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 [**li_can_2000**] suggests that texture borders can be localized through lateral neural interactions in V1, specifically iso-oriented inhibition and co-linear excitation (see Fig. 1.I). However, [**popple_testing_2001**] 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 [**popple_testing_2001**]'s experiment allowed confounding feedback from visual areas beyond V1 [**liu_testing_2019**]. And according to the Central-Peripheral-Dichotomy, central viewing of the stimuli in [**popple_testing_2001**]'s experiment allows further confounding feedback [**zhaoping_feedback_2017**].

Objectives　The objective of this lab rotation was to replicate the V1 model by [**li_can_2000**]. I subsequently aimed to adapt [**popple_testing_2001**]'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 V1 model (see Fig. 1.I) was implemented using the documentation of the original model [**li_neural_1998**, **zhaoping_understanding_2014**] 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 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 $C = 16$ as in 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 [**zhaoping_understanding_2014**]. 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 vertically oriented texture, and negative if the most salient column is part of the horizontally oriented texture.

Experimental stimuli   The experiment, see Fig. 2.A, was implemented in PsychoPy [**peirce_psychopy2_2019**] 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 horizontally and vertically oriented 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 made up "orientation textures", i.e. Gabor patches of equal orientation were arranged in square grids (with spacing of 6σ). In total, the stimulus consisted of four orientation textures, referred to by their relative location on the display: one upper left, one upper right, one lower left and one lower right texture. The upper and lower textures were vertically separated by 1.72. The patches in the left and right textures were of opposite orientation, thus separated by a vertical border defined by a 90 orientation contrast. The upper and lower textures were mirror images of one another, e.g. if the upper right texture contained vertical patches, so the upper left texture contained horizontal patches, then the lower left texture contained vertical patches and the lower right texture horizontal patches; and vica verca. Together, the textures extended over the whole screen horizontally, here around 32 columns. The number of rows varied randomly between 1 or extending over the whole screen vertically, here around 6 and 11 rows for the upper and lower textures respectively. The position of the center of each Gabor patch was uniformly randomized by up to $\pm 0.6\sigma$ in both vertical and horizontal direction. The horizontal displacement of the upper textures relative to the lower

---

[1] The chosen size of the Gabor patches is motivated by [**popple_testing_2001**] scaled up to account for lower peripheral visual acuity by a factor of $(1 + \frac{e}{e_2}) \approx 4$ (rounded up) for $e_2 = 3.3$ and eccentricity of the border $e = \sqrt{e_x^2 + e_y^2} \approx 11$ [**zhaoping_feedback_2017**]

textures 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, with an eccentricity of approx. $e_y = +5$ vertically (to avoid any confounding effects of the blind spot) and $e_x \pm 10$ horizontally from fixation. The fixation target remained on screen on top of the stimulus, together with the white ellipsoid outline framing a stimulus exclusion zone to avoid saliently appearing stimuli within central vision. The stimulus was presented 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.

Experimental task   The task was to indicate on each trial whether the border between the upper left and right texture is located "left" or "right" of the border between the lower left and right texture through pressing the "←" or "→" key respectively. Trials were blocked in sessions of maximally 15 minutes duration, following 12 instruction and maximally 30 practice trials per session. During one session, all trials showed stimuli with the border location only in the 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 over the whole screen.

Adaptive staircase   The horizontal displacement of the upper texture border was selected during the experiment with the adaptive staircase procedure QUEST+ [**watson_questplus_201**
python_questplus_package]. I used two QUEST+ staircases per session, one for all responses and stimuli consisting of multiple texture rows (condition $i = 1$) and one for all responses and stimuli consisting of single texture rows (condition $i = 2$). QUEST+ scanned horizontal displacements $d_i \in [-3, 3]$ columns, in steps of 0.25 columns. For each staircase the response r_i is defined as 1 if the subject located the texture border towards the upper vertical texture, i.e. if the upper right texture is vertical and the subject responds "right" or if the upper left texture is vertical and the subject responds "left". The response is $r_i = 0$ if the subject pressed the key in the opposite direction towards the upper horizontal texture. Then, each staircase used a binomial likelihood $r_i \sim \text{Binom}(p(d_i))$ parametrized by the psychometric function

$$p(d_i|\mu_i, \sigma_i, \delta_{lower,i}, \delta_{upper,i}) = \delta_{lower,i} + (1 - \delta_{lower,i} - \delta_{upper,i}) \cdot \psi(d_i|\mu_i, \sigma_i) \tag{4}$$

$$\psi(d|\mu_i, \sigma_i) = \int_{-\infty}^{d} \mathcal{N}(x|\mu_i, \sigma_i)dx, \tag{5}$$

where μ_i denotes the mean, σ_i the standard deviation and $\delta_{lower,i}, 1 - \delta_{upper,i}$ the lower and upper asymptotes, which are non-zero if there are random errors or guesses. The posterior was computed assuming a uniform prior for parameters $\mu_i \in [-2.5, 2.5]$ in steps of 0.25 columns, $\sigma_i \in [0.1, 5]$ in steps of 0.1 columns, and symmetric lapse rates of $\delta_{lower,i} = \delta_{upper,i} \in [0, 0.45]$ in steps of 0.05. After each trial, QUEST+ updates the posterior with the collected displacement-response pair and selects the horizontal displacement for the next trial which maximizes the expected posterior information gain. In the last two sessions, the staircases were initialized with all displacement-response

pairs from previous sessions.

Data analysis   Data from all trials with stimuli of equal number of texture rows, regardless of the border being shown in the left or right peripheral visual field, were pooled together. As shown in Fig. 3, for each texture row condition $i \in \{1, 2\}$, psychometric parameters μ_i, σ_i, $\delta_{lower,i}$ and $\delta_{upper,i}$ were fitted to the relationship between responses $r_i \in \{0, 1\}$ and horizontal displacement d_i by maximizing the likelihood parametrized by eq. 4

$$\prod_{n=1}^{N_i} p(d_{i,n}|\mu_i, \sigma_i, \delta_{lower,i}, \delta_{upper,i})^{r_{i,n}} \cdot (1 - p(d_{i,n}|\mu_i, \sigma_i, \delta_{lower,i}, \delta_{upper,i}))^{1-r_{i,n}}. \tag{6}$$

assuming independent binomial-distributed responses for all trials $n \in [1, N_i]$. The point of subjective equivalence (PSE) is defined as the horizontal displacement $d_{PSE,i}$ for which the upper texture border is with equal chance perceived towards the upper vertical texture and in the opposite direction towards the upper horizontal vertical texture (after accounting for guesses or errors independent of the stimulus displacement), $\psi(d_{PSE,i}|\mu_i, \sigma_i) = 0.5$, which implies $d_{PSE,i} = \mu_i$. The border localization bias is therefore $b_i = -\frac{\mu_i}{2}$. 95% confidence intervals for maximum likelihood estimates of the psychometric parameters were estimated by bootstrapping with $B = 10000$ samples. The estimated parameters for single vs. multiple texture rows were compared using a one-sided permutation test.

3   Results

Model replication   The successful replication of the V1 model was verified by visualizing all model parameters, including the neural connections (see Fig. 1.I), and comparison with expectation and original model [**li_neural_1998**]. Further, the following were replicated:

- model responses to calibration inputs, see Fig. 1.II

- filling-in and avoiding leaking-out [**li_computational_2001**, **zhaoping_understanding_2014**]

- temporal evolution of responses to texture border [**zhaoping_understanding_2014**]

- the figure ground and medial axis effect [**zhaoping_v1_2003**, **zhaoping_understanding_2014**]

- the summation curve for grating disc inputs showing cross-orientation enhancement [**zhaoping_v1_2003**, **zhaoping_understanding_2014**]

- dynamical properties of the reduced two-point excitatory-inhibitory network [**li_computational_**
 **zhaoping_understanding_2014**]

Model predictions As shown in Fig. 2.B, for medium to high input intensities, the V1 model predicts a border localization bias towards the vertical texture for high number of texture rows, but a border localization bias in opposite direction towards the horizontal texture for a lower number of texture rows, confirming the simplified argument by [**popple_testing_2001**]. Therefore, the model predicts $b_1 > 0$, $b_2 < 0$ and $b_1 > b_2$. Additionally, the saliency of the texture border is higher for a higher number of texture rows. Intuitively, for lower numbers of texture rows, the co-linear excitation between the horizontal texture elements dominates over the iso-oriented inhibition between vertical texture elements. If a border is less salient, a subject might increasingly guess randomly, i.e. independent of the horizontal displacement. Or a subject might rather rely more on more salient aspects such as co-linearly excited horizontal texture elements further away from the border to perform the task, i.e. potentially dependent on the horizontal displacement. Therefore, for textures with less rows, the model predicts either increased asymptotes, $\delta_{lower,1} + \delta_{upper,1} < \delta_{lower,2} + \delta_{upper,2}$, or a decreased sensitivity for border localization, i.e. increased slope $\sigma_1 < \sigma_2$.

Preliminary data For textures with multiple rows I find a border localization bias towards the vertical texture $b_1 \approx 0.42 \pm 0.06$ significantly above zero, which is in agreement with the V1 model prediction and the results by [**popple_testing_2001**]. Further, low asymptotes $\delta_{left,1} = 0.00^{+0.04}_{-0.00}, \delta_{right,1} = 0.00^{+10^{-6}}_{-0.00}$ and slope $\sigma_1 = 0.8 \pm 0.1$ indicate high saliency of the border for multiple texture rows. I find significantly higher asymptotes ($p_0 < 10^{-5}$) for single row textures (and no significant difference for slopes), also in agreement with the V1 model prediction. Assuming $\delta_{lower} \neq \delta_{upper}$, I find for a single texture row a border localization bias $b_2 \approx 0.1^{+0.5}_{-0.2}$ not significantly larger than zero, and significantly lower than for multiple texture rows ($p_0 \approx 3 \cdot 10^{-4}$). Both results are in agreement with the model predictions, but I find a dependence on the fit. Assuming instead $\delta_{lower} = \delta_{upper}$ leads to a bias $b_2 \approx 0.4 \pm 0.2$ significantly larger than zero for single texture rows, not significantly different from the bias b_1 for multiple texture rows ($p_0 \approx 0.43$). A preliminary comparison of the goodness of fit considering also the different number of fit parameters remains inconclusive, since for single texture rows the Aikake information criterion (AIC) gives a lower value for the $\delta_{lower} \neq \delta_{upper}$ model, but the Bayesian information criterion (BIC) the opposite.

4   Discussion

Summary   In this study, the V1 model by [**li_can_2000**] was successfully replicated, with suggestions for minor improvements of the documentation. For orientation textures with vertical borders, I confirmed that for decreasing number of texture rows, the V1 model predicts a switch of a border localization bias from towards the vertical texture to the horizontal texture. I adapted and implemented an experiment to test this prediction with less confounding top-down feedback from visual areas beyond V1. In a pilot study, I find a significant localization border bias for textures with multiple rows, in agreement with the model predictions as well as the experimental results by [**popple_testing_2001**]. I further find a significantly higher lapse rate for single texture
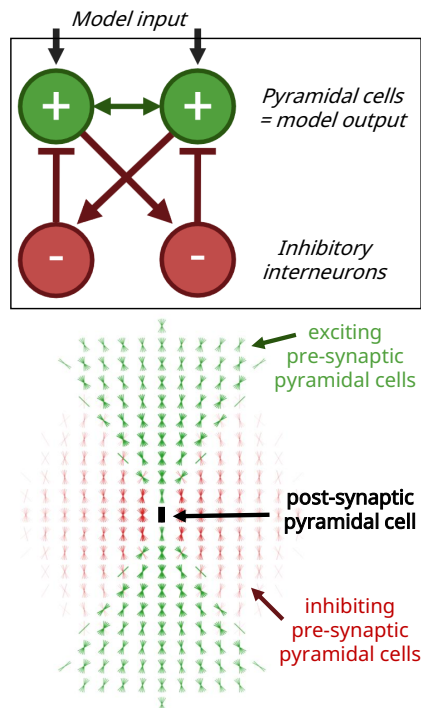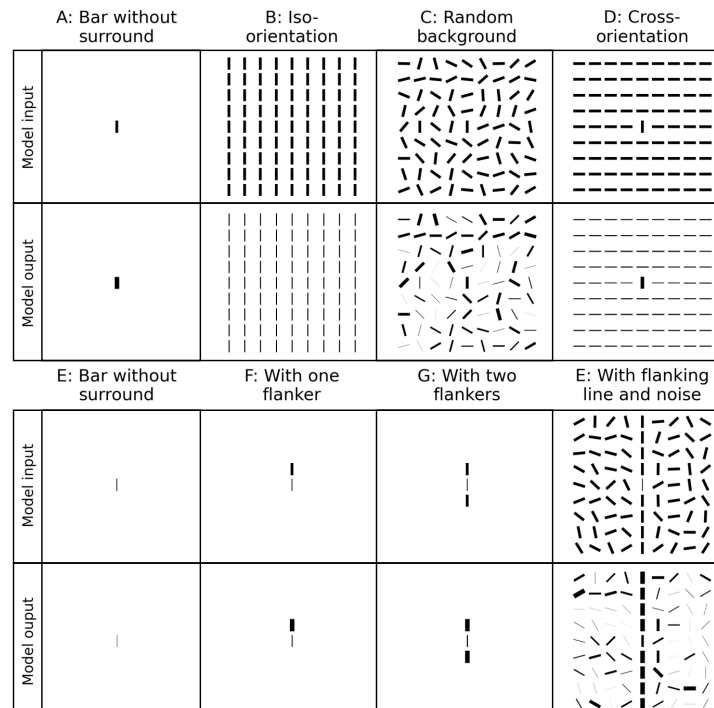
Figure 1: Model replication. I: Sketch of neural circuit model of primary visual cortex (V1). Top: The model consists of two cell types, excitatory pyramidal cells, which are the in- and output of the model, and inhibitory interneurons. Bottom: Orientation and location of bars depict tuning and receptive 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.
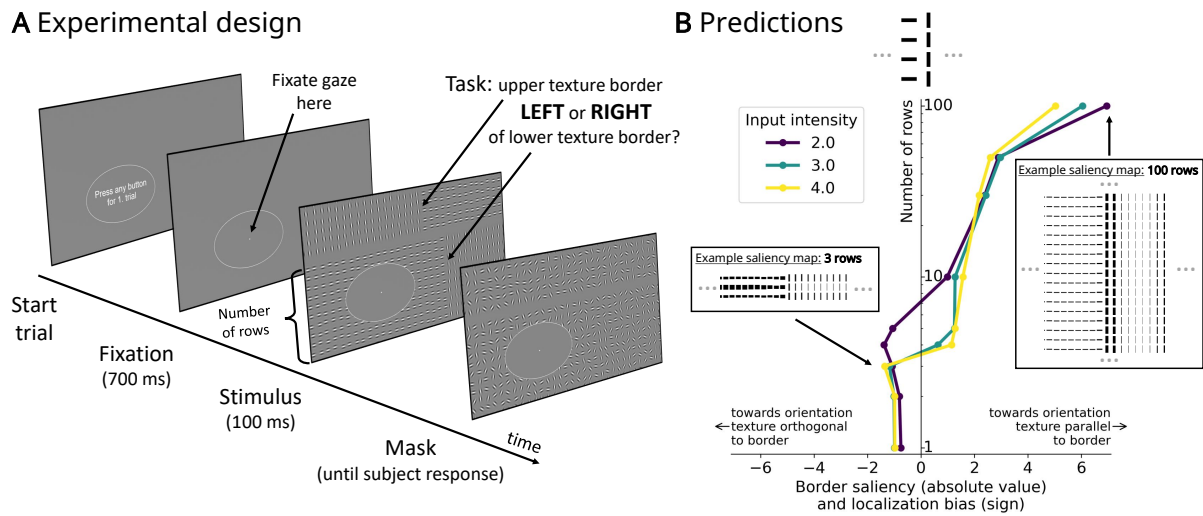
Figure 2: Experimental design and predicitions. A: One example trial of the proposed experiment. After subjects fixate for 700 ms, the stimulus with two texture borders of varying horizontal displacement and number of rows is shown in the periphery for 100 ms. Visual backward masking afterwards reduces 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.
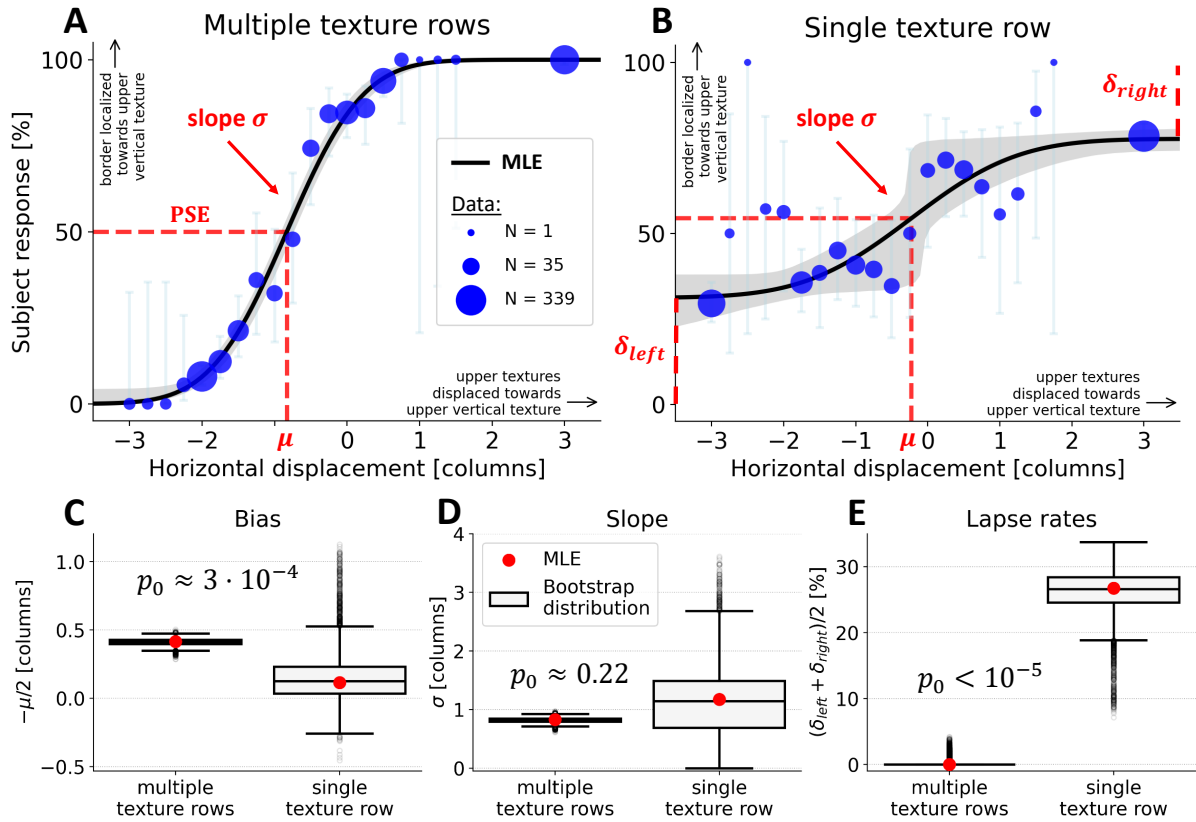
Figure 3: Preliminary experimental results. A-B: Psychometric fit for stimuli textures consisting of multiple rows (A) or a single row (B). The x-axes show the horizontal displacement of the upper textures varied during the experiment, in units of the texture column width. Positive displacements are 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 $\delta_{left} \neq \delta_{right}$.

rows, in agreement with the lower saliency predicted by the V1 model.

Limitations   The V1 model predictions were made using a heuristic read-out of the saliency map for the texture border localization, and only for input with a single texture border, not based on an optimal observer solving the actual experimental task by sampling locations from the saliency distribution predicted by the V1 model for visual input consisting of two texture borders. Data was only collected for a single, non-naive subject. Further, while low lapse rates and precision of the bias measurement indicate that the experimental task, at least for multiple texture rows, was easy enough for one non-naive subject, preliminary data collected in a single session from a naive subject showed responses not higher than chance. This suggests that the task might be too difficult without training.

Further work   Further improvements of the experimental design and data analysis are necessary to robustly estimate the bias for single texture rows. More data collections with multiple subjects are needed. A further decreased viewing durations of up to 20 ms beyond the scope of this study might further decrease confounding feedback. The biases measured in future work might falsify the V1 model explanation and support alternative neural mechanisms of visual texture segmentation. Else, assuming a visual size estimation consistent withe border localization, the V1 model might 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 V1and higher visual areas to not only visual texture segmentation but also size perception.

5   Acknowledgements