayesian principle and a simple generic prior has proven surprisingly powerful at accounting for low-level brightness illusions.

6 Future Work

Because this model is easily extensible within the Gaussian linear framework, our agenda is straightforward: we will retain the model's simplicity while making firmer contact with the physiology, in order to sharpen the model's predictions and broaden its range of applicability:

- Two dimensional stimuli.

- Construction of novel brightness illusions via optimization of measure of non-veridicality.

- Calibration of the model using realistic ganglion cell receptive fields.

- Physiologically realistic levels of shot noise in ganglion cell outputs, which will make the noise of a ganglion cell vary with its activity level.

- Inclusion of color.

6

- Incorporation of physiologically plausible saturation effects, since saturation of a neuron constitutes an additional source of noise.

- Incorporation of brightness and local contrast adaptation.

Once the simple model has been exhausted, one could consider moving beyond the Gaussian framework by including for example an edge process in the prior.

Acknowledgements

References

Adelson, E. (1993). Perceptual organization and the judgment of brightness. *Science*, 262:2042–2044.

Adelson, E. H. (2000). Lightness perception and lightness illusions. In Gazzaniga, M., editor, *The New Cognitive Neurosciences*, chapter 24, pages 339–351. MIT Press, 2nd edition.

Albright, T. D. (1994). Why do things look as they do? *Trends in Neurociences*, 17(5).

Albright, T. D. (1995). My most true mind thus makes mine eye untrue. *Trends in Neurociences*, 18(8):331–333.

Atick, J. J., Li, Z., and Redlich, A. N. (1992). Understanding retinal color coding from first principles. *Neu. Comp.*, 4(4):559–572.

Atick, J. J. and Redlich, A. N. (1990a). Predicting ganglion and simple cell receptive field organizations. *Int. J. of Neural Systems*, 1:305.

Atick, J. J. and Redlich, A. N. (1990b). Towards a theory of early visual processing. *Neu. Comp.*, 2(3):308–320.

Bernhard, C. G. (1967). Presentation speech by member of the Nobel committee for physiology or medicine of the Royal Caroline Institute.

Hartline, H. K., Wagner, H. G., and Ratliff, F. (1956). Inhibition in the eye of Limulus. *Journal of General Physiology*, 39(5):651–673.

Li, Z. (1996). A theory of the visual motion coding in the primary visual cortex. *Neu. Comp.*, 8(4):705–730.

Linsker, R. (1993). Deriving receptive fields using an optimal encoding criterion. In *Adv. in Neu. Info. Proc. Sys. 5*, pages 953–960. Morgan Kaufmann.

Newsome, W. T., Britten, K. H., and Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, 341:52–54.

Ratliff, F. (1965). *Mach bands: Quantitative studies on neural networks in the retina*. Holden Day, San Francisco.

Rossi, A. F. and Paradiso, M. A. (1996). Temporal limits of brightness induction and mechanisms of brightness perception. *Vision Res.*, 36(10):1391–1398.

Rossi, A. F., Rittenhouse, C. D., and Paradiso, M. A. (1996). The representation of brightness in primary visual cortex. *Science*, 273:1104–1107.

Sun, J. and Perona, P. (1996a). Early computation of shape and reflectance in the visual system. *Nature*, 379:165–168.

Sun, J. and Perona, P. (1996b). Preattentive perception of elementary three-dimensional shapes. *Vision Research*, 36:2515–2529.

van Beers, R. J., Baraduc, P., and Wolpert, D. M. (2002). Role of uncertainty in sensorimotor control. *Phil. Trans. R. Soc. Lond. B*, 357:1137–1145.

Weiss, Y., Simoncelli, E. P., and Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nat. Neurosci.*, 5(6):598–604.