Paper in submission to SPIE Defense + Commercial Sensing: Unmanned Systems Technology XXIV (Conference 12124-20, 6 April 2022) https://spie.org/SI22/conferencedetails/unmanned-systems-technology?SSO=1

HDR luminance normalization via contextual facilitation in highly recurrent neuromorphic spiking networks

Alexander J. White [a], Chou P. Hung [b#], Chung-Chuan Lo [a,c#]
[a] Institute of Systems Neuroscience, National Tsing Hua University, Hsinchu, Taiwan
[b] DEVCOM Army Research Laboratory, Aberdeen Proving Ground, MD, USA 21005
[c] Brain Research Center, National Tsing Hua University, Hsinchu, Taiwan

ABSTRACT

Previous research has shown that many luminance normalization mechanisms are engaged when viewing scenes with high dynamic range (HDR) luminance. In one such phenomenon, areas of similar luminance contextually facilitate the perception of ambiguous textures. Using inspiration from biological circuitry, we developed a recurrent spiking neural network that reproduces experimental results of contextual facilitation in HDR images. The network uses correlations between luminance and texture to correctly classify and segment ambiguous textures in images. While many deep neural networks can successfully perform many types of image analysis, they have limited ability to process images under naturalistic HDR illumination, requiring millions of neurons and power hungry GPUs. It is an open question if a recurrent spiking neural network can minimize the number of neurons required to perform HDR image segmentation based on texture. To that end, we designed a biologically inspired proof-of-concept recurrent SNN that can perform such a task. The network is implemented using leaky integrate-and-fire neurons, with CuBa synapses. We use the Nengo LOIHI API to simulate the network, so it can be run on Intel's LOIHI neuromorphic hardware. The network uses a highly recurrent structure to both group image elements based on luminance and texture, and to seamlessly combine these modalities to correctly segment ambiguous textures. Furthermore, we can continuously modulate how much luminance or texture contribute to the segmentation. We surmise that further development of this network will improve the resilience of optical flow computations under environments with complex naturalistic illumination.

Keywords: Luminance, texture, neuromorphic, artificial neural network, contextual processing, normalization

1. INTRODUCTION

Navigating a visual environment requires making sense of areas of high illumination and dark shadows simultaneously [1, 2]. These so-called high dynamic range (HDR) luminance visual scenes often pose a challenge to standard dynamic range (SDR, 8-bit) cameras, as their sensors are limited to scenes in which the brightest and darkest pixels differ by less than 255-to-1 luminance ratio. For HDR scenes, SDR sensors either capture limited light and miss darker parts of the scene, or they become oversaturated at brighter regions [3]. A common practice to obtain HDR images from SDR sensors is to digitally blend the sensor images across multiple exposures [3-5]. However, this requires offline processing [3] or large power hungry convolutional neural networks [4, 5], and it is insufficient to address the inverse problem of recovering 3D shape by estimating surface reflectance and slant under uneven illumination. In contrast, biology solves this problem with fewer resources and limited time, using divisive normalization and other processes for contrast gain control, to compensate for uneven HDR illumination and for gaze-dependent contrast changes[6] in structurally and photically dynamic environments [2]. Moreover, the visual system makes use of contextual information, via lateral, recurrent, and feedback connections in early visual brain areas, to improve its ability to correctly adjust contrast based on surrounding context [7-12]. Controlling contrast in HDR environments is key for downstream computations such as optical flow, a critical feature of dynamic vision used in both biology and engineering. This is because uncontrolled contrast can result in errors in overestimating the speed of the flow, as extreme changes in contrast will lead the algorithm to believe objects are moving faster than reality. Thus, contrast gain control in HDR ('tonemapping') is a key preprocessing step in real-time video processing.

Previous research in humans and animals has shown that the luminance properties of nearby contours matters for estimating orientation of a target surface, and that the spatial arrangement of nearby contours ('articulation') affects the perceived lightness of a central patch [13-15]. It is thought that the brain performs luminance-feature untangling by inferring

CH: chou.p.hung.civ@army.mil, CL: cclo@mx.nthu.edu.tw

Figure 1: The need for HDR normalization.
In a high dynamic range (HDR) environment, luminance can vary over 10,000 fold (10k:1) from light areas to dark areas. Because of the low (256:1) dynamic range of standard 8-bit photos, the tunnel is either too dark at low exposure, or too saturated at high exposure. The goal of HDR normalization is to balance the luminance contrast across the image so that more of the image is visible. Our neuromorphic network is a step towards developing brain-like computations for HDR normalization for autonomous navigation.

that nearby surfaces likely receive similar illumination, then computing the shape feature of a target surface by grouping related elements [1]. It is known that contrast is strongly modulated by the luminance of surrounding elements under SDR and HDR conditions [16-18]. We recently extended this understanding of luminance-edge interactions to the HDR domain, by discovering a psychophysical phenomenon whereby the perceived orientation of a central luminance patch is biased by nearby co-oriented patches with *similar* luminance, consistent with similarity-based contextual facilitation [1, 19]. The effect was strongest for HDR images, and it depended on the specific interaction between patch luminance and flanker orientation, such that the effect was abolished if flankers had dissimilar luminance. The contextual specificity of these luminance-orientation interactions suggested that they could inform the construction of a computational model for context-dependent grouping for contrast gain control. We speculated that these behavioral phenomena could be useful signatures of the brain's computational processes, reflecting statistical regularities in visual scenes such as the tendency of nearby surfaces to receive similar illumination, which the brain then harnesses for contrast gain control for downstream processes such as optical flow and inferring shape from shading.

To build a neural network model that reproduces the empirical observations, we first revisit several findings from previous cortical research. The fact that these contextual effects are highly orientation-sensitive suggests that their computational processes begin in the primary visual cortex (V1), the first brain area with orientation-selective neurons. V1 also has tonic and phasic luminance-sensitive neurons (responding to luminance level and to relative luminance change) [20, 21], and it is thought that these luminance-sensitive and orientation-selective neurons interact to encode a higher-order percept of surface texture and possibly surface brightness and reflectance [8, 22, 23]. Besides this evidence of functional circuitry between luminance and orientation neurons, there is also physiological and anatomical evidence of strongly clustered horizontal fiber networks in V1 that could support grouping by common features, e.g. by co-orientation [24-26] or by similarity in luminance or color [27-29], but such specificity may not extend to long-range horizontal inhibitory connections [30].

Grouping of like stimuli is a consequence of the cortical microstructure [31] and inhibitory interneurons [31-33], including lateral [12, 34, 35], feedback inhibition [36, 37], and mutual inhibition [38]. Here the mutual inhibition of neurons allows cortical regions to segment similar visual information into different distinct sets of neurons with distinct neural activity [32, 38, 39]. More specifically, this is accomplished because inhibitory neurons 'decide' to what group a particular visual precept belongs [12, 33, 35, 38-40]. Moreover, once the visual precepts have been segmented into groups, they stay in these groups. This is because the segmentation process is robust to perturbation (like shifting luminance in high contrast situations), as the visual precepts are "stuck" in their basin of attraction [31, 32, 38, 39].

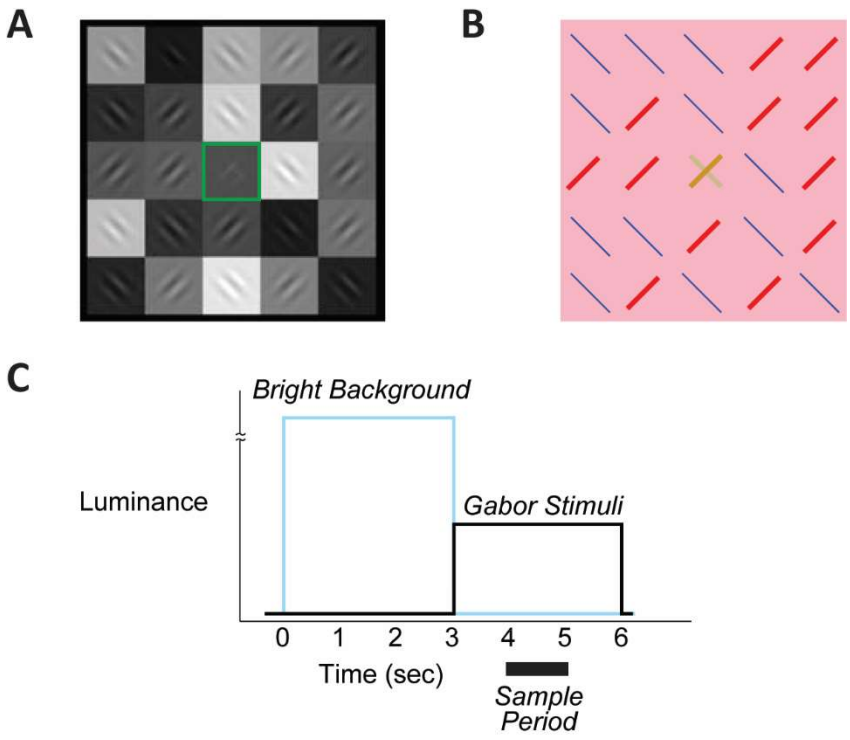

Figure 2. Experimental protocol used to test the circuit
A grid of 25 Gabor filters arranged either at a 45° angle, or 135°, with randomized luminance pedestals B: The central patch has a contrast mixture of two Gabors at 45° and 135°, and is much weaker than surrounding Gabors (average 2% contrast at center, vs. 100% contrast for surrounding Gabors). The central patch's Gabor contrast mixture was tested at regular intervals from 40%/60% to 60%/40% A/B, where A is 45° and B is 135°. C: There is a preceding 3 s bright (maximum luminance) background before the Gabor stimuli are presented for 3 s. The activity of the units is sampled between 4 and 5 s.

With this understanding of the basic computational structure for edge-luminance interactions, we sought to develop a model to explain our recent findings of contextual modulation under HDR luminance. Here, we construct a spiking neural network (SNN) to reproduce the experimental results of contextual facilitation from our recent study [1]. Our SNN makes use of LOIHI neuromorphic hardware emulated with Nengo [41, 42]. Specifically, we show that our SNN can reproduce the observed facilitation under HDR conditions, and that the facilitation disappears when the HDR condition disappears. We make use of three neural mechanisms: all-to-all excitation to boost weak ambiguous signals, global inhibition to implement divisive normalization, and local inhibitory neurons to classify textures and luminances. We show that this network architecture is able to reproduce the human experimental results.

2. METHODS

2.1 Presented Stimuli

To understand the brain circuits underlying resilient surface perception under HDR luminance, Hung et al. [1] presented human subjects with a difficult discrimination task: a full-field bright (400 cd/m^2) adapting blank stimulus was presented, followed by a collection of 25 Gabor filters with different luminance pedestals and orientations (Figure 2A and 2B). The

luminance pedestals of the 25 Gabor patches ranged from 0.4 cd/m^2 to 40 cd/m^2, and each Gabor's pixels ranged from 0.3× to 3.3× its patch's pedestal luminance, resulting in a 1111-to-1 dynamic range (3333-to-1 including the adapting blank stimulus). The subject's task was to correctly guess the central patch's Gabor orientation. The central patch had a contrast mixture of two Gabors (45° and 135°) presented simultaneously, and depending on the condition the contrast mixture was either 40%/60% or 60%/40% of the total amplitude for the 45°/135° Gabors, respectively. The total amplitude of the central Gabor was set at 34% that of flanker Gabors, resulting in a central Gabor peak-to-peak amplitude of 1.25% that of the adapting background.

The central target's perceived orientation depended on how the brain grouped that target with the background texture and luminance patterns, e.g. is the perception biased by the orientation of nearby contours with similar luminance to the target, or biased by the contours with the brightest luminance? The addition of the bright adapting blank stimulus effectively created a situation where determining the central Gabor filter's orientation was difficult, forcing the brain to reveal its mechanisms for contextual grouping for contrast gain control and tonemapping (Figure 2C).

2.2 Simulation Environment

In this paper, we use Nengo, the Python based scripting language for neural networks [41]. Nengo is able to emulate the LOIHI chip spiking neural network (SNN) developed by Intel [41, 42] . LOIHI is a neuromorphic SNN containing LIF neurons connected by Current Based (CuBa) synapses [42]. The LIF neuron is a differential equation that captures the essence of a biological neuron's action potential and its synaptic dynamics. They can be described by a system of differential equations:

$$C_m \frac{dV_i}{dt} = g_L(V_i - E_L) + u_i \quad \text{when } V_i > V_{th} \text{ then } V_i \to V_{rest}$$

$$u_i(t) = g_{ij}\, \alpha(t) * \sum_j \quad \delta_j(t - t_{spike})$$

where V is the membrane voltage, C_m is the membrane capacitance, g_L is the leak conductance, E_L is the leak current, V_{th} is the voltage threshold, and V_{rest} is the reset voltage. Here $u_i\ (t)$ is the synaptic current. Whenever V exceeds the voltage threshold V_{th}, the neuron fires a spike. This is used to calculate the synaptic current in the postsynaptic neuron. There are several variants of the mathematical form of the synaptic current. Here we use the one that is implemented in LOIHI: the delta pulse given by the equation $u_i(t) = g_{ij}\, \alpha(t) * \sum_j \quad \delta_j(t - t_{spike})$. Here the current is a convolution with exponential decay term $\alpha(t) \quad = \frac{e^{\frac{t}{\tau}}}{\tau}$ and the sum of all the presynaptic spikes $\sum_j \quad \delta_j(t - t_{spike})$. The synapse weight g_{ij} is excitatory when $g_{ij} > 0$, and it is inhibitory otherwise.

Nengo provides an easy environment for both constraining the parameters and simulation of the network. We use the standard parameters for LOIHI which are default settings for Nengo [41, 42].

2.3 Network Architecture

We construct the network using the aforementioned LIF model as implemented by LOIHI. Our goal is to take advantage of recurrent neural network structure to reproduce the experimental results in our previous study [1]. Specifically, we split up the visual field into 25 patches as a 5×5 array. Within each patch we process two types of sensory information, average luminance which we will call the luminance pedestal, and texture, which we will refer to as Gabor orientation.

Here Gabor orientation is computed as a convolution of an entire patch with a Gabor filter oriented at 45° or 135°. Here the wavelength is $\lambda = 20$ pixels. The standard deviation of the filter was set at $\sigma = 20$. Likewise, the luminance pedestal is calculated as the average luminance over the entire patch. We segregate dark and light luminance pedestals

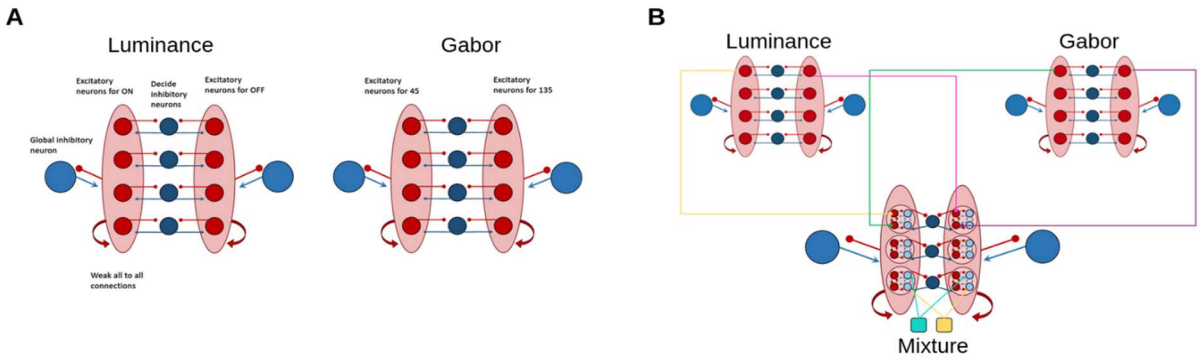

Figure 3. Neuromorphic circuit for luminance-edge integration and normalization
A: The first layer of the neural circuit, consisting of two copies of the circuit; one for pedestal luminance, and the other for Gabor filter orientation. Each circuit has two distinct groups of all-to-all excitatory neurons (red) that represent one location on the grid. Each grid location has an inhibitory neuron (blue) that only allows one stimulus group to be active at a given time. There is a global inhibitory neuron responsible for maintaining a reasonable firing rate in the all-to-all network. B: The second layer of the network combines the information from the first layer and has a very similar network structure to that in the first layer. However, each excitatory neuron is replaced with a 4-neuron circuit that has two excitatory and two inhibitory neurons. These neurons help integrate the two different modalities together. Two excitatory control neurons (cyan and yellow) excite one of the inhibitory neurons in the 4-neuron circuit and titrate the amount that a particular modality from the first layer contributes.

with ON/OFF neurons that have sigmoidal activation curves. For ON neurons the sigmoid increases with luminance, and for OFF neurons the sigmoid decreases with luminance. The threshold and steepness of the sigmoid is the same for all patch. The steepness is modulated by the variance x_{var} of all patches (if variance is zero it is set at a minimum x_{min}= 4 cd), and the threshold x_{avg} is modulated by the average luminance of all patches. $I_{gabor} = P_{patch}(x, y) * [\cos(\frac{2\pi}{\lambda}(x \pm y))\, e^{\frac{x^2+y^2}{2\sigma^2}}] =$

$$I_{off} = \frac{1}{1 + e^{\frac{x - x_{avg}}{x_{var} + x_{min}}}} \qquad I_{on} = \frac{1}{1 + e^{\frac{-(x - x_{avg})}{x_{var} + x_{min}}}}$$

Each aspect of the visual information is processed by its own similar subcircuit (see Figure 3A). We can classify stimuli across the visual field into distinct orientation groups (45° or 135°). For each of the 25 Gabor patches, there is a separate 3-neuron group. Within each group, two excitatory neurons connect to a central inhibitory neuron, which in turn inhibits the excitatory neurons. This creates a decision network, where only one of the excitatory neurons can be active at a time. Here each excitatory neuron has a Gabor filter as its receptive field. Therefore, one excitatory neuron represents a simple cell oriented for 45° stimuli, whilst the other excitatory neuron represents a different simple cell whose receptive field is oriented for 135° stimuli. In order to take into account the weaker central Gabor filter's orientation, we group the 25 excitatory neurons that are selective for a specific orientation. Within this group all excitatory neurons are connected to all other excitatory neurons in two all-to-all networks, which can boost weak inputs of similar stimuli. We also have a global inhibitory neuron to control the firing rate of the excitatory populations.

We can use a copy of the exact same circuit for the luminance pedestals. However, here the receptive fields are the ON and OFF neurons for luminance pedestals. The information similarly segregates the patchess into dark and bright patches. While the first layer captures the boosting of weak stimuli and HDR via gain normalization, they do not capture the effect of contextual modulation of similar luminance pedestals. Specifically, the behavioral experiment had a counterbalanced design in which luminance-orientation relationships of the flanking Gabor patches were defined according to two different groups of luminance pedestals and used this information to modulate the perceived orientation of the central Gabor patch. This requires integrating these two different modalities (luminance and orientation).

To do this we use a similar decision network in layer two. However the two excitatory neurons of the first layer are replaced with two copies of a well-studied 4-neuron circuit (Figure 3B) The 4-neuron circuit is a highly flexible motif that is easy to control [43]. The 4-neuron circuit has two mutually connected excitatory neurons and two mutually connected inhibitory neurons, and the inhibitory neurons provide inhibitory feedback to the excitatory neurons. This gives it the ability to flexibly integrate information, and make decisions based on the information presented to it. Moreover, the circuit is controllable with top down control into the inhibitory neurons (represented by the yellow and cyan bias currents). This allows the 4-neuron circuit to titrate which stimuli property (Gabor orientation or pedestal luminance) to use to classify information. In the case that one property is ambiguous (e.g. orientation), the other property (e.g. luminance) can be used to make the classification.

Like layer 1 the classification is mediated by a central inhibitory neuron that inhibits the 4 neuron-circuit's excitatory neurons. In this way, at least one of the 4-neuron circuits will be active, thus segregating the image into two populations. Just like layer 1, there are also two global inhibitory neurons to help constrain firing rates.

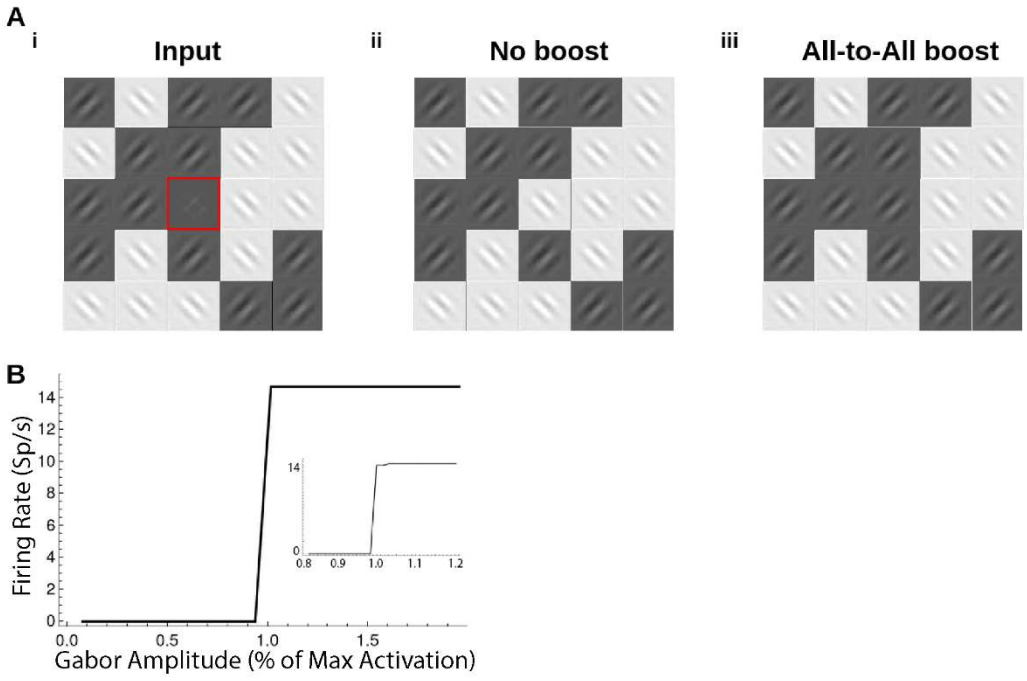

Figure 4. Neuromorphic Gabor filter output

A. Visualization of neuromorphic Gabor filters outputs. (Ai) Raw input into the Gabor Orientation network. In this example, 45° Gabors are at darker patches, and 135° Gabors are at lighter patches. The central patch (highlighted by red box) contains both 45° and 135° Gabors at weaker 2% amplitude, resulting in an ambiguous orientation input. Of the 24 surrounding patches, the 12 45° patches have similar luminance as the central target, whereas the remaining 12 135° patches have a different luminance. (Aii and Aiii) The output shows how the decision network classified the input orientation. (Aii) shows the results without all-to-all boosting. Note the misclassification of the central patch orientation as 135°. (Aiii) shows the results with all-to-all boosting, with facilitation of the 45° orientation consistent with psychophysical phenomenon of lateral 'facilitation' of a weak co-oriented central Gabor by flankers with similar luminance as the target. B. A parameter sweep shows the sensitivity to central Gabor amplitude required to elicit a boosted response. The central Gabor amplitude is swept from 0.1% amplitude to 2% amplitude. (B inset) Shows an expanded view of the transition amplitude. Note the sudden increase in firing rate at about 1% amplitude.

3. RESULTS

In order to test the network, we first examined whether the network is able to boost the signal of a weak ambiguous pair of Gabors. We presented the network with a square grid of 25 patches with 12 Gabors oriented at 45° and 12 Gabors oriented at 135°. The central patch is a contrast mixture of two Gabors oriented at 45° and 135°. A sample input to the network is shown in Figure 4Ai. The central Gabor has 2% of the amplitude of that of the other 24 Gabors. We consider two conditions. First, we consider the network without all-to-all connections between the excitatory neurons, by setting their all-to-all connection weights to zero (Figure 4Aii). Without boosting, the network is unable to detect and correctly classify the central Gabor's signal. As such both the 135° and 45° neurons remain OFF, or with injected noise, they are both equally activated. However, with the synaptic weights of the all-to-all connections set to non-zero, the network is suddenly able to classify the weak central Gabor and correctly assign it to 45°, due to co-excitation from neighboring 45° patches that have similar luminance to the central target (Figure 4Aiii). These results are consistent with thepsychophysical phenomenon that flankers with similar orientation as a weak central target can facilitate and boost that target's detectability.

Next we determine how sensitive the network is to weak Gabors, and attempt to quantify the minimum amplitude required to boost the signal. To test this we tested the network with a 60%:40% mixture of two Gabors at 45° and 135°, and incrementally increased its maximum amplitude from 0.1% to 2.0%. As shown in Figure 4B, we see a sudden jump at approximately 0.98% the amplitude of the flanker Gabors. Interestingly, we see a sudden jump in the firing rate of the neurons. This hints at an underlying saddle-node bifurcation that is common in all-to-all excitatory networks [32, 43].

After establishing that all-to-all excitation can boost weak signals, we tested whether the whole network is able to facilitate identification of an ambiguous central Gabor in the Luminance Similarity condition and cannot facilitate identification of an ambiguous central Gabor in the Luminance Brightest condition. This mirrors results presented in our previous human behavioral study [1]. Also consistent with our human results, facilitation was only present when there was a preceding flash, tested at 50 to 1000 times brighter than the luminance pedestal (Figure 2C).

In both conditions all luminance pedestals are ranked-ordered from brightest to least brightest. The central Gabor is always the 13th brightest. In the 'Similar' luminance condition (Figure 5A), the luminance pedestals of all Gabors oriented 45° (135°) ranged from 7th brightest to 18th brightest. That is, the central Gabor's orientations match patches with similar brightness. In the 'brightest' condition, the luminance pedestals of all Gabors oriented 45° (135°) range from 1st brightest to 13th brightest. That is, the central Gabor's orientation matches the brightest patchess.

To test the network's performance, we measured the average firing rate, during 1-2 second period post Gabor onset, of both excitatory neurons in the 4-neuron circuit associated with the correct answer. We also measured the average firing rate of both excitatory neurons in the 4-neuron circuit associated with the incorrect answer. We considered the network output to be correct if there were a positive difference between the correct and incorrect answer. As expected, we found that in the Luminance Similarity condition the network was more likely to classifying the central Gabor correctly (Figure 5B). Moreover, it facilitates the classification of ambiguous stimuli of 45%:55% (Figure 5C left, blue curve). That is, the network was able to resolve the ambiguity, and used correlations in brightness information to choose the correct orientation. Furthermore, target stimuli that are perfectly ambiguous (50%:50% contrast mixture) are about 15% more likely than chance to be classified as 45° orientations.

However, in the Luminance Brightest condition the network was unable to resolve the ambiguity, as the correlation between flanker and target orientation was perfectly balanced between bright and dark. As a result, brightness provided no useful information to help classify the stimuli. This was evident as for ambiguous target contrast mixtures near 50%:50% (Figure 5C left, yellow), the target was less likely to be identified as 45°, and the Luminance Similarity condition was about 30% more likely to be identified as 45° compared to the Luminance Brightest condition.

As expected from our previous human behavioral results, when the preceding flash was removed, the facilitation effect was abolished (Fig. 5C right). This is evidenced by the overlap of the curves for the Luminance Similarity and Luminance Brightest conditions. Specifically, when the preceding flash was removed, there is no more facilitation (as evidenced by a lack of leftward shift away from the 50%/50% mark). Moreover, the difference between the two conditions was statistically insignificant.

As a final demonstration we tested the ability of the first layer of the network to interpret a synthetic image. Our image (Figure 6A) was composed of 45° sinusoidal gratings at a range of contrasts (0 to 50% Michelson contrast against a gray background), spanning the entire image, with the strongest contrasts along the bottom-left to top-right diagonal. At the

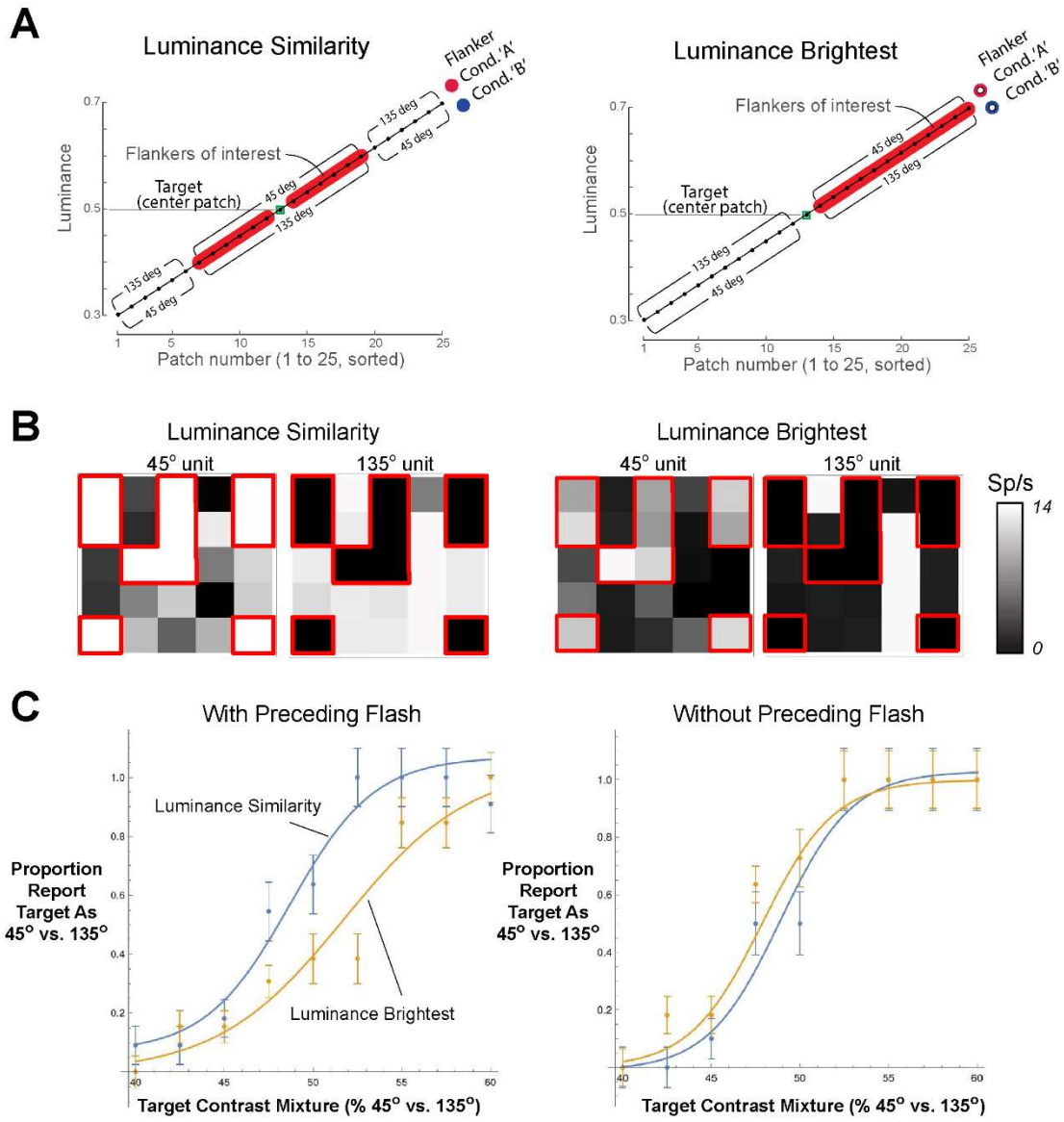

Figure 5. Neuromorphic facilitation depends on luminance similarity

A. Schematic of 25 patches (abscissa) sorted by luminance level (ordinate). The central Gabor orientation can match either the orientation of the brightest patches ('Luminance Brightest' condition, right) or the orientation of the patches with the most similar luminance ('Luminance Similarity' condition). In each of these conditions, Cond. A is when the relevant flanker patches are oriented 45°, and Cond. B is when the relevant flanker patches are oriented 135°. B. Example activity levels of 45° and 135° units across Luminance Similarity and Luminance Brightest conditions for Cond. A. Gray shading indicates unit activity level, not patch luminance. C. Psychometric curve showing proportion of trials in which the unit for the central patch reports 45° (ordinate) across variations in the actual contrast mixture of the central target for 45° vs 135° (abscissa). Each point is based on 25 trials. *Left*, As expected, the target reports are correct for the extreme conditions at left and right (over 60%:40 or less than 40%:60% contrast mixture). For ambiguous stimuli, the output is biased towards 45° in the Luminance Similarity condition (blue), when the relevant flankers are also 45°. The bias is abolished for the Luminance Brightest condition (yellow). *Right*, The facilitation effect is also abolished in the absence of the preceding luminance flash. These results are consistent with our recent human behavioral study (Hung et al. 2020).

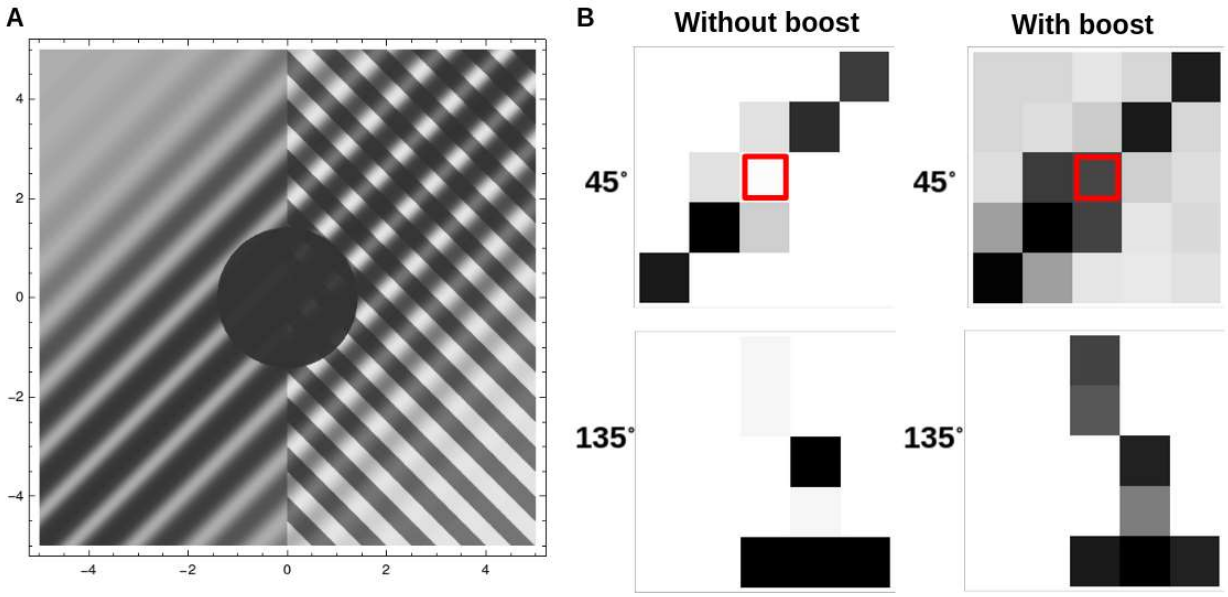

Figure 6. Neuromorphic facilitation of a cast shadow region
A sample image with gratings processed by the network. A: The image is comprised of 45° sinusoidal gratings with contrast varying across the image, with additive 135° square-wave biphasic gratings on the right side of the image. The center of the image has a shadow disc that reduces grating contrast to 5%. B: Results of the first layer of the network for 45° units (top row) and 135° units (bottom row). Darker shading indicates stronger unit activity. Without boosting (left column), neither grating is detected under the shadow. Boosting (right column) enhances the detection of the 45° grating under the circular shadow (red square), consistent with the more prevalent 45° grating in surrounding regions.

right half of the image, these are additively combined with 135° square-wave gratings of fixed 50% contrast (symmetrical biphasic duty cycle). The appearance is that of a wrinkled sheet with physical undulations oriented 45° but textured stripes oriented 135°. The center of the image is overlaid by a dark opaque disc, by dividing the luminance of each pixel by 10, resulting in a contrast of 5% for each grating. This disc has the appearance of a cast shadow that greatly weakens the apparent contrast of the two gratings, but they both remain discernable. The network divides the image into a 5×5 set of regions ('patches') and pre-processes the image by computing the luminance contrast magnitude of each of the 45° and 135° orientations within each patch, by convolving with a sinusoidal filter. The goal of the network was to report the apparent contrast strength at 45° and 135° orientations in each patch.

Fig 6B shows the results of the network with and without the all-to-all boosting. In both cases, the neurons sensitive to 45° stimuli detected the strongest sinusoidal gratings, and the 135° neurons detected the square-wave gratings in areas not overlaid by the sinusoidal gratings. In regions where both orientations are strong, the network result is dominated by 45° because of its stronger sensitivity to sinusoidal gratings. The effect of boosting can be seen at the central patch. Without boosting, the network fails to identify both 45° and 135°. With boosting, the network detected the 45° grating but not the 135° grating, consistent with perceptual facilitation by the more prevalent 45° flankers. Boosting also elevated the sensitivity to the low contrast 45° grating at the upper left and bottom right corners. At the top middle of the image, boosting also elevated the sensitivity to the 135° grating when it overlaid a medium-contrast 45° grating.

4. CONCLUSION

We constructed a LOIHI-implementable spiking neural network that reproduced human behavioral results showing that grouping by HDR luminance similarity can bias orientation perception. Moreover, we have shown that we can process and extract textural information from complex stimuli using a small and energy efficient circuit. Using a recurrent neural circuit, we were able to reproduce the observed effect of facilitation of ambiguous stimulus orientation, depending on the specific combination of flanker orientation and luminance similarity. This implies that the neural mechanisms of weak all-

to-all excitations, divisive normalization from global inhibitory neurons, and inhibition-based local decision making are sufficient to explain the phenomenon of HDR luminance-based contextual grouping after sudden darkening. Although our computations are based on SDR images, we suggest that it is very feasible to extend this tonemapping and feature grouping effect to HDR images by incorporating logarithmic luminance sensors as a front-end [44].

5. DISCUSSION AND FUTURE WORK

While our preliminary neuromorphic results are able to reproduce the psychophysical phenomenon of 'facilitation' (boosting weak stimuli via feature grouping), and while our results show extensibility to HDR contrast gain normalization and tonemapping, we have not yet attempted to replicate all of our previous psychophysical effects of contextual modulation. Specifically, our previous behavioral experiment tested for contextual grouping effects across three different groups of luminance pedestals (three different types of 'articulation' comprised of different combinations of up to 25 different luminance pedestals). However, our model only classifies luminance pedestals into two groups, and does not consider more fine grained classification of stimuli.

To incorporate more groupings, we need to increase the number of luminance groups to three or more. This is relatively easy by simply expanding the number of excitatory groups we have in the first layer, as one decision inhibitory neuron can easily service multiple excitatory neurons. In order to correctly synthesize this information in the second layer, we will need to introduce a simplified version of a dendritic arbor [45]. Here the dendritic arbor is responsible for multimodal integration, as each branch of the dendritic tree represents one possible group of pedestal luminances. An inhibitory control neuron uses shunting inhibition to gate which dendritic branch is silenced. This is based on the local connectivity of VIP and SST neurons that is key for incorporating contextual information in the cortex [33]. This neural architecture could allow the network to handle arbitrary synthesis and perception of groups while still reproducing the particular results of the experiment. Moreover, because of the dendritic tree, this model could scale easily to other modalities such as color, shape, or texture.

Nonetheless, our modeling results confirm that such that our network is capable of grouping stimuli based on the luminance and texture. Moreover, it can handle spatial variability (variability in Gabor amplitude and luminance brightness). However, we have not dealt with the temporal variability. Moving forward, we will need to construct decision networks that integrate evidence over a short period of time. Specifically making use of basins of attraction that are not too sensitive to random fluctuations over a short period of time. This has been well worked out in the literature [31].

Our circuit has not been tested with natural images. Extending these results to real natural images will be a key goal moving forward, for tasks such as autonomous flying[46]. Environments such as wooded areas with sun shining and wind blowing through a canopy of leaves, or an unilluminated building with a sunlit outside are common HDR environments. However, extending this network to natural scenes such as this will require a large increase in neuron number and number of layers. Moreover, any investigation of the usefulness of our networks needs to compare them to feedforward models such as convolutional neural networks (CNN) [47, 48] and connectionist models [49, 50]. Furthermore, it would be nice to compare the statistics of these models to psychophysics results. A good framework for achieving such a comparison with maximal likelihood has been laid out by others [51].

However, given any optical flow algorithm, there is often an error associated with changing luminance and with estimating shape under uneven illumination [1, 2, 52] (Figure 1). This is because optical flow is generally computed frame to frame. If luminance changes suddenly (as would occur from moving shadows in a windy forest), there could be widely changing optical flow values. This can make tasks such as horizon tracking [53] and object tracking and avoidance difficult. The problem is compounded under uneven illumination, because estimating surface slant and shape is interdependent with estimating surface reflectance. This 'inverse problem' of estimating 3D shape from a 2D image requires biological visual systems to make simplifying assumptions, and it is thought that one assumption is that nearby or related surfaces are under similar illumination. Our neuromorphic circuit linking orientation and luminance computations provides a possible structure for untangling the interdependencies of nearby surface contrasts.

In the future, we plan to apply our model to improve optical flow performance in high-contrast environments by creating stable representation of contrast and texture across different lighting levels for realistic videos. Moreover, we plan to use the neuromorphic LOIHI chip as a module that can be coupled with optical flow algorithms to reduce the error in

HDR environments. More specifically, our circuit has the potential to act as a top-down control (e.g. to selectively emphasize certain modalities) for the preprocessing step for video information. By contextually normalizing and stabilizing the contrast and texture in each frame, ideally any optical flow algorithm will show improvement in estimated optical flows.

6. ACKNOWLEDGMENTS

We thank Andre Harrison and MaryAnne Fields for helpful discussions. This material is based on work supported by the International Technology Center Indo-Pacific (ITC IPAC), Army Research Office, and Office of Naval Research-Global. It was accomplished under Contract No. FA5209-21-P-0189 and Cooperative Agreement Number FA5209-21-R-0019. The views, findings, and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the DEVCOM Army Research Laboratory, ITC IPAC, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

REFERENCES

[1] C. P. Hung, C. Callahan-Flintoft, P. D. Fedele *et al.*, “Abrupt darkening under high dynamic range (HDR) luminance invokes facilitation for high contrast targets and grouping by luminance similarity,” J Vision, 20, 9 (2020).

[2] A. A. Lazar, N. H. Ukani, and Y. Zhou, “Sparse identification of contrast gain control in the fruit fly photoreceptor and amacrine cell layer,” The Journal of Mathematical Neuroscience, 10(1), 1-35 (2020).

[3] D. Hafner, O. Demetz, and J. Weickert, "Simultaneous HDR and optic flow computation," 2014 22nd International Conference on Pattern Recognition. 2065-2070 (2014).

[4] W. Wu, M. Kan, X. Liu *et al.*, "Recursive spatial transformer (rest) for alignment-free face recognition," Proceedings of the IEEE International Conference on Computer Vision. 3772-3780 (2017).

[5] Y. Fang, B. Zhan, W. Cai *et al.*, "Locality-constrained spatial transformer network for video crowd counting," 2019 IEEE International Conference on Multimedia and Expo (ICME). 814-819 (2019).

[6] J. E. Niemeyer, and M. A. Paradiso, “Contrast sensitivity, V1 neural activity, and natural vision,” Journal of neurophysiology, 117(2), 492-508 (2017).

[7] E. H. Adelson, "Lightness Perception and Lightness Illusions", in [The New Cognitive Neurosciences, 2nd ed.] MIT Press, Cambridge, MA, 24, pp. 339-351, (2000).

[8] C. P. Hung, B. M. Ramsden, and A. W. Roe, “A functional circuitry for edge-induced brightness perception,” Nature neuroscience, 10(9), 1185-1190 (2007).

[9] T. D. Albright, and G. R. Stoner, “Contextual influences on visual processing,” Annual review of neuroscience, 25(1), 339-379 (2002).

[10] C. D. Gilbert, and W. Li, “Top-down influences on visual processing,” Nature Reviews Neuroscience, 14(5), 350-363 (2013).

[11] T. S. Lee, “Contextual influences in visual processing,” (2008).

[12] Y.-S. Lin, C.-C. Chen, and M. W. Greenlee, “The role of lateral modulation in orientation-specific adaptation effect,” Journal of Vision, 22(2), 13-13 (2022).

[13] C.-C. Chen, and C. W. Tyler, “Lateral modulation of contrast discrimination: Flanker orientation effects,” Journal of Vision, 2(6), 8-8 (2002).

[14] I. K. Zemach, and M. E. Rudd, “Effects of surround articulation on lightness depend on the spatial arrangement of the articulated region,” JOSA A, 24(7), 1830-1841 (2007).

[15] F. A. Kingdom, “Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy,” Vision Research, 51(7), 652-673 (2011).

[16] J. Xing, and D. J. Heeger, “Measurement and modeling of center-surround suppression and enhancement,” Vision research, 41(5), 571-583 (2001).

[17] S. P. MacEvoy, and M. A. Paradiso, “Lightness constancy in primary visual cortex,” Proceedings of the National Academy of Sciences, 98(15), 8827-8831 (2001).

[18] S. R. Allred, A. Radonjic, A. L. Gilchrist *et al.*, "Lightness perception in high dynamic range images: local and remote luminance effects," J Vis, 12(2), 7, 1-16 (2012).
[19] C. P. Hung, C. Callahan-Flintoft, A. J. Walker *et al.*, "A 100,000-to-1 high dynamic range (HDR) luminance display for investigating visual perception under real-world luminance dynamics," Journal of Neuroscience Methods, 108684 (2020).
[20] C. P. Hung, B. M. Ramsden, L. M. Chen *et al.*, "Building surfaces from borders in Areas 17 and 18 of the cat," Vision research, 41(10-11), 1389-1407 (2001).
[21] J. Dai, and Y. Wang, "Representation of surface luminance and contrast in primary visual cortex," Cerebral cortex, 22(4), 776-787 (2012).
[22] A. W. Roe, H. D. Lu, and C. P. Hung, "Cortical processing of a brightness illusion," Proceedings of the National Academy of Sciences, 102(10), 3869-3874 (2005).
[23] D. A. Ruff, D. H. Brainard, and M. R. Cohen, "Neuronal population mechanisms of lightness perception," Journal of Neurophysiology, 120(5), 2296-2310 (2018).
[24] D. Ts'o, and C. D. Gilbert, "The organization of chromatic and spatial interactions in the primate striate cortex," Journal of Neuroscience, 8(5), 1712-1727 (1988).
[25] R. Malach, R. B. Tootell, and D. Malonek, "Relationship between orientation domains, cytochrome oxidase stripes, and intrinsic horizontal connections in squirrel monkey area V2," Cerebral Cortex, 4(2), 151-165 (1994).
[26] K. E. Schmidt, R. Goebel, S. Löwel *et al.*, "The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex," European Journal of Neuroscience, 9(5), 1083-1089 (1997).
[27] M. M. Chernov, R. M. Friedman, G. Chen *et al.*, "Functionally specific optogenetic modulation in primate visual cortex," Proc Natl Acad Sci U S A, 115(41), 10505-10510 (2018).
[28] C. C. J. Chu, P. F. Chien, and C. P. Hung, "Tuning dissimilarity explains short distance decline of spontaneous spike correlation in macaque V1," Vision Res, 96, 113–132 (2014).
[29] M. S. Livingstone, and D. H. Hubel, "Specificity of intrinsic connections in primate primary visual cortex," Journal of Neuroscience, 4(11), 2830-2835 (1984).
[30] Z. F. Kisvárday, and U. T. Eysel, "Functional and structural topography of horizontal inhibitory connections in cat visual cortex," European Journal of Neuroscience, 5(12), 1558-1572 (1993).
[31] X.-J. Wang, "Decision making in recurrent neuronal circuits," Neuron, 60(2), 215-234 (2008).
[32] A. J. White, P.-H. B. Liu, and C.-C. Lo, "Multi-functional microcircuits: Robust, flexible control of functionality in small recurrent networks," Annual Meeting of the Society for Neuroscience. Nanosymposium 728.11. (2019).
[33] X.-J. Wang, and G. R. Yang, "A disinhibitory circuit motif and flexible information routing in the brain," Current opinion in neurobiology, 49, 75-83 (2018).
[34] U. Polat, and D. Sagi, "Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments," Vision research, 33(7), 993-999 (1993).
[35] J. R. Cavanaugh, W. Bair, and J. A. Movshon, "Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons," Journal of neurophysiology, 88(5), 2530-2546 (2002).
[36] Y. Petrov, and S. P. McKee, "The effect of spatial configuration on surround suppression of contrast sensitivity," Journal of vision, 6(3), 4-4 (2006).
[37] L. Schwabe, K. Obermayer, A. Angelucci *et al.*, "The role of feedback in shaping the extra-classical receptive field of cortical neurons: a recurrent network model," Journal of Neuroscience, 26(36), 9117-9129 (2006).
[38] N. Kogo, F. B. Kern, T. Nowotny *et al.*, "Dynamics of a Mutual Inhibition Circuit between Pyramidal Neurons Compared to Human Perceptual Competition," Journal of Neuroscience, 41(6), 1251-1264 (2021).
[39] M. Koyama, and A. Pujala, "Mutual inhibition of lateral inhibition: a network motif for an elementary computation in the brain," Current opinion in neurobiology, 49, 69-74 (2018).
[40] A. Angelucci, M. Bijanzadeh, L. Nurminen *et al.*, "Circuits and mechanisms for surround modulation in visual cortex," Annual review of neuroscience, 40, 425-451 (2017).
[41] T. Bekolay, J. Bergstra, E. Hunsberger *et al.*, "Nengo: a Python tool for building large-scale functional brain models," Frontiers in neuroinformatics, 7, 48 (2014).
[42] M. Davies, N. Srinivasa, T.-H. Lin *et al.*, "Loihi: A neuromorphic manycore processor with on-chip learning," Ieee Micro, 38(1), 82-99 (2018).
[43] B. Liu, A. J. White, and C.-C. Lo, "Augmenting Flexibility: Mutual Inhibition Between Inhibitory Neurons Expands Functional Diversity," bioRxiv, 2020.11. 08.371179 (2021).

[44] J. Li, O. Skorka, K. Ranaweera *et al.*, "Novel real-time tone-mapping operator for noisy logarithmic cmos image sensors," Electronic Imaging, 2016(12), 1-13 (2016).
[45] S. Li, N. Liu, X. Zhang *et al.*, "Dendritic computations captured by an effective point neuron model," Proceedings of the National Academy of Sciences, 116(30), 15244-15252 (2019).
[46] P. P. Parlevliet, A. Kanaev, C. P. Hung *et al.*, "Autonomous Flying With Neuromorphic Sensing," Frontiers in Neuroscience, 15, 672161 (2021).
[47] M. Barni, A. Costanzo, E. Nowroozi *et al.*, "CNN-based detection of generic contrast adjustment with JPEG post-processing," 2018 25th IEEE International Conference on Image Processing (ICIP). 3803-3807 (2018).
[48] G. Cao, Y. Zhao, R. Ni *et al.*, "Contrast enhancement-based forensics in digital images," IEEE transactions on information forensics and security, 9(3), 515-525 (2014).
[49] J. D. Cohen, K. Dunbar, and J. L. McClelland, "On the control of automatic processes: a parallel distributed processing account of the Stroop effect," Psychological Review, 97(3), 332 (1990).
[50] P. Hoffman, J. L. McClelland, and M. A. Lambon Ralph, "Concepts, control, and context: A connectionist account of normal and disordered semantic cognition," Psychological Review, 125(3), 293-328 (2018).
[51] N. Prins, and F. A. Kingdom, "Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox," Frontiers in Psychology, 9, 1250 (2018).
[52] P. O. Moore, [Nondestructive Testing Handbook Vol. 9 Visual Testing], American Society for Nondestructive Testing, Inc., (2010).
[53] G. Bebis, R. Boyle, B. Parvin *et al.*, [Advances in Visual Computing: 7th International Symposium, ISVC 2011, Las Vegas, NV, USA, September 26-28, 2011. Proceedings, Part I], Springer, (2011).