Context-dependent efficient sensory coding underlying the tilt-illusion in human visual cortex and artificial neural networks

Ling-Qi Zhang (zhangl5@janelia.hhmi.org)
Janelia Research Campus, Howard Hughes Medical Institute
Ashburn, VA 20147, United States

Geoffrey K. Aguirre (aguirreg@pennmedicine.upenn.edu)
Department of Neurology, University of Pennsylvania
Philadelphia, PA 19104, United States

Alan A. Stocker (astocker@psych.upenn.edu)
Department of Psychology, University of Pennsylvania
Philadelphia, PA 19104, United States

Abstract

Human perception of basic visual attributes is often influenced by spatial context. A canonical example is the tilt illusion, in which the perceived orientation of a stimulus is altered by the presence of a spatially oriented surround. We hypothesize that surround effects originate from changes in neural representation that maximize coding efficiency based on spatial context. We simultaneously collected perceptual and fMRI data from human participants in a tilt-illusion experiment. We found that orientation encoding in the absence of a surround reflects natural scene statistics both in behavior and in the neural responses of the visual cortex. In the presence of an orientated surround, encoding accuracy was significantly increased at the surround orientation. The pattern of change in coding accuracy is consistent with the surround-conditioned orientation statistics of spatially adjacent regions in natural images. Furthermore, we found the same orientation encoding characteristics and contextual modulation in convolutional neural networks trained on natural images. Our results suggest that efficient coding based on spatial context is a general mechanism in visual processing of natural images.

Keywords: efficient coding; tilt illusion; Fisher information; convolutional neural network; natural scene statistics

Introduction

Human perception of orientation is shaped by spatial context. For example, in the tilt illusion, the perceived orientation of a center stimulus can be altered by the presence of an oriented surround (Gibson & Radner, 1937). Explanations of the tilt illusion have primarily focused on how the surround context changes the neural representation of orientation. Research has shown that neurons in the early visual cortex suppress their responses to, and shift their tuning preferences away from, the contextual orientation (Dragoi, Rivadulla, & Sur, 2001; Benucci, Saleem, & Carandini, 2013). Other studies have demonstrated that neural responses evoked by stimuli within the receptive field (RF) exhibit complex dependencies on content outside the RF (Angelucci & Bressloff, 2006).

Connecting these neural findings to perceptual behavior is challenging, however, due to the lack of a coherent theoretical framework. More broadly, there is still a need for a functional and teleological account of the tilt illusion. In this study, we aim to address these issues by studying the tilt illusion during simultaneous measurements of behavior and neural activity. We interpret these data from an efficient coding perspective, which relates changes in neural representation to natural scene statistics. We also demonstrate how the same efficient coding principle applies to artificial neural networks trained to perform tasks involving natural images.

Methods

Experimental Design Subjects (n = 10) performed a delayed orientation estimation task conducted in the fMRI scanner (Fig. 1). On each trial, an orientation stimulus was presented for 1.5 sec within an annular surround of either non-oriented noise, or gratings with one of two fixed orientations ($\pm$ 35 degrees off vertical). After a blank delay period, a line probe appeared, and subjects used a two-button pad to rotate the probe and report their estimates. Each subject completed a total of 1,200 trials (400 trials for each surround condition).

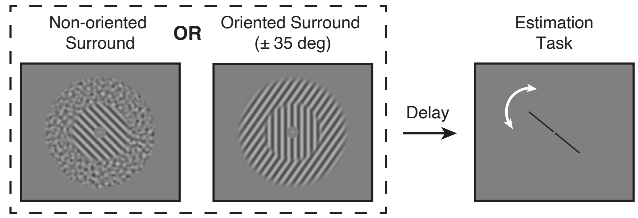

Figure 1: Subjects performed a orientation estimation task of garting stimuli, presented within an annular surround of either non-oriented noise, or one of two fixed orientations.

Theoretical Framework We model orientation perception as an encoding-decoding process (Stocker & Simoncelli, 2006). Stimulus orientation θ is encoded as a noisy neural measurement m, described by the encoding model $p(m|\theta)$. The Fisher Information (FI) of the encoding is defined as $J(\theta) = E[(\frac{\partial}{\partial\theta}\log p(m|\theta))^2|\theta]$, which quantifies the precision of the encoding as a function θ. Importantly, it has been suggested that for a neural population that encodes information efficiently under resource constraints, there is a direct relationship between the stimulus prior $p(\theta)$ and the FI of encoding as $p(\theta) \propto \sqrt{J(\theta)}$ (Wei & Stocker, 2015, 2016). In our analysis, we infer the FI of orientation encoding independently from the behavioral data (referred to as Behavioral FI); fMRI neural data (referred to as Neural FI); and convolutional neural networks (referred to as ConvNet FI). We will examine how FI varies between surround conditions, and how it relates to orientation priors of the natural environment.

Data Analysis *Behavioral FI:* Given an encoding model $p(m|\theta)$ with FI $J(\theta)$, the Cramer-Rao Lower Bound states the following: $J(\theta) \geq [1+b'(\theta)]^2/\sigma^2(\theta)$ (Casella & Berger, 2021). Here $b(\theta)$ is the bias, and $\sigma^2(\theta)$ is the variance. We assume the lower bound is tight, allowing us to infer FI from the bias and variance of subjects' response data (Noel, Zhang, Stocker, & Angelaki, 2021). *Neural FI:* We model the voxel activity based on a probabilistic encoding model developed previously (Van Bergen, Ji Ma, Pratte, & Jehee, 2015). The model defines a multivariate normal distribution over voxel responses $N(\mathbf{m};\ \mu(\theta), \Omega)$. To obtain the neural FI, we directly compute the average negative second derivative of the log-likelihood function using the encoding model. A cross-validation procedure was employed where the model parameters and the FI were estiamted from separate trials. *ConvNet FI:* We denote the responses at a given layer of a convolutional neural network (CNN) to an orientation stimulus as $r = f(\theta)$. The ConvNet FI for independent Gaussian noise is $J(\theta) = ||\partial f/\partial\theta||_2^2$ (A. Benjamin, Qiu, Zhang, Kording,

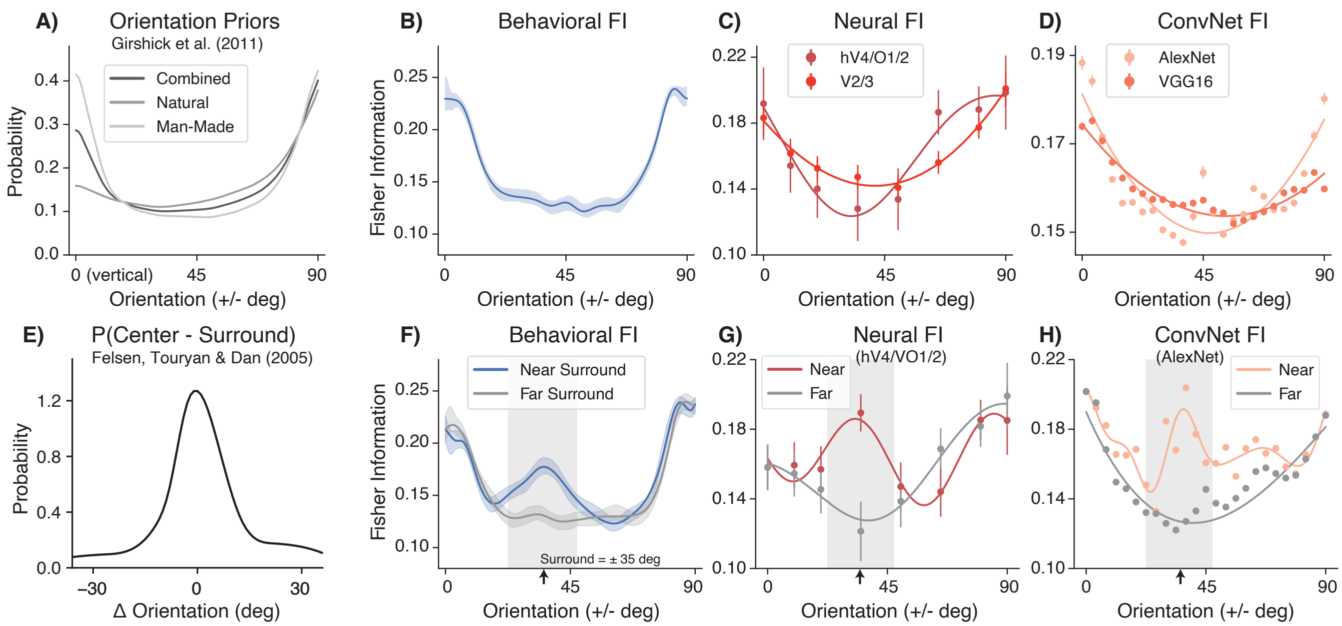

Figure 2: Orientation priors and orientation encoding FI. **A)** Orientation prior measured from photographic images of different environments, reproduced from Girshick et al. (2011). **B)** Behavioral FI, **C)** Neural FI of two ROIs in early visual cortex, and **D)** ConvNet FI, for stimulus with non-oriented surround. **E)** Distribution of the differences in orientation between the center and surround regions in natural images reproduced from Felsen et al. (2005). **F) - H)** Comparison of the **F)** Behavioral, **G)** Neural, and **H)** ConvNet encoding FI between the near-surround side (colored line) and far-surround side (gray line) in the oriented surround condition. Solid lines are obtained from Fourier series fits. Shaded areas and error bars indicate +/- SEM.

& Stocker, 2019). Lastly, to compare FI directly between conditions, and with orientation priors, we report the normalized, square root of FI: $\tilde{J}(\theta) = \sqrt{J(\theta)}/\int_{\theta}\sqrt{J(\theta)}d\theta$, which can be interpreted as an orientation prior inferred from neural coding.

Results

We first examine orientation encoding FI in the non-oriented surround condition. Note that we assumed vertical symmetry and combined the data from the negative and positive range. We found that orientation encoding at baseline is anisotropic (Fig. 2B): Behavioral FI is highest at the cardinal orientations, and lowest at the obliques. This pattern of non-uniform encoding is also reflected in the neural activity in areas of the early visual cortex (Fig. 2C). Consistent with previous studies (Henderson & Serences, 2021; A. Benjamin et al., 2019; A. S. Benjamin, Zhang, Qiu, Stocker, & Kording, 2022), a similar pattern of encoding was obtained from two CNN models trained on natural images (Fig. 2D). Thus, as predicted by the efficient coding hypothesis, the encoding FI resembles the environmental priors of orientation (Fig. 2A).

We now consider the tilt illusion by examining encoding in the oriented surround conditions. At the behavioral level, a surround context induces a strong repulsive bias near the surround orientation and a subtle attraction further away in the perceptual estimates (Clifford, 2014). To understand the changes in encoding, we compare the encoding FI close to the surround orientation (near surround) to that further away from the surround (far surround). We observed a significant increase in behavioral FI close to the surround orientation, while the overall pattern, such as the cardinal bias, remains unchanged (Fig. 2F). Consistent with the behavioral measure, we found a significant effect of surround modulation in the hV4/VO1/2 areas (Fig. 2G). A smaller, but still significant, effect was also observed in V2/V3. Lastly, when presented with stimuli in the oriented surround condition, a comparable pattern of surround modulation was revealed in the CNN models (Fig. 2H; similar results were obtained in VGG16).

Why should the visual system increase encoding precision close to the surround orientation? Efficient coding theory suggests that the increase in encoding FI should correspond to an increase in the probability of orientations. Indeed, we found that the change in encoding FI closely resembles the probability distribution of orientation differences between center and surround regions in natural images (Fig. 2E). That is, when a particular surround orientation is observed, it signals a significant increase in the probability of the center orientation being similar. Therefore, more encoding resources should be allocated to orientations closer to the surround.

Summary

Our results suggest that the tilt illusion emerges naturally from a dynamic coding strategy that efficiently reallocates neural coding resources based on the spatial stimulus context. Furthermore, the same mechanism may also be applicable to artificial systems trained to process natural images. Future work will aim to generalize our finding to other contextual phenomenon, such as the tilt aftereffect (Schwartz, Sejnowski, & Dayan, 2009).

Acknowledgments

This work was funded by a Collaborative Research Grant from MindCORE of the University of Pennsylvania. L.-Q.Z. is supported by the Howard Hughes Medical Institute. We thank Huseyin O. Taskin and Marlie Tandoc for their assistance with fMRI data collection.

References

Angelucci, A., & Bressloff, P. C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate v1 neurons. *Progress in brain research*, *154*, 93–120.

Benjamin, A., Qiu, C., Zhang, L.-Q., Kording, K., & Stocker, A. (2019). Shared visual illusions between humans and artificial neural networks. In *2019 conference on cognitive computational neuroscience* (Vol. 10, pp. 2019–1299).

Benjamin, A. S., Zhang, L.-Q., Qiu, C., Stocker, A. A., & Kording, K. P. (2022). Efficient neural codes naturally emerge through gradient descent learning. *Nature Communications*, *13*(1), 7972.

Benucci, A., Saleem, A. B., & Carandini, M. (2013). Adaptation maintains population homeostasis in primary visual cortex. *Nature neuroscience*, *16*(6), 724–729.

Casella, G., & Berger, R. L. (2021). *Statistical inference*. Cengage Learning.

Clifford, C. W. (2014). The tilt illusion: Phenomenology and functional implications. *Vision research*, *104*, 3–11.

Dragoi, V., Rivadulla, C., & Sur, M. (2001). Foci of orientation plasticity in visual cortex. *Nature*, *411*(6833), 80–86.

Felsen, G., Touryan, J., & Dan, Y. (2005). Contextual modulation of orientation tuning contributes to efficient processing of natural stimuli. *Network: Computation in Neural Systems*, *16*(2-3), 139–149.

Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. i. quantitative studies. *Journal of experimental psychology*, *20*(5), 453.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature neuroscience*, *14*(7), 926–932.

Henderson, M., & Serences, J. T. (2021). Biased orientation representations can be explained by experience with nonuniform training set statistics. *Journal of vision*, *21*(8), 10–10.

Noel, J.-P., Zhang, L.-Q., Stocker, A. A., & Angelaki, D. E. (2021, 05). Individuals with autism spectrum disorder have altered visual encoding capacity. *PLOS Biology*, *19*(5), 1-21. doi: 10.1371/journal.pbio.3001215

Schwartz, O., Sejnowski, T. J., & Dayan, P. (2009). Perceptual organization in the tilt illusion. *Journal of Vision*, *9*(4), 19–19.

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature neuroscience*, *9*(4), 578–585.

Van Bergen, R. S., Ji Ma, W., Pratte, M. S., & Jehee, J. F. (2015). Sensory uncertainty decoded from visual cortex predicts behavior. *Nature neuroscience*, *18*(12), 1728–1730.

Wei, X.-X., & Stocker, A. A. (2015). A bayesian observer model constrained by efficient coding can explain'anti-bayesian'percepts. *Nature neuroscience*, *18*(10), 1509–1517.

Wei, X.-X., & Stocker, A. A. (2016). Mutual information, fisher information, and efficient coding. *Neural computation*, *28*(2), 305–326.