

Mimicking Fly Motion Tracking and Fixation Behaviors with a Hybrid Visual Neural Network

Qinbing Fu

Computational Intelligence Lab (CIL)
School of Computer Science
University of Lincoln, Lincoln, UK, LN6 7TS
Email: qifu@lincoln.ac.uk

Shigang Yue

Computational Intelligence Lab (CIL)
School of Computer Science
University of Lincoln, Lincoln, UK, LN6 7TS
Email: syue@lincoln.ac.uk

Abstract—How do animals like insects perceive meaningful visual motion cues involving directional and locational information of moving objects in visual clutter accurately and efficiently? In this paper, with respect to latest biological research progress made in underlying motion detection circuitry in the fly’s preliminary visual system, we conduct a novel hybrid visual neural network, combining the functionality of two bio-plausible, namely the motion and the position pathways, for mimicking motion tracking and fixation behaviors. This modeling study extends a former direction selective neurons model to the higher level of behavior. The motivated algorithms can be used to guide a system that extracts location information of moving objects in a scene regardless of background clutter, using entirely low-level visual processing. We tested it against translational movements in synthetic and real-world scenes. The results demonstrated the following contributions: (1) The proposed computational structure fulfills the characteristics of a putative signal tuning map of the fly’s physiology. (2) It also satisfies a biological implication that visual fixation behaviors could be simply tuned via the position pathway; nevertheless, the motion-detecting pathway improves the tracking precision. (3) Contrary to segmentation and registration based computer vision techniques, its computational simplicity benefits the building of neuromorphic visual sensor for robots.

I. INTRODUCTION

Fast motion tracking is still a pronounced challenge in computer vision and robotic applications nowadays. From biology to computational intelligence, nature has given us a lot of inspirations and solutions for building artificial vision systems. The ability to process visual information in an efficient and accurate manner, becomes more desirable for a practical system in mobile machines, like autonomous robots.

As the result of hundreds of millions of years of evolution, motion vision plays a critically important role for animals’ survival. In recent ten years, much biological progress has been made in revealing the preliminary motion detection pathways in insects, e.g., flies [1]–[8] and locusts [9], [10]. Compared to mammals, insects can handle the complexity of real worlds with a relatively small amount of neurons. However, mapping the underlying mechanisms and circuits to neural processing of the higher level of behavior still challenges scientists.

To simulate insects’ vision, many biologically inspired neural networks have been conducted for varied application areas, like the collision detection in robot navigation (e.g. [11]–[14]), the translational movements perception (e.g. [15],

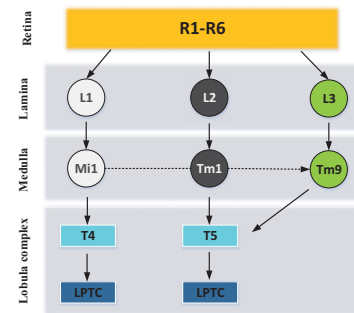


Fig. 1. The preliminary visual processing pipeline throughout four neuropil-layers in the fly’s visual brain: photoreceptors in the retina layer convey motion information to three parallel pathways. Routes starting from L1 and L2 neurons to lobula plate tangential cells (LPTCs) indicate the ON and OFF motion-detecting (motion) pathways respectively. The route starting from L3 to T5 forms the position pathway. Dashed lines denote the putative interactions between interneurons in the motion and the position pathways.

[16]), the small target movements detection [17] and so on. Those modeling studies, all inspired by insects physiology, provide suggestions or solutions for guiding the building of cheap, quick and reliable motion detectors.

Different visual features of a moving object, such as the position and the direction, are crucial to elicit two well-studied behaviors for animals, i.e., the optomotor and the fixation [1]. Motion tracking is vital for animals to possess the ability to extract useful motion cues from visual clutter timely, then evoke advisable behaviors, like the turning response, for maintaining moving targets within their receptive fields. The visual fixation response, first observed in flies, was proposed one of the most important follow-up behaviors after the motion detection [1]: when an object of interest appears in the view, a fly tends to keep it near the center of frontal view, no matter the direction in which the object or the background is moving.

More specifically, in the fly physiology, the tracking and fixation behaviors were demonstrated to be mediated by parallel, the motion and the position visual pathways [1], [4], [7]. In addition, a biological study [1] implicated that the fixation behavior could be tuned by only the position pathway whilst the motion pathway likely corresponds to the optomotor response. However, it appears that both pathways give rise to collaborative effects on shaping the fixation behavior.

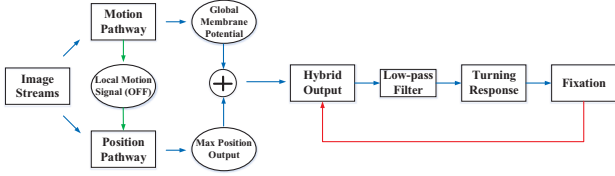


Fig. 2. The schematic illustration of signal processing in the hybrid visual neural network: Blue arrows specify the feed-forward processing flowchart towards the fixation response. Green arrows indicate the local inputs from the motion to the position pathway. A red arrow designates a feedback control.

A schematic signal tuning map is represented in Fig. 1 inspired by the physiological postulate of the fly’s preliminary visual neural network [3], [4], [7], [8]. In general, the motivated framework involves three visual pathways, which are computationally conducted as the motion and the position pathways. Concretely, the ON and OFF parallel pathways make up the motion-detecting (or motion) pathway whilst an extra class of neurons with wide-field properties constitute the position pathway providing location instead of direction information. It is important to state that a relevant modeling study on direction selective neurons was proposed very recently [15], which combines the functionality of ON and OFF pathways with biological plausibility for constructing the proposed motion pathway. In this research, for the first time we extended the former neural network to the higher level of behavior, via incorporating in the framework a neural network realizing the functionality of position pathway. Moreover, we designed a hybrid neural network (Fig. 2) in order to mediate the fixation behavior, by integrating all pathways with a feedback control and more important the interactions (local motion information) between the motion and the position pathways underlying the OFF-motion sensitivity across a wide receptive field in the fly’s visual system [4], [7], [8].

In the following sections, the hybrid visual neural network architecture with algorithms and parameters setting will be presented in Section II. Followed by are the systematic experiments with results and analysis in Section III. Finally we give a conclusion with the future work in Section IV.

II. THE VISUAL MODEL ARCHITECTURE

In this section, we will present the hybrid visual neural network with the motion and the position pathways as depicted in Fig. 2 and 3. It is necessary to clarify the concrete modeling with algorithms of the motion pathway is illustrated fully in a partial research [15], which is briefly introduced in this section. We highlight the functionality of new-built position pathway and the hybrid neural network design for shaping the fixating response. In addition, it is also worth emphasizing that contrary to the traditional tracking strategies like the regression based and search/segmentation based models, the biologically motivated neural network is guided by low-level visual processing that is only interested in motion information with the direction and the magnitude properties. To simplify, the acronyms of visual model components in Fig. 3 and corresponding algorithms are all listed in Table I.

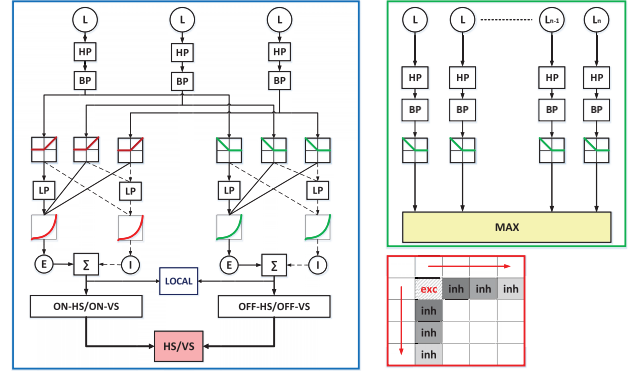


Fig. 3. The illustration of the motion (in the blue box) and the position (in the green box) pathways. The connection of three cells are shown for instance in the motion pathway with more details in [15]. For each local neuron in the medulla and lobula layers, the lateral multi-connections of the same-sign polarity cells along two directions with dynamic delays (in the red box) form the excitation and inhibitions at the starting and adjacent cells in sampling distance respectively. Full-names of model components reference Table. I.

A. The Motion Pathway

Retinal layer - In the motion pathways, the first retina layer involves photoreceptors arranged in a 2D matrix form, the number (n in Fig. 3) of which corresponds to the resolution of input visual streams. Each photoreceptor captures gray-scaled luminance then relays it to a simplified high-pass filter in order to get the luminance change between successive frames:

$$P_{x,y,t} = L_{x,y,t} - L_{x,y,t-1} \quad (1)$$

After that, for each local pixel, we apply a band-pass filter in spatial for mimicking the center-surrounding antagonism found in insects’ visual system. It is represented by a ‘Difference of Gaussians’ (DoG) algorithm, which enhances the motion edge selectivity as suggested in [18], and also removes redundant environmental noise so that maximizing relayed visual information transmission to the following layers [15].

Lamina layer - In the second lamina layer, the first-order interneurons of ON and OFF transient cells encode onset and offset response by luminance increment and decrement respectively, and split band-pass filtered signals into separated ON and OFF channels forming the starting points of ON and OFF motion pathways. Each photoreceptor corresponds to a pairwise ON and OFF cells. Such mechanisms are expressed by the ‘half-wave’ rectifiers as follows:

$$\begin{aligned} LA_{x,y,t}^{ON} &= (P_{x,y,t} + |P_{x,y,t}|)/2, \\ LA_{x,y,t}^{OFF} &= (|P_{x,y,t} - P_{x,y,t}|)/2 \end{aligned} \quad (2)$$

In addition, we employ a ‘Lipetz transfer’ function to transform the analog value of luminance to the membrane potential in a roughly logarithmic manner [15].

Medulla layer - In the third medulla layer, the signals in either polarity ON or OFF pathways form two kinds of flows - the excitation and the inhibition. As depicted in Fig. 3, the computational form of each pairwise combination of same-sign cells reconciles that of the symmetric Reichardt detectors.

TABLE I
THE ACRONYMS OF VISUAL MODEL COMPONENTS

P	photoreceptor	HS	horizontally sensitive system
HP	high-pass filter	VS	vertically sensitive system
LP	low-pass filter	LOCAL	local motion detector
BP	band-pass filter	L	gray-scale luminance
MAX	max operation	E/I	excitation/inhibition
LA	lamina layer	TR	turning response
ME	medulla layer	MP/PP	motion/position pathway output

Importantly, we build the temporal dynamics within the dual-pathways, i.e., the delay represented by the first-order low-pass filtering depends on the sampling distance. More specifically, we put forth the longest time span in the combination at the shortest sampling distance, and then gradually reduce it as the distance growing along both directional multi-connections. We take one cell in the horizontally sensitive system as an instance to show the forming of excitation and inhibition and their linear competition:

$$ME_{x,y,t}^{HS} = \sum_{i=d}^{d \cdot N_{con}} (D_{x,y,t} \cdot LA_{x+i,y,t} - W_i \cdot D_{x+i,y,t} \cdot LA_{x,y,t})$$

where, $d/dt\{D_{x,y,t}\} = 1/\tau_s(LA_{x,y,t} - D_{x,y,t})$ (3)

where N_{con} denotes the number of connected polarity cells, d is the increasing step in sampling distance. τ_s indicates the dynamic time constant in milliseconds. W_i is a bias to form a partially balanced model with stronger response to the preferred over null directional motion. Similarity for the computations in the vertically sensitive system and for both ON and OFF motion-detecting pathways.

There are also local motion detectors (*LOCAL* in Fig. 3) combining local excitations from ON and OFF channels in a supralinear manner, with regard to the computation in [15]. More importantly, in the proposed hybrid neural network, they are additional inputs to the position pathway indicating helpful local motion information as shown in Fig. 1 and 2.

Lobula complex layer - In the final layer of motion pathway, four groups of LPTCs linearly integrate all the directionally specific excitations of both ON and OFF pathways constituting the global membrane potential, then exponentially transfer them as the HS and VS outputs towards the hybrid pathway [15]. Positive outputs of the motion pathway will be generated stimulated by the preferred-directional (front-to-back and downward) motion while negative outputs via the null-directional (back-to-front and upward) motion.

B. The Position Pathway

As illustrated in Fig 2 and 3, in parallel with the motion pathway, the first layer of the position pathway shares the same input of visual streams, modeled by a 2D array of photoreceptors as well. On the contrary, there are no lateral interactions between neighboring interneurons in the position pathway. We also employ a high-pass filtering process expressed as:

$$P'_{x,y,t} = \sigma_{hp} \cdot (P'_{x,y,t-1} + L_{x,y,t} - L_{x,y,t-1}),$$

where, $\sigma_{hp} = \tau_1/(\tau_1 + \tau_i)$ (4)

where τ_1 denotes a time constant in milliseconds and τ_i indicates the time interval between successive frames. After that, the filtered signals also go through spatial band-pass filtering represented by the DoG algorithm, as well as the ‘half-wave’ rectifying, pertaining to the OFF-motion edges selectivity along with filtering out onset responses. A max operation subsequently combines the location of maximum response occurs in the position pathway, with the output of maximum modulus from local motion detectors (abbreviated as LM) of both HS and VS systems of the motion pathway:

$$LM_{x',y',t} = \max_{(x,y) \in \Omega(max_x, max_y)} \|LM'_{x,y,t}\|_2,$$

where, $LM'_{x,y,t} = LM_{x,y,t,HS}^2 + LM_{x,y,t,VS}^2$ (5)

The output is the maximum local motion signal with position information (x', y') in a neighboring field $\Omega(max_x, max_y)$ centered by (max_x, max_y) of the maximum offset response elicited by the position pathway, and the radius of the field corresponds to the max sampling distance ($d \cdot N_{con}$) in the motion pathway. It is important to state that in this research, we only demonstrate the motion tracking in horizontal direction using x' to activate the position pathway via an exponential transformation as follows:

$$PP_t^{HS} = \begin{cases} 1/e^{-((x' - x^{vc})/(C/4))^2} - 1, & \text{if } x' - x^{vc} \geq 0 \\ 1 - 1/e^{-((x' - x^{vc})/(C/4))^2}, & \text{else} \end{cases}$$
 (6)

where x^{vc} is the horizontal location of image view center (vc) and C is the number of columns in the receptive field.

C. The Hybrid Pathway

In the hybrid pathway, as illustrated in Fig. 2, the separated outputs - the directionally membrane potential from the motion pathway and the max location output from the position pathway are integrated to form the hybrid turning response in a purely linear manner:

$$TR_t = \sigma_m \cdot MP_t^{HS} + \sigma_p \cdot PP_t^{HS}, \quad \frac{d\{TR'_t\}}{dt} = \frac{1}{\tau_2}(TR_t - TR'_t)$$
 (7)

where σ_m and σ_p are two gain factors. The output of hybrid pathway - the ‘turning response’ is also delayed by a low-pass filtering with a time constant τ_2 in milliseconds. Taken this response of behavioral level, we can simulate updating of the fly’s view center via:

$$x^{vc} = x^{vc} + TR'_t$$
 (8)

Therefore, we demonstrate that a successful visual fixation behavior should satisfy the following condition:

$$\lim_{t \rightarrow t_0} \|x'_t - x_t^{vc}\| \leq \gamma$$
 (9)

where γ is a predefined threshold which is normally set equally as the sampling distance in the motion pathway. As shown in Fig. 2, we also design a quick feedback pathway for the purpose of adjusting the gain factor (σ_p) of the position pathway for more quickly meeting the requirements of fixating:

$$\sigma_p = \sigma_p + \sigma_c, \quad \text{if } |x'_t - x_t^{vc}| > \gamma \ \& \ \frac{d\{|x'_t - x_t^{vc}\}}{dt} \geq 0$$
 (10)

TABLE II
THE VISUAL MODEL PARAMETERS SETTING

Name	Value	Name	Value	Name	Value
N_{con}	8	d	2	W_i	0.89
σ_c	10	τ_s	5 ~ 200	C	adaptable
γ	N_{con}	σ_m	3	σ_p	10
τ_1	20	τ_2	10	τ_i	adaptable

D. Parameters Setting

The chosen parameters in Table II were decided empirically based on consideration of the optimization of functionality and implementation of proposed framework for fast and precise motion tracking. It possesses a feed-forward low-level visual processing structure without any parameters training methods. The adaptable parameters correspond to the resolution and the sampling frequency of input visual streams. More detailed parameters set-up of the motion pathway is suggested in a partial research [15]. Importantly, a shortcoming of the visual model is that the combination of gain factors in the hybrid pathway greatly influences its fixating performance: increasing either gain factors, especially that of the position pathway, could accelerate the process to fit the requirements of a successful fixation; however, as the neural network is also sensitive to the velocity of translational motion, it may also bring about fluctuations of the relative position between the moving objects and the view center. Therefore, a robust learning method is badly needed in the near future.

We hope the follow-up experiments will provide useful conclusions or suggestions for designing artificial motion tracking system, and exploring the potential of biologically neural networks for utility in intelligent robots.

III. EXPERIMENTAL EVALUATION

In this section, we will present the systematic off-line experiments, which can be categorized into two kinds of tests, i.e., challenged by translational motion embedded in synthetic and real-world scenes respectively¹. All the input visual streams were converted to the gray-scale with the resolution of $320 \cdot 240$ and $432 \cdot 240$ for synthetic and real physical scenes respectively. We show experiments results via the outputs of relative position between the translating object(s) and the simulated fly’s view center (VC), during each tracking and fixating process. We also investigate and compare the different fixation responses between neural networks with the motion-blocked and the intact-pathways systems in the synthetic tests motivated by a biological study [1].

A. The Synthetic Visual Stimuli Tests

In the first kind of tests, we tested the proposed model against synthetic translational movements on both horizontal directions. The visual stimuli include a single darker or lighter object translating and elongating (and shortening) against the clean background, as well as two gray-scaled objects moving

¹An attached demo video shows all the off-line visual stimuli with results.

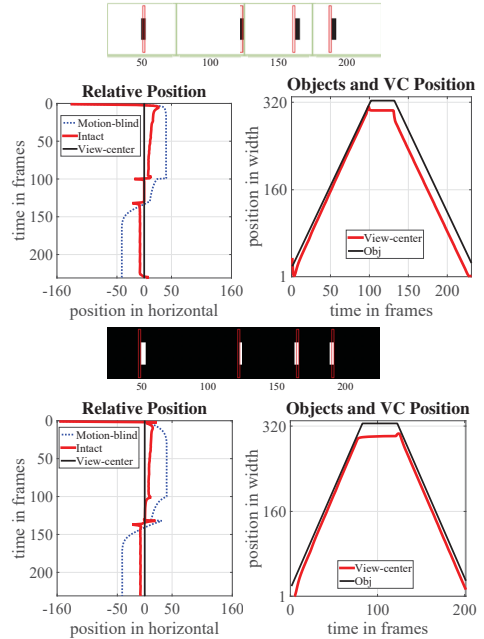


Fig. 4. The outputs represented by horizontal positions of the simulated view-center (VC) and the translating object: some snapshots with the VC (represented by the red-frame) are shown at the top of each pairwise results. Both the motion-blocked and intact-pathways neural networks are tested.

concurrently at the same or different constant speeds. There was no background noise in those simulated scenarios.

The results illustrated in Fig. 4 and 5 allow the following conclusions: first, when tested by either darker or lighter objects translating in either horizontal directions, the tracking and fixation behaviors elicited by both neural networks with the motion pathway blocked and the intact pathways are well achieved. During each tracking process, the outputs of relative position quickly converge within a small range, i.e., the simulated fly’s view center is dramatically guided to close in the position of translating object matching the fixation behavior perfectly. In addition, it appears that the visual model with complete pathways achieves more precise tracking performance with relatively smaller relative positions.

Interestingly, when challenged by two dark objects translating simultaneously at an identical speed-level, the updated view center of fixation behavior is always following the darker object movements. The results reveal the contrast sensitivity of the proposed framework with the preference to stronger offset response caused by darker motion. In addition, when the translating objects have different moving speeds, the view center of fixation response initially accompanies the darker object moving and then quickly jumps to the less-darker one, once the darker object stops moving. Importantly, the results also provide a profound implication that the motion-detecting is essential for the proposed hybrid visual neural network to elicit the fixation behavior.

With similar ideas, we also examine its performance against two dark objects elongating and shortening with only a single edge moving (Fig. 6). The visual fixation responses

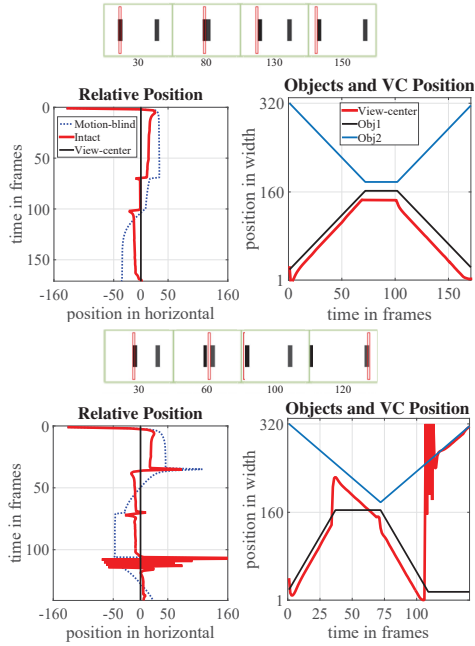


Fig. 5. The outputs under synthetic translation movements of two darker objects moving at the same and different constant-speeds respectively.

are well realized as expected. However, the results point out the offset-response selectivity of the proposed framework: the simulated fly’s view center is always following the OFF-edge elongating (offset) rather than shortening (onset). Furthermore, the statistics shown in Fig. 7 demonstrate that the motion-blocked system represents similar turning response compared to that elicited by the intact-pathways system during motion tracking, blocking the motion pathway nevertheless leads to larger outputs of relative position at all tested translating velocities, i.e., the tracking precision is much reduced.

B. The Real-world Visual Stimuli Tests

In the second kind of off-line tests, we inspected its performance challenged by real-world translational movements in visual clutter. Compared to the synthetic tests, there were also much environment noise in the real physical scenes. The visual stimuli involve the person-crossing and five gray-scaled objects translating, all embedded in the busy background. Satisfactory results (Fig. 8) demonstrate the proposed visual model successfully mimics the fly motion tracking and fixation behaviors regardless of the cluttered background and environmental noise. It appears that without translational motions within the receptive field, the simulated view center is rigorously affected by the background noise, wandering intensely within the receptive field. However, if translating objects appear, the proposed ‘motion’ sensitive neural networks can guide the simulated fly’s view center to follow translational movements, both in a timely and reliable manner.

IV. CONCLUDING REMARKS

In this paper, we propose a hybrid visual neural network inspired by the fly’s preliminary vision system, mimicking

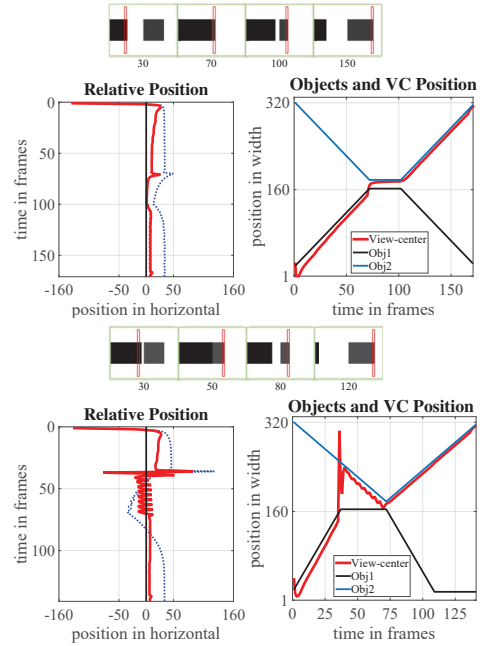


Fig. 6. The outputs stimulated by synthetic elongation and shortening movements of two darker objects at the same and different speeds respectively.

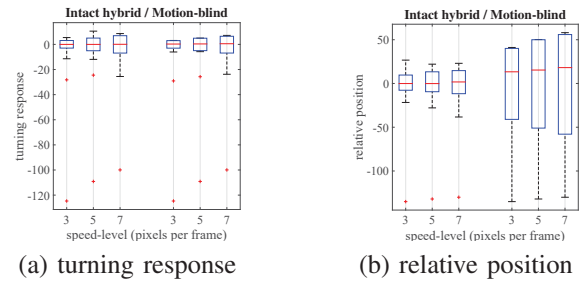


Fig. 7. The statistical results of the turning response (a) and the relative position (b) during each entire tracking course for both the intact-pathways and motion-blocked neural networks, tested by a single darker object translating at three constant velocity levels separately.

motion tracking and fixation behaviors. The motivated framework extends a former motion-detecting model to the higher level of behavior. Its feed-forward structure is fully guided by low-level visual processing strategies. The motion and the position pathways explored by biologists are computationally conducted to provide parallel outputs. The visual fixation behavior is shaped by a hybrid pathway, integrating the outputs of both pathways and eliciting the turning response in order to simulate the updating of view-center during fixating along with translational movements. Informative off-line tests results demonstrate the proposed neural networks match the underlying functionality of fly’s visual pathways perfectly, which can cope with motion tracking in a fast and reliable manner, even against busy backgrounds. Moreover, the results well reconcile with a biological finding that the position pathway contributes more significantly in mediating the visual fixation whilst the motion pathway rigorously improves the precision and efficiency of tracking.

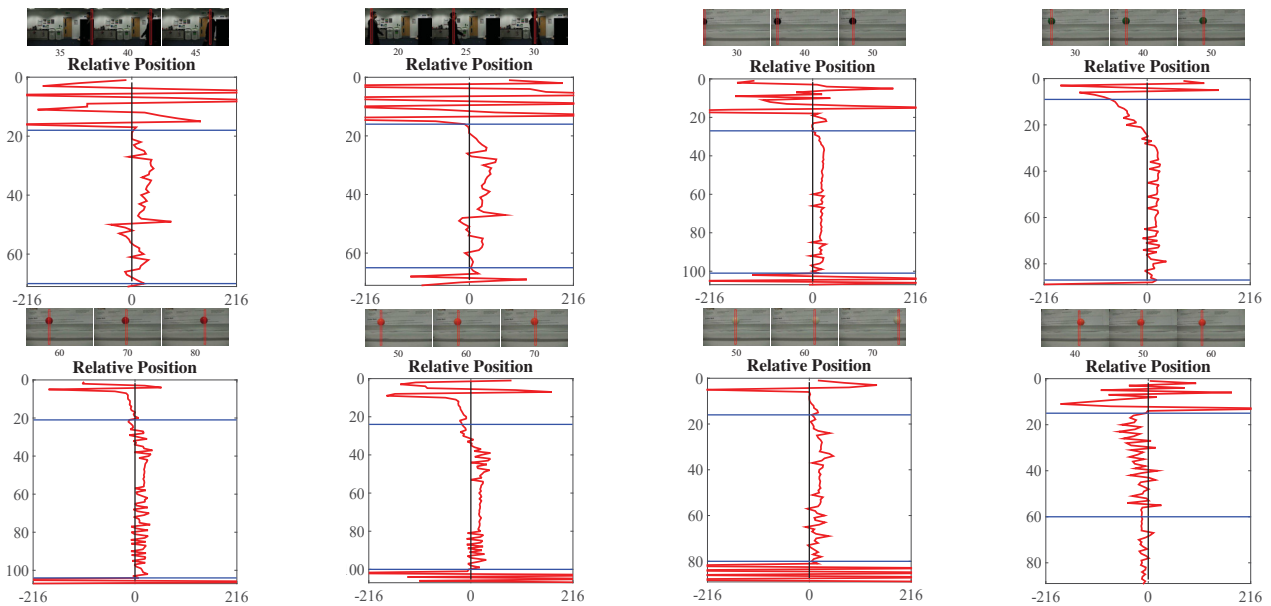


Fig. 8. The outputs of the relative position under the real-world tests: two horizontal blue lines specify the time window between the target appearing and leaving the receptive field. X-axis denotes the horizontal relative-position and Y-axis designates the time sequence. Some snapshots with labeled frames and the depicted view-centers are shown at the top of each result. Only the neural network with intact pathways is tested.

This modeling study also opens several directions for future research. First, its computational simplicity and robustness also have great potential to build the neuromorphic sensor for utility in robotic vision guiding real-time tasks of translational movements detection and tracking mixed with multiple robot agents and/or humans. Second, we will investigate its internal characteristics in real-time motion tracking, and compare its performance with other state-of-the-art tracking methods. Moreover, we also expect to combine its functionality with the biologically visual collision detectors mimicking insects' motion detection in more complex scenarios.

ACKNOWLEDGMENT

This work was partly supported by the grants of EU Horizon 2020 project STEP2DYNA(691154).

REFERENCES

- [1] A. Bahl, G. Ammer, T. Schilling, and A. Borst, "Object tracking in motion-blind flies," *Nat Neurosci*, vol. 16, no. 6, pp. 730–738, 2013.
- [2] A. Borst and M. Helmstaedter, "Common circuit design in fly and mammalian motion vision," *nature neuroscience*, vol. 18, pp. 1067–1076, 2015.
- [3] A. Borst and T. Euler, "Seeing things in motion: models, circuits, and mechanisms," *Neuron*, vol. 71, no. 6, pp. 974–994, 2011.
- [4] 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.
- [5] M. Joesch, B. Schnell, S. V. Raghun, D. F. Reiff, and A. Borst, "On and off pathways in drosophila motion vision," *Nature*, vol. 468, no. 7321, pp. 300–304, 2010.
- [6] M. S. Maisak, J. Haag, G. Ammer, E. Serbe, M. Meier, A. Leonhardt, T. Schilling, A. Bahl, G. M. Rubin, A. Nern, B. J. Dickson, D. F. Reiff, E. Hopp, and A. Borst, "A directional tuning map of drosophila elementary motion detectors," *Nature*, vol. 500, no. 7461, pp. 212–216, 2013.
- [7] J. Rister, D. Pauls, B. Schnell, C.-Y. Ting, C.-H. Lee, I. Sinakevitch, J. Morante, N. J. Strausfeld, K. Ito, and M. Heisenberg, "Dissection of the peripheral motion channel in the visual system of drosophila melanogaster," *Neuron*, vol. 56, no. 1, pp. 155–170, 2007.
- [8] N. Vogt and C. Desplan, "The first steps in drosophila motion detection," *Neuron*, vol. 56, no. 1, pp. 5–7, 2007.
- [9] S. Wernitznig, F. C. Rind, P. Polt, A. Zankel, E. Pritz, D. Kolb, E. Bock, and G. Leitinger, "Synaptic connections of first-stage visual neurons in the locust schistocerca gregaria extend evolution of tetrad synapses back 200 million years," *J Comp Neurol*, vol. 523, no. 2, pp. 298–312, 2015.
- [10] F. C. Rind, S. Wernitznig, P. Polt, A. Zankel, D. Gutl, J. Sztarker, and G. Leitinger, "Two identified looming detectors in the locust: ubiquitous lateral connections among their inputs contribute to selective responses to looming objects," *Scientific Reports*, p. 35525, 2016.
- [11] S. Yue and F. C. Rind, "A collision detection system for a mobile robot inspired by locust visual system," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2005, Conference Proceedings, pp. 3843–3848.
- [12] Q. Fu, C. Hu, and S. Yue, "Bio-inspired collision detector with enhanced selectivity for ground robotic vision system," in *British Machine Vision Conference 2016*, 2016, Conference Proceedings, pp. 6.1–6.13.
- [13] S. Bermudez i Badia, U. Bernardet, and P. F. Verschure, "Non-linear neuronal responses as an emergent property of afferent networks: a case study of the locust lobula giant movement detector," *PLoS Comput Biol*, vol. 6, no. 3, p. e1000701, 2010.
- [14] Q. Fu, C. Hu, T. Liu, and S. Yue, "Collision selective lgmds neuron models research benefits from a vision-based autonomous micro robot," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2017, Conference Proceedings, pp. 3996–4002.
- [15] Q. Fu and S. Yue, "Modeling direction selective visual neural network with on and off pathways for extracting motion cues from cluttered background," in *The 2017 International Joint Conference on Neural Networks*, 2017, Conference Proceedings, pp. 831–838.
- [16] S. Yue and F. C. Rind, "Visual motion pattern extraction and fusion for collision detection in complex dynamic scenes," *Computer Vision and Image Understanding*, vol. 104, no. 1, pp. 48–60, 2006.
- [17] 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.
- [18] D. A. Clark, L. Bursztyn, M. A. Horowitz, M. J. Schnitzer, and T. R. Clandinin, "Defining the computational structure of the motion detector in drosophila," *Neuron*, vol. 70, no. 6, pp. 1165–1177, 2011.