A Neural Model of Crossover Inhibition for Looming Motion Detection

Lixiao Liang[1], Ziyan Qin[2], Shaobing Gao[3], Qinbing Fu[*,1]
[1] School of Mathematics and Information Science, Guangzhou University, China
[2] School of Mathematics and Systems Science, Guangdong Polytechnic Normal University, China
[3] College of Computer Science, Sichuan University, China
[*] Corresponding author: qifu@gzhu.edu.cn

***Abstract*— Looming motion detection is a prerequisite for collision avoidance of both biological and artificial entities. In primate retinas, crossover inhibition could play a key role in detecting looming motion by mediating inhibitory interactions between parallel ON and OFF pathways. This mechanism enhances sensitivity to transient luminance changes and supports reliable perception of approaching objects. However, most existing bio-inspired models either neglect or oversimplify crossover inhibition, which limits their ability to distinguish approaching motion from receding or translating motion and leads to unreliable responses in noisy environments. To address this limitation, we propose a neural model of crossover inhibition for looming motion detection. The model adopts a biologically plausible ON/OFF-pathway framework encoding brightness increment (ON)/decrement (OFF), separately. It firstly enhances motion edges using a spatial operator of Difference of Gaussians, then incorporates crossover inhibition to regulate interactions between ON-OFF/OFF-ON pathways, suppressing redundant luminance fluctuations and non-approaching motion responses. Experimental results show that the proposed model achieves strong selectivity for approaching motion while effectively inhibiting responses to receding and translating stimuli. Quantitative evaluations demonstrate performance improvements of 10–22% over comparative bio-inspired models in real-world scenarios, with stable behavior maintained under severe synthetic noise. These findings confirm the functional importance of crossover inhibition in looming motion detection and demonstrate its potential for improving robustness in machine vision systems.**

***Index Terms*—Neural system modeling, ON/OFF pathways, Crossover inhibition, Looming detection, Approaching selectivity**

I. Introduction

The ability to sense and respond to objects approaching on a collision course–referred to as *looming motion detection*–is a fundamental survival mechanism across the animal kingdom. Among various sensory modalities, vision provides the richest and most informative cues for detecting such threats. Robust visual looming detection systems enable animals to navigate complex natural environments and even engage in large-scale collective migration over long distances while largely avoiding collisions.

In the development of artificial looming motion detection systems, biological vision has long served as a rich source of inspiration. Neuroscience observation suggests that the ability to detect approaching objects can emerge at very early stages of visual information processing [1], [2]. In particular, visual signals in primate retinas are processed through two parallel pathways: the ON pathway, which is sensitive to increases in luminance (ON contrast), and the OFF pathway, which responds to decreases in luminance (OFF contrast) [3]. This early-stage segregation of visual information enables efficient encoding of local contrast changes and provides a robust foundation for looming motion detection under both bright and dark environmental conditions.

Recent bio-inspired looming detection models have incorporated the segregated ON/OFF pathway architecture, enabling improved sensitivity to rapidly approaching objects in complex and dynamic environments [4]. However, these models often still exhibit pronounced responses to non-approaching motions, such as translation and recession, which leads to false detections and reduced robustness under realistic conditions. Consequently, reliably distinguishing approaching motion from non-approaching motion remains a central challenge for bio-inspired visual collision perception systems.

Beyond the segregation of visual signals, physiological studies have shown that the ON and OFF pathways interact through a crossover inhibition (CI) mechanism mediated by intermediate retinal neurons [5], as illustrated in Fig. 1. This mechanism introduces antagonistic interactions between pathways of opposite polarity, namely ON-to-OFF and OFF-to-ON inhibition. Previous studies have demonstrated that CI effectively suppresses redundant responses to symmetric luminance changes, thereby enhancing sensitivity to motion-related signals at early stages of visual processing [6]. Importantly, CI has also been implicated in the neural processing of looming motion [7], [8]. Despite its biological relevance, CI has rarely been explicitly incorporated into bio-inspired looming perception systems.

In consequence, we build upon the classical ON/OFF pathway framework and propose a CI–driven looming motion detection model (CI-LMD). In this model, CI introduces bidirectional antagonistic interactions between the ON and OFF pathways, such that activity in one pathway suppresses redundant responses in the opposite pathway. This interaction selectively enhances spatial-temporal cues associated with approaching motion while effectively attenuating symmetric luminance changes arising from translation or background fluctuations. Experimental results demonstrate that CI-LMD consistently outperforms existing bio-inspired models, achiev-

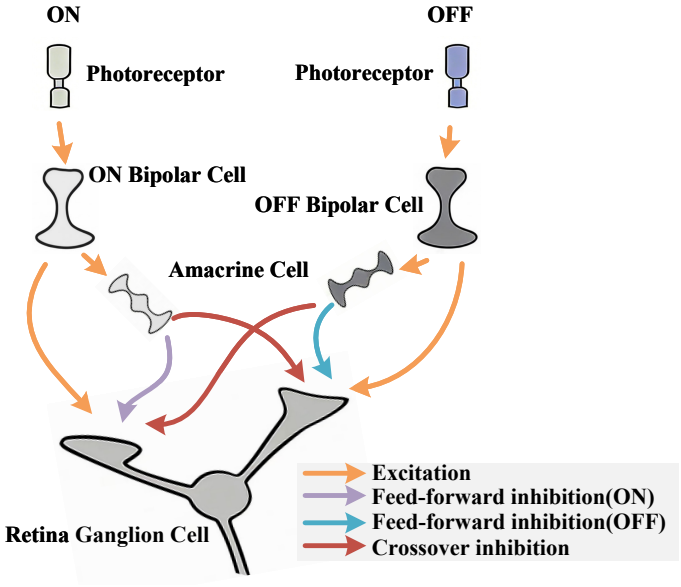


Fig. 1: Schematic illustration of crossover inhibition (CI) mechanism in the primate retinas. Within the parallel ON/OFF pathways, orange arrows represent excitatory pathways from photoreceptors to bipolar cells, as well as from bipolar cells to amacrine cells and retinal ganglion cells (RGCs). Red arrows denote crossover inhibition pathways acting on RGCs: the ON pathway receives inhibitory inputs mediated by OFF amacrine cells, while the OFF pathway is inhibited by ON amacrine cells. The remaining arrows indicate inhibitory pathways acting on ganglion cells via ON/OFF-type feed-forward inhibition.

ing up to 22% performance improvement in real-world aerial drone scenarios and maintaining stable performance under artificial noise. These results, on the other hand, validate the critical functional role of CI in neural processing for looming motion detection.

The remainder of this paper is structured as follows. Section II reviews related work with a focus on crossover inhibition. Section III presents the formulation of the proposed model. Section IV reports and analyzes the experimental results. Finally, Section V summarizes the main findings and outlines directions for future research. [1]

II. Related Work

In 1986, Wässle et al. experimentally demonstrated McGuire et al.'s (1984) [5] "push-pull hypothesis" using cat retinal preparations, establishing the core framework of CI for the first time, showcasing ON and OFF pathways interact through bidirectional cross-pathway inhibition [9]. Specifically, the ON pathway receives inhibitory signals from OFF amacrine cells, while the OFF pathway simultaneously receives inhibitory signals from ON amacrine cells (see Fig. 1).

As research progressed, the focus shifted toward exploring the role of CI in visual information processing [10], [11]. It was discovered that CI enhances the visual system response to transient luminance changes [12]. Besides, CI suppresses responses to continuous local contrast changes of opposite polarity [13]. Furthermore, Cafaro et al. found that introducing cross-pathway inhibition could improve the system's temporal sensitivity to spatial stimuli [14]. Notably, early CI research primarily focused on its general visual processing functions [15], without specifically addressing the scenario of "motion detection".

Until 2018, Manookin et al. overturned the traditional notion that "motion computation begins in the visual cortex" through primate retinal electrophysiology experiments [6]. They confirmed that the CI is the core source of retinal motion sensitivity, providing crucial research for the direct link between the CI mechanism and motion perception. In 2020, Appleby et al. advanced this understanding, demonstrating CI direct involvement in encoding approaching stimuli (characterized by spatial expansion) [8]. By dynamically modulating interactions between ON and OFF pathways, CI can specifically capture the spatiotemporal features of approaching motion [7].

However, existing studies on CI have largely been confined to biological experiments, with limited systematic computational modeling to validate its functional role in looming perception or to explore its potential in machine vision. To address this gap, this study develops a neural model grounded in the core mechanisms of CI. The proposed model provides principled computational validation of CI and offers new insights into its fundamental contribution to looming perception.

III. Method Formulation

This section presents the formulation of the proposed computational framework, illustrated in Fig. 2. Inspired by the encoding properties of primate retinal ganglion cells in response to looming stimuli, the model decomposes visual inputs into parallel ON and OFF pathways that selectively process increasing and decreasing luminance signals, respectively. Crucially, a biologically grounded CI mechanism is incorporated to mediate interactions between these two pathways. By constructing an interactive circuit that dynamically modulates oppositely polarized inputs, the proposed framework selectively amplifies responses to approaching motion while effectively suppressing non-approaching stimuli, such as translational and receding motion.

A. Early Visual Encoding

The first layer simulates the response characteristics of retinal photoreceptor cells. It consists of photoreceptors arranged in a matrix to capture gray-scale luminance and compute changes. Let the input image stream be $L(x, y, t)$, the temporally partial derivative is calculated as

$$P(x, y, t) = \frac{\partial L(x, y, t)}{\partial t}, \tag{1}$$

where $L(x, y, t)$ is the gray-scale luminance. This step simulates the transient response of photoreceptors to changes in luminance, enabling the filtering out of static background.

Then, a spatial operator vDOG is formulated to simulate the center-surround antagonistic of the receptive field [16]. The vDOG enhances motion edge features with ON/OFF-contrast

[1] Source at https://github.com/fuqinbing/CI-model-for-looming-detection

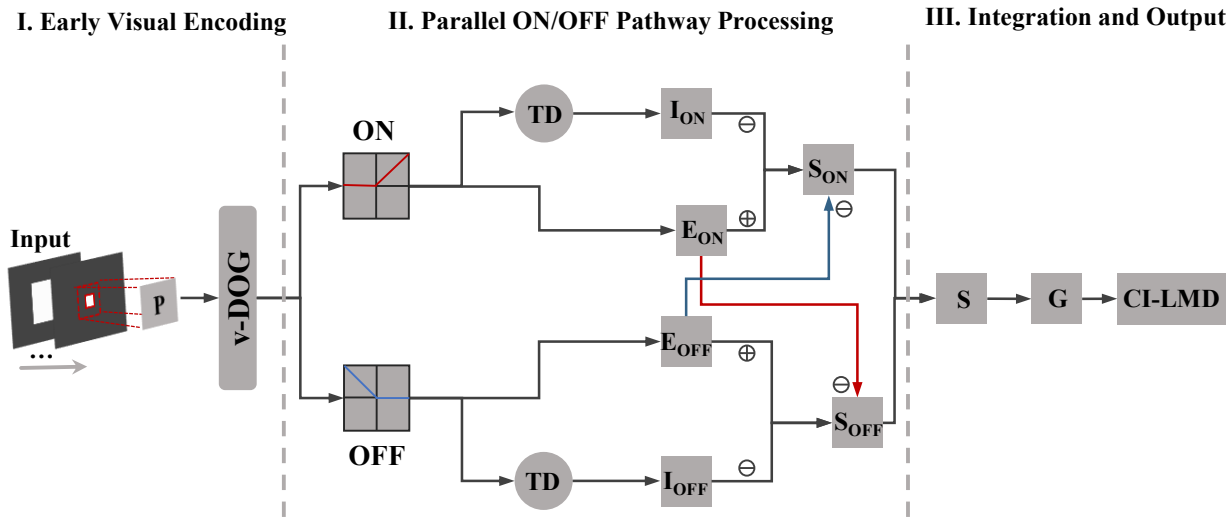


Fig. 2: Framework of the proposed model for looming detection. Abbreviations: P – photoreceptor, vDOG – variant Difference of Gaussians operator, E – excitation unit, I – inhibition unit, TD – time delay unit, S – summation unit, G – grouping unit, CI-LMD: output unit. The model consists of three core layers: (1) early visual encoding – capturing luminance changes and enhances motion edges, (2) parallel ON/OFF pathway – splitting visual signals into ON/OFF pathways, and modulates excitatory signals via feed-forward inhibition and crossover inhibition between the two pathways, (3) integration and output – combining polarized excitation in the S unit, suppressing noise and amplifies continuous motion edges in the G unit, finally integrating all remaining excitation to generate the final spatiotemporal output.

selectivity. The center excitatory field $P_e(x,y,t)$ and surround inhibitory field $P_i(x,y,t)$ are computed as

$$P_e(x,y,t) = \iint G_{\sigma_e}(u,v)P(x-u,y-v,t)\,du\,dv, \quad (2)$$

$$P_i(x,y,t) = \iint G_{\sigma_i}(u,v)P(x-u,y-v,t)\,du\,dv, \quad (3)$$

$$G(\sigma,u,v) = \frac{1}{2\pi\sigma^2}\exp(-\frac{(u^2+v^2)}{2\sigma^2}), \quad (4)$$

where σ_e and σ_i are the standard deviations of excitatory and inhibitory fields, respectively.

The photoreceptor response $RF(x,y,t)$ is obtained by combining center excitation and surround inhibition as

$$RF(x,y,t) = \begin{cases} |P_e(x,y,t) - P_i(x,y,t)|, & \text{s.t. } P_e(x,y,t) \geq 0, P_i(x,y,t) \geq 0 \\ -|P_e(x,y,t) - P_i(x,y,t)|, & \text{s.t. } P_e(x,y,t) < 0, P_i(x,y,t) < 0. \end{cases} \quad (5)$$

B. Parallel ON/OFF Pathway Processing

1) Signal Bifurcation: Next, the RF response is split into ON and OFF parallel pathways, responding respectively to stimuli of increasing and decreasing luminance. This mechanism can be realized through half-wave rectification, defined as follows:

$$ON(x,y,t) = [RF(x,y,t)]^+, \quad (6)$$

$$OFF(x,y,t) = -[RF(x,y,t)]^-. \quad (7)$$

Here, $[x]^+$ and $[x]^-$ denote $\max(0,x)$ and $\min(x,0)$, respectively. $ON(x,y,t)$ denotes the luminance increase as ON contrast in the ON pathway, while $OFF(x,y,t)$ denotes the luminance decrease as OFF contrast in the OFF pathway. This structure ensures independent processing of opposite polarity information, which is the basis for the crossover inhibition mechanism.

In the primate retina, bipolar cells receive inputs from photoreceptor cells, convert them into neurotransmitter release signals, and transmit these signals to downstream retinal ganglion cells (RGCs) and amacrine cells. To simplify this procedure, the rectified ON and OFF signals are transmitted to the corresponding ON and OFF bipolar cell units within the model. These bipolar cells transmit the input signals and generate excitatory outputs:

$$E_{ON}(x,y,t) = ON(x,y,t), \quad (8)$$

$$E_{OFF}(x,y,t) = OFF(x,y,t). \quad (9)$$

Here, $E_{ON}(x,y,t)$ and $E_{OFF}(x,y,t)$ represent the excitatory outputs of ON and OFF bipolar cells, respectively.

2) Feed-Forward Inhibition within Pathways: Within parallel ON/OFF pathway processing, feed-forward inhibition (FFI) mechanisms are involved within each pathway. Taking the ON pathway as an example, ON amacrine cells receive excitatory input from retinal bipolar cells corresponding to the ON polarity. Subsequently, they provide feedforward inhibitory input to RGCs, thereby forming a local inhibitory loop within the ON pathway. Notably, this FFI loop exhibits a synaptic delay [17]. The OFF pathway undergoes a similar processing mechanism. Consequently, a delayed FFI signal $I(x,y,t)$ is generated within the pathway, mathematically expressed as follows:

$$I_{ON}(x,y,t) = \iint W_I(u,v)E_{ON}(x-u,y-v,t-\Delta t)\,du\,dv, \quad (10)$$

$$I_{OFF}(x,y,t) = \iint W_I(u,v)E_{OFF}(x-u,y-v,t-\Delta t)\,du\,dv, \quad (11)$$

where $I_{ON}(x,y,t)$ and $I_{OFF}(x,y,t)$ represent the FFI of ON and OFF pathway. $[W_I]$ is a spatial convolution kernel,

simulating FFI dominated by neighboring amacrine cells at different distances.

3) Crossover Inhibition between Pathways: Beyond the FFI within each pathway, the proposed CI mediates suppressive interactions between the ON and OFF pathways. As illustrated in Fig. 2, this mechanism simulates the bidirectional antagonistic connection in the retina where ON pathways inhibit OFF pathways via OFF-type amacrine cells, while OFF pathways simultaneously inhibit ON pathways via ON-type amacrine cells, forming a cross-pathway temporal modulation network. The mathematical model for crossover inhibition is defined as

$$CI_{ON}(x,y,t) = E_{OFF}(x,y,t), \tag{12}$$

$$CI_{OFF}(x,y,t) = E_{ON}(x,y,t). \tag{13}$$

where, $CI_{ON}(x,y,t)$ denotes the CI transmitted by OFF amacrine cells, and $CI_{OFF}(x,y,t)$ denotes the crossover inhibition transmitted by ON amacrine cells.

Subsequently, each channel linearly integrates the excitatory and inhibitory signs, that is,

$$\begin{aligned} S_{ON}(x,y,t) = E_{ON}(x,y,t) - \omega_1 I_{ON}(x,y,t) \\ - \beta_1 CI_{ON}(x,y,t), \end{aligned} \tag{14}$$

$$\begin{aligned} S_{OFF}(x,y,t) = E_{OFF}(x,y,t) - \omega_2 I_{OFF}(x,y,t) \\ - \beta_2 CI_{OFF}(x,y,t), \end{aligned} \tag{15}$$

where ω and β represent the coefficients for FFI and CI, used to adjust the intensity of different inhibition, respectively.

C. Integration and RGC Output

To integrate the responses of the ON and OFF channels, we introduce S units (summation Units) as the core processing module. In the S units, the ON and OFF pathway signals are integrated through super-linear summation [18], calculated as follows

$$S(x,y,t) = \theta_1 \cdot S_{ON} + \theta_2 \cdot S_{OFF} + \theta_3 \cdot S_{ON} S_{OFF}, \tag{16}$$

where $\{\theta_1,\theta_2,\theta_3\}$ denotes the coefficient for each term, enabling to achieve a certain balance between different interactions in the ON and OFF channels. Additionally, nonlinear terms are used to enhance the approaching selectivity.

After that, the output of S unit undergoes a simplified grouping operator, denoted as $G(x,y,t)$. Signals within local neighborhoods are weighted and integrated, thereby amplifying spatially continuous edge signals while suppressing isolated noise [19]. The G unit processing is defined as follows:

$$G(x,y,t) = \iint W_g(u,v) S(x-u, y-v, t)\, du\, dv, \tag{17}$$

$$\hat{G}(x,y,t) = \begin{cases} G(x,y,t) & \text{if } G(x,y,t) \geq T_g, \\ 0 & \text{else.} \end{cases} \tag{18}$$

$\left[W_g\right]$ is an equal-weighted kernel, T_g denotes the local activation threshold.

Finally, the model simulates the integration of all local input signals by RGC. It applies a nonlinear sigmoid function to convert the integrated signal into a membrane potential signal, thereby approximating the response characteristics of RGC. That is,

$$RGC(t) = \iint \hat{G}(x,y,t)\, dx\, dy, \tag{19}$$

$$CI\text{-}LMD(t) = \frac{1}{1 + \exp\left(\frac{-RGC(t)}{n_{cell}}\right)}, \tag{20}$$

where $RGC(t)$ is the integrated signal, n_{cell} denotes the total number of retinal cells participating in neural computation, and $CI\text{-}LMD(t)$ is the membrane potential signal as the output.

TABLE I: Parameters of the proposed neural model

Parameter	Description	Value
$\{\sigma_e,\sigma_i\}$	standard deviation in vDoG operator	{2, 4}
Δt	sampling time interval	1(frame)
$\{\omega_1,\omega_2\}$	coefficients for FFI	{0.4, 0.6}
$\{\beta_1,\beta_2\}$	coefficients for CI	{0.25, 0.3}
$\{\theta_1,\theta_2,\theta_3\}$	combination of term coefficients	{1, 0.9, 0.5}
T_g	local activation threshold	10

D. Setting the Parameters

The parameters of the proposed neural model are summarized in Table I. Parameter selection is guided by a balance between biological plausibility and detection accuracy. All adjustable parameters are directly related to the physical characteristics of the input visual stream, particularly the spatial resolution of the images and the temporal sampling rate of the video sequences derived from both synthetic and real-world stimuli. Notably, this study departs from conventional learning-based collision detection approaches by eliminating the need for a training process, instead identifying approaching motion through visual looming cues characterized by expanding edge structures. Unless otherwise stated, all parameters are fixed across experiments to ensure fair and consistent comparisons with related methods.

IV. Experimental Evaluation

In this section, we conduct comprehensive evaluations of the proposed model from both mechanism and application perspectives. Firstly, we illustrate the basic model response to simple scenario and compare the proposed model with representative bio-inspired looming detection models, including the neural model based on dynamic neural field theory [20] (abbreviated as S-DNF), and two insect inspired models with ON/OFF pathways: the LGMD1-based model proposed by Fu et al. [4] (abbreviated as Fu-LGMD1), and the LGMD1-based model which eliminates the response to translating motion proposed by Lei et al. [21] (abbreviated as Lei-LGMD1). Subsequently, we perform ablation experiments to examine the functional role and the effectiveness of the proposed CI mechanism. Finally, we evaluate the model robustness and practical performance using aerial vehicle footage and simple scene video data containing artificial noise.

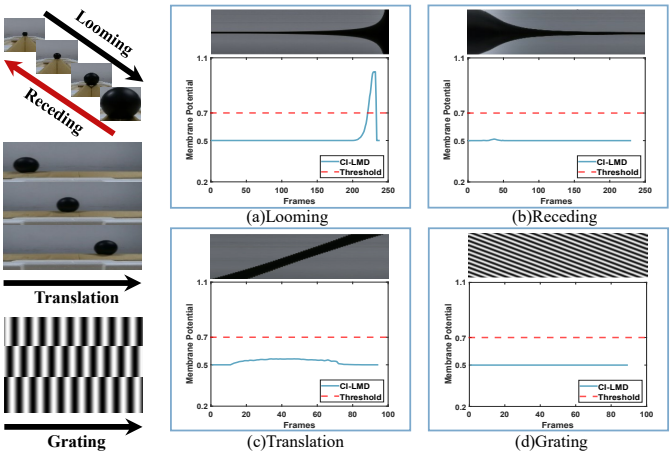


Fig. 3: Response characteristics of the proposed model to various motion stimuli. Snapshots of approaching, receding, and translating ball and grating stimuli in the simple scene are used in the experiment (left panel). Response characteristics of the proposed model in the simple scene (right panel): each sub-figure consists of two rows – the first row illustrates the temporal evolution of the tested visual stimulus, the second row shows the membrane potential output of the proposed model along with a fixed threshold indicator (0.7) used for collision risk alarm and models comparison. If the model output consistently exceeds such threshold over a period of time, it indicates a strong looming motion, signifying a high collision risk.

A. Experimental Setting and Evaluation Metric

This study constructed two datasets for model validation. The simple scenario ball motion dataset includes 79 videos covering three types of stimuli: approaching, receding, and translation motion. All videos have a resolution of 480×720 pixels with a frame rate of 60 fps. The drone scene dataset comprises 138 real-world recorded videos, all with a resolution of 480×854 pixels and a frame rate of 30 fps. Building on the simple scenario described above, we further introduce three types of artificial noise–Gaussian, salt-and-pepper, and Poisson–at mild, moderate, and severe intensity levels to construct independent datasets for testing the model robustness in noisy environments.

To quantitatively assess the performance of comparative models, we employ F1-score as the evaluation metric, calculated as follows:

$$\mathrm{F1} = 2\times\frac{\mathrm{TP}}{2\mathrm{TP}+\mathrm{FN}+\mathrm{FP}}\times 100\%, \tag{21}$$

which basically denotes a harmonic mean of precision and recall. TP (true positive) represents the number of test samples where the model correctly generates collision signals in true approaching datasets. FP (false positive) denotes the number of test samples where the model erroneously generates collision signals in non-approaching datasets. FN (false negative) denotes the number of test samples where no collision signals are generated in true approaching datasets. TN (true negative) denotes the number of test samples where no collision signals are generated in non-approaching datasets.

By using this metric, the overall performance of a model in predicting positive samples can be more reliably evaluated. A higher F1-score indicates more robust and balanced model performance.

B. Basic Characteristics and Comparative Analysis

1) Model Responses and Comparisons: We firstly evaluate the proposed model responses to approaching, receding, and translating motions of a dark object in a bright background, and grating motion, as shown in Fig. 3. The proposed model exhibits a strong selectivity for approaching motion, while responses to both receding and translating motions remain at a low level. The model output consistently exceeds the threshold indicator only during approaching stimuli, whereas it stays below the indicator for non-approaching motions, including receding, translation, and grating stimuli.

Specifically, Table II summarizes the responses of different models against various motion stimuli in the simple scenario. All models exhibited selectivity for approaching stimuli and showed no response to stripe patterns. The Fu-LGMD1, S-DNF, and Lei-LGMD1 models exhibit responses to receding or translating stimuli under certain conditions. In contrast, the proposed CI-LMD model shows no response to receding and translating motions across both dark and light object conditions. These results indicate that the proposed model produces weaker responses under non-approaching motion stimuli compared with the other evaluated models.

TABLE II: Responses of comparative models to simple scenario stimuli

Stimuli	CI-LMD	Fu-LGMD1	S-DNF	Lei-LGMD1
Dark Approaching	✓	✓	✓	✓
Light Approaching	✓	✓	✓	✓
Dark Receding	×	✓	×	✓
Light Receding	×	✓	×	✓
Dark translating	×	✓	✓	×
Light translating	×	✓	✓	×
Dark elongation	×	✓	✓	✓
Light elongation	×	✓	✓	✓
Grating	×	×	×	×

[1] ✓ indicates that the model responds to the stimulus, while × denotes non-response.

2) Mechanism Analysis and Feature Visualization: Subsequently, to further elucidate the mechanistic roles of excitation, FFI, and CI in suppressing translation motion responses, we analyze and visualize the intermediate responses of the proposed model in spatiotemporal feature maps when challenged by a black object translating across a bright background, as illustrated in Fig. 4. In the ON/OFF pathway, translation motion generates symmetric excitation along object contrast change edge, reflecting luminance changes caused by motion rather than true spatial expansion. Although FFI attenuates local excitatory responses to some extent, it is insufficient to suppress the global activation induced by translation.

Interestingly, when CI is incorporated, antagonistic interactions between the ON and OFF pathways effectively inhibit these excitatory signals. As shown in the summation module responses, the incorporation of CI significantly reduces residual excitation that persists after FFI alone. Following

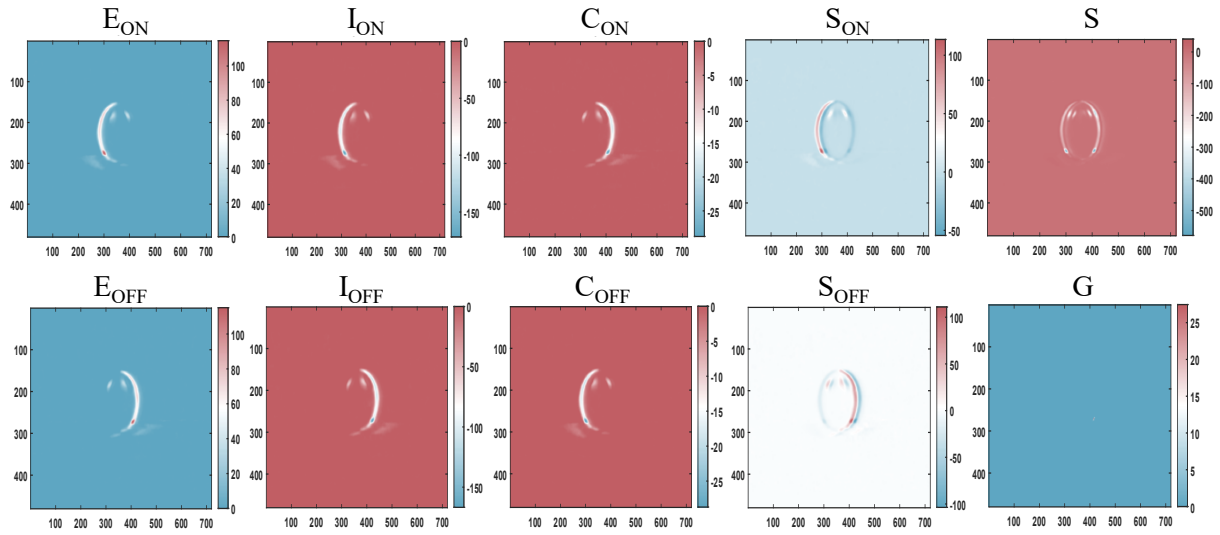


Fig. 4: Mechanistic analysis and visualization of excitation, FFI, and CI under translating motion stimulus. The top row shows responses in the ON pathway, including excitation (E), FFI (I), CI (C), the Summation (S) and Grouping (G) response. The bottom row shows the corresponding responses in the OFF pathway. The stimulus is a black moving object translating across a bright background. Signal intensity is indicated by the color bar where positive values represent excitation, while negative values represent inhibition.

summation and grouping, the excitation is suppressed at all as visualized, preventing false positives of looming detection to translation motion.

3) Ablation Study: We conduct ablation experiments by removing CI connections, while retaining other model components unchanged (referred to as NCI), to investigate the functional basis of CI. As shown in Fig. 5, the results show that in the absence of CI, the NCI model retains its selectivity for looming stimuli but exhibits a marked increase in sensitivity to receding and translation stimuli, generating strong responses to these non-approaching motions. In contrast, the intact CI-LMD model maintains stable response characteristics comparable to the results in Fig. 3.

Notably, this approaching motion selectivity of the proposed model persists consistently under both bright and dark object conditions, indicating that the model may not rely on specific luminance polarity but instead captures spatiotemporal features associated with approaching object. In contrast, when the input consists of translation motion or grating stimuli, the model response consistently remains at a low level. It is owing to the association of FFI and CI successfully suppresses non-approaching motion responses triggered by continuous opposite polarity changes or periodic variations (grating movements).

In summary, experimental results in simple scenarios demonstrate that the CI mechanism further shapes the model's overall response characteristics. Through this regulated cross-pathway inhibition, the CI-LMD model enhances sensitivity to looming stimuli while effectively suppressing irrelevant responses triggered by non-approaching stimuli. These observations provide the basis for further evaluation in more complex and real-world scenarios.

C. Robustness Tests in Challenging Scenarios

In this subsection, we employ real drone video sequences and dynamic scenes with complex noise interference to further evaluate the robustness and practical applicability of the proposed model. Compared to the simpler scenarios used previously, these test scenarios include interference factors such as UAV camera shake and background texture noise. This provides more rigorous validation conditions for verifying the effectiveness of the crossover inhibition mechanism in complex environments.

This category of experiments consists of two main components, both using F1 as the evaluation metric. It compares the proposed model against the CI-ablated NCI model, S-DNF model, Fu-LGMD1 model, and Lei-LGMD1 model in complex aerial vehicle scenarios, and under artificial noise interference.

1) Results under Aerial Vehicle Scenarios: To visually compare performance in real-world scenarios, we evaluated the proposed CI-LMD model against representative existing methods using aerial drone recordings captured in urban and natural environments, as shown in Fig. 6. The experimental results demonstrate that, under complex and dynamic backgrounds, CI-LMD exhibits minimal responses to non-approaching motions such as translation and recession. In contrast, the other models–lacking explicit cross-pathway inhibition mechanisms–produce pronounced erroneous responses under the same conditions. Notably, the Lei-LGMD1 model also shows partial robustness to translating stimuli in aerial recordings, as it employs a divisive interaction between the ON and OFF pathways, which to some extent suppresses responses to translational motion.

When evaluated on datasets comprising both urban and natural scenes recorded by aerial drones, the overall quantitative results are summarized in Table III. In urban environments, despite challenges such as complex building textures and camera-induced motion artifacts, the proposed CI-LMD achieved an F1-score of 80.41%, outperforming other bio-inspired visual models. In natural scenes featuring obstacles such as trees and birds, as well as uneven illumination conditions, CI-

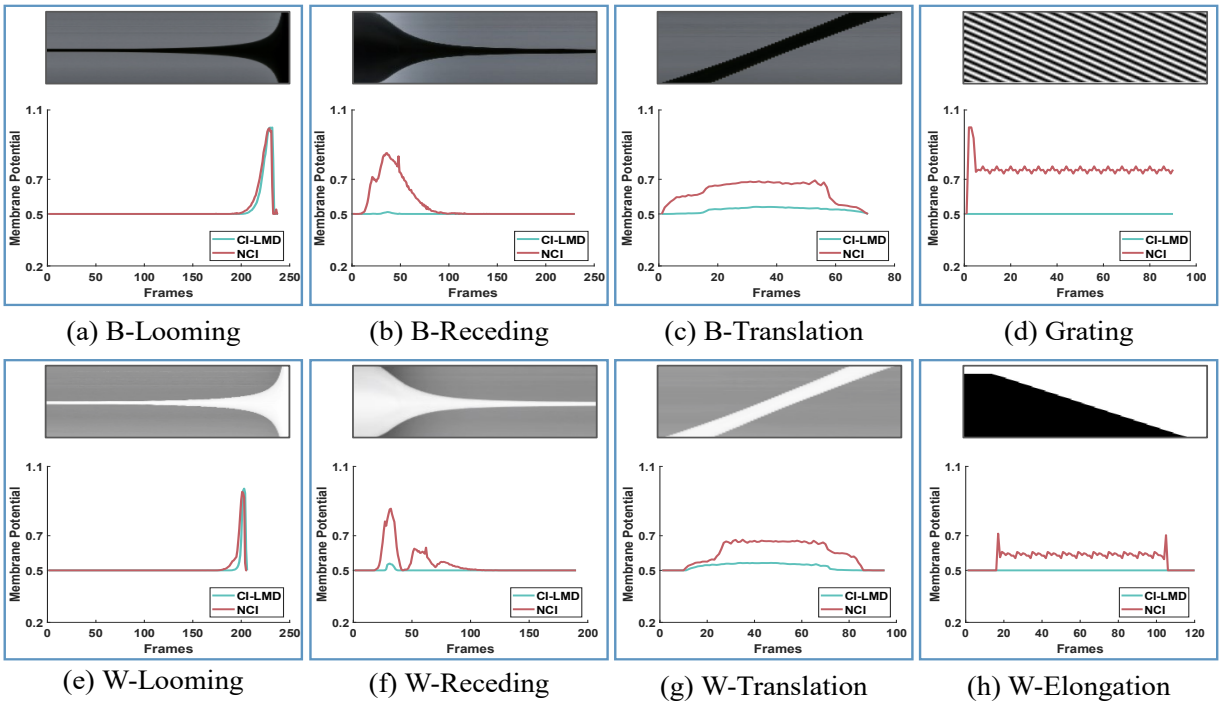


Fig. 5: Ablation study on the role of crossover inhibition in the CI-LMD model. The responses of the intact CI-LMD model and its ablated model (without CI, denoted as NCI) to various stimuli are compared. Stimuli include approaching, receding, and translating motions of a black ball (B-) on a bright background, a white ball (W-) on a dark background, grating motion, and a light bar (W-) elongation. Each sub-figure follows the same format as Fig. 3). The intact model exhibited low response to all non-approaching stimuli, whereas the ablated model showed pronounced responses to non-approaching motions.

LMD maintained relatively stable and reliable detection performance. In contrast, the CI-ablated NCI variant exhibited a pronounced degradation in performance, attributable to the absence of CI. This comparison clearly demonstrates that the CI mechanism plays a critical role in enhancing robustness against diverse and dynamic environmental motions.

TABLE III: F1 score of Various Models to UAV scenes

Stimuli	CI-LMD	NCI	S-DNF	Lei-LGMD1
Urban Scenes	**80.41%**	65.22%	64.29%	70.21%
Natural Scenes	**61.95%**	38.64%	43.90%	50.51%

2) Results under Artificial Noise: As summarized in Table IV, the proposed CI-LMD model consistently demonstrated superior robustness. Specifically, its F1-score remained at 83.72% under mild noise conditions; under moderate noise, the F1 score exhibited only a marginal decrease while remaining above 80.1%; and under severe noise, performance stabilized at approximately 77.05%. Notably, the performance degradation of CI-LMD was substantially smaller than that observed in the other models, including the CI-ablated NCI variant, highlighting the effectiveness of CI in enhancing noise resilience.

TABLE IV: F1 score of Various Models Under Synthetic Noise

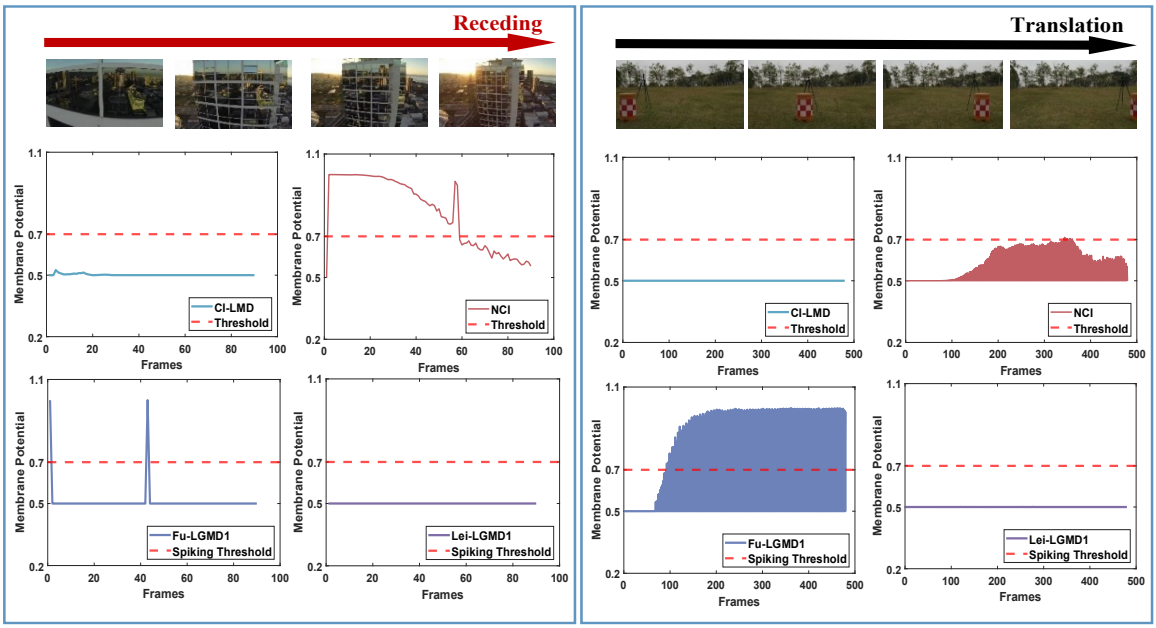
Stimuli	CI-LMD	NCI	S-DNF	Lei-LGMD1
Mild Noise	**83.72%**	55.17%	57.14%	63.49%
Moderate Noise	**80.1%**	45.86%	34.48%	44.44%
Severe Noise	**77.05%**	40.58%	39.51%	35.0%

V. Conclusion and Discussion

This study proposes a brain-inspired visual looming motion detection model. Compared to previous bio-inspired models, the proposed framework no longer treats ON and OFF pathways as independent information processing structures. By incorporating crossover inhibition as the mechanism interacting pathways with opposite polarities, the model achieves improved selectivity to approaching motion. Its effectiveness was validated through experiments across synthetic and real-world scenarios, further demonstrating its robustness against diverse environmental motions.

Systematical experiments with quantitative results indicated that crossover inhibition plays a crucial role in shaping looming motion selectivity. Crossover inhibition effectively suppresses responses triggered by overall luminance changes and translation motion by temporally adjusting ON-to-OFF and OFF-to-ON inhibition. This regulatory mechanism enhances the model sensitivity to rapid, continuous spatial expansion while suppressing its response to other categories of movement patterns. Ablation study further validated the effectiveness of crossover inhibition.

However, the proposed model still exhibits several limitations. Under low-contrast stimulus conditions, the model selectivity is somewhat impacted. Furthermore, the CI mechanism is implemented in a relatively simple way without consideration of its temporal dynamics. Future research will



(a) Urban scene (b) Natural scene

Fig. 6: Responses of various models to UAV scenes. The evaluated scenes are divided into Urban and Natural scenes: (a) UAV receding a building, (b) UAV translating away from a barrier. The proposed CI-LMD model is compared with three representative bio-inspired looming motion detection models, including the ablation model NCI, Lei-LGMD1, and Fu-LGMD1.

conduct parameter sensitivity analyses, optimize the implementation of CI, compare with computer vision and learning-based methods, and validate on robotic platforms to enhance its robustness, practicality, and generalization capabilities in real-world environments.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China under grant nos. 62376063 and 12571558.

REFERENCES

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

[2] Q. Fu, "Motion perception based on on/off channels: A survey," *Neural Networks*, vol. 165, pp. 1–18, 2023.

[3] P. H. Schiller, J. H. Sandell, and J. H. R. Maunsell, "Functions of the on and off channels of the visual system," *Nature*, vol. 322, no. 6082, pp. 824–825, Aug. 1986.

[4] 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, Apr. 2018.

[5] B. Mcguire, J. Stevens, and P. Sterling, "Microcircuitry of bipolar cells in cat retina," *The Journal of neuroscience*, vol. 4, pp. 2920–38, Jan. 1985.

[6] M. B. Manookin, S. S. Patterson, and C. M. Linehan, "Neural mechanisms mediating motion sensitivity in parasol ganglion cells of the primate retina," *Neuron*, vol. 97, no. 6, pp. 1327–1340.e4, Mar. 2018.

[7] J. M. Rosa, S. Ruehle, H. Ding, and L. Lagnado, "Crossover inhibition generates sustained visual responses in the inner retina," *Neuron*, vol. 90, no. 2, pp. 308–319, Apr. 2016.

[8] T. Appleby and M. Manookin, "Selectivity to approaching motion in retinal inputs to the dorsal visual pathway," *eLife*, vol. 9, Feb. 2020.

[9] H. Wässle, I. Schäfer-Trenkler, and T. Voigt, "Analysis of glycinergic inhibitory pathway in the cat retina," *The Journal of neuroscience*, vol. 6, pp. 594–604, Mar. 1986.

[10] J.-J. Pang, F. Gao, and S. M. Wu, "Cross-talk between ON and OFF channels in the salamander retina: Indirect bipolar cell inputs to ON–OFF ganglion cells," *Vision Research*, vol. 47, no. 3, pp. 384–392, Feb. 2007.

[11] M. B. Manookin, D. L. Beaudoin, Z. R. Ernst, L. J. Flagel, and J. B. Demb, "Disinhibition combines with excitation to extend the operating range of the off visual pathway in daylight," *J. Neurosci.*, vol. 28, no. 16, pp. 4136–4150, Apr. 2008.

[12] B. Roska, A. Molnar, and F. S. Werblin, "Parallel processing in retinal ganglion cells: How integration of space-time patterns of excitation and inhibition form the spiking output," *Journal of Neurophysiology*, vol. 95, no. 6, pp. 3810–3822, Jun. 2006.

[13] F. Werblin, "Six different roles for crossover inhibition in the retina: Correcting the nonlinearities of synaptic transmission," *Visual neuroscience*, vol. 27, pp. 1–8, Apr. 2010.

[14] J. Cafaro and F. Rieke, "Regulation of spatial selectivity by crossover inhibition," *The Journal of neuroscience*, vol. 33, pp. 6310–20, Apr. 2013.

[15] A. Molnar, H.-A. Hsueh, B. Roska, and F. S. Werblin, "Crossover inhibition in the retina: Circuitry that compensates for nonlinear rectifying synaptic transmission," *J Comput Neurosci*, vol. 27, no. 3, pp. 569–590, Dec. 2009.

[16] A. Angelucci, M. Bijanzadeh, L. Nurminen, F. Federer, S. Merlin, and P. C. Bressloff, "Circuits and mechanisms for surround modulation in visual cortex," *Annu. Rev. Neurosci.*, vol. 40, no. 1, pp. 425–451, Jul. 2017.

[17] J. S. Diamond, "Inhibitory interneurons in the retina: Types, circuitry, and function," *Annu. Rev. Vis. Sci.*, vol. 3, no. 1, pp. 1–24, Sep. 2017.

[18] J. B. Demb and J. H. Singer, "Functional circuitry of the retina," *Annu. Rev. Vis. Sci.*, vol. 1, no. 1, pp. 263–289, Nov. 2015.

[19] Shigang Yue and F. Rind, "Collision detection in complex dynamic scenes using an lgmd-based visual neural network with feature enhancement," *IEEE Trans. Neural Netw.*, vol. 17, no. 3, pp. 705–716, May 2006.

[20] Z. Qin, Q. Fu, and J. Peng, "A computationally efficient and robust looming perception model based on dynamic neural field," *Neural Networks*, Jun. 2024.

[21] F. Lei, Z. Peng, M. Liu, J. Peng, V. Cutsuridis, and S. Yue, "A robust visual system for looming cue detection against translating motion," *IEEE Transactions on Neural Networks and Learning Systems*, vol. PP, pp. 1–15, Feb. 2022.