

# An Improved LPTC Neural Model for Background Motion Direction Estimation

Hongxin Wang  
School of Computer Science  
University of Lincoln  
Lincoln, LN6 7TS, UK  
13488337@students.lincoln.ac.uk

Jigen Peng  
School of Mathematics and Statistics  
Xi'an Jiaotong University  
Xi'an, Shaanxi, 710049, P.R. China  
jgpeng@mail.xjtu.edu.cn

Shigang Yue  
School of Computer Science  
University of Lincoln  
Lincoln, LN6 7TS, UK  
syue@lincoln.ac.uk

**Abstract**—A class of specialized neurons, called lobula plate tangential cells (LPTCs) has been shown to respond strongly to wide-field motion. The classic model, elementary motion detector (EMD) and its improved model, two-quadrant detector (TQD) have been proposed to simulate LPTCs. Although EMD and TQD can percept background motion, their outputs are so cluttered that it is difficult to discriminate actual motion direction of the background. In this paper, we propose a max operation mechanism to model a newly-found transmedullary neuron Tm9 whose physiological properties do not map onto EMD and TQD. This proposed max operation mechanism is able to improve the detection performance of TQD in cluttered background by filtering out irrelevant motion signals. We will demonstrate the functionality of this proposed mechanism in wide-field motion perception.

## I. INTRODUCTION

When flies search for and track prey or conspecifics, their own motion generates displacement of the visual surroundings, inducing wide-field background motion across the retina [1]. A class of specialized neurons, called lobula plate tangential cells (LPTCs), has been shown to respond strongly to wide-field motion. LPTCs can be broadly divided into a vertical system (VS) and a horizontal system (HS), which signal wide-field motion in vertical and horizontal directions, respectively [2].

The classic correlation model, elementary motion detector (EMD) [3] and its improved model, two-quadrant detector (TQD) [4], [5] have been proposed to simulate LPTC neurons. These two models show strong responses to wide-field motion and have a clear mapping onto neural circuits of fly visual system. Although EMD and TQD are able to detect background motion, detection performances of these two models are always unsatisfying, especially in cluttered environment. Due to indiscriminate signal correlation, both EMD and TQD always have four outputs which do not show much differences in the strength, representing lobula plate tangential cells' (LPTCs) neural responses along four cardinal directions (up, down, left, right). In some cases, model outputs along actual motion direction is even weaker than model outputs along other directions. Therefore, the actual direction of target motion cannot be determined by simply comparing the strengths of model outputs along four cardinal directions.

Recently, biologists have identified a transmedullary neuron, Tm9 whose physiological properties do not map onto classic EMD and TQD models, but is required for motion perception [6]. Further research indicates that the receptive field of Tm9 is much larger than that of its downstream neurons T5. Besides, signals from multi columns are converging at the level of Tm9. These findings are surprising especially when we consider that only signals from two adjacent photoreceptors are needed for motion computation in both EMD and TQD models. Based on these findings, we infer that Tm9 neurons are able to inform downstream neurons about local points in a wide receptive field. This property of Tm9 may help flies effectively avoid confusion caused by incorrect signal correlation while perceiving wide-field motion.

In this paper, we propose a max operation mechanism to simulate Tm9 neurons in order to improve the detection performance of TQD in cluttered background. This mechanism which acts on signals after ON-OFF channel separation of TQD is able to inform downstream neuron T4 and T5 about spatial maximum of ON and OFF signals in a local neighborhood. These local maximal signals are then temporally-delayed and integrated using the same method with classic TQD model. In the following paper, we will present modeling details of the improved TQD model, meanwhile demonstrating that the improved TQD model is able to overcome shortages of classic TQD model.

## II. MODELING

Fig. 1 shows the schematic of the improved TQD model. For showing the difference between classic TQD and the improved TQD model, the connection between Tm9 and T4, T5 neurons was not plotted in Fig. 1. We introduce the improved TQD model layer by layer in the following paper.

### A. Retina Layer

In order to simulate the signal processing of photoreceptors, we start by representing visual stimuli as varying luminance values, noted  $I(x, y, t) \in \mathbf{R}$ , where  $x, y$  and  $t$  are spatial and temporal field positions. Then, the functionality of photoreceptors is described by the following equation,

$$L(x, y, t) = \iint I(u, v, t) G_{\sigma_1}(x - u, y - v) dudv \quad (1)$$

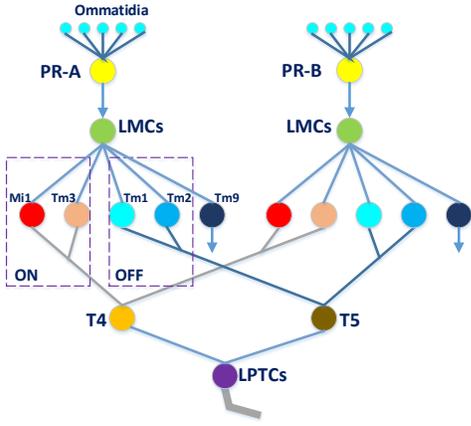


Fig. 1. The schematic of the improved TQD model, Each colorized disk denotes a neuron. PR-A, PR-B, LMCs, LPTCs are the abbreviation of photoreceptor A, Photoreceptor B, Large Monopolar Cells, Lobula Plate Tangential Cells, respectively.

where  $L(x, y, t)$  is the output of photoreceptors. Particularly, for a given spatial position  $(x_0, y_0)$ ,  $I(x_0, y_0, t)$  represents luminance information received by the ommatidia located in spatial position  $(x_0, y_0)$  at time  $t$  while  $L(x_0, y_0, t)$  denotes the output of the photoreceptor located in position  $(x_0, y_0)$  at time  $t$ .  $G_{\sigma_1}(x, y)$  is a Gaussian function, defined as

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

### B. Lamina Layer

Photoreceptors synapse on large monopolar cells (LMCs) located in lamina layer which are able to remove redundancy contained in input signals  $(L(x, y, t))$  and maximize information about illumination change. Here, we implement a temporal contrast detector on input signal  $(L(x, y, t))$  so as to simulate neural responses of LMCs. That is,

$$P(x, y, t) = \int L(x, y, s)H(t-s)ds \quad (3)$$

$$H(t) = \Gamma_{n_1, \tau_1}(t) - \Gamma_{n_2, \tau_2}(t) \quad (4)$$

where  $P(x, y, t)$  and  $\Gamma_{n, \tau}(t)$  are the output of LMCs, Gamma function, respectively.  $\Gamma_{n, \tau}(t)$  is defined as

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

Before LMCs relay processed signals  $(P(x, y, t))$  to medulla layer, LMCs receive lateral inhibition from adjacent neurons. In accord with classic lateral inhibition mechanism, we convolve signal  $P(x, y, t)$  with an inhibition kernel  $W_1(x, y, t)$ . That is,

$$P_I(x, y, t) = \iiint P(u, v, s)W_1(x-u, y-v, t-s)dudvds \quad (6)$$

where  $P_I(x, y, t)$  is laterally inhibited signal and  $W_1(x, y, t)$  is defined using the following equations,

$$W_1(x, y, t) = W_S^{Pos}(x, y)W_T^{Pos}(t) + \dots + W_S^{Neg}(x, y)W_T^{Neg}(t). \quad (7)$$

In this paper, we set  $W_S^{Pos}(x, y)$ ,  $W_S^{Neg}(x, y)$ ,  $W_T^{Pos}(t)$ ,  $W_T^{Neg}(t)$  as

$$W_S^{Pos} = [G_{\sigma_2}(x, y) - G_{\sigma_3}(x, y)]^+ \quad (8)$$

$$W_S^{Neg} = [G_{\sigma_2}(x, y) - G_{\sigma_3}(x, y)]^-, \quad \sigma_3 = 2 * \sigma_2 \quad (9)$$

$$W_T^{Pos} = \frac{1}{\alpha_1} \exp\left(-\frac{t}{\alpha_1}\right) \quad (10)$$

$$W_T^{Neg} = \frac{1}{\alpha_2} \exp\left(-\frac{t}{\alpha_2}\right), \quad \alpha_2 > \alpha_1. \quad (11)$$

where  $[x]^+$ ,  $[x]^-$  denote  $\max(x, 0)$  and  $\min(x, 0)$ , respectively.  $G_{\sigma_2}(x, y)$  and  $G_{\sigma_3}(x, y)$  are Gaussian functions.

### C. Medulla Layer

Previous research identified two parallel pathways which selectively respond to brightness increments (ON pathway) and decrements (OFF pathway) in medulla layer [7], [8]. These two parallel pathways are implemented by four intermediate neurons, i.e., Tm1, Tm2, Tm3 and Mi1. To be more precise, Mi1 and Tm3 constitute ON pathway whereas Tm1 and Tm2 form OFF pathway, shown in Fig. 1. Besides, compared to the output of Tm3, the output of Mi1 is temporally delayed. Similarly, the output of Tm1 has a temporal delay compared to that of Tm2.

Based on these biological findings, TQD which is the improved model of EMD firstly separate laterally inhibited signal  $P_I(x, y, t)$  into ON and OFF channels, shown in Fig. 1. That is,

$$S^{ON}(x, y, t) = \frac{(|P_I(x, y, t)| + P_I(x, y, t))}{2} \quad (12)$$

$$S^{OFF}(x, y, t) = \frac{(|P_I(x, y, t)| - P_I(x, y, t))}{2} \quad (13)$$

where  $S^{ON}$  and  $S^{OFF}$  denote the output of Tm3 and Tm2, respectively.

Due to small temporal delay exists between the outputs of Mi1 and Tm3 (the outputs of Tm1 and Tm2), signal  $S^{ON}$  ( $S^{OFF}$ ) is convolved with a Gamma function so as to obtain the delayed signal  $S_D^{ON}$  ( $S_D^{OFF}$ ). This process can be described by the following equations,

$$S_D^{ON}(x, y, t) = \int S^{ON}(x, y, s)\Gamma_{n_3, \tau_3}(t-s)ds \quad (14)$$

$$S_D^{OFF}(x, y, t) = \int S^{OFF}(x, y, s)\Gamma_{n_3, \tau_3}(t-s)ds \quad (15)$$

where  $S_D^{ON}$  and  $S_D^{OFF}$  are time-delayed signals, corresponding to the output of Mi1 and Tm1, respectively.

However, recent research has identified a transmedullary neuron Tm9 whose physiological properties do not map onto classic TQD model but which is required for motion perception [6]. Compared to other neurons, such as Tm1, Tm2, Tm3 and Mi1, Tm9 has much larger receptive field which conflicts with the view that downstream neurons, like T4 and T5, only require signals from two neighboring photoreceptors in a relatively small receptive field. Because Tm9 has a larger receptive field, signals from multi columns (or photoreceptors) can be integrated in Tm9. Obviously, this functionality cannot

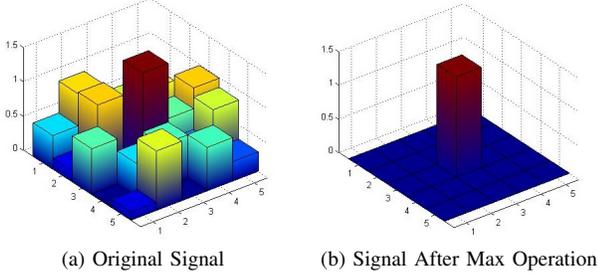


Fig. 2. Schematic illustration of the proposed max operation mechanism.

be accomplished by other neurons, like Tm1, Tm2, Tm3 and Mi1, which always only receive signal from a single column (or photoreceptor). Based on these biological findings, we assume that Tm9 is able to compare the strength of signals received from multi columns and find a local maximum. In order to account for properties of Tm9, we propose a max operation mechanism acting on ON and OFF signals ( $S^{ON}$  and  $S^{OFF}$ ). This max operation mechanism can be described by the following equations,

$$\tilde{S}^{ON}(x_0, y_0, t) = \begin{cases} 0 & \text{if flag}^{ON} = 0 \\ S^{ON}(x_0, y_0, t) & \text{if flag}^{ON} = 1 \end{cases} \quad (16)$$

$$\tilde{S}^{OFF}(x_0, y_0, t) = \begin{cases} 0 & \text{if flag}^{OFF} = 0 \\ S^{OFF}(x_0, y_0, t) & \text{if flag}^{OFF} = 1 \end{cases} \quad (17)$$

where  $\text{flag}^{ON}$  and  $\text{flag}^{OFF}$  are defined by the following equations,

$$\text{flag}^{ON} = \begin{cases} 0 & \text{if } S^{ON}(x_0, y_0, t) \neq \max_{(x,y) \in \Omega(x_0, y_0)} S^{ON} \\ 1 & \text{if } S^{ON}(x_0, y_0, t) = \max_{(x,y) \in \Omega(x_0, y_0)} S^{ON} \end{cases} \quad (18)$$

$$\text{flag}^{OFF} = \begin{cases} 0 & \text{if } S^{OFF}(x_0, y_0, t) \neq \max_{(x,y) \in \Omega(x_0, y_0)} S^{OFF} \\ 1 & \text{if } S^{OFF}(x_0, y_0, t) = \max_{(x,y) \in \Omega(x_0, y_0)} S^{OFF} \end{cases} \quad (19)$$

where  $\Omega(x_0, y_0)$  is a local neighborhood centered at  $(x_0, y_0)$ .

In order to clearly illustrate this max operation mechanism, an example is shown in Fig. 2. As we can see from Fig. 2,  $S^{ON}(3, 3, t_0)$  (or  $S^{OFF}(3, 3, t_0)$ ) is the local maximum in the local neighborhood  $\Omega_{(3,3)}$ , so  $S^{ON}(3, 3, t_0)$  (or  $S^{OFF}(3, 3, t_0)$ ) is preserved after max operation. However, because other signals, for example  $S^{ON}(1, 1, t_0)$ ,  $S^{ON}(1, 2, t_0)$ ,  $S^{ON}(2, 2, t_0)$  and so on, are not local maximum in the local neighborhood, these signals are set as 0 after max operation. Obviously, this max operation mechanism is able to decrease clutter and increase sparsity of input signals. Due to the increment of sparsity, incorrect signal correlation will be effectively avoided in signal-correlation step.

After max operation, signal  $\tilde{S}^{ON}(x, y, t)$  and  $\tilde{S}^{OFF}(x, y, t)$  are temporally delayed. This step is similar with time-delay operation shown in Eq.(14) and (15).

$$\tilde{S}_D^{ON}(x, y, t) = \int \tilde{S}^{ON}(x, y, s) \Gamma_{n_3, \tau_3}(t - s) ds \quad (20)$$

$$\tilde{S}_D^{OFF}(x, y, t) = \int \tilde{S}^{OFF}(x, y, s) \Gamma_{n_3, \tau_3}(t - s) ds \quad (21)$$

where  $\tilde{S}_D^{ON}(x, y, t)$  and  $\tilde{S}_D^{OFF}(x, y, t)$  are temporally-delayed signals.

#### D. Lobula Layer

In lobula layer, various high-order neurons integrate signals relayed from ON and OFF pathways, then respond selectively to specific visual stimuli. For example, small target motion detectors (STMDs) show exquisite selectivity for small target motion. Lobula plate tangential cells (LPTCs) are sensitive to wide-field motion. Elementary small target motion detector (ESTMD) [9], [10] and two-quadrant detector (TQD) [4] have been proposed to simulate STMD and LPTC neurons, respectively. In this paper, we focus on LPTC neuron modeling for background motion direction detection.

T4 and T5 neurons which are pre-synaptic neurons of LPTCs integrate signals relayed from medulla layer. More precisely, T4 responds selectively to ON signals while T5 is specialized for OFF signals. Let  $D^{T4}(x, y, t; \theta)$  and  $D^{T5}(x, y, t; \theta)$  denote T4 and T5 neural responses at spatial-temporal coordinate  $(x, y, t)$  along direction  $\theta$ , respectively. Then, in classic TQD model,  $D^{T4}(x, y, t; \theta)$  and  $D^{T5}(x, y, t; \theta)$  are given by the following equations,

$$D^{T4}(x, y, t; \theta) = S^{ON}(x, y, t) S_D^{ON}(x', y', t) \quad (22)$$

$$D^{T5}(x, y, t; \theta) = S^{OFF}(x, y, t) S_D^{OFF}(x', y', t) \quad (23)$$

where  $\theta \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$ .

Because T4 and T5 also receive signals from Tm9, T4 and T5 have corresponding outputs  $\tilde{D}^{T4}$ ,  $\tilde{D}^{T5}$ ,

$$\tilde{D}^{T4}(x, y, t; \theta) = \tilde{S}^{ON}(x, y, t) \tilde{S}_D^{ON}(x', y', t) \quad (24)$$

$$\tilde{D}^{T5}(x, y, t; \theta) = \tilde{S}^{OFF}(x, y, t) \tilde{S}_D^{OFF}(x', y', t) \quad (25)$$

where  $\theta \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$ .

Lobula plate tangential cells (LPTCs) further integrate signals provided by T4 and T5 neurons.

**1) Classic TQD Model:** For classic TQD model, LPTC output  $F(x, y, t; \theta)$  is defined by the following equation,

$$F(x, y, t; \theta) = D^{T4}(x, y, t; \theta) + D^{T5}(x, y, t; \theta). \quad (26)$$

**2) Improved TQD Model:** For improved TQD model, LPTC output  $\tilde{F}(x, y, t; \theta)$  is defined by the following equation,

$$\tilde{F}(x, y, t; \theta) = \tilde{D}^{T4}(x, y, t; \theta) + \tilde{D}^{T5}(x, y, t; \theta). \quad (27)$$

The direction of background motion is determined by comparing the strength of LPTCs' neural responses at different directions. That is,

$$\Theta(t) = \arg \max_{\theta} \iint \tilde{F}(x, y, t; \theta) dx dy \quad (28)$$



Fig. 3. The 840th frame of the first image sequence. The red arrow and  $V_B$  denote motion direction and velocity of background, respectively.

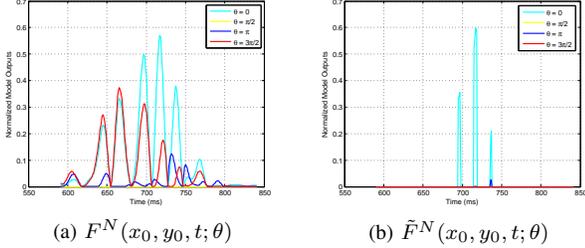


Fig. 4. The normalized model outputs of a spatial coordinate  $(x_0, y_0)$  of the first image sequence during time period  $[550, 850]$  ms, i.e.,  $F^N(x_0, y_0, t; \theta)$ ,  $\tilde{F}^N(x_0, y_0, t; \theta)$ ,  $t \in [550, 850]$  ms,  $\theta \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$ .

where  $\Theta(t)$  denotes the motion direction of background at time  $t$ .

Numerous methods have been developed for optic flow estimation. Particularly, structure-tensor based methods construct the tensor for each pixel within its neighborhood, then convert the optic flow estimation problem to an eigenvalue analysis problem [11], [12]. Compared to structure-based methods, the improved TQD model offers a totally different way to estimate background motion. Based on biological findings, TQD model detects background motion by correlating signals relayed from two photoreceptors. Although this correlation method is relatively simple, it reflects the signal processing mechanism and neural circuits of fly visual system.

### III. EXPERIMENT

In this section, three synthetic image sequences were used to evaluate detection performance of classic TQD and improved TQD models. The sampling frequencies of these three image sequences are all set as 1000 Hz. Fig. 3 shows a frame of the first image sequence which is 500 (in horizontal) by 250 (in vertical) pixels. As we have mentioned before, background motion is caused by flies' ego-motion. Therefore, in this paper, background is in one of four cardinal direction motion (rightward, leftward, upward, downward) so as to simulate the displacement of flies' head. For example, in Fig. 3, background is in rightward motion where red arrow and  $V_B$  denote motion direction and velocity of background, respectively.

In order to compare classic TQD model output  $F(x, y, t; \theta)$  and the improved TQD model output  $\tilde{F}(x, y, t; \theta)$ ,  $F(x, y, t; \theta)$  and  $\tilde{F}(x, y, t; \theta)$  are firstly normalized. That is,

$$F^N(x, y, t; \theta) = \frac{F(x, y, t; \theta)}{\max_{x, y, \theta} F(x, y, t; \theta)} \quad (29)$$

$$\tilde{F}^N(x, y, t; \theta) = \frac{\tilde{F}(x, y, t; \theta)}{\max_{x, y, \theta} \tilde{F}(x, y, t; \theta)}. \quad (30)$$

Then, the normalized model outputs of a spatial coordinate  $(x_0, y_0)$  during a time period  $[550, 850]$  ms, i.e.,  $F^N(x_0, y_0, t; \theta)$ ,  $\tilde{F}^N(x_0, y_0, t; \theta)$ ,  $t \in [550, 850]$  ms,  $\theta \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$ , are shown in Fig. 4. As it is shown in Fig. 4a, model output at direction  $\frac{3\pi}{2}$ , i.e.,  $F^N(x_0, y_0, t; \frac{3\pi}{2})$ , is higher than model output at direction 0, i.e.,  $F^N(x_0, y_0, t; 0)$ , during time period  $[620, 680]$  ms. This result conflicts with that TQD model should show the strongest response along actual motion direction, so motion direction can be inferred by determining the direction of the strongest model response. However, as we have mentioned before, for classic TQD model, incorrect signal correlation will cause confusion in motion-direction determination, especially in cluttered background. This confusion reflects in that model output along actual motion direction is not significantly higher or even lower than model outputs along other directions. Because background is in rightward motion (direction 0), the expected result should be that  $F^N(x_0, y_0, t; 0)$  is higher than  $F^N(x_0, y_0, t; \frac{3\pi}{2})$  during time period  $[550, 850]$  ms. Obviously, as we can see from Fig. 4a, confusion has arisen in classic TQD model outputs  $F^N(x_0, y_0, t; \theta)$ . Compared to Fig. 4a, in Fig. 4b, all model outputs at four cardinal directions of the improved TQD model, i.e.,  $\tilde{F}^N(x_0, y_0, t; \theta)$ ,  $\theta \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$ , are close to 0 during time period  $[620, 680]$  ms. This is because  $S^{ON}(x_0, y_0, t)$  and  $S^{OFF}(x_0, y_0, t)$  are not local maximum in the local neighborhood  $\Omega_{(x_0, y_0)}$  during time period  $[620, 680]$  ms. After max operation mechanism, all signals which are not local maximum in neighborhood  $\Omega_{(x_0, y_0)}$ , will be set as 0. Therefore, model outputs of spatial coordinate  $(x_0, y_0)$  will be close to 0 during this time period. However, we should mention that although model outputs of the improved TQD model are 0 during time period  $[620, 680]$  ms, motion direction of the background in neighborhood  $\Omega_{(x_0, y_0)}$  can be inferred by local maximum in this neighborhood. This is feasible because a local maximum must exist in each local neighborhood.

In order to intuitively present detection performance of classic TQD model and the improved TQD model, model outputs of these two models corresponding to Fig. 3, i.e.,  $F^N(x, y, 840; \theta)$  and  $\tilde{F}^N(x, y, 840; \theta)$ , are projected onto X-Y plane. Here, we should indicate that when a projection threshold  $\gamma$ , time  $t$  and direction  $\theta$  are given, spatial coordinate  $(x, y)$  whose corresponding model outputs  $F^N(x, y, t; \theta)$  and  $\tilde{F}^N(x, y, t; \theta)$  are larger than projection threshold  $\gamma$ , can be shown on X-Y plane. Projection results of Fig. 3 are presented in Fig. 5 and Fig. 6, where projection threshold  $\gamma$  is set as 0.05. As we can see from Fig. 5, classic TQD model not only show strong response along actual motion direction ( $\theta = 0$ ), but also along other three directions. Obviously, this is not the result of what we expect. Because background is in rightward motion shown in Fig. 3, TQD model should show the strongest response to actual motion direction ( $\theta = 0$ ), but much weaker or even no responses to other directions. However, due to incorrect signal correlation mentioned before,

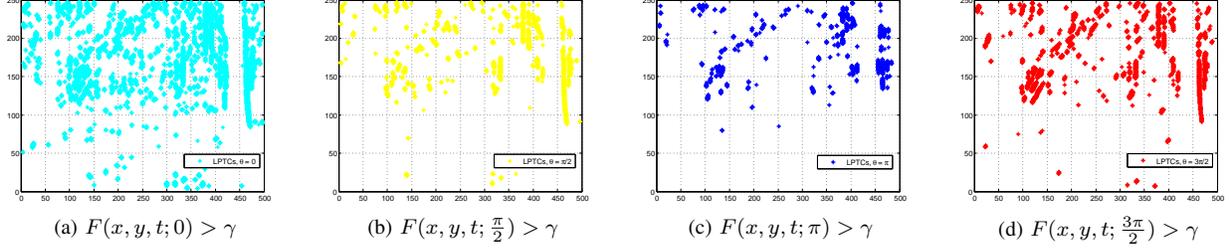


Fig. 5. Projection results of normalized model outputs  $F^N(x, y, t; \theta)$  of the first image sequence, where projection threshold is set as 0.05 and  $t$  is equal to 840 ms.

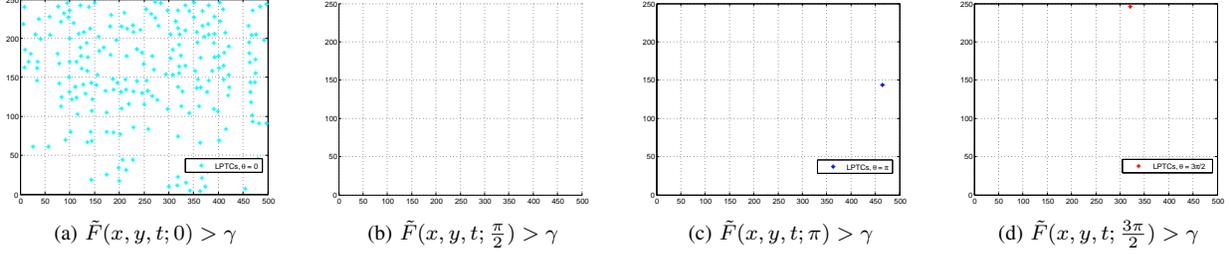


Fig. 6. Projection results of normalized model outputs  $\tilde{F}^N(x, y, t; \theta)$  of the first image sequence, where projection threshold is set as 0.05 and  $t$  is equal to 840 ms.

classic TQD model may have four strong responses along four cardinal directions at a spatial coordinate. In this case, motion direction of the background cannot be obtained by determining the direction of the strongest model outputs. For this reason, confusion will arise when we determine motion direction of the background in a local region. The output of the improved TQD model is much clearer than the output of classic TQD model. In comparison with Fig. 5, in Fig. 6, the improved TQD model shows strong response to actual motion direction ( $\theta = 0$ ), but do not respond to other directions. Therefore, we can effectively infer motion direction of the background in a local region by the direction of the strongest response of the local maximum in this local region.

In the following paper, two evaluation indexes are defined so as to quantitatively evaluate detection performance of classic TQD model and the improved TQD model. Firstly, a set of projection threshold  $\{\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_n\}$  where  $\gamma_0 = 0.01$  and  $\gamma_{i+1} > \gamma_i, i \geq 0$ , should be given. Then, for a projection threshold  $\gamma_i$ , time  $t$  and direction  $\theta$ , we can obtain the number of points  $(x, y)$  whose corresponding output  $F^N(x, y, t; \theta)$  (or  $\tilde{F}^N(x, y, t; \theta)$ ) is higher than projection threshold  $\gamma_i$ , denoted by  $N(\gamma_i, \theta, t)$ . Here, we define detection rate  $DR(\gamma_i, t)$  and the normalized number of detected points  $NP(\gamma_i, t)$  by the following equations,

$$DR(\gamma_i, t) = \frac{N(\gamma_i, \theta_0, t)}{\sum_{\theta} N(\gamma_i, \theta, t)} \quad (31)$$

$$NP(\gamma_i, t) = \frac{N(\gamma_i, \theta_0, t)}{\sum_{\gamma_i} N(\gamma_i, \theta_0, t)} \quad (32)$$

where  $\theta_0$  is the motion direction of background.

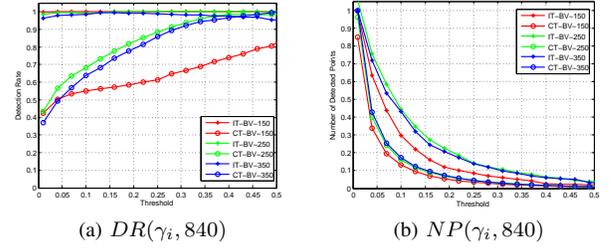


Fig. 7. Detection rate  $DR(\gamma_i, 840)$  and the normalized number of detected points  $NP(\gamma_i, 840)$  of the 840th frame of the first image sequence. Horizontal axis denotes projection threshold while vertical axis represents detection rate or the number of detected points. Legend IT-BV-150 and CT-BV-150 denote the result of the improved TQD model (IT) and classic TQD model (CT) when background velocity (BV) is set as 150. Similarly for other legends.

For Fig. 3, i.e., 840th frame of the first image sequence, we set background velocity as 150, 250, 350 and corresponding results of classic TQD model and the improved TQD model are shown in Fig. 7. As it is shown in Fig. 7, for classic TQD model, detection rate will increase as the rise of projection threshold while the normalized number of the detected points will decrease. However, for the improved TQD model, although the normalized number of detected points will decrease as the increase of projection, detection rate shows no significant change. More precisely, detection rates of the improved TQD model are close to 1 in despite of background and projection threshold. In general, we hope to obtain a higher detection rate at relatively low projection threshold, because the number of detected points can also reach a higher value at this time. Higher number of detected points always means that



Fig. 8. The 840th frame of the second image sequence. The red arrow and  $V_B$  denote motion direction and velocity of background, respectively.



Fig. 10. The 840th frame of the third image sequence. The red arrow and  $V_B$  denote motion direction and velocity of background, respectively.

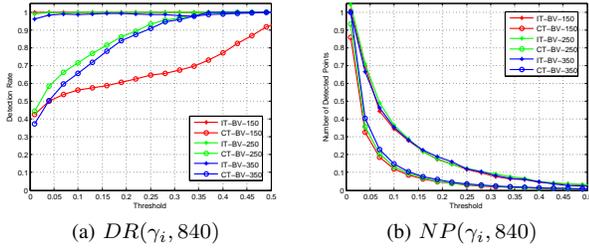


Fig. 9. Detection rate  $DR(\gamma_i, 840)$  and the normalized number of detected points  $NP(\gamma_i, 840)$  of the 840th frame of the second image sequence.

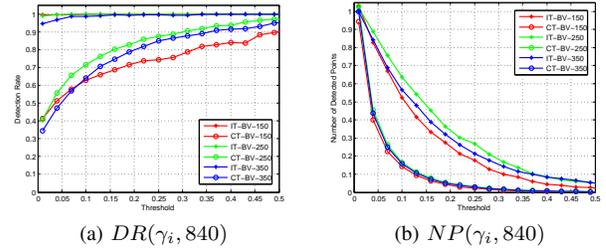


Fig. 11. Detection rate  $DR(\gamma_i, 840)$  and the normalized number of detected points  $NP(\gamma_i, 840)$  of the 840th frame of the third image sequence.

motion direction of the background can be inferred in a wider receptive field. However, as we can see from Fig.7, detection rates of classic TQD model are much lower than that of the improved TQD model at relatively low projection threshold in despite of background velocity.

The second and the third image sequences were also used to evaluate detection performance of these two models. The 840th frame of these two image sequences are presented in Fig. 8 and Fig. 10, respectively. In Fig. 8, background is in leftward motion while in Fig. 10, background is in rightward motion. Relevant results are shown in Fig. 9 and Fig. 11. As we can see from Fig. 9 and Fig. 11, variation trends of curves do not show significantly different from that of Fig. 7. Therefore, similar conclusion can be obtained from Fig. 9 and Fig. 11. That is, optic-flow perception performance of the improved TQD model is much better than that of classic TQD model.

#### IV. CONCLUSION

In this paper, a max operation mechanism is proposed to simulate physiological properties of a newly-identified intermediate neuron, Tm9. The functionality of Tm9 neuron was not reflected in previous correlation models, such as EMD and TQD. This max operation mechanism which acts on ON and OFF signals after signal rectification, is able to improve detection performance of classic TQD model in wide-field motion perception. Synthetic visual stimuli experiments demonstrate that this max operation mechanism can help TQD model avoid confusion in model outputs caused by incorrect signal correlation.

#### ACKNOWLEDGMENT

This research was supported by EU FP7-IRSES Project EYE2E (269118), LIVCODE (295151), HAZCEPT (318907), HORIZON project STEP2DYNA (691154) and ENRICHME (643691).

#### REFERENCES

- [1] J. L. Fox and M. A. Frye, "Figure-ground discrimination behavior in drosophila. ii. visual influences on head movement behavior," *Journal of Experimental Biology*, vol. 217, no. 4, pp. 570–579, 2014.
- [2] Y.-J. Lee, H. O. Jönsson, and K. Nordström, "Spatio-temporal dynamics of impulse responses to figure motion in optic flow neurons," *PLoS one*, vol. 10, no. 5, p. e0126265, 2015.
- [3] B. Hassenstein and W. Reichardt, "Systemtheoretische analyse der zeit-, reihenfolgen- und vorzeichenauswertung bei der bewegungsperzeption des rüsselkäfers chlorophanus," *Zeitschrift für Naturforschung B*, vol. 11, no. 9-10, pp. 513–524, 1956.
- [4] H. Eichner, M. Joesch, B. Schnell, D. F. Reiff, and A. Borst, "Internal structure of the fly elementary motion detector," *Neuron*, vol. 70, no. 6, pp. 1155–1164, 2011.
- [5] Q. Fu, S. Yue *et al.*, "Modeling direction selective visual neural network with on and off pathways for extracting motion cues from cluttered background," 2017.
- [6] Y. E. Fisher, J. C. Leong, K. Sporar, M. D. Ketkar, D. M. Gohl, T. R. Clandinin, and M. Silies, "A class of visual neurons with wide-field properties is required for local motion detection," *Current Biology*, vol. 25, no. 24, pp. 3178–3189, 2015.
- [7] R. Behnia, D. A. Clark, A. G. Carter, T. R. Clandinin, and C. Desplan, "Processing properties of on and off pathways for drosophila motion detection," *Nature*, vol. 512, no. 7515, pp. 427–430, 2014.
- [8] M. Joesch, B. Schnell, S. V. Raghu, D. F. Reiff, and A. Borst, "On and off pathways in drosophila motion vision," *Nature*, vol. 468, no. 7321, pp. 300–304, 2010.
- [9] H. Wang, J. Peng, and S. Yue, "Bio-inspired small target motion detector with a new lateral inhibition mechanism," in *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE, 2016, pp. 4751–4758.
- [10] S. D. Wiederman, P. A. Shoemaker, and D. C. O'Carroll, "A model for the detection of moving targets in visual clutter inspired by insect physiology," *PLoS one*, vol. 3, no. 7, p. e2784, 2008.
- [11] H. Liu, R. Chellappa, and A. Rosenfeld, "Accurate dense optical flow estimation using adaptive structure tensors and a parametric model," *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1170–1180, 2003.
- [12] L. Kratz and K. Nishino, "Tracking with local spatio-temporal motion patterns in extremely crowded scenes," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 693–700.