



# A biologically-based computational model of visual cortex that overcomes the X-junction illusion

Parvin Zarei Eskikand <sup>a,\*</sup>, Tatiana Kameneva <sup>a,b</sup>, Michael R. Ibbotson <sup>c</sup>, Anthony N. Burkitt <sup>a</sup>, David B. Grayden <sup>a</sup>

<sup>a</sup> NeuroEngineering Laboratory, Department of Biomedical Engineering, The University of Melbourne, Parkville, Australia

<sup>b</sup> Faculty of Science, Engineering and Technology, Swinburne University of Technology, Australia

<sup>c</sup> National Vision Research Institute, Australian College of Optometry, Carlton, Australia

## ARTICLE INFO

### Article history:

Received 28 September 2017

Received in revised form 24 January 2018

Accepted 9 February 2018

Available online 16 February 2018

### Keywords:

Vision

Neural model

Motion perception

Middle temporal (MT)

X-junction illusion

## ABSTRACT

The end-points of a moving bar (intrinsic terminators) contain unambiguous information that can be used to extract the bar's correct direction of motion, regardless of the orientation of the bar. However, extrinsic terminators, formed at the intersection of two overlapping bars, can result in motion signals with conflicting directions compared to those of the intrinsic terminators. Using a computational model, we propose that interactions between form and motion information may assist neurons in the motion-specific regions of primate cortex to differentiate intrinsic from extrinsic terminators. The motion processing model has two stages. The first stage is a model of V1 complex neurons, including end-stopped neurons. The resulting first stage motion signals are transmitted to the second stage, which is a model of MT neurons. In the proposed model, MT neurons additionally receive form information from neurons in V1 that are not orientation or direction selective but respond strongly to the contrast of the stimulus. These neurons have polarity-dependent center-surround receptive fields, as found in layer 4 of V1 in primates. As the inhibitory surrounds of these neurons are less activated at the intrinsic terminators, the signals generated by the end-points of the objects are stronger than the signals from the extrinsic terminators, which are inhibited by strong suppression from the surround. Therefore, the excitatory inputs received by integration MT neurons from center-surround V1 neurons enhance the unambiguous motion signals at the intrinsic terminators, which therefore dominate over the local motion signals generated at X-junctions. The results show that, despite the inability of V1 end-stopped neurons to distinguish between the two different types of terminators, center-surround V1 neurons provide the capacity for the second stage of the model to preferentially respond to the intrinsic terminators and, therefore, predict the true directions of the crossing bars.

© 2018 Elsevier Ltd. All rights reserved.

## 1. Introduction

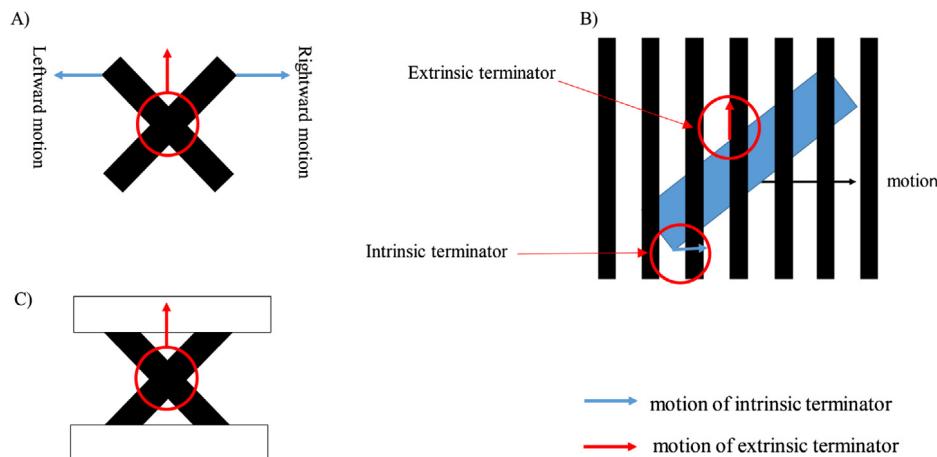
V1 neurons that respond to the edge of a moving stimulus are not always able to measure the correct direction of the motion because of their small receptive fields. They only respond to the component of the motion that is perpendicular to the edge of the stimulus. This is known as the aperture effect. The end-points of a bar (e.g. corners) are known as intrinsic terminators and carry information allowing the prediction of the correct direction of bar motion, irrespective of its orientation. In contrast, the motion signals resulting from extrinsic terminators, which are formed at the corners where bars overlap with each other, transmit conflicting directional motion information (Berzhanskaya, Grossberg, &

Mingolla, 2007; Hubel & Wiesel, 1965; Pack, Livingstone, Duffy, & Born, 2003). These extrinsic terminators need to be differentiated from the intrinsic terminators to determine the correct directions in which the individual bars are moving. For example, Fig. 1A illustrates that when two overlapping bars that are oriented 90° apart (in this case, +45° and -45° relative to horizontal) are moving in opposite directions in the horizontal plane, the central junction will appear to move up or down. If the intrinsic terminators of the bars are occluded, then the bars are perceived as a single object whose motion is in the direction of apparent movement of the central X-junction (Fig. 1C). This phenomenon is referred to as the chopstick illusion (Anstis, 1990).

As illustrated in Fig. 1B, the direction of movement of most objects that are passing behind a picket fence can be perceived with no difficulty. This shows that the visible occluders do not interfere with the perception of movement, and the visual system can integrate the different sections of the stimulus to identify the correct

\* Corresponding author.

E-mail address: [pzarei@unimelb.edu.au](mailto:pzarei@unimelb.edu.au) (P.Z. Eskikand).



**Fig. 1.** The effects of intrinsic and extrinsic terminators on motion perception. (A) The X-junction formed as the result of two overlapped crossing bars moving in opposite directions. (B) The chopstick illusion. The endpoints of a diagonal bar, known as intrinsic terminators, carry unambiguous motion information, shown by the blue arrows. The terminators formed as the result of overlapping with other objects are known as extrinsic terminators. The global motion information is not affected by the local motion signals representing different directions of motion. Therefore, these local motion signals at the extrinsic terminators may be suppressed by MT neurons before their propagation to other regions. The red arrows show the directions of local motions at the extrinsic terminators. (C) The bars are perceived as a single object moving in the upward direction when the end-points of the bars are occluded. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

direction of motion of single moving objects (Berzhanskaya et al., 2007; Chey, Grossberg, & Mingolla, 1997; Dresp, Durand, & Grossberg, 2002; Grossberg, 1994; Grossberg, Mingolla, & Viswanathan, 2001). The global perception of the direction of movement of a bar behind the occluders, which matches the direction of motion of the intrinsic terminators, is not affected by the local directions of motion that result from the extrinsic terminators apparent at the intersections of the bar with the occluders.

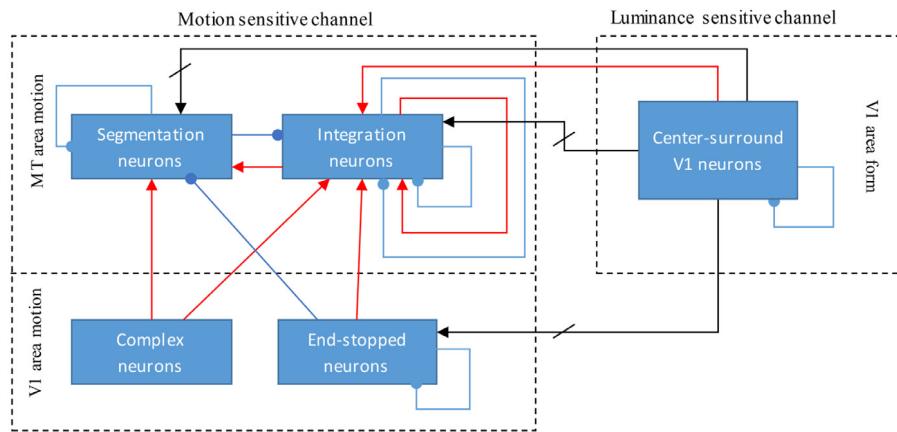
It is often thought that motion information is processed exclusively in the dorsal visual pathway and form information in the ventral pathway (Goodale & Milner, 1992; Ibbotson, 2007; Landau & Jackendoff, 1993; Ungerleider, 1982). However, there is electrophysiological evidence showing that the division of the form and motion processing pathways is not strict and form information may be involved in the perception of motion in the dorsal stream. For example, Kourtzi and Kanwisher (2000) demonstrated activity of neurons in MT and MST when stimulated by static images that imply motion, such as a snapshot of a human or animal in action. The results of this experiment demonstrate the interactions of the neurons in areas MT and MST with form processing mechanisms (Kourtzi & Kanwisher, 2000). Krekelberg, Dannenberg, Hoffmann, Bremmer, and Ross (2003) also recorded the activity of MT and MST neurons in response to static Glass patterns in which the arrangement of the dots generates an illusion of movement. The neurons fire in response to Glass patterns, with the directional tuning curves matching the responses of the same neurons to actual motion. This provides further evidence of the combination of form and motion information in the dorsal pathway (Krekelberg et al., 2003).

There are several modeling studies that proposed the interaction of form and motion information in the visual cortex for the accurate perception of motion direction. The models of Berzhanskaya et al. (2007) and Grossberg (1994, 2015) use an interaction between form and motion information that leads to the perception of 3D form from a 2D image when an object occludes another. The foundation of these models is based on the FASCADE (Form And Color And Depth) theory, which explains how the boundaries of an object occluded by another object are completed, thus generating the perception of a single object (Chey et al., 1997; Grossberg, 1994). The FASCADE model is based on interactions between motion, form, and depth features. It is made of two separate systems to cover the properties of the surface of the input image and boundary

segmentation. This is a hierarchical model of different areas of the visual cortex that is composed of three systems: feature-ground separation to discriminate the occluders from the occluding object properties, a motion processing stream, and a system to integrate motion and form information (Chey et al., 1997; Grossberg, 1994).

Tlapale, Masson, and Kornprobst (2007) proposed a model to deal with the X-junction problem. It uses interactions between form and motion information and is an extension of Bayerl and Neumann's model (Tlapale et al., 2007). The process for estimating the correct direction of motion arises through the feedback connections from MT to V1, which results in the propagation of unambiguous motion signals from terminators to other regions in V1. However, the difference is that the authors used luminance information to change the level of integration of local motion signals. If the level of the luminance of two neighboring pixels is similar, MT neurons integrate the motion signals more robustly than in locations where the luminance does not match (Tlapale et al., 2007). Although this is an interesting method to segregate moving objects that have different luminance, it would fail to segregate overlapping objects with the same contrast levels and, therefore, propagate the motion signals of the extrinsic terminators to other regions. In addition, the process of measuring the similarity between the luminance of the pixels is not referred to in the model and is not associated with any particular type of known visual receptive field.

Here, we present a model that combines form and motion information to generate signals that represent the correct direction of motion in the presence of occluders. Their suppressive surrounds influence the activities of these neurons. The Hermann illusion (Brewster, 1844) is evidence for the existence of visual neurons that have center-surround interactions. It is commonly thought that this illusion is caused by the receptive field characteristics of the ON- and OFF-center receptive fields found in the retina and lateral geniculate nucleus (Spillmann, 1994; Wolfe, 1984). However, Schiller and Carvey (2005) demonstrated that the appearance of this illusion requires the involvement of cortical V1 neurons. This is due to the orientation selective and size invariant nature of the illusion, which cannot be explained by the receptive fields of retinal/geniculate receptive fields; the illusion must involve orientation selective and multiple-size tuned cells, which imply cortical involvement (Schiller & Carvey, 2005). Experiments show that the appearance of this illusion is not size dependent and the effect



**Fig. 2.** The structure of the model. The red arrows represent excitatory connections, blue lines ending in filled circles are inhibitory connections between neurons, and black arrows are connections that generate a contrast-dependent threshold that regulates neural activity. Visual information is processed through two different pathways. The motion pathway has two different stages. The initial motion information is provided by two sets of V1 neurons: standard complex and end-stopped neurons. There are excitatory and inhibitory interconnections between integration MT neurons and segmentation MT neurons. MT neurons receive excitatory input from the center-surround neurons generating form-related information in V1. The activities of these luminance sensitive center-surround neurons determine the level of the threshold that regulates the activity of the neurons responding to the movement. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

greatly diminishes when there is a distortion of the orientation information in the grids (Schiller & Carvey, 2005). The influence of the area beyond the excitatory receptive field of V1 neurons on their activities has been demonstrated in many studies (Allman, Miezin, & McGuinness, 1985; Cavanaugh, Bair, & Movshon, 2002a; Hubel & Wiesel, 1968) that provide neurophysiological evidence of the suppressive effect of the surround in some V1 neurons when the stimulus in the surround has the same characteristics as the stimulus in the center (i.e., both are bright or both are dark).

The model presented here suggests that contrast sensitive, center-surround neurons are utilized to calculate the correct direction of motion of overlapping bars that move in different directions. The responses of these neurons at the extrinsic terminators are significantly lower than the responses of neurons at the intrinsic terminators. It is proposed that this assists high level motion processing neurons to differentiate the responses to the extrinsic terminators from the unambiguous motion signals at the intrinsic terminators, and prevent extrinsic terminator signals from propagating to the whole of the moving object. The proposed models by Beck and Neumann (2010, 2011) and Tlapale et al. (2007) use feedback connections from areas MT and V1 to propagate unambiguous activities from MT to V1 neurons, suggesting additional processing that evolves over time. However, neurophysiological evidence shows that the activity of V1 neurons does not change over time, making such feedback connections controversial. In contrast to these previous studies, our model demonstrates the necessity of only center-surround V1 neurons that are sensitive to stimulus contrast.

## 2. Methods

In the model presented here, the processing of motion information is performed in two stages that represent areas V1 and MT. A model of complex V1 neurons, which has spatiotemporal filters tuned to eight different directions, computes the initial motion information. The activities of complex V1 neurons generate ambiguous motion information because of their small receptive fields, a phenomenon known as the aperture problem. When the receptive fields of these neurons are along the edges of objects, they respond to the component of the motion that is moving orthogonal to the edge of the stimulus. However, there are some V1 cells, known as end-stopped V1 neurons that exclusively respond to the terminators (such as the corners of bars) and so represent

accurate directional information (Hubel & Wiesel, 1965). At the next stage, the model of MT neurons receives initial information from end-stopped and standard complex V1 neurons (Eskikand, Kameneva, Ibbotson, Burkitt, & Grayden, 2016).

MT neurons are categorized into two different types depending on their center-surround interaction: integration and segmentation neurons (Lidén & Pack, 1999). Fig. 2 is a schematic diagram of the model showing the different types of neurons and the excitatory and inhibitory interactions between them. The end-stopped neurons respond with high levels of activity at the corners of the stimulus regardless of the type of terminators (i.e., extrinsic or intrinsic). Therefore, for the neurons in area MT to be able to discriminate the extrinsic terminators from the intrinsic terminators, they must have access to additional information about the form of the stimulus. In this study, form information is represented by a set of non-direction-selective, center-surround neurons in V1 (described in Section 2.1). The threshold level for the MT neurons is set by the activity of the center-surround neurons. The activity of the MT neurons is highly dependent on the contrast of the stimulus. Therefore, the defined threshold on the activities of MT neurons needs to adapt to changes in the contrast level of the stimulus. The contrast information is provided by the activity of the center-surround V1 neurons.

### 2.1. Center-surround neurons

The center-surround neurons have a high level of activity when there is a change in the luminance of the input image in the receptive fields of the neurons, but they are not selective to the direction of movement of the stimulus over time (Cavanaugh, Bair, & Movshon, 2002b). The activity of these neurons is influenced by stimulation of the receptive field surround (blue, negative regions in Fig. 3). This antagonistic interaction with the surround results in the suppression of activity of these V1 neurons in regions where there is coherency in the stimulus between center and surround. The suppressive surround of these neurons is a key element in the differentiation of the intrinsic from the extrinsic terminators. An effect of this configuration is that center-surround V1 neurons are not inhibited substantially at the corner of a single object (Fig. 3A) but are inhibited strongly at a corner that is formed when two objects overlap (Fig. 3B).

To model the suppressive effect of the surround, the receptive fields of the center–surround V1 neurons are modeled as a difference of Gaussians applied to the input stimulus (Rodieck, 1965),

$$R_{x,y} = \frac{A_C \exp\left(\frac{-(x^2+y^2)}{2\sigma_{x,c}^2+2\sigma_{y,c}^2}\right)}{\sqrt{2\pi\sigma_{x,c}^2+2\pi\sigma_{y,c}^2}} - \frac{A_S \exp\left(\frac{-(x^2+y^2)}{2\sigma_{x,s}^2+2\sigma_{y,s}^2}\right)}{\sqrt{2\pi\sigma_{x,s}^2+2\pi\sigma_{y,s}^2}}, \quad (1)$$

where  $\sigma_{x,s}$  is the size of the surround,  $\sigma_{x,c}$  is the size of the center, and  $A_S$  is the gain of the surround, which is less than the gain of the center,  $A_C$ .

The activities of the form neurons,  $v_{x,y}^f$ , are computed by convolving the spatial filter (1) with the intensity level of the input stimulus,  $I_{x,y}$ , according to

$$v_{x,y}^f = R_{x,y} * I_{x,y}, \quad (2)$$

where  $*$  denotes two-dimensional convolution. It is assumed that the polarity of the contrast is positive for the black stimulus with the white background in all of the following simulations. The activities of these center–surround V1 neurons are transmitted to integration MT neurons as excitatory inputs to assist MT neurons to emphasize the activity of the intrinsic terminators compared to any extrinsic terminators. The center–surround V1 neurons respond with a higher level of activity at the intrinsic terminators compared to the extrinsic terminators. The transmission of the activity of these V1 neurons as an excitatory input assists integration MT neurons to resolve the X-junction illusion.

## 2.2. Complex V1 neurons

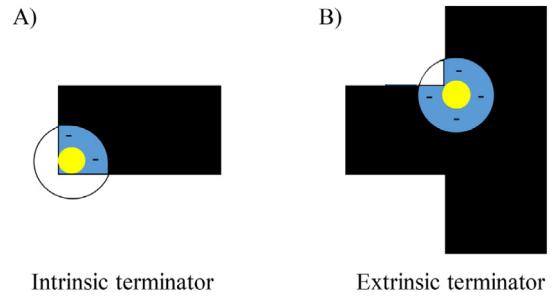
The activities of complex V1 neurons are modeled by motion energy filters that are composed of a combination of spatiotemporal filters (Adelson & Bergen, 1985; Eskikand et al., 2016; Van Santen & Sperling, 1985). The spatial filters are Gabor functions and the temporal filter is a multi-stage low-pass filter, which is described below as

$$g_n(t) = (t/\tau_g)^n \exp(-t/\tau_g) \left[ \frac{1}{n!} - \frac{(t/\tau_g)^2}{(n+2)!} \right]. \quad (3)$$

$\tau_g$  is the time constant of the filter and  $n$  takes different values of 6 and 9 to simulate the delay between two different temporal filters. The output of the spatiotemporal filters are combined as

$$\begin{aligned} r_{x,y,\theta}(t) &= [-I_{x,y}(t) * \sin(2\pi f x_\theta) \exp\left(\frac{-x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right) * g_6(t) \\ &+ (I_{x,y}(t) * \cos(2\pi f x_\theta) \exp\left(\frac{-x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right) * g_9(t)]^2 \\ &+ [I_{x,y}(t) * \sin(2\pi f x_\theta) \exp\left(\frac{-x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right) * g_9(t) \\ &+ (I_{x,y}(t) * \cos(2\pi f x_\theta) \exp\left(\frac{-x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right) * g_6(t)]^2, \end{aligned} \quad (4)$$

$$\begin{aligned} l_{x,y,\theta}(t) &= [I_{x,y}(t) * \sin(2\pi f x_\theta) \exp\left(\frac{-x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right) * g_6(t) \\ &+ (I_{x,y}(t) * \cos(2\pi f x_\theta) \exp\left(\frac{-x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right) * g_9(t)]^2 \\ &+ [-I_{x,y}(t) * \sin(2\pi f x_\theta) \exp\left(\frac{-x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right) * g_9(t) \\ &+ (I_{x,y}(t) * \cos(2\pi f x_\theta) \exp\left(\frac{-x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right) * g_6(t)]^2, \end{aligned} \quad (5)$$



**Fig. 3.** The activities of center–surround V1 neurons to luminance boundaries in the input stimulus. The suppressive surround of the receptive fields of the neurons results in more inhibition at the extrinsic terminators (A) compared to the intrinsic terminators (B). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where  $I_{x,y}(t)$  is the input stimulus at the spatial location  $(x, y)$ ,  $f$  is the spatial frequency of the sinusoidal carrier,  $(x_\theta, y_\theta)$  represents the oriented coordinate in direction  $\theta$ , and  $\sigma_x$  and  $\sigma_y$  are standard deviations. The symbol  $*$  represents convolution. The output of complex V1 neurons is computed by applying the resulting motion-energy filters on the input stimulus.

End-stopped V1 neurons, which respond only to the motion of stimulus terminators, are modeled with additional inhibitory interconnections between neighboring V1 neurons (Eskikand et al., 2016; Pack et al., 2003; Tsui, Hunter, Born, & Pack, 2010). The inhibitory connection is effective when the activity levels of the neighboring V1 neurons are above a threshold level. To improve the robustness of the model to changes in the contrast of the stimuli, the level of the threshold is defined to be a function of the contrast of the input stimulus. The threshold,  $\rho_{x,y}^{cx}$ , is computed by applying a Gaussian filter to the activity of center–surround V1 neurons and normalizing the output,

$$\rho_{x,y}^{cx} = \frac{c^{cx} v_{x,y}^f * \exp(-\sqrt{(x^2+y^2)})}{\sqrt{\sum_x \sum_y (v_{x,y}^f * \exp(-\sqrt{x^2+y^2}))^2}}, \quad (6)$$

where  $c^{cx}$  is a constant. The temporal dynamics of the end-stopped neurons is

$$\frac{d}{dt} v_{x,y,\theta}^{es}(t) = (1 - v_{x,y,\theta}^{es}(t)) (G_{es}^{cx1} v_{x,y,\theta}^{cx}(t)) - v_{x,y,\theta}^{es} (\tau_{es} + G_{es}^{cx2} \Gamma_{x,y,\theta}(t)), \quad (7)$$

where  $v_{x,y,\theta}^{es}$  is the activity of an end-stopped cell selective to direction  $\theta$  located at the coordinate  $(x, y)$ ,  $v_{x,y,\theta}^{cx}$  is the activity of the complex neuron in the same location and direction,  $\tau_{es}$  is a decay rate, and  $G_{es}^{cx1}$  and  $G_{es}^{cx2}$  are constant gains.  $\Gamma_{x,y,\theta}$  is the inhibition that the neuron receives from complex neurons when the activity of the neighboring standard complex neurons is above the threshold,  $\rho_{x,y}^{cx}$ ,

$$\Gamma_{x,y,\theta} = \begin{cases} \sum_{i=-8}^8 \sum_{j=-8}^8 \mu_{i,j} v_{x+i,y+j,\theta}^{cx} & v_{x+i,y+j,\theta}^{cx} |_{i=3,j=3}^{i=-3,j=-3} > \rho_{x,y}^{cx} \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

where  $\mu_{i,j}$  is the inhibitory connectivity matrix which extends across a patch of 8 by 8 neighboring neurons and has a discretized Gaussian shape (Eskikand et al., 2016).

## 2.3. MT neurons

As indicated in the schematic diagram of the model (Fig. 2), MT neurons are divided into two types: integration and segmentation neurons. Integration neurons propagate motion information

through excitatory interconnections from the terminators of the stimulus along the bar. Segmentation neurons respond to discontinuities in the stimulus and suppress the activity of integration neurons via inhibitory connections (Eskikand et al., 2016). The additional effect of form information provided by center-surround V1 neurons is inserted in the model as input to the integration MT neurons. Interactions between integration neurons have three components: (1) excitation from neighboring neurons selective to the same direction, (2) inter-directional inhibition between neurons at the same spatial location, and (3) long-range inhibition from distant neighboring neurons. The behavior of the integration neurons,  $v_{x,y,\theta}^{ig}$ , is determined by

$$\begin{aligned} \frac{dv_{x,y,\theta}^{ig}(t)}{dt} = & h(G_{ig}^{cx} v_{x,y,\theta}^{cx}(t) + G_{ig}^{es} v_{x,y,\theta}^{es}(t) \\ & + G_{ig}^{ig2} \lambda_{x,y,\theta}(t) + G_{ig}^f \kappa_{x,y,\theta}(t) - \\ & G_{ig}^{ig1} \gamma_{x,y,\theta}(t - T^{ig}) - G_{ig}^{ig3} \zeta_{x,y,\theta}(t - T^{ig}) \\ & - G_{ig}^{sg} v_{x,y,\theta}^{sg}(t) - \tau_{ig} v_{x,y,\theta}^{ig}(t)), \end{aligned} \quad (9)$$

where  $\lambda_{x,y,\theta}$  is the excitation from neighboring integration neurons selective to the same direction  $\theta$ ,  $\kappa_{x,y}$  is excitatory input from center-surround V1 neurons.  $\gamma_{x,y,\theta}$  is inter-directional inhibition,  $\zeta_{x,y,\theta}$  is long-range inhibition, and  $\tau_{ig}$  defines the value of the decay rate. These neurons receive excitatory input from V1 complex and end-stopped neurons,  $v_{x,y,\theta}^{cx}$  and  $v_{x,y,\theta}^{es}$ , respectively. They also receive inhibitory input from segmentation neurons,  $v_{x,y,\theta}^{sg}$ . Finally,  $h()$  is a piece-wise linear saturation function that keeps the level of activity within a specified range (between 0 and 1),

$$h(x) = \begin{cases} 1, & \text{if } x \geq 1 \\ x, & \text{if } 0 \leq x < 1 \\ 0, & \text{if } x < 0. \end{cases} \quad (10)$$

The effectiveness of the excitatory and inhibitory interconnections are determined by constant gain factors,  $G_{ig}^{cx}$ ,  $G_{ig}^{es}$ ,  $G_{ig}^{ig2}$ ,  $G_{ig}^f$ ,  $G_{ig}^{ig1}$ ,  $G_{ig}^{ig3}$ ,  $G_{ig}^{sg}$ ,  $G_{ig}^f$ ; their values are shown in Table 1. The model is not very sensitive to changes in the values of these parameters. The values of the parameters for the core model were optimized using a modified version of the genetic algorithm described by Eskikand et al. (2016). The functional role of the excitation from neighboring integration neurons is to propagate unambiguous signals from the terminators along the object. To overcome the aperture problem, neurons receive excitation only from neighboring integration neurons with higher levels of activity; i.e.,  $v_{x+i,y+j,\theta}^{ig} - v_{x,y,\theta}^{ig} > \rho_{x,y}^{ig}$ , where  $\rho_{x,y}^{ig}$  is the threshold level set by the difference between the activity level of neighboring integration neurons. The level of the threshold depends on the contrast level of the input stimulus at location  $(x, y)$ , which is computed in a similar way to Eq. (6), but with a different constant value,  $c^{ig}$ . This interaction between MT neurons is modeled as

$$\lambda_{x,y,\theta} = \begin{cases} \sum_{i=-l}^l \sum_{j=-l}^l v_{x+i,y+j,\theta}^{ig} H(\rho_{x,y}^{sg} - v_{x,y,\theta}^{sg}), & \text{if } v_{x+i,y+j,\theta}^{ig} - v_{x,y,\theta}^{ig} \\ & > \rho_{x,y}^{ig} \\ 0, & \text{otherwise,} \end{cases} \quad (11)$$

where  $(x + i, y + j)$  is a location index,  $l$  indicates the extent of the excitatory connections, and  $H()$  is the Heaviside step function.

The excitatory connections between integration neurons are gated by the level of the activity of segmentation neurons. Therefore, it will be effective only when the activity levels of the segmentation neurons are below the threshold,  $\rho_{x,y}^{sg}$ .  $\rho_{x,y}^{sg}$  depends on the contrast level of the stimulus at location  $(x, y)$ , which is

computed in a similar way to Eq. (6), but with the constant value of  $c^{sg}$ .

Segmentation MT neurons limit the activity of integration MT neurons at the discontinuities of the input stimulus, where they are activated by effectively shunting excitation from neighboring integration neurons.

$$\frac{dv_{x,y,\theta}^{sg}}{dt} = h(G_{sg}^{cx} v_{x,y,\theta}^{cx} - G_{sg}^{es} v_{x,y,\theta}^{es} + G_{sg}^{ig} \eta_{x,y} + G_{sg}^{sg} \chi_{x,y,\theta}^e - G_{sg}^{sg} \chi_{x,y,\theta}^i - \tau_{sg} v_{x,y,\theta}^{sg}), \quad (12)$$

where  $\eta_{x,y}$  is the excitatory input received from integration MT neurons, which is defined by

$$\eta_{x,y} = \sum_{\theta=1}^8 v_{x,y,\theta}^{ig}. \quad (13)$$

$\chi_{x,y,\theta}$  in Eq. (12) represents the interconnection between segmentation neurons as the result of center-surround interactions described by

$$\chi_{x,y,\theta} = \begin{cases} \sum_{i \in \Pi} \sum_{j \in \Pi} v_{x+i,y+j,\theta}^{sg}, & \text{if } v_{x,y,\theta}^{sg} > \rho_{sg} \text{ and } v_{x+i,y+j,\theta}^{sg} > \rho_{sg}, \\ 0, & \text{otherwise,} \end{cases} \quad (14)$$

where  $\Pi$  indicates the surround of an MT neuron and  $\rho_{sg}$  is the threshold level on the activity of the segmentation MT neurons. All of the remaining parameters are constant and their values are given in Table 1.

The activities of center-surround V1 neurons are transmitted to integration MT neurons as excitatory inputs. This input increases the level of activity at the intrinsic terminators relative to the extrinsic terminators. The level of excitatory input is gated by the activity of end-stopped neurons to prevent suppression of its effect by the inter-directional inhibitory connections between MT neurons,

$$\kappa_{x,y,\theta} = \begin{cases} v_{x,y}^f & \text{if } v_{x,y,\theta}^{es} > \rho_{x,y}^{es} \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

where  $\rho_{x,y}^{es}$  is the threshold of the activity of the end-stopped neurons, which depends on the contrast level of the stimulus at location  $(x, y)$ . Similar to the model described in Eskikand et al. (2016), an inhibitory connection,  $\gamma_{x,y,\theta}$ , is defined between neurons selective to different directions at the same spatial location,

$$\gamma_{x,y,\theta} = \sum_{\phi \neq \theta} v_{x,y,\phi}^{ig}, \quad (16)$$

where  $\phi$  is the preferred direction of other neurons at the same location. Other sources of inhibition,  $\zeta_{x,y,\theta}$ , come from distant neighboring neurons selective to other directions,

$$\zeta_{x,y,\theta} = \sum_{\phi \neq \theta} \sum_{i \in \Phi} \sum_{j \in \Phi} v_{x+i,y+j,\phi}^{ig}, \quad (17)$$

where  $\Phi$  is the set of locations in the effective area of long-range inhibition. The integration neurons receive inhibitory connections from neurons selective to all directions,  $\phi$ , that are three positions away from reference neurons, except for the neurons with the same directional preference,  $\theta$ . The inhibitory connections between integration neurons have a longer time delay,  $T^{ig}$ , compared to the excitatory connections.

### 3. Results

The responses of the simulated neurons at different stages of the model are examined using a stimulus consisting of two bars moving in opposite directions. A bar with 45° orientation is moving

**Table 1**

The constant parameters used in the model, their values, and their units.

Description	Parameter	Value	Unit
Connection strength of input to the end-stopped neurons	$G_{es}^{cx1}$	2	–
Connection strength of inhibitory connections on end-stopped neurons	$G_{es}^{cx2}$	3	–
Connection strength of complex V1 neurons to integration neurons	$G_{ig}^{cx}$	0.7	–
Connection strength of end-stopped V1 neurons to integration neurons	$G_{es}^{ig}$	1	–
Connection strength of center-surround V1 neurons to integration neurons	$G_{ig}^f$	0.5	–
Connection strength of complex V1 neurons to segmentation neurons	$G_{sg}^{cx}$	2	–
Connection strength of end-stopped V1 neurons to segmentation neurons	$G_{sg}^{es}$	1	–
Connection strength of excitatory connections to integration neurons	$G_{ig}^{ig2}$	1	–
Connection strength of inter-directional inhibitory connections	$G_{ig}^{ig1}$	1	–
Connection strength of long-range inhibitory connections	$G_{ig}^{ig3}$	1	–
Connection strength of inhibition from segmentation neurons	$G_{ig}^{sg}$	0.1	–
Connection strength of excitation from integration on segmentation neurons	$G_{sg}^{ig1}$	0.5	–
Connection strength of surround suppression on segmentation neurons	$G_{sg}^{sg}$	0.2	–
Number of neurons at each location-selective to different directions	$N$	8	–
Constant value for the threshold on the activity of complex V1 neurons	$c^{cx}$	0.13	–
Constant value for the threshold on the activity of segmentation neurons	$c^{sg}$	0.01	–
Constant value for the threshold on the activity of integration neurons	$c^{ig}$	0.05	–
Constant value for the threshold on the activity of end-stopped neurons	$c^{es}$	0.3	–
Decay rate of the activity of integration neurons	$\tau_{ig}$	0.2	–
Decay rate of the activity of segmentation neurons	$\tau_{sg}$	0.2	–
Decay rate of the activity end-stopped neurons	$\tau_{es}$	0.01	–
Simulation time step	$\Delta t$	0.01	ms
Time constant of the temporal filter	$\tau_g$	0.01	ms
Time delay of inhibition between MT integration neurons	$T_{ig}$	0.4	ms
Spatial frequency	$f$	1.1	cyc/deg
Standard deviation of horizontal spatial Gaussian filter	$\sigma_x$	0.5	–
Standard deviation of vertical spatial Gaussian filter	$\sigma_y$	0.5	–
Standard deviation of center portion of horizontal spatial Gaussian filter	$\sigma_{x,c}$	0.15	–
Standard deviation of center portion of vertical spatial Gaussian filter	$\sigma_{y,c}$	0.15	–
Standard deviation of surround portion of horizontal spatial Gaussian filter	$\sigma_{x,s}$	0.5	–
Standard deviation of surround portion of vertical spatial Gaussian filter	$\sigma_{y,s}$	0.5	–
Strength of center portion of spatial Gaussian filter	$A_c$	1	–
Strength of surround portion of spatial Gaussian filter	$A_s$	0.9	–

to the left and another bar with 135° orientation is moving to the right. In the first section of the results, the bars have the same level of contrast. The activity of the standard complex V1 neurons selective to different directions, at a time when the bars overlap, is illustrated in Fig. 4. Although the activities of the neurons at the intrinsic terminators represent accurate directions of motion, their activity levels are lower than the neurons with ambiguous motion information because of the aperture problem and the neurons at the extrinsic terminators of the stimulus, which express an upward direction of movement.

Fig. 5 shows the activity of the end-stopped neurons selective to different directions in response to these stimuli. End-stopped neurons have a high level of activity at the terminators of the stimulus. However, they are not capable of differentiating and suppressing the activity at the extrinsic terminators of the input stimulus.

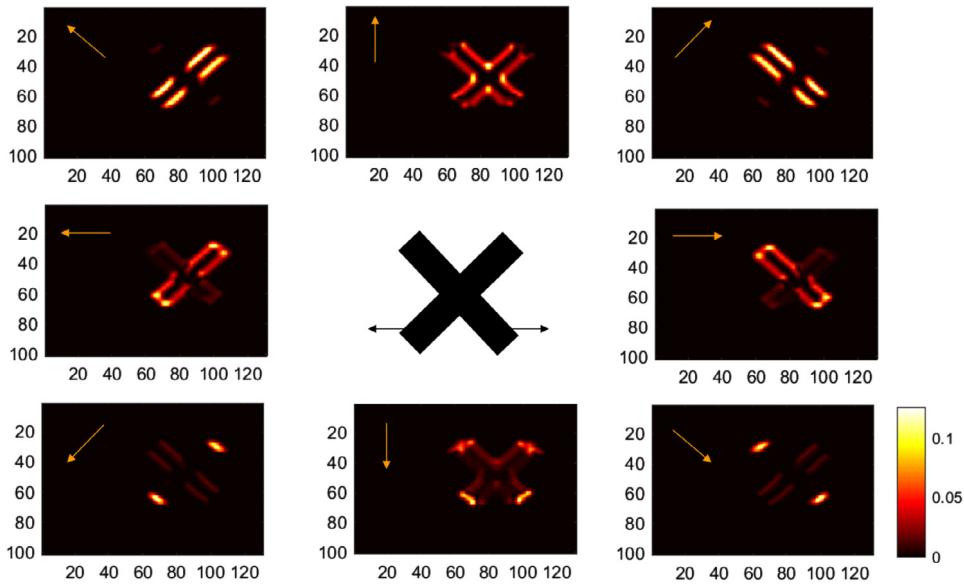
Fig. 6 shows the activity of the center-surround V1 neurons in response to the same stimuli. These V1 neurons have higher activity at the intrinsic terminators compared to the extrinsic terminators as they receive a strong level of inhibition from their surround at the extrinsic terminator locations.

The activities of the integration MT neurons, shown in Fig. 7, demonstrate their ability to estimate the correct directions of the global motion directions of both bars. The neurons selective to rightward motion have a high level of activity in response to the rightward moving bar and those selective to the left direction have a high level of activity to the leftward moving bar. At the regions where the bars overlap, the inter-directional inhibition suppresses the activity of the neurons.

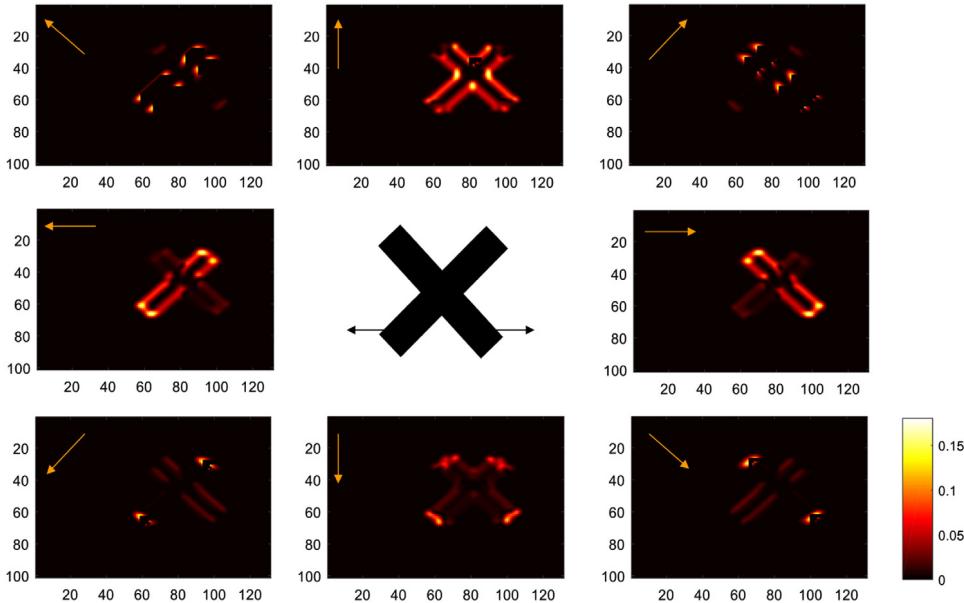
To investigate the robustness of the model to changes in contrast, the responses of the neurons to stimuli that consisted of two crossing bars with different levels of contrast were evaluated. The activity level of these neurons depends on the luminance of the stimulus. The strength of the activity decreases for the low contrast stimulus compared to the stimulus with a high level of contrast. The neurons have a high level of activity at the edges compared to the intrinsic terminators of each stimulus because of the aperture problem. Their activities are also high at the crossing points where extrinsic terminators are located compared to the activities of neurons at the intrinsic terminators of the bar with low contrast.

The center-surround V1 neurons respond with a lower level of activity in response to the low contrast stimulus. However, they have the highest level of activity at the intrinsic terminators, as the surrounds of the neurons at the intrinsic terminators are not stimulated and they do not have a suppression effect on these neurons. The neurons at the extrinsic terminators are not active as they are inhibited by their suppressive surrounds.

Fig. 8 shows the activity of the integration MT neurons. Integration MT neurons are able to estimate the directions of the actual movements of the bars. The neurons selective to the rightward direction have a high level of activity at the border of the bar that is moving rightward, and the activity of the neurons selective to the left direction is high at the border of the bar that is moving to the left. At the crossing points of the bars, the illusory depth results in one of the bars appearing to move in front of the other. Therefore, the activities of the integration MT neurons at the intersection are more continuous along the borders of the high contrast bar.



**Fig. 4.** The activities of complex V1 neurons in response to the stimuli, which are two crossing bars with the same contrast moving in opposite directions. Each graph shows the activities of all V1 neurons in the model that are selective to the same direction (as indicated by the arrow in each box). The axes represent the location of each V1 neuron in the array and the colors represent the strengths of activities generated by the stimulus, as indicated by the color bar. The stimulus is shown in the center of the figure, which has one bar with 45° orientation is moving to the left and another bar with 135° orientation moving to the right. Among the neurons selective to the right, the end-points of the 45° oriented bar have the highest level of activity and, similarly among the neurons selective to the left direction, the endpoints of the 135° orientation represent the accurate direction of motion. The activity level at the extrinsic terminators (formed at the X-junction) has the highest value for neurons selective to the upward direction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



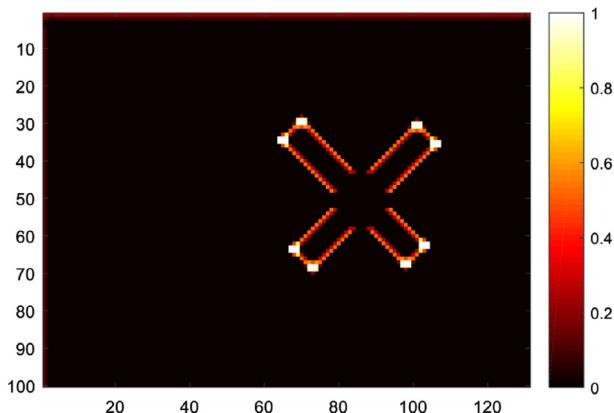
**Fig. 5.** The activities of the end-stopped neurons selective to different directions in response to the crossing bar stimulus (representation as for Fig. 4). Each graph shows the activity of V1 neurons selective to the same direction. In each box, the arrow indicates the preferred direction of the array of neurons. The end-stopped neurons have a high level of activity at the intrinsic and extrinsic terminators. The inhibitory interconnections between end-stopped neurons have suppressed the ambiguous activity of the neurons along the edges caused by the aperture problem.

The responses of MT neurons to the stimuli when the low contrast bar was in front were also examined. The purpose was to confirm that the dominance of the activity of the neurons selective to the movement of the high contrast bar is not because of the stronger activity of the neurons to the high contrast bar. Fig. 9 illustrates the sum of activities of complex V1 neurons over different directions in response to overlapped crossing bars with different contrast levels. It shows the response to the stimulus with different levels of contrast when the black bar covers the gray bar

(Fig. 9A) compared to the case when the lighter bar covers the black bar (Fig. 9B).

#### 4. Discussion

The model presented here uses form information to achieve an accurate measure of the direction of motion of an object. This is performed by compensating for the effect of extrinsic terminators. In the model, interconnections between MT neurons result in



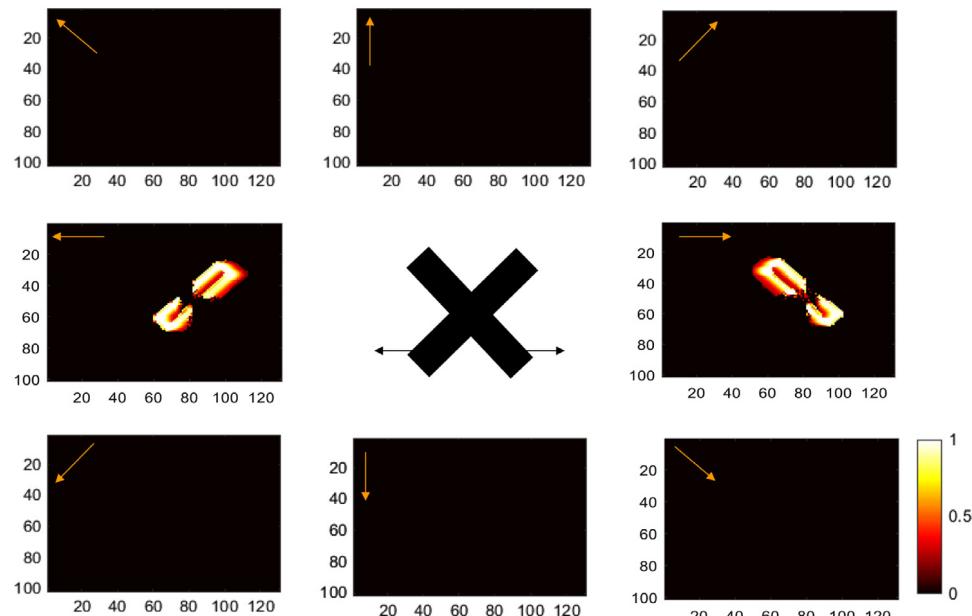
**Fig. 6.** The activities of center–surround V1 neurons in response to the crossing bar stimulus. These neurons have a high level of activity at the intrinsic terminators compared to the extrinsic terminators, which are suppressed by their inhibitory surround. The axes represent the location neurons in the array and the color bar represents the level of activity of center–surround V1 neurons. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the propagation of unambiguous motion information from the intrinsic terminators to other regions of the stimulus. This process solves the aperture problem, thus allowing an estimation of the correct direction of motion. In the case of two moving bars that overlap, the extrinsic terminators formed at the crossing junctions between bars represent a different, conflicting estimation of the global motion direction. We have shown that the excitatory input provided by the activities of center–surround V1 neurons may assist MT neurons in discriminating the extrinsic from the intrinsic terminators. The inhibitory effect of the surround results in a suppression of the activity of these V1 neurons at the extrinsic terminators, where the surround of the neurons is more strongly stimulated compared to the intrinsic terminators. At the extrinsic terminators, the end-stopped V1 neurons also have a high level of activity, but the center–surround V1 neurons are inactive at these

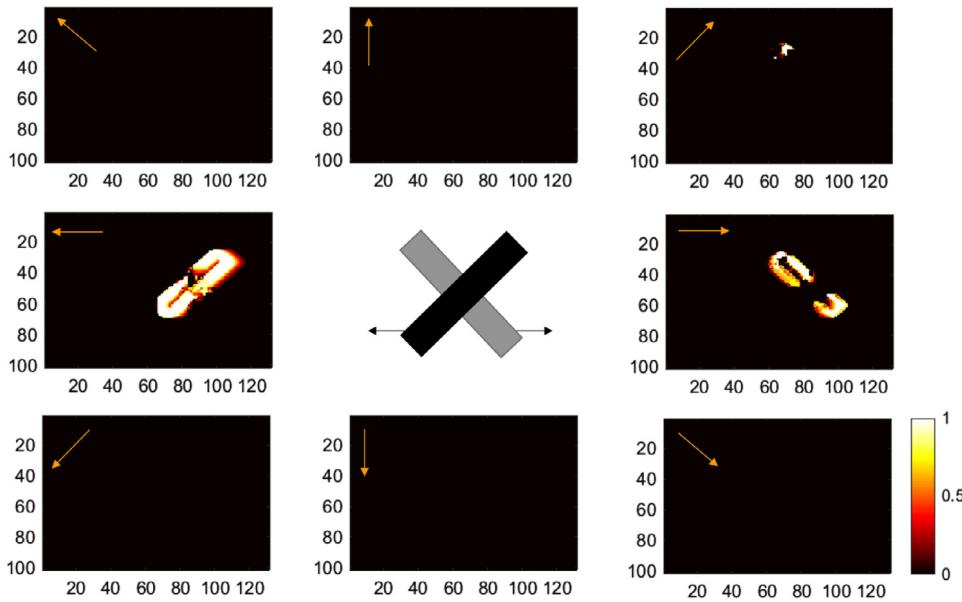
regions. Therefore, the high values of these excitatory inputs at the intrinsic terminators of the stimuli assist MT neurons to overcome the ambiguous information resulting from both the X-junction and the aperture problems.

The results of our model demonstrate the successful estimation of the global motion direction for individual components of the overlapping stimuli. However, without form-related information, the excitatory connections between integration neurons propagate the motion signals at the extrinsic terminators over the whole object in a similar way to the activity of the neurons at the intrinsic terminators. In the case of overlapped crossing bars with the same contrast, the inter-directional inhibition between MT neurons strongly suppresses the activity of the neurons at the crossing junctions where motion information exists for multiple directions at the same location. Therefore, MT neurons have a very low level of activity at the crossing junctions when they are stimulated by overlapping moving bars with the same contrast. The model was previously shown to be robust to the injection of large amounts of neuronal noise (Eskikand et al., 2016). In that set of simulations, stochastic noise with a Gaussian distribution was added to the input to the model. The results showed that the model was able to estimate the correct direction of motion with high levels of noise, although errors in motion detection did grow with increases in the noise level.

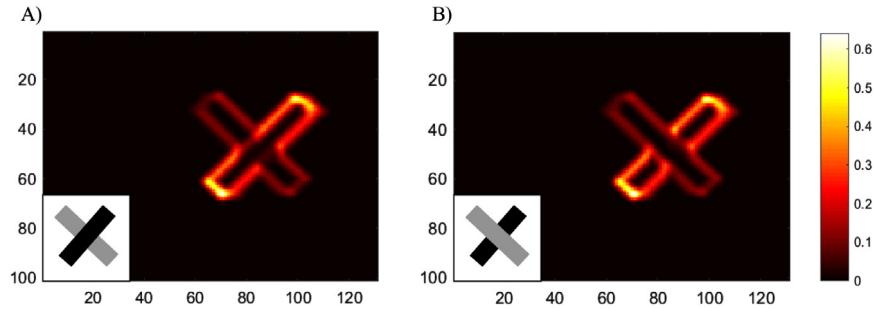
The results of our model are consistent with electrophysiological evidence (Snowden, Treue, Erickson, & Andersen, 1991), which demonstrates the existence of inter-directional inhibition between MT neurons, resulting in the suppression of the activity of the neurons when stimulated with simultaneous motions in several directions (Snowden et al., 1991). Snowden et al. (1991) showed that V1 neurons are able to represent signals about opposing directions at the same time and spatial location. However, MT neurons average and smooth the opposing motion information and their activities are suppressed in response to transparent stimuli. These findings support the existence of division-like inhibition between neurons that are activated by movements in opposite directions; i.e., neurons in which the level of inhibition increases with increases in the response amplitude of the neuron (Snowden et al., 1991). Experimental data from Mikami, Newsome, and Wurtz



**Fig. 7.** The activities of integration MT neurons selective to different directions in response to the stimuli with the same level of contrast (representation as for Fig. 4). Each graph shows the activity of V1 neurons selective to the same direction. The arrow in each box indicates the preferred direction of that array of neurons. The neurons selective to the right and left direction represent the directions of motion of the moving bars. The activities of the neurons selective to other directions are suppressed.



**Fig. 8.** The activity of the integration MT neurons for the stimulus shown in the middle in which the black bar moves to the left (representation as for Fig. 4). The neurons selective for rightward motion have a high level of activity in response to the gray bar, which is moving to left. The neurons selective to the left direction reveal a high level of activity in response to the leftward moving bar. The estimation of the edges of the black bar are more coherent compared to the other bar.



**Fig. 9.** The sum of the activities of standard complex V1 neurons. The axes represent the location and the colors represent the levels of activity of the neurons. (A) The activities of V1 neurons in response to two crossing bars with different contrasts when the high contrast bar covers the low contrast bar. The stimulus is shown in the bottom-left corner of the figure. There is a coherency in the border of the high contrast bar at the crossing junction. This shows the capability of standard V1 neurons in determining the boundary ownership of the stimuli. (B) The activity of standard V1 neurons in response to the case when the low contrast bar covers the black bar. There is a coherency in the border of the low contrast bar. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(1986) on MT neurons also suggest the existence of inhibitory interactions between these neurons (Mikami et al., 1986).

When overlapped moving stimuli have different levels of contrast, illusory depth is perceived by human observers, i.e., it appears that one of the stimuli is moving in front of the other. In our model, contrast-selective center-surround V1 neurons are able to determine the boundary ownership of the stimulus, which should assist in determining which stimulus is moving in front of the other. Therefore, the excitatory input transmitted by them to MT neurons results in the dominance of the activity of neurons selective to the direction of the movement of the stimulus that appears to be in front at the location of the crossing junctions. The results show a coherency in the motion detected by MT neurons for the foremost stimulus, demonstrating a high level of accuracy in determining border ownership. The boundary ownership of the overlapping stimuli is successfully determined by the activity of the standard complex V1 neurons. This suggests a role for complex V1 neurons in the processing of motion information. Although the motion information provided by complex V1 neurons is ambiguous because of their small receptive fields, they may play a key role in estimating the illusory depth observed with overlapping stimuli that have different contrast levels.

The effect of contrast on the activity of neurons has been investigated in several studies. Experiments by Cavanaugh et al. (2002a, b) and Sceniak, Ringach, Hawken, and Shapley (1999) showed that the suppressive surround of neurons is highly dependent on the luminance of the input stimulus. Cavanaugh et al. (2002a, b) showed that not all V1 neurons have a suppressive surround and the level of the suppression of the surround increases with contrast level. Sceniak et al. (1999) showed that increasing the contrast level of the stimulus results in a reduction in the size of the excitatory receptive field. The ratio of the inhibitory to excitatory connections decreases with increases in contrast (Sceniak et al., 1999). The neurophysiological evidence by Yazdanbakhsh and Livingstone (2006) shows the dependence of the activity of end-stopped neurons on the contrast of the stimulus (Yazdanbakhsh & Livingstone, 2006). Although the activity of the center-surround V1 neurons in the model is gated by the activity of the end-stopped neurons, supplying form information only to end-stopped neurons instead of the direct excitatory input to MT neurons needs to be considered as an option in future work.

Several modeling studies suggest the necessity of interactions with form information for the coherent perception of form and motion of overlapping stimuli. Berzhanskaya et al. (2007) and Grossberg (1994, 2015) investigated how motion information and

form cues cooperate to determine figure-ground segregation and achieve depth perception in 2D images. In these models, form information is transmitted to the neurons responding to the motion by including a model of bipole neurons in V2. Bipole neurons have two separated receptive fields and are able to respond to illusory contours (Berzhanskaya et al., 2007; Dresp et al., 2002; Grossberg, 1994, 2015). An illusory contour is an edge that appears in a blank space between two aligned edges. There is neurophysiological and anatomical evidence for the existence of bipole neurons that demonstrates their ability to connect parts of the stimulus that appear to belong to the same coherent object (Bosking, Zhang, Schofield, & Fitzpatrick, 1997; Von Der Heydt, Peterhans, & Baumgartner, 1984). Therefore, the main role of form information in the models proposed by Berzhanskaya et al. (2007) and Grossberg (1994, 2014, 2015) is connecting the discontinuities in the stimulus formed by occlusion by other objects so that the stimulus is perceived as a single coherent moving feature. Therefore, bipole neurons in these models do not play a significant role in calculating the accurate direction of motion; their role is to generate a united image of the object in the presence of occluders to determine the perception of depth from a 2D stimulus (Berzhanskaya et al., 2007; Chey et al., 1997; Dresp et al., 2002; Grossberg, 1994, 2015). It is not known whether the global form of the stimulus is perceived in areas MT and MST or this perception is generated in higher levels of the visual cortex.

In a proposed model by Raudies and Neumann (2010), a hierarchical network is presented for motion detection and figure-ground segregation by modeling the responses of neurons in areas V1, V2, MT and MST, along with V4 to generate modulating feedback connections that arise from spatial attention signals. In their model, initial motion information and motion boundaries are determined at different stages of the dorsal pathway. Information regarding the shape of the objects defined by their boundaries is processed through the ventral pathway (e.g. areas V2 and V4). The dorsal and ventral pathways interact with each other through mutual interconnection and by modulating attentional feedback via areas in higher stages of the visual cortex (via V4 neurons). At the final stage, form and motion information are integrated in the lateral intraparietal (LIP) area, which generates a decision based on the available information. The cross-pathway interaction between neurons results in coherent form and motion boundaries (Raudies & Neumann, 2010). In these models, feedback connections from higher stages with larger receptive fields result in evolution of the initial rough estimation of motion in V1, which ultimately leads to the correct estimation of motion (Beck & Neumann, 2010, 2011; Raudies & Neumann, 2010). The form information is transmitted to the dorsal pathway by the activity of bipole neurons in V2, similar to the related models described earlier (Berzhanskaya et al., 2007; Chey et al., 1997; Dresp et al., 2002; Grossberg, 1994, 2014, 2015), and their role is mainly in the guidance of the integration processes of MT neurons. The proposed scheme used for processing motion information in these models is controversial, as there is no evidence to support the evolution of the activity of V1 neurons over time from the ambiguous motion information for the accurate estimations of the direction of whole-object movement. Conversely, there is neurophysiological evidence that demonstrates the temporal evolution of MT neuron activity to overcome ambiguous motion information (Pack & Born, 2001). However, there is no experimental study showing that the responses of V1 neurons to intrinsic terminators are propagated to other neurons in V1.

In the proposed models by Berzhanskaya et al. (2007), Dresp et al. (2002) and Grossberg (1994, 2015), the model of bipole neurons in V2 assists MT neurons to determine the illusory depth of overlapping stimuli. The extensive receptive fields of bipole neurons predict the global form of the stimulus by connecting the illusory contours that form as the result of overlaps with other objects

(Berzhanskaya et al., 2007; Grossberg, 1994, 2014, 2015). However, in our model, the interaction with form information is essential for an accurate estimation of the global motion of stimuli that overlap with other objects, by suppressing the effect of extrinsic terminators. Our model predicts that it is not essential for MT neurons to have a perception of the global form of the stimulus and, according to existing neurophysiological evidence (Snowden et al., 1991), MT neurons suppress opposing motion directions at the same location, which is in contrast to previous modeling work (Berzhanskaya et al., 2007; Dresp et al., 2002; Grossberg, 1994, 2014, 2015).

The interaction of form and motion information in the models proposed by Raudies and Neumann (2010) is applied to achieve the coherent boundaries of form and motion, which is in contrast to our model that applies the form-related information to differentiate the extrinsic from the intrinsic terminators. In their models, the feedback connections from MT with larger receptive fields to V1 neurons disambiguate the initial rough estimation of motion by V1 neurons. Although, experimental evidence exists for the gradual disambiguation of the activity of MT neurons, the temporal dynamic of the activity of the V1 neurons proposed in this model is not in accord with the neurophysiological data, which does not show gradual disambiguation of the complex V1 neurons (Hubel & Wiesel, 1965; Pack & Born, 2001; Pack et al., 2003).

In the model proposed by Lidén and Pack (1999), which has a similar general structure to our model, there is no mechanism embedded to differentiate the unambiguous motion information at the intrinsic or extrinsic terminators from the ambiguous motion signals at the edges of the stimulus (Lidén & Pack, 1999). Therefore, this model fails to simultaneously overcome the aperture problem and respond to X-junctions, and the activities of the neurons at the edges of the stimulus dominate over the activities of all other regions.

Our model suggests the activities of the center-surround neurons have a key role in the estimation of the motion direction of individual components of overlapping stimuli. It shows that the complex V1 neurons provide necessary information for determining the boundary ownership of overlapping, moving stimuli, while the form information supplied by center-surround V1 neurons is also necessary to strengthen the activity of the neurons responding to intrinsic terminators compared to extrinsic terminators.

## Acknowledgments

This research was supported by the Australian Research Council through Discovery Grants [DE120102210, DP140104533] and the ARC Centre of Excellence for Integrative Brain Function [CE140100007]. Parvin Zarei Eskikand acknowledges a postgraduate scholarship from the National Information and Communication Technology Australia (NICTA- <http://www.nicta.com.au/>). NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy (<http://www.communications.gov.au/>) and the Australian Research Council through the ICT Centre of Excellence program.

## References

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(2), 284–299.
- Allman, J., Miezin, F., & McGuinness, E. (1985). Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception*, 14(2), 105–126.
- Anstis, S. (1990). Imperceptible intersections: The chopstick illusion. *AI and the Eye*, 105–117.
- Beck, C., & Neumann, H. (2010). Interactions of motion and form in visual cortex—a neural model. *Journal de Physiologie (Paris)*, 104(1), 61–70.

- Beck, C., & Neumann, H. (2011). Combining feature selection and integration—a neural model for MT motion selectivity. *PLoS One*, 6(7), e21254.
- Berzhanskaya, J., Grossberg, S., & Mingolla, E. (2007). Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision*, 20(4), 337–395.
- Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 17(6), 2112–2127.
- Brewster, D. (1844). A notice explaining the cause of an optical phenomenon observed by the rev W Selwyn. Report of the Fourteenth Meeting of the British Association for the Advancement of Science, (Pt. 2).
- Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002a). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of Neurophysiology*, 88(5), 2530–2546. <http://dx.doi.org/10.1152/jn.00692.2001>.
- Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002b). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *Journal of Neurophysiology*, 88(5), 2547–2556. <http://dx.doi.org/10.1152/jn.00693.2001>.
- Chey, J., Grossberg, S., & Mingolla, E. (1997). Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. *Journal of the Optical Society of America A*, 14(10), 2570–2594.
- Dresp, B., Durand, S., & Grossberg, S. (2002). Depth perception from pairs of overlapping cues in pictorial displays. *Spatial Vision*, 15(3), 255–276.
- Eskikand, P. Z., Kameneva, T., Ibbotson, M. R., Burkitt, A. N., & Grayden, D. B. (2016). A possible role for end-stopped V1 neurons in the perception of motion: A computational model. *PLoS One*, 11(10), e0164813.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, 15(1), 20–25.
- Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception & Psychophysics*, 55(1), 48–121.
- Grossberg, S. (2014). How visual illusions illuminate complementary brain processes: Illusory depth from brightness and apparent motion of illusory contours. *Frontiers in Human Neuroscience*, 8, 854. <http://dx.doi.org/10.3389/fnhum.2014.00854>.
- Grossberg, S. (2015). Cortical dynamics of figure-ground separation in response to 2D pictures and 3D scenes: How V2 combines border ownership, stereoscopic cues, and gestalt grouping rules. *Frontiers in Psychology*, 6, 270–293.
- Grossberg, S., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*, 41(19), 2521–2553.
- Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28, 229–289.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1), 215–243.
- Ibbotson, M. R. (2007). Reshaping the binding problem of form and motion vision. *The Journal of Physiology*, 585(2), 319–319.
- Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, 12(1), 48–55.
- Krekelerberg, B., Dannenberg, S., Hoffmann, K., Bremner, F., & Ross, J. (2003). Neural correlates of implied motion. *Nature*, 424(6949), 674–677.
- Landau, B., & Jackendoff, R. (1993). Whence and whither in spatial language and spatial cognition?. *Behavioral and Brain Sciences*, 16(02), 255–265.
- Lidén, L., & Pack, C. (1999). The role of terminators and occlusion cues in motion integration and segmentation: A neural network model. *Vision Research*, 39(19), 3301–3320.
- Mikami, A., Newsome, W. T., & Wurtz, R. H. (1986). Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *Journal of Neurophysiology*, 55(6), 1308–1327.
- Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, 409(6823), 1040–1042.
- Pack, C. C., Livingstone, M. S., Duffy, K. R., & Born, R. T. (2003). End-stopping and the aperture problem: Two-dimensional motion signals in macaque V1. *Neuron*, 39(4), 671–680.
- Raudies, F., & Neumann, H. (2010). A neural model of the temporal dynamics of figure-ground segregation in motion perception. *Neural Networks*, 23(2), 160–176.
- Rodieck, R. W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, 5(12), 583–601.
- Sceniak, M. P., Ringach, D. L., Hawken, M. J., & Shapley, R. (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neuroscience*, 2(8), 733–739.
- Schiller, P. H., & Carvey, C. E. (2005). The hermann grid illusion revisited. *Perception*, 34(11), 1375–1397.
- Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 11(9), 2768–2785.
- Spillmann, L. (1994). The hermann grid illusion: A tool for studying human perspective field organization. *Perception*, 23(6), 691–708.
- Tlapale, É., Masson, G., & Kornprobst, P. (2007). Research Report No. Research Report, RR-6293. INRIA.
- Tsui, J. M., Hunter, J. N., Born, R. T., & Pack, C. C. (2010). The role of V1 surround suppression in MT motion integration. *Journal of Neurophysiology*, 103(6), 3123–3138. <http://dx.doi.org/10.1152/jn.00654.2009>.
- Ungerleider, L. G. (1982). Two cortical visual systems. *Analysis of Visual Behavior*, 549–586.
- Van Santen, J. P., & Sperling, G. (1985). Elaborated reichardt detectors. *Journal of the Optical Society of America A*, 2(2), 300–321.
- Von Der Heydt, R., Peterhans, E., & Baurgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, 224, 1260–1262.
- Wolfe, J. M. (1984). Global factors in the hermann grid illusion. *Perception*, 13(1), 33–40.
- Yazdanbakhsh, A., & Livingstone, M. (2006). Contrast-sign selectivity of end-stopping and length-summation. *Journal of Vision*, 6(6), 687–687.