

Tunisian Republic

Ministry of National Defense

Air Force

Borj el Amri Aviation School

**Final project report**

**Specialty**

**Telemechanic**

By

**SLT Boukary DERRA**

**Bio-inspired small target motion detection against cluttered backgrounds using optical sensors**

Jury president : ….

Reporter : …..

Supervisor : LCL Tijeni Delleji

Academic Year: 2022 – 2023

Dedication:

*I dedicate this work,*

*To my family,*

*To my friends and loved ones for their support and encouragement,*

*To all those anonymous individuals who have contributed in any way to my growth,*

*To LCL Tijeni Delleji, without whom this project would not have been possible.*

*Boukary DERRA*

Thanks:

*Firstly, I would like to thank the* ***Major Colonel Commandant*** *of the Borj El Amri Aviation School as well as the* ***Colonel Director*** *of Studies for all the support and help they have given me.*

*I would also like to express my gratitude to the* ***Director of University Studies*** *for his encouragement and wise advice throughout the realization of my work.*

*Last but not least, I would like to thank the honorable* ***Jury Members*** *who kindly agreed to discuss this work and provided valuable criticism and suggestions.*

*At the end of this work, it is my duty to warmly and sincerely thank all those who contributed to the realization of this modest work. Words are not enough to express the gratitude I have towards* ***LCL Tijeni Delleji****, Professor at ENSIT, who supported and supervised me in this work and provided me with the desirable scientific and moral support. I personally thank him for giving me his trust and guiding me with his valuable advice.*

*I would also like to thank all the personnel of the Borj El Amri Aviation School with whom I spent these 5 wonderful years*

Contents

[Acronyms: 6](#_Toc134923102)

[GENERAL INTRODUCTION 8](#_Toc134923103)

[1. Traditional motion detection methods 10](#_Toc134923104)

[1.1. Optical flow 10](#_Toc134923105)

[1.2. Background subtraction 13](#_Toc134923106)

[1.3. Frame differencing 15](#_Toc134923107)

[1.4. Temporal differencing 16](#_Toc134923108)

[2. Formulation of the bio-inspired system 19](#_Toc134923109)

[2.1. Introduction 19](#_Toc134923110)

[2.2. Biological background of Small target motion detectors (STMD) 20](#_Toc134923111)

[2.3. Bio-inspired small target motion detector with a new lateral inhibition mechanism 25](#_Toc134923112)

[2.3.1. Introduction 25](#_Toc134923113)

[2.3.2. Modeling 26](#_Toc134923114)

[2.3.3. Experiment and conclusion 36](#_Toc134923115)

[2.4. Feedback ESTMD 37](#_Toc134923116)

[2.5. DSTMD 37](#_Toc134923117)

[2.6. STMD+ 37](#_Toc134923118)

[2.7. apg-STMD 37](#_Toc134923119)

[3. ESTMD implementation, results and discussion 38](#_Toc134923120)

[3.1. Introduction 38](#_Toc134923121)

[3.2. Gaussian function 45](#_Toc134923122)

[3.2.1. Two-Dimensional Gaussian Function 46](#_Toc134923123)

[3.2.2. Discrete form: The Gaussian kernel 49](#_Toc134923124)

[3.3. Retina Layer 51](#_Toc134923125)

[3.3.1. Photoreceptor 51](#_Toc134923126)

[3.3.2. Lipetz transformation 52](#_Toc134923127)

[3.4. Lamina Layer 54](#_Toc134923128)

[3.4.1. Low pass filter 54](#_Toc134923129)

[3.4.2. Large Monopolar Cell 54](#_Toc134923130)

[GENERAL CONCLUSION 58](#_Toc134923131)

[BIBLIOGRAPGIE 59](#_Toc134923132)

[ANNEXES 63](#_Toc134923133)

List of figures

[Figure 1: Détection de mouvement par flux optique 12](#_Toc130817459)

[Figure 2: Détection de mouvement par soustraction de fond 14](#_Toc130817460)

[Figure 3: Détection de mouvement par différenciation de cadre 16](#_Toc130817461)

[Figure 4: Détection de mouvement par différenciation temporelle 18](#_Toc130817462)

[Figure 5: Détection de mouvement par infrarouge 19](#_Toc130817463)

[Figure 6: Un sphinx 21](#_Toc130817464)

[Figure 7: Un syrphe 22](#_Toc130817465)

[Figure 8: Une libellule 22](#_Toc130817466)

[Figure 9: 24](#_Toc130817467)

[Figure 10 : 25](#_Toc130817468)

[Figure 11: Modèles et Application du système visuel des insectes [7] 26](#_Toc130817469)

[Figure 12: Schéma de ESTMD avec un nouveau mécanisme d’inhibition latéral. [9] 29](#_Toc130817470)

[Figure 13: Masque de convolution Gaussien. (3) 30](#_Toc130817471)

[Figure 14: Matrice de convolution. 30](#_Toc130817472)

[Figure 15: Fonction de Lipetz. (4) 30](#_Toc130817473)

[Figure 16: Filtre passe-bas de Lij(t). (5) 30](#_Toc130817474)

[Figure 17: Filtre passe- bas de Pij(t) 30](#_Toc130817475)

[Figure 18: Filtre spatiale Gaussien dans le cas continu. (1) 32](#_Toc130817476)

[Figure 19: Fonction Gaussienne. (2) 32](#_Toc130817477)

# Acronyms:

**BPF**: Band Pass Filtering

**DCMD**: Descending Contra-lateral Movement Detector

**DSNs**: Direction Selective Neuron(s)

**EMDs**: Elementary Motion Detector(s)

**FDSR**: fast depolarization, slow repolarization

**GB**: Gaussian Blur

**HPF**: High Pass Filtering

**HW-R**: half wave rectifier

**LGMDs**: Lobula Giant Movement Detector(s)

**LI**: Lateral Inhibition

**LIM**: Lateral Inhibition Mechanism

**LPF**: Low Pass Filtering

**LPNM**: Looming Perception Neuron Models

**LPTCs**: Lobula Plate Tangential Cell(s)

**LSM**: Looming Sensitive Models

**ND**(s): null or non-preferred direction(s)

**ODE**: ordinary differential equation

**OF**: Optic Flow

**PD**(s): preferred direction(s)

**STMDs**: Small Target Motion Detector(s)

**STMSM**: Small Target Motion Sensitive Models

**TSM**: Translation Sensitive Models

**UAV**: Unmanned Aerial Vehicle

# GENERAL INTRODUCTION

For any living organism, being able to detect a moving object is very important, and often vital. This operation is even more crucial for certain animals, such as insects, which must be able to analyze the actions around them, follow their peers, avoid predators, or even attack prey.

In the era of technical and technological progress, there is the same problem of detecting a moving object, especially in areas such as robotics, autonomous vehicles, air defense, etc. This detection becomes even more complicated when the size of the target is small or when the background in which it is located is relatively complex or cluttered. The goal of this project is to meet this challenge while drawing inspiration from biological models, particularly insects.

Indeed, with millions of years of evolution, the visual system of insects is among the most effective for detecting movement. For example, grasshoppers can fly for hundreds of kilometers in a very dense environment without colliding; bees can follow the movements of their partners despite their speed; and praying mantises can monitor small moving prey in a complex environment.

But before we go any further, it is legitimate to ask ourselves the following question: Why draw inspiration from a biological model when there are detection techniques that are less complex and easy to implement? Traditionally, there are several methods for detecting moving targets, including Infrared Small Target Motion Detector, Optical flow, background subtraction, temporal differencing, etc.

However, all these methods become obsolete when the target being targeted is very small (on the order of a few pixels) or when the background in which it is located is too complex and has characteristics similar to the target. Hence the need to turn to new means of detection, new sources of inspiration such as biology, particularly insects.

Insects are very effective in searching for their partners or tracking prey that often appear as simple spots in a complex field of vision. Several models based on insects have been developed in recent years.

Among these bio-inspired models [7], we have:

* Looming sensitive models (LSMs), mainly used for collision detection;
* Translation sensitive models (TSMs), mainly used to determine the directions taken by objects;
* And small target motion sensitive models (STMSMs). These models are characterized by STMD detectors, which will be the subject of our project.

The main objective of our project is to detect the motion of small targets (which can be of the order of a few pixels) in a relatively complex and possibly moving background.

We will therefore organize the rest of the work as follows:

In Section I, we will review conventional motion detection models and their limitations in the context of our project;

In Section II, we will delve into the heart of the project, which consists of studying the different models and applications of the insect visual system;

For Section III, we will develop in detail the directional model (DSTMDs) that we have chosen after the comparative study in Section II.

Finally, we will conclude the work with the limitations of the chosen model, the possibilities, and the perspectives in Section IV.

# Traditional motion detection methods

There are several techniques for detecting the motion of objects in nature, and each of these techniques is important depending on the type of target and the complexity of the environment.

## Optical flow

Motion detection using optical flow is a technique that involves analyzing the motion of pixels in an image or video sequence. Optical flow is a mathematical algorithm that estimates the movement of pixels between frames in a video. This movement can be used to detect small targets or objects that are moving in a video sequence.

The optical flow algorithm calculates the direction and magnitude of movement for each pixel in an image or video sequence. This is done by comparing the intensity values of pixels in two consecutive frames. The resulting flow vectors represent the displacement of each pixel between the two frames.

To use optical flow for small target motion detection, the flow vectors can be analyzed to identify regions of the image or video sequence where there is significant movement. These regions can be further processed to detect small targets.

There are several different optical flow algorithms that can be used for motion detection. Some popular algorithms include Lucas-Kanade, Horn-Schunck, and Farneback. These algorithms vary in their complexity and accuracy, and the choice of algorithm will depend on the specific application and requirements. [14] .

The general process of motion detection using optical flow can be summarized as follows:

1. Capture two consecutive frames: The first step is to capture two consecutive frames from a video sequence. These frames should be captured with a fixed time interval between them.
2. Compute optical flow: The next step is to compute the optical flow between the two frames using an optical flow algorithm such as Lucas-Kanade or Horn-Schunck. The optical flow algorithm estimates the movement of pixels between the two frames, producing a dense flow field.
3. Analyze the flow field: The flow field can be analyzed to identify regions of the image where there is significant motion. This can be done by thresholding the magnitude of the flow vectors or by looking for areas where the flow vectors are aligned in a particular direction.
4. Refine the flow field: The flow field can be further refined to remove noise or artifacts that may be present. This can be done by smoothing the flow field using techniques such as Gaussian filtering or by applying a median filter.
5. Detect moving objects: Once the flow field has been analyzed and refined, moving objects can be detected by identifying regions of the image where there is consistent motion over time. This can be done by comparing the flow field to a background model or by using techniques such as connected component analysis.
6. Track moving objects: Finally, the moving objects can be tracked over time by updating their position based on the flow field in subsequent frames. This can be done using techniques such as Kalman filtering or particle filtering.

The exact implementation of these steps will depend on the specific application and requirements. However, this general process provides an overview of the key steps involved in motion detection using optical flow.

Optical flow has several advantages and disadvantages for motion detection applications. Here are some of the main ones:

Advantages:

* High accuracy: Optical flow can provide highly accurate estimates of the motion of pixels between frames. This makes it well-suited for applications that require precise motion measurements.
* Dense flow field: Optical flow provides a dense flow field that covers all pixels in an image or video sequence. This allows for detailed analysis of motion patterns and can help identify small or subtle changes in motion.
* Real-time performance: Many optical flow algorithms can be computed in real-time, making it possible to use them for real-time applications such as video surveillance and robotics.
* Robustness: Optical flow can be used in noisy or cluttered environments, and can handle occlusions and other complex motion patterns.

Disadvantages:

* Sensitivity to brightness changes: Optical flow is sensitive to brightness changes between frames, which can lead to errors in the flow estimates. This can be mitigated by using robust optical flow algorithms or by preprocessing the images to reduce the effects of brightness changes.
* Limited accuracy for large motions: Optical flow may not be accurate for large motions or when there is significant occlusion or clutter in the scene. This can limit its usefulness for some applications.
* Computational complexity: Some optical flow algorithms can be computationally expensive, especially when dealing with large images or video sequences. This can limit the speed and scalability of optical flow-based motion detection systems.
* Lack of depth information: Optical flow only provides information about the 2D motion of pixels in an image or video sequence. It does not provide any information about the depth or 3D structure of the scene. This can limit its usefulness for some applications that require 3D information.

In summary, Optical flow is a popular technique for motion detection that provides high accuracy, real-time performance, and a dense flow field covering all pixels in an image or video sequence. It is also robust in noisy or cluttered environments and can handle occlusions and other complex motion patterns. However, optical flow can be sensitive to brightness changes, may not be accurate for large motions or when there is significant occlusion or clutter, can be computationally expensive, and does not provide any information about the depth or 3D structure of the scene. The advantages and disadvantages of optical flow should be carefully considered when choosing a motion detection technique for a specific application.



Figure 1: Motion detection using optical flow

## Background subtraction

Motion detection using background subtraction is a technique that involves separating the foreground objects from the background in an image or video sequence by subtracting a reference or background image from the current frame. [15]

The general process can be summarized as follows:

1. Capture a reference image: The first step is to capture a reference image of the scene with no moving objects. This reference image represents the background of the scene.
2. Obtain the current frame: The next step is to obtain the current frame of the video sequence.
3. Subtract the reference image: The reference image is subtracted from the current frame to obtain the difference or delta image. This image highlights the regions of the frame where there are changes in pixel values, which correspond to moving objects.
4. Threshold the difference image: The difference image is then thresholded to create a binary mask that separates the foreground objects from the background.
5. Filter the binary mask: The binary mask can be filtered to remove noise or artifacts that may be present. This can be done using techniques such as morphological operations or median filtering.
6. Detect moving objects: Once the binary mask has been filtered, moving objects can be detected by identifying the connected components in the binary mask.
7. Track moving objects: Finally, the moving objects can be tracked over time by updating their position based on the binary mask in subsequent frames. This can be done using techniques such as Kalman filtering or particle filtering.

Motion detection using background subtraction has several advantages and disadvantages. Advantages include its simplicity, real-time performance, and ability to work well in simple scenes with static backgrounds. Disadvantages include its sensitivity to changes in lighting conditions, its inability to handle complex backgrounds or moving cameras, and the need for manual adjustment of parameters for each scene.

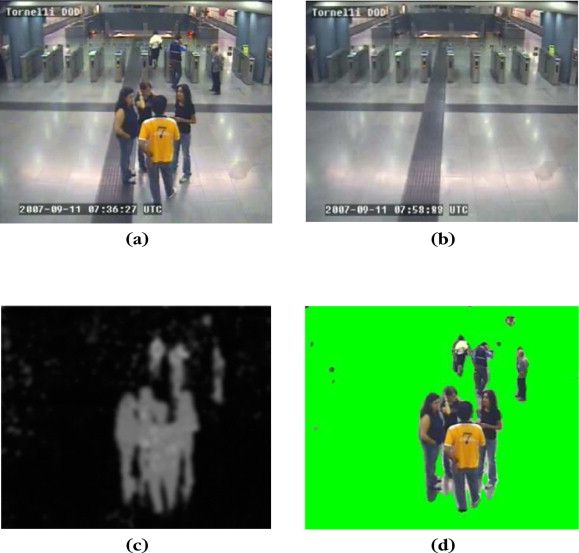


Figure 2: Motion detection using background subtraction

## Frame differencing

Motion detection using frame differencing is a technique that involves subtracting two consecutive frames of an image or video sequence to obtain the difference or delta image. [16]

The general process can be summarized as follows:

1. Obtain the current frame: The first step is to obtain the current frame of the video sequence.
2. Obtain the previous frame: The next step is to obtain the previous frame of the video sequence, which was captured just before the current frame.
3. Subtract the previous frame: The previous frame is subtracted from the current frame to obtain the difference or delta image. This image highlights the regions of the frame where there are changes in pixel values, which correspond to moving objects.
4. Threshold the difference image: The difference image is then thresholded to create a binary mask that separates the foreground objects from the background.
5. Filter the binary mask: The binary mask can be filtered to remove noise or artifacts that may be present. This can be done using techniques such as morphological operations or median filtering.
6. Detect moving objects: Once the binary mask has been filtered, moving objects can be detected by identifying the connected components in the binary mask.
7. Track moving objects: Finally, the moving objects can be tracked over time by updating their position based on the binary mask in subsequent frames. This can be done using techniques such as Kalman filtering or particle filtering.

Motion detection using frame differencing has several advantages and disadvantages. Advantages include its simplicity, real-time performance, and ability to work well in simple scenes with static backgrounds. Disadvantages include its sensitivity to changes in lighting conditions, its inability to handle complex backgrounds or moving cameras, and the need for manual adjustment of parameters for each scene.

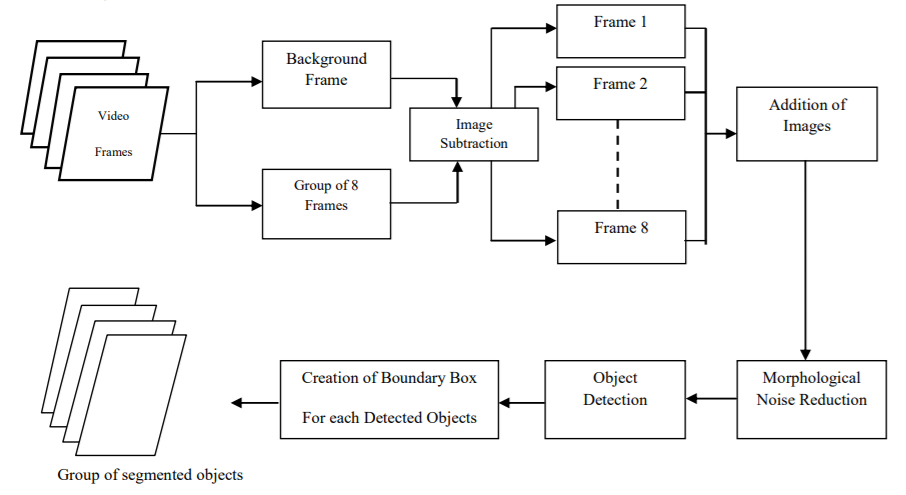


Figure 3: Motion Detection using frame differencing

## Temporal differencing

Motion detection using temporal differencing is a technique that involves subtracting two or more frames from different times in a video sequence to obtain the difference or delta image. The general process can be summarized as follows:

1. Obtain the current frame: The first step is to obtain the current frame of the video sequence.
2. Obtain the previous frame(s): The next step is to obtain one or more previous frames of the video sequence, which were captured at different times before the current frame.
3. Subtract the previous frame(s): The previous frame(s) are subtracted from the current frame to obtain the difference or delta image. This image highlights the regions of the frame where there are changes in pixel values, which correspond to moving objects.
4. Threshold the difference image: The difference image is then thresholded to create a binary mask that separates the foreground objects from the background.
5. Filter the binary mask: The binary mask can be filtered to remove noise or artifacts that may be present. This can be done using techniques such as morphological operations or median filtering.
6. Detect moving objects: Once the binary mask has been filtered, moving objects can be detected by identifying the connected components in the binary mask.
7. Track moving objects: Finally, the moving objects can be tracked over time by updating their position based on the binary mask in subsequent frames. This can be done using techniques such as Kalman filtering or particle filtering.

Motion detection using temporal differencing has several advantages and disadvantages. Advantages include its ability to handle more complex backgrounds and lighting conditions compared to frame differencing, and its ability to work with moving cameras. Disadvantages include the need to use multiple frames for temporal differencing, which can result in more computational complexity and slower performance, and the potential for ghosting artifacts when objects are moving at different speeds or directions.

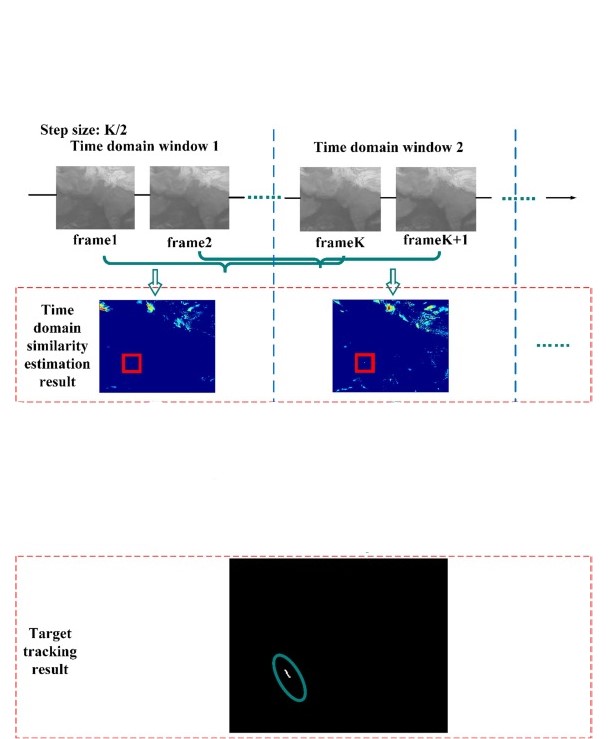


Figure 4: Motion detection using temporal differencing

# Formulation of the bio-inspired system

## Introduction

In the chapter I, we highlighted the limitations of traditional models for detecting the movement of small objects. These limitations are mainly the complexity and/or mobility of the backgrounds, as well as the very small size of the targets.

Therefore, to overcome these problems, we will focus on bio-inspired models. Specifically, models for detecting motion based on the visual system of insects.

Figure 5: Block schema of the different motion patterns

In general, artificial visual system based on insects’ visual system can be classified into 3 main models [7]. As shown in Figure 5, these bio-inspired models are as follows:

* Looming sensitive models (LSMs):

The looming sensitive neuronal models is mainly use for collision-detecting systems more precisely for mobile ground robots, UAVs and ground vehicles. It is inspired by the locust visual systems. These include two neuronal models of the LGMD1 and the LGMD2.

* Translation sensitive models (TSMs):

The translation sensitive motion detectors is used to detect directions taken by objects. It’s inspired by the directionally selective motion detecting neurons in locust, namely the locust direction selective neurons (DSN).

* And small target motion sensitive models (STMSMs):

These models are characterized by small target motion detectors (STMD) and use for small target motion detection against cluttered and moving background.

For the next part, we will focus on the third point, which is the subject of our project.

## Biological background of Small target motion detectors (STMD)

The small target motion detectors (STMD) are a type of motion-sensitive neuron found in insect visual systems. They exhibit a remarkable ability to detect small target motion and have been identified in various insect species such as hawk moths, hover flies and dragonflies, which are shown in Fig. 6. Over the past twenty years, there have been numerous studies examining the anatomy and physiology of STMD neurons, with a significant body of research published on the topic. Some notable studies include those by researchers numbered [10, 11, 12, and 13].

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

Figure 6: (a) A hawk moth. (b) A hover fly. (c) A dragonfly.

Compared to other motion-sensitive neurons with wide-field perception, such as LGMD, DSN, LPTC, etc., the small target motion detector (STMD) has a specific size selectivity for small-field movements. Specifically, STMD neurons respond most strongly to targets that are between 1-3 degrees in the field of view and do not respond to larger bars or wide-field grating stimuli, as observed in studies [11, 17]. Fig. 9 provides a clear demonstration of the size selectivity of STMD neurons by showing the neural response to targets of varying heights. As shown in Fig. 9a, smaller targets with heights of 0.8◦ and 3◦ can elicit stronger neural responses, while the response to the larger target with a height of 15◦ is weaker and indistinguishable from spontaneous activity [7]. Fig. 9b further illustrates the selectivity of STMD to target height, showing an optimal size sensitivity that corresponds to the strongest neural response. If the target height is higher or lower than the optimal one, the neural response will significantly decrease. In summary, the STMD is unique among motion-sensitive neurons in its size selectivity, responding most strongly to small-field movements within a specific range of sizes.

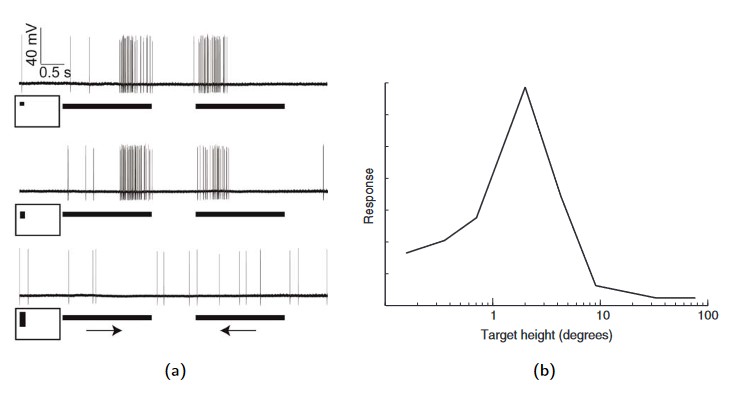


Figure 7: STMD neuronal raw responses: (a) Neuronal responses to motion of three different-sized targets (0.8◦, 3◦, or 15◦ high by 0.8◦ wide) drifted against bright backgrounds: the horizontal bars indicate the movement duration and the arrows denote the direction of target motion; from [7]. (b) The response of an STMD to targets of varying height; from [7].

Directionally selective STMD neurons have been observed in some studies [10, 11], which respond strongly to small target motion oriented along the preferred direction (PD), but show weaker or no response to non-preferred direction (ND) motion.

Fig. 10a illustrates the responses of a directionally selective STMD neuron to three targets of different sizes, showing that the larger target with a height of 15◦ cannot activate the STMD neuron, even when moved in the PD direction. However, the STMD neuron responds strongly to the PD motion of smaller targets with heights of 0.8◦ and 3◦, while remaining inactive when the smaller targets move in the ND direction. Further research [11, 12] has shown that the size and direction selectivity of STMD neurons are independent of background motion. In other words, STMD neurons respond robustly to small target motion against visually cluttered backgrounds, regardless of the direction and velocity of background motion.

Fig. 10b shows that the STMD neuron responds strongly to small targets moving in the PD direction (downward), but weakly to those moving in the ND direction (upward). This response is consistent regardless of the direction or velocity of the background motion.

In summary, STMD neurons are directionally selective and can recognize small target motion even in the absence of relative motion between the moving objects and the background.

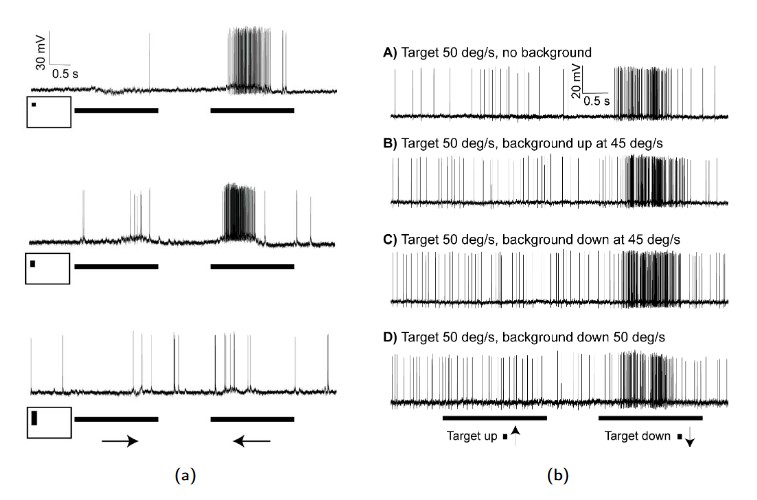


Figure 8 : STMD neuronal directionally responses: (a) Raw responses of the directionally selective STMD neuron which prefers target motion to left, tested by motion of three different-sized targets (0.8◦, 3◦, or 15◦ high by 0.8◦ wide) drifted against bright backgrounds: the horizontal bars indicate the stimuli duration and the arrows denote the direction of target motion; from [7]. (b) Responses of the STMD neuron which prefers target motion downward, to targets drifted against cluttered backgrounds; from [7].

Several computational models have been proposed in the last decade to simulate the STMD based on biological findings.

One of the earliest model is the elementary small target motion detector (ESTMD) developed by Steven D. Wiederman, Patrick A. Shoemaker, and David C. O’Carrol. This model is account for the size selectivity. However, the ESTMD model is unable to realize the DS of the STMD revealed by biologist.

To account for the DS, two hybrid models: the ESTMD-EMD and the EMD-  
ESTMD were proposed for achieving the DS of the STMD [2]. More specifically, the  
ESTMD-EMD indicates that the ESTMD cascades with the EMD, while the EMD-ESTMD indicates that the EMD cascades with the ESTMD. These two hybrid models have been successfully used for target tracking against cluttered backgrounds in an autonomous mobile ground robot [22, 23, and 24]

In order to make faster progress, we will focus more on the models developed by Wang and his teams [3, 4, 5, 6, 7, and 9]. These models are based on the works of Widerman et al. and have the advantage of being more recent.

Figure 9: Computational models of STMD [7]

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

### Introduction

Based on the previous ESTMD model, Wang at al. proposed a new model with a new lateral inhibition mechanism. Inspired by the biological visual process in the fly, their proposed model is composed of four neural layers: retina, lamina, medulla and lobula.

In preceding chapters, we have observed that detecting small moving targets amidst cluttered backgrounds is a crucial task for animals in nature. Recently, biologists have discovered a specific type of neuron in the lobula complex, known as STMDs which exhibit extraordinary selectivity in detecting small targets amidst visual clutter. Additionally, some researchers contend that lateral inhibition plays a significant role in differentiating target motion from background motion and may even account for many higher-order visual neurons' tuning characteristics. Inspired by the discovery that complete lateral inhibition occurs only when the central and peripheral regions move identically, Wang et al. have suggested a new lateral inhibition mechanism that employs motion velocity and direction to enhance the ESTMD model's (elementary small target motion detector) performance. This chapter introduces the ESTMD model with the innovative lateral inhibition mechanism.

As noted in [1] and [11], the lateral inhibition mechanism could play a critical role in the visual processing system of insects and aid in shaping the response tuning to small targets. However, implementing it has not been clear. Wang et al. suggest that lateral inhibition is a pervasive biological mechanism in the visual processing system of insects, and recent studies on lateral inhibition [18] have revealed that certain neurons respond to local motion on the retina only when the motion trajectory is different from that in a large surrounding region. Specifically, the author found that these neurons respond when an object in their receptive field center moves relative to the background, but are nearly completely suppressed when the object moves together with the background. Additionally, results from [13] demonstrate that lateral inhibition is velocity and direction-selective. In fact, lateral inhibition is most pronounced when the object and distracter target move with the same velocity and direction within a certain range but remains weak or silent when the object and distracter target move with the same velocity but in different motion directions.

Although ESTMD proposed by [1] contains two lateral inhibition mechanisms, located in the lamina layer and medulla layer, respectively, these two lateral inhibition mechanisms do not entirely align with the aforementioned biological findings. In fact, according to these two lateral inhibition mechanisms, the object receives the same amount of lateral inhibition regardless of whether there is relative motion between the object and background or not. The motion of the object can be significantly weakened by the motion of the background even when the object motion differs from the background motion. Consequently, the detection performance of ESTMD is unstable, particularly when the object moves through a moving cluttered background.

Inspired by the velocity and motion direction-dependent lateral inhibition phenomenon discussed above, Wang et al. have enhanced the ESTMD model proposed by [1] with a new lateral inhibition mechanism that considers velocity and motion direction. They have demonstrated that their new lateral inhibition outperforms the lateral inhibition proposed in [1] and can improve target detection performance.

### Modeling

Wang et al. have presented a new model that builds upon the previous ESTMD model [7] by incorporating a novel lateral inhibition mechanism. The proposed model is inspired by the visual system of flies and comprises four neural layers: retina, lamina, medulla, and lobula. Figure 1 illustrates the model's schematic, which will be discussed further in the subsequent paper.

***Image***

***Preprocessing***

Photoreceptor

Luminance or Gray Scale Image

Input Image

**I**

**L**

Low Pass Filter

Lipetz

Transformation

**L**

**Retina**

**Layer**

**Lc**

Low Pass Filter

**P**

High Pass Filter

(LMCs)

**Lamina**

**Layer**

**x**

**YON**

**YOFF**

FDSR

FDSR

**-SOFF**

**YON**

**-SON**

**YOFF**

**SON**

**Medulla**

**Layer**





**FON**

**FOFF**

HW-R

HW-R

**HWOFF->FOFF**

**HWON->FON**



LI

LI

**Lobula**

**Layer**



**FON**

FDSR

**LobOFF**



**O**

Figure 10: Schematic illustration of Wang at al. proposed model

Haut du formulaire

Bas du formulaire

#### Retina Layer

The layer of the eye called the retina is made up of a matrix of M rows and N columns of photoreceptors, each of which corresponds to a pixel point. These photoreceptors receive luminance or gray levels from consecutive images. To simulate the **spatial blur** caused by fly optics, Wang et al. convolve the intensity of each pixel  in an image frame at time t, denoted by, with a Gaussian convolution mask.

That is:

 (1)

Where:

 (2)

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

 (3)

 in the equation (3) is the low-pass filtered version of  and satisfies the following relationship.



 (4)

Where  is the time constant.



(4) is an ordinary differential equation (ODE) of the first order to the form of:



With  and  ; In fact, is a constant for each t.

And the solution of this equation is:



Where  is a constant.

Retina

Layer

|  |
| --- |
| image composed |
|  |
| of pixels |

|  |
| --- |
| image composed |
|  |
| of pixels then |

***Blur effect then membrane potential.***

Figure 11: Retina Layer

#### Lamina Layer

There is a minor delay when the output of retina layer B is transmitted to the lamina layer. This delay is modelled by:

  (5)

Where  is the time constant.

(5) is an ODE of the first order.

The delayed signal  is employed as input for large monopolar cells (LMCs) situated in the lamina layer. Research on LMCs suggests that they can eliminate redundant information and enhance information transmission [19], [20]. In essence, the functionality of LMCs can be expressed using the following equation:



 (6)

 (7)

Where  is the output of LMCs and is the first-order low-pass filtered version of while τ3 is the time constant.

Lamina

Layer

|  |
| --- |
| image composed |
|  |
| of pixels |

|  |
| --- |
| image composed |
|  |
| of pixels then |

***Slight delay then deletion of redundant information and the maximization of transmission.***

Figure 12: Lamina Layer

W

LPF



#### Medulla Layer

The output of LMCs () is transmitted to the medulla layer, where it is divided into ON and OFF channels in the initial portion of the medulla layer. This operation can be mathematically represented as:

 (8)

 (9)

Where  and  are the signal of ON and OFF channels, respectively.

To simulate credible biophysical processes, the ON and OFF channels for each pixel (i, j) are transformed into an "adaptation state" by subjecting them to a non-linear low pass filter featuring a rapid depolarizing and slow repolarizing attribute [21]. Known as the FDSR (fast depolarization, slow repolarization) mechanism, it effectively suppresses texture information that changes quickly and enhances the contrast of novel changes.

We denote  ,  as the signal of ON and OFF channels after FDSR respectively, then





Let  and 



 (10)

Where  and  are the time constant and satisfy.

Similarly, we have

 (11)

Then, the filtered signal  ,  are subtractive by the original signals  ,.

 (12)

 (13)

Where,  are the output of medulla layer.

Once the FDSR mechanism is applied, the resulting signals  and  are transmitted to a half-wave rectifier (HW-R). The ON and OFF channels are then denoted as  and  after the half-wave rectification process. Then

 (14)

 (15)

To simplify, Wang et al. use  and  to refer to the signal after half wave rectification. Previous research proposed a lateral inhibition mechanism where central ON and OFF channels are inhibited by surrounding channels of the same polarity [1]. However, this mechanism seems to contradict recent findings, which show that lateral inhibition is only strong when central and peripheral regions move at the same velocity and direction [13], [18]. Therefore, a more reasonable mechanism should consider motion velocity and direction. In the paper [9], Wang, et al. propose a new lateral inhibition mechanism based on motion velocity and direction, and calculate the velocity vector of each pixel. While the EMD model can detect motion and direction, it has contrast and velocity dependence, making their responses ambiguous. Hence, they use a general matching algorithm to calculate motion vector for every pixel.

The matching criteria is defined as:

 (16)

Where  is the size of the search window and  is the input image at time t.

The motion vector, which is the translation vector, is derived by identifying the minimum value of the D parameter.

 (17)

Where  and  is the search range.

By utilizing the general matching algorithm mentioned earlier, we can calculate the motion vector  for every pixel  , where  and  represent the horizontal and vertical components of the motion vector, respectively. Consequently, we can define the motion vector matrix  and  for an input image as

 (18)

 (19)

To compute the velocity difference between the central and peripheral regions, we convolve U and V with H, which can be expressed as:

 (20)

 (21)

Where  is convolution operator and

 (22)

To clarify the function of  and , Wang, et al. begin by defining a neighborhood (PR) for the pixel  ,

 (23)

Where  ,,  , are decided by the size of small target.

If the motion vector  of pixel  is the same as the motion vector  of a pixel s in the peripheral region (PR), i.e.

 (24)

Then after convolving  and  with  ,  and  become zero. Conversely, if there is a velocity difference between pixel  and its peripheral region,  and  become non-zero. The magnitude of and  (and  ) is an indicator of the velocity difference between the central and peripheral regions (). The higher the magnitude, the greater the velocity difference. The total velocity difference between pixel  and its neighborhood is given by the equation

 (25)

Wang, et al. proposed lateral inhibition mechanism is implemented by multiplying the signal of ON and OFF channels by  .

 (26)

 (27)

The reason for this mechanism is to inhibit background motion and enhance small target motion, as it is more biologically plausible. The new mechanism is based on the observation that if a pixel belongs to the background, its motion vector will be the same as the motion vector of its peripheral region. In this case, the signal of the pixel should be strongly inhibited. However, if a pixel belongs to a moving target, then its motion vector will not be the same as the motion vector of other pixels in its peripheral region, unless the small target and background have the same motion. In this case, the signal of the pixel should be enhanced. The magnitude of the lateral inhibition depends on the value of  , and can be adjusted by parameter . This mechanism enhances the saliency of moving small targets smaller than  and inhibits false positives caused by the motion of the moving background, thus improving the model's performance.

|  |
| --- |
| image composed |
|  |
| of pixels |

|  |
| --- |
|  |
|  |
|  |

Medulla

Layer

Figure 13: Medulla Layer

#### Lobula Layer

The lobula layer exhibits correlation between the OFF channel, which has a delay, and the un-delayed ON channel. The delay in the OFF channel is obtained through a first order low pass filter, and the length of delay is dependent on the size and velocity of the small target.

 (28)

And the final output of the lobula layer is

 (29)

Lobula

Layer

|  |
| --- |
| image composed |
|  |
| of pixels |

|  |
| --- |
| and |
|  |
|  |
|  |

***Exhibits correlation between OFF and ON channels.***

Figure 14:Lobula Layer

### Experiment and conclusion

## Other STMD

### Feedback ESTMD

After the ESTMD with a new LIM, Wang et al. have introduced a feedback neural network to detect small target motion in the presence of naturally cluttered backgrounds. The model employs a feedback loop, wherein the model output is delayed in time and used as feedback to the preceding neural layer. Through extensive experiments, it was demonstrated that this feedback neural network outperforms existing STMD-based models for small target motion detection.

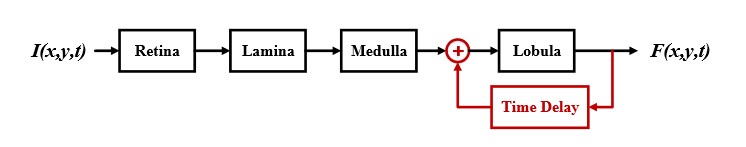


Figure 15: Schematic illustration of feedback STMD

### DSTMD

Another recent model, the directionally selective small target motion detector (DSTMD), was proposed by Wang et al. to achieve direction selectivity in STMD. All of these models are shown in Figure 30.

### STMD+

### apg-STMD

# 3. ESTMD implementation, results and discussion

## 3.1. Introduction

In this chapter, we will create a Python class that we will call STMD. This class will be based on the various equations and layers of the STMD models studied in Chapter 2. Our task here will be to implement these equations, vary the different parameters to obtain interesting results. The STMD class should allow for the detection of small target motion in a relatively complex and possibly moving background. The different methods of the Python STMD class will represent the layers of the STMD detectors.

Moving forward, we will use the ATOM editor to write our codes [25]; and Windows PowerShell for package installations, virtual environment management, and scripts execution [26].

For the various tests, we utilized two 5-second videos, each of which was divided into 151 frames using the following code.

The goal was to gauge the performance of the model depending on the complexity of the background.

Results:

* [30]: video and 151 frames of a small targets in more complex and moving background.
* [31]: video and 151 frames of a small target in less complex and moving background.

For testing on the frames, we will use frames 62 and 63 as follows:

|  |  |
| --- | --- |
| Figure 26: frame 63 | Figure 27: complex frame 63 |
| Figure 28: frame 62 | Figure 29: complex frame 62 |

* Fig. 26 is the frame that we will use to test the detection of a drone against a less complex and moving background; and Fig. 28 is the frame that precedes it.
* The same applies for the Fig. 27 that we will use to test the detection of a drone against a more complex and moving background; and Fig. 29 is the frame that precedes it.

All the source code is available on our GitHub account.

<https://github.com/boukary-derra/pfe/tree/main/code>

## 3.2. Gaussian function

Given the utility of the Gaussian function for the model (within the lamina layer), we will start by introducing it.

The Gaussian function is a mathematical function widely used in various fields of science, including physics, engineering, and mathematics. It is named after the mathematician Carl Friedrich Gauss, who first introduced it in the early 19th century.

The function is defined as follows:



Where:

* x is the independent variable
* A is the amplitude of the curve
* μ is the mean or center of the curve
* σ is the standard deviation of the curve (or the width)

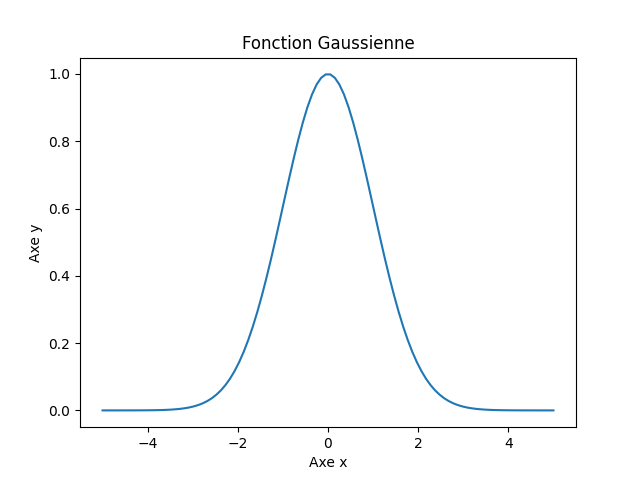


Figure 30: Gaussian function 3D with (A, μ, σ) (1, 0, 2) [28]

The curve of the Gaussian function is bell-shaped and symmetrical with respect to its mean. It is characterized by its amplitude, mean, and standard deviation.

The amplitude is the maximum value of the function, reached at the mean. The standard deviation determines the width of the curve.

The Gaussian function has several important properties that make it useful in various applications. One of its most important properties is that it is a solution to the diffusion equation. This means that it describes the diffusion of a quantity, such as temperature, in time and space. This property makes it useful for modeling various physical phenomena, such as heat diffusion and electromagnetic radiation.

The Gaussian function is also widely used in probability theory and statistics. In these fields, it is often used to model the distribution of random variables. For example, the normal distribution, which is a widely used probability distribution in statistics, is a special case of the Gaussian function.

In addition to its applications in physics, engineering, and mathematics, the Gaussian function is also used in various other fields, such as image processing, signal processing, and machine learning. In these fields, it is often used to smooth or filter data, model noise or error, or estimate the parameters of a model.

### 3.2.1. Two-Dimensional Gaussian Function

The two-dimensional Gaussian function (i.e., with two variables x and y) or two-dimensional Gaussian function is generally used in image processing, signal processing, computer vision, and machine learning.

The general form of the two-dimensional Gaussian function is as follows:



Where:

* A is the amplitude of the curve
* x0 and y0 are the means or centers of the curve along x and y, respectively
* σx and σy are the standard deviations of the curve along x and y, respectively

For the rest, we will assume that the curve of the function is centered at 0 along x as well as along y; so  and.

And that the standard deviations of the curve along x and y are the same; so.

The Gaussian function now becomes:

****

Our objective is to use this function to apply blur effects on images.

This blurring operation aims to:

* **Noise reduction**: Blurring can help smooth and reduce the impact of random noise or small variations in pixel intensity values.
* **Feature extraction**: By blurring, fine details in an image can be averaged, making it easier to detect more important and significant features in the image.
* **Image resizing**: Gaussian blurring can be used as an image resizing method, where the image is resized to a smaller size and then enlarged, producing a smoother and visually more appealing result.

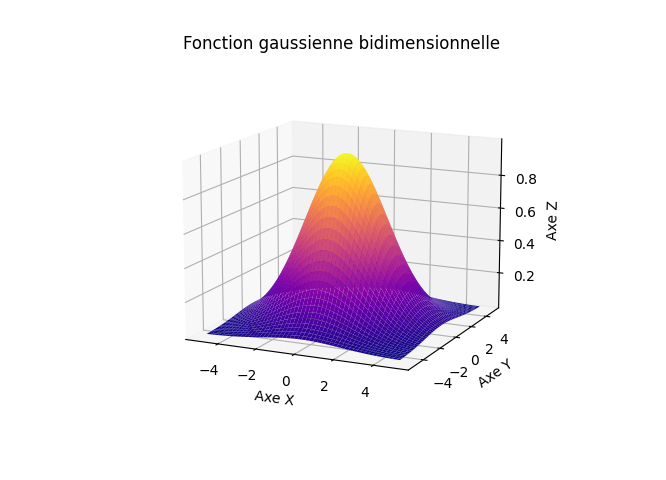


Figure 31: Gaussian function 3D with (A, μ, σ) (1, 0, 2) [28]

For the amplitude A of the curve, we will take: 

Indeed, the amplitude of the curve in the two-dimensional Gaussian function is often defined to normalize the curve and facilitate its comparison with other curves. The amplitude is chosen so that the integral over the entire surface of the curve is equal to 1 ( ). The mathematical justification for this amplitude can be found using **Fubini**'s theorem.



Let denote

By substituting  by his expression in the following relation, we obtain:







Then, we integrate this function g(x) with respect to x from minus infinity to plus infinity:



By equating this expression to 1 and solving for A, we obtain: 

This demonstrates that if we choose the amplitude of the Gaussian function to be, the integral over the entire surface of the curve is equal to 1.

Finally, we obtain:



Now that we have the final mathematical form of the Gaussian function, we need to write a code that utilizes this function to determine the Gaussian kernel.

### 3.2.2. Discrete form: The Gaussian kernel

A Gaussian kernel is a pixel matrix of size determined by the standard deviation () of the Gaussian function, which is used to weigh the neighboring pixels of the image during convolution. The Gaussian kernel is a discrete representation of the multidimensional Gaussian function, which is used to model the probability distribution of certain visual features in image processing. The Gaussian kernel is often used in the field of computer vision to reduce noise in an image, improve its quality, or perform a blurring operation.

The Gaussian kernel is a square matrix of size (2k+1) x (2k+1) where k is a positive integer. The coefficients of this matrix are calculated from the Gaussian function. More specifically, the coefficient located at the position of the Gaussian kernel is given by:

(30)



Where () is the standard deviation of the Gaussian distribution.

The size of the Gaussian kernel is generally chosen based on the standard deviation sigma. The larger the standard deviation, the flatter the Gaussian curve and the larger the kernel size needed to capture the entire curve. Conversely, if the standard deviation is small, the curve is sharper and a small kernel size may suffice.

The use of the Gaussian kernel in image filtering is based on the fact that convolving an image with a Gaussian kernel blurs the image while preserving the contours and important details. This is possible because the Gaussian function is a probability density function that evenly distributes values around the mean. Thus, by using a Gaussian kernel, more weight is given to the neighboring pixels of the image, while reducing the importance of more distant pixels.

In our code, we have created a method that allows us to obtain the Gaussian kernel based on the desired sigma and filter size.

|  |  |  |
| --- | --- | --- |
| * Size = 3 * = 0.8 |  | [[0.05212608 0.11385382 0.05212608]  [0.11385382 0.2486796 0.11385382]  [0.05212608 0.11385382 0.05212608]] |

Figure 33: Gaussien kernel value for sigma=0.8 and size=3

The features and equations of the different layers of the ESTMD model were described and demonstrated in Chapter 2. Therefore, in this section, we will focus on implementing these mathematical equations to view the results in a concrete manner and discuss them.

## 3.3. Retina Layer

### 3.3.1. Photoreceptor

**Implementation**

**Results**

|  |  |
| --- | --- |
|  |  |
| PHOTORECEPTOR | Blur effect |
|  |  |

Figure 35: Results after photoreceptor

Following the photoreceptor stage, we indeed achieved the blurring effect on our test images.

**Discuss:**

### 3.3.2. Lipetz transformation

**Implementation:**

**Results**:

|  |  |
| --- | --- |
|  |  |
| LIPETZ TRANSFORMATION | To membrane potential |
|  |  |

**Discuss:**

## 3.4. Lamina Layer

### 3.4.1. Low pass filter

|  |  |
| --- | --- |
|  |  |
| LOW PASS FILTER | Slight delay |
|  |  |

### 3.4.2. Large Monopolar Cell

|  |  |
| --- | --- |
|  |  |
| LMC | . Remove redundant information;  . Maximize information transmission |
|  |  |

## 3.5. Medulla Layer

In the first part of medulla Layer, we separate the output of LMCs using the following method:

### 3.5.1. FDSR

Implementation:

Results :

1. ON channel

|  |  |
| --- | --- |
|  |  |
| + last\_frame (frame 62)  LMC | + last\_frame (frame 62)  . Remove redundant information;  . Maximize information transmission |
|  |  |

1. OFF channel

|  |  |
| --- | --- |
|  |  |
| + last\_frame (frame 62)  LMC | + last\_frame (frame 62)  . Remove redundant information;  . Maximize information transmission |
|  |  |

Discuss

# GENERAL CONCLUSION

# BIBLIOGRAPGIE

[1] 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. 1–11, Jul. 2008.

[2] S. D. Wiederman and D. C. O’Carroll, “Biologically inspired feature  
detection using cascaded correlations of off and on channels,” Dec. 2013.

[3] H. Wang, J. Peng, and S. Yue, “A directionally selective small target  
motion detecting visual neural network in cluttered backgrounds,” IEEE  
Trans. Cybern., vol. 50, no. 4, pp. 1541–1555, Apr. 2020.

[4] H. Wang, J. Peng, X. Zheng, and S. Yue, “A robust visual system for  
small target motion detection against cluttered moving backgrounds,”  
IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 3, pp. 839–853,  
Mar. 2020.

[5] Ho. Wang, J. Zhao, Hu. Wang, J. Peng and S. Yue, “Attention and Prediction Guided Motion Detection for Low-contrast Small Moving Targets,”  
IEEE Trans, Dec. 2021.

[6] H. Wang, J. Peng, and S. Yue, “A Feedback Neural Network for Small Target Motion Detection in Cluttered Backgrounds,”  
July. 2018.

[7] Q. Fu, H. Wang, C. Hu, and S. Yue, “Towards Computational Models and Applications of Insect Visual Systems for Motion Perception: A Review,”  
IEEE Trans, Apr. 2019.

[8] Warrant, E.J.: Matched filtering and the ecology of vision in insects. In: The

Ecology of Animal Senses, pp. 143–167. Springer (2016)

[9] H. Wang, J Peng and S. Yue, “Bio-inspired Small Target Motion Detector with a new Lateral Inhibition Mechanism,”  
Conference Paper - July 2016.

10. Nordstro ̈m K, O’Carroll DC (2006) Small object detection neurons in female  
hoverflies. P Roy Soc B-Biol Sci 273: 1211–1216.

11. Nordstro ̈m K, Barnett PD, O’Carroll DC (2006) Insect detection of small targets  
moving in visual clutter. PLOS Biol 4: 378–386.

12. Barnett PD, Nordstro ̈m K, O’Carroll DC (2007) Retinotopic organization of  
small-field-target-detecting neurons in the insect visual system. Curr Biol 17:  
569–578.

[13] Bolzon, D. M., Nordstr ̈om, K., & O’Carroll, D. C. (2009). Local and large-  
range inhibition in feature detection. Journal of Neuroscience, 29 (45),  
14143–14150.

[14] "Computer Vision: Algorithms and Applications" by Richard Szeliski.

[15] "Background subtraction techniques: A review" by Radke et al.

[16] "Moving Object Detection Using Background Subtraction Techniques and Frame Difference" by Fatima et al.

[17] Nordstr ̈om, K. (2012). Neural specializations for small target detection in  
insects. Current Opinion in Neurobiology, 22 (2), 272–278.

[18] B. P. ̈Olveczky, S. A. Baccus, and M. Meister, “Segregation of object  
and background motion in the retina,” Nature, vol. 423, no. 6938, pp.  
401–408, 2003.

[19] 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.

[20] 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.

[21] 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 Informa-  
tion Processing (ISSNIP), 2011 Seventh International Conference on.  
IEEE, 2011, pp. 143–148.

[22]. Bagheri, Z. M., Cazzolato, B. S., Grainger, S., OʼCarroll, D. C., & Wiederman, S. D. (2017). An autonomous robot inspired by insect neurophysiology pursues moving features in natural environments. Journal of Neural Engineering, 14(4), 046030.

[23]. Bagheri, Z. M., Wiederman, S. D., Cazzolato, B. S., Grainger, S., & OʼCarroll, D. C. (2015). Properties of neuronal facilitation that improve target tracking in natural pursuit simulations. Journal of the Royal Society Interface, 12(108), 20150083.

[24]. Bagheri, Z. M., Wiederman, S. D., Cazzolato, B. S., Grainger, S., & OʼCarroll, D. C. (2017). Performance of an insect-inspired target tracker in natural conditions. Bioinspiration & Biomimetics, 12(2), 025006.

[25] [To download atom editor.](https://atom.en.uptodown.com/windows)

[26] PowerShell is available by default on Windows; it is also possible to use Terminal, which is available on Mac and Unix systems.

[27] Gaussian function 2D source code: <https://github.com/boukary-derra/pfe/blob/main/code/gauss/plot_gauss_1d.py>

[28] Gaussian function 3D source code : <https://github.com/boukary-derra/pfe/blob/main/code/gauss/plot_gauss_2d.py>

[29] <https://github.com/boukary-derra/pfe/tree/main/code>

[30] <https://github.com/boukary-derra/pfe/tree/main/code>

# ANNEXES





When 



Figure 16: Python class STMD / \_\_init\_\_ () method, inputs and constants



Figure 17: Python class STMD / Gaussian kernel method



Figure 18: Python model STMD: Retina LAYER



Figure 19: Python class STMD / Lamina LAYER



Figure 20: Python class STMD / Medulla LAYER



Figure 21 : Python class STMD / Lobula LAYER



Figure 22: Python class STMD / Other methods



Figure 23: Python class STMD / functions



Figure 24: Python class STMD / LI functions



Figure 25: Python code to convert a video into frames.





Figure 32: Gaussian kernel method in our python class STMD based on equation (30)

Figure 34: Python class STMD / photoreceptor method based on equation [1]



Figure 36: Python class STMD / lipetz transformation method based on equation [3] and 4



Figure 37: Python class STMD / low pass filter method based on equation [5]



Figure 38: Python class STMD / LMCs method based on equation [6] and [7]



Figure 39: Python class STMD / ON and OFF signals separation method based on equation [8] and [9]

