Graduate School of Computational Neuroscience
Graduate Training Center of Neuroscience
University of Tübingen

Neural Circuit Dynamics of Primary Visual Cortex

Testing Neural Circuit Mechanisms Underlying Visual Illusions of Texture Border and Size

Laboratory Report

presented by

Kai Rothe

The study was supervised by

Prof. Dr. Li Zhaoping

Dpt. for Sensory and Sensorimotor Systems
Max-Planck Institute for Biological Cybernetics

Duration of the lab rotation: 9.10.2025 - 23.12.2025

Deadline for submission: 03.01.2026

Abstract

How does the human brain recognize and localize objects? One strategy is to localize the borders between distinct visual textures. Li's (2000) neural circuit model of primary visual cortex (V1) proposes that texture borders are localized through lateral neural interactions in V1, specifically iso-oriented inhibition and co-linear excitation. However, Popple (2003) measured border localizations which are inconsistent with the biases predicted by the V1 model. So far, it remained unclear if the limitations of Popple's psychophysical experiment, specifically central viewing and prolonged stimulus duration of 150 ms, allowed confounding feedback from visual areas beyond V1. In this experiment we test if the biases in localizing borders between orientation textures contradict the V1 model even in the case of stimuli presented for durations as short as 20 ms in peripheral vision. We replicate and extend Popple's experiment, and compare our results with predictions simulated by a replicated version of the V1 model. The biases measured in this experiment might falsify the V1 model explanation and support alternative neural mechanisms of visual texture segmentation. Else, the V1 model and our results predict a new visual size illusion inverse to the well-known Helmholtz illusion (1867). Future comparative study of both illusions might open up a new experimental and theoretical avenue to understand the contributions of V1 and higher visual areas to not only visual texture segmentation but also size perception.

1 Introduction

Background How does the human brain recognize and localize objects? One strategy is to localize the borders between distinct visual textures. A neural circuit model of primary visual cortex (V1) by [1] suggests that texture borders can be localized through lateral neural interactions in V1, specifically iso-oriented inhibition and co-linear excitation. However, [2] measured border localizations which are inconsistent with the biases predicted by the V1 model. So far, it remained unclear if the prolonged viewing duration of 150 ms, without visual backward masking, in [2]'s experiment allowed confounding feedback from visual areas beyond V1 [3]. And according to the Central-Peripheral-Dichotomy, central viewing of the stimuli in [2]'s experiment allows further confounding feedback [4].

Objectives The objective of this lab rotation was to replicate the V1 model by [1]. I subsequently aimed to adapt [2]'s experiment to shorter stimuli durations viewed in the periphery and with visual backward masking. In a pilot experiment, I collected preliminary data on myself to test the experimental design and falsify predictions simulated with the replicated V1 model.

2 Methods

V1 model replication The model was implemented using the documentation of the original model [5, 6] and is available at `https://github.com/iakioh/V1SH_model`. For simulations I used the Euler method with a simulation time of $T = 12\alpha$, where α denotes the membrane time constant, and a typical step size of $\Delta t = 0.01\alpha$. I used the same model parameters as the original model, with one exception: The orientation-unspecific normalization for an excitatory pyramidal cell at location i and tuned to orientation θ (following the original notation) was computed as

$$I_o = 0.85 - 2.0 \left[\frac{\sum_{j \in S_i} \sum_{\theta'} g_x(x_{j\theta'})}{C} \right] \tag{1}$$

with $C = 16$ following the original implementation of the V1 model, not the documentation where $C = \sum_{j \in S_i} 1 = 25$ when using a Manhatten grid for sampling visual input. Simulations started from the initial condition

$$y_0 = I_c/\alpha = 1.0 \tag{2}$$

$$x_0 = \left(I_o - g_y(y_0) \cdot \sum_{\theta'} \psi(\theta') \right) / \alpha = -1.65 \tag{3}$$

which corresponds to the neural resting state with neither visual input nor top-down feedback.

V1 model predictions The saliency of a vertical border between two orientation textures was defined as the highest saliency of a grid column close, i.e. of maximally 3 columns horizontally distant, to the horizontal midpoint between two orientation textures. The saliency of a grid column was computed as the average saliency of all column locations. The saliency of a location is defined as the highest (temporally averaged) pyramidal response to visual input at this location [6]. The saliency of a border was normalized by z-scoring over all locations i of an input texture.
The predicted localization bias for a vertical border was defined as the horizontal distance between the most salient column close to the border and the horizontal midpoint between the textures. By convention, the bias is positive if the most salient column is part of the texture oriented parallel to the border, i.e. here always vertical, and else negative.

Experimental design The experiment, see Fig. 2.A, was implemented in PsychoPy [7] and is available at `https://github.com/iakioh/V1SH_experiment`. In the pilot experiment one subject, the experimentator himself, with visual acuity corrected to normal, participated. Stimuli were generated by and viewed on a HP ProBook 450 G8 Notebook PC (60 Hz frame rate), from an approximate distance of 60 cm. Each trial started with a text prompting the subject to press any button to start the next trial. The text was shown at the same location as the fixation target appearing after a button press. The text was framed by a white ellipsoid outline of 10 width and 5 height, which was shown on screen throughout the whole remaining trial. The subsequent fixation target was a white filled circle with diameter of 0.1 and was presented for 700 ms. The stimulus consisted of Gabor patches with luminance $l \propto \exp(-(x^2+y^2)/2\sigma^2) \cdot \sin(2\pi f x)$ from mid-grey background (x: displacement from Gabor center along axis of sinusoidal modulation, y: along perpendicular axis), with 100% Mitchell contrast, a frequency of $f = 3$ cpd and a standard deviation of $\sigma = 0.16$ [1]. The Gabor patches were aranged in a square grid texture with a spacing of 6σ. The number of rows varied randomly between 1 and extending over the whole screen vertically, here around 6 and 11 rows for the upper and lower texture respectively. The texture extened over over the whole screen horizontally, here around 32 columns. The position of the center of each Gabor patch was uniformly jittered by up to $\pm 0.6\sigma$ in both vertical and horizontal direction. Two orientation textures vertically separated by 1.72 were shown. In the center of each texture was a vertical border defined by a 90 orientation contrast. The upper and lower textures were mirror images of one another, e.g. if upper texture contained vertical patches left of border and horizontal patches right of border, then the lower texture contained horizontal patches left of border and vertical patches right of border. The orientation of Gabor patches varied randomly between 0 vs. 90. The horizontal displacement of upper texture relative to lower texture was shifted between trials. The border was located in the right or left upper peripheral visual field of the subject fixating on the fixation target, shifted by an eccentricity of approx. $+5$ vertically (to avoid any confounding effects of the blind spot) and ± 10 horizontally from fixation. The fixation target remained on screen on top of stimulus, as well as the white ellipsoid outline framing a stimulus exclusion zone to avoid

[1] The size is motivated by [2] scaled up to account for lower resolution in the periphery by a factor of $(1 + \frac{e}{e_2}) \approx 4$ (rounded up) for border eccentricity $e = \sqrt{e_x^2 + e_y^2} \approx 11$ and $e_2 = 3.3$ [4]

saliently appearing stimuli within central vision. The stimulus was presented in this pilot experiment for $100ms$, followed by a visual backward mask. The mask was a grid of Gabor patches of identical position as the stimulus, but with randomized orientations. The mask was shown until the subject performed the task. The task was to indicate on each trial whether the upper border of the shown stimulus is located left or right of the lower border through pressing the "←" or "→" key respectively. Trials were blocked in sessions of maximally 15 minutes duration, following 12 instruction and and maximally 30 practice trials. During one session, all trials showed stimuli with the border location only in left or only in the right peripheral field. In total 8 sessions with 2823 test trials were collected, with 1411 trials for stimulus with textures consisting of single rows, and 1412 trials for stimulus with textures consisting of multiple rows extending the screen. The horizontal displacement of the upper texture border was selected during the experiment through the adaptive staircase procedure QUEST+ [**watson_questplus_2017**, **python_questplus_package**]. QUEST+ scanned horizontal displacements in $[-3, 3]$ in steps of 0.25 columns. A uniform prior and a binomial likelihood Binom(p) parametrized by a psychometric cumulative normal distribution

$$p(d|\mu, \sigma, \delta_{left}, \delta_{right}) = \delta_{left} + (1 - \delta_{left} - \delta_{right}) \int_{-\infty}^{d} \mathcal{N}(x|\mu, \sigma) dx \tag{4}$$

where d denotes the horizontal displacement towards the vertically oriented upper texture, μ the mean or threshold, σ the standard deviation or slope, and $\delta_{left}, \delta_{right}$ the left and right asymptote, also called lapse or guess rates. The posterior was computed for parameters $\mu \in [-2.5, 2.5]$ in steps of 0.25 columns, $\sigma \in [0.1, 5]$ in steps of 0.1 columns, and assuming symmetric lapse rates of $\delta_{left} = \delta_{right} \in [0, 0.45]$ in steps of 0.05. QUEST+ updates the posterior with the collected stimulus-response pair after each trial, and selects the horizontal displacement for the next trial which maximizes the expected information gain (i.e. minimizes the expected entropy of the posterior after the upcoming trial). I used two QUEST+ staircases per session, one for each number of row condition. In the last two sessions, the staircases were initialized using the previous session data.
- pooled all data of the same number of texture rows over left and right peripheral field
- Maximum likelihood fit of responses in dependence of horizontal distance of texture borders with Cumulative Gaussian distribution $\mathcal{N}(\mu, \sigma)$ - PSE -> bias $b = -\mu/2$ - bootstrapping for confidence intervals - one-sided permutation test to compare parameters fitted to multiple vs single texture row

3 Results

1. Model replication: - successful replication of V1 model verified by visualization of all model parameters, including connections, and comparison with expectation and original model [5] as well as replication of - model responses to calibration inputs, see Fig. 1.II, - filling-in and avoiding leaking-out [6, 8] - temporal evolution of responses to texture border [6], - the figure ground and medial axis effect [6, 9] - summation curve for grating disc inputs showing cross-orientation enhancement [6, 9], - dynamical properties of reduced two-point EI network [6, 10]

2. Predictions: - see Fig. 2.B, for medium to high input intensities - for higher number of rows border localization bias positive, i.e. towards texture parallel to border, confirming simulation by [1], but negative for lower number of rows; confirming simplified argument by [2] - additionally, saliency of border higher for higher number of rows, because for lower rows co-linear excitation of texture vertically oriented to border dominates over iso-oriented inhibition between small number of rows. independent of displacement of the texture.
3. Data: - border localization bias towards vertical texture $b \approx 0.42 \pm 0.06$ significantly above zero for textures with multiple rows - low lapse rates $\delta_{left} = 0.00^{+0.04} - 0.00$, $\delta_{right} = 0.00^{+10^{-6}} - -0.00$ and slope $\sigma = 0.8 \pm 0.1$ for textures with multiple rows - significantly higher lapse rates ($p_0 < 10^{-5}$) for single row textures - for $\delta_{right} \neq \delta_{left}$, no significant border localization bias $b \approx 0.1^{+0.5}_{-0.2}$ for single texture row, and significantly lower than for mutliple texture rows ($p_0 \approx 3 \cdot 10^{-4}$) - the latter result depends highly on the fit. assuming $\delta_{right} = \delta_{left}$ leads to a significant bias $b \approx 0.4 \pm 0.2$ for single row texture, not significantly different from multiple texture rows ($p_0 \approx 0.43$).

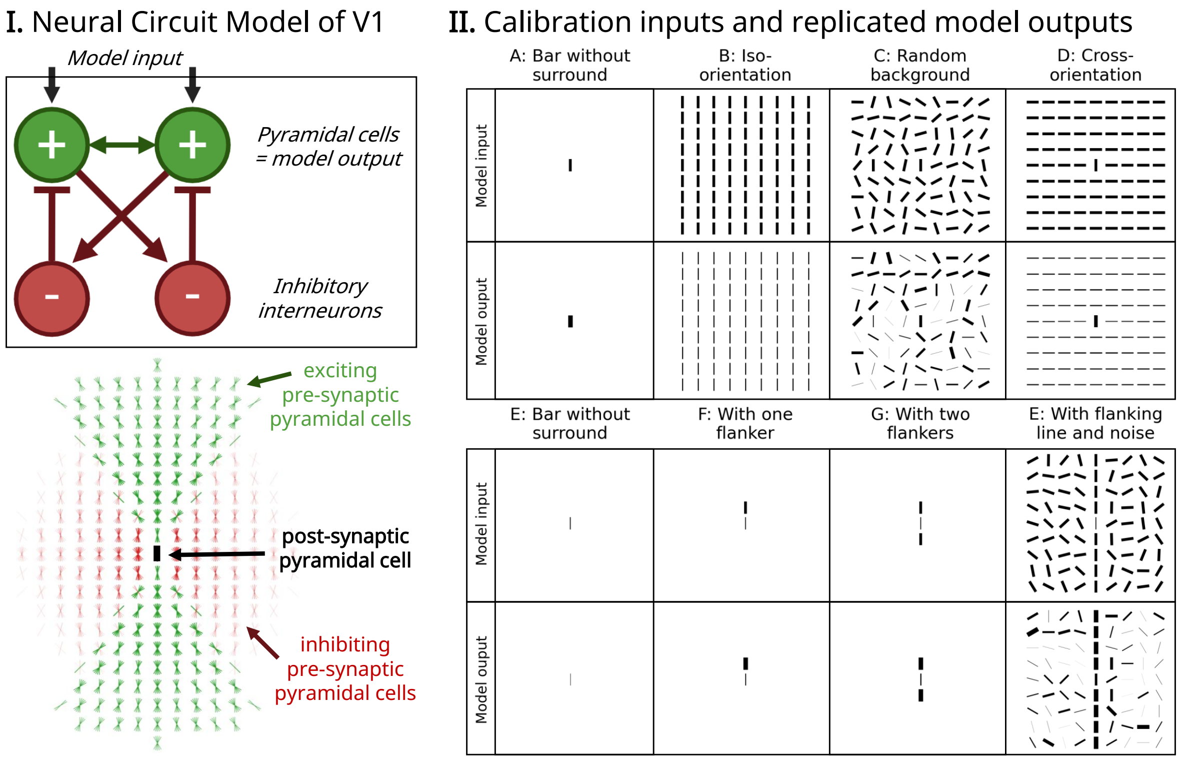

Figure 1: Model replication. I: Sketch of neural circuit model of primary visual cortex (V1). Top: The model consists of two cell types, excitatory pyramidal cells, which are the in- and output of the model, and inhibitory interneurons. Bottom: Orientation and location of bars depict tuning and receptive field of edge / bar detector neurons in V1. The excitatory co-linear and inhibitory iso-oriented interactions mediate contextual influences underlying saliency computations in V1. II: Inputs and replicated outputs for model calibration. A-D show contextual effects dominated by iso-orientation supression. E-H show contextual effects dominated by co-linear facilitation.

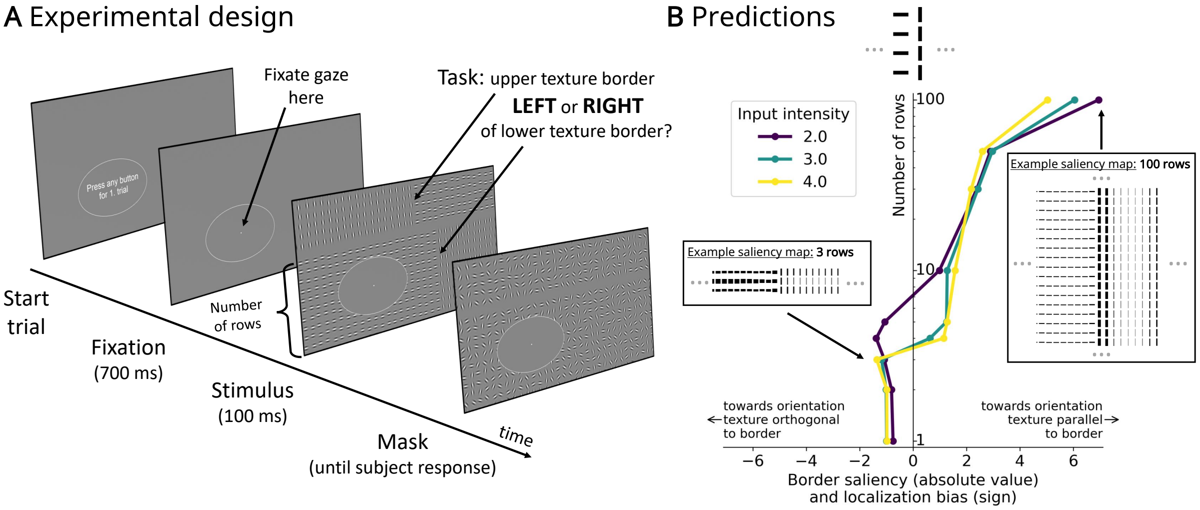

Figure 2: Experimental design and predicitions. A: One example trial of the proposed experiment. After subjects fixate for 700 ms, the stimulus with two texture borders of varying horizontal displacement and number of rows is shown in the periphery for 100 ms. Visual backward masking afterwards reduces confounding extrastriatal feedback. Subjects are forced to decide between two alternative choices: is the upper texture border left or right of the lower texture border? B: Predicted saliency and localization of borders given orientation textures of varying height. The absolute value and sign of the x-axis show border saliency (z-scored, here found to be positive for all predictions) and the direction of the localization bias as indicated by the sketched texture border above the y-axis, respectively. The y-axis shows the number of rows of the input texture, with examplatory predicted saliency maps for the case of 3 and 100 rows. For medium to high input intensity and increasing number of rows, the border saliency increases after a sign switch of the border localization bias. Due to wrap-around boundary conditions and an model input size of 100 rows, 100 rows indicates the case of texture ends beyond the visual input field.

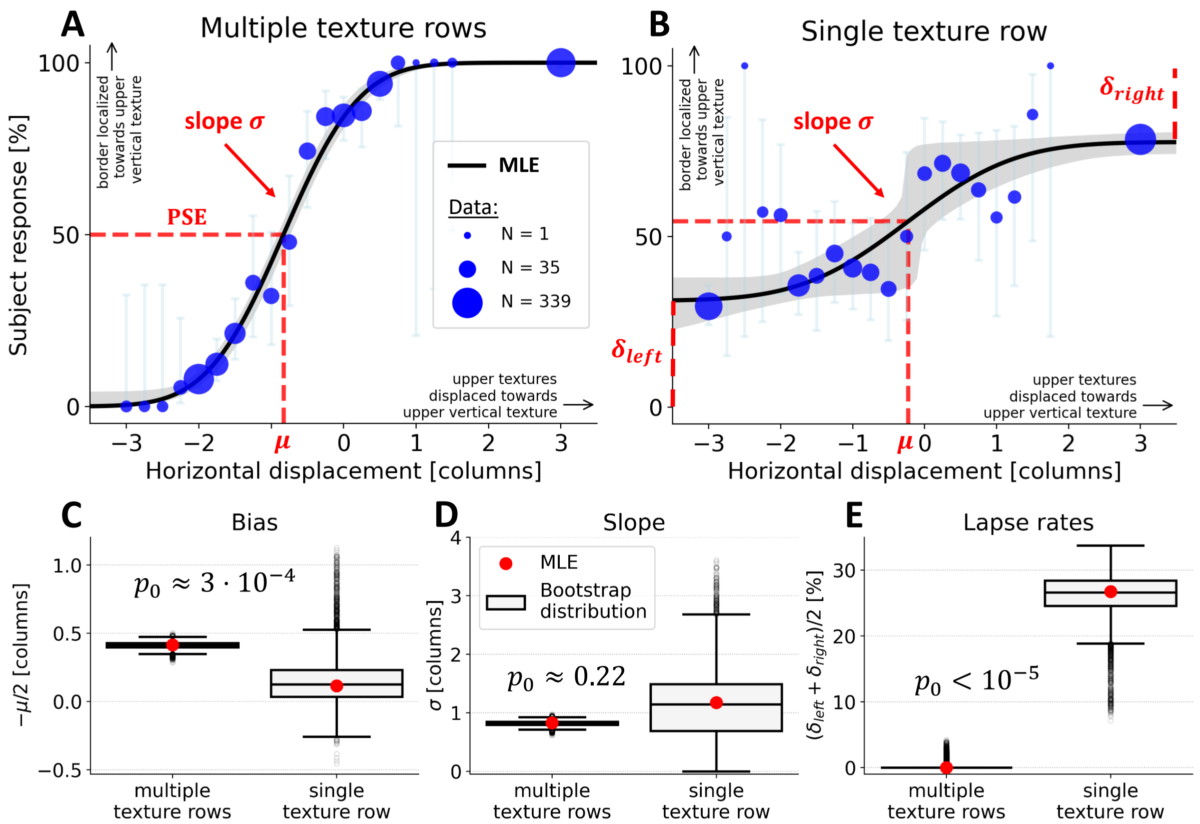

Figure 3: Preliminary experimental results. A-B: Psychometric fit for stimuli textures consisting of multiple rows (A) or a single row (B). The x-axes show the horizontal displacement of the upper textures varied during the experiment, in units of the texture column width. Positive displacements are towards the upper vertical texture. The y-axes show the percentage of trials for which the pilot subject localizes the border towards the upper vertical texture. The data (blue dots; diameter indicating number of trials; error bars indicating Wilson score interval) was fitted via maximum binomial likelihood parametrized by eq. 4 (black line; grey outline indicates bootstrapped 95% CI). The parameters of the fitted psychometric likelihood are indicated in red: asymptotes or lapse rates $\delta_{left}, \delta_{right}$, mean μ at the point of subjective equivalence (PSE) $(1 - \delta_{left} - \delta_{right})/2$ and slope σ. C-E: fitted psychometric parameters of A and B in comparison. The red dots indicate the maximum likelihood estimate (MLE). Boxplots visualize the bootstrap distribution. P-values p_0 are derived from an one-side permutation test under the null hypothesis that textures of multiple and single row lead to interchangable responses. C: bias $-\mu/2$, D: slope σ. E: average lapse / guess rates $(\delta_{left} + \delta_{right})/2$. Important note of caution: the significance and results of B-D highly depend on the fit, especially the assumption $\delta_{left} \neq \delta_{right}$.

4 Discussion

- important findings: - V1 model successfully replicated, with minor improvements to documentation - confirmed switch in localization bias and predicted decreased saliency of texture border - designed and implemented improved experiment to test this prediction - tested experiment in and analyzed data of a pilot study, finding - significant localization border bias towards vertically oriented texture for textures with mulitple rows, in agreement with the predictions by V1 model and with [2] - significantly higher lapse rate for textures of single row, in agreement with lower saliency predicted by V1 model
- limitations: - leaking-out in replicated and original model in dependence of numerical accuracy - predictions made for approximate case, not real stimuli - single subject, preliminary pilot data - open problem: significance of border localization bias for single rows depends highly on assumptions on fit - further, while low lapse rates and slope indicate the task for multiple texture rows to be easy enough for the experimentator, a single session with a naive subject with responses not higher than chance show task to be too difficult to do without training -> train or make easier, e.g. only shifting a single texture. possible confounding factor: non-naive subjects might use the fact that only upper texture border moves and approximate location of lower texture border -> also shift lower texture border
- implications of the work, open questions / future studies: - further improvements of experimental design and data analysis to rebustly estimate bias for single row texture, more subjects, lower viewing durations of up to 20ms - biases and sensitivity might falsify V1 model explanation -> support alternative neural mechanisms, e.g. in higher visual areas - else, assuming visual size estimation consistent with border localization, model predicts visual size illusion inverse to well-known Helmholtz illusion [11] - future comparative study of both illusions for new experimental and theoretical avenue to understand contributions of V1 and higher visual areas to visual texture segmentation and size perception

5 Acknowledgements

- supervision and funding by Li Zhaoping - support by fellow lab members Vladislav Aksiotis, Fani Lohrmann and Maria Pavlovic.

References

[1] Z. Li. “Can VI Mechanisms Account for Figure-Ground and Medial Axis Effects?” In: Thirteenth Annual Neural Information Processing Systems Conference (NIPS 1999). June 2000, pp. 136–142.

[2] Ariella Popple and Zhaoping Li. “Testing a V1 model —- Perceptual biases and saliency effects”. In: *Journal of Vision* 1.3 (Dec. 1, 2001), p. 148.

[3] Y. Liu and L. Zhaoping. "Testing a prediction of the central-peripheral dichotomy in visual inference: visual backward masking is weaker in the peripheral visual field". In: 42nd European Conference on Visual Perception (ECVP 2019). Sept. 2019, p. 188.

[4] Li Zhaoping. "Feedback from higher to lower visual areas for visual recognition may be weaker in the periphery: Glimpses from the perception of brief dichoptic stimuli". In: *Vision Research* 136 (July 1, 2017), pp. 32–49.

[5] Z. Li. "A neural model of contour integration in the primary visual cortex". In: *Neural Computation* 10.4 (May 15, 1998), pp. 903–940.

[6] Li Zhaoping. *Understanding Vision: Theory, Models, and Data*. 1st ed. May 8, 2014.

[7] Jonathan Peirce et al. "PsychoPy2: Experiments in behavior made easy". In: *Behavior Research Methods* 51.1 (Feb. 1, 2019), pp. 195–203.

[8] Zhaoping Li. "Computational Design and Nonlinear Dynamics of a Recurrent Network Model of the Primary Visual Cortex*". In: *Neural Computation* 13.8 (Aug. 1, 2001), pp. 1749–1780.

[9] Li Zhaoping. "V1 mechanisms and some figure–ground and border effects". In: *Journal of Physiology-Paris*. Neuroscience and Computation 97.4 (July 1, 2003), pp. 503–515.

[10] Zhaoping Li and Peter Dayan. "Computational differences between asymmetrical and symmetrical networks". In: *Advances in Neural Information Processing Systems* 11 (1998).

[11] Hermann von Helmholtz. *Handbuch der physiologischen Optik*. In collab. with University of California Libraries. 1867. 926 pp.