# On the Ensemble of Collision Perception Neuron Models Towards Ultra-Selectivity

Jiatao Li[1], Xuelong Sun[1], Haiyang Li[1], Jigen Peng*,[1], Qinbing Fu*,[1]
[1] Machine Life and Intelligence Research Center, School of Mathematics and Information Science, Guangzhou University, China

***Abstract*—This paper investigates the coordination of multiple looming sensitive neuronal models to get closer to the ultra-selectivity represented by organisms, i.e., responding to merely approaching objects. In locust's brain, a group of lobula giant movement detectors (LGMDs) cooperate very well in collision perception. Although the existing single neuron computation can portray individual LGMDs relatively well, it seems to be insufficient to handle all complex situations as stimuli like receding and translating still greatly stimulate the models. Compounding is a method of biological problem-solving, therefore, this paper conjectures that there is signal deepening between LGMDs. We propose a composite model by deepening the forward dendritic circuitry of existing LGMD models. Two types of ensemble herein are investigated with which the LGMD1 cascaded with LGMD2 is named LGMD1-LGMD2, and the LGMD2 cascaded with LGMD1 is named LGMD2-LGMD1. We carry out systematic experiments to test our hypothesis. The results demonstrate that compared with the single neuron models, the proposed composite model features extreme looming selectivity to only approaching objects over other categories of movements. Specifically, the selectivity of either LGMD1 or LGMD2 has been mutually enhanced built upon their original selectivity, i.e., the LGMD1-LGMD2 ensemble is sensitive to approaching white and dark objects, not to any receding or translating stimuli; while the LGMD2-LGMD1 ensemble only responds to approaching dark objects. Accordingly, this research provides a novel case study on the ensemble of motion perception neuronal models to achieve stronger selectivity that a single neuron model cannot accomplish.**

***Index Terms*—Looming perception, Ultra-selectivity, Neural ensemble, Mutual enhancement, LGMD**

## I. Introduction

Fast-and-reliable collision detection is indispensable for the survival of living things and the safe operation of machines. There are many excellent collision detection methods available, including those in deep learning, such as dynamic soft-attention proposed by Chan et al. [1] in the LSTM model. Bao et al. [2] proposed to learn spatiotemporal representation through graph convolution network and recurrent neural network. Among them, there is a class of methods inspired by nature that are effective, simple, and low-power. More and more researchers are thinking about how to mine inspiration from living things to solve problems [3] [4].

Organisms are able to recognize collisions very robustly and are not interested in other motions because they show extreme selectivity to objects that approach eyes. Nowadays, modeling individual neurons can achieve their respective selectivity, but the selectivity is still far from that exhibited by organisms. Accordingly, single neuron computation has its limit to achieve the capability exhibited by organisms, while an ensemble of similar neurons can reach closer to that. Just like the compounding of biological systems, we believe that models need to be compounded to enhance their ability. How neurons coordinate is currently an open question with different considerations. In this paper, we exemplify an interesting approach to coordinate two looming sensitive neuron models inspired by locust's visual systems.

### A. Related Work and Proposed Hypothesis

Based on the dissection of locust's lobula structures, biologists have discovered that locusts use a set of special neurons called lobula giant motion detectors (LGMDs) to perceive looming motion [5] [6], helping them to avoid collisions in high-speed swarm flight. LGMDs that have been identified by biologists include LGMD1 and LGMD2 [7], which are physically very close to each other, have similar physiological functionality but different **looming selectivity**: LGMD1 detects the approach of dark or bright objects, while LGMD2 responds only to dark, looming objects.

One approach to modeling LGMDs is to build a single-pathway model [8] [9]. Biologists revealed that in the initial motion detection circuits of many animals, luminance change is split into ON and OFF pathways [10] [11]. Researchers began to introduce the polarity pathways into LGMDs neurons and constructed many different bio-plausible models [12]. The LGMD1 neural network with ON/OFF pathways was established by Fu et al. by adding spatiotemporal computations to multiple layers of each pathway [13]. Lei et al. proposed an LGMD1 model against the translation motion to improve the looming selectivity [14]. Fu et al. developed a dual-channel LGMD2 neural network, which achieved the specific selectivity of LGMD2 for the first time [15]. The looming selectivity of LGMD1 and LGMD2 can already be completely distinguished by incorporating ON/OFF pathways, well demonstrated in mobile machines including UAV [16] and ground robot [17].

These models can well portray the looming selectivity of a single neuron. However, the capability of collision detection is still far from the ultra-selectivity exhibited by organisms. Like biologists, we think that biological ultra-selectivity is the result of a group of neurons working collaboratively, but biologists have no answer for now as to how they function together.

With regard to collision perception, very few people have studied the complementary functionality and selectivity of

neurons working together. Previously, Fu et al. attempted to connect the LGMDs models of locusts and the neuronal models of flies in parallel, using the coordination and competition between the two sets of neurons to construct the model [18] [19]. However, this approach only coordinates the outputs of individual neuron models and would respond to translating movements when relying only on the parallel combination of LGMDs. Differently to previous studies, we herein propose an interesting hypothesis: **there could be signals deepened within the forward circuits of LGMD neurons, with which the selectivity built upon a single neuron could be further sharpened up**.

### B. Contribution

In this paper, the looming selectivity of the composite models is investigated through systematic and comparative experiments, and the effectiveness is verified using both synthetic stimuli and complex real-world data. The main contributions of this paper are threefold: (1) The modeling provides a new perspective that can be borrowed for how to computationally composite multiple neurons from the biological nervous system. (2) By compounding two neurons, LGMD1 and LGMD2, they mutually enhance their selectivity to only approaching objects built upon their original selectivity, a step closer to ultra-selectivity represented by organisms. (3) The enhanced selectivity can be maintained in very complex vehicle and UAV scenarios.

The rest of this paper is organized as follows: Section II introduces the proposed method; Section III conducts comparative experiments with analysis; Section IV recalls our hypothesis and discusses future work; Section V summarizes this paper.

## II. Formulation of the Model

In this section, we first present the core structure of this composite model and then introduce each layer of the neural network separately in the following subsections.

As shown in Fig.1, the composite model consists of two subneural networks and the Leaky Integrate-and-Fire. The signals received from the visual field are computed spatiotemporally in two subneural networks. Each sub-neural network consists of four layers of basic structure, including photoreceptor (P), excitation (E), inhibition (I), and summation (S), with different choices of convolution kernels, inhibition coefficients, and channel weights capable of building either LGMD1 or LGMD2 models. Thus, mathematically, we can build different types of models using the architecture of this composite model, including LGMD1-LGMD2 and LGMD2-LGMD1. They differ in the order of signal deepening transfer, which leads to a difference in their looming selectivity .

Finally, the membrane potential is output using Leaky Integrate-and-Fire and the potential collision threat is indicated by the generated spikes.

### A. Initial Photoreceptor Layer

The first computational layer consists of photoreceptors arranged in a matrix, which capture the grayscale luminance $L(x, y, t)$ in the field of view.

$$M(x, y, t) = \alpha \left( L(x, y, t) - L(x, y, t-1) + M(x, y, t-1) \right) \tag{1}$$

where $x$ and $y$ respectively represent the horizontal and vertical positions in the camera coordinates and $t$ means the moment of the input video. $\alpha$ is a decay coefficient, including time delay constant $\tau_1$ and the time interval $\tau_{in}$ between the consecutive frames of digital signals in milliseconds, and is calculated by $\alpha = \tau_1/(\tau_1 + \tau_{in})$.

Next, Gaussian Blur is applied to process the received information to remove the isolated noise points on the image, the output of the P cell is given by the equation:

$$P^k(x, y, t) = \sum_{i=-1}^{1} \sum_{j=-1}^{1} M(x+i, y+j, t) W_p(i+1, j+1) \tag{2}$$

where $W_p$ is a normalized Gaussian Blur kernel of size 3*3, standard deviation is 1.

To simulate the inhibition exhibited by neurons in the face of large-scale luminance changes in the visual field, we use the adaptive inhibition mechanism (AIM) to adjust the time-varying bias of the ON and OFF channels [15].

$$PM(t) = \sum_{x=1}^{R} \sum_{y=1}^{C} |M(x, y, t)| \cdot (C \cdot R)^{-1} \tag{3}$$

$$\hat{PM}(t) = \left( PM(t), \hat{PM}(t-1), \hat{PM}(t-2) \right) \cdot \vec{\theta} \tag{4}$$

$$\omega_1(t) = \max(\omega_3, \frac{\hat{PM}(t)}{T_{PM}}), \omega_2(t) = \max(\omega_4, 2\omega_1(t)) \tag{5}$$

where $R$ and $C$ denote the rows and columns of the video frame, $\omega_3$ and $\omega_4$ indicate different baselines of bias. The coefficient vector $\vec{\theta}$ is $(0.6, 0.3, 0.1)'$.

### B. Partial Neural Network

The PNN receives signal input from the initial P layer or the S layer of the previous neuron, flows the information into different channels and computes it through the ON/OFF mechanism, and finally fuses the information from different pathways together as the output. The corner markers $k \in \{1, 2\}$ represent the type of PNN, including LGMD1 and LGMD2.

$$P^k_{ON}(x, y, t) = \left[ P^k(x, y, t) \right]^+ + \beta \cdot P^k_{ON}(x, y, t-1) \tag{6}$$

$$P^k_{OFF}(x, y, t) = -\left[ P^k(x, y, t) \right]^- + \beta \cdot P^k_{OFF}(x, y, t-1) \tag{7}$$

where $[x]^+$ and $[x]^-$ denote $\max(x, 0)$ and $\min(x, 0)$.

Local excitation $E^k_{ON}$ and $E^k_{OFF}$ are responsible for receiving signals from the corresponding channels in the P layer and doing convolution on spatial locations., where the convolution kernel $W_e$ is defined by

$$[W_e] = \frac{1}{8} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{8}$$

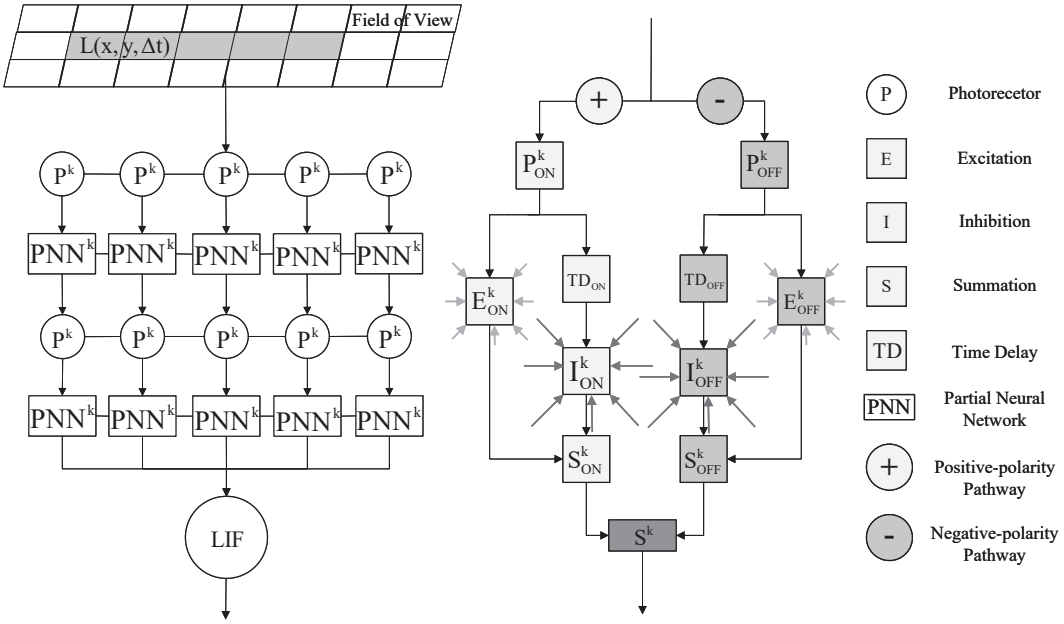



Fig. 1. The left part illustrates the overall architecture of the composite model, and the right part shows the structure of the local neural network. By choosing different groups of blocks, different types of composite models can be constructed. The arrow part indicates that the convolution is done at the spatial location, and the arrow length is positively correlated with the radius of the convolution kernel. LIF stands for leaky-integrate-and-fire neuron.

The first-order low-pass filtering method is used to add time-delayed units to the local excitation, and the local inhibition is formed by the influence of the elements around the spatial convolution. The whole process can be defined as

$$D^k_{ON}(t) = \left(E^k_{ON}(t), D^k_{ON}(t-1), D^k_{ON}(t-2)\right) \cdot \vec{\alpha}^k_{ON} \quad (9)$$

$$D^k_{OFF}(t) = \left(E^k_{OFF}(t), D^k_{OFF}(t-1), D^k_{OFF}(t-2)\right) \cdot \vec{\alpha}^k_{OFF} \quad (10)$$

$$I^k_{ON}(x,y,t) = \sum_{i=-2}^{2}\sum_{j=-2}^{2} D^k_{ON}(x+i, y+j, t) W^k_{I_{on}}(i+1, j+1) \quad (11)$$

$$I^k_{OFF}(x,y,t) = \sum_{i=-2}^{2}\sum_{j=-2}^{2} D^k_{OFF}(x+i, y+j, t) W^k_{I_{off}}(i+1, j+1) \quad (12)$$

where the values of $\vec{\alpha}^1_{ON}$ and $\vec{\alpha}^1_{OFF}$ are both $(0.5, 0.3, 0.2)'$, indicating that the delay of the two channels in LGMD1 is the same. The value of $\vec{\alpha}^2_{ON}$ is $(0.7, 0.2, 0.1)'$, and the value of $\vec{\alpha}^2_{OFF}$ is $(0.5, 0.3, 0.2)'$, which means that the ON channel of LGDM2 shorens delay and improves local inhibition. They all satisfy the characteristic that the closer the time is, the larger the value taken and the component sum is 1. The inhibited convolution kernel of each of the two PNN is defined as

$$\left[W^k_{I_{off}}\right] = \frac{1}{16}\begin{bmatrix} 1 & 2 & 4 & 2 & 1 \\ 2 & 4 & 8 & 4 & 2 \\ 4 & 8 & 16 & 8 & 4 \\ 2 & 4 & 8 & 4 & 2 \\ 1 & 2 & 4 & 2 & 1 \end{bmatrix} \quad (13)$$

$$\left[W^1_{I_{on}}\right] = \left[W^1_{I_{off}}\right], \left[W^2_{I_{on}}\right] = 2 * \left[W^2_{I_{off}}\right] \quad (14)$$

After both ON/OFF channels generate local excitation and inhibition, they will perform a purely linear calculation at their respective summation cells, described as

$$S^k_{ON}(x,y,t) = \left[E^k_{ON}(x,y,t) - \omega^k_{ON}(t) \cdot I^k_{ON}(x,y,t)\right]^+ \quad (15)$$

$$S^k_{OFF}(x,y,t) = \left[E^k_{OFF}(x,y,t) - \omega^k_{OFF}(t) \cdot I^k_{OFF}(x,y,t)\right]^+ \quad (16)$$

where $\omega^k_{ON}(t)$ and $\omega^k_{OFF}(t)$ are time-varying local biases to control the strength of the inhibited flows, obtained through AIM, defined as

$$\omega^1_{ON}(t) = \omega^1_{OFF}(t) = \omega^2_{OFF}(t) = \omega_1(t), \omega^2_{ON}(t) = \omega_2(t) \quad (17)$$

The summation units of ON and OFF channels are integrated by the following equation, where the values of $\gamma_1$ and $\gamma_2$ are 0.5. As the models are deepened, the signal is generally weakened, which may lead to some small signals being filtered by the threshold, making the extracted signals disappear. Therefore, we enhance the small signals by a power function with an exponent between 0 and 1, which is

$$S^k(x,y,t) = \left[S^k_{ON}(x,y,t)\right]^{\gamma_1} + \left[S^k_{OFF}(x,y,t)\right]^{\gamma_2} \quad (18)$$

This method is able to reduce the numerical differences between pixel points while maintaining their original size relationships, allowing multiple features extracted to be preserved as much as possible.

### C. Intermediate Photoreceptor Layer

This layer is responsible for connecting the two PNNs, which receives the output from the $S$ layer of the previous neural network and calculates the change in brightness over two consecutive frames as input to the next neural network, described as

$$P^k(x,y,t) = S^j(x,y,t) - S^j(x,y,t-1) \quad (19)$$

### D. Leaky Integrating-and-Firing for Collision Detection

After two PNNs, we use a normalized Gaussian kernel of radius 5 and standard deviation 1 to do the Gaussian Blur

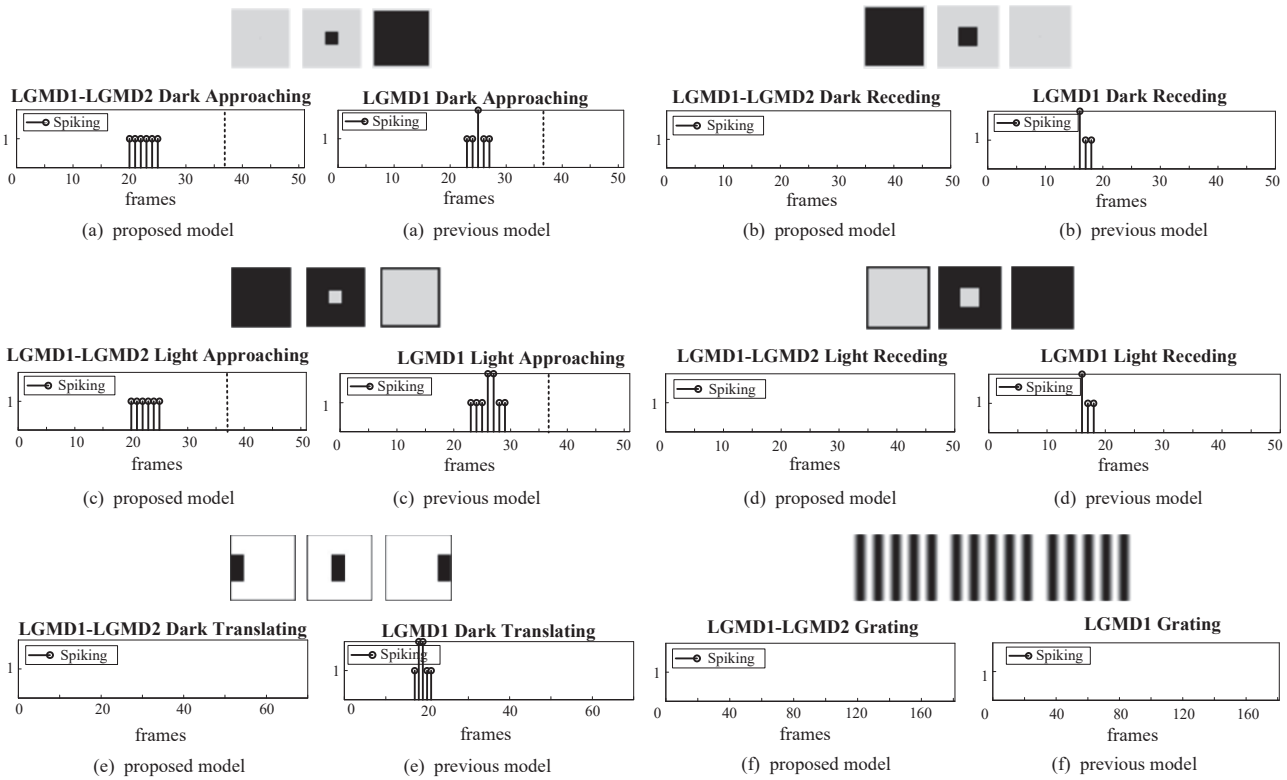



Fig. 2. Variation of spike following time for LGMD1-LGMD2 under different stimulations. (a) Dark looming. (b) Dark recession. (c) Bright looming. (d) Bright recession. (e) Dark translation. (f) Grating movement. The three frames captured include the beginning, middle, and end of the video. The dashed line marks the moment when the object is no longer moving, indicating the time to collision.

at the center of each pixel point. Then, we add the dynamic contrast normalization mechanism [20] to process the signal to reduce the effect of contrast on the model. Next, we perform the convolution with an equal-weight kernel of radius 1, which satisfies the sum of all elements to 1. Last, we introduce the spike frequency adaptation (SFA) mechanism [15] to further improve the selectivity of the model and output the result $\hat{S}(x, y, t)$ to Leaky Integrate-and-Fire.

The output neuron integrates the current stimulus across all pixels and emits spikes to modulate the membrane potential output.

$$\hat{g}(t) = \sum_{x=1}^{R} \sum_{y=1}^{C} \hat{S}(x, y, t) \tag{20}$$

$$spi(t) = \begin{cases} 1 & \text{if } \hat{g}(t) + \hat{g}(t-1) \geq V_{th} \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

$$\nu(t) = \hat{g}(t-1) \cdot e^{-\frac{t-t_1}{\tau_{in}}} + \hat{g}(t) - spi(t) \cdot V_{th} + V_{rest} \tag{22}$$

where $V_{th}$ indicates the spiking threshold, $t_1$ is the time of the last emission spike and $V_{rest}$ denote resting potential. Finally, the following equation is used to represent the potential collision threat in the experiments, defined as

$$Col(t) = \begin{cases} True & \text{if } \sum_{i=t-n_{ts}}^{t} spi(t) \geq n_{sp} \\ False & \text{otherwise} \end{cases} \tag{23}$$

where $n_{sp}$ represents the threshold for the number of spikes within time window $n_{ts}$.

## III. Experimental Evaluation

### A. Setting the Parameters

The parameters are set as shown in table I, where $T_{PM}$ is responsible for regulating the weights of the inhibition layer and $V_{th}$ influences the number of spikes. These two parameters are extremely important for collision warning, with relatively small values in simple scenes and large in complex.

Biological research has brought us the structure of neural networks and the mechanisms of signaling, turning the structure of neurons into semi-transparent black boxes so that we can select parameters by previous experience, and it is in this way that the proposed model selects parameters. Each block of the composite model already has well-established parameters, but as learning is not part of the current work, the parameters will be further optimized by learning in the future.

### B. Setting the Experiments

The proposed model framework, the model test, and the visualization of channel features are implemented in MATLAB 2018b. The experimental design is composed of two parts: one experiment with a single stimulus to test the model's response to different types, and another experiment consisting of car crash simulations and offline testing with UAV to assess whether the model can provide reasonable warning signals in real-world scenarios. We set the warning signal to emit four consecutive spikes.

All visual stimulus videos were captured at a frame rate of 30 frames per second, with varying sample image sizes.

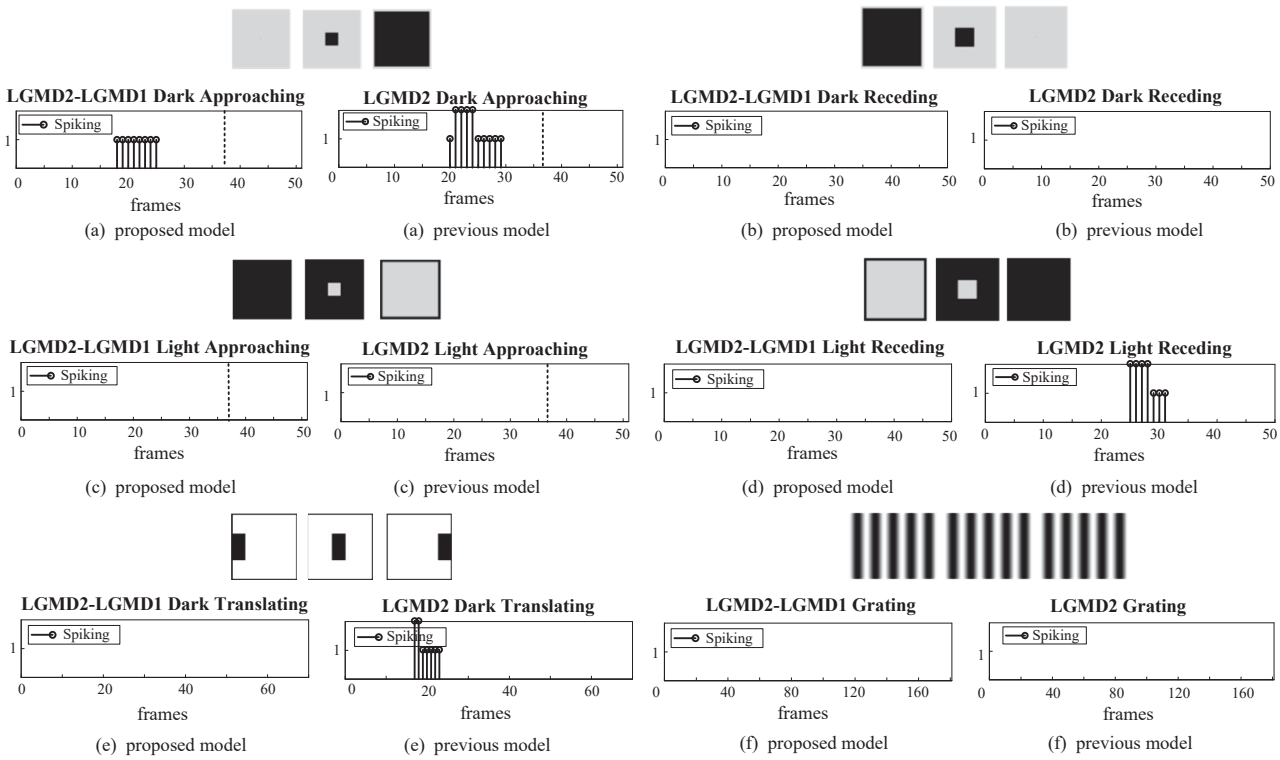



Fig. 3. Variation of spike following time for LGMD2-LGMD1 under different stimulations. Stimulus types and labeling are kept the same as in Fig. 2.

TABLE I
NETWORK PARAMETERS

| Parameter | Description | Value |
|---|---|---|
| $\tau_1$ | time delay constant | 100 |
| $\tau_{in}$ | time interval of input digital signal | $30 \sim 50$ |
| $\beta$ | coefficient in half-wave function | 0.1 |
| $T_{PM}$ | threshold in AIM | $10 \sim 50$ |
| $\omega_3, \omega_4$ | bias baseline | 0.6, 0.3 |
| $\gamma_1, \gamma_2$ | coefficient in signal combination | 0.5, 0.5 |
| $Ccn$ | baseline value of contrast normalization | $3 \sim 10$ |
| $T_{sfa}$ | threshold constant of SFA mechanism | $0.003 \sim 0.1$ |
| $V_{th}$ | threshold constant of spiking | $10 \sim 50$ |
| $V_{rest}$ | resting potential | -4 |
| $n_{ts}$ | time window for spikes | $3 \sim 6$ |
| $n_{sp}$ | number of spikes within $n_{ts}$ | $4 \sim 6$ |

Specifically, in the single-stimulus experiments, the sample image size was $100 \times 100$ and $320 \times 240$ pixels for the object translation and grating videos, respectively, and $600 \times 600$ pixels for the remaining videos. In the car crash experiment, a sample image size of $426 \times 240$ pixels was used for all videos. In the UAV offline test, all videos were captured at a resolution of $720 \times 480$ pixels. Our code and testing stimuli are contributed as open source at Github.

### C. LGMD2's Network Deepening LGMD1's

Our composite model uses LIF output spike signals instead of an exponential mapping mechanism in the previous LGMD model, producing at most one spike at any moment. Consequently, at each moment, we only consider spike production, not quantity. As shown in Fig. 2, experiments were conducted using synthetic stimuli to test the LGMD1-LGMD2 composite model and a previous model from Fu et al. [13] under the same conditions. The previous model had been successful in simulating the responses of individual LGMD1 neurons.

The experiments reveal three important findings: (1) The LGMD1-LGMD2 model provides an earlier warning time for looming objects than the previous model, although the warning occurs before collision. (2) The LGMD1-LGMD2 model shows selective responses to looming for bright and dark objects, exhibiting a continuous spike signal, and is not sensitive to receding, translating, and grating stimuli. (3) The selectivity demonstrated by the LGMD1-LGMD2 model is closer to the biological responses compared to those of the single LGMD1 neuron model.

### D. LGMD1's Network Deepening LGMD2's

The experiments utilized the single stimuli described in Subsection III-B to test the responses of both the LGMD2-LGMD1 composite model and the previous LGMD2 model [15]. The results of these experiments are presented in Fig. 3.

The experimental findings reveal three significant points: (1) The LGMD2-LGMD1 model only responds to dark objects looming and can provide an earlier warning signal than the previous model before the collision occurs. (2) In contrast to the previous LGMD2 model, the LGMD2-LGMD1 model does not respond to bright objects receding. This result suggests that the LGMD2-LGMD1 model maintains the sensitivity of LGMD2 neurons to movements with reduced luminance

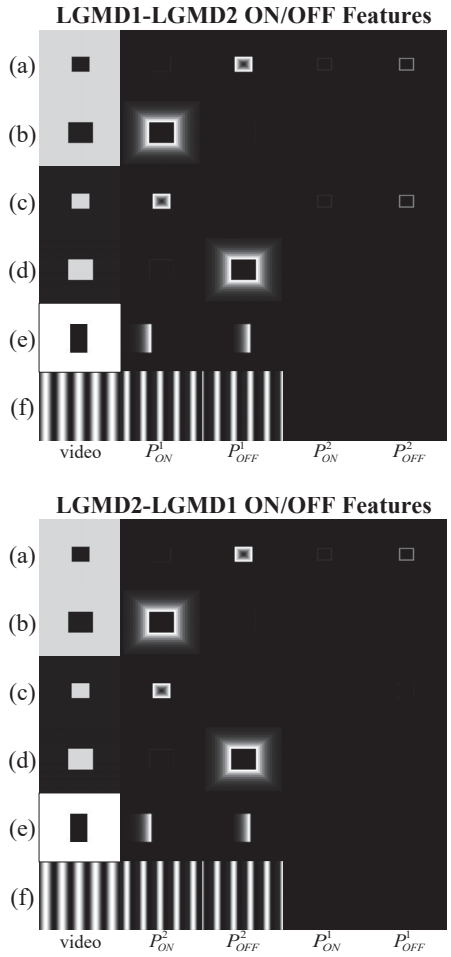



Fig. 4. The four channels $P^k_{ON}$ and $P^k_{OFF}$ visualizations of LGMD1-LGMD2 (top) and LGMD2-LGMD1 (bottom). The visualization is a kind of heat-map representation where the brighter the pixel point means the value is larger. The selected frame number is the middle moments of the video, and the stimulus type is consistent with that in Fig. 2 .

values while being able to distinguish between looming and receding stimuli. (3) The LGMD2-LGMD1 model does not respond to receding, translating, and grating stimuli.

### *E. Representation of Features Entering ON/OFF Channels*

The two PNNs of the model accept the brightness changes of the field of view and the S layer of the previous PNN respectively, and flow the signals into the ON and OFF channels of the P layer. We visualized the $P^k_{ON}$ and $P^k_{OFF}$ channels and enhanced the image with features.

From the Fig.4 we can be found that : (1) The inputs of the first PNN are the same for both different types of composite models. (2) Even for the simplest looming and receding, unlike the single neuron model, the composite model is a joint determination of selectivity by both channels. (3) LGMD1-LGMD2 responds to bright object looming, while LGMD2-LGMD1 does not. In the LGMD1-LGMD2 model, the S layer of LGMD1 receives information for the bright object looming and performs frame difference to the signal at the P layer of LGMD2. As the edges expand, there is a signal input to the $P^2_{OFF}$ channel, which allows the model to issue a response. However, in the LGMD2-LGMD1 model, the bright object looming is suppressed by LGMD2, so LGMD1 cannot receive the signal, resulting in no response from the model.

### *F. Effectiveness in Vehicle Crash Detection*

To test whether the LGMD1-LGMD2 and LGMD2-LGMD1 models can adapting to complex environments, three car crash experiments were conducted on each of them. Fig.5 and Fig.6 demonstrate the performance of both models in the face of visual challenges, with the models still able to give early warning signals before a collision occurs, regardless of whether the background is brightly lit during the day or dimly lit at night. The visualization shows that this composite model is able to focus attention well on the vehicle with the threat of collision, and the overall performance is satisfactory, although there may be interference from surrounding noise.

Of course, we have only done preliminary tests to demonstrate the feasibility of the composite model in real-life scenarios. To adapt the model to more complex scenarios, learning methods will be added in the future to optimize the experimental parameters, and in due course, the composite model will be compared with deep learning methods.

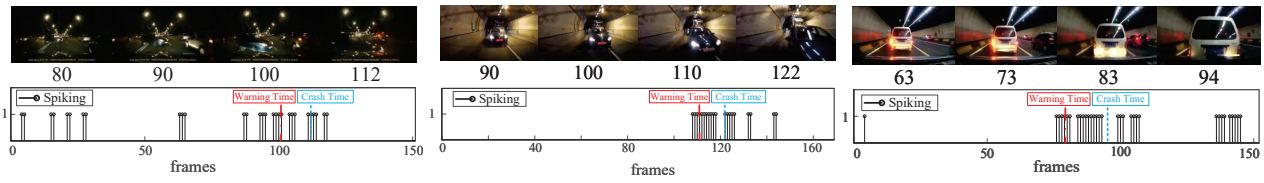
### *G. Offline UAV Testing*

In our future research, we may consider applying the composite model to real-time systems of UAVs. Therefore, we conducted offline tests of the composite model from the first-person perspective of UAVs to verify the feasibility of integrating the model with UAVs.

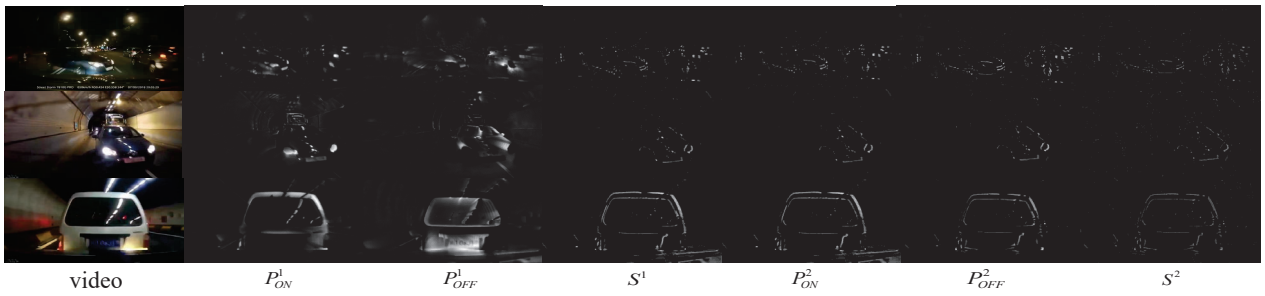
As shown in the Fig. 7, we validated the two types of models using three different types of stimuli, and the experiments mainly revealed the following two points: (1) Both models exhibited extreme selectivity. The models emitted continuous and strong spike signals as the UAVs approached the targets, and the spikes became sparse and discrete as the UAVs moved away from the targets. The models also provided warning signals for objects that suddenly appeared in the field of view, but did not respond to the relative translation of the UAVs with respect to the objects. (2) The models demonstrated a certain degree of noise resistance. Since the videos were collected outdoors, there were many noise inputs to the models. However, the models were still able to function normally, indicating their practicality.

## IV. Discussion

As shown in the Fig.8(a), LGMD1 and LGMD2 are neighboring neurons, physically close to each other [21], so it is natural for us to associate whether there will be deepened transmission of signals between the dendrites or axons of the two neurons. Through computer modeling and experimentation, we found that two neurons with their selectivity could enhance each other's selectivity through deepening the forward dendritic circuitry, achieving ultra-selectivity for approaching objects, see Fig.8(b). However, the composite model also has problems: one is that it is less sensitive to movements than the previous models [13] [15], as evidenced by the earlier warning

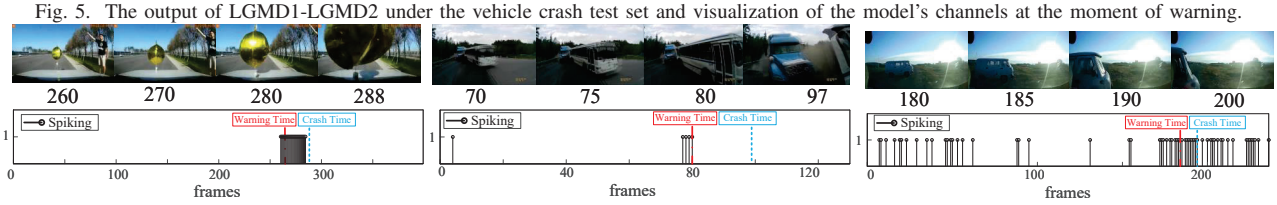



(a) sudden turning-around vehicle (b) night-time rear-end collision (c) vehicle following crash

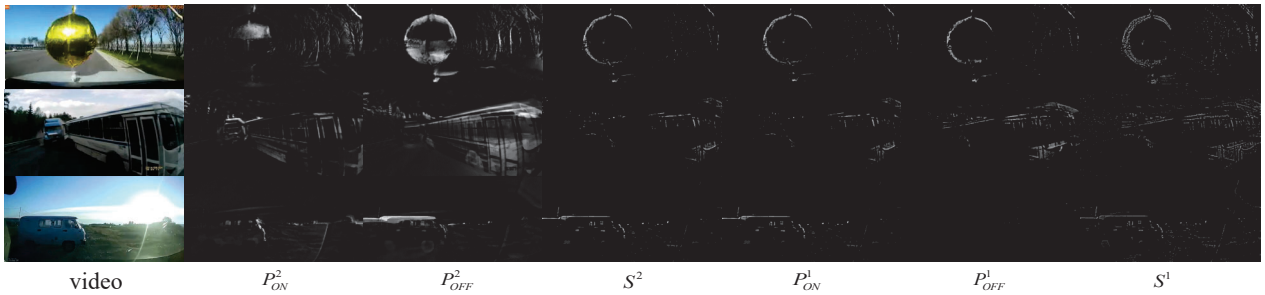



(d) LGMD1-LGMD2 Channel Feature Visualization

Fig. 5. The output of LGMD1-LGMD2 under the vehicle crash test set and visualization of the model's channels at the moment of warning.




(a) head-on collision with a balloon (b) serial vehicle collision (c) crossing car collision in strong light




(d) LGMD2-LGMD1 Channel Feture Visualization

Fig. 6. The output of LGMD2-LGMD1 under the vehicle crash test set and visualization of the model's channels at the moment of warning.

signal; the other is that its complexity is higher than that of each building block but still acceptable for now.

Both models have enhanced looming selectivity, but there are still shortcomings. Firstly, our model lacks a rigorous mathematical description to support it, and it may be possible to select more appropriate parameters mathematically to achieve better results. Secondly, the testing dataset used to verify the proposed composite model is relatively limited; for future investigation, systematic testing using a wider range of videos with various contrast and statistical properties is necessary. Furthermore, real-time online verification using mobile robots [4] to implement the composite model will be more challenging and convincing.

## V. Concluding Remarks

In this paper, we provide an idea of multiple neurons functioning in concert, through deepening signal transmission between dendritic anterior circuits to form composite models. We test our hypothesis through computational modeling and comparative experiments using both synthetic stimuli and real-world scenarios. Our proposed method of compositing different LGMD neurons to form ultra-selectivity to approaching objects not only enhances mutually the two LGMDs' selectivity but also exemplifies a novel perspective of neural ensemble in the context of visual motion perception.

## Acknowledgment

This research has received funding from the National Natural Science Foundation of China under the Grant No. 12031003, No. 12271117, No. 12211540710, No. 62206066, and the Social Science Fund of the Ministry of Education of China under the Grant No. 22YJCZH032. *Correspondence: Qinbing Fu and Jigen Peng* {qifu,jgpeng}@gzhu.edu.cn

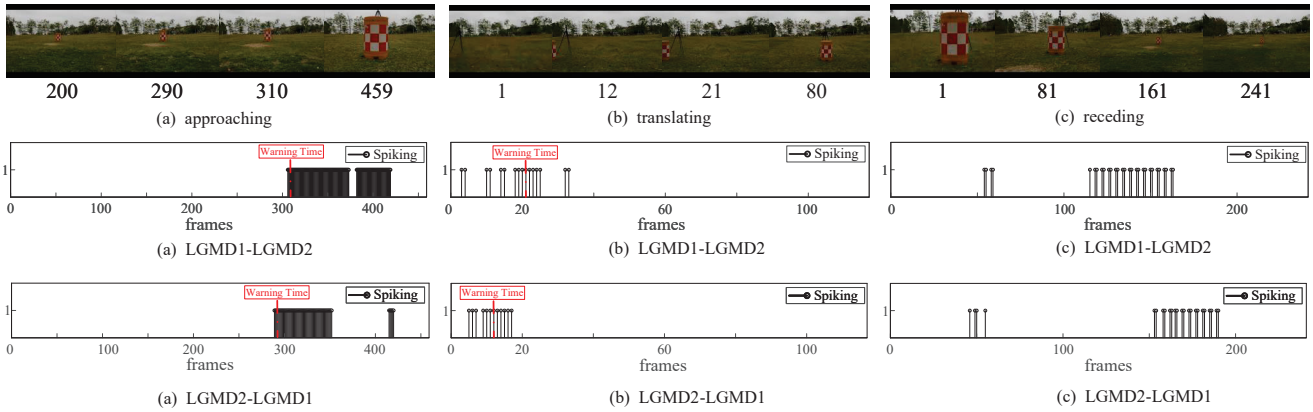
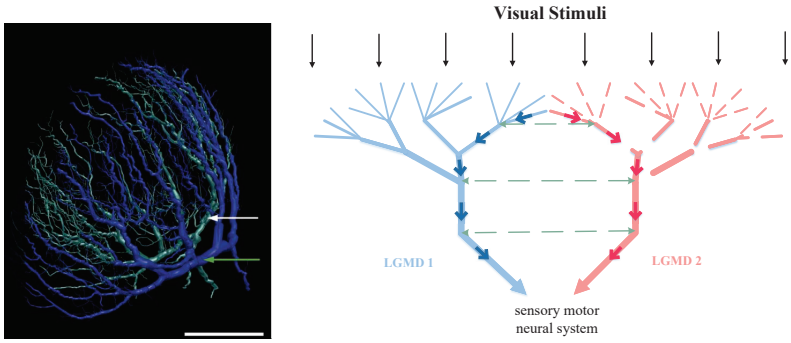



Fig. 7. Spike responses of composite models LGMD1-LGMD2 and LGMD2-LGMD1 in offline testing with UAV under different stimuli




(a) neuromorphology (b) our hypothesis of signal deepening

Fig. 8. Figure (a) is a morphological schematic of two neighbour LGMD neurons with the white arrow pointing to LGMD1 and the green arrow pointing to LGMD2. Scale bar: 50 μm, image courtesy of [21]. Figure (b) is a schematic of the signal transmission of the composite model. The green dashed line indicates the possible presence of signals deepening between two neural circuits, not yet been identified by biologists.

## REFERENCES


[1] F. Chan, Y. Chen, and M. Sun, "Anticipating accidents in dashcam videos," *ACCV 2016*, vol. 10114, pp. 136–153, 2017.

[2] W. Bao, Q. Yu, and Y. Kong, "Uncertainty-based traffic accident anticipation with spatio-temporal relational learning," *Proceedings of the 28th ACM International Conference on Multimedia*, pp. 2682–2690, 2020.

[3] B. Hu, Z. Zhang, and L. Li, "Lgmd-based visual neural network for detecting crowd escape behavior," *IEEE International Conference on Cloud Computing and Intelligence Systems*, p. 772778, 2018.

[4] C. Hu, Q. Fu, and S. Yue, "Colias IV: The affordable micro robot platform with bio-inspired vision," in *Towards autonomous robotic systems conference*, 2018, pp. 197–208.

[5] F. Rind and D. Bramwell, "Neural network based on the input organization of an identified neuron signaling impending collision," *Journal of Neurophysiology*, vol. 75, no. 3, pp. 967–985, 1996.

[6] S. Yue and F. Rind, "Collision detection in complex dynamic scenes using an lgmd-based visual neural network with feature enhancement," *IEEE Transactions on Neural Networks*, vol. 17, no. 3, pp. 705–716, 2006.

[7] P. J. Simmons and F. C. Rind, "Responses to object approach by a wide field visual neurone, the lgmd2 of the locust: Characterization and image cues," *Comparative Physiology A*, vol. 180, pp. 203–214, 1997.

[8] A. C. Silva, J. Silva, and C. P. dos Santos, "A modified lgmd based neural network for automatic collision detection," *ICINCO*, vol. 283, pp. 217–233, 2014.

[9] S. Zhang, G. Lei, and X. Liang, "A single-pathway biomimetic model for potential collision prediction," *Pattern Recognition and Computer Vision*, vol. 13536, pp. 165–178, 2022.

[10] A. Borst and T. Euler, "Seeing things in motion: Models, circuits, and mechanisms," *Neuron*, vol. 71, no. 6, pp. 975–994, 2011.

[11] A. Borst and M. Helmstaedter, "Common circuit design in fly and mammalian motion vision," *nature neuroscience*, vol. 18, no. 8, pp. 1067–1076, 2015.

[12] M. S. Keil, E. Roca-Moreno, and A. Rodriguez-Vazquez, "A neural model of the locust visual system for detection of object approaches with real-world scenes," *IASTED*, pp. 340–345, 2004.

[13] Q. Fu, C. Hu, J. Peng, and S. Yue, "Shaping the collision selectivity in a looming sensitive neuron model with parallel on and off pathways and spike frequency adaptation," *Neural networks*, vol. 106, no. 21, pp. 127–143, 2018.

[14] F. Lei, Z. Peng, V. Cutsuridis, M. Liu, Y. Zhang, and S. Yue, "Competition between on and off neural pathways enhancing collision selectivity," *International Joint Conference on Neural Networks*, pp. 1–8, 2020.

[15] Q. Fu, C. Hu, J. Peng, F. C. Rind, and S. Yue, "A robust collision perception visual neural network with specific selectivity to darker objects," *IEEE Transactions on Cybernetics*, vol. 50, no. 12, pp. 5074–5088, 2020.

[16] L. Salt, G. Indiveri, and Y. Sandamirskaya, "Obstacle avoidance with lgmd neuron: Towards a neuromorphic uav implementation," *IEEE International Symposium on Circuits and Systems*, pp. 1–4, 2017.

[17] Q. Fu, H. Wang, J. Peng, and S. Yue, "Improved collision perception neuronal system model with adaptive inhibition mechanism and evolutionary learning," *IEEE Access*, vol. 8, pp. 108 896–108 912, 2020.

[18] Q. Fu and S. Yue, "Complementary visual neuronal systems model for collision sensing," *International Conference on Advanced Robotics and Mechatronic*, pp. 609–615, 2020.

[19] Q. Fu, X. Sun, T. Liu, C. Hu, and S. Yue, "Robustness of bio-inspired visual systems for collision prediction in critical robot traffic," *Frontiers in Robotics and AI*, vol. 8, pp. 1–19, 2021.

[20] Q. Fu, Z. Li, and J. Peng, "Harmonizing motion and contrast vision for robust looming detection," *Array*, vol. 17, p. 100272, 2022.

[21] J. Sztarker and F. Rind, "A look into the cockpit of the developing locust: Looming detectors and predator avoidance," *Developmental Neurobiology.*, vol. 74, no. 11, pp. 1078–1095, 2014.