Bio-inspired small target motion detector with a new lateral inhibition mechanism

Conference Paper · July 2016

DOI: 10.1109/IJCNN.2016.7727824

3 authors, including:

Hongxin Wang
University of Lincoln
28 PUBLICATIONS **165** CITATIONS
SEE PROFILE

Shigang Yue
University of Lincoln
231 PUBLICATIONS **3,013** CITATIONS
SEE PROFILE

Bio-inspired Small Target Motion Detector with a New Lateral Inhibition Mechanism

Hongxin Wang
School of Computer Science
University of Lincoln
Lincoln, LN6 7TS, UK
hongxin.w@hotmail.com

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*—In nature, it is an important task for animals to detect small targets which move within cluttered background. In recent years, biologists have found that a class of neurons in the lobula complex, called STMDs (small target motion detectors) which have extreme selectivity for small targets moving within visual clutter. At the same time, some researchers assert that lateral inhibition plays an important role in discriminating the motion of the target from the motion of the background , even account for many features of the tuning of higher order visual neurons. Inspired by the finding that complete lateral inhibition can only be seen when the motion of the central region is identical to the motion of the peripheral region, we propose a new lateral inhibition mechanism combined with motion velocity and direction to improve the performance of ESTMD model (elementary small target motion detector). In this paper, we will elaborate on the biological plausibility and functionality of this new lateral inhibition mechanism in small target motion detection.**

I. Introduction

Discriminating objects which amidst in the cluttered moving background is an important task for animals searching for and tracking prey or conspecifics. However, due to the fact that the spatial resolution of most insects eyes is relatively coarse [1], it is more challenging for insects to complete this task. Actually, even the best insect eyes those of predatory dragonflies, have a spatial resolution of only about $0.25 - 0.5^{\circ}$ in the acute zone [2], [3]. Therefore, the research in the sophisticated visual mechanisms of insects that underlie rapid detection and tracking of targets has received increasing attention.

Recent researches have found that a number of neurons located in the lobula complex, called small target motion detectors (STMDs), which are selective for small moving targets. According to the proposed results [4], [5], [6], [7], these neurons (STMDs) are strongly excited by the motion of small, black targets (0.8 square) within a fronto-dorsal receptive field. And larger targets ($> 3^{\circ}$) elicit weaker responses which fell to spontaneous levels for target subtending more than 10°. Besides, it has been shown that this size selectivity is independent of the location and shape of the receptive field [7]. As mentioned in [6], [7], lateral inhibition mechanism could play an important role in visual processing system of insects and is helpful to shape the response tuning to small targets. However, it was still not clear how to implement it.

Lateral inhibition is a pervasive biological mechanism in visual processing system of insects. Recent research on lateral inhibition [8] found that some neurons respond to local motion on the retina only if the motion trajectory differs from that in a large surrounding region. More precisely, by recording the spike trains of ganglion cells, the author found that these neurons respond when an object in its receptive field centre moves relative to the background, but is almost completely suppressed when the object moves together with the background. In addition to this, results proposed in [9] also demonstrate that lateral inhibition is velocity- and direction-selective. In fact, lateral inhibition reaches its maximum when the object and the distracter target move with the same velocity and direction in certain range but remains weak or silent when the object and the distracter target move with the same velocity to different motion directions.

Although ESTMD (elementary small target motion detector) proposed by [7] contains two lateral inhibition mechanisms which are located in lamina layer and medulla layer respectively, these two lateral inhibition mechanisms do not seem to be in accord with aforementioned biological findings completely. Actually, according to these two lateral inhibition mechanisms, the object receives the same amount of lateral inhibition no matter whether there exist relative motion between the object and background or not. The motion of the object can be strongly attenuated by the motion of background even when object motion is different from background motion. For this reason, the detection performance of ESTMD is unstable especially when the object go through the moving cluttered background.

Inspired by aforementioned biological findings on velocity and motion directions determined lateral inhibition phenomenon, we improve ESTMD model proposed by [7] with a new lateral inhibition mechanism which takes velocity and motion direction into consideration. And we will show that compared with the lateral inhibition proposed in [7], our new lateral inhibition can improve the performance of target detection.

II. Modeling

Based on the previous ESTMD model [7], we proposed a new model with a new lateral inhibition mechanism. Inspired

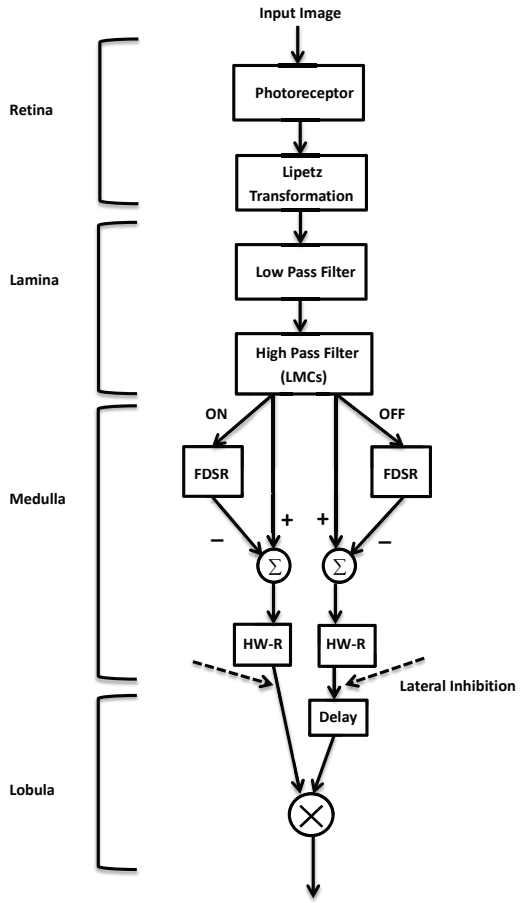

Fig. 1. Schematic of our proposed model

by the biological visual process in the fly, our proposed model is composed of four neural layers: retina, lamina, medulla and lobula. Fig 1 gives the schematic of our proposed model, and will be elaborated in the following paper.

A. Retina Layer

The retina layer is composed of $M \times N$ photoreceptors arranged in a matrix form with M rows and N columns. Each photoreceptor corresponds to a pixel point and receives luminance or gray levels from successive images.

Let $I_{ij}(t)$ denote the intensity of pixel (i, j) for an image frame at time t and we mimic the spatial blur of fly optics by using $I_{ij}(t)$ convolves with a Gaussian convolution mask. That is,

$$L_{ij}(t) = \sum_{u=-1}^{1} \sum_{v=-1}^{1} w_{uv} \cdot I_{i+u,j+v}(t) \tag{1}$$

where

$$W = \begin{pmatrix} \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \end{pmatrix}. \tag{2}$$

After the spatial blur, photoeceptors transform the input luminance to membrane potential. This process can be implemented by using Lipetz function with the exponent u set as 0.7.

$$P_{ij}(t) = \frac{(L_{ij}(t))^u}{(L_{ij}(t))^u + (L^c_{ij}(t))^u}. \tag{3}$$

$L^c_{ij}(t)$ in the equation (3) is the low-pass filtered version of $L_{ij}(t)$ and satisfies the following relationship

$$\frac{dL^c_{ij}(t)}{dt} = \frac{1}{\tau_1}(L_{ij}(t) - L^c_{ij}(t)) \tag{4}$$

where τ_1 is the time constant.

B. Lamina Layer

The output of retina layer ($P_{ij}(t)$) is provided to lamina layer with a slight delay,

$$\frac{dx_{ij}(t)}{dt} = \frac{1}{\tau_2}(P_{ij}(t) - x_{ij}(t)) \tag{5}$$

where τ_2 is the time constant.

Then, the delayed signal $x_{ij}(t)$ is used as the input of large monopolar cells (LMCs) located in lamina layer. On the basis of previous research in large monopolar cells, it is believed that LMCs can remove redundant information and maximize information transmission [10], [11]. Generally, the functionality of LMCs can be described as following equation:

$$\frac{dX^{LMC}_{ij}(t)}{dt} = \frac{1}{\tau_3}(x_{ij}(t) - X^{LMC}_{ij}(t)) \tag{6}$$

$$Y^{LMC}_{ij} = x_{ij}(t) - X^{LMC}_{ij}(t) \tag{7}$$

where Y^{LMC}_{ij} is the output of LMCs and $X^{LMC}_{ij}(t)$ is the first-order low-pass filtered version of $x_{ij}(t)$ while τ_3 is the time constant.

C. Medulla Layer

LMCs located in the lamina layer provide their output to medulla layer. In the first part of medulla layer, the output of LMCs ($Y^{LMC}_{ij}(t)$) is separated into ON and OFF channels. This process can be expressed in the language of mathematics, i.e.

$$Y^{ON}_{ij}(t) = (Y^{LMC}_{ij}(t) + |Y^{LMC}_{ij}(t)|)/2, \tag{8}$$

$$Y^{OFF}_{ij}(t) = (Y^{LMC}_{ij}(t) - |Y^{LMC}_{ij}(t)|)/2 \tag{9}$$

where $Y^{ON}_{ij}(t)$ and $Y^{OFF}_{ij}(t)$ are the signal of ON and OFF channels, respectively.

For approximating plausible biophysical mechanisms, each independent channel (ON or OFF channel for every pixel (i, j)) is formed into an "adaptation state" through the application of a non-linear low pass filter with a fast depolarizing, slow repolarizing characteristic [12]. This FDSR (fast depolarization, slow repolarization) mechanism is able to suppress rapidly changed texture information and enhance novel contrast change.

We denote $S^{ON}_{ij}, S^{OFF}_{ij}$ as the signal of ON and OFF channels after FDSR respectively, then

$$\frac{dS_{ij}^{ON}(t)}{dt} = \begin{cases} \frac{1}{\tau_{fast}}(S_{ij}^{ON}(t) - Y_{ij}^{ON}(t)), \\ \qquad \text{if } Y_{ij}^{ON}(t) > Y_{ij}^{ON}(t-1) \\ \frac{1}{\tau_{slow}}(S_{ij}^{ON}(t) - Y_{ij}^{ON}(t)), \\ \qquad \text{if } Y_{ij}^{ON}(t) \leq Y_{ij}^{ON}(t-1) \end{cases} \tag{10}$$

where τ_{fast} and τ_{slow} are the time constant and satisfy $\tau_{fast} < \tau_{slow}$.

Similarly, we have

$$\frac{dS_{ij}^{OFF}(t)}{dt} = \begin{cases} \frac{1}{\tau_{fast}}(S_{ij}^{OFF}(t) - Y_{ij}^{OFF}(t)), \\ \qquad \text{if } Y_{ij}^{OFF}(t) > Y_{ij}^{OFF}(t-1) \\ \frac{1}{\tau_{slow}}(S_{ij}^{OFF}(t) - Y_{ij}^{OFF}(t)), \\ \qquad \text{if } Y_{ij}^{OFF}(t) \leq Y_{ij}^{OFF}(t-1) \end{cases} \tag{11}$$

Then, the filtered signal $S_{ij}^{ON}, S_{ij}^{OFF}$ are subtractive by the original signal $Y_{ij}^{ON}(t), Y_{ij}^{OFF}(t)$,

$$F_{ij}^{ON}(t) = Y_{ij}^{ON}(t) - S_{ij}^{ON}(t) \tag{12}$$

$$F_{ij}^{OFF}(t) = Y_{ij}^{OFF}(t) - S_{ij}^{OFF}(t) \tag{13}$$

where $F_{ij}^{ON}(t), F_{ij}^{OFF}(t)$ are the output of medulla layer.

After FDSR mechanism, $F_{ij}^{ON}(t), F_{ij}^{OFF}(t)$ are provided to a half wave rectifier(HW-R). We denote the signal of ON and OFF channel after half wave rectification as $HW_{ij}^{ON}(t), HW_{ij}^{OFF}(t)$, then

$$HW_{ij}^{ON}(t) = \max(0, F_{ij}^{ON}(t)) \tag{14}$$

$$HW_{ij}^{OFF}(t) = \max(0, F_{ij}^{OFF}(t)). \tag{15}$$

For simplicity, we still use $F_{ij}^{ON}(t), F_{ij}^{OFF}(t)$ to denote the signal after half wave rectification. Following this, S.D.Wiederman,et al. [7] proposed a second-order local inhibitory interactions between the same channel polarity. That is, central ON channel ($F_{ij}^{ON}(t)$) is subtractively inhibited by surrounding ON channels and similarly for OFF channels ($F_{ij}^{OFF}(t)$). But this lateral inhibition mechanism seems to be inconsistent with proposed findings [8], [9]. According to this lateral inhibition mechanism proposed by [7], central ON channel can be inhibited by surrounding ON channels even when there is differential motion between central region and peripheral region. However, recent biological findings [8], [9] demonstrate the greatest lateral inhibition can only be seen when central motion is identical to peripheral motion. If there is differential motion between central region and peripheral region, lateral inhibition is much weaker and even disappear. These biological findings indicae that lateral inhibition is strongly velocity- and direction- selective. Therefore, a more reasonable lateral inhibition mechanism should take motion velocity and direction into consideration.

In the following paper, we propose a new lateral inhibition mechanism which is more in accordance with above-mentioned biological findings. Our proposed lateral inhibition mechanism is based on motion velocity and direction, so we have to calculate velocity vector of each pixel firstly. EMD (elementary motion detector) has been studied extensively and is regarded as a reasonable model to account for why insects can infer the velocity of moving target. However, although EMD (elementary motion detector) can explain the ability of neurons to detect motion and direction of an object, the contrast dependance and velocity dependance of EMDs make their responses ambiguous with respect to a representation of the retinal velocity [13]. Therefore, we adopt a general matching algorithm to calculate motion vector for every pixel. We define the matching criteria as:

$$D = \sum_{x=1}^{m}\sum_{y=1}^{n} |I_{xy}(t) - I_{x+u,y+v}(t-1)| \tag{16}$$

where $m \times n$ is the size of the search window and $I(t)$ is the input image at time t.

The translation vector, namely the motion vector, is obtained when the D value finds its minimum.

$$(u', v') = \arg\min D_{(x,y)}(u, v) \tag{17}$$

where $(u, v) \in \{(u, v)| - R \leq u, v \leq R\}$ and R is the search range.

Using above general matching algorithm, we can obtain motion vector $\mathbf{V_{ij}} = (u_{ij}, v_{ij})$ for each pixel (i, j), where u_{ij}, v_{ij} are horizontal component and vertical component of motion vector, respectively. Then for a input image, we have motion vector matrix U, V defined as

$$U = (u_{i,j})_{M\times N} \tag{18}$$

$$V = (v_{i,j})_{M\times N} \tag{19}$$

For computing the velocity difference between central region and peripheral region, U and V are convolved by H. That is,

$$U^c = U * H \tag{20}$$

$$V^c = V * H \tag{21}$$

where $*$ is convolution operator and

$$H = \begin{pmatrix} b & b & \cdots & b & b \\ b & 0 & \cdots & 0 & b \\ \vdots & \vdots & a & \vdots & \vdots \\ b & 0 & \cdots & 0 & b \\ b & b & \cdots & b & b \end{pmatrix}_{p\times q}, \ a + 4(q-1)b = 0. \tag{22}$$

To explain the role of U^c and V^c, we firstly define a neighbourhood(PR) for pixel (i, j),

$$\begin{aligned} \text{PR} = \{&(i - ro, j - co : j + co), \\ &(i + ro, j - co : j + co), \\ &(i - ro : i + ro, j - co), \\ &(i - ro : i + ro, j + co)\} \end{aligned} \tag{23}$$

where $ro = \text{round}(p/2)$, $co = \text{round}(q/2)$, p, q are decided by the size of small target.

If the motion vector (u_{ij}, v_{ij}) of pixel (i,j) is identical to the motion vector (u_s, v_s) of pixel s in peripheral region (PR),i.e.

$$u_{ij} = u_s, v_{ij} = v_s, \quad \forall s \in \text{PR} \tag{24}$$

then we have $U^c(i,j) = 0, V^c(i,j) = 0$ after U, V convolve with H. Conversely, we can obtain $U^c(i,j) \neq 0$ or $V^c(i,j) \neq 0$ if there is velocity difference between pixel (i,j) and peripheral region. It also should be noted that the bigger velocity difference between (i,j) and its peripheral region(PR), the higher $|U^c(i,j)|, |V^c(i,j)|$. Therefore, $|U^c(i,j)|$ and $|V^c(i,j)|$ can be used as an indicator to determine the existence of velocity difference between central region and peripheral region.

Once we get $U^c = (u^c_{ij})_{M\times N}$ and $V^c = (v^c_{ij})_{M\times N}$, we define w_{ij} by following equation

$$w_{ij}(t) = \sqrt{(u^c_{ij}(t))^2 + (v^c_{ij}(t))^2}. \tag{25}$$

$w_{i,j}(t)$ is able to show the total velocity difference between pixel (i,j) and its neighbourhood by combining $U^c(t)$ and $V^c(t)$. The higher value of $w_{i,j}(t)$ means the bigger velocity difference between pixel (i,j) and its neighbourhood.

Then, we implement our new lateral inhibition mechanism by multiplying $F^{ON}_{ij}(t)$ and $F^{OFF}_{ij}(t)$ by $w_{ij}(t)$, respectively.

$$\tilde{F}^{ON}_{ij}(t) = k_1 F^{ON}_{ij}(t) + k_2 F^{ON}_{ij}(t) w_{ij}(t) \tag{26}$$

$$\tilde{F}^{OFF}_{ij}(t) = k_1 F^{OFF}_{ij}(t) + k_2 F^{OFF}_{ij}(t) w_{ij}(t) \tag{27}$$

where k_1, k_2 are the constant, $\tilde{F}^{ON}_{ij}$ and $\tilde{F}^{OFF}_{ij}$ are the signal of ON and OFF channel after lateral inhibition, respectively.

The reason why we propose this new lateral inhibition mechanism is not only it is more biologically plausible, but also it can inhibit background motion and enhance small target motion effectively. The motivation of our proposed new lateral inhibition mechanism is based on the observation that if a pixel (i,j) belongs to background, its motion vector would be identical to the counterpart of its peripheral region.In this case, the signal of pixel (i,j) should be strongly inhibited. However, if a pixel (i,j) belongs to a independently moving target, then its motion vector would not equal to the motion vector of other pixels located in its peripheral region, except when there is not different motion between the small target and background. In this case, the signal of pixel (i,j) should be enhanced. As we can see from Eq (26) and (27), $\tilde{F}^{ON}_{ij}(t) = k_1 F^{ON}_{ij}(t), k_1 < 1$ when $w_{ij}(t) = 0$. At this moment, central ON channel receives the strongest lateral inhibition from peripheral ON channels. However, when $w_{ij}(t) > 0$, $\tilde{F}^{ON}_{ij}(t) = k_1 F^{ON}_{ij}(t) + k_2 F^{ON}_{ij}(t) w_{ij}(t) > k_1 F^{ON}_{ij}(t)$. At this moment, the lateral inhibition which depends on the value of $w_{ij}(t)$ is much less, so laterally inhibited signal($\tilde{F}^{ON}_{ij}(t)$) is larger than $k_1 F^{ON}_{ij}(t)$. We even can enhance the signal of ON or OFF channel by adjusting parameter k_2 appropriately when $w_{ij}(t) > 0$.

Overall, this kind of lateral inhibition is capable of enhancing the saliency of the moving small target whose size is smaller than $p \times q$ and inhibiting false positives caused by the motion of moving background so as to improve the performance of model.

D. Lobula Layer

In the lobula layer, the delayed OFF channel is correlated with the un-delayed ON channel. A first order low pass filter is used to achieve the delay of OFF channel and the delay time is determined by the size and velocity of the small target.

$$\frac{dLob^{OFF}_{ij}(t)}{dt} = \frac{1}{\tau_4}(\tilde{F}^{OFF}_{ij}(t) - Lob^{OFF}_{ij}(t)) \tag{28}$$

The final output of the lobula layer is

$$O_{ij}(t) = \tilde{F}^{ON}_{ij}(t) \times Lob^{OFF}_{ij}(t) \tag{29}$$

III. Experiment

In this section, we perform a simulation experiment to evaluate the properties of our proposed algorithm. Its performance is compared with that of elementary small target motion detector (ESTMD) proposed by [7]. We firstly use Vision Egg [14] to produce a series of vision stimuli. Vision Egg which is a free, open-source library has been widely used by biologists in biological experiments to study properties of insects' visual system. Due to the fact that spatial frequency and power of natural scenes satisfy a statistical relationship of $1/f^2$ [15], natural images which satisfy the above relationship are used as background in produced vision stimuli.

Fig. 2 shows one frame of the input image sequence. As it can be seen from Fig. 2, a small target located in the red circle is moving from right to left along the horizontal midline. The red left arrow (V_t) denotes the motion direction of small target. In contrast, the cluttered background is moving from left to right and its motion direction is denoted by the red right arrow (V_b). The size and luminance of the moving target are 5×5 and 0, respectively.

Fig. 2. A small moving target amidst in the moving cluttered background

In the following paper, the performance of our proposed algorithm is compared with that of ESTMD model by using the same input stimuli. Fig. 3 is the 182th frame of the input vision stimuli. Fig. 4 shows the output of ESTMD model while Fig. 5 is the output of our proposed algorithm. The white pixels in Fig. 4 and Fig. 5 denote pixels whose output ($O_{ij}(t)$) is larger than a given threshold.

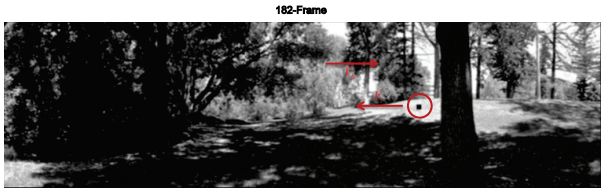

Fig. 3. The 182th frame of the input visual stimuli

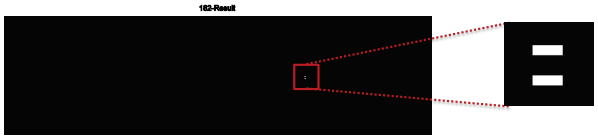

Fig. 4. The output of ESTMD

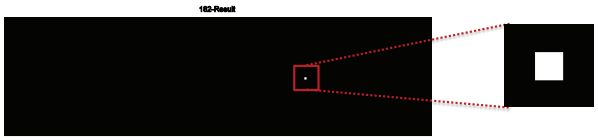

Fig. 5. The output of our proposed algorithm

In Fig. 3, high contrast between the small target and its neighborhood can be seen. In this case, both ESTMD model and our proposed algorithm can detect the small moving target easily and obtain the best detection performance. However, there are still some differences between the outputs of two models. The response of ESTMD model to a small moving target is a pair of moving edges while that of our proposed algorithm is a moving rectangle. The main reason that cause the difference between the outputs of two models is because in ESTMD model [7], [12], the signal receives lateral inhibition from surrounding area after passing the high pass filter located in lamina layer. That is, high-pass filtered signal (Y^{LMC}) convolve with a kernel M,

$$M = \begin{pmatrix} -1/9 & -1/9 & -1/9 \\ -1/9 & 8/9 & -1/9 \\ -1/9 & -1/9 & -1/9 \end{pmatrix}. \tag{30}$$

After convolving with M, the edges of ON region where luminance increase will be enhanced , especially for edges which are parallel to motion direction of the target. However, the center of ON region will be weaken. Similarly, the center of OFF region (where luminance decrease) will be weaken while its edges will be enhanced. In the following medulla and lobula layer, the response of ESTMD model to ON region's (OFF region's) edges remains larger than the response to ON region's (OFF region's) center. Therefore, the final response of ESTMD model to the small target motion is always a pair of moving edges. However, due to the fact that we propose a new lateral inhibition mechanism which is significantly different from the lateral inhibition mechanism used by [7], [12], ON region (OFF region) can preserve rectangular shape after lateral inhibition.

Fig. 6 shows the 189th frame of the input vision stimuli. Fig. 7 and Fig. 8 are the output of ESTMD model and our proposed model, respectively.

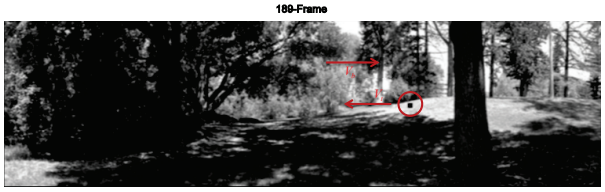

Fig. 6. The 189th frame of the input visual stimuli

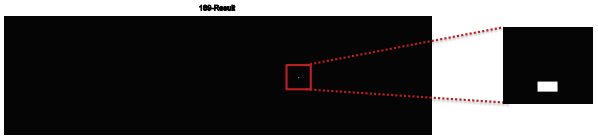

Fig. 7. The output of ESTMD

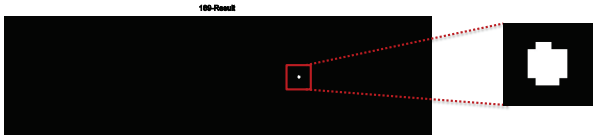

Fig. 8. The output of our proposed algorithm

As shown in Fig. 7, ESTMD model only detect a moving edge in this frame rather than a pair of edge compared with Fig. 4. This is because the second lateral inhibition mechanism proposed by [7], [12] do not take motion velocity and direction into consideration. Thus, the ON region and OFF region caused by the target motion can be strongly inhibited by the background motion even when there is velocity difference between the small target and the background. In this case when only an edge is detected, it is difficult for us to determine the existence of small target motion, because this edge may be the response of ESTMD to noise or background motion. In contrast, our proposed lateral inhibition mechanism is able to reduce the amount of lateral inhibition from surrounding area if there is velocity difference. Therefore, the response to the ON region and OFF region caused by target motion can be preserved when the small target go through the cluttered background. As it can be seen from Fig. 8, although the output of our proposed algorithm is slightly different from the best detection result shown in Fig.5, it still can be used as a criterion for existence of small target motion.

Fig. $9, 10, 11$ present the 200th frame of the input vision stimuli, the output of ESTMD model and the output of our algorithm, respectively.

Fig. 9. The 200th frame of the input visual stimuli

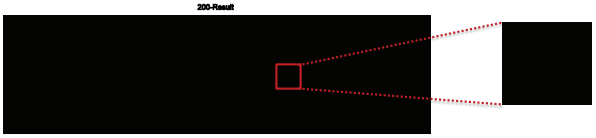

Fig. 10. The output of ESTMD

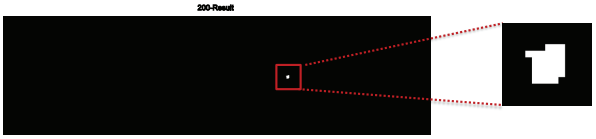

Fig. 11. The output of our proposed algorithm

Obviously, ESTMD model shows no response to the motion of small target in Fig. 10, because the ON and OFF region caused by target motion are completely inhibited by background motion. On the contrary, since our proposed lateral inhibition mechanism can determine the amount of inhibition based on the velocity difference between the target and the background, the ON and OFF region receive much less inhibition from surrounding area in this frame. Therefore, as we can see from Fig. 11, our proposed algorithm still provide a clear response to the small target motion.

Fig. $12, 13, 14$ present the 210th frame of vision stimuli, the output of ESTMD model and the output of our algorithm, respectively.

Fig. 12. The 210th frame of the input visual stimuli

As it is shown in Fig. 13, the output of ESTMD model is just a pair of incomplete moving edges. However, our proposed model shows a better performance of motion detection in Fig. 14.

For more clear and valid comparison between two models' detection performance, we firstly define a N value for the output of two models. N represents the number of white pixels

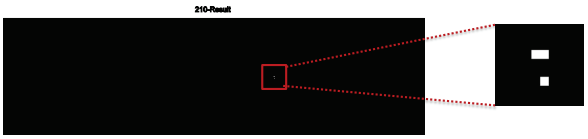

Fig. 13. The output of ESTMD

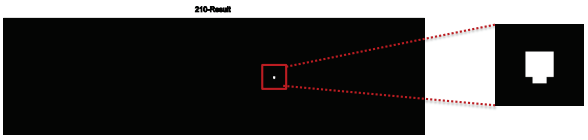

Fig. 14. The output of our proposed algorithm

located in the given red box. For example, as we can see from Fig. 5, the white pixels located in the red rectangle is 12, then the N value of our proposed algorithm for this frame is defined as 12.

Then, we compare the N value of two models from the 180th to the 250th frame. Fig. 15 and Fig. 16 present the N value of ESTMD model and our proposed model from the 180th to the 200th frame, respectively. The horizontal axis denotes the input frame and the vertical axis is the number of pixel which is located in red box, i.e. N value. The red lines in Fig. 15 and Fig. 16 denote the least N value (which is set as 6 for ESTMD model, 12 for our proposed algorithm) to determine the existence of target motion and is used as a reference line.

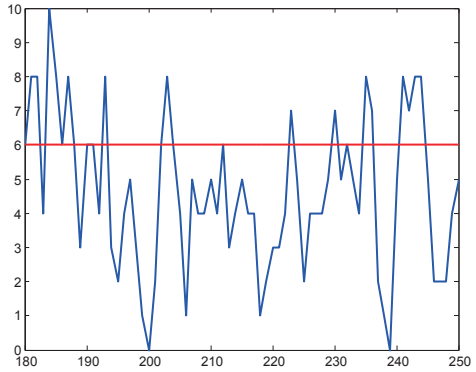

Fig. 15. The N value of ESTMD from the 180th frame to the 250th frame. The red line denotes the least number of white pixels which are needed to determine the existence of target motion.

As we can see from Fig. 15, the most frames' N value is lower than the red reference line which is set as 6 for ESTMD model. This is because when the small target go through the moving cluttered background, the ON and OFF region caused by target motion receive strong lateral inhibition from surrounding area. As a consequence of strong lateral inhibition, the response of ESTMD model to small target motion will be weaken. That is, the number of white pixels will

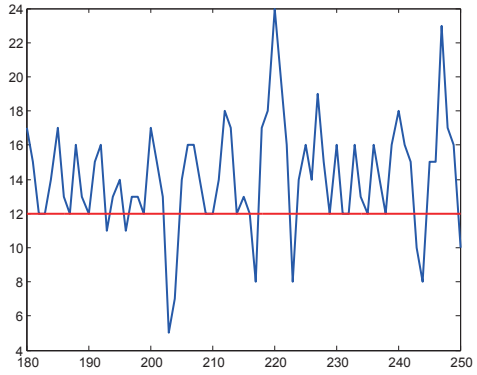

Fig. 16. The N value of our proposed algorithm from the 180th frame to the 250th frame.The red line denotes the least number of white pixels which are needed to determine the existence of target motion.

decrease when the small target moves through the background. If the N value is lower than the reference value, we can not distinguish the small target motion from background motion and noise easily. However, the N value of most frames are higher than the red reference line which is set as 12 for our proposed algorithm in Fig 16. Compared with the output of ESTMD model in Fig. 15, our proposed model can detect motion of small target in most frames.

In the following paper, we use two additional image sequences to test the proposed method. Fig. 17 and Fig. 20 are representative images of image sequence 1 and 2, respectively.

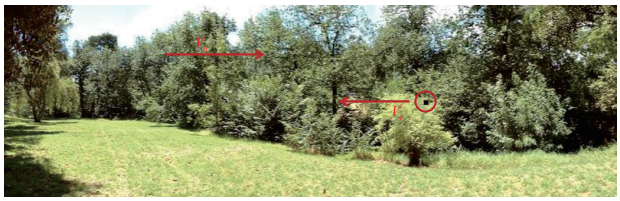

Fig. 17. Image Sequence 1, a small moving target amidst in the moving cluttered background

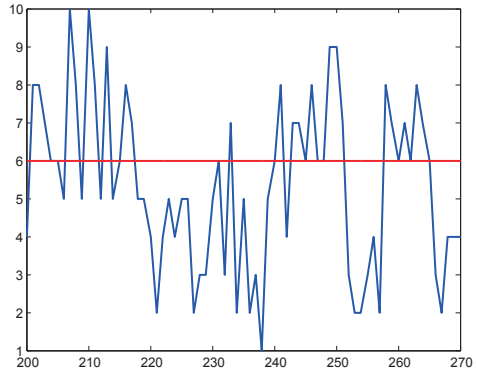

Fig. 18. The N value of ESTMD from the 200th frame to the 270th frame for the input image sequence 1. The red line denotes the least number of white pixels which are needed to determine the existence of target motion.

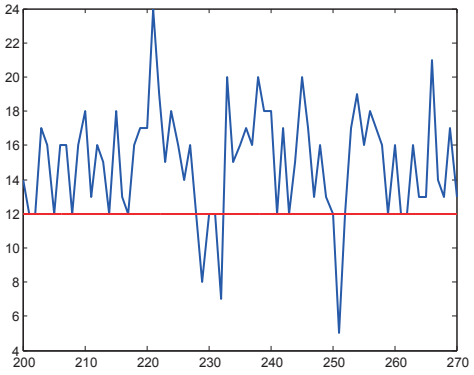

Fig. 19. The N value of our proposed algorithm from the 200th frame to the 270th frame for the input image sequence 1.The red line denotes the least number of white pixels which are needed to determine the existence of target motion.

Fig. 18 and Fig.19 show the N value of ESTMD model and our proposed algorithm from the 200th frame to the 270th frame for the input image sequence 1. Similarly, when the image sequence 2 is used as the input of two models, Fig.21 and Fig. 22 represent the N value of two models (ESTMD and our proposed algorithm), respectively.

Fig. 20. Image Sequence 2, a small moving target amidst in the moving cluttered background

Fig.18 and Fig.21 are similar to Fig.13. When the small target moves through the cluttered moving background, because of the strong lateral inhibition from its peripheral region, the detection performance is unstable and the N value of most frames are lower than reference value which is set as 6 for ESTMD model. However, due to the existence of velocity difference between the small target and the cluttered background, the amount of lateral inhibition of our proposed algorithm is much less than that of ESTMD model. Therefore, as it is shown in Fig.19 and Fig.22, the detection performance of the proposed algorithm is more stable and the N value is larger than reference value which is set as 12 for our proposed algorithm in most frames.

IV. Conclusion

To summarize, inspired by the biological findings, we propose a new lateral inhibited STMD model in this paper which can inhibit the response to background, but also highlight the motion of small target. We find that although ESTMD model can detect small moving target, the detection result is not very stable due to the influence of background motion. As demonstrated in the experiments, our proposed model has

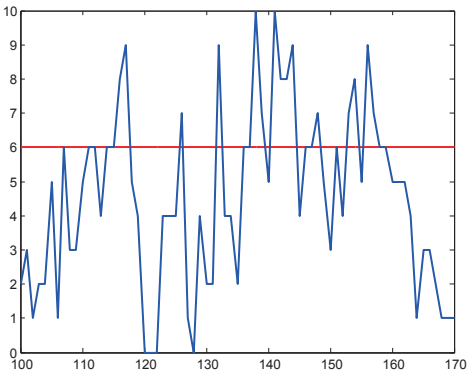

Fig. 21. The N value of ESTMD from the 100th frame to the 170th frame for the input image sequence 2. The red line denotes the least number of white pixels which are needed to determine the existence of target motion.

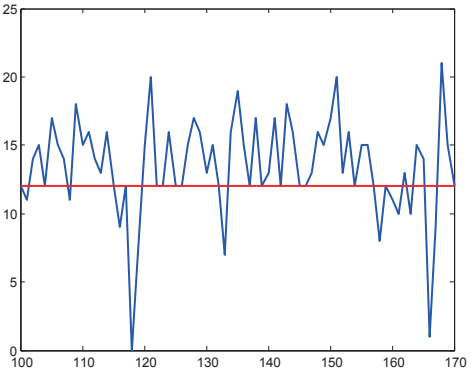

Fig. 22. The N value of our proposed algorithm from the 100th frame to the 170th frame for the input image sequence 2.The red line denotes the least number of white pixels which are needed to determine the existence of target motion.

efficiently overcome the deficiency of ESTMD model and obtained a better detection performance than ESTMD model.

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. R. Dunbier, S. D. Wiederman, P. Shoemaker, D. C. O'Carroll *et al.*, "Modelling the temporal response properties of an insect small target motion detector," in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2011 Seventh International Conference on*. IEEE, 2011, pp. 125–130.

[2] G. A. Horridge, "The separation of visual axes in apposition compound eyes," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 285, no. 1003, pp. 1–59, 1978.

[3] M. F. Land, "Visual acuity in insects," *Annual review of entomology*, vol. 42, no. 1, pp. 147–177, 1997.

[4] K. Nordström and D. C. O'Carroll, "Small object detection neurons in female hoverflies," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 273, no. 1591, pp. 1211–1216, 2006.

[5] P. D. Barnett, K. Nordström, and D. C. O'Carroll, "Retinotopic organization of small-field-target-detecting neurons in the insect visual system," *Current Biology*, vol. 17, no. 7, pp. 569–578, 2007.

[6] K. Nordstrom, P. D. Barnett, and D. C. O Carroll, "Insect detection of small targets moving in visual clutter," *PLoS biology*, vol. 4, no. 3, p. 378, 2006.

[7] 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, pp. 1–12, 2008.

[8] B. P. Ölveczky, S. A. Baccus, and M. Meister, "Segregation of object and background motion in the retina," *Nature*, vol. 423, no. 6938, pp. 401–408, 2003.

[9] D. M. Bolzon, K. Nordström, and D. C. O'Carroll, "Local and large-range inhibition in feature detection," *The Journal of Neuroscience*, vol. 29, no. 45, pp. 14 143–14 150, 2009.

[10] M. V. Srinivasan, S. B. Laughlin, and A. Dubs, "Predictive coding: a fresh view of inhibition in the retina," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 216, no. 1205, pp. 427–459, 1982.

[11] J. Van Hateren, "Theoretical predictions of spatiotemporal receptive fields of fly lmcs, and experimental validation," *Journal of Comparative Physiology A*, vol. 171, no. 2, pp. 157–170, 1992.

[12] K. J. Halupka, S. D. Wiederman, B. S. Cazzolato, and D. C. O'Carroll, "Discrete implementation of biologically inspired image processing for target detection," in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2011 Seventh International Conference on*. IEEE, 2011, pp. 143–148.

[13] M. Egelhaaf, R. Kern, and J. P. Lindemann, "Motion as a source of environmental information: a fresh view on biological motion computation by insect brains," *Frontiers in neural circuits*, vol. 8, 2014.

[14] A. D. Straw, "Vision egg: an open-source library for realtime visual stimulus generation." *Frontiers in neuroinformatics*, vol. 2, no. November, p. 4, 2008.

[15] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *JOSA A*, vol. 4, no. 12, pp. 2379–2394, 1987.