



OPEN ACCESS

EDITED BY

Hong Qiao,
University of Chinese Academy of
Sciences, China

REVIEWED BY

Mu-Yen Chen,
National Cheng Kung University,
Taiwan
Jie Yang,
Dalian University of Technology, China
Junyan Hu,
University College London,
United Kingdom
Kefan Wu,
The University of Manchester,
United Kingdom, in collaboration with
reviewer JH

*CORRESPONDENCE

Haiyang Li
fplihaiyang@126.com
Jigen Peng
jgpeng@gzhu.edu.cn

RECEIVED 02 July 2022

ACCEPTED 26 August 2022

PUBLISHED 20 September 2022

CITATION

Ling J, Wang H, Xu M, Chen H, Li H
and Peng J (2022) Mathematical study
of neural feedback roles in small target
motion detection.
Front. Neurobot. 16:984430.
doi: 10.3389/fnbot.2022.984430

COPYRIGHT

© 2022 Ling, Wang, Xu, Chen, Li and
Peng. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Mathematical study of neural feedback roles in small target motion detection

Jun Ling¹, Hongxin Wang^{2,3}, Mingshuo Xu¹, Hao Chen¹,
Haiyang Li^{1*} and Jigen Peng^{1*}

¹School of Mathematics and Information Science, Guangzhou University, Guangzhou, China,

²Machine Life and Intelligence Research Center, Guangzhou University, Guangzhou, China,

³Computational Intelligence Lab (CIL), School of Computer Science, University of Lincoln, Lincoln, United Kingdom

Building an efficient and reliable small target motion detection visual system is challenging for artificial intelligence robotics because a small target only occupies few pixels and hardly displays visual features in images. Biological visual systems that have evolved over millions of years could be ideal templates for designing artificial visual systems. Insects benefit from a class of specialized neurons, called small target motion detectors (STMDs), which endow them with an excellent ability to detect small moving targets against a cluttered dynamic environment. Some bio-inspired models featured in feed-forward information processing architectures have been proposed to imitate the functions of the STMD neurons. However, feedback, a crucial mechanism for visual system regulation, has not been investigated deeply in the STMD-based neural circuits and its roles in small target motion detection remain unclear. In this paper, we propose a time-delay feedback STMD model for small target motion detection in complex backgrounds. The main contributions of this study are as follows. First, a feedback pathway is designed by transmitting information from output-layer neurons to lower-layer interneurons in the STMD pathway and the role of the feedback is analyzed from the view of mathematical analysis. Second, to estimate the feedback constant, the existence and uniqueness of solutions for nonlinear dynamical systems formed by feedback loop are analyzed *via* Schauder's fixed point theorem and contraction mapping theorem. Finally, an iterative algorithm is designed to solve the nonlinear problem and the performance of the proposed model is tested by experiments. Experimental results demonstrate that the feedback is able to weaken background false positives while maintaining a minor effect on small targets. It outperforms existing STMD-based models regarding the accuracy of fast-moving small target detection in visual clutter. The proposed feedback approach could inspire the relevant modeling of robust motion perception robotics visual systems.

KEYWORDS

visual system modeling, small target motion detection, feedback mechanism, existence of solutions, fixed point iteration

1. Introduction

Small target motion detection is one of the most important problems in computer vision, and it has been widely applied in underwater robot vision (Xu et al., 2018), security monitoring (Escobar-Alvarez et al., 2019), unmanned driving (Li et al., 2017), military interception (Bosquet et al., 2018), etc. However, a small target occupies few pixels in an image so that it hardly displays physical features. Moreover, the complex dynamic environment always contains a great number of small-target-like features (such as leaves, flowers, and shadows), and there often exists ego-motion during sampling *via* camera. These phenomena mentioned above will bring great difficulties to small target motion detection.

The traditional computer vision methods for objects detection, including background subtraction (Saleemi and Shah, 2013), temporal differencing (Shuigen et al., 2009; Javed et al., 2018), and optical flow (Fortun et al., 2015). These conventional methods achieve sufficiently good performance in detecting large objects (such as pedestrians, cars, and vehicles) in static background. However, their detection performance to detect a small target with the dynamic background will decrease significantly due to the small target always concealing in background clutter and hardly displaying visual features. In addition, some machine learning methods, such as convolutional neural networks (Redmon et al., 2016) and support vector machines (Tang et al., 2017) can be used for object detection. These machine learning methods perform well in detecting objects with high resolution, clear appearance, and structure from which the discriminative features can be learned. However, they may fail to detect small objects with only one or a few pixels in size, since rich representations are difficult to learn from their poor-quality appearance and structure.

In nature, insects, such as dragonfly, hoverfly, and drosophila, display exquisite sensitivity to small target motion and are able to pursue potential mates and small prey with success rates greater than 97% in complex dynamic environments (Mischiati et al., 2015). Electrophysiological experiments have identified a class of special neurons in the brain of insects, called Small Target Motion Detectors (STMDs) (Nordström and O'Carroll, 2006; Nordstrom and O'Carroll, 2009; Barnett et al., 2007; Keleş and Frye, 2017), which make insects sensitive to small target motion. More precisely, the STMDs respond strongly to small moving targets subtending $1^\circ \sim 3^\circ$ of the visual field, but weakly to larger objects subtending more than 10° (Nordström et al., 2006) or wide-field grating stimuli. In addition, STMD neurons also respond robustly to small targets moving in the complex dynamic background. These distinctive functions of STMDs provide reliable support for the design of small target motion detection models.

Inspired by the insect's vision system, some models have been proposed to imitate the functions of STMDs. For example, as pioneering work, a computational model, called Elementary STMD (ESTMD), was designed to implement the size selectivity of the STMDs and detect the moving small target (Wiederman et al., 2008). The ESTMD displays strong responses to the presence of small moving targets, while weak or no responses to a large moving object. In order to estimate motion directions, the Elementary Motion Detector designed by Hassenstein and Reichardt (Hassenstein and Reichardt, 1956) was incorporated into ESTMD (Wiedermann and O'Carroll, 2013a,b), generating two new models, named EMD-ESTMD and ESTMD-EMD, respectively. Recently, Wang et al. developed a directionally selective STMD (DSTMD) (Wang et al., 2018), which makes use of correlation mechanism of two locations to detect positions and motion directions of small moving targets. On the base of DSTMD, Wang et al. (2019) exploited a direction contrast pathway to filter out fake features, where the resulted model is called STMD-plus. These models mentioned above process information in a feed-forward manner to detect small target motion. Despite the success of these feed-forward models in small target motion detection, the detection performances of these models in the complex dynamic background are unsatisfying and their detection results contain a number of background false positives.

The feedback mechanism plays an important role in modulating visual stimuli in animals' visual systems (Lamme et al., 1998; Bastos et al., 2015; Clarke and Maler, 2017), and it can potentiate the abilities of visual neurons to detect the motion of targets in a complex dynamic environment (Klink et al., 2017; Mohsenzadeh et al., 2018; Nurminen et al., 2018). Further biological research studies have also revealed feedback loops in insect visual systems. For instance, insects used closed-loop control in their brains to distinguish whether or not the changes in the environment are caused by their own behavior (Paulk et al., 2015). Insects selectively responded to salient visual stimuli *via* a close-loop in the brain (Paulk et al., 2014), which guided behavioral choices made by these insects. The feedback connection was discovered in the binocular stereopsis of praying mantis (Rosner et al., 2019), which calculated the distances from disparities between the two retinal images *via* feedback to trigger a raptorial strike of their forelegs when prey is within reach. In recent years, feedback mechanism has been proven to effectively improve model performance in many studies, including medical image segmentation (Soker, 2016) and object recognition (Wang and Huang, 2015). However, their connection patterns and functional roles in small target motion detection pathway still remain unclear. Based on the feedback mechanism, Wang et al. (2018) developed a feedback STMD model (Feedback STMD) for small target motion detection. The Feedback STMD was modeled by transmitting the output layer information to medulla layer neurons to filter out fake features. Although the Feedback STMD

performed well in detecting the small target, its detection results contain a number of background false positives in complex dynamic backgrounds. In addition, the feedback model only verified the effectiveness *via* experiments, and the feedback constant was chosen by empiric rule, lacking mathematical proof analysis.

In this paper, we proposed a new time-delay feedback model, called \mathcal{F} STMD, for detecting small target motion in complex dynamic backgrounds. The \mathcal{F} STMD forms a nonlinear dynamic system by propagating the output of higher-layer neurons to the lower-layer neurons for weakening background fake features. The main contributions of this paper can be summarized as follows:

- 1). We develop a time-delay feedback STMD model for small target motion detection by transmitting the STMD neurons outputs to the lamina neurons to weaken the background false positives.
- 2). The functional role of the feedback is revealed by comparing the outputs of with and without feedback to different velocities.
- 3). To estimate the strength of the feedback, we prove the existence and uniqueness of the solutions to the nonlinear dynamic system *via* Schauder's fixed point theorem and contraction mapping theorem.
- 4). We design an iterative algorithm to find the approximate solution of a nonlinear system and verify the effectiveness of the algorithm *via* experiments. The results of our experiments demonstrate that our proposed model is unable to improve the detection performance in detecting small target motions on a complex background.

The article is structured as follows. In Section 2, some related works are introduced. In Section 3, the \mathcal{F} STMD model and the working mechanism of feedback are introduced in detail. In Section 4, the existence and uniqueness of the solutions to the nonlinear dynamic system are proved and an iterative algorithm is proposed. In Section 5, experimental results demonstrate the advantage of the proposed algorithm over other STMD-based algorithms in detecting motion of small targets. Finally, some conclusions are given in Section 6 and the proofs of the theorems are provided in [Appendix](#).

2. Related work

In this section, some related works are introduced, mainly including some STMD-based models, feedback mechanism, and infrared small target detection.

2.1. The STMD-based models

Small target motion detectors are a special kind of motion-sensitive neurons, which respond strongly to small target motion even in complex dynamic backgrounds. Motivated by the superior properties of the STMD neurons, some STMD-based models have been developed for small target motion detection. For instance, as a pioneer, [Wiederman et al. \(2008\)](#) first proposed an Elementary Small Target Motion Detector (ESTMD) which well matches to the size selectivity of the STMDs and can detect the presence of small moving targets. However, the ESTMD model is not directionally selective and cannot estimate the direction of motion. In order to model the directional selectivity of STMDs, [Wiedermann and O'Carroll \(2013b\)](#) developed two mixed models, including EMD-ESTMD and ESTMD-EMD, which are designed by the Elementary Motion Detector ([Hassenstein and Reichardt, 1956](#)) combined with ESTMD ([Wiedermann and O'Carroll, 2013a,b](#)). However, the motion direction is only divided into four directions, i.e., right/up and left/down. Recently, [Wang et al. \(2018\)](#) developed a directionally selective STMD (DSTMD), which makes use of the correlation mechanism of two locations to detect positions and estimate motion directions of small moving targets systematically. On the base of DSTMD, [Wang et al. \(2019\)](#) exploited a direction contrast pathway to filter out most of the fake features. These models mentioned above all process visual information in a feedforward manner. However, feedback is a common regulatory mechanism in biology and has been not investigated in the STMD pathway.

2.2. Feedback mechanism

Feedback is a fundamental mechanism which regulates visual signals in the biological visual system. It refers to the process of returning the output of the system to the input and changing the input in some way and affecting the system function. The feedback mechanism has been applied in the field of artificial intelligence to achieve higher performance. For example, [Carreira et al. \(2016\)](#) proposed a framework to pose estimation. It enhances the expressive ability of the hierarchical feature extractors *via* top-down feedback and shows excellent performance in articulated pose estimation. [Cao et al. \(Cao et al., 2018\)](#) developed a Feedback CNN model, which could implement the selectivity mechanism of neuron activation and localize and segment the interested objects accurately in images. [Zhang et al. \(2018\)](#) introduced multi-path recurrent feedback to enhance salient target detection. It enhanced effective feature learning by introducing multi-path recurrent feedback to transfer global semantic information from the top-level convolution layer to the shallower layer and performed favorably against the state-of-the-art approaches. In addition, feedback has also been used extensively in the

nonlinear dynamic system to pursue the stability problem of nonlinear systems and design feedback controller over the past decade (Liu and Tong, 2015; Sahoo et al., 2015; Brunton et al., 2016). Although feedback mechanisms have achieved great success in many fields, the feedback connection mode and function to STMD neural pathways still remain unclear.

2.3. Infrared small target detection

There are many conventional computer vision methods for objects detection, including background subtraction (Saleemi and Shah, 2013), temporal differencing (Shuigen et al., 2009; Javed et al., 2018), and optical flow (Fortun et al., 2015). Most of conventional computer vision methods achieve sufficiently good performance in detecting large objects, such as pedestrians, cars, and vehicles. However, these methods are powerless to detect small targets. The reason is that the small targets hardly show features such as shape, color, and structure. In addition, the current small target movement detection mainly focuses on infrared images. For example, an infrared image patch-image model was proposed by formulating an optimization problem of recovering low-rank, which achieved superior performance for different target sizes and signal-to-clutter ratio values (Gao et al., 2013). Bai et al. proposed the derivative entropy-based contrast measure for small target detection under various contrast measure for small target detection under various complex background clutters. It applied the derivative entropy-based contrast measure to enhance the infrared small target detection and suppress background clutter (Bai and Bi, 2018). Deng et al. applied a special ring Top-Hat transformation to suppress the complex background and developed a novel local entropy for capturing local features and target enhancement to enhance the infrared small target detection (Deng et al., 2021). These methods have performed well to detect small targets on infrared images, which are heavily dependent on temperature differences between the background and small targets. However, they are impotent for small target motion detection on natural cluttered backgrounds.

3. Formulation of the model and the working mechanism of feedback

The proposed model is composed of four neural layers and a feedback pathway (see Figure 1). The four neural layers are retina, lamina, medulla, and lobula. The four neural layers contain a number of specialized visual neurons coordinated together to detect small target motion. Specifically, visual information is captured by the ommatidia. The large monopolar cells (LMCs) receive the ommatidia output and feedback signal to calculate the change of brightness over time. The outputs of LMCs are further processed by Tm1 and Tm3 neurons in parallel. Finally, STMDs integrate the outputs of

Tm1 and Tm3 neurons to detect the motion of small targets. In the following, the proposed model will be introduced in detail.

3.1. Retina

The retinal layer contains a number of ommatidia (Sanes and Zipursky, 2010; Borst and Helmstaedter, 2015), as shown in Figure 1. Each ommatidium receives visual stimuli from a small region of visual field (Warrant, 2017; Meglič et al., 2019). In \mathcal{F} STMD, the ommatidium is modeled by a spatial Gaussian filter (see Figure 2) to smooth the luminance signal of each pixel. To be more precise, let $I(x, y, t) \in R$ denote the brightness value captured by each ommatidium, where x, y , and t are spatial and temporal field positions. Then, the output of an ommatidium $P(x, y, t)$ is described by

$$P(x, y, t) = [I(\cdot, \cdot, t) * G_{\sigma_1}](x, y), \quad (1)$$

where $*$ denotes convolution; $G_{\sigma_1}(x, y)$ is a Gaussian function, defined as

$$G_{\sigma_1}(x, y) = \frac{1}{2\pi\sigma_1^2} e^{-\frac{(x^2+y^2)}{2\sigma_1^2}}, \quad (2)$$

where σ_1 is the standard deviation of the Gaussian function.

3.2. Feedback pathway

As shown in Figure 2, the LMC input signal is defined by adding the time-delay feedback signal with the ommatidium output, that is

$$P_{TF}(x, y, t) = P(x, y, t) + D_F(x, y, t), \quad (3)$$

where $D_F(x, y, t)$ denotes the time-delay feedback signal. It is defined by convolving the STMD output $D(x, y, t)$ with a Gamma kernel $\Gamma_{n_4, \tau_4}(t)$, that is

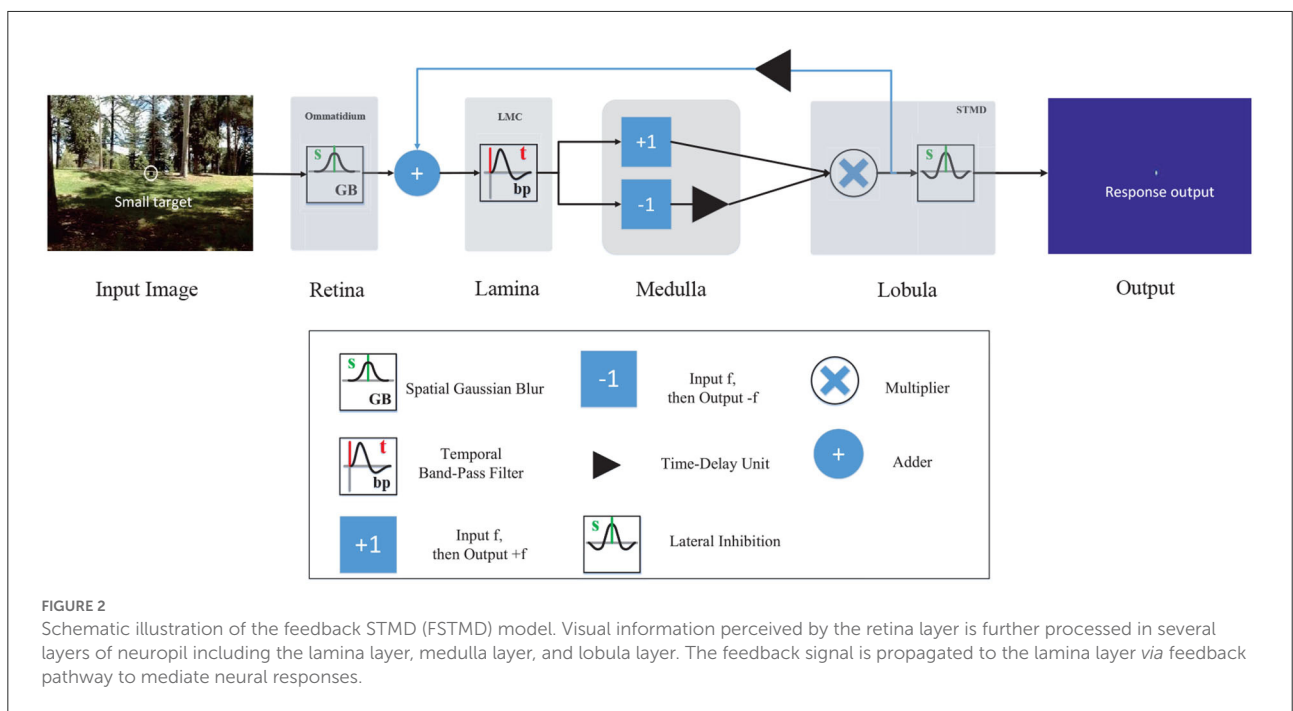
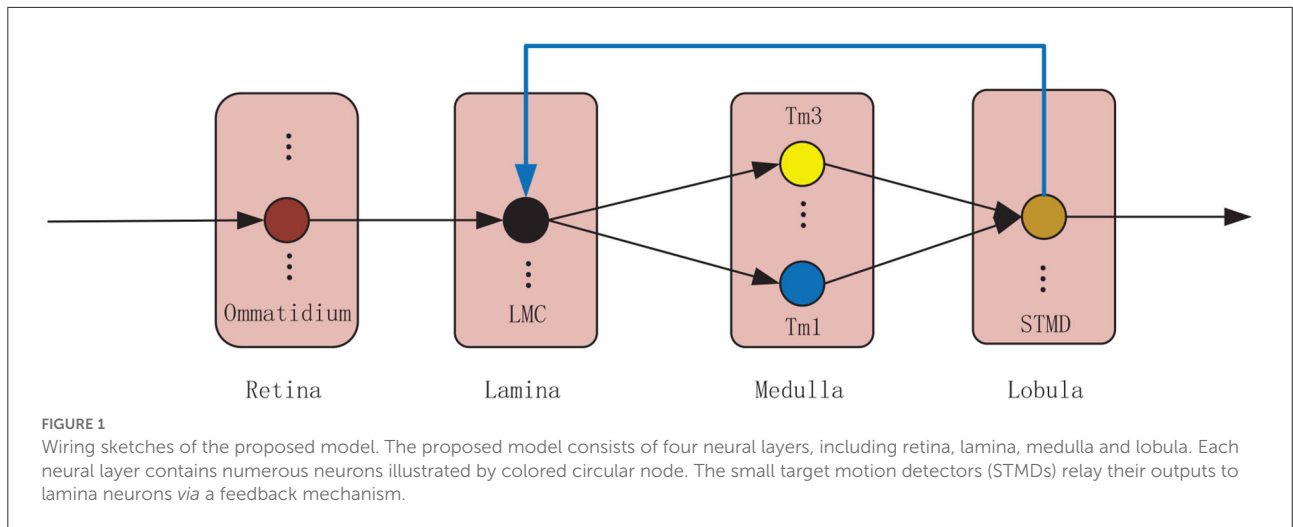
$$D_F(x, y, t) = a(D(x, y, \cdot) * \Gamma_{n_4, \tau_4})(t), \quad (4)$$

$$\Gamma_{n, \tau}(t) = (nt)^n \frac{e^{-\frac{nt}{\tau}}}{(n-1)! \cdot \tau^{n+1}}, \quad (5)$$

where a ($0 < |a| < 1$) is the feedback constant; n_4 and τ_4 are the order and time constants of the Gamma kernel (De Vries and Principe, 1991), respectively.

3.3. Lamina

As shown in Figure 1, the lamina layer consists of large monopolar cells (LMCs) (Borst, 2009), which are postsynaptic



neurons of the ommatidia and are sensitive to changes in brightness (Tuthill et al., 2013; Clark and Demb, 2016). They show strong responses to increase and decrease of brightness (Freifeld et al., 2013; Behnia et al., 2014; Li et al., 2017). In the proposed vision model, LMC is simulated by a time domain band-pass filter to extract brightness changes from the input signal. Mathematically, the LMC output $L(x, y, t)$ is calculated by convolving the output of ommatidium $P_{TF}(x, y, t)$ with a kernel $H(t)$, that is

$$L(x, y, t) = (P_{TF}(x, y, \cdot) * H)(t), \tag{6}$$

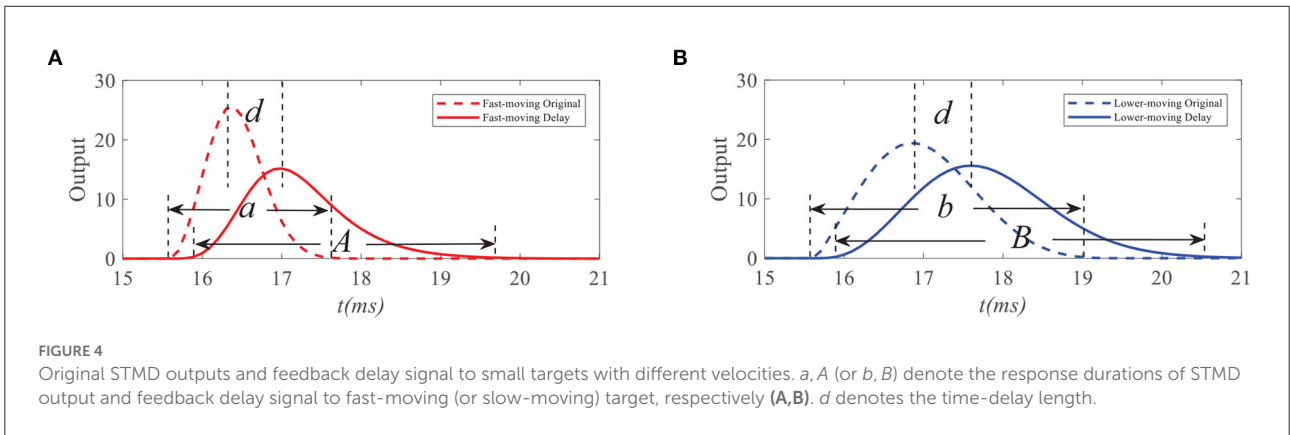
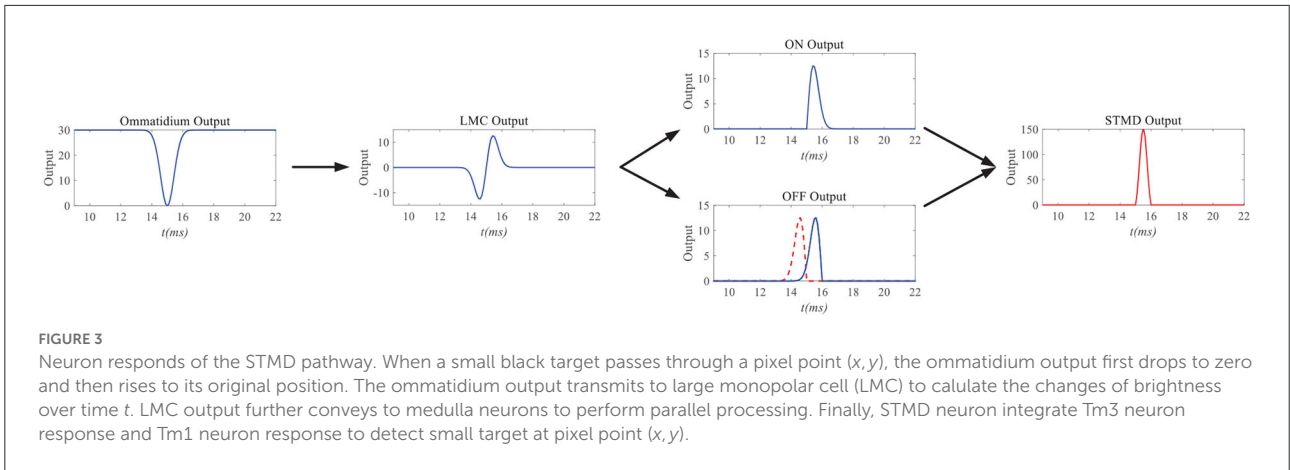
where $H(t)$ is defined by

$$H(t) = \Gamma_{n_1, \tau_1}(t) - \Gamma_{n_2, \tau_2}(t), \tag{7}$$

where n and τ denote the order and time constants of the Gamma kernel.

3.4. Medulla

The medulla neurons, including Tm1 and Tm3 (Takemura et al., 2013; Fu et al., 2019), as illustrated in Figure 1. Tm3 neuron responds to brightness increase (Joesch et al., 2010; Clark et al.,



2011); on the contrary, the Tm1 neuron responds to the decrease of brightness, and the response of the Tm1 neuron is relative to the Tm3 neuron with time delay at the same spatial positions. In the \mathcal{F} STMD, Tm1 and Tm3 neurons are simulated by half-wave rectifiers. The output of Tm3 neuron $S_{ON}(x, y, t)$ is defined by the positive part of the output of LMC $L(x, y, t)$, that is

$$S_{ON}(x, y, t) = [L(x, y, t)]^+ \tag{8}$$

Meanwhile, the output of Tm1 neuron $S_{OFF}(x, y, t)$ is defined by convolving the negative part of the output of LMC $L(x, y, t)$ with a kernel $\Gamma_{n_3, \tau_3}(t)$, that is

$$S_{OFF}(x, y, t) = ([L(x, y, \cdot)]^- * \Gamma_{n_3, \tau_3})(t) \tag{9}$$

3.5. Lobula

As can be seen from Figure 1, the lobula layer contains plenty of STMDs which integrate the signal from medulla neurons including Tm3 and Tm1 (Geurten et al., 2007; Clark et al., 2011). In the proposed visual system, the output of STMD neuron is defined by multiplying the output of Tm3 neuron $S_{ON}(x, y, t)$

with the output of Tm1 neuron $S_{OFF}(x, y, t)$ to detect small target motion, that is

$$D(x, y, t) = S_{ON}(x, y, t) \times S_{OFF}(x, y, t) \tag{10}$$

Figure 3 simulates the neuron responses to the STMD pathway. When a small black target passes through a pixel point (x, y) , the ommatidium output first drops to zero and then rises to its original position. It is worth noting that the drop and rise of ommatidium output are caused by the arrival and departure of the small target at pixel point (x, y) . The ommatidium output transmits to LMC to calculate the changes of brightness over time t . LMC output further conveys to medulla neurons to perform parallel processing. Tm3 neuron responds to brightness increase; on the contrary, the Tm1 neuron responds to the decrease in brightness, and the response of the Tm1 neuron is relative to the Tm3 neuron with time delay at the same spatial positions. The time-delay length is defined as the ratio of the small target width to its velocity v . Finally, the STMD neuron integrates the Tm3 neuron response and Tm1 neuron response to detect a small target at pixel point (x, y) .

To demonstrate the working mechanism of feedback, we first analyze the STMD outputs and feedback signals to the small

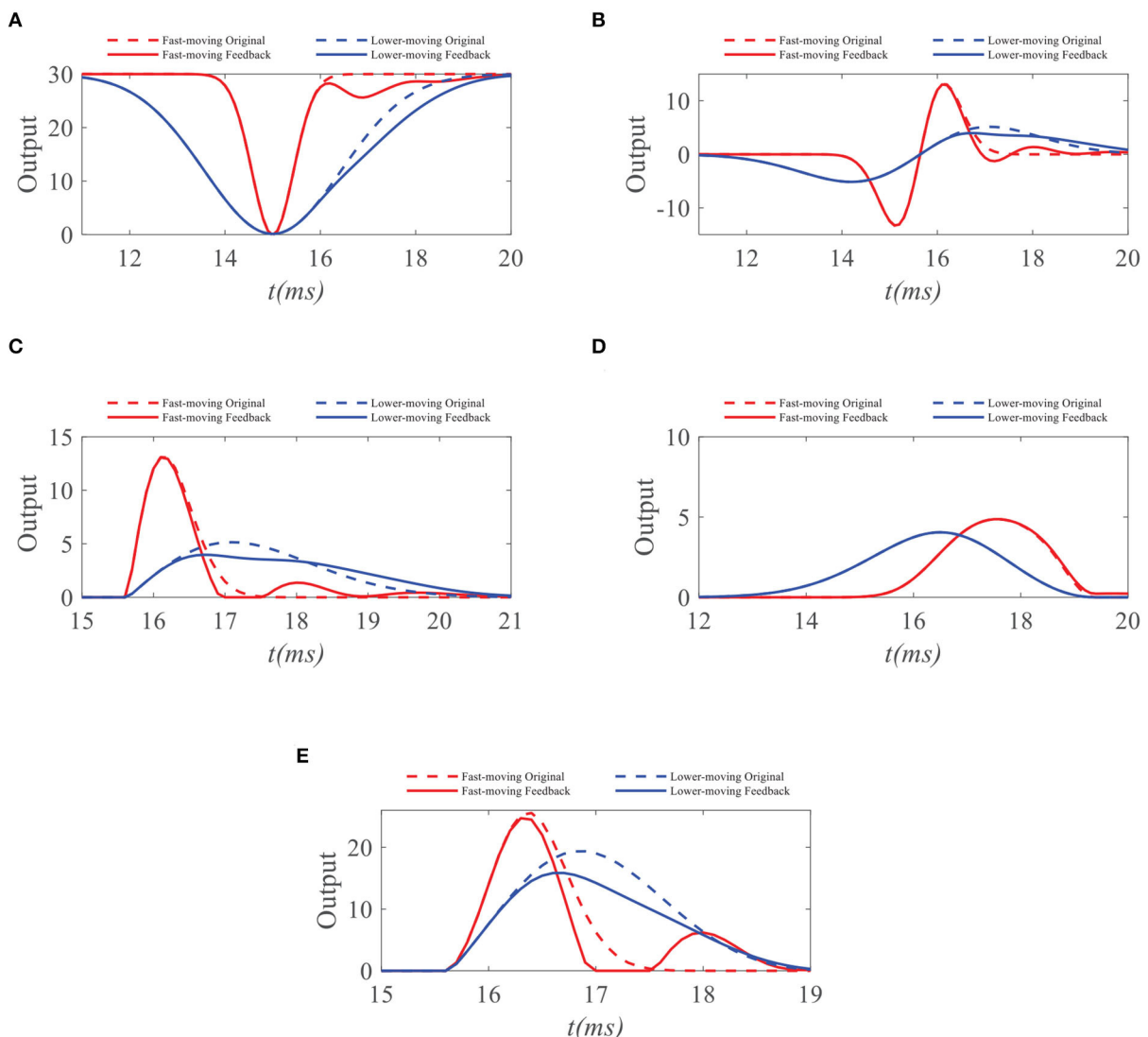


FIGURE 5

(A) Original ommatidia outputs and feedback ommatidia outputs to small targets with different velocities. The feedback signal weakens the ommatidium output to the slow-moving target in the second half of its response duration. In contrast, the ommatidium output to the fast-moving target receives minor inhibition from the feedback signal. Meanwhile, feedback signals have no effect on small targets with different velocities in the first half of their response durations, since STMD neuron has not yet responded when the small target enters the pixel point. (B) Original LMC outputs and feedback LMC outputs to small targets with different target velocities. The peak value of LMC output for fast-moving target maintains unchanged and the peak value of LMC output for slow-moving target has declines. (C) The original ON outputs and feedback ON outputs to small targets with different target velocities. The peak value of ON output to lower-moving target reduces, but the peak values of ON outputs to fast-moving target hardly change. (D) The original OFF outputs and feedback OFF outputs to small targets with different target velocities. (E) The peak values of OFF outputs to target with different velocities hardly change. The original STMD outputs and feedback STMD outputs to small targets with different target velocities. The STMD output reduction with a fast-moving target is significantly lower than a lower-moving target. Notice that there is a small perturbation in the STMD output of a fast-moving target. However, the peak value of small perturbation is lower than the maximum value of STMD output. Therefore, the small perturbation will be filtered by the region maximization operation.

target with different velocities. We simulate response outputs of four neural layers and feedback signals to small targets with different velocities, as shown in Figures 4, 5. Figures 4A,B shows the response durations of the STMD and feedback signals to small targets with different velocities. The response duration of

STMD is defined by the time of the object completely covering a pixel point. When the target velocity is v , the response duration of STMD is controlled by $1/v$. A longer response duration means it takes longer for the object to cover the pixel, which means the target velocity is slower. We express this relationship as $a =$

$f(1/v)$, where f is an increasing function. The feedback signal is a time-delayed form of STMD output, where the response duration and time-delay length are controlled by the parameters n_4 and τ_4 , respectively. Notice that the response duration of the feedback signal is larger than that of the STMD output (i.e., $A > a$) because of the convolutional operation.

We further compare the neural outputs of each neural layer to small targets with different velocities after negative feedback. As can be seen from Figure 5A, the feedback signal weakens the ommatidium output to a slow-moving target in the second half of its response duration. In contrast, the ommatidium output to the fast-moving target does not receive any inhibition from the feedback signal in the second half of its response duration, although a little disturbance appears outside the response duration. Meanwhile, feedback signals have no effect on small targets with different velocities in the first half of their response durations. The reason for this is that the STMD neuron has not yet responded when the small target enters the pixel point. In Figure 5B, we provide the LMC outputs to small targets with different velocities. We can see that the peak value of LMC output to slow-moving target has a decline, but maintains unchanged to fast-moving target. Figures 5C,D displays the medulla neurons the outputs to small targets with different velocities. Since the ON outputs are defined by the positive part of LMC outputs, the maximum value of ON output to a slow-moving target is suppressed by the feedback signal. In comparison, the maximum values of OFF outputs with and without feedback remain identical, because the feedback signal has a minor effect on the negative part of LMC outputs. Finally, we compare the STMD neural outputs with and without feedback in Figure 5E. As can be seen, the maximum value of the STMD output for a slow-moving target significantly decreases after feedback, whereas that of the STMD output for a fast-moving target hardly changes. Although there is a small perturbation in the STMD output to fast-moving target, the peak value of the perturbation is lower than the maximum value of STMD output and can be filtered out by the max-pool operation.

Comparing the STMD outputs to small targets with different velocities, we can find that if the time-delay length d is larger than half of the response duration of the feedback signal (i.e., $d > A/2$), the STMD output will maintain their maxima after subtracting the feedback signal. Thus, we obtain

$$d > \frac{A}{2} > \frac{a}{2} = \frac{f(1/v)}{2}, \quad (11)$$

$$v > \frac{1}{f^{-1}(2d)}, \quad (12)$$

It means that when the time-delay length d is fixed, the feedback signal has a small effect to targets with velocity $v > 1/f^{-1}(2d)$; meanwhile, if velocity $v < 1/f^{-1}(2d)$, the STMD output will be significantly reduced by feedback signal.

As shown in Figure 2, \mathcal{F} STMD implements a lateral inhibition mechanism on $D(x, y, t)$ for size selectivity, that is

$$D_w(x, y, t) = [(D(\cdot, \cdot, t) * W)(x, y)]^+, \quad (13)$$

$$W(x, y) = A[g(x, y)]^+ + B[g(x, y)]^-, \quad (14)$$

$$g(x, y) = G_{\sigma_2}(x, y) - eG_{\sigma_3}(x, y) - \rho, \quad (15)$$

where $W(x, y)$ represents the lateral inhibition kernel; $[x]^+$ and $[x]^-$ denote $\max(x, 0)$ and $\min(x, 0)$, respectively; A, B, e, ρ are constants.

Finally, we give a detection threshold θ and compare it with the model output $D_w(x, y, t)$. If the output $D_w(x, y, t)$ is greater than the threshold θ , then we believe that a small target is detected at position (x, y) and time t .

4. Existence and uniqueness of solutions and algorithm

We propose the time-delay feedback vision system for small target motion detection. It designs a feedback pathway by transmitting information from output-layer neurons to lower-layer interneurons, forming a mathematical nonlinear dynamical system in infinite dimensional space. In order to estimate the range of the feedback constant a and propose the convergent algorithm, we need to prove the existence and uniqueness of the solutions to the nonlinear dynamical system. In the following, the existence and uniqueness of solutions to the nonlinear dynamic system will be analyzed by applying Schauder's fixed point theorem and contraction mapping theorem. We first introduce some basic facts to be used in this paper.

Let $L^p(\Omega)$ be the space of real valued measurable functions u such that $|u(t)|^p$ is Lebesgue integrable and the corresponding norm is given by

$$\|u\|_{L^p(\Omega)} = \begin{cases} \left(\int_{\Omega} |u(t)|^p dt \right)^{\frac{1}{p}}, & \text{for } 1 \leq p < \infty, \\ \text{ess sup}_{t \in \Omega} |u(t)|, & \text{for } p = \infty. \end{cases}$$

Definition 1. (Completely Continuous Operator Li et al., 2006) Let X be a Banach space and let $T: X \rightarrow X$ be an operator. T is said to be compact if it maps bounded sets of X into a sequentially compact set. Moreover, T is said to be completely continuous, if it is continuous and compact.

Theorem 1. (Kolmogorov-Riesz Hanche-Olsen and Holden, 2010) A subset F of $L^2([0, T], \mathbb{R}_+)$ is a sequentially compact set if, and only if,

(i) F is bounded,

(ii) for every $\epsilon > 0$ there is some $\delta > 0$ so that, for every $f \in F$ and h with $0 < h < \delta$,

$$\int_{[0, T]} |f(t+h) - f(t)|^2 dt < \epsilon. \quad (16)$$

Next, we recall two fixed point theorems, which help us analyze the existence and uniqueness of the solutions to the equation.

Theorem 2. (Banach Fixed Point Theorem [Sousa et al., 2019](#)) Let X be a Banach space and $T : X \rightarrow X$ is a contraction mapping. Then, T has a unique fixed point in X .

Theorem 3. (Schauder's Point Theorem [Khan et al., 2011](#)) Let K be a closed convex subset of a Banach space X . If $T : K \rightarrow K$ is continuous and $T(K)$ is relatively compact, then T has a fixed point in K .

The feedback system can be formulated as $D(x, y, t) = F_3(F_2(F_1(P(x, y, t) + D(x, y, t)))) = F(P(x, y, t) + D(x, y, t))$, where $P(x, y, t)$ is the retina output signal, $D(x, y, t)$ is the output signal, and F denotes the composition of three functions F_1, F_2 , and F_3 , where F_1, F_2 , and F_3 represent the system response functions of three neural layers lamina, medulla, and lobula, respectively. For the sake of convenience, we define the operator F on $L^2(R^2 \times [0, T], R_+)$ by

$$(FD)(x, y, t) = [(I * G(x, y) * H)(t) + a(D(x, y, t) * \Gamma_4 * H)(t)]^+ \times \{[(I * G(x, y) * H) + aD(x, y, t) * \Gamma_4 * H]^- * \Gamma_3(t)\}, \quad (17)$$

where

$$\Gamma_{n,\tau}(t) = \begin{cases} (nt)^n \frac{e^{-\frac{t}{\tau}}}{(n-1)! \cdot \tau^{n+1}}, & \text{for } t \geq 0, \\ 0, & \text{for } t < 0. \end{cases} \quad (18)$$

Using $f^+ = \frac{(|f|+f)}{2}$ and $f^- = \frac{(|f|-f)}{2}$, the Equation (17) is equivalent to

$$(FD)(t) = [(P * H)(t) + a(D * V)(t)]^- * \Gamma_3(t) \times [(P * H)(t) + a(D * V)(t)]^+ = \frac{1}{2} \{[(P * H)(t) + a(D * V)(t)] * \Gamma_3(t) - [(P * S)(t) + a(D * K)(t)]\} \times \frac{1}{2} \{[(P * H)(t) + a(D * V)(t)] + (P * H)(t) + a(D * V)(t)\}, \quad (19)$$

where $H(t) = \Gamma_1(t) - \Gamma_2(t)$, $V(t) = (\Gamma_4 * H)(t)$, $K(t) = (\Gamma_3 * V)(t)$, $S(t) = (\Gamma_3 * H)(t)$, $P(x, y, t) = (I(\cdot, \cdot, t) * G)(x, y)$ and $\Gamma(t), H(t), V(t), K(t), S(t) \in L^p(R)(1 \leq p \leq \infty)$.

Before investigating the main result in this paper, we list essential conditions:

$$(H_1): P(t) \in L^2([0, T]).$$

Based on (H_1) , we can derive the following theorem. Theorem 4 illustrates that the operator F defined on L^2 is well-defined.

Theorem 4. Suppose that (H_1) holds, then $F(L^2([0, T], R_+)) \subset L^2([0, T], R_+)$.

Proof. See the **Proof of Theorem 4** in the [Appendix](#).

The following theorem illustrates that operator F is continuous, which is very important for the proof of existence and uniqueness of the solution to the equation.

Theorem 5. Assume that (H_1) holds, then operator $F : L^2([0, T], R_+) \rightarrow L^2([0, T], R_+)$ is continuous.

Proof. See the **Proof of Theorem 5** in the [Appendix](#).

From the previous discussion, it is clear that the operator F is continuous. To get the existence of the solution to the equation, we also need to illustrate that the operator F is a compact operator.

Theorem 6. Suppose that (H_1) holds, then $F : L^2([0, T], R_+) \rightarrow L^2([0, T], R_+)$ is a compact operator.

Proof. See the **Proof of Theorem 6** in the [Appendix](#).

According to Theorem 5 and Theorem 6, we immediately obtain that the operator F is a completely continuous operator. In the following theorems, the existence and uniqueness of solution are verified by applying Schauder's fixed point theorem and contraction mapping theorem. We first present the existence of solutions.

Theorem 7. Suppose that (H_1) holds and $a < 0$, then the following results hold:

(i). If $Q^2 - 4N'\varphi = 0$ and $aQ + 1 > 0$, then $a \in [-\frac{1}{2Q}, 0)$ and the Equation (19) has at least one solution on $B_r = \{D \in L^2([0, T], R_+), \|D\|_{L^2} \leq r\}$, where

$$\frac{(aQ + 1) - \sqrt{2Qa + 1}}{2a^2N'} \leq r \leq \frac{(aQ + 1) + \sqrt{2Qa + 1}}{2a^2N'}. \quad (20)$$

(ii). If $Q^2 - 4N'\varphi > 0$ and $aQ + 1 > 0$, then $a \in (-\frac{1}{Q}, \frac{-1}{Q+2\sqrt{N'\varphi}}]$ and the Equation (19) has at least one solution on $B_r = \{D \in L^2([0, T], R_+), \|D\|_{L^2} \leq r\}$, where

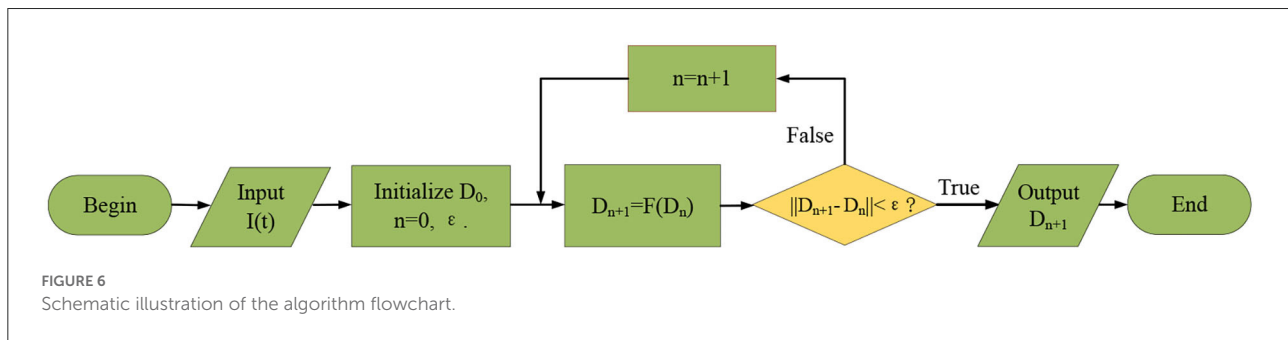
$$\frac{(aQ + 1) - \sqrt{(aQ + 1)^2 - 4a^2N'\varphi}}{2a^2N'} \leq r \leq \frac{(aQ + 1) + \sqrt{(aQ + 1)^2 - 4a^2N'\varphi}}{2a^2N'}. \quad (21)$$

(iii). If $Q^2 - 4N'\varphi < 0$ and $aQ + 1 > 0$, then $a \in [\frac{-1}{Q+2\sqrt{N'\varphi}}, 0)$ and the Equation (19) has at least one solution $D \in B_r = \{D \in L^2([0, T], R_+), \|D\|_{L^2} \leq r\}$, where

$$\frac{(aQ + 1) - \sqrt{(aQ + 1)^2 - 4a^2N'\varphi}}{2a^2N'} \leq r \leq \frac{(aQ + 1) + \sqrt{(aQ + 1)^2 - 4a^2N'\varphi}}{2a^2N'}. \quad (22)$$

Input: Input $I(t), n_1, \tau_1, n_2, \tau_2, n_3, \tau_3, n_4, \tau_4, \sigma_1$;
 Initial value of iteration: $D_0 = D_E$ (without feedback output).
 1: Calculate the $P(t)$ by (1);
 2: **for** $n = 0, 1, 2, \dots$ **do**
 3: Calculate $D_{n+1} = F(D_n)$ by (19);
 4: **end for** when some stopping criterion is met.

Algorithm 1. \mathcal{F} STMD Iterative algorithm.



Proof. See the **Proof of Theorem 7** in the [Appendix](#).

The following theorem gives the unique results of the solution to the nonlinear Equation (19).

Theorem 8. Assume that (H_1) holds, then the following results hold:

(i). If $Q^2 - 4N'\varphi = 0$, $a \in \left[-\frac{1}{2Q}, 0\right)$ and $r = \frac{(aQ+1) - \sqrt{2Qa+1}}{2a^2N'}$, then there is unique one solution to the Equation (19) on $L^2([0, T], R_+)$.

(ii). If $Q^2 - 4N'\varphi < 0$, $a \in \left[\frac{-1}{Q+2\sqrt{N'\varphi}}, 0\right)$ and $r = \frac{(aQ+1) - \sqrt{(aQ+1)^2 - 4a^2N'\varphi}}{2a^2N'}$, then there is unique one solution to the Equation (19) on $L^2([0, T], R_+)$.

Proof. See the **Proof of Theorem 8** in the [Appendix](#).

In the following, we design a fixed point iteration algorithm to solve the nonlinear dynamic system (19), where the pseudocodes of our algorithm are listed in [Algorithm 1](#), and the flowchart of the algorithm is shown in [Figure 6](#). In the iterative process, when the error between $D_{n+1}(x, y, t)$ and $D_n(x, y, t)$ is lower than the preset error threshold 10^{-3} , we stop the iteration and assume that $D_{n+1}(x, y, t)$ is model response output.

Feedback Constant Estimation: From the theorem 8, we prove the nonlinear system has unique solution, which means the proposed algorithm to solve dynamic systems is stable. To determine the feedback constant a , we first estimate Q , N' , and φ by Theorem 7, and then substitute them into $a \in \left[\frac{-1}{Q+2\sqrt{N'\varphi}}, 0\right)$ revealed in the condition (ii) of Theorem 8.

5. Experiment

In this section, we demonstrate the effectiveness of the proposed algorithm for detecting small target's motion against cluttered backgrounds. We evaluate the proposed \mathcal{F} STMD on the Vision Egg data set ([Straw, 2008](#)) and the RIST data set RIST Data Set¹. The Vision Egg data set includes a number of synthetic image sequences, each of which displays a small target moving against a natural background image (see [Figures 7A, 11A](#)). Each synthetic video contains one or multiple small target motions, whose resolution and sampling frequency equate to 300×250 pixels and 1,000 Hz, respectively. The RIST data set includes 16 videos, which were captured in real environments by using an action camera (GoPro Hero 6) at 240 fps. Each video contains an object that is moving in cluttered scenarios. All the experiments are performed under Windows 10 and MATLAB (R2017a) running on a computer with a Core i5 CPU at 3.10GHz with 16GB of memory. The parameters of the \mathcal{F} STMD model are shown in [Table 1](#). The parameters of model consist of two parts, including the parameters of the four feed-forward neural layers and the feedback pathway. The parameters of the four feed-forward neural layers and their effect on the performance have been investigated and analyzed in the reference ([Wang et al., 2018](#)). In this paper, we choose the parameters to make the model satisfy the basic properties of STMD neurons such as size selectivity and velocity selectivity. The parameters of the feedback pathway n_4 and τ_4 can be tuned so that the appropriate optimal velocity and preferred velocity range of the model can be selected.

¹ <https://sites.google.com/view/hongxinwang-personalsite/download/> (accessed on April 6, 2020).

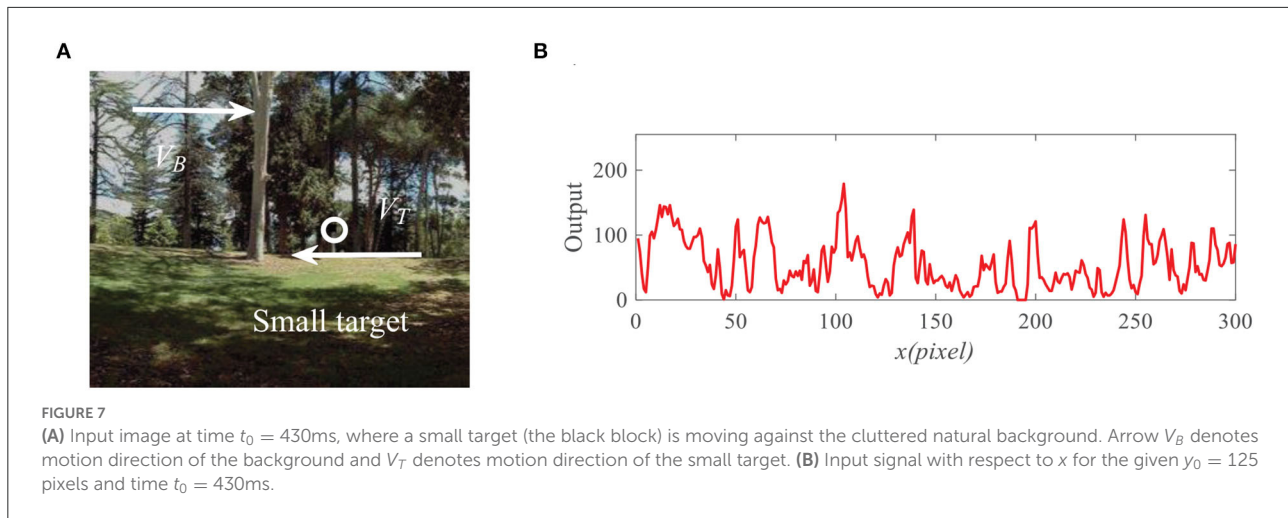


TABLE 1 Parameters of the feedback small target motion detectors (FSTMDs).

Equations	Parameters
(1)	$\sigma_1 = 1$
(6)	$n_1 = 2, \tau_1 = 3, n_2 = 6, \tau_2 = 9$
(9)	$n_3 = 5, \tau_3 = 25$
(4)	$n_4 = 5, \tau_4 = 10$
(13)	$\sigma_2 = 1.5, \sigma_3 = 3, e = 1, \rho = 0$

To demonstrate the effectiveness of the proposed model, we present the neural outputs with and without feedback. **Figure 7A** shows the input image at time $t_0 = 430$ ms, where a small black target (250 pixels/s) is moving against a cluttered background (150 pixels/s). We fix $y_0 = 125$ pixels then display the input luminance signal $I(x, y_0, t_0)$ with respect to x at time t_0 in **Figure 7B**. It can be seen from **Figure 7B** that the input signal is quite complex and small target is submerged in background clutters. As shown in **Figure 8A**, the outputs of ommatidia are a smooth version of the input luminance signal *via* Gaussian blur. **Figure 8B** shows the feedback ommatidia results which are calculated by subtracting the feedback signal from the original ommatidia outputs. From **Figure 8B**, we can find that the ommatidia outputs with feedback decrease after experiencing negative feedback at $x = 41, 54, 198,$ and 248 . The outputs of LMCs in **Figures 8C,D** are given to show luminance change at time t_0 at each pixel. In particular, a positive value of LMC output indicates luminance increases, while a negative value of LMC output represents a decrease in luminance at time t_0 .

We further compare the outputs of the medulla neurons. **Figure 9** displays the outputs of ON and OFF with and without feedback. We can see that the ON outputs decrease *via* feedback, while the OFF outputs hardly alter at $x = 41, 54, 198,$ and

248. This is because that the feedback suppresses the positive part of LMC output to targets but has minor effect on the negative part of LMC outputs. In addition, the reduction of the ON output $R = 0.9$ at $x = 198$ is smaller than that of background false targets locations 1.9, 2.3, and 2.6. The outputs of STMDs are defined by multiplying the outputs of Tm3 neurons and Tm1 neurons and implementing a lateral inhibition. From **Figure 10A**, the outputs of STMDs with feedback are significantly lower than the original outputs of STMDs. In **Figure 10B**, we provide the inhibition percentage (IP) of small target positions and background false target locations. It can be seen from **Figure 10B** that the IP of the real target location is generally much smaller than that of background fake target locations. In particular, the IP of the fast-moving small target at position $x = 198$ is 2.7%, and the IPs of the slow-moving background false targets at position $x = 41, 54,$ and 248 are 79.1, 37.5, and 51.8% respectively. The above results demonstrate that feedback significantly suppresses slow-moving background features.

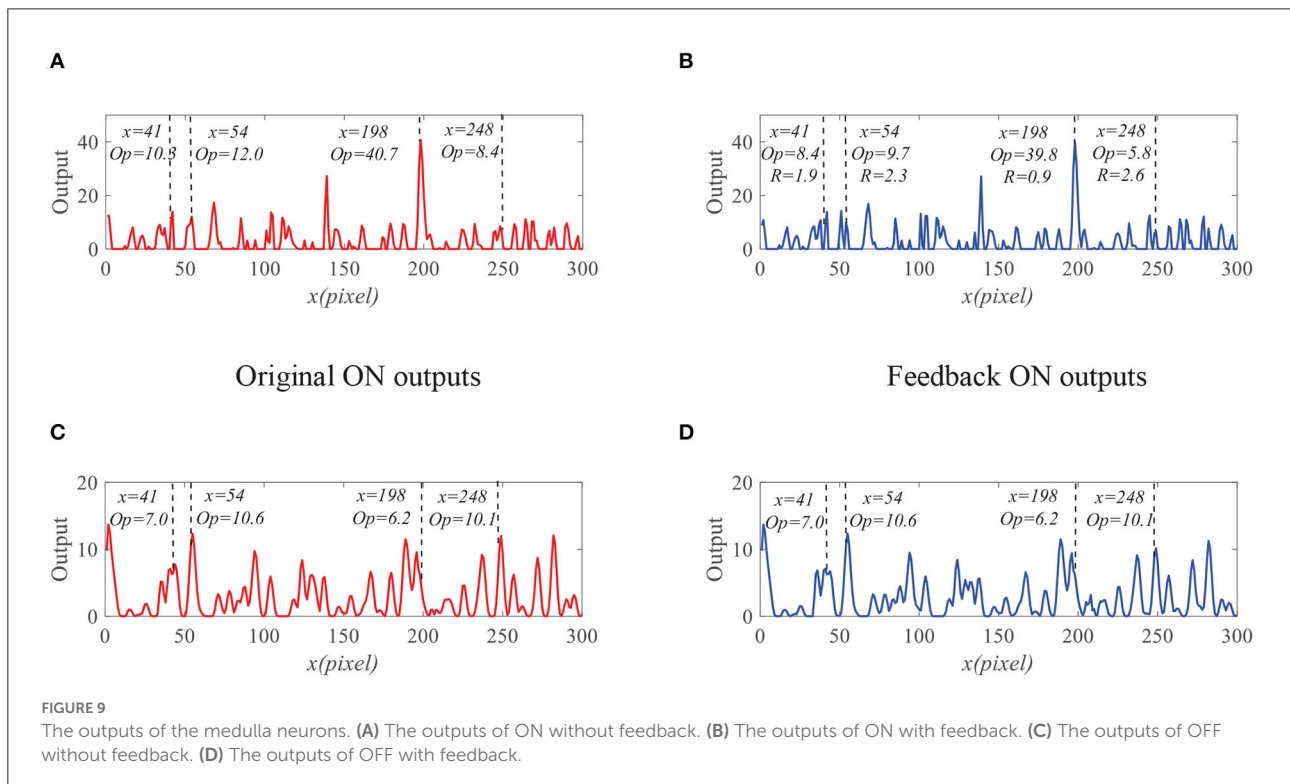
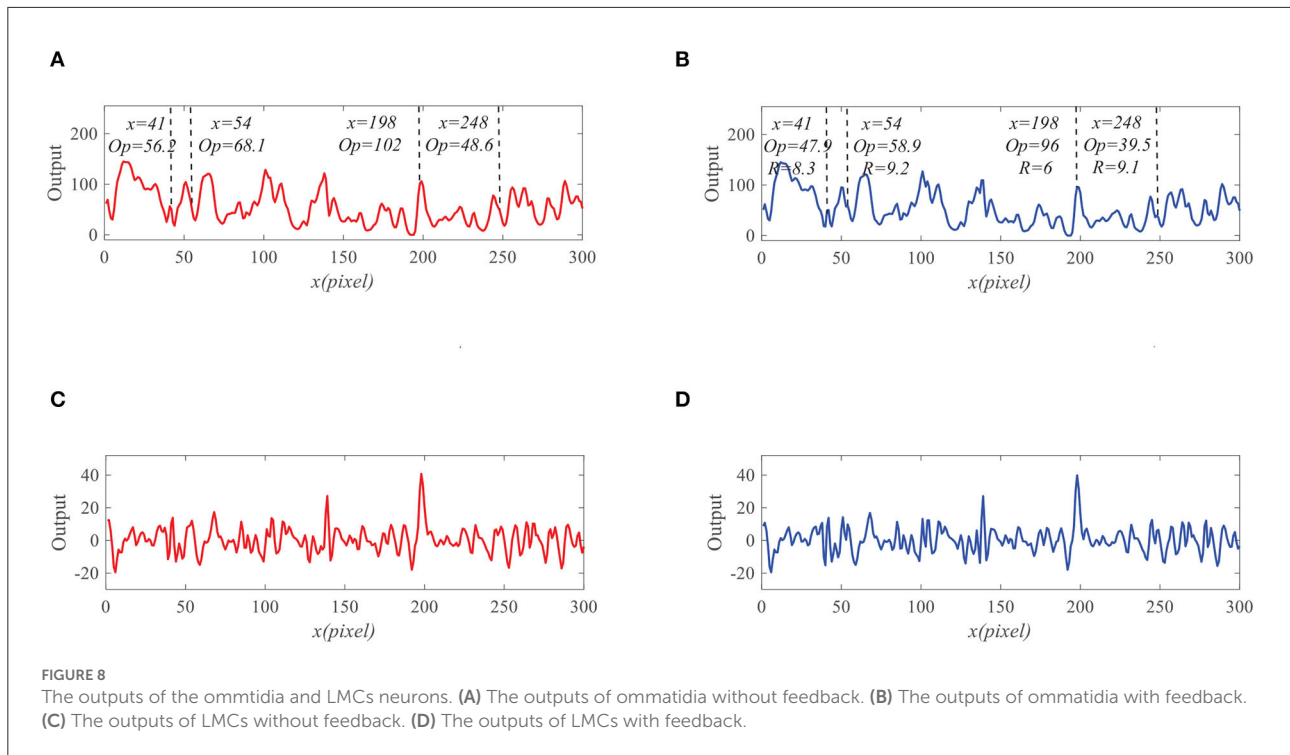
To quantitatively compare with the existing STMD-based model, we define two metrics (Wang et al., 2018),

$$D_R = \frac{\text{number of true detections}}{\text{number of actual targets}}, \quad (23)$$

$$F_A = \frac{\text{number of false detections}}{\text{number of images}}, \quad (24)$$

where D_R denotes detection rate and F_A represents false alarm rate. If the pixel distance between the actual position of the target and the detection result is within a threshold (5 pixels), then we consider the detection result is correct.

We compare the performance of the proposed FSTMD with ESTMD (Wiederman et al., 2008) and DSTMD (Wang et al., 2018) models in terms of different target velocities, target sizes, target luminance, and background velocities. The details of the synthetic image sequences are listed in **Table 2**. **Figure 11B** shows



the $D_R - F_A$ curves of three models for the initial synthetic image sequence. It can be seen that the feedback model has higher detection rates compared with ESTMD and DSTMD for

any false alarm rates. **Figure 12A** provides the detection result curves of the proposed model for different feedback constant a . These $D_R - F_A$ curves demonstrate that with the decrease

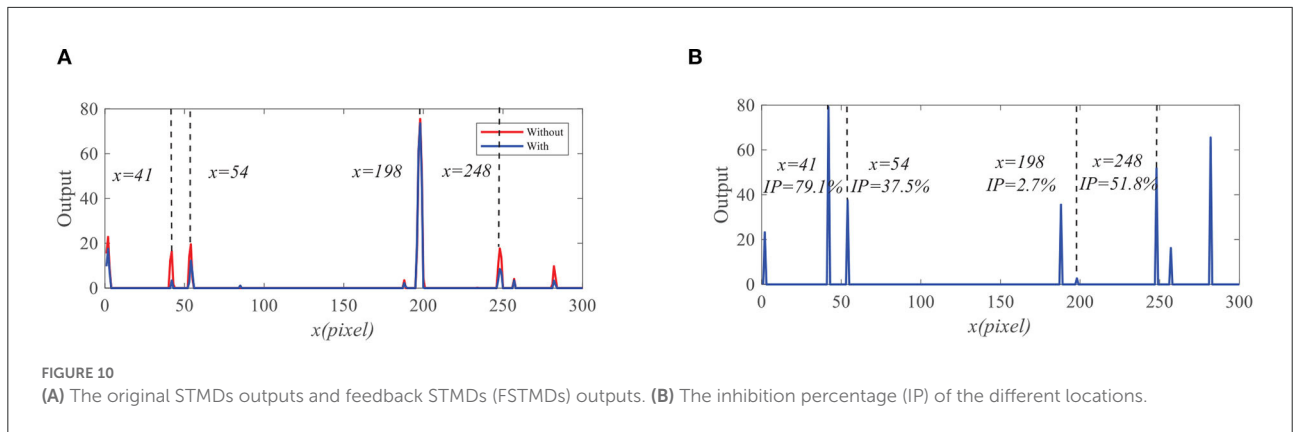


FIGURE 10 (A) The original STMDs outputs and feedback STMDs (FSTMDs) outputs. (B) The inhibition percentage (IP) of the different locations.

TABLE 2 Details of the synthetic image sequences.

Image parameter	Initial sequence	Test 1	Test 2	Test 3	Test 4	Test 5
Target size (pixel × pixel)	5 × 5	1 × 1 ~ 11 × 11	5 × 5	5 × 5	5 × 5	5 × 5
Target luminance	0	0	0 ~ 0.1	0	0	0
Target velocity (pixel/s)	250	250	250	100 ~ 330	250	250
Background velocity (pixel/s)	150	150	150	150	50 ~ 180	50 ~ 180
Background motion direction	rightward	rightward	rightward	rightward	rightward	leftward
Background image	Figure 11A	Figure 11A	Figure 11A	Figure 11A	Figure 11A	Figure 11A

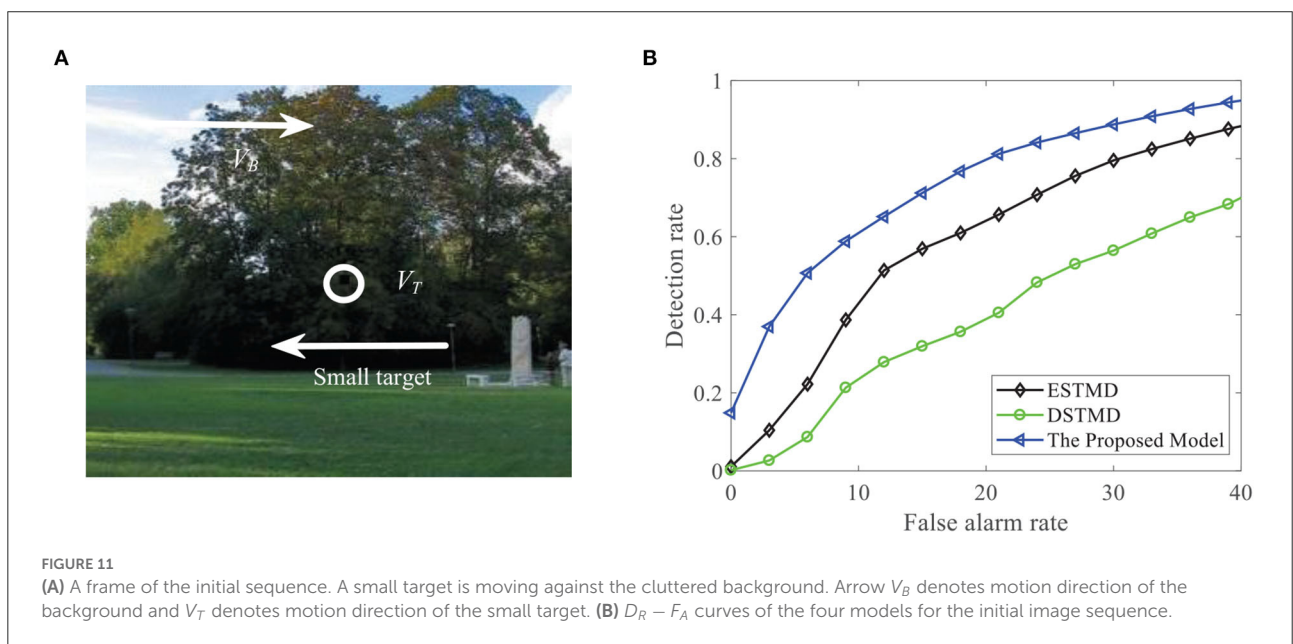
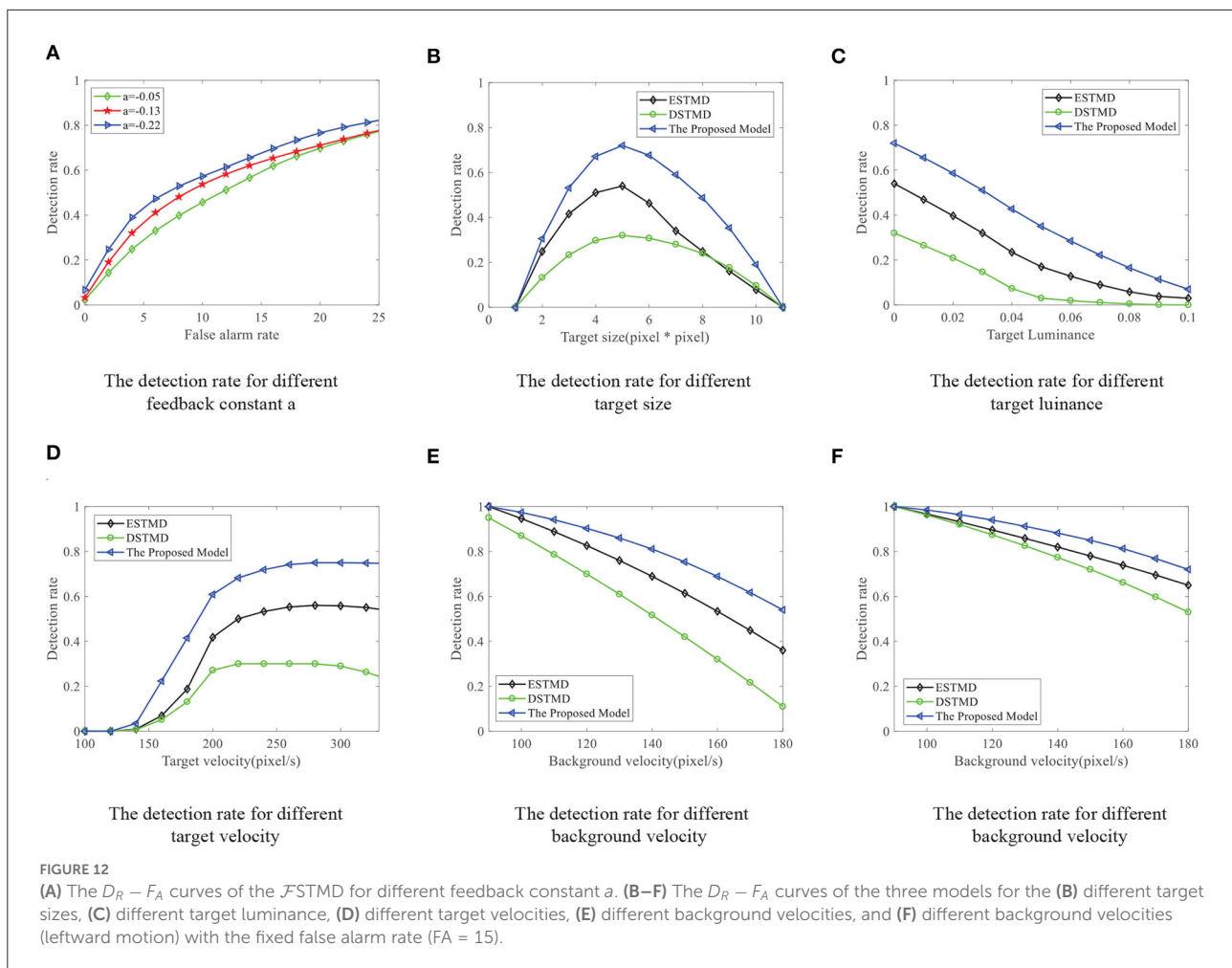


FIGURE 11 (A) A frame of the initial sequence. A small target is moving against the cluttered background. Arrow V_B denotes motion direction of the background and V_T denotes motion direction of the small target. (B) $D_R - F_A$ curves of the four models for the initial image sequence.

of feedback constant a , the detection performance of \mathcal{F} STMD model becomes better. The reason for this is that with the decrease of feedback constant a , the feedback signal enhances inhibition to slow-moving background false targets but has a small effect on fast-moving small targets. Figures 12B–F displays the detection rates of the three models concerning target size,

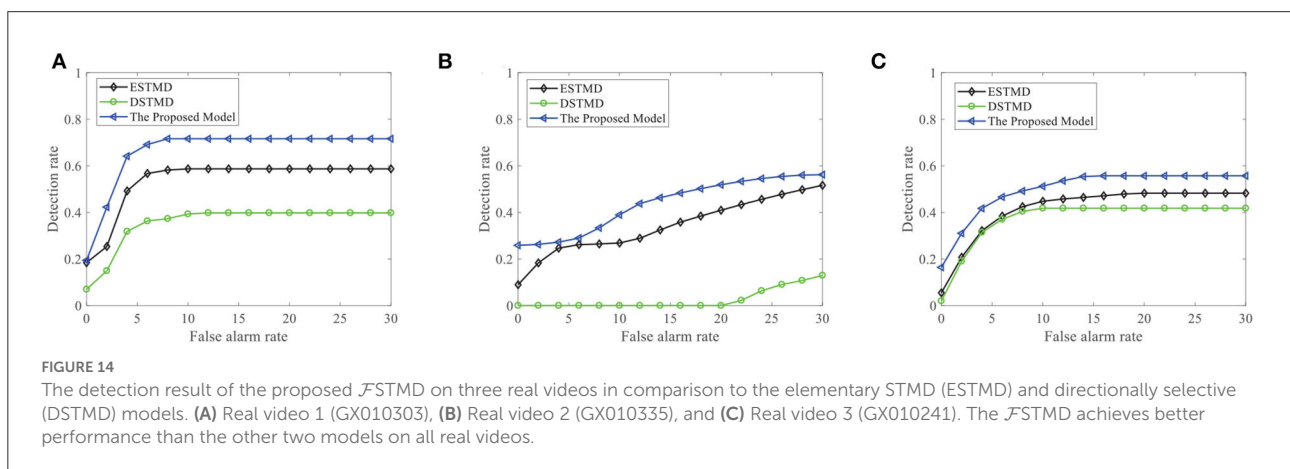
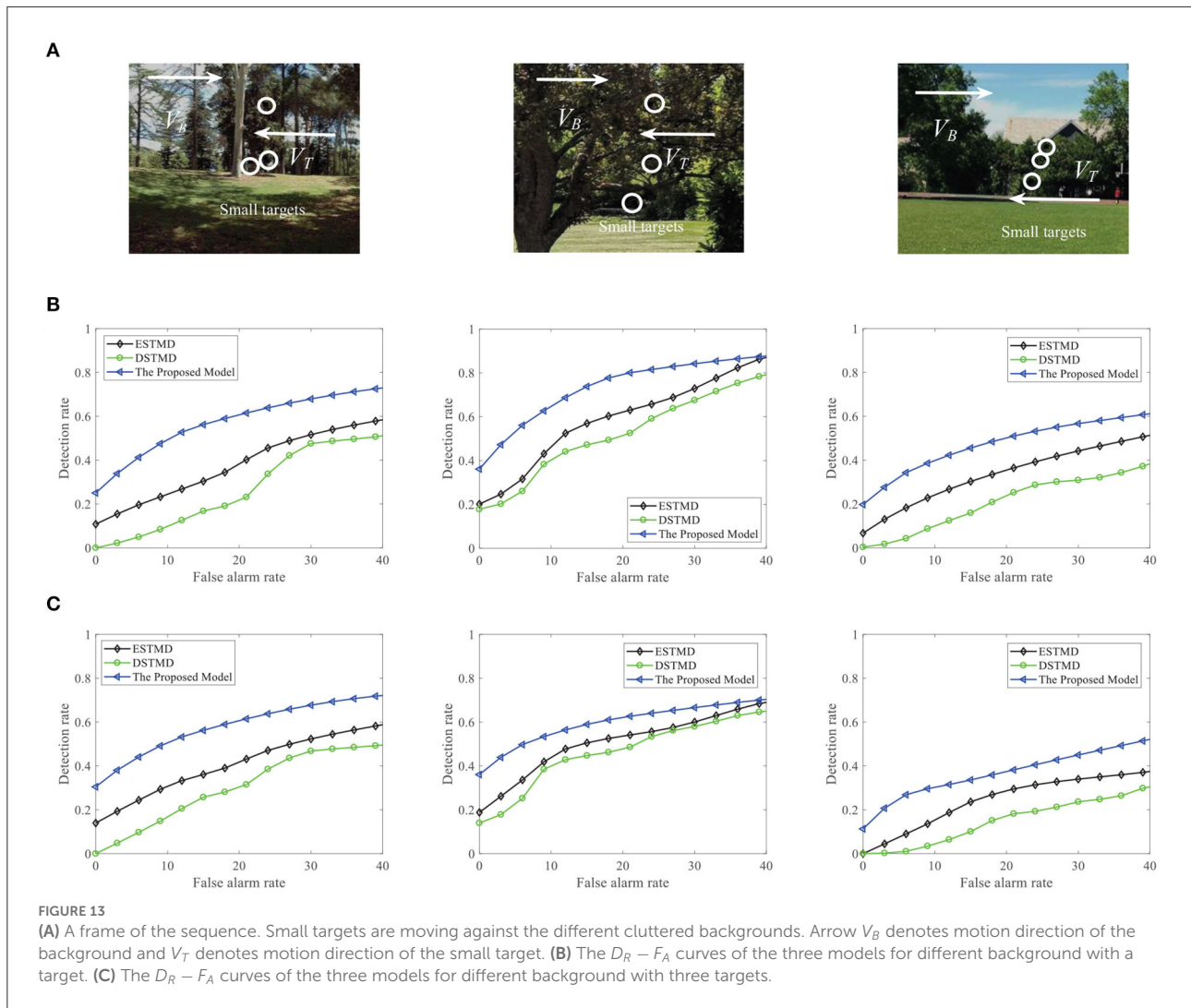
target luminance, target velocity, and background velocity when the false alarm rate is equal to 15. The detection results of the three models for different target sizes are displayed in Figure 12B. It can be seen that the detection rates of three models show a significant decrease after reaching the highest point when the target size increases from 1×1 to 11×11 . However,



the feedback model achieves higher detection rates than other models with the increase in target size. In [Figure 12C](#), we can observe that in the three models, all decrease with the increase in target luminance. It is also worth noting that the proposed model has much better detection performance than the ESTMD and DSTMD when the target luminance is less than 0.1. The results of detection for target velocity are presented in [Figure 12D](#). It can be seen that the detection rates of DSTMD first go up and then slightly decline when the target velocity increases from 100 pixels/s to 330 pixels/s. The detection rates of feedback model and ESTMD increase with the increase of target velocity. However, the performance of feedback model in discriminating the motion of small target from the dynamic background has a superior advantage when target velocity exceed 200 pixels/s. The reason for this is that with the increase of target velocity, the suppression of feedback progressively weakens to a fast-moving target. [Figures 12E,F](#) show the results of the detection of the three models for different background velocities. [Figures 12E,F](#) show that the proposed model improves detection rates when the background velocity v_B is lower than 180 pixels/s. The reason

is that when the background is moving slower than the target, the background false targets will receive stronger suppression by the feedback signal, which consequently leads to the detection rate of \mathcal{F} STMD that is higher than the baseline models.

We further demonstrate the performance of the \mathcal{F} STMD model by changing the background image (see [Figure 13A](#)). [Figure 13B](#) displays the $D_R - F_A$ curves of the three models to detect the motion of a small target in different backgrounds. Meanwhile, the $D_R - F_A$ curves of the three models to detect the motion of three small targets in different backgrounds are presented in [Figure 13C](#). From [Figures 13B,C](#), we can see that the \mathcal{F} STMD has a better performance than the existing models for different backgrounds and numbers of the target. In addition, we also evaluate the models on three real videos and display their detection result curves in [Figure 14](#). The videos are randomly selected from the RIST data set, each of which displays a moving small target against the cluttered background. The corresponding video numbers are given in the caption. As can be seen, the \mathcal{F} STMD outperforms the ESTMD and DSTMD models in all three videos. Specifically, its detection rate is always



higher than those of the other two models at any false alarm rate. Combined with the above experimental results, it shows that our algorithm is practical and feasible.

6. Conclusion

In this paper, we have proposed a time-delay feedback vision system for detecting a small moving target in cluttered backgrounds. Our model contains four neural layers and a time-delay feedback pathway. The four neural layers process motion information in a feed-forward manner, and the feedback pathway propagates the output of the model to lower-layer neurons for weakening slow-moving background false positive responses. In order to determine the feedback intensity, we prove the existence and uniqueness of the solution to the nonlinear dynamic system formed by the feedback loop and designed an iterative algorithm to solve the approximate solution of the model. To demonstrate the advantages of the proposed model, we apply the proposed model to detect small target motions in cluttered backgrounds. The experimental results show that the proposed model maintains a minor effect to fast-moving objects, while significantly suppressing those with lower velocities. Finally, we compare the performance of the proposed model with the existing models, experimental results show that the proposed model is able to improve detection performance for small targets with velocities higher than that of the complex background.

However, the performance of STMD is crucially dependent on the contrast between the background and target. Insects pursue prey based on the contrast between the background and target. The proposed model also uses the contrast between background and target to detect small targets. It conforms to the laws of biological vision. When the contrast between the background and target is small, the performance of the model is low. In the future, some contrast enhancement methods, such as histogram equalization and gray transformation, can be integrated into the model to overcome the problem of low contrast. In addition, this paper only analyzes the fixed time-delay feedback system, which is formed by transmitting the information from the output layer to the lamina layer in the STMD pathway. There may be some problems that have not been considered in the STMD vision system. For example, we may consider the time-varying feedback cases and other neural layers' feedback types in the STMD pathway.

References

Bai, X., and Bi, Y. (2018). Derivative entropy-based contrast measure for infrared small-target detection. *IEEE Trans. Geosci. Remote Sens.* 56, 2452–2466. doi: 10.1109/TGRS.2017.2781143

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

JL: writing, experiments, and theoretical analysis. HL and JP: ideas of the project and review. HW, MX, and HC: experiments and writing-review. All authors contributed to the article and approved the submitted version.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 12031003 and 11771347, in part by the China Postdoctoral Science Foundation under Grant No. 2021M700921.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnbot.2022.984430/full#supplementary-material>

Barnett, P. D., Nordström, K., and O'carroll, D. C. (2007). Retinotopic organization of small-field-target-detecting neurons in the insect visual system. *Curr. Biol.* 17, 569–578. doi: 10.1016/j.cub.2007.02.039

- Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., et al. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* 85, 390–401. doi: 10.1016/j.neuron.2014.12.018
- Behnia, R., Clark, D. A., Carter, A. G., Clandinin, T. R., and Desplan, C. (2014). Processing properties of on and off pathways for drosophila motion detection. *Nature* 512, 427–430. doi: 10.1038/nature13427
- Borst, A. (2009). Drosophila's view on insect vision. *Curr. Biol.* 19, R36–R47. doi: 10.1016/j.cub.2008.11.001
- Borst, A., and Helmstaedter, M. (2015). Common circuit design in fly and mammalian motion vision. *Nat. Neurosci.* 18, 1067. doi: 10.1038/nn.4050
- Bosquet, B., Mucientes, M., and Brea, V. M. (2018). "Stdnet: a convnet for small target detection," in *BMVC* (Northumbria), 253.
- Brunton, S. L., Brunton, B. W., Proctor, J. L., and Kutz, J. N. (2016). Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLoS ONE* 11, e0150171. doi: 10.1371/journal.pone.0150171
- Cao, C., Huang, Y., Yang, Y., Wang, L., Wang, Z., and Tan, T. (2018). Feedback convolutional neural network for visual localization and segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 1627–1640. doi: 10.1109/TPAMI.2018.2843329
- Carreira, J., Agrawal, P., Fragkiadaki, K., and Malik, J. (2016). "Human pose estimation with iterative error feedback," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV: IEEE), 4733–4742.
- Clark, D. A., Bursztyl, L., Horowitz, M. A., Schnitzer, M. J., and Clandinin, T. R. (2011). Defining the computational structure of the motion detector in drosophila. *Neuron* 70, 1165–1177. doi: 10.1016/j.neuron.2011.05.023
- Clark, D. A., and Demb, J. B. (2016). Parallel computations in insect and mammalian visual motion processing. *Curr. Biol.* 26, R1062–R1072. doi: 10.1016/j.cub.2016.08.003
- Clarke, S. E., and Maler, L. (2017). Feedback synthesizes neural codes for motion. *Curr. Biol.* 27, 1356–1361. doi: 10.1016/j.cub.2017.03.068
- De Vries, B., and Prince, J. C. (1991). "A theory for neural networks with time delays," in *Advances in Neural Information Processing Systems* (Denver), 162–168.
- Deng, L., Zhang, J., Xu, G., and Zhu, H. (2021). Infrared small target detection via adaptive m-estimator ring top-hat transformation. *Pattern Recognit.* 112, 107729. doi: 10.1016/j.patcog.2020.107729
- Escobar-Alvarez, H. D., Ohradzansky, M., Keshavan, J., Ranganathan, B. N., and Humbert, J. S. (2019). Bioinspired approaches for autonomous small-object detection and avoidance. *IEEE Trans. Rob.* 35, 1220–1232. doi: 10.1109/TRO.2019.2922472
- Fortun, D., Bouthemy, P., and Kervrann, C. (2015). Optical flow modeling and computation: a survey. *Comput. Vision Image Understand.* 134, 1–21. doi: 10.1016/j.cviu.2015.02.008
- Freifeld, L., Clark, D. A., Schnitzer, M. J., Horowitz, M. A., and Clandinin, T. R. (2013). Gabaergic lateral interactions tune the early stages of visual processing in drosophila. *Neuron* 78, 1075–1089. doi: 10.1016/j.neuron.2013.04.024
- Fu, Q., Wang, H., Hu, C., and Yue, S. (2019). Towards computational models and applications of insect visual systems for motion perception: a review. *Artif. Life* 25, 263–311. doi: 10.1162/artl_a_00297
- Gao, C., Meng, D., Yang, Y., Wang, Y., Zhou, X., and Hauptmann, A. G. (2013). Infrared patch-image model for small target detection in a single image. *IEEE Trans. Image Process.* 22, 4996–5009. doi: 10.1109/TIP.2013.2281420
- Geurten, B. R., Nordström, K., Sprayberry, J. D., Bolzon, D. M., and O'Carroll, D. C. (2007). Neural mechanisms underlying target detection in a dragonfly centrifugal neuron. *J. Exp. Biol.* 210, 3277–3284. doi: 10.1242/jeb.008425
- Hanche-Olsen, H., and Holden, H. (2010). The kolmogorov-riesz compactness theorem. *Exposit. Math.* 28, 385–394. doi: 10.1016/j.exmath.2010.03.001
- Hassenstein, B., and Reichardt, W. (1956). Systemtheoretische analyse der zeit, reihenfolgen-und vorzeichenbewertung bei der bewegungsperzeption des rüsselkäfers chlorophanus. *Zeitschrift für Naturforschung B* 11, 513–524. doi: 10.1515/znb-1956-9-1004
- Javed, S., Mahmood, A., Al-Maadeed, S., Bouwmans, T., and Jung, S. K. (2018). Moving object detection in complex scene using spatiotemporal structured-sparse rpa. *IEEE Trans. Image Process.* 28, 1007–1022. doi: 10.1109/TIP.2018.2874289
- Joesch, M., Schnell, B., Raghu, S. V., Reiff, D. F., and Borst, A. (2010). On and off pathways in drosophila motion vision. *Nature* 468, 300–304. doi: 10.1038/nature09545
- Keleş, M. F., and Frye, M. A. (2017). Object-detecting neurons in drosophila. *Curr. Biol.* 27, 680–687. doi: 10.1016/j.cub.2017.01.012
- Khan, R. A., Rehman, M., and Henderson, J. (2011). Existence and uniqueness of solutions for nonlinear fractional differential equations with integral boundary conditions. *Fract. Differ. Calc* 1, 29–43. doi: 10.7153/fdc-01-02
- Klink, P. C., Dagnino, B., Gariel-Mathis, M.-A., and Roelfsema, P. R. (2017). Distinct feedforward and feedback effects of microstimulation in visual cortex reveal neural mechanisms of texture segregation. *Neuron* 95, 209–220. doi: 10.1016/j.neuron.2017.05.033
- Lamme, V. A., Super, H., and Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Curr. Opin. Neurobiol.* 8, 529–535. doi: 10.1016/S0959-4388(98)80042-1
- Li, F., Li, Y., and Liang, Z. (2006). Existence of solutions to nonlinear hammerstein integral equations and applications. *J. Math. Anal. Appl.* 323, 209–227. doi: 10.1016/j.jmaa.2005.10.014
- Li, J., Liang, X., Wei, Y., Xu, T., Feng, J., and Yan, S. (2017). "Perceptual generative adversarial networks for small object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Honolulu, HI: IEEE), 1222–1230.
- Liu, Y.-J., and Tong, S. (2015). Adaptive fuzzy control for a class of unknown nonlinear dynamical systems. *Fuzzy Sets Syst.* 263, 49–70. doi: 10.1016/j.fss.2014.08.008
- Meglić, A., Ilić, M., Pirić, P., Škorjanc, A., Wehling, M. F., Kreft, M., et al. (2019). Horserfly object-directed polarotaxis is mediated by a stochastically distributed ommatidial subtype in the ventral retina. *Proc. Natl. Acad. Sci. U.S.A.* 116, 21843–21853. doi: 10.1073/pnas.1910807116
- Mischianti, M., Lin, H.-T., Herold, P., Imler, E., Olberg, R., and Leonardo, A. (2015). Internal models direct dragonfly interception steering. *Nature* 517, 333–338. doi: 10.1038/nature14045
- Mohsenzadeh, Y., Qin, S., Cichy, R. M., and Pantazis, D. (2018). Ultra-rapid serial visual presentation reveals dynamics of feedforward and feedback processes in the ventral visual pathway. *Elife* 7, e36329. doi: 10.7554/eLife.36329
- Nordström, K., Barnett, P. D., and O'Carroll, D. C. (2006). Insect detection of small targets moving in visual clutter. *PLoS Biol.* 4, e54. doi: 10.1371/journal.pbio.0040054
- Nordström, K., and O'Carroll, D. C. (2006). Small object detection neurons in female hoverflies. *Proc. R. Soc. B Biol. Sci.* 273, 1211–1216. doi: 10.1098/rspb.2005.3424
- Nordstrom, K., and O'Carroll, D. C. (2009). Feature detection and the hypercomplex property in insects. *Trends Neurosci.* 32, 383–391. doi: 10.1016/j.tins.2009.03.004
- Nurminen, L., Merlin, S., Bijanzadeh, M., Federer, F., and Angelucci, A. (2018). Top-down feedback controls spatial summation and response amplitude in primate visual cortex. *Nat. Commun.* 9, 1–13. doi: 10.1038/s41467-018-04500-5
- Paulk, A. C., Kirszenblat, L., Zhou, Y., and van Swinderen, B. (2015). Closed-loop behavioral control increases coherence in the fly brain. *J. Neurosci.* 35, 10304–10315. doi: 10.1523/JNEUROSCI.0691-15.2015
- Paulk, A. C., Stacey, J. A., Pearson, T. W., Taylor, G. J., Moore, R. J., Srinivasan, M. V., et al. (2014). Selective attention in the honeybee optic lobes precedes behavioral choices. *Proc. Natl. Acad. Sci. U.S.A.* 111, 5006–5011. doi: 10.1073/pnas.1323297111
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV: IEEE), 779–788.
- Rosner, R., von Hadeln, J., Tarawneh, G., and Read, J. C. (2019). A neuronal correlate of insect stereopsis. *Nat. Commun.* 10, 1–9. doi: 10.1038/s41467-019-10721-z
- Sahoo, A., Xu, H., and Jagannathan, S. (2015). Neural network-based event-triggered state feedback control of nonlinear continuous-time systems. *IEEE Trans. Neural Netw. Learn. Syst.* 27, 497–509. doi: 10.1109/TNNLS.2015.2416259
- Sallemi, I., and Shah, M. (2013). Multiframe many-many point correspondence for vehicle tracking in high density wide area aerial videos. *Int. J. Comput. Vis.* 104, 198–219. doi: 10.1007/s11263-013-0624-1
- Sanes, J. R., and Zipursky, S. L. (2010). Design principles of insect and vertebrate visual systems. *Neuron* 66, 15–36. doi: 10.1016/j.neuron.2010.01.018
- Shuigen, W., Zhen, C., and Hua, D. (2009). "Motion detection based on temporal difference method and optical flow field," in *2009 Second International Symposium on Electronic Commerce and Security*, Vol. 2 (Nanchang: IEEE), 85–88.
- Soker, N. (2016). The jet feedback mechanism (jfm) in stars, galaxies and clusters. *New Astron. Rev.* 75, 1–23. doi: 10.1016/j.newar.2016.08.002
- Sousa, J. V. d., C., Kucche, K. D., and De Oliveira, E. C. (2019). Stability of ψ -hilfer impulsive fractional differential equations. *Appl. Math. Lett.* 88, 73–80. doi: 10.1016/j.aml.2018.08.013

- Straw, A. D. (2008). Vision egg: an open-source library for realtime visual stimulus generation. *Front. Neuroinform.* 2, 4. doi: 10.3389/neuro.11.004.2008
- Takemura, S.-y., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P. K., et al. (2013). A visual motion detection circuit suggested by drosophila connectomics. *Nature* 500, 175–181. doi: 10.1038/nature12450
- Tang, J., Tian, Y., Zhang, P., and Liu, X. (2017). Multiview privileged support vector machines. *IEEE Trans. Neural Netw. Learn. Syst.* 29, 3463–3477. doi: 10.1109/TNNLS.2017.2728139
- Tuthill, J. C., Nern, A., Holtz, S. L., Rubin, G. M., and Reiser, M. B. (2013). Contributions of the 12 neuron classes in the fly lamina to motion vision. *Neuron* 79, 128–140. doi: 10.1016/j.neuron.2013.05.024
- Wang, C., and Huang, K.-Q. (2015). Vfm: visual feedback model for robust object recognition. *J. Comput. Sci. Technol.* 30, 325–339. doi: 10.1007/s11390-015-1526-1
- Wang, H., Peng, J., and Yue, S. (2018). A directionally selective small target motion detecting visual neural network in cluttered backgrounds. *IEEE Trans. Cybern.* 50, 1541–1555. doi: 10.1109/TCYB.2018.2869384
- Wang, H., Peng, J., Zheng, X., and Yue, S. (2019). A robust visual system for small target motion detection against cluttered moving backgrounds. *IEEE Trans. Neural Netw. Learn. Syst.* 31, 839–853. doi: 10.1109/TNNLS.2019.2910418
- Warrant, E. J. (2017). The remarkable visual capacities of nocturnal insects: vision at the limits with small eyes and tiny brains. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160063. doi: 10.1098/rstb.2016.0063
- Wiederman, S. D., and O'Carroll, D. C. (2013a). "Biomimetic target detection: modeling 2 nd order correlation of off and on channels," in *2013 IEEE Symposium on Computational Intelligence for Multimedia, Signal and Vision Processing (CIMSIVP)* (Singapore: IEEE), 16–21.
- Wiederman, S. D., Shoemaker, P. A., and O'Carroll, D. C. (2008). A model for the detection of moving targets in visual clutter inspired by insect physiology. *PLoS ONE* 3, e2784. doi: 10.1371/journal.pone.0002784
- Wiedermann, S., and O'Carroll, D. C. (2013b). Biologically inspired feature detection using cascaded correlations of off and on channels. *J. Artif. Intell. Soft Comput. Res.* 3, 5–14. doi: 10.2478/jaiscr-2014-0001
- Xu, F., Ding, X., Peng, J., Yuan, G., Wang, Y., Zhang, J., et al. (2018). "Real-time detecting method of marine small object with underwater robot vision," in *2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)* (Kobe: IEEE), 1–4.
- Zhang, X., Wang, T., Qi, J., Lu, H., and Wang, G. (2018). "Progressive attention guided recurrent network for salient object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT: IEEE), 714–722.