



OPEN The effects of distractors on brightness perception based on a spiking network

Weisi Liu^{1,2} & Xinsheng Liu²✉

Visual perception can be modified by the surrounding context. Particularly, experimental observations have demonstrated that visual perception and primary visual cortical responses could be modified by properties of surrounding distractors. However, the underlying mechanism remains unclear. To simulate primary visual cortical activities in this paper, we design a k-winner-take-all (k-WTA) spiking network whose responses are generated through probabilistic inference. In simulations, images with the same target and various surrounding distractors perform as stimuli. Distractors are designed with multiple varying properties, including the luminance, the sizes and the distances to the target. Simulations for each varying property are performed with other properties fixed. Each property could modify second-layer neural responses and interactions in the network. To the same target in the designed images, the modified network responses could simulate distinguishing brightness perception consistent with experimental observations. Our model provides a possible explanation of how the surrounding distractors modify primary visual cortical responses to induce various brightness perception of the given target.

Brightness perception is a fundamental function of the primary visual cortex and has been explored by previous studies^{1–5}. Brightness perception of a given stimulus has been demonstrated to depend on the surrounding contexts^{1,6}. Particularly, it is a common situation that the surrounding context consists of distractors which could modify visual perception of the given stimulus^{7–10}. However, the underlying mechanism of how the distractors modulate visual cortical responses to generate different brightness perception of the same stimulus remains unclear.

In brightness perception of the same stimulus, neural responses are modulated by contexts around the stimulus. Contextual modulation is a common phenomenon during visual perception, including brightness perception^{2,6,11–16}. Perceived brightness of a stimulus is affected by the surrounding context^{2,6}. A recent experimental study has suggested that illusory brightness perception known as simultaneous brightness contrast is associated with low-level visual process prior to binocular fusion². In experiments with complex stimuli, brightness perception of the target is influenced strongly by perceptual organizations and considered to be relative with high-level visual system⁶. The previous study on brightness perception has found that primary visual cortical neurons accomplish spatial integration of contextual information rather than response to illumination of the retina strictly³. Over extracellular recordings made in the retina, the lateral geniculate nucleus and the striate cortex during experiments, neural responses in the striate cortex were found to be correlated with brightness perception under all the designed conditions⁴. Besides brightness perception, contextual modulation also occurs during other visual experiments^{11,12}. Through experiments with oriented flanking stimuli, short-range connections within local neural circuitry have been found to mediate contextual modulation of neural responses to angular visual features¹¹. In experiments, the varying relative contrast of stimuli in neural receptive visual fields can cause contextual effects and impact the neural stimulus-selectivity¹². Contextual modulation also appears in visual experiments of the figure-ground segregation, the shape aftereffect, the neural sensitivity to naturalistic image and the recognitions to associated objects^{13–16}. Particularly, in the surrounding context, properties of distractors have been demonstrated in experiments with influences on neural responses and visual perception^{17,18}. For example, the sizes of distractors, the contrast between distractors and the stimulus, as well as the distance from distractors to the stimulus, have effects on visual perception^{7–9}. Through analyses of visual neural responses and the contrast of images, Rodriguez and Granger present a generalized formulation of visual contrast energy and provide an explanatory framework for performances of identifying targets when distractors have the varying

¹School of Mathematics and Information Sciences, Yantai University, Yantai 264005, China. ²State Key Laboratory of Mechanics and Control of Mechanical Structures, Institute of Nano Science and Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China. ✉email: xsliu@nuaa.edu.cn

flanking distance and number⁷. Through experiments with designed flanking stimuli, Levi and Carney has found that the bigger flanks can lead to the smaller crowding phenomena⁹. Recently, with primary visual cortical neural responses of macaque monkeys to oriented stimuli, the influence of distances between distractors and the visual stimulus has been explained as a kind of spatial contextual effect on primary visual cortical activities⁸. In this way, effects of surrounding distractors on visual perception might be interpreted as a kind of contextual modulation of primary visual neural responses.

The primary visual cortex is a basic biological structure associated with a broad range of visual researches, including sparse responding, object recognition, contextual modulation and brightness perception^{2,4,11–13,19–28}. The unsupervised learning within the primary visual cortex has been considered widely in neuroscience research, including the emergence of visual misperception^{29–32}. In contextual modulation, both neural responses and neural interactions in the primary visual cortex could be modified^{11,13,33}. Based on experimental observations, various models have been presented to explore visual perception. Batard and Bertalmio have improved a type of image processing technique which is based on differential geometry with properties of the human visual system and covariant derivatives, applying the image processing technique in exploration of brightness perception and color images correction⁵. The networks have performed as the effective models in exploring visual recognition and perception^{32,34–36}. Considering interactions between boundary and feature signals in brightness perception, Pessoa, Mingolla and Neumann have developed a neural network which can implement boundary computations and provide a new interpretation of feature signals through the explicit representation of contrast-driven and luminance-driven information³⁵. To simulate the visual cortex combined binocular information, Grossberg and Kelly have designed a binocular neural network which computes image contrasts in each monocular channel and fuses these discounted monocular signals with nonlinear excitatory and inhibitory signals to represent binocular brightness perception³⁶. Considering influences of neural interactions, both contextual modulation and brightness perception in the visual cortex could be explored through the neural networks^{37,38}. To explore the influence of distractors in brightness perception, contextual modulation of primary visual cortical neural responses and interactions might be the possible factors.

Probabilistic inference performs as the feasible exploring method for the mechanism of visual perception and neural coding^{39–44}. Murray has presented a probabilistic graphical model with assumptions about lighting and reflectance to infer globally optimal interpretations of images, exhibiting partial lightness constancy, contrast, glow, and articulation effects³⁹. To estimate brightness perception of images with spatial variation considered, Allred and Brainard have developed a Bayesian algorithm with the prior distributions designed to allow spatial variation in both illumination and surface reflectance and explored changes in image luminance from the aspect of spatial location⁴⁴. Based on probabilistic inference, the winner-take-all (WTA) networks are designed for researches on visual cortical computational functions^{45–48}. For simplification, these WTA networks are designed in the two-layer structure^{45–47}. With neural spiking sequences generated, the plastic weights in these WTA networks have the learning rules as the Hebbian spike-timing-dependent plasticity (STDP) which is a kind of unsupervised learning⁴⁹. With the Hebbian STDP, primary visual cortical sparse responses could be simulated from the WTA network^{28,46}. The temporal structure of the neural spiking train could contribute to the information transformation compared to the rate-based neural coding⁵⁰. The WTA networks consist of excitatory pyramidal neurons with lateral inhibitions, following neural microcircuits observed in layers 2/3⁵¹. Besides, the k-winner-take-all (k-WTA) network could simulate the observed simultaneous activities of multiple neurons in experiments⁴⁵. Constructing the k-WTA spiking networks, our previous studies have explored the primary visual cortical neural responding variability and illusory stereo perception^{52,53}. Yet, these WTA networks have not considered the influence of distractors on cortical neural responses and brightness perception.

In this paper, the plastic two-layer k-WTA spiking network is constructed to explore how varying properties of distractors induce contextual modulation of neural responses to generate different brightness perception of the given stimulus. Connective weights have the STDP learning rule. Afferent neurons in the first layer of the network transform the visual information into feedforward Poisson spiking trains towards the second layer. There are inter-connected excitatory and inhibitory neurons in the second-layer network. Both excitatory and inhibitory neurons in the second layer generate responses in the stochastic methods. The network recognizes outside images depending on second-layer excitatory responses. In simulations, visual images are designed with the given target stimulus and various surrounding distractors. Distractors are designed with different grayscale values, sizes and distances to the target. The influence of one varying property is explored through simulations with other properties fixed. Modulations on neural responses and lateral interactions induced by each property are measured over simulations. Brightness perception is simulated with responses from the network. Our simulations show that, to designed stimuli, varying properties of distractors could modulate second-layer neural responses and lateral interactions, inducing different brightness perception of the same target.

This article is structured as follows. In Materials and methods, the network and the corresponding probabilistic model are introduced. In the section of Results, images with various distractors perform as visual stimuli. For the designed network, neural responding modulations and brightness perception induced by distractors are investigated. With the same target stimulus and the same background, varying distractors can induce modulations of second-layer excitatory neural spiking responses and interactions, leading to different brightness perception. Finally, a conclusion is given.

Materials and methods

The structure of the spiking network. In this subsection, the construction of our k-WTA network was introduced (Fig. 1). The network was designed following the basic neural circuit in the layer 2/3 for simulating primary visual cortical neural activities^{45,51}. In the network, there were N first-layer afferent excitatory neurons, K second-layer excitatory neurons and J second-layer inhibitory neurons. In the second layer, inter-connections

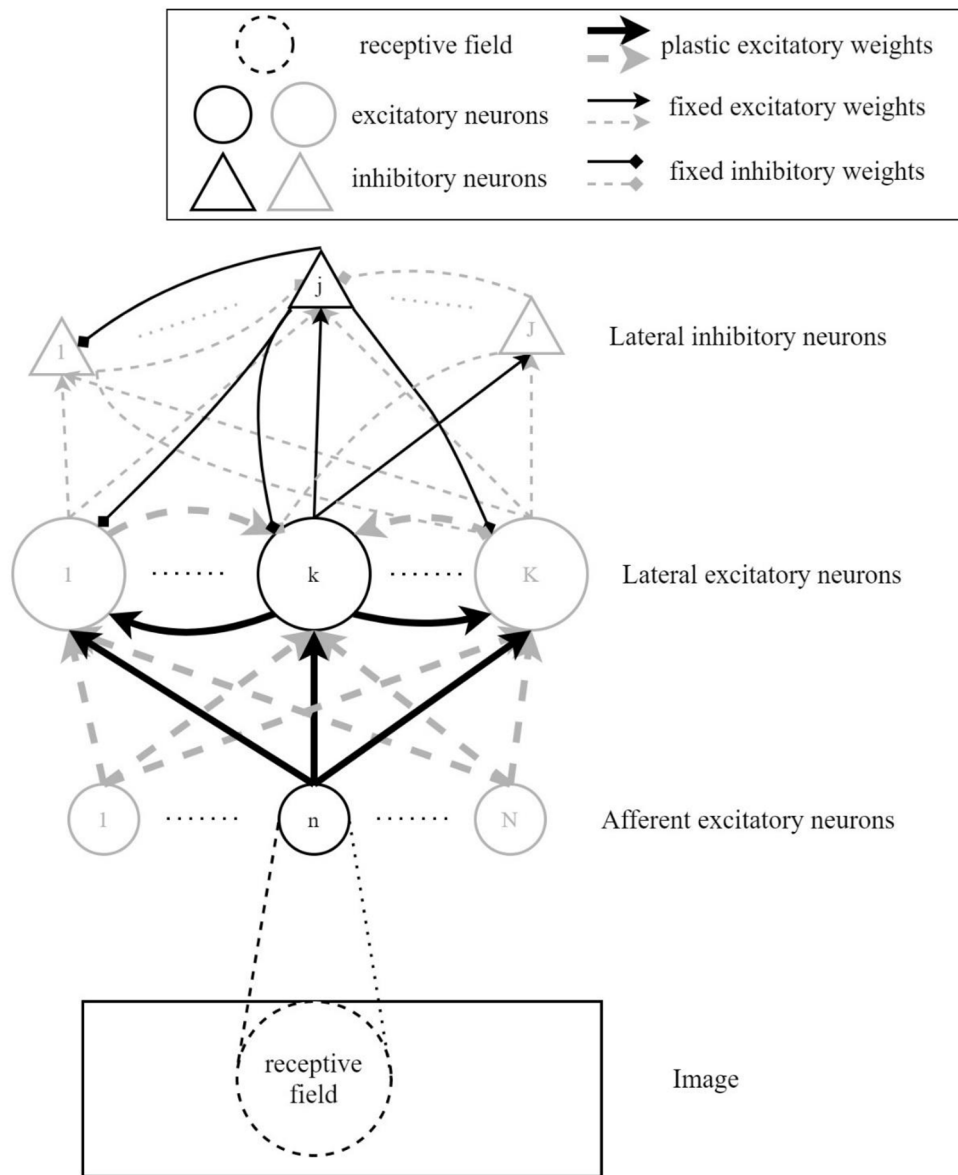


Figure 1. The construction of the designed k-WTA spiking network. Solid circles indicated excitatory neurons. Solid triangles indicated inhibitory neurons. The network contained first-layer afferent excitatory neurons, second-layer excitatory neurons and second-layer inhibitory neurons. Lines with arrows indicated excitatory connections. Lines with squares indicated inhibitory connections. Plastic connections were represented by thick lines. Fixed connections were represented by thin lines. The dotted circular region within the bottom image indicated the afferent neural receptive field. Afferent neurons were induced by stimuli within receptive fields to generate Poisson spikes.

among neurons were designed according to specific connective probabilities in the previous experiment and the study^{45,51}. With connective probabilities, the structure of the network accorded with the observed neural circuit in the layer 2/3. Influences of both neural interactions and connective plasticity in visual perception were considered in the network^{54,55}.

In Fig. 1, the black circle and triangle performed as example neurons while the gray ones indicated the other neurons. The black lines presented for the synaptic connections from the example neurons while the gray lines indicated the other connections. The excitatory and inhibitory connective weights were presented by arrows and squares, respectively. The thin lines indicated the fixed weights. The plastic excitatory connective weights were expressed by the thick lines with arrows.

Each second-layer excitatory neuron received excitatory inputs from both first-layer afferent neurons and other second-layer lateral excitatory neurons. Matrices $W \in R_{K \times N}$ and $V \in R_{K \times K}$ were excitatory feedforward and lateral connective weights, respectively. Each second-layer excitatory neuron had random connections from a subset of second-layer inhibitory neurons. With $p_{EI} = 0.6$ as the special connective probability, each excitatory

neuron was randomly decided to be connected by an inhibitory neuron or not. With the inputs mentioned above, the k th second-layer excitatory neuron had the temporal membrane potential at the t th timestep as:

$$u_{ik}^z = \sum_n w_{kn} \tilde{x}_{tn} + \sum_{k'} v_{kk'} \tilde{z}_{tk'} - \sum_{j \in J_k} v^{EI} \tilde{y}_{tj} + b_k, \tag{1}$$

where u_{ik}^z was the temporal neural membrane potential, $w_{kn} \tilde{x}_{tn}$ and $v_{kk'} \tilde{z}_{tk'}$ determined values of feedforward excitatory postsynaptic potentials (EPSP) induced by the n th first-layer afferent neuron and the k' th second-layer excitatory neuron, respectively. w_{kn} was the excitatory feedforward weight from the n th neuron in the first layer. $v_{kk'}$ was the excitatory lateral weight from the k' th second-layer excitatory neuron. The feedforward weights, as well as the lateral weights between excitatory neurons in the second layer were plastic and limited in $(0, 1)$. The excitability of the k th second-layer excitatory neuron was controlled by the parameter b_k . In our simulations, b_k was sampled from the uniform distribution of $(0, 1)$. Indices of second-layer inhibitory neurons projecting to the k th second-layer excitatory neuron were collected into a set as J_k . At the t th timestep, the j th inhibitory neuron among J_k induced the inhibitory postsynaptic potential (IPSP) as $v^{EI} \tilde{y}_{tj}$. Lateral connective weights projected to excitatory neurons from inhibitory neurons were fixed in our network and expressed as a common parameter v^{EI} . Because excitatory weights were limited in $(0, 1)$, v^{EI} was equal to the mean excitatory connective strength of 0.5 in our simulations. With neural spiking trains from corresponding neurons, \tilde{x}_{tn} , $\tilde{z}_{tk'}$ and \tilde{y}_{tj} were the synaptic traces at the t th timestep. If the k' th second-layer excitatory neuron generated its spikes at $\{t_{k'}^{(1)}, t_{k'}^{(2)}, t_{k'}^{(3)}, \dots\}$, its temporal synaptic trace was given as:

$$\tilde{z}_{tk'} = \sum_f \varepsilon(t - t_{k'}^{(g)}),$$

$$\varepsilon(t - t_{k'}^{(g)}) = \begin{cases} \exp[-(t - t_{k'}^{(g)})/\tau_r] - \exp[-(t - t_{k'}^{(g)})/\tau_f], & \text{if } t \geq t_{k'}^{(g)} \\ 0, & \text{else} \end{cases}, \tag{2}$$

where $\varepsilon(\cdot)$ was the synaptic response kernel, $t_{k'}^{(g)}$ was the generated time of the g th neural spike. The rise-time constant and the fall-time constant in the synaptic response kernel were set as $\tau_r = 1$ timestep, $\tau_f = 10$ timesteps⁴⁵. With the common kernel $\varepsilon(\cdot)$, \tilde{x}_{tn} and \tilde{y}_{tj} in Eq. (1) could be expressed similarly.

At the timestep t , whether a neuron generated a spike or not was presented by another variable as the temporal neural active state. In this paper, 1 ms in experiments was represented by a timestep in simulations. It was assumed in the simulations that only one spike could be generated by each neuron at most during each timestep. Under this assumption, the k th excitatory neural active state at t was defined as a variable $z_{kt} \in \{0, 1\}$. Depending exponentially on the temporal neural membrane potential, the distribution of z_{kt} could be expressed as:

$$p(z_{ik} | u_{ik}^z; \Theta) = \frac{\exp(u_{ik}^z z_{ik})}{\sum_{z'_{ik} \in \{0,1\}} \exp(u_{ik}^z z'_{ik})}, \tag{3}$$

where the neural active state was denoted as z_{ik} at t , u_{ik}^z was the temporal neural membrane potential in Eq. (1). For normalization, another variable z'_{ik} in the denominator took all the possible values of the neural active state. At each timestep t , a sample was generated from the distribution in Eq. (3) as the value of z_{ik} . $z_{ik} = 1$ indicated that the k th excitatory neuron emitted a spike at t . $z_{ik} = 0$ indicated the silence of this neuron. If the k th excitatory neuron generated a spike, it had a refractory period of 5 timesteps and was set to be silent⁵⁶. Neural active states reflected neural spiking responses generated from the network.

The network could simulate identifications of stimuli at each timestep t through the comparisons of the vector of excitatory neural responses $\tilde{z}_t = (z_{t1}, \dots, z_{tK})'$ in the second layer and the clustering sets⁵³. If there were S stimuli in simulations, the network identified the temporal stimulus through its temporal action $a_t \in \{1, 2, \dots, S\}$. With the s th outside stimulus presented, the identification from the network was correct if $a_t = s$ or incorrect otherwise. The network could receive the temporal reward as $r_t = 1$ from the correct identification and $r_t = 0$ from the incorrect identification. From previous correct identifications to each stimulus, excitatory neural responses in the second layer were collected into the clustering set for comparisons and identifications in the next. To S stimuli, clustering sets were marked as $C = \{c_s, s = 1 \dots, S\}$ with a matrix c_s containing previous neural responses to the s th stimulus. The network received temporal rewards from identifications to update both clustering sets and plastic connective weights, which was introduced in the next subsection and the supplementary material.

In the second layer, the inhibitory neural spikes were generated in the stochastic method. With the frequency-current (f -I) curve, the instantaneous firing rate of the j th second-layer inhibitory neuron was given as⁴⁵:

$$\rho_{tj}^y = \sigma_{rect}(u_{tj}^y), u_{tj}^y = \sum_{k \in \varphi_j} v^{IE} \tilde{z}_{tk} - \sum_{j' \in \mathcal{S}_j} v^{II} \tilde{y}_{tj'}, \tag{4}$$

where u_{tj}^y was the membrane potential of the j th lateral inhibitory neuron at t . ρ_{tj}^y was the temporal neural firing rate. $\sigma_{rect}(\cdot)$ was the linear rectifying function. For $u_{tj}^y \geq 0$, $\sigma_{rect}(u_{tj}^y) = u_{tj}^y$. Otherwise, $\sigma_{rect}(u_{tj}^y) = 0$. With $p_{IE} = 0.575$ ($p_{II} = 0.55$) as the specific connective probability, each excitatory (inhibitory) neuron in the second layer was decided to have the lateral connection to the given inhibitory neuron or not. φ_j (\mathcal{S}_j) included the indices of second-layer excitatory (inhibitory) neurons those connected to the j th second-layer inhibitory neuron. Through stochastic decisions of second-layer lateral connections, the network could have the structure

as the experimental observations⁵¹. Second-layer excitatory (inhibitory) neurons connected to each inhibitory neuron with the common lateral connective weight as v^{IE} (v^{II}). Same with v^{EI} in Eq. (1), $v^{IE} = 0.5$ ($v^{II} = 0.5$) in our simulations. In Eq. (4), the temporal synaptic trace of the k th lateral excitatory neuron (the j 'th lateral inhibitory neuron) was expressed as $\tilde{z}_{tk}(\tilde{y}_{ij}^y)$. With the temporal neural firing rate ρ_{ij}^y , Poisson spikes of the j th second-layer inhibitory neuron were generated. The absolute refractory period of 3 timesteps was set for inhibitory neurons⁴⁵. The temporal active state of the j th second-layer inhibitory neuron was expressed as $y_{ij} \in \{0, 1\}$.

Through Poisson spikes, the afferent neurons transmitted visual stimuli towards the second-layer network. For the visual images in simulations, the stimulus projected into the receptive field could induce afferent neural responses. Various kinds of Gaussian filters had been applied to explore brightness perception^{57,58}. In this paper, the receptive field of the n th afferent neuron was modeled with the Difference-of-Gaussians filter⁵⁹:

$$\begin{aligned} f_n(\vec{x}, \vec{x}_n) &= [g_n^c(\vec{x}, \vec{x}_n) - \phi \cdot g_n^s(\vec{x}, \vec{x}_n)], \\ g_n^c(\vec{x}, \vec{x}_n) &= \exp\left\{-\frac{[(x_1 - x_{n1})^2 + (x_1 - x_{n2})^2]}{\sigma_n^c}\right\}, \\ g_n^s(\vec{x}, \vec{x}_n) &= \exp\left\{-\frac{[(x_1 - x_{n1})^2 + (x_1 - x_{n2})^2]}{\sigma_n^s}\right\}, \end{aligned} \tag{5}$$

where $\vec{x} = (x_1, x_2)$ was the point in the image, $\vec{x}_n = (x_{n1}, x_{n2})$ was the center of the n th afferent receptive field. The n th afferent neural receptive field included center and surround Gaussian functions as $g_n^c(\cdot)$ and $g_n^s(\cdot)$. Two Gaussian functions had the spatial radii as σ_n^c and σ_n^s . The ratio between two Gaussian functions was described as the parameter ϕ . Parameters in DoG functions were generated from the distributions in⁵⁹. With a spacing distance of 2 pixels, the afferent neural receptive fields were positioned on a grid to cover the image cooperatively⁶⁰.

Visual images in our simulations were 60×60 -pixel squares with different distractors which were introduced in Results in detail. Because of designed spacing distances between afferent receptive fields, there were 900 afferent neurons in the first-layer network. The second-layer network consisted of 100 excitatory neurons and 50 inhibitory neurons.

Probabilistic model and unsupervised identifications of images. Through probabilistic inference, the network generated neural responses and identified the outside images. The probabilistic model and unsupervised identifications were similar with those in our previous study⁵³. The detail introduction was presented in the supplementary material.

In simulations, the image performed as the visual stimulus Sti . The network simulated neural responses in the second layer through the Hidden Markov model (HMM)⁴⁶. At the timestep t , the temporary observed pattern \vec{o}_t was the instantaneous feedforward input traces $\vec{x}_t = (\tilde{x}_{t1}, \dots, \tilde{x}_{tN})^T$. Second-layer neural responses $\{\tilde{z}_t, \tilde{y}_t\}$ and lateral synaptic traces $\{\tilde{z}_t, \tilde{y}_t\}$ were included in the temporal hidden state \vec{h}_t . Plastic parameters in the model were collected as $\Theta = \{W, V, C\}$. $W \in R_{K \times N}$ and $V \in R_{K \times K}$ were matrices of plastic connections. Clustering sets to stimuli were presented as $C = \{c_s, s = 1 \dots, S\}$.

For the timestep t , $\vec{O}_t = (\vec{o}_1, \vec{o}_2, \dots, \vec{o}_t)$ included occurred observations and $\vec{H}_{t-1} = (\vec{h}_1, \vec{h}_2, \dots, \vec{h}_{t-1})$ presented hidden states up to the timestep $t - 1$. The stochastic dynamics of the k-WTA network implemented a forward sampling process and sampled a new hidden active state h_t forward in time based on \vec{O}_t and \vec{H}_{t-1} ⁴⁶. In the sampling of h_t , excitatory and inhibitory neural responses in the second layer were generated by Eqs. (3) and (4), which were introduced in detail in our previous study⁵². At each timestep, the network generated an action a_t to identify stimuli with second-layer excitatory neural spikes and obtained a temporal reward $r_t \in \{0, 1\}$.

Biologically, the primary visual cortical neural responses are modified by dopaminergic rewards⁶¹. The previous WTA network had considered the reward-modified Hebbian learning⁶². In this paper, modulations of plastic weights at each timestep depended on the temporal reward r_t . $r_t = 1$ at the timestep t indicated that the network made a correct identification. Then, plastic weights were updated at t according to generated neural responses. Up to t , sequences of observed variables and hidden variables were marked as $\vec{O}_t = (\vec{o}_1, \vec{o}_2, \dots, \vec{o}_t)$, $\vec{H}_t = (\vec{h}_1, \vec{h}_2, \dots, \vec{h}_t)$. For $r_t = 1$, the connective modulations are impacted by several sub-sequences of dynamics with their lengths as $T \in \{1, \dots, t\}$. Up to t , the T -step sub-sequences of observed variables and hidden variables were marked as $\vec{O}_T = (\vec{o}_{t-T+1}, \dots, \vec{o}_t)$, $\vec{H}_T = (\vec{h}_{t-T+1}, \dots, \vec{h}_t)$. With the discount factor $\gamma \in (0, 1)$, the contribution of this pair of T -step sub-sequences to r_t was given as $\alpha(T) = (1 - \gamma)\gamma^{T-1}$. With all the sub-sequences of neural dynamics considered, the likelihood function $L(\Theta)$ was expressed as⁶²:

$$L(\Theta) = \sum_{\tau=1}^t \left[\alpha(\tau) \cdot \langle r_t \cdot \log p(\vec{O}_\tau, \vec{H}_\tau, Sti; \Theta) \rangle_{p(\vec{O}_\tau, \vec{H}_\tau, Sti, r_t)} \right]. \tag{6}$$

The joint distributions $p(\vec{O}_\tau, \vec{H}_\tau, Sti, r_t)$ and $p(\vec{O}_\tau, \vec{H}_\tau, Sti; \Theta)$ could be factorized with the assumption of Hidden Markov model:

$$\begin{aligned}
 p(\bar{O}_\tau, \bar{H}_\tau, Sti, r_t) &= p(r_t | \bar{H}_\tau) \cdot p(\bar{H}_\tau | \bar{O}_\tau) \cdot p(\bar{O}_\tau | Sti) \cdot p(Sti), \\
 p(\bar{O}_\tau, \bar{H}_\tau, Sti; \Theta) &= p(Sti | \bar{O}_\tau, \bar{H}_\tau; \Theta) p(\bar{O}_\tau, \bar{H}_\tau; \Theta) \\
 &= p(Sti | \bar{O}_\tau, \bar{H}_\tau; \Theta) \cdot \prod_{t'=\tau+1}^t p(\bar{o}_{t'} | \bar{h}_{t'}; \Theta) p(\bar{h}_{t'} | \bar{h}_{t'-1}; \Theta), \\
 p(\bar{O}_\tau, \bar{H}_\tau; \Theta) &= \sum_{t'=\tau+1}^t p(\bar{o}_{t'} | \bar{h}_{t'}; \Theta) p(\bar{h}_{t'} | \bar{h}_{t'-1}; \Theta).
 \end{aligned}
 \tag{7}$$

The likelihood function $L(\Theta)$ had the equivalent representation with the factorization in Eq. (7) as:

$$\begin{aligned}
 L(\Theta) &= \sum_{\tau=1}^t \left\{ \alpha(\tau) \cdot \left\langle r_t \cdot \log \left[p(Sti | \bar{O}_\tau, \bar{H}_\tau; \Theta) \cdot \prod_{t'=\tau+1}^t p(\bar{o}_{t'} | \bar{h}_{t'}; \Theta) p(\bar{h}_{t'} | \bar{h}_{t'-1}; \Theta) \right] \right\rangle_{p(\bar{O}_\tau, \bar{H}_\tau, Sti, r_t)} \right\} \\
 &= \sum_{\tau=1}^t \alpha(\tau) \cdot \left\langle r_t \cdot \log p(Sti | \bar{O}_\tau, \bar{H}_\tau; \Theta) + r_t \cdot \sum_{t'=\tau+1}^t \log \left[p(\bar{o}_{t'} | \bar{h}_{t'}; \Theta) p(\bar{h}_{t'} | \bar{h}_{t'-1}; \Theta) \right] \right\rangle_{p(\bar{O}_\tau, \bar{H}_\tau, Sti, r_t)}.
 \end{aligned}
 \tag{8}$$

For the likelihood function $L(\Theta)$ in Eq. (8), the network updated plastic connective weights through an online variant of the Expectation–Maximization algorithm. In this paper, the E-step was to estimate the expectation with a single sample of neural responses⁴⁶. With the single sample from $p(\bar{O}_\tau, \bar{H}_\tau, Sti, r_t)$, $L(\Theta)$ was rearranged and approximated as:

$$L(\Theta) \simeq \sum_{\tau=1}^t \alpha(\tau) \cdot r_t \cdot \left[\log p(Sti | \bar{O}_\tau, \bar{H}_\tau; \Theta) + \sum_{t'=\tau+1}^t \log p(\bar{o}_{t'} | \bar{h}_{t'}; \Theta) p(\bar{h}_{t'} | \bar{h}_{t'-1}; \Theta) \right].
 \tag{9}$$

Then, with directions given by partial derivatives of $L(\Theta)$, the network updated its plastic weights in the M-step⁶³. In this paper, the directions for updating w_{kn} and $v_{kk'}$ were given as^{52,53}:

$$\begin{aligned}
 \frac{\partial}{\partial w_{kn}} L(\Theta) &\simeq r_t \sum_{t'=1}^t \gamma^{t-t'} \cdot z_{t'k} \cdot \left\{ \tilde{x}_{t'n} - \left[1 - \frac{1}{w_{kn}} + \frac{1}{\exp(w_{kn}) - 1} \right] \right\}, \\
 \frac{\partial}{\partial v_{kk'}} L(\Theta) &\simeq r_t \sum_{t'=1}^t \gamma^{t-t'} \cdot z_{t'k'} \cdot \left\{ \tilde{z}_{t'k'} - \left[1 - \frac{1}{v_{kk'}} + \frac{1}{\exp(v_{kk'}) - 1} \right] \right\}.
 \end{aligned}
 \tag{10}$$

The modification of each weight depended on the pre- and postsynaptic responses, as well as the temporal reward r_t . r_t controlled connective modifications. Connective weights only would be updated with the correct identification. In simulations, the discount factor $\gamma = 0.9$.

To identify the outside stimuli, the network applied the unsupervised method⁶⁴. The detail generations of a_t and r_t were introduced in our previous study⁵³. The brief description of identifications was introduced in this subsection. Previous neural responses had been identified and collected into clustering sets for latter identifications. For temporal second-layer excitatory neural spikes \bar{z}_t and α , we calculated energy distances between \bar{z}_t and clustering sets $\{c_s, s = 1, \dots, S\}$ to estimate parameters $\{e_{accept}(\alpha, c_s, t), s = 1, \dots, S\}$ through R re-samples⁶⁴. For $\{c_s, s = 1, \dots, S\}$, a common parameter was set as $e_{accept}(\alpha, t) = \min \{e_{accept}(\alpha, c_s, t), s = 1, \dots, S\}$. The likelihoods between \bar{z}_t and clustering sets could be quantified by the probabilities $\{p(\alpha, c_s, t) = p(e(\alpha, c_s, t) \leq e_{accept}(\alpha, t)), s = 1, \dots, S\}$.

To the outside stimulus, a_t at the timestep t performed as the temporal identification. The distribution of a_t was given as:

$$p(a_t = s') = \exp(p(\alpha, c_{s'}, t)) / \left[\sum_s \exp(p(\alpha, c_s, t)) \right].
 \tag{11}$$

If a_t took a value equal to the serial number of the presented stimulus, the temporal identification was correct. In a First-In-First-Out (FIFO) manner, the clustering set would be updated at a timestep with the correct identification. In each update, \bar{z}_t would be added to the end of the corresponding clustering set. After updating, the clustering set would delete redundant components from its beginning if its size exceeds $n_{cluster}$. Detail introductions of updating clustering sets were introduced in our previous study⁵³. In our simulations, $n_{cluster} = 20$, $\alpha = 0.05$, $R = 50$.

Results

Similar with stimuli in the previous study⁷, gray square images with different contexts are designed in simulations. This section is to explore how distractors modify brightness perception of the target stimulus.

The phenomenon of simultaneous brightness contrast. This subsection is to test whether our network can simulate distinguishing brightness perception of the same stimulus upon opposite backgrounds. In simulations, two sets of images are designed as visual stimuli.

In the first set, two images are designed by combining a grey square with darker and lighter backgrounds, respectively⁶⁵. Each image is a 60×60 -pixel square. The target stimulus is a 20×20 -pixel square in the center of each image (Fig. 2A). In the second set, 60×60 -pixel square images have the combined backgrounds. A 20×20 -pixel target stimulus belongs to different parts in images (Fig. 2F). In both two sets of images, the grey target stimulus has a grayscale value of 0.5, the darker background has a grayscale value of 0.2 and the lighter background has a grayscale value of 0.8.

For stimuli in Fig. 2A, visual images are encoded by afferent neural spikes. In each learning simulation, two images are rearranged into a random sequence. In 200 learning simulations, connective weights and clustering sets are updated as introduced in Materials and Methods. Initial values of plastic feedforward and lateral connective weights are sampled independently from a uniform distribution of (0.001, 1). During modifications, weights are limited in (0, 1). The testing phase consists of 100 simulations with weights and clustering sets fixed. A sequence of images is presented as in Fig. 2A. Neural responses over simulations are measured to explore contextual modulation induced by distractors. For images in Fig. 2F, our network is trained and tested in the same way. Besides, learning simulations and testing simulations in all the following subsections are designed similarly.

Brightness perception of the target is simulated through reconstructions of images. Over testing simulations, a point (x_1, x_2) in each image is reconstructed with averaged neural responses as:

$$\begin{aligned} I_{rec}(x_1, x_2) &\equiv \sum_{k,n} \bar{z}_k \cdot w_{kn} \cdot \bar{y}_n \cdot I_n^{rec}(x_1, x_2) \\ I_n^{rec}(x_1, x_2) &= a_n^{rec} \cdot f_n(\vec{x}, \vec{x}_n) \\ a_n^{rec} &= \sum_{(x'_1, x'_2)} I(x'_1, x'_2) \cdot f_n(\vec{x}', \vec{x}_n), \end{aligned} \quad (12)$$

where $\vec{x} = (x_1, x_2)$ is a point in the image, $f_n(\cdot)$ is the Difference-of-Gaussian filter in Eq. (5) with its center as $\vec{x}_n = (x_{n1}, x_{n2})$. w_{kn} is the modified feedforward weight. \bar{z}_k, \bar{y}_n are averaged spiking rates of corresponding neurons over timesteps and simulations. $I(x'_1, x'_2)$ is the gray-scale value of a point $\vec{x}' = (x'_1, x'_2)$. $I_{rec}(x_1, x_2)$ is the reconstructed gray-scale value of (x_1, x_2) . The reconstruction in Eq. (12) performs as averaged brightness perception. Reconstructions of two images are normalized before comparison. Reconstructed gray-scale values of all the points in two images are collected together into a set. The mean value and standard deviation of this set are common parameters for normalizations. In this way, perception of the same target upon different contexts can be simulated and compared. For the first set, the target in Sti. 1 appears to be lighter (Fig. 2B). For the second set, the target in Sti. 4 appears to be lighter (Fig. 2G). Simulated brightness perception has the relationship as same as that in experiments⁶⁵.

The contextual modulation index (CMI) and the cross-correlation are used to qualify contextual modifications of excitatory neural responses and excitatory neural lateral interactions, respectively. For instance, the CMI of the k th second-layer excitatory neuron to Sti. 1 and Sti. 2 is calculated as³⁷:

$$CMI_k = \frac{r_{k1} - r_{k2}}{r_{k1} + r_{k2}}, \quad (13)$$

where r_{k1} and r_{k2} are spiking counts of the k th excitatory neuron to Sti. 1 and Sti. 2, respectively. If $r_{k1} = r_{k2}$, responses of this neuron to two stimuli are same and $CMI_k = 0$. If $r_{k1} \neq r_{k2}$, two stimuli induce different responses of this neuron and $CMI_k \neq 0$. In this way, a non-zero CMI indicates contextual modulation of neural responses. CMI in Eq. (13) is limited within $[-1, 1]$ by the denominator. With an absolute value closer to 1, a CMI reflects larger contextual modulation induced by different stimuli. For a neural population, contextual modulation of neural responses can be reflected by the histogram of CMIs. For second-layer excitatory neurons, non-zero CMIs in Fig. 2C,H indicate that our network can simulate contextual modulation of neural responses induced by different backgrounds.

The cross-correlation of each pair of excitatory neurons in the second layer quantify their interactions⁵⁵. Cross-correlations are calculated with the time lag of 0. To reflect influences on neural responding cross-correlations, differences between neural responding cross-correlations to each set of images are calculated. For a pair of excitatory neurons, a non-zero difference of neural cross-correlation indicates the context can induce the modulation on their interactions. As shown in Fig. 2D,I, our network can simulate contextual modulation of neural interactions.

Our spiking network can simulate sparse neural responses to images^{45,46}. To two sets of designed images in Fig. 2A,F, temporal averaged firing rates of second-layer excitatory neurons to images are calculated over testing simulations. As shown in Fig. 2E,J, neural responding strengths are expressed by the varying color. For the convenient observation, the sequence of neurons is re-arranged according to their responding strengths to each set of images. At each timestep, the darker rectangular block indicates the stronger responding strength of a neuron. The white blocks stand for the minimum responding strength of 0. To an image, some excitatory neurons emit spikes intensely while other excitatory neurons generate few spikes. Different images can induce intense neural responses of different subsets of second-layer excitatory neurons. After learning, our network can simulate sparse and distinguishing responses to different images.

We also provide brief simulations of anchoring⁶⁶. Images in Fig. 2K perform as another set of visual stimuli. The grey target stimulus has a grayscale value of 0.5. The darker background has a grayscale value of 0.2. Other lighter backgrounds have grayscale values of 0.8, 0.9 and 1, respectively. Brightness perception simulated from the network and the anchoring function in the previous study⁶⁶ are shown in Fig. 2L. With the background becoming

Figure 2. Perception of simultaneous brightness contrast from the network. (A) Images used as the first set of visual stimuli in simulations. (B) Averaged brightness perception of target squares with neural responses from our network. Perception is averaged over simulations, timesteps and pixels. Perception of the same gray square is lighter upon the darker background. (C) The histogram of contextual modulation indices (CMIs) about excitatory neural responses to Sti. 1 and Sti. 2. A non-zero CMI indicates contextual modulation of neural responses. (D) Difference of excitatory neural cross-correlations to Sti. 1 and Sti. 2. A non-zero difference indicates contextual modulation of interactions between paired neurons. (E) Temporal excitatory neural firing rates to Sti. 1 and Sti. 2 averaged over simulations. Color of each block indicates neural responding strength. Different images can induce intense responses of different subset of neurons. (F) Images used as the second set of visual stimuli in simulations. (G) Averaged brightness perception of target squares. (H) The histogram of CMIs about excitatory neural responses to Sti. 3 and Sti. 4. (I) Difference of second-layer excitatory neural cross-correlations to Sti. 3 and Sti. 4. (J) Temporal second-layer excitatory neural firing rate to Sti. 3 and Sti. 4 averaged over simulations. (K) Images used as the third set of visual stimuli. (L) Averaged brightness perception of the target from both the neural network and the anchoring function. With the background becoming lighter, varying trends of estimated brightness perception in subplots are similar.

lighter, varying trends of brightness perception in subplots are similar. Our network can provide simulations of the basic phenomenon of anchoring.

Simulations in this subsection show that our network can reflect contextual modulation of neural responses and interactions, simulating opposite brightness perception of the same target.

Modulations and perception induced by distances between distractors and targets. Recent study has explored illusory brightness perception induced by different orders of presentations of two stimuli⁶⁷. In this paper, the additional gray-scale squares are designed around the target and perform as distractors. It is to explore how distractors modify brightness perception of the given target.

In this subsection, distractors are designed to have different distances to targets in visual images. In our simulations, the gray patch is upon different contexts in Fig. 3A. Each image is a 60×60 -pixel square. The 20×20 -pixel grey target has a grayscale value of 0.5, the darker background has a grayscale value of 0.2 and the lighter background has a grayscale value of 0.8. 10×10 -pixel distractors have a common grayscale value of 1. Distractors in Sti. 3 locate along boundaries of the image. Distractors in Sti. 4 become closer to the target with both horizontal and vertical moving steps of 5 pixels.

With images in Fig. 3A, 200 learning simulations and 100 testing simulations are designed for our network. Over testing simulations, a reconstruction of an image is given by Eq. (12). Average perception of the target is calculated and shown in Fig. 3B. Brightness perception of the target stimulus is darker when the distance from distractors to the target is shorter. The varying perception can reflect influences of distances between distractors and the target. The contextual modulation index (CMI) and the cross-correlations are used to qualify influences on the excitatory neural responses and lateral interactions, respectively. To calculate modulations induced by different distances, neural responses to Sti. 1 are collected for comparison. Non-zero CMIs in histograms reflect induced responding modifications of second-layer excitatory neurons (Fig. 3C). When distractors become closer to the target, there are more CMIs having the absolute value close to 1. It indicates that, for designed images, a shorter distance between distractors and the target can induce larger neural responding modulations.

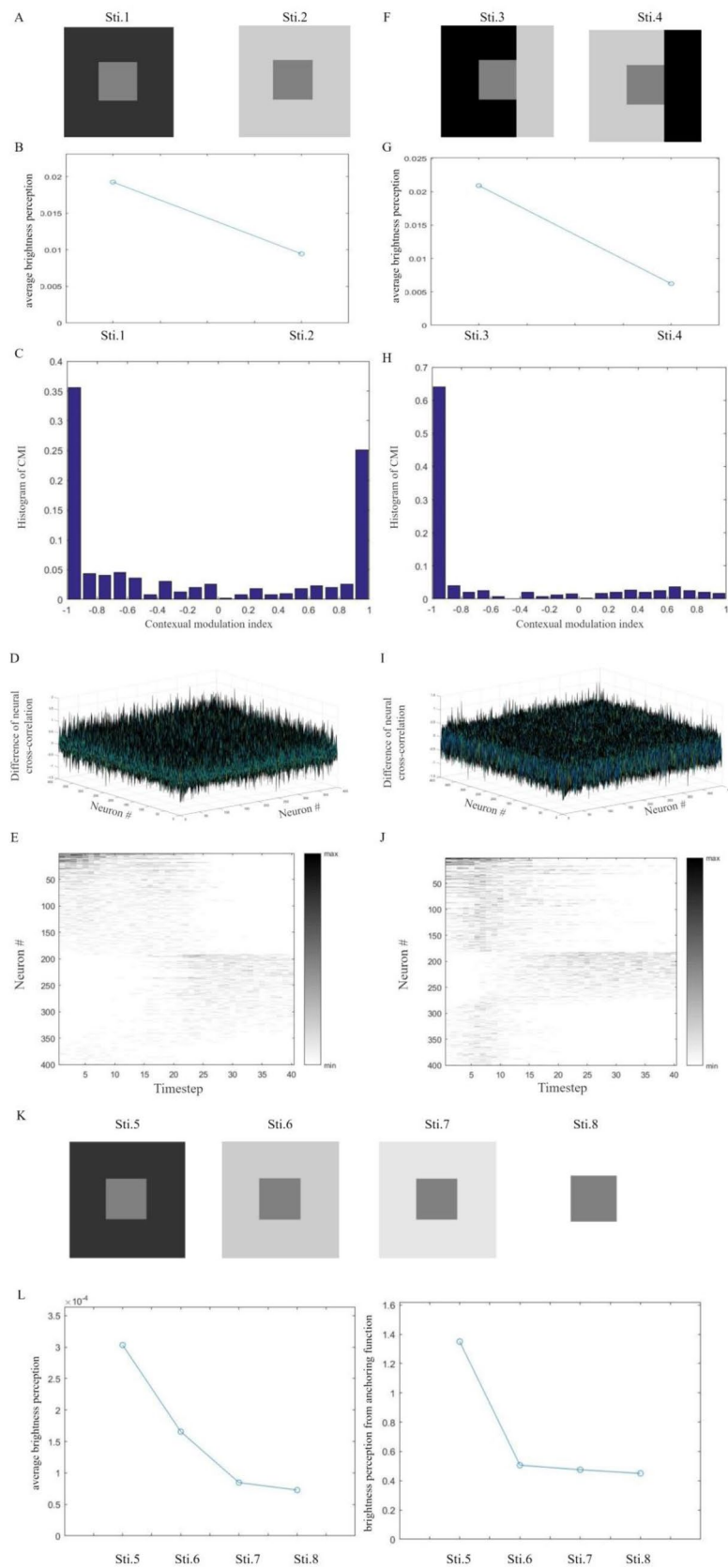
Differences between neural responding cross-correlations to Sti. 1 and other stimuli are calculated. As shown in Fig. 3D, a non-zero difference of neural cross-correlation indicates the influence of the varying distance on neural interactions. With simulations of Sti. 1 for comparison, different amplitudes of neural interactive modification reflect influences of distances between distractors and the target.

Simulations show that our network can reflect contextual modification induced by distances between distractors and the target. Modified neural activities simulate different perception of the same target. For designed images in Fig. 3A a shorter distance makes our network simulate darker perception of the same target.

Modulations and perception induced by grayscale values of distractors. In this subsection, distractors are designed to have different grayscale values. Four 60×60 -pixel images are designed as in Fig. 4A. The target and backgrounds are the same with those in Fig. 3A. The 10×10 -pixel distractors in Sti. 3 has a grayscale value of 0.9. The distractors in Sti. 4 has a grayscale value of 1.

200 learning simulations and 100 testing simulations are designed for the network. Over testing simulations, brightness perception of the target is simulated and shown in Fig. 4B. Our simulations show that the target with lighter distractors appears to be darker. With Sti. 1 for comparison, CMIs modulated by other 3 stimuli are measured for each neuron. Induced by different distractors, neural responding modulations are shown in Fig. 4C. Simulations show that, to designed images in Fig. 4A, lighter distractors can induce larger neural responding modulations. The network can reflect modifications of excitatory neural responses induced by grayscale values of distractors. Differences of neural responding cross-correlations are calculated and shown in Fig. 4D. A non-zero difference of neural cross-correlation shows how grayscale values of distractors affect neural interactions. Compared to simulations of Sti. 1, interactive modulations between each pair of neurons are induced by stimuli and reflected by amplitudes of non-zero differences.

This subsection explores how grayscale values of distractors modify the neural activities of the network. With modulated responses, our network can simulate distinguishing brightness perception induced by grayscale values of distractors. For designed images in Fig. 4A, lighter distractors make our network simulate darker perception of the target.



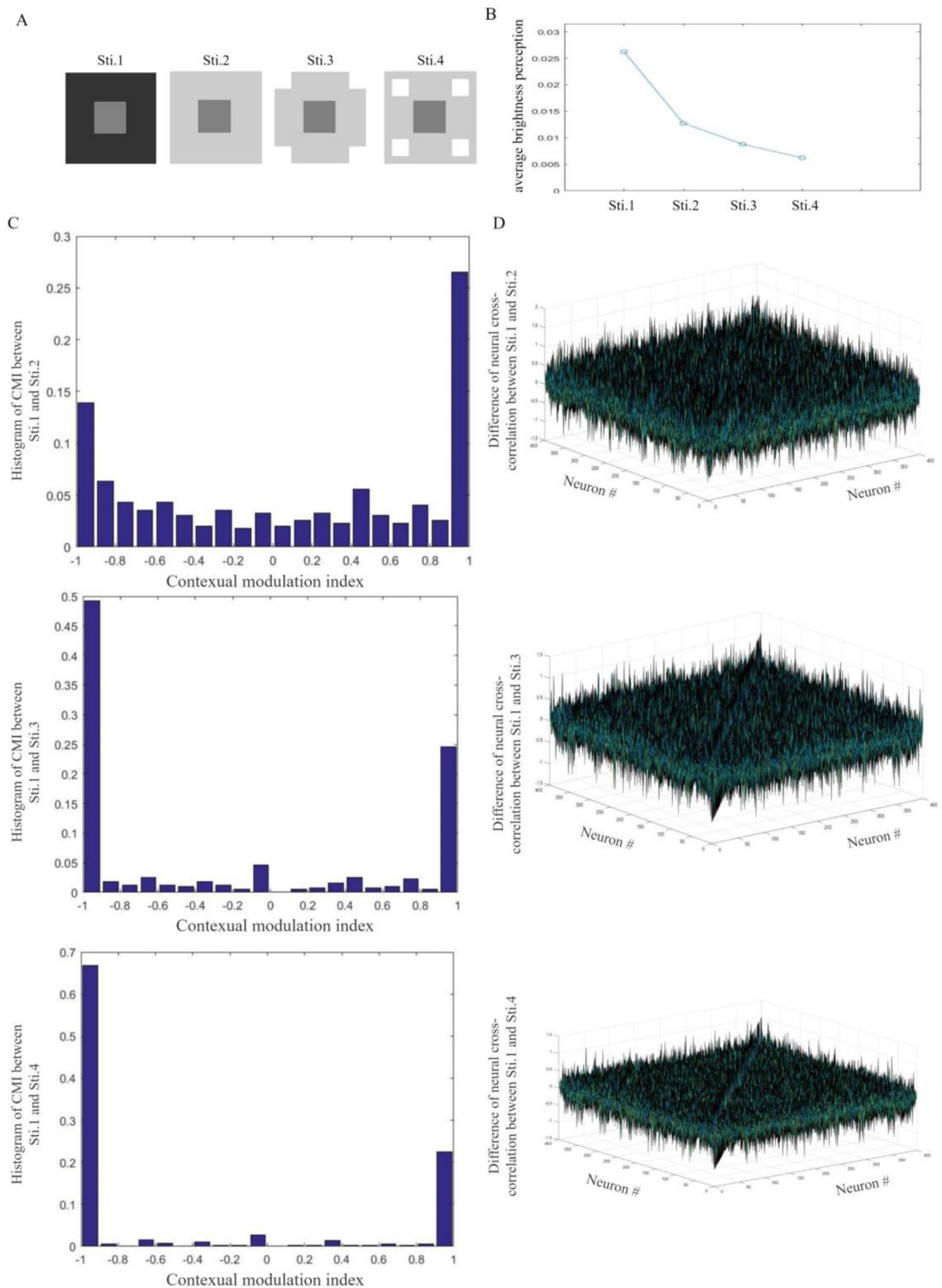


Figure 3. Brightness perception induced by distances between distractors and targets. **(A)** Images used as visual stimuli in our simulations. **(B)** Averaged brightness perception of target squares. For designed stimuli, distractors with a shorter distance induce darker brightness perception of the target. **(C)** The histogram of excitatory neural CMIs to stimuli. A non-zero CMI indicates the neural responding modification induced by the varying distance between distractors and the target. For a neural population, the histogram of CMIs reflects contextual modulation of neural responses. For designed stimuli, distractors with a shorter distance can induce larger neural responding modulations. **(D)** Difference of excitatory neural responding cross-correlation to stimuli. A non-zero difference indicates the modulation of neural interactions induced by the varying distance. With simulations of Sti. 1 for comparison, different amplitudes of neural interactive modification reflect influences of the varying distance.

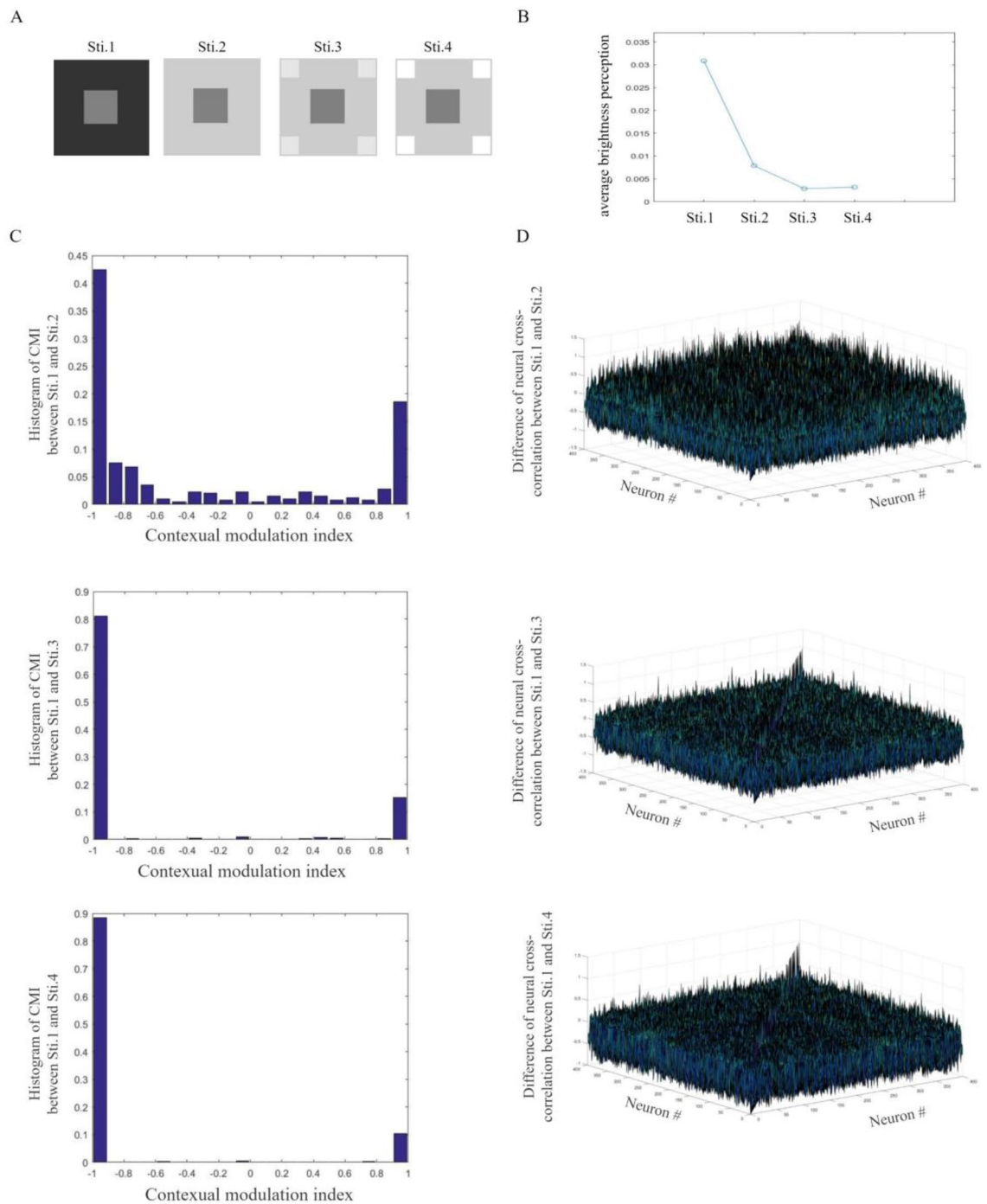


Figure 4. Brightness perception induced by grayscale values of distractors. (A) Images used as visual stimuli in our simulations. (B) Averaged brightness perception of target squares. For designed stimuli, the target with lighter distractors appears to be darker. (C) The histogram of CMIs of excitatory neurons to images. Non-zero CMIs indicate neural responding modifications induced by grayscale values of distractors. For designed stimuli, lighter distractors can induce larger neural responding modulations. (D) Differences of excitatory neural responding cross-correlation to stimuli. For each pair of neurons, amplitudes of non-zero differences reflect influences of the varying grayscale value of distractors.

Modulations and perception induced by sizes of distractors. In this subsection, distractors are designed to have different sizes. The 60×60 -pixel images in Fig. 5A have the same target and backgrounds as those in Fig. 3A. With a grayscale value of 1, square distractors have different sizes. In our simulations, Sti. 3 has 5×5 -pixel distractors while Sti. 4 has 10×10 -pixel distractors.

With stimuli in Fig. 5A, the learning phase and the testing phase are designed. Over testing simulations, reconstructions of the target given by Eq. (12) perform as brightness perception. As shown in Fig. 5B, for

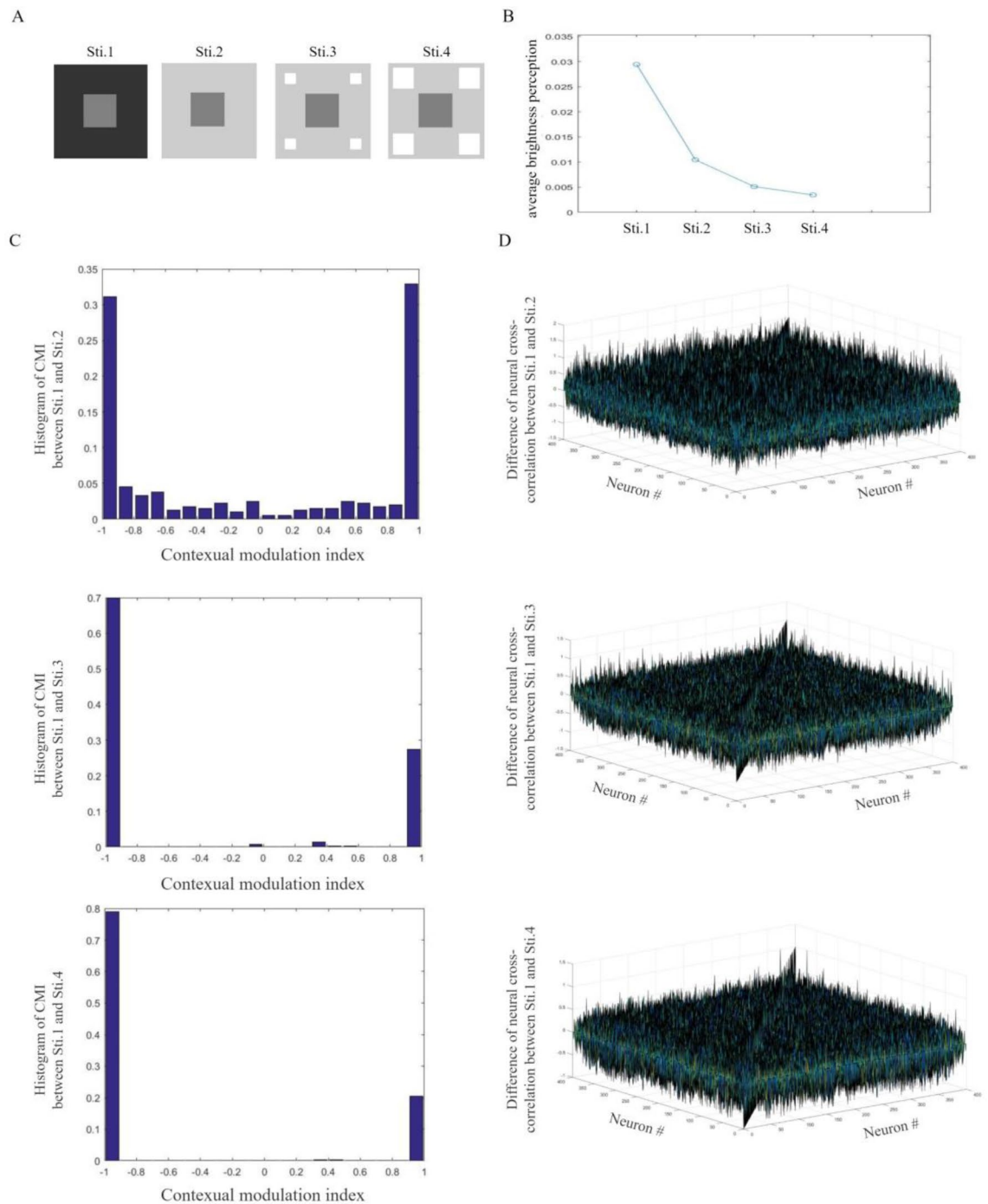


Figure 5. Brightness perception induced by sizes of distractors. **(A)** Images used as visual stimuli in our simulations. **(B)** Averaged brightness perception of target squares. For designed stimuli, perception of the same gray square is darker with bigger distractors. **(C)** The histogram of CMIs to stimuli. For designed stimuli, larger distractors could induce larger neural responding modulations. **(D)** Differences of excitatory neural responding cross-correlation to stimuli. A non-zero difference indicates the modulation of neural interactions induced by the varying size of distractors.

designed images in simulations, the stimulus with bigger distractors appears to be darker. With Sti. 1 for comparison, CMIs and differences of neural responding cross-correlations are measured for other three stimuli. The varying size of distractors could induce neural responding modulations (Fig. 5C). Distractors with larger size can lead to larger modulations of neural responses. A non-zero difference of neural cross-correlation indicates contextual modulation of neural interactions induced by the varying size (Fig. 5D).

Neural responses and interactions modified by sizes of distractors are measured and reflected from the network. With modulated responses, our network can simulate distinguishing brightness perception induced by the

varying size. For designed images in Fig. 5A, bigger distractors make our network simulate darker perception of the target.

Conclusion

This paper has provided a probabilistic framework to explore how distractors modify primary visual cortical responses and induce different brightness perception of the same stimulus. Recent experimental study has demonstrated that the phenomenon of simultaneous brightness contrast is associated with primary visual cortical neural responses². Besides, contextual effects are also associated with the primary visual cortex^{13,33,38}. While brightness perception has been studied for a long time, how distractors modify primary visual cortical neural responses to induce different brightness perception of the same stimulus is still not clear. To explore the corresponding mechanism, we design a stochastic spiking network with plastic connections in this study. Visual images are designed to control varying properties of distractors, excluding undesired factors in simulations.

Our network is constructed with two layers. With neural receptive fields as Difference-of-Gaussians filters, first-layer afferent neurons generate Poisson spiking responses to received stimuli. Through forward sampling in the Hidden Markov Model, second-layer excitatory and inhibitory neurons generate their spiking responses and communicate with each other. With multi-dimensional excitatory neural spiking responses in the second layer, the network identifies the presented stimulus and receives rewards which control connective modulations.

Applications of afferent receptive fields in this paper remove the strict limitation on the number of neurons in previous WTA networks while simulating sparse neural responses^{45,46}. Compared to neural populational coding with Gaussian functions⁵⁸, this model considers neural interactions and synaptic plasticity in the primary visual cortex. In this way, besides neural responses, the plasticity-based influence is also considered to explore illusory brightness perception induced by distractors⁵⁴. Compared to the feedforward network model for visual perception with distractors, our recurrent network considers influences of neural lateral interactions⁶⁸. In biology, the dopaminergic reward has been found to participate in the synaptic plasticity in the primary visual cortex⁶¹. Compared to previous networks on brightness perception^{5,35,36,57}, this model provides unsupervised identifications of stimuli and considers influences of rewards on learning. This unsupervised method only depends on generated neural spikes which are easy to obtain from simulations. Without limiting the dimension of neural responses, the unsupervised identification can provide online rewards to control connective modifications.

This paper explores how properties of distractors modulate neural responses and induce different brightness perception of the given target. For images as visual stimuli in simulations, distractors are designed with the varying grayscale value, the size and the distance to the target. Over simulations, both brightness perception of the target stimulus and neural responding modulations are measured. Our network can reflect modulations on both neural responses and interactions induced by each property of distractors, simulating different brightness perception with modulated responses.

Recent experimental observations have localized brightness perception to a site preceding binocular fusion². Following the associated biological structure, our network can simulate both brightness perception and contextual modifications at the same time, providing a theoretical framework based on probabilistic inference to explore how distractors modulate neural responses and lead to different brightness perception of the same target. Our model provides an alternative method to explore brightness perception from contextual modification of primary visual cortical neural responses and interactions.

Data availability

Supporting codes and data will be made available upon request to the corresponding author.

Received: 7 July 2022; Accepted: 17 January 2023

Published online: 27 January 2023

References

- White, M. The effect of the nature of the surround on the perceived lightness of grey bars within square-wave test gratings. *Perception* **10**, 215–230 (1981).
- Sinha, P. *et al.* Mechanisms underlying simultaneous brightness contrast: Early and innate. *Vision Res.* **173**, 41–49 (2020).
- Rossi, A. F., Rittenhouse, C. D. & Paradiso, M. A. The representation of brightness in primary visual cortex. *Science* **273**, 1104–1107 (1996).
- Rossi, A. F. & Paradiso, M. A. Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *J. Neurosci.* **19**, 6145–6156 (1999).
- Batard, T. & Bertalmio, M. A geometric model of brightness perception and its application to color images correction. *J. Math. Imaging Vis.* **60**, 849–881 (2018).
- Adelson, E. H. Perceptual organization and the judgment of brightness. *Science* **262**, 2042–2044 (1993).
- Rodriguez, A. & Granger, R. On the contrast dependence of crowding. *J. Vis.* **21**, 4 (2021).
- Henry, C. A. & Kohn, A. Spatial contextual effects in primary visual cortex limit feature representation under crowding. *Nat. Commun.* **11**, 1687 (2020).
- Levi, D. M. & Carney, T. Crowding in peripheral vision: Why bigger is better. *Curr. Biol.* **19**, 1988–1993 (2009).
- Chicherov, V., Plomp, G. & Herzog, M. H. Neural correlates of visual crowding. *Neuroimage* **93**, 23–31 (2014).
- Das, A. & Gilbert, C. D. Topography of contextual modulations mediated by short-range interactions in primary visual cortex. *Nature* **399**, 655–661 (1999).
- Levitt, J. B. & Lund, J. S. Contrast dependence of contextual effects in primate visual cortex. *Nature* **387**, 73–76 (1997).
- Rossi, A. F., Desimone, R. & Ungerleider, L. G. Contextual modulation in primary visual cortex of macaques. *J. Neurosci.* **21**, 1698–1709 (2001).
- Gheorghiu, E. & Kingdom, F. A. A. Dynamics of contextual modulation of perceived shape in human vision. *Sci. Rep.* **7**, 43274 (2017).
- Ziamba, C. M., Freeman, J., Simoncelli, E. P. & Movshon, J. A. Contextual modulation of sensitivity to naturalistic image structure in macaque V2. *J. Neurophysiol.* **120**, 409–420 (2018).

16. Quek, G. L. & Peelen, M. V. Contextual and spatial associations between objects interactively modulate visual processing. *Cereb. Cortex* **30**, 6391–6404 (2020).
17. Pelli, D. G. & Tillman, K. A. The uncrowded window of object recognition. *Nat. Neurosci.* **11**, 1129–1135 (2008).
18. Whitney, D. & Levi, D. M. Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends Cogn. Sci.* **15**, 160–168 (2011).
19. Ozeki, H. *et al.* Relationship between excitation and inhibition underlying size tuning and contextual response modulation in the cat primary visual cortex. *J. Neurosci.* **24**, 1428–1438 (2004).
20. Franceschiello, B., Sarti, A. & Citti, G. A neuromathematical model for geometrical optical illusions. *J. Math. Imaging Vis.* **60**, 94–108 (2018).
21. Mahmoodi, S. Linear neural circuitry model for visual receptive fields. *J. Math. Imaging Vis.* **54**, 138–161 (2016).
22. Baspinar, E., Citti, G. & Sarti, A. A geometric model of multi-scale orientation preference maps via Gabor functions. *J. Math. Imaging Vis.* **60**, 900–912 (2018).
23. Adjarian, P. *et al.* Induced visual illusions and gamma oscillations in human primary visual cortex. *Eur. J. Neurosci.* **20**, 587–592 (2004).
24. King, J. L. & Crowder, N. A. Adaptation to stimulus orientation in mouse primary visual cortex. *Eur. J. Neurosci.* **47**, 346–357 (2018).
25. Bharmauria, V. *et al.* Network-selectivity and stimulus-discrimination in the primary visual cortex: Cell-assembly dynamics. *Eur. J. Neurosci.* **43**, 204–219 (2016).
26. Dai, J. & Wang, Y. Contrast coding in the primary visual cortex depends on temporal contexts. *Eur. J. Neurosci.* **47**, 947–958 (2018).
27. Ghodrati, M., Alwis, D. S. & Price, N. S. C. Orientation selectivity in rat primary visual cortex emerges earlier with low-contrast and high-luminance stimuli. *Eur. J. Neurosci.* **44**, 2759–2773 (2016).
28. Olshausen, B. A. & Field, D. J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**, 607–609 (1996).
29. Storrs, K. R., Anderson, B. L. & Fleming, R. W. Unsupervised learning predicts human perception and misperception of gloss. *Nat. Hum. Behav.* **5**, 1402–1417 (2021).
30. Metzger, A. & Toscani, M. Unsupervised learning of haptic material properties. *eLife* **11**, 64876 (2022).
31. Fleming, R. W. & Storrs, K. R. Learning to see stuff. *Curr. Opin. Behav. Sci.* **30**, 100–108 (2019).
32. Flachot, A. & Gegenfurtner, K. R. Color for object recognition: Hue and chroma sensitivity in the deep features of convolutional neural networks. *Vision Res.* **182**, 89–100 (2021).
33. Zipsper, K., Lamme, V. A. F. & Schiller, P. H. Contextual modulation in primary visual cortex. *J. Neurosci.* **16**, 7376–7389 (1996).
34. Kerr, D., McGinnity, T. M., Coleman, S. & Clogenson, M. A biologically inspired spiking model of visual processing for image feature detection. *Neurocomputing* **158**, 268–280 (2015).
35. Pessoa, L., Mingolla, E. & Neumann, H. A contrast- and luminance-driven multiscale network model of brightness perception. *Vision Res.* **35**, 2201–2223 (1995).
36. Grossberg, S. & Kelly, F. Neural dynamics of binocular brightness perception. *Vision Res.* **39**, 3796–3816 (1999).
37. Keller, A. J. *et al.* A disinhibitory circuit for contextual modulation in primary visual cortex. *Neuron* **108**, 1181–1193.e8 (2020).
38. Ursino, M. & La Cara, G. E. A model of contextual interactions and contour detection in primary visual cortex. *Neural Netw.* **17**, 719–735 (2004).
39. Murray, R. F. A model of lightness perception guided by probabilistic assumptions about lighting and reflectance. *J. Vis.* **20**, 28 (2020).
40. Orbán, G., Berkes, P., Fiser, J. & Lengyel, M. Neural variability and sampling-based probabilistic representations in the visual cortex. *Neuron* **92**, 530–543 (2016).
41. Zemel, R. S. Probabilistic interpretation of population codes. *Neural Comput.* **10**, 403–430 (1998).
42. Lloyd, K. & Leslie, D. S. Context-dependent decision-making: A simple Bayesian model. *J. R. Soc.* **10**, 20130069 (2013).
43. Ye, R. & Liu, X. How the known reference weakens the visual oblique effect: A Bayesian account of cognitive improvement by cue influence. *Sci. Rep.* **10**, 20269 (2020).
44. Allred, S. R. & Brainard, D. H. A Bayesian model of lightness perception that incorporates spatial variation in the illumination. *J. Vis.* **13**, 18 (2013).
45. Jonke, Z., Legenstein, R., Habenschuss, S. & Maass, W. Feedback inhibition shapes emergent computational properties of cortical microcircuit motifs. *J. Neurosci.* **37**, 8511–8523 (2017).
46. Kappel, D., Nessler, B. & Maass, W. STDP installs in winner-take-all circuits an online approximation to hidden Markov model learning. *PLoS Comput. Biol.* **10**, e1003511 (2014).
47. Klampfl, S. & Maass, W. Emergence of dynamic memory traces in cortical microcircuit models through STDP. *J. Neurosci.* **33**, 11515–11529 (2013).
48. Abadi, A. K., Yahya, K., Amini, M., Friston, K. & Heinke, D. Excitatory versus inhibitory feedback in Bayesian formulations of scene construction. *J. R. Soc. Interface* **16**, 20180344 (2019).
49. van Rossum, M. C., Bi, G. Q. & Turrigiano, G. G. Stable Hebbian learning from spike timing-dependent plasticity. *J. Neurosci.* **20**, 8812–8821 (2000).
50. Van Rullen, R. & Thorpe, S. J. Rate coding versus temporal order coding: What the retinal ganglion cells tell the visual cortex. *Neural Comput.* **13**, 1255–1283 (2001).
51. Avermann, M., Tomm, C., Mateo, C., Gerstner, W. & Petersen, C. C. H. Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *J. Neurophysiol.* **107**, 3116–3134 (2012).
52. Liu, W. & Liu, X. The effects of eye movements on the visual cortical responding variability based on a spiking network. *Neurocomputing* **436**, 58–73 (2021).
53. Liu, W. & Liu, X. Depth perception with interocular blur differences based on a spiking network. *IEEE Access* **10**, 11957–11978 (2022).
54. Hussain, Z., Webb, B. S., Astle, A. T. & McGraw, P. V. Perceptual learning reduces crowding in amblyopia and in the normal periphery. *J. Neurosci.* **32**, 474–480 (2012).
55. Hata, Y., Tsumoto, T., Sato, H. & Tamura, H. Horizontal interactions between visual cortical neurones studied by cross-correlation analysis in the cat. *J. Physiol.* **441**, 593–614 (1991).
56. Masquelier, T., Guyonneau, R. & Thorpe, S. J. Competitive STDP-based spike pattern learning. *Neural Comput.* **12**, 1259–1276 (2009).
57. Ding, J. & Levi, D. M. Binocular combination of luminance profiles. *J. Vis.* **17**, 4 (2017).
58. Blakeslee, B., Cope, D. & McCourt, M. E. The Oriented Difference of Gaussians (ODOG) model of brightness perception: Overview and executable Mathematica notebooks. *Behav. Res. Methods* **48**, 306–312 (2016).
59. Benardete, E. A. & Kaplan, E. The receptive field of the primate P retinal ganglion cell, I: linear dynamics. *Vis. Neurosci.* **14**, 169–185 (1997).
60. Segal, I. Y. *et al.* Decorrelation of retinal response to natural scenes by fixational eye movements. *PNAS* **112**, 3110–3115 (2015).
61. Arsénault, J. T., Nelissen, K., Jarraya, B. & Vanduffel, W. Dopaminergic reward signals selectively decrease fMRI activity in primate visual cortex. *Neuron* **77**, 1174–1186 (2013).

62. Rueckert, E., Kappel, D., Tanneberg, D., Pecevski, D. & Peters, J. Recurrent spiking networks solve planning tasks. *Sci. Rep.* **6**, 21142 (2016).
63. Legenstein, R., Jonke, Z., Habenschuss, S. & Maass, W. A probabilistic model for learning in cortical microcircuit motifs with data-based divisive inhibition. *ArXiv arXiv:1707.05182v1* (2017).
64. Heinerman, J., Haasdijk, E. & Eiben, A. E. Unsupervised identification and recognition of situations for high-dimensional sensori-motor streams. *Neurocomputing* **262**, 90–107 (2017).
65. Agostini, T. & Proffitt, D. R. Perceptual organization evokes simultaneous lightness contrast. *Perception* **22**, 263–272 (1993).
66. Economou, E., Zdravkovic, S. & Gilchrist, A. Anchoring versus spatial filtering accounts of simultaneous lightness contrast. *J. Vis.* **7**, 2 (2007).
67. Zhou, H. *et al.* Spatiotemporal dynamics of brightness coding in human visual cortex revealed by the temporal context effect. *Neuroimage* **205**, 116277 (2020).
68. Chen, N., Bao, P. & Tjan, B. S. Contextual-dependent attention effect on crowded orientation signals in human visual cortex. *J. Neurosci.* **38**, 8433–8440 (2018).

Acknowledgements

This work was supported by National Natural Science Foundation of China (61374183, 51535005), National Key Research and Development Program of China (2019YFA0705400), the Research Fund of State Key Laboratory of Mechanics and Control of Mechanical Structures (MCMS-I-0419K01), the Fundamental Research Funds for the Central Universities (NJ2020003, NZ2020001), A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Author contributions

W.L. and X.L. developed the design of the study, performed analyses, discussed the results and contributed to the text of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-28326-4>.

Correspondence and requests for materials should be addressed to X.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023