DISSERTATION

Research on artificial visual system for motion direction detection and motion speed perception based on HRC model

HRCモデルに基づく物体運動方向および速度検出の人工視覚システムに関する研究

Graduate School of
Natural Science and Technology
Kanazawa University

Division of Electrical Engineering and Computer Science

Student ID: 2024042001
Name: Chenyang Yan
Chief Advisor: Yuki Todo
Date of Submission: January 2023

Acknowledgement

I sincerely appreciate all those who have given me helpful advice and kind assistance during my doctoral research. This dissertation could be completed all respect to their concern and help.

Firstly, I am very grateful for the help of my chief advisor Yuki Todo. It is her inspiration that made me realize the charm of artificial intelligence. With her usual guidance and constant encouragement, I could survive the downturn and finally obtained this degree.

Secondly, I want to express my gratitude to Professor Zheng Tang for giving me his help. With his kind support, I could overcome difficulties throughout the period of my PhD program.

Thirdly, I want to appreciate all members of my laboratory. When I encountered noisy neighbors and had to sleep in the laboratory, it is their understanding that support me to complete my research.

Last but not least, I would like to show respect to my parents. It is their unrequited support that allows me to live in a foreign country without the burden of economics. Their deep love is the most valuable treasure in my life.

Abstract

The computer vision has been greatly improved with the development of neural networks. As one of the most popular research trends, searching for simple and efficient neural architectures is meaningful to solve more and more complex tasks. During my PhD program, I devoted myself to the research on bio-inspired artificial visual systems, which are described as follows:

Firstly, we cited the concept of simple cells, and designed the local motion-sensitive directionally detective neurons based on the core computation of the HRC model. With reference to the main excitatory ON motion signals processing in the *Drosophila*, we proposed the Full-neurons scheme motion detection mechanism for global motion direction detection. Considering the findings that the fruit fly ON motion pathway is sufficient to drive the optomotor response at high pattern contrast, a series of experiments have been conducted under the condition that the value of visual signals is 1 and the value of the background is 0. The experimental results have proved the reliability of the proposed mechanism. Furthermore, we applied the mechanism to an artificial visual system and compare the performance of our AVS with the time-considered CNN and EfficientNetB0 under the same experimental conditions. The comparison results further proved that our AVS is not only capable of global motion direction detection tasks, but also has an excellent performance in noise resistance.

Secondly, we cited the concept that temporal delay and sampling base determine the optimal velocity of a correlation-type motion detector and the characteristics of single unidirectional motion detector of the HRC model, proposed the local velocity-sensitive directionally detective neurons. With reference to the image speed can be detected by arrays of single unidirectional motion detectors and the fruit fly can receive more visual information than the inter-ommatidial angle at a specific

speed, we proposed the Temporal-based multi-neurons scheme motion detection mechanism for global motion speed perception. We applied the mechanism to the proposed artificial visual system and validate the reliability of our AVS through a series of experiments. The comparison results with 2-channels CNN have further proved that our AVS is not only capable of global motion speed perception tasks, but also has excellent performance in noise resistance.

The rest of my dissertation is structured as follows: Chapter 1 presents a detailed introduction to the background of this research. Chapter 2 describes the mechanism of the HRC model. Chapter 3 introduces the local motion-sensitive directionally detective neurons, the Full-neurons scheme motion detection mechanism and the performance of our AVS for global motion direction detection. Chapter 4 introduces the local velocity-sensitive directionally detective neurons, the Temporal-based multi-neurons scheme motion detection mechanism, and the performance of our AVS for global motion speed perception. Chapter 5 summarizes the conclusion and provides some future research.

Contents

List of Abbreviations

AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
HRC	Hassenstein-Reichardt Correlator
AVS	Artificial Visual System
CNN	Convolutional Neural Network
DSNs	Direction-Selective Neurons
DSGC	Direction-Selective Ganglion Cell
LPTC	Lobula Plate Tangential Cell
SAC	Starburst Amacrine Cell
EM	Electron Microscopy

Chapter 1 Introduction

1.1 Background

The phrase “Artificial Intelligence (AI)” was coined at the Dartmouth Summer Research Project on Artificial Intelligence in 1956 [1]. In the decade that followed, the AI field has gone through ups and downs [2], eventually numerous AI applications can be observed in our daily life nowadays. Computer vision, as one of the most popular AI research trends, has been greatly advanced with the development of Machine Learning (ML) and its sub-field Deep Learning (DL) in recent years [3]. However, the rapidly increasing number of computations and learning costs seems lead a phenomenon: the more computational architectures you applied in a neural network (NN), the better performance you got in feature extraction. To undo this phenomenon, searching for simple and efficient models has become the focus of computer science, which includes approach to collaborate with the field of neuroscience.

The human brain contains around 10^{11} neurons with 10^{15} connections between them [4]. It has long been considered as a hypercomplex neural network and many computational models have been proposed with reference to its functions [5,6]. As one major sensory input, visual plays a critical role in receiving information from surrounding environment [7]. In the past decades, scientists have discovered at least three separate visual processing systems (shape, color, and movement), and each system is dedicated to processing its own features [8]. Among these functional processing systems, motion perception is considered the most basic visual capability because the visual signals projected onto animal retina are rarely still, and this function is widespread presence in the visual system ranging from vertebrates to invertebrates [9]. Research on motion perception can be dated back to

1894, when Exner first presented a drawing of motion detection neural network [10]. In 1912, Wertheimer introduced the Phi phenomenon and revealed that motion could be observed in two spatially-stationary flashes [11]. The discovery that optic nerve fibers' activities can be recorded by the discharge of impulse makes it possible for biologists to conduct research on single ganglion cells [12]. After the observation of many retinal ganglion cells in the receptive field with light specific responses [13,14], neurons that respond to the movement in a direction-selective way have been observed in both the mammalian retina (direction-selective ganglion cells, DSGCs) and the insect optic lobe (lobula plate tangential cells, LPTCs) [15,16]. As advances in techniques, more details about neuronal circuits in both mammals and flies have been investigated [17]. Direction-selective neurons (DSNs), whether they are in the mammalian retina or the insect optic lobe, have been proved to firstly gather local motion signals from the same array of photoreceptors [18]. Then, motion signals are processed in separate channels depending on the characteristics of the light increments or light decrements and received by different types of DSNs [19]. In mammals, two main types of DSGCs have been observed that respond to light increments or both light increments and decrements. They receive excitatory input from glutamatergic bipolar cells through cones and both excitatory input and inhibitory input from starburst amacrine cells (SACs) [20,21]. Such neurons can also be found in the optic lobe of fruit fly in which T4 neurons and T5 neurons are necessary for processing ON motion signals and OFF motion signals [22]. The LPTCs can both receive direct excitatory input from T4/T5 neurons and indirect inhibitory input from lobula plate intrinsic (LPi) cells [23]. Last but not least, motion signals from local region can be integrated to drive the optomotor response. Research on the fruit fly motion pathway points out that LPTCs sum signals from medulla columns and generate a wide-field response [24]. Similarly, the concept of simple cells and complex cells which originated from physiological experiments on cats have also supported this perspective [25].

These investigations encourage researchers to construct artificial visual systems (AVS) from the perspective of neural computations. Till now, a good number of bio-inspired motion detection AVS have been proposed and many of them have been applied in robots [26].

The recovery of motion direction is considered to be the first step to extract apparent-motion information [27]. Therefore, scientists have proposed a variety of motion-detector models to understand the mechanism underlying signals processing in direction-selective (DS) way. Theoretically, a biological-based correlation-type motion-detector model has to obey at least three fundamental rules: (1) The motion-detector model needs at least two inputs. (2) The motion-detector model needs a non-linear interaction between the input signals. (3) The motion-detector model needs to be structural asymmetry [28]. In 1956, Hassenstein and Reichardt based on the optomotor behavior of beetle *Chlorophanus viridis*, proposed the first computationally based motion-detector model, the so-called HRC model [29]. It is an excitatory scheme correlation-type model which consists of two mirror-symmetrical subunits, and each subunit computes it motion direction by preferred-direction enhancement [30]. In 1965, Barlow et al. analyzed the responses of direction-selective ganglion cells in rabbit retina, proposed an inhibitory scheme motion-detector model which computes motion direction by null-direction suppression [31]. In 1985, Adelson et al. scientifically studied human vision and proposed the spatiotemporal energy model [32]. These computational models are considered as three dominant algorithms that have long been used to interpret the motion detection in insects, vertebrate retina, and vertebrate visual cortex [33]. Although they differ in the details of calculating the direction of motion, they all need to fulfill the above three fundamental rules [34].

Despite many research results have been achieved in understanding direction selectivity at the cellular level, the full systemic mechanism of motion detection in animal brains remains elusive [35].

Besides, it is difficult to investigate the mechanisms underlying visual motion pathways only through limit neuronal inputs and physiological experiments [36]. Therefore, searching for reliable motion detection mechanisms is important not only for future research in neuroscience, but also for the development of computer science.

1.2 Motivation

The main purpose of this research is to validate the mechanisms we proposed for global motion direction detection and global motion speed perception in a two-dimensional view. In order to obtain the local motion information, we cited the core computation of the HRC model and designed the local motion-sensitive directionally detective neurons and local velocity-sensitive directionally detective neurons. In the study of AVS for global motion direction detection, we propose the Full-neurons scheme motion detection mechanism for detecting the direction of global motion. Through a series of experiments, we prove the reliability of our AVS and conclude that the mechanism is suitable for global motion direction detection in a two-dimensional view. In the study of AVS for global motion speed perception, we design the local velocity-sensitive directionally detective neurons based on both the characteristics of single unidirectional motion detector of the HRC model and the biological findings. The Temporal-based multi-neurons scheme motion detection mechanism is proposed for detecting the direction and speed of global motion and has been applied to an artificial visual system. We test the performance of our AVS in the datasets those have different conditions of background and compare it with a 2-channels CNN. The experimental results have proved that the mechanism we proposed is suitable for global motion speed perception in a two-dimensional view.

1.3 Outline

The dissertation is organized as follows:

In Chapter 1, we present the background of this research.

In Chapter 2, we describe the mechanism of Hassenstein-Reichardt Correlator model.

In Chapter 3, we introduce the realization of AVS for global motion direction detection.

In Chapter 4, we introduce the realization of AVS for global motion speed perception.

In Chapter 5, we conclude our studies and introduce the future works.

Chapter 2 Hassenstein-Reichardt Correlator model

The nervous system in fly, because its far fewer neurons than mammalian nervous system, has always been a popular choice for functional studies of visual system [37]. In the studies of fruit fly, two types of columnar cells have been identified to detect the direction of local motion (T4 neurons that respond primarily to ON motion signals and T5 neurons that respond primarily to OFF motion signals) [38]. They are considered to be the first neurons that represent directional selectivity in the visual motion pathways and depolarize (hyperpolarize) during the objects moving in their preferred direction (null direction) [39]. As shown in Figure 2.1, either columnar T4 neurons or columnar T5 neurons comprise four main subtypes and each subtype is sensitive to one of four cardinal directions. These findings have been revealed with the application of calcium imaging [40]. How does LPTCs gather the local motion information to produce a wide-field response? It arises the focus on studies of fruit fly neuronal circuits in the past decades [41-43]. With the development of experimental approaches, we now know that: (1) At high pattern contrast, either ON motion pathway or OFF motion pathway is sufficient to drive the optomotor response [44]. (2) Lamina cells L1 receives signals from photoreceptors and is one main input for processing the ON motion signals [45]. (3) Medulla cells Mi1 and Tm3 are two main pathways bridging between L1 and T4, together contribute most synaptic inputs from L1 to T4 [46]. (4) Both Mi1 and Tm3 exhibit to process the excitatory inputs onto T4, blocking them will lead the Drosophila nearly lose the optomotor response of ON motion stimuli [47]. (5) Intracellular dynamics within columnar T4 neurons are suggested as a critical role in detecting the direction of fast-motion [48]. (6) LPTCs sum signals from medulla columns and generate a wide-field response [24]. These findings provide evidence for motion direction could be detected by excitatory inputs enhancement in the preferred direction, which coincides with the mechanism of HRC model [49].

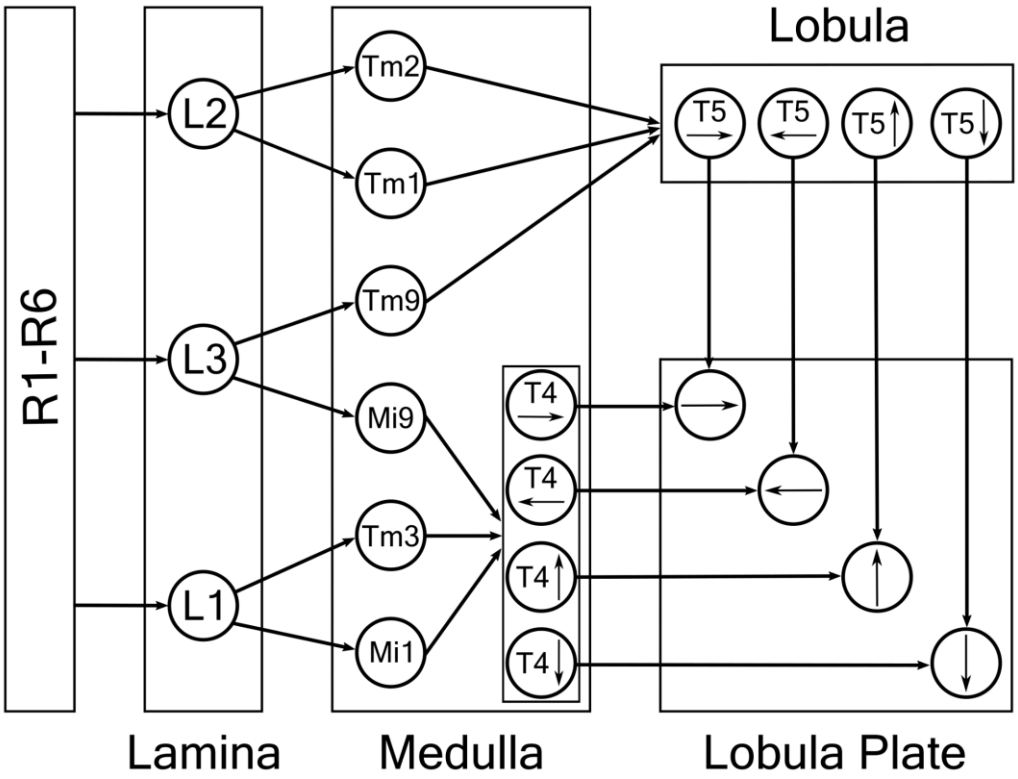

Figure 2.1 General motion-detection neural circuits in *Drosophila*.

2.1 HRC for motion direction detection

The HRC model consists of two mirror-symmetrical subunits and each subunit has its own preferred direction (as shown in Figure 2.2a). In its most basic form, the HRC model receives inputs from two spatially separated receptors in which one of the inputs has been temporally delayed. Figure 2.2b shows the rightwards motion detector of the HRC model. The detector will respond when there comes a rightwards shift but evoke nearly no response to a leftwards one. For further explanation, in Figure 2.2c, we use a yellow dot to indicate the visual signal at time T and a red dot to indicate that signal at time T+ΔT. If we mathematically define the value of signals as 1, the computation in its preferred direction at time T+ΔT is $1 \times 1 = 1$. Thus, the rightwards motion detector will output the signal. In contrast, the detector will fail to evoke in detecting the leftwards shift because the

calculation at time T+ΔT is $0 \times 1 = 0$.

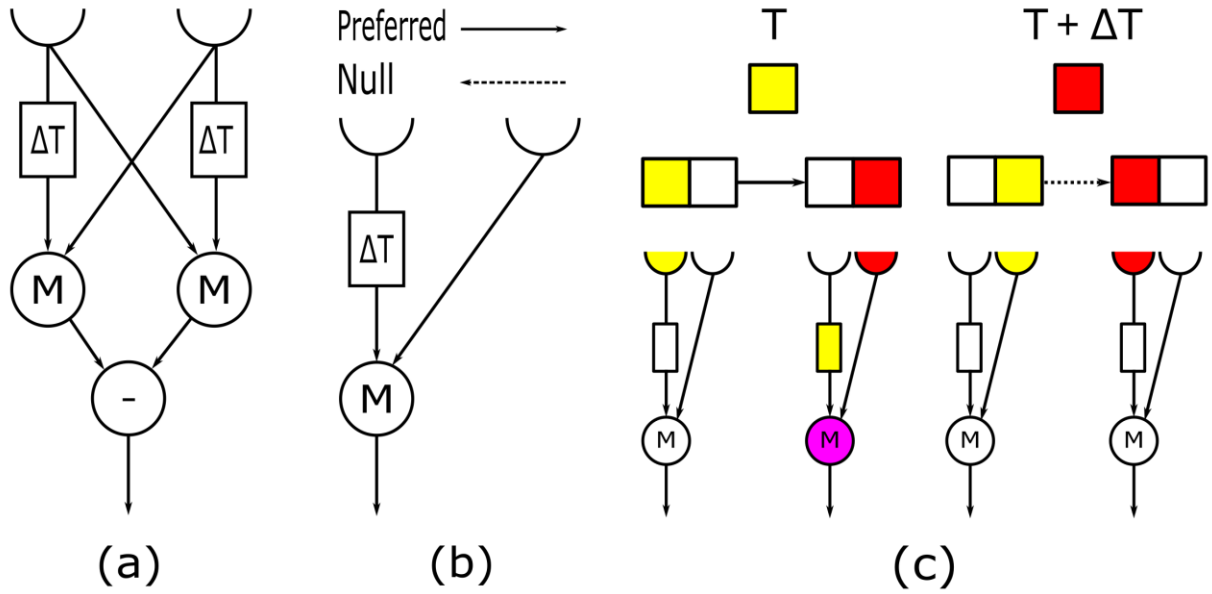

Figure 2.2 (**a**) Schematic of the HRC model. (**b**) Schematic of the rightwards motion detector of the HRC model. (**c**) Process of the rightwards motion detector in detecting its preferred direction and null direction. In its preferred direction, the activation signal caused by the visual signal at time T (yellow) will be delayed until time T+ΔT and then multiply with the activation signal caused by the visual signal at time T+ΔT (red), thus, the rightwards motion detector will respond (pink). In converse, the rightwards motion detector will evoke no response in its null direction detection because the signal from the receptor at time T has not been temporally delayed.

2.2 HRC for motion speed perception

A single motion detector of the HRC model has long been discussed that respond best not only at a preferred speed, but also at a certain intensity. In 1989, Egelhaaf et al. discussed the properties of the half-detector and pointed out the measurement of speed can be done with simple mechanism [50]. Subsequently, investigations of image speed can be detected by arrays of single unidirectional motion detectors have been conducted [51,52]. Figure 2.3 shows the visual signals processing in different rightwards motion detectors. We use yellow dots to indicate the light spot moving in the

visual field from left to right and the processing signals in motion detectors are indicated by yellow color. If we mathematically define the motion detectors are activated by the intensity of value 1, the core computation in their preferred direction is $1 \times 1 = 1$. With different spatial structures and temporal frequency, we can measure the image velocity of 1 (ΔX=1, D=1), 2 (ΔX=2, D=1) and 1/2 (ΔX=1, D=2).

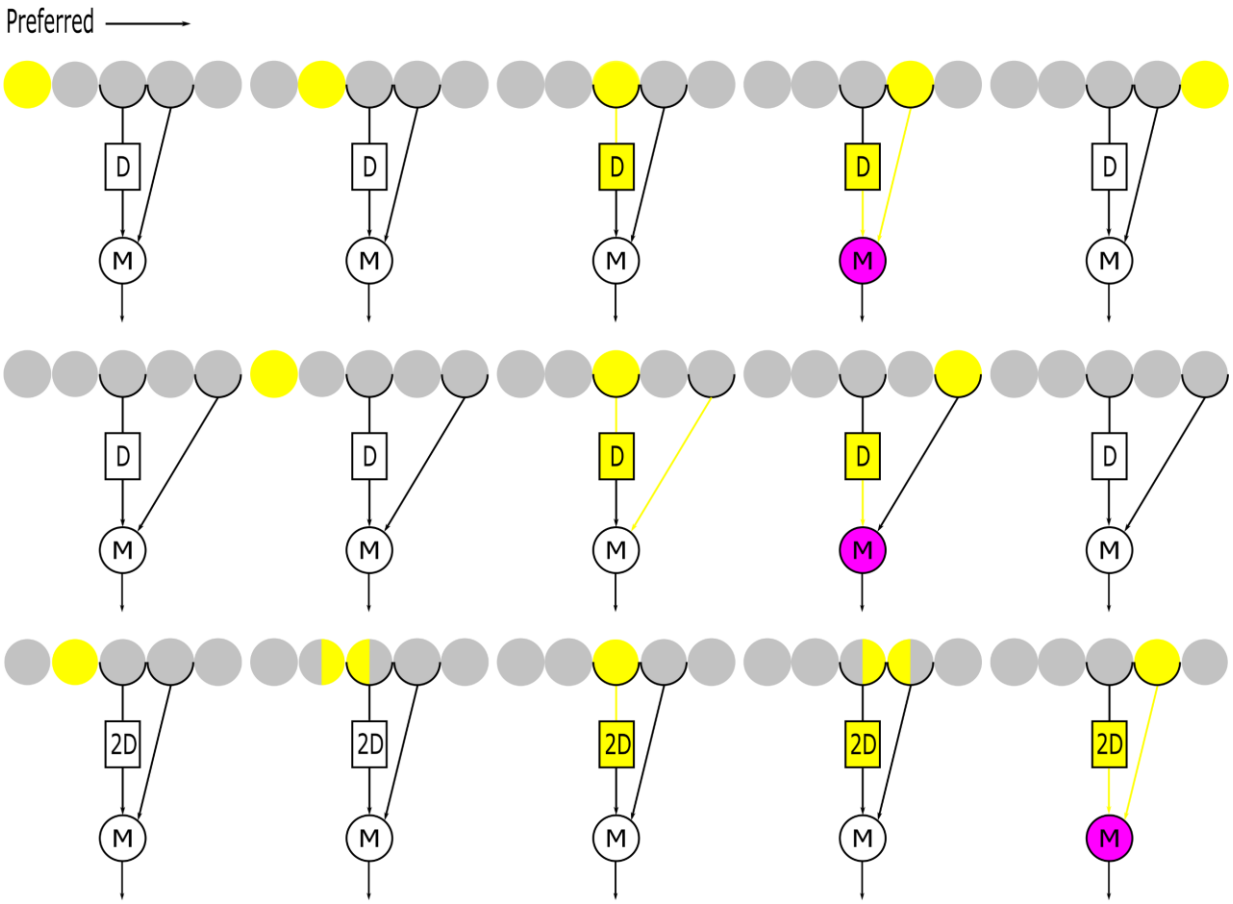

Figure 2.3 Schematic of different rightwards motion detectors based on the core computation of the HRC model.

Chapter 3 AVS for global motion direction detection

In this chapter, we introduce the realization of an AVS for global motion direction detection systematically.

3.1 Method

Neurons are the fundamental units of visual nervous system. Since 1962, Hubel and Wiesel introduced the concept of simple cells and complex cells, single neuron can complete a specific task has been acknowledged by the studies of biological neural networks. Combining the findings that signals from simple cells are collected by complex cells in the Hubel-Wiesel model and local motion information from medulla columns are gathered by LPTCs to produce a wide-field response in the fruit fly motion pathways, in this research, we employ the local motion-sensitive directionally detective neurons to detect the direction of local motion and gather them to generate the global motion direction with the Full-neurons scheme motion detection mechanism.

Considering each simple cell has a fixed receptive field, we designed our local motion-sensitive directionally detective neurons with the fixed size of 3×3, for simplicity, eight neurons are employed to detect eight potential direction of local motion separately (0°, 45°, 90°, 135°, 180°, 225°, 270° and 315°). Inspired by signals from the same position in the visual field can be transmitted by different ommatidia in the compound eyes, we assume that each photoreceptor in the visual field has its own eight local motion-sensitive directionally detective neurons. With the similarity to LPTCs sum signals from medulla columns and generate the wide-field response, we accumulate the activated number of neurons in detecting the same direction and mark the maximum value as the global motion direction.

3.2 Realization

3.2.1 Local motion-sensitive directionally detective neurons

A biological motion detector is not only response to the motion signals, but also sensitive to the temporal brightness modulations of a stationary stimulus [50], thus, we design our local motion-sensitive directionally detective neurons through spatiotemporally separated photoreceptors to receive light signals directly from the visual filed. Considering the fruit fly ON motion pathway is sufficient to drive the optomotor response at high pattern contrast, in this research, we assign the value 1 to the visual signals and value 0 to the background. Figure 3.1a shows the theoretical structure of a rightwards local motion-sensitive directionally detective neuron. If we use notations $X_{(i, j, T)}$ to indicate the signal from the centrally located photoreceptor $P_{(i, j)}$ at time T, and $X_{(i+1, j, T+\Delta T)}$ denote the signal from its right-side photoreceptor $P_{(i+1, j)}$ at time $T+\Delta T$. The calculation in Figure 3b can be expressed by the following equation:

$$Y_{R(i, j)} = X_{(i, j, T)} \cdot X_{(i+1, j, T+\Delta T)} \tag{3-1}$$

If - and only if – the activation result of $Y_{R(i, j)}$ is 1, the rightwards local motion-sensitive directionally detective neuron located at $P_{(i, j)}$ will be activated.

Moreover, we extend our proposed neurons to two-dimensional multi-directions detection. With reference to the receptive field of simple cells, we employ eight local motion-sensitive directionally detective neurons for detecting eight directions, respectively. As shown in Figure 3.1c, they are 0°-detective neuron (Rightward, R), 45°-detective neuron (Upper Rightward, UR), 90°-detective neuron (Upward, U), 135°-detective neuron (Upper Leftward, UL), 180°-detective neuron (Leftward, L), 225°-detective neuron (Lower Leftward, LL), 270°-detective neuron (Downward, D), and 315°-detective neuron (Lower Rightward, LR).

The equations of extended neurons' activation results can be expressed as follows:

$$Y_{UR(i,j)} = X_{(i,j,T)} \cdot X_{(i+1,j+1,T+\Delta T)} \tag{3-2}$$

$$Y_{U(i,j)} = X_{(i,j,T)} \cdot X_{(i,j+1,T+\Delta T)} \tag{3-3}$$

$$Y_{UL(i,j)} = X_{(i,j,T)} \cdot X_{(i-1,j+1,T+\Delta T)} \tag{3-4}$$

$$Y_{L(i,j)} = X_{(i,j,T)} \cdot X_{(i-1,j,T+\Delta T)} \tag{3-5}$$

$$Y_{LL(i,j)} = X_{(i,j,T)} \cdot X_{(i-1,j-1,T+\Delta T)} \tag{3-6}$$

$$Y_{D(i,j)} = X_{(i,j,T)} \cdot X_{(i,j-1,T+\Delta T)} \tag{3-7}$$

$$Y_{LR(i,j)} = X_{(i,j,T)} \cdot X_{(i+1,j-1,T+\Delta T)} \tag{3-8}$$

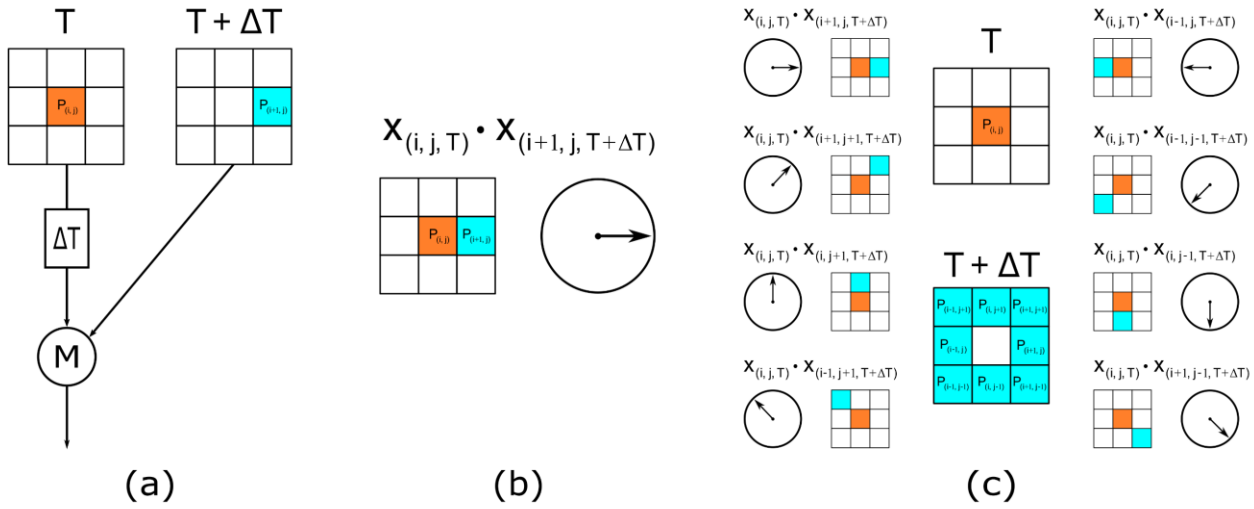

Figure 3.1 (a) The theoretical structure of the rightwards local motion-sensitive directionally detective neuron. (b) Schematic of 0°-detective neuron. The rightwards local motion-sensitive directionally detective neuron will respond only if the photoreceptor P(i, j) at time T and the photoreceptor P(i+1, j) at time T+ΔT receive signals. (c) Schematic of eight local motion-sensitive directionally detective neurons.

3.2.2 Full-neurons scheme motion detection mechanism

With the development of biological science, neuronal circuits of the fruit fly have been well-studied. We now know that the visual processing in *Drosophila* begins in the optic lobe, and the principal regions of the optic lobe are four neuropils (the lamina, medulla, lobula, and lobula plate). Here, we demonstrate the main excitatory process of ON motion signals detected by *Drosophila*. When an ON-motion stimuli appears in the visual field, the photoreceptor neurons (R1–R6) in the retinal ommatidia will first receive visual signals and transmit them to lamina monopolar cells in the first neuropil (the lamina). Then, medulla columnar neurons (e.g., Mi1 and Tm3) receive signals from the lamina neurons and project inputs to the dendritic arbors of T4 in the M10 layer of the second neuropil (medulla). Finally, LPTCs receive signals from T4 cells, summed to generate wide-field motion responses.

Inspired by biophysical investigations of *Drosophila*, we propose an artificial visual system using the Full-neurons scheme motion detection mechanism. Firstly, we hypothesize that each light spot in the visual field can be received by different local motion-sensitive directionally detective neurons. In this research, we assume that each light spot at time T can be received by 8 local motion-sensitive directionally detective neurons. Secondly, we design our local motion-sensitive directionally detective neurons with the core computation of the HRC model. Each local motion-sensitive directionally detective neuron will be activated when it receives signals from both the null direction side photoreceptor at time T and the preferred direction side photoreceptor at time $T+\Delta T$ respectively. Thirdly, we accumulate the number of the activated neurons with the same preferred direction and regard it as the activation strength in that direction. The AVS will give the detection result based on the maximum value of activation strength in eight directions.

For a visual field of size $M \times N$ region, we use $P_{(m, n)}$ to denote the photoreceptor located at the

m-th row and n-th column ($1 \leqslant m \leqslant M$, $1 \leqslant n \leqslant N$). The Full-neurons scheme motion detection mechanism can be expressed by the following equations:

$$Z_R = \sum_{m=1}^{M} \sum_{n=1}^{N} Y_{R(m, n)} \tag{3-9}$$

$$Z_{UR} = \sum_{m=1}^{M} \sum_{n=1}^{N} Y_{UR(m, n)} \tag{3-10}$$

$$Z_U = \sum_{m=1}^{M} \sum_{n=1}^{N} Y_{U(m, n)} \tag{3-11}$$

$$Z_{UL} = \sum_{m=1}^{M} \sum_{n=1}^{N} Y_{UL(m, n)} \tag{3-12}$$

$$Z_L = \sum_{m=1}^{M} \sum_{n=1}^{N} Y_{L(m, n)} \tag{3-13}$$

$$Z_{LL} = \sum_{m=1}^{M} \sum_{n=1}^{N} Y_{LL(m, n)} \tag{3-14}$$

$$Z_D = \sum_{m=1}^{M} \sum_{n=1}^{N} Y_{D(m, n)} \tag{3-15}$$

$$Z_{LR} = \sum_{m=1}^{M} \sum_{n=1}^{N} Y_{LR(m, n)} \tag{3-16}$$

$$Out = \max\{Z_R, Z_{UR}, Z_U, Z_{UL}, Z_L, Z_{LL}, Z_D, Z_{LR}\} \tag{3-17}$$

Figure 3.2 shows the detection process of the AVS for a 1-pixel object. We use notations $P_{(1, 1)}$, $P_{(1, 2)}$, $P_{(1, 3)}$, …, $P_{(5, 5)}$ to indicate the positions of photoreceptors in the visual field, so the detection process in Figure 4 can be described as: Firstly, the light spot at time T (yellow dot) is received by the centrally located photoreceptor $P_{(3, 3)}$ of the 5 × 5 visual field. Secondly, the outputs of each

photoreceptor are sent to 8 local motion-sensitive directionally detective neurons, which will be delayed until time T+ΔT and received by the center point of each local motion-sensitive directionally detective neuron's 3 × 3 local receptive field. Thirdly, each local motion-sensitive directionally detective neuron will detect whether there exists a signal from the photoreceptor located in its preferred direction at time T+ΔT (orange dot center, blue dots surrounding it). In this flowchart, the photoreceptor located at $P_{(3, 3)}$ receives the light signal at time T, and $P_{(3, 4)}$ receives the signal at time T+ΔT (red dot), so the 0°-detective neuron is activated (pink dot). Finally, the AVS gives the detection result based on the maximum value of activation strength in eight directions. Since the 0°-detective neuron located at $P_{(3, 3)}$ is the only activated neuron during this detection, the global motion direction of this 1-pixel object is Rightward.

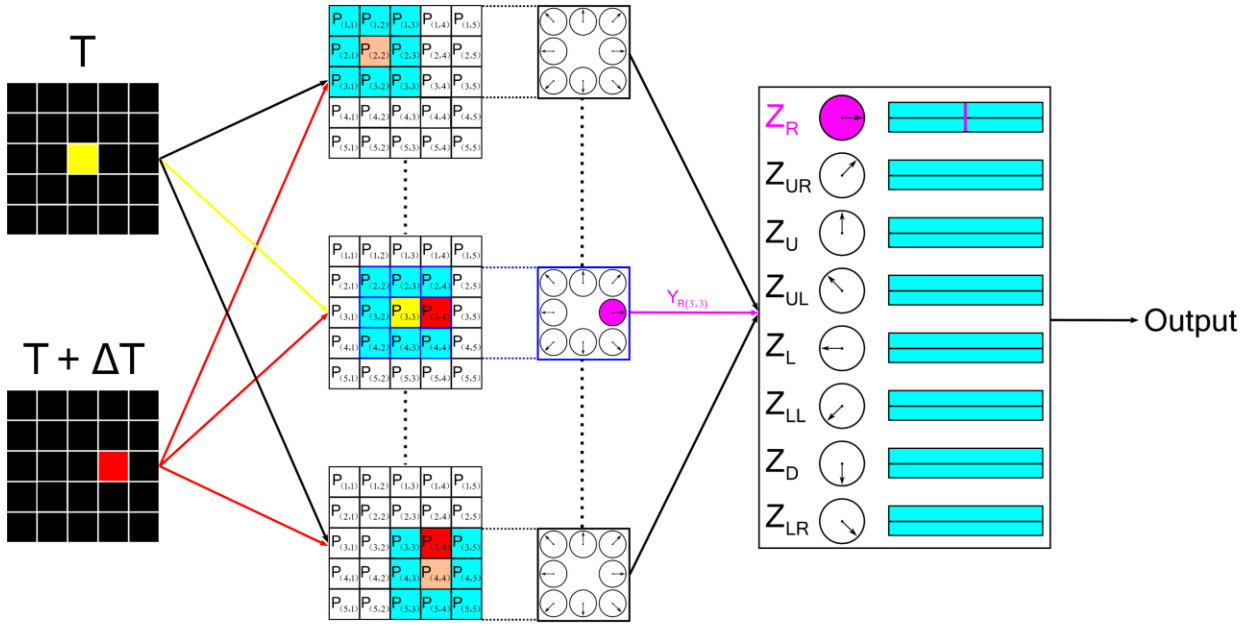

Figure 3.2 Flowchart of the AVS in detecting 1-pixel object.

A more complex process of the AVS for a 3-pixel object is presented in Figure 3.3. During this detection, a total of 5 local motion-sensitive directionally detective neurons have been activated.

They are 0°-detective neuron and 45°-detective neuron located at $P_{(3, 2)}$, 45°-detective neuron located at $P_{(4, 2)}$, 45°-detective neuron, and 90°-detective neuron located at $P_{(4, 3)}$. As the 45°-detective neurons achieve the maximum value of the activation strength, the global motion direction of this 3-pixel object is Upper Rightward.

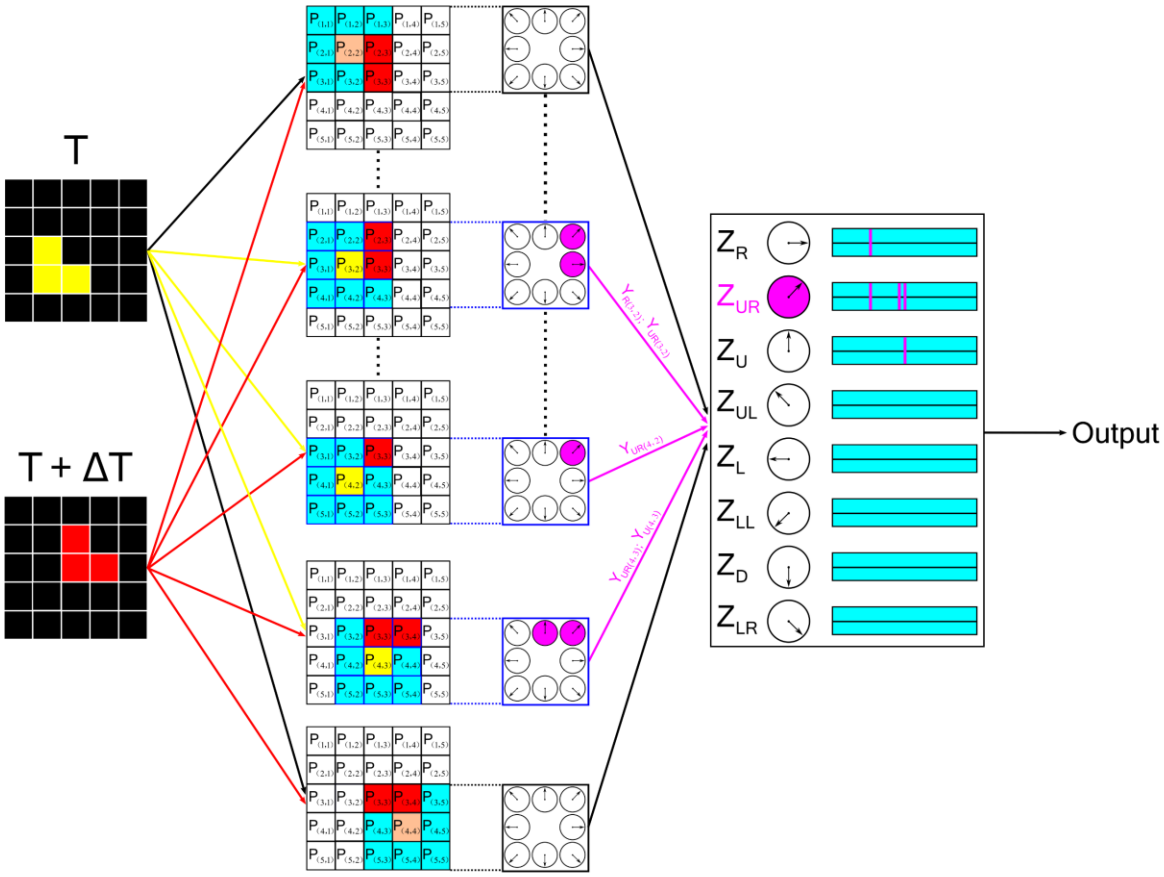

Figure 3.3 Flowchart of the AVS in detecting 3-pixels object.

3.3 Simulation and Result

To validate the reliability of our AVS, we conduct a series of experiments with a background size of 32 × 32, 1,024-pixel datasets. The objects in all experiments have random sizes ranging from 1 to 128 pixels, and their shapes and positions are set to arbitrary.

In the first series of experiments, we test our AVS within the background with no noise. The

dataset has totally 192,000 images (each size of object has 24,000 images) and the detection results are presented in Table 3.1. As all responses in the dataset are caused by the moving objects, so the detection results of our AVS are 100%. The activation plots of the AVS for a 16-pixels object motion direction detection within the background with no noise have been shown in Fig 3.4. The activation numbers of 8 local motion-sensitive directionally detective neurons are: 0°-detective neurons 6, 45°-detective neurons 5, 90°-detective neurons 6, 135°-detective neurons 4, 180°-detective neurons 9, 225°-detective neurons 9, 270°-detective neurons 16 and 315°-detective neurons 9. Since the 270°-detective neurons get the maximum value of activation strength, so the global direction of this 16-pixels object's motion is Downward.

Table 3.1 Detection results of the AVS within the background with no noise.

Model	**1**	**2**	**4**	**8**	**16**	**32**	**64**	**128**
AVS	100%	100%	100%	100%	100%	100%	100%	100%

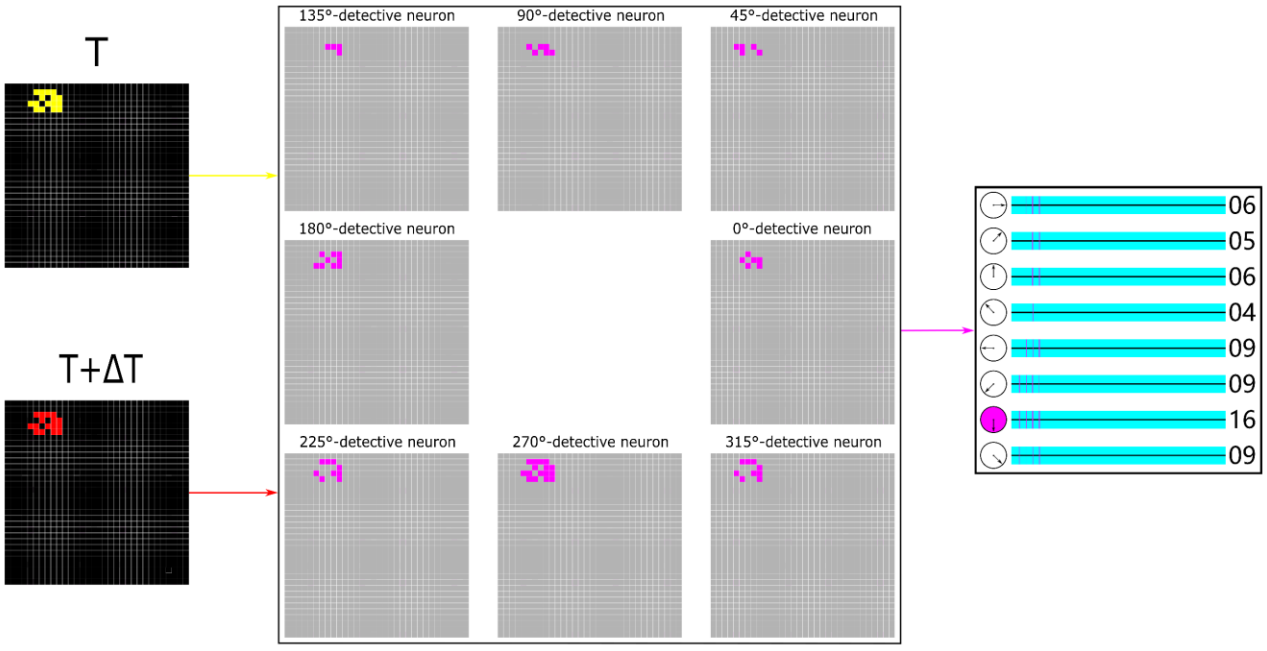

Figure 3.4 The activation plots of the AVS for 16-pixels object motion direction detection within the background with no noise.

In the second series of experiments, we test our AVS within the background with 1% to 10% separated noises. As we mentioned above, we designed our local motion-sensitive directionally detective neurons based on the core computation of a classically biological motion detector, so the effect of stationary stimulus should be considered during the global motion direction detection. In this research, we defined the light spot which does not have spatial motion between the time T and the time T+ΔT as noise. The dataset has totally 192,000 images (each size of object has 24,000 images) and the detection results are presented in Table 3.2. The reason why separated noises have no effect on the detection results is because no local motion-sensitive directionally detective neurons have been activated by noises. Therefore, all responses in the dataset are still caused by the moving objects. Figure 3.5 shows an example of the activation plots of the AVS for a 32-pixels object motion direction detection within the background with 10% separated noises. The activation numbers of 8 local motion-sensitive directionally detective neurons are: 0°-detective neurons 10, 45°-detective neurons 7, 90°-detective neurons 13, 135°-detective neurons 13, 180°-detective neurons 32, 225°-detective neurons 13, 270°-detective neurons 13 and 315°-detective neurons 13. Since the 180°-detective neurons get the maximum value of activation strength, so the global direction of this 32-pixels object's motion is Leftward.

Table 3.2 Detection results of the AVS within the background with separated noises.

Model	**1**	**2**	**4**	**8**	**16**	**32**	**64**	**128**
1%	100%	100%	100%	100%	100%	100%	100%	100%
2%	100%	100%	100%	100%	100%	100%	100%	100%
5%	100%	100%	100%	100%	100%	100%	100%	100%
10%	100%	100%	100%	100%	100%	100%	100%	100%

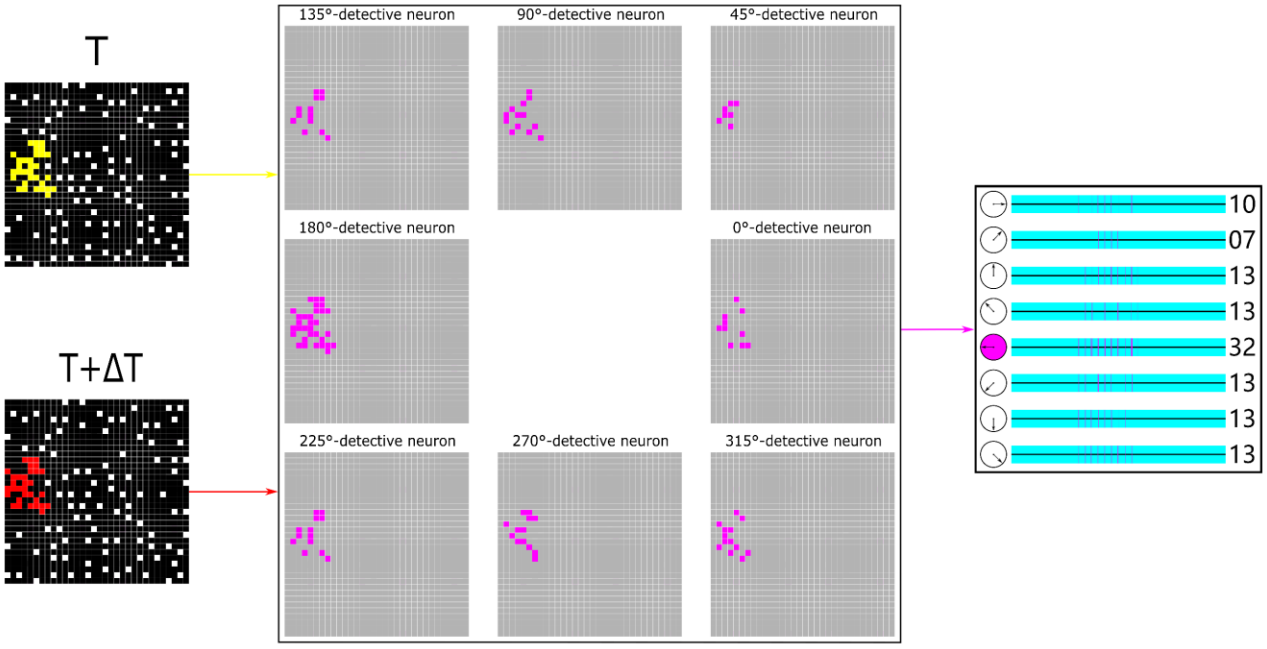

Figure 3.5 The activation plots of the AVS for 32-pixels object motion direction detection within the background with 10% separated noises.

In the third series of experiments, we test our AVS within the background with 1% to 10% connected noises. The dataset has totally 192,000 images (each size of object has 24,000 images) and the detection results are presented in Table 3.3. Through observation, it can be found that the connected noises do have impacts on the detection results, especially when the object's size is small. However, as the size of the object increases, the detection results will return to high accuracy. The activation plots of the AVS for a 64-pixels object motion direction detection within the background with 2% connected noises have been presented in Fig 3.6, and the error activations due to noise are represented by the purple dots. The activation numbers of 8 local motion-sensitive directionally detective neurons are: 0°-detective neurons 38, 45°-detective neurons 36, 90°-detective neurons 35, 135°-detective neurons 30, 180°-detective neurons 30, 225°-detective neurons 36, 270°-detective neurons 43 and 315°-detective neurons 64. Since the 315°-detective neurons get the maximum value of activation strength, so the global direction of this 64-pixels object's motion is Lower Rightward.

Table 3.3 Detection results of the AVS within the background with connected noises.

Model	1	2	4	8	16	32	64	128
1%	81.6%	96.0%	99.8%	100%	100%	100%	100%	100%
2%	56.7%	84.0%	97.9%	99.9%	100%	100%	100%	100%
5%	36.6%	52.1%	75.0%	95.1%	99.8%	100%	100%	100%
10%	30.7%	37.8%	52.3%	74.1%	94.5%	99.8%	100%	100%

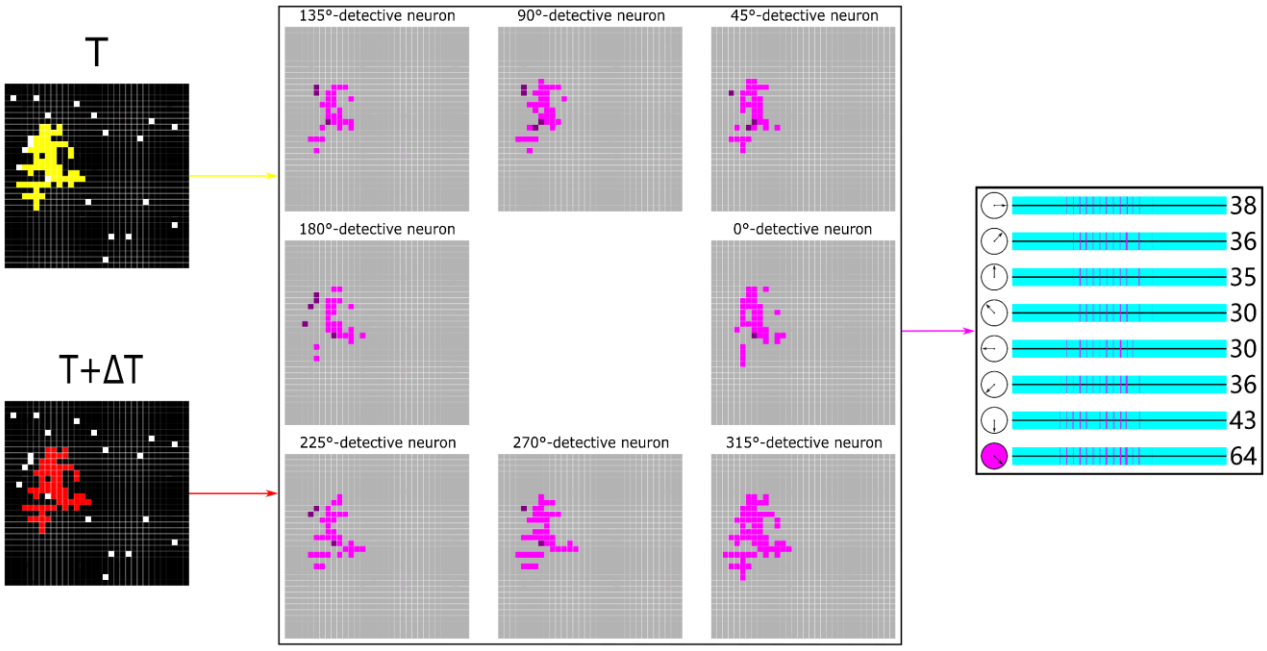

Figure 3.6 The activation plots of the AVS for 64-pixels object motion direction detection within the background with 2% connected noises.

In the last series of experiments, we compare the performances of the AVS, the time-considered CNN and the EfficientNetB0. In recent research, the convolutional architectures are proved to perform well in tasking audio synthesis and the convolutional networks are claimed to be regarded as the natural starting point for sequence modeling tasks [53]. Therefore, we propose a time-considered CNN and train it for the task of detecting the global direction of motion. The time-considered CNN is made up of 8 convolution kernels (3 × 3), 8 feature maps (32 × 32), and 8 outputs. We also compare the performance of our AVS with the EfficientNetB0 and the structure of the time-

considered CNN and EfficientNetB0 are shown in Figure 3.7.

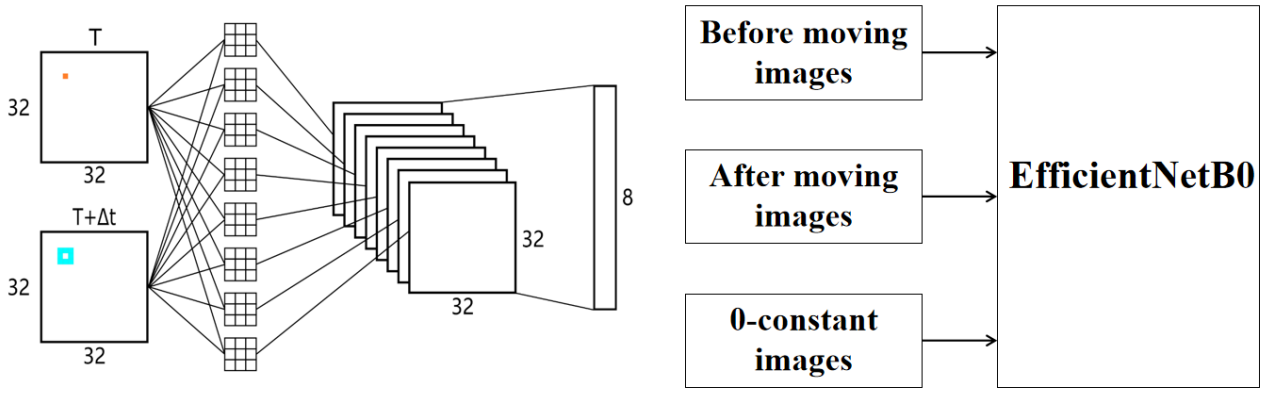

Figure 3.7 Schematic of the time-considered CNN and EfficientNetB0.

At first, we use the dataset with no noises in the background and randomly picked 3,000 images out of 24,000 images from each size of object's dataset to form a new dataset. Then, we pick 15,000 images in the new dataset for training the CNN and the EfficientNetB0, use 5,000 of the remaining images to form the test dataset. Finally, we save the CNN and the EfficientNetB0 when the training accuracy reached 100% and compared their performances with the proposed AVS. The accuracy rates on the test dataset are: AVS 100%, the time-considered CNN 63.6% and the EfficientNetB0 100%. Furthermore, we do the comparison among three models in detecting the dataset with separated noises and the dataset with connected noises. The detection results are shown in Table 3.4. We can easily observe from the table that our AVS has better performance than the time-considered CNN and the EfficientNetB0. In order to further compare these three systems, we use the dataset which the background with no noises and only includes the size of object is 1 for training. 18,000 images are used as the training data and the rest of them are used for testing. The detection results of three systems on test data are all 100% and the three systems in detecting other sizes of objects within the background with no noises are present in Table 3.5. Through observation, we can conclude that our AVS is reliable for global motion direction detection. Since our AVS cites the spatial

structure of the basic form of the HRC model, the requirements on the dataset are almost none, while the CNN and the EfficientNetB0 need to learn different sizes of objects and cost a lot of data for training.

Table 3.5 Detection results of three systems within the background with no noise.

Models	2	4	8	16	32	64	128
AVS	100%	100%	100%	100%	100%	100%	100%
CNN	91.7%	69.5%	39.1%	27.3%	21.0%	12.5%	12.5%
EfficientNetB0	74.0%	67.4%	61.7%	56.0%	53.7%	54.0%	53.5%

3.4 Summary

In this research, we validate the reliability of our AVS with a series of experiments. The high accuracy experimental results have proved that the Full-neurons scheme motion detection mechanism is suitable for global motion direction detection in a two-dimensional view. Moreover, we compare the performance of our AVS with the time-considered CNN and EfficientNetB0 under the same experimental conditions. The comparison results further proved that our AVS is not only capable of global motion direction detection tasks, but also has an excellent performance in noise resistance.

Table 3.4 Detection results of AVS, the time-considered CNN and EfficientNetB0.

Background Noises	Object Sizes	Models / Noise Types / Accuracy (%)	AVS		Time-considered CNN		EfficientNetB0	
			Separated	Connected	Separated	Connected	Separated	Connected
1%	1		100	81.6	14.8	15.2	29.3	27.8
	2		100	96.0	19.7	17.4	36.9	34.2
	4		100	99.8	20.0	20.3	43.8	43.0
	8		100	100	25.4	25.7	50.1	51.2
	16		100	100	36.0	36.4	58.4	58.6
	32		100	100	55.6	55.9	64.8	63.1
	64		100	100	78.6	79.9	69.9	72.5
	128		100	100	95.0	95.1	81.5	82.8
2%	1		100	56.7	13.8	13.6	15.0	15.4
	2		100	84.0	14.7	14.7	18.4	18.2
	4		100	97.9	16.1	15.8	22.1	22.2
	8		100	99.9	18.4	18.8	28.4	30.8
	16		100	100	23.9	23.9	37.2	36.0
	32		100	100	36.7	37.2	43.9	42.7
	64		100	100	58.2	62.1	52.2	52.4
	128		100	100	86.9	88.9	66.8	66.1
5%	1		100	36.6	12.9	12.9	14.0	13.8
	2		100	52.1	13.5	13.4	14.5	14.1
	4		100	75.0	13.7	14.1	15.7	13.4
	8		100	95.1	14.4	15.0	17.2	16.9
	16		100	99.8	16.9	16.5	22.6	20.4
	32		100	100	21.3	20.8	29.8	28.2
	64		100	100	28.2	33.8	42.2	38.8
	128		100	100	55.8	65.7	58.0	55.4
10%	1		100	30.7	12.7	12.7	12.9	13.3
	2		100	37.8	12.9	13.0	11.9	13.0
	4		100	52.3	13.3	13.4	12.1	14.1
	8		100	74.1	13.5	13.9	12.7	17.4
	16		100	94.5	15.6	15.3	15.2	20.3
	32		100	99.5	18.6	17.9	18.9	26.4
	64		100	100	21.1	25.3	23.6	33.0
	128		100	100	37.8	50.6	30.5	39.9

Chapter 4 AVS for global motion speed perception

In this chapter, we introduce the realization of an AVS for global motion speed perception systematically.

4.1 Method

Neurons that not only exhibit direction selectivity, but also respond maximum to their preferred speed have been reported their existence in both mammals and insects. Such neurons have been observed through biophysical experiments on honeybee [54], rabbits [55], cats [56,57], and human [58]. In 1975, Levinson and Sekuler analyzed the experimental results of human vision and proposed the additive model and the substituted model [59,60]. In 1989, Egelhaaf et al. discussed the properties of the half-detector and pointed out the image speed can be measured by mechanism of correlation-type model. In 1998, Zanker et al. further studied the elementary motion detectors and figured out that the subunit of HRC model could account for velocity.

With reference to findings that the temporal delay and sampling base determine the optimal velocity of a motion detector, in this research, we employ three kinds of local velocity-sensitive directionally detective neurons for detecting not only the direction of local motion, but also the velocity of it. Consistent with the concept of simple cells, each neuron is designed to have its own preferred direction and preferred velocity. For simplicity, eight major directions and three types of velocity are taken into account. Inspired by the image speed can be detected by arrays of unidirectional motion detectors, we propose the Temporal-based multi-neurons scheme motion detection mechanism to achieve the global motion information. The local motion information is aggregated by summation and the global motion depends on the activation strength of each neuron.

4.2 Realization

4.2.1 Local velocity-sensitive directionally detective neurons

Based on the characteristic that single motion detector of the HRC model is not only sensitive to its preferred direction, but also respond best to its preferred speed and signal intensity, in this research, we proposed three kinds of local velocity-sensitive directionally detective neurons based on the core computation of the HRC model. As introduced above, the detection is under the condition that the visual signals have high contrast to the background (value 1 to the visual signals, value 0 to the background). We consider the effect of stationary stimulus to a biological motion detector, so design our local velocity-sensitive directionally detective neurons through spatiotemporally separated photoreceptors to receive light signals directly from the visual filed. Figure 4.1 shows the theoretical structures of rightwards local velocity-sensitive directionally detective neurons. Considering the recent findings that the fruit fly R1-R6 photoreceptors' encoding capability is maximized to fast high-contrast bust [61], we design our local velocity-sensitive directionally detective neurons with different theoretical structures and assume the sampling base are only increase at the high temporal frequency.

If we use notations $X_{(i, j, T)}$ to indicate the signal from the centrally located photoreceptor $P_{(i, j)}$ at time T, and $X_{(i+1, j, T+\Delta T)}$ denote the signal from its right-side photoreceptor $P_{(i+1, j)}$ at time $T+\Delta T$. The calculation in Figure 4.1a can be expressed by the following equation:

$$V1_{R(i, j)} = X_{(i, j, T)} \cdot X_{(i+1, j, T+\Delta T)} \tag{4-1}$$

If - and only if – the activation result of $V1_{R(i, j)}$ is 1, the V1 rightwards local velocity-sensitive directionally detective neuron will be activated.

Similarly, the calculation in Figure 4.1b and Figure 4.1c can be expressed as follows:

$$V2_{R(i,j)} = X_{(i,j,T)} \cdot X_{(i+2,j,T+\Delta T)} \tag{4-2}$$

$$V1/2_{R(i,j)} = X_{(i,j,T)} \cdot X_{(i+1,j,T+2\Delta T)} \tag{4-3}$$

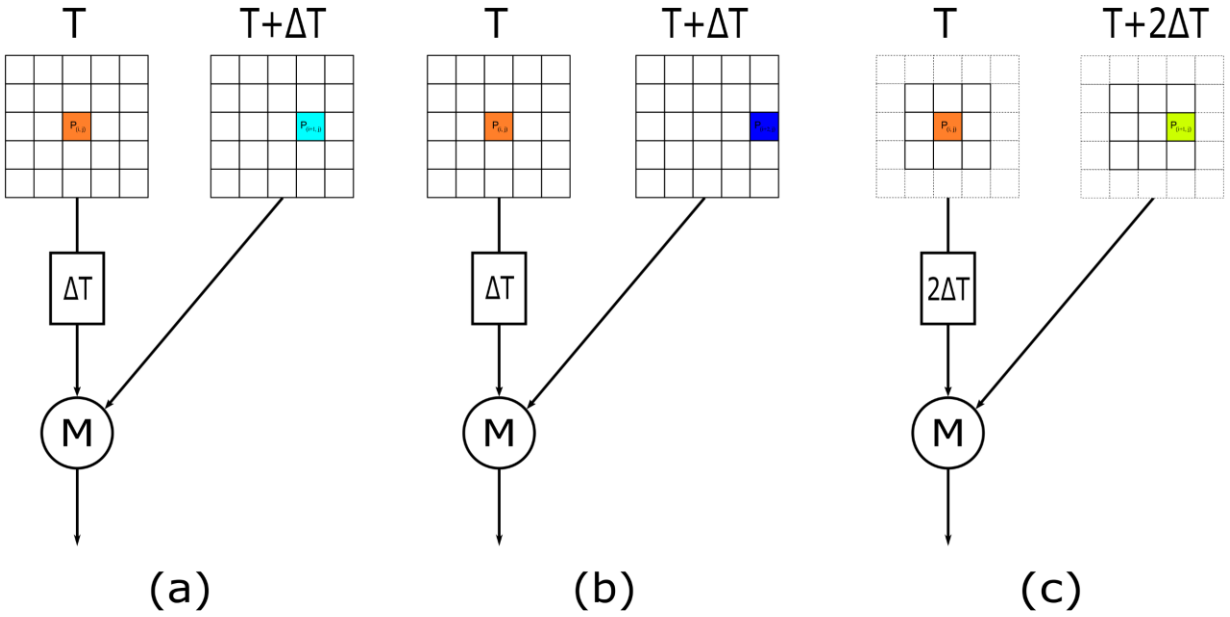

Figure 4.1 (a) The theoretical structure of V1 rightwards local velocity-sensitive directionally detective neuron. (b) The theoretical structure of V2 rightwards local velocity-sensitive directionally detective neuron. (c) The theoretical structure of V1/2 rightwards local velocity-sensitive directionally detective neuron.

Furthermore, the proposed neurons can be easily extended to detect multi-directions in a two-dimensional view. For simplicity, in this research, we limit our discussion to the objects are only moving in eight major directions. As shown in Figure 4.2, they are: Rightward ($V1_R$, $V2_R$, $V1/2_R$), Upper Rightward ($V1_{UR}$, $V2_{UR}$, $V1/2_{UR}$), Upward ($V1_U$, $V2_U$, $V1/2_U$), Upper Leftward ($V1_{UL}$, $V2_{UL}$, $V1/2_{UL}$), Leftward ($V1_L$, $V2_L$, $V1/2_L$), Lower Leftward ($V1_{LL}$, $V2_{LL}$, $V1/2_{LL}$), Downward ($V1_D$, $V2_D$, $V1/2_D$), and Lower Rightward ($V1_{LR}$, $V2_{LR}$, $V1/2_{LR}$). 16 neurons are output signals at time T+ΔT and the rest 8 neurons are output signals at time T+2ΔT. The equations of extended neurons' activation results are expressed as follows:

$$V1_{UR(i,j)} = X_{(i,j,T)} \cdot X_{(i+1,j+1,T+\Delta T)} \quad (4\text{-}4)$$

$$V2_{UR(i,j)} = X_{(i,j,T)} \cdot X_{(i+2,j+2,T+\Delta T)} \quad (4\text{-}5)$$

$$V1/2_{UR(i,j)} = X_{(i,j,T)} \cdot X_{(i+1,j+1,T+2\Delta T)} \quad (4\text{-}6)$$

$$V1_{U(i,j)} = X_{(i,j,T)} \cdot X_{(i,j+1,T+\Delta T)} \quad (4\text{-}7)$$

$$V2_{U(i,j)} = X_{(i,j,T)} \cdot X_{(i,j+2,T+\Delta T)} \quad (4\text{-}8)$$

$$V1/2_{U(i,j)} = X_{(i,j,T)} \cdot X_{(i,j+1,T+2\Delta T)} \quad (4\text{-}9)$$

$$V1_{UL(i,j)} = X_{(i,j,T)} \cdot X_{(i-1,j+1,T+\Delta T)} \quad (4\text{-}10)$$

$$V2_{UL(i,j)} = X_{(i,j,T)} \cdot X_{(i-2,j+2,T+\Delta T)} \quad (4\text{-}11)$$

$$V1/2_{UL(i,j)} = X_{(i,j,T)} \cdot X_{(i-1,j+1,T+2\Delta T)} \quad (4\text{-}12)$$

$$V1_{L(i,j)} = X_{(i,j,T)} \cdot X_{(i-1,j,T+\Delta T)} \quad (4\text{-}13)$$

$$V2_{L(i,j)} = X_{(i,j,T)} \cdot X_{(i-2,j,T+\Delta T)} \quad (4\text{-}14)$$

$$V1/2_{L(i,j)} = X_{(i,j,T)} \cdot X_{(i-1,j,T+2\Delta T)} \quad (4\text{-}15)$$

$$V1_{LL(i,j)} = X_{(i,j,T)} \cdot X_{(i-1,j-1,T+\Delta T)} \quad (4\text{-}16)$$

$$V2_{LL(i,j)} = X_{(i,j,T)} \cdot X_{(i-2,j-2,T+\Delta T)} \quad (4\text{-}17)$$

$$V1/2_{LL(i,j)} = X_{(i,j,T)} \cdot X_{(i-1,j-1,T+2\Delta T)} \quad (4\text{-}18)$$

$$V1_{D(i,j)} = X_{(i,j,T)} \cdot X_{(i,j-1,T+\Delta T)} \quad (4\text{-}19)$$

$$V2_{D(i,j)} = X_{(i,j,T)} \cdot X_{(i,j-2,T+\Delta T)} \quad (4\text{-}20)$$

$$V1/2_{D(i,j)} = X_{(i,j,T)} \cdot X_{(i,j-1,T+2\Delta T)} \quad (4\text{-}21)$$

$$V1_{LR(i,j)} = X_{(i,j,T)} \cdot X_{(i+1,j-1,T+\Delta T)} \quad (4\text{-}22)$$

$$V2_{LR(i,j)} = X_{(i,j,T)} \cdot X_{(i+2,j-2,T+\Delta T)} \quad (4\text{-}23)$$

$$V1/2_{LR(i,j)} = X_{(i,j,T)} \cdot X_{(i+1,j-1,T+2\Delta T)} \quad (4\text{-}24)$$

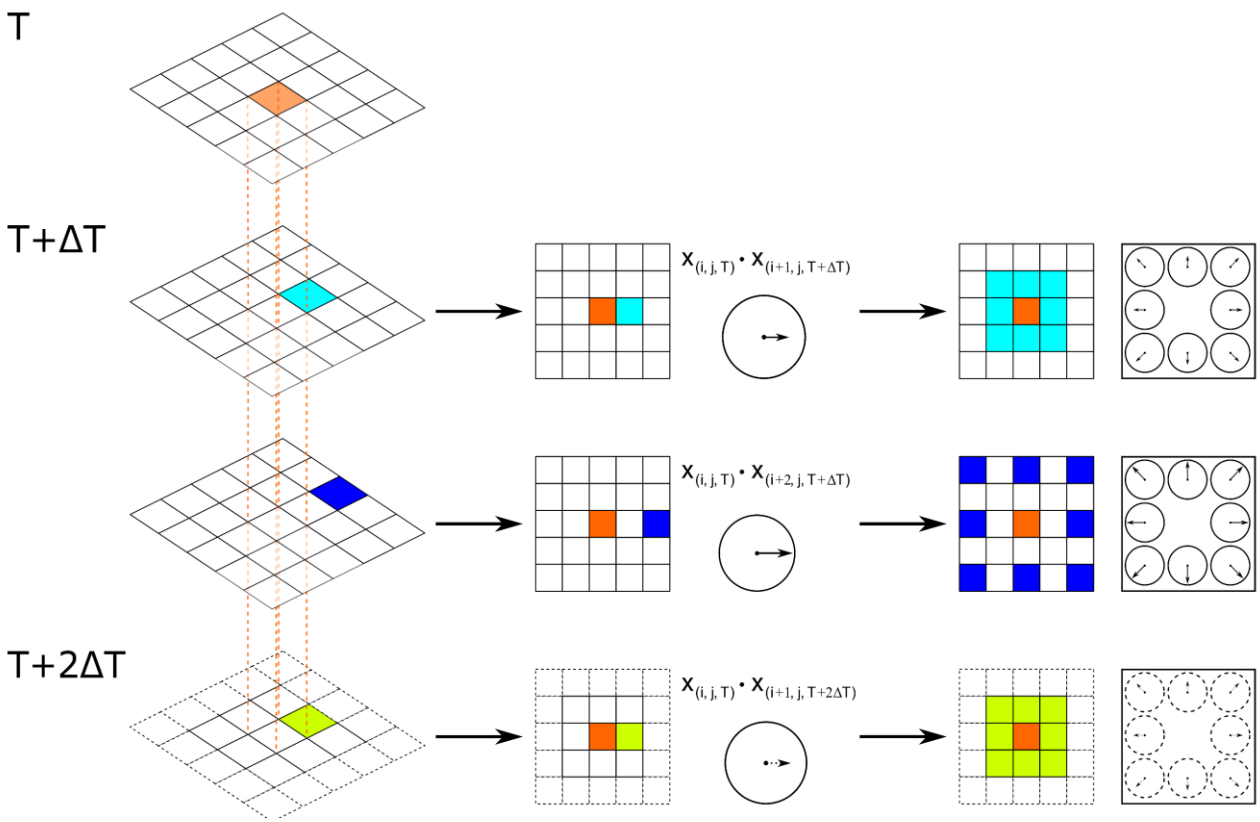

Figure 4.2 Schematic of local velocity-sensitive directionally detective neurons.

4.2.2 Temporal-based multi-neurons scheme motion detection mechanism

As previously stated, the temporal delay and sampling base determine the optimal velocity of a correlation-type motion detector and the image speed can be detected by arrays of unidirectional motion detectors. Moreover, recent research on the fruit fly proved evidence that *Drosophila* can receive more angles of visual information than the structures of ommatidia at a specific speed [62]. Therefore, we propose an artificial visual system which uses the Temporal-based multi-neurons scheme motion detection mechanism to detect the speed and direction of global motion. Figure 4.3 shows the AVS in detecting 1-pixel object at time T+ΔT. Without loss of generality, we use 16 neurons scan every region simultaneously from $P_{(1, 1)}$ to $P_{(5, 5)}$ over the two-dimensional receptive field at time T and respond to yield the local motion directions of the regions. During this detection,

the $V2_R$ local velocity-sensitive directionally detective neuron is the only activation neuron which is activated once time at the location $P_{(3, 3)}$, thus, the global motion direction of this object is Rightward and the velocity of it is 2.

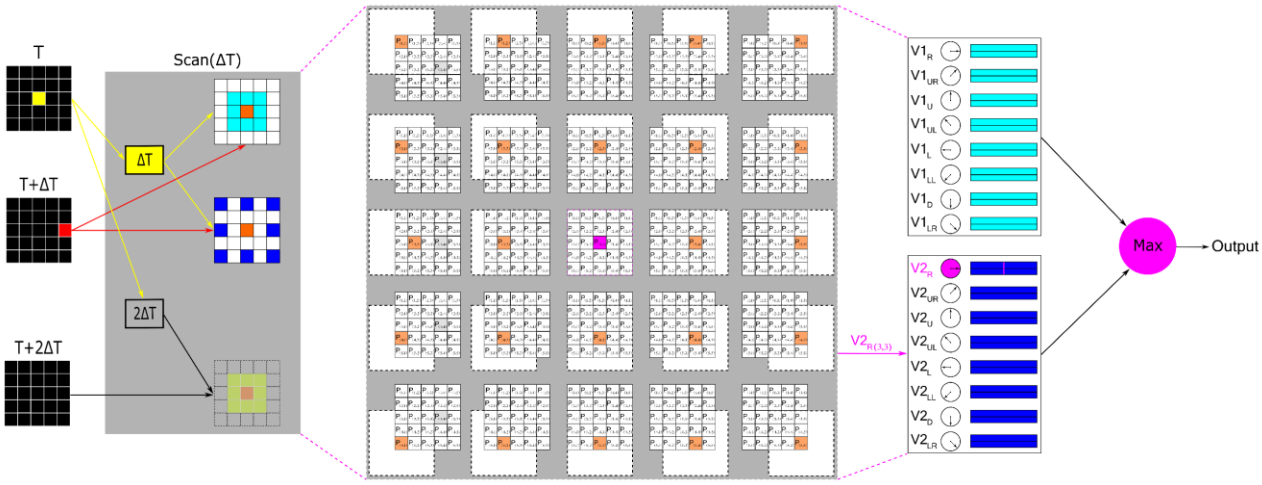

Figure 4.3 Flowchart of the AVS in detecting 1-pixel object.

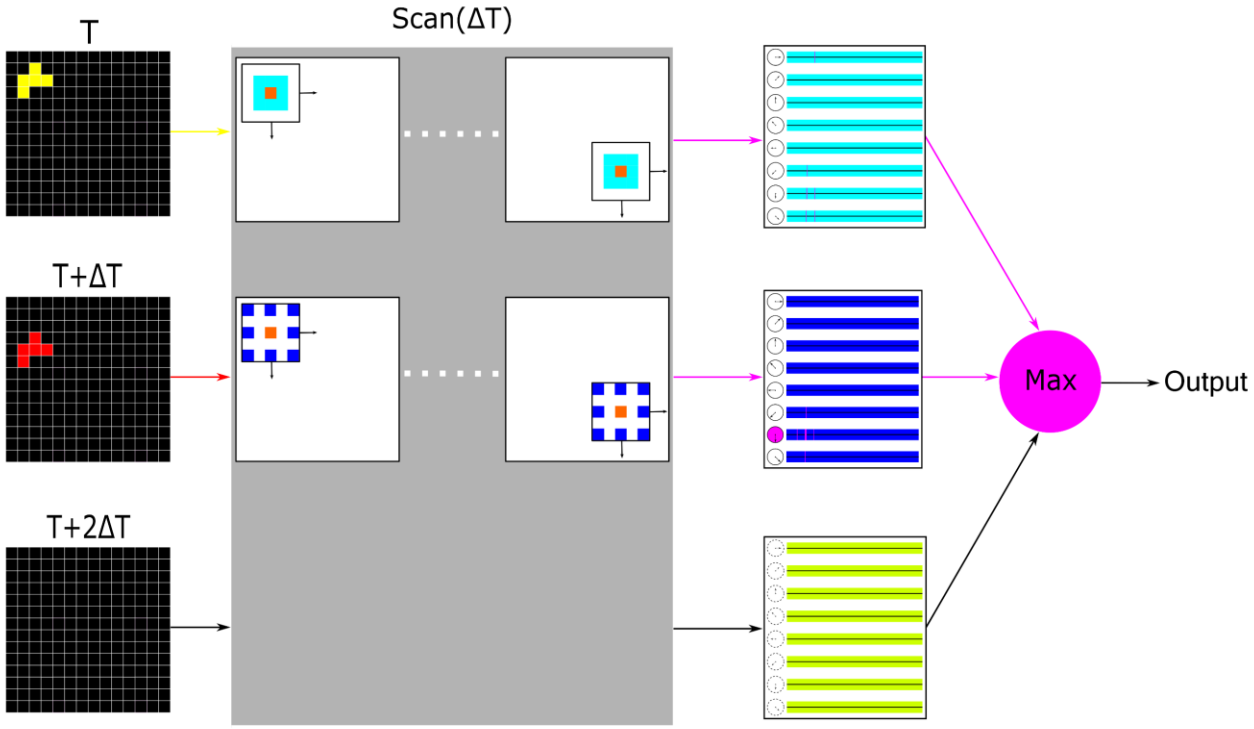

Figure 4.4 Flowchart of the AVS in detecting 4-pixels object.

A more complex process of the AVS for a 4-pixels object in receptive field's size of 14 × 14 is presented in Figure 4.4. During this detection, the $V2_D$ local velocity-sensitive directionally detective neuron get the maximum value of activation strength, so the global motion direction of this object is Downward and the velocity of it is 2.

4.3 Simulation and Result

To validate the reliability of our AVS, we conduct a series of experiments with a background size of 32 × 32, 1,024-pixel datasets. The objects in all experiments have random sizes ranging from 1 to 128 pixels, and their shapes and positions are set to arbitrary. In order to better differentiation of the velocity, we design three types of speed dataset (Dataset V1, Dataset V2 and Dataset V1/2) and totally 3 × 240,000 images are tested to examine the performance of our AVS.

In the first series of experiments, we test our AVS within the background with no noise. The dataset has totally 240,000 images (each size of object has 10,000 images and each type of velocity dataset has totally 80,000 images) and the detection results are presented in Table 4.1.

Table 4.1 Detection results of the AVS within the background with no noise.

Dataset	1	2	4	8	16	32	64	128
V1	100%	100%	100%	100%	100%	100%	100%	100%
V2	100%	100%	100%	100%	100%	100%	100%	100%
V1/2	100%	100%	100%	100%	100%	100%	100%	100%

The activation plots of the AVS for a 16-pixels object motion direction detection within the background with no noise (Dataset V1) have been shown in Fig 4.5. As the detection occurs at time T+ΔT, 16 local velocity-sensitive directionally detective neurons are applied to scan over every point in the visual field. The activation strength of each neuron is: $V1_R$-neuron 7, $V2_R$-neuron 5, $V1_{UR}$-neuron 4, $V2_{UR}$-neuron 2, $V1_U$-neuron 4, $V2_U$-neuron 0, $V1_{UL}$-neuron 3, $V2_{UL}$-neuron 1, $V1_L$-

neuron 8, $V2_L$-neuron 6, $V1_{LL}$-neuron 16, $V2_{LL}$-neuron 9, $V1_D$-neuron 10, $V2_D$-neuron 6, $V1_{LR}$-neuron 6, and $V2_{LR}$-neuron 3. Since the $V1_{LL}$-neuron get the maximum value of activation strength, the global direction of this 16-pixels object is Lower Leftward and the velocity of it is 1.

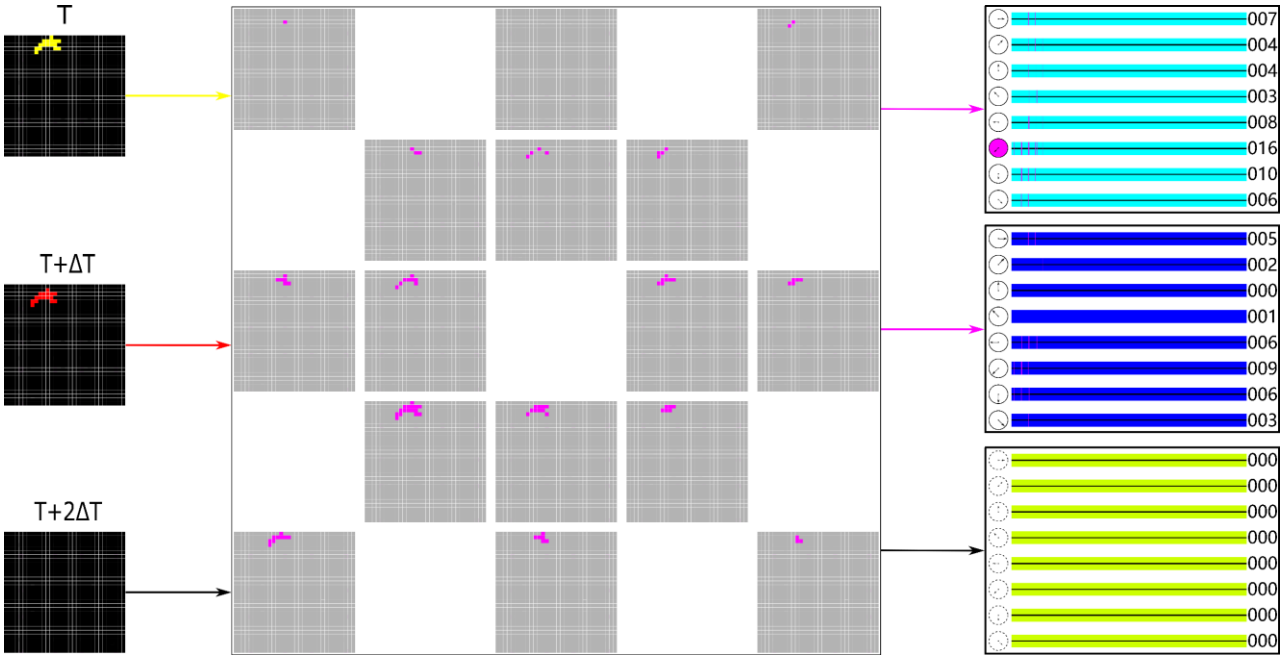

Figure 4.5 The activation plots of the AVS for 16-pixels object motion direction detection within the background with no noise (Dataset V1).

The activation plots of the AVS for a 64-pixels object motion direction detection within the background with no noise (Dataset V2) have been shown in Fig 4.6. As the detection occurs at time T+ΔT, 16 local velocity-sensitive directionally detective neurons are applied to scan over every point in the visual field. The activation strength of each neuron is: $V1_R$-neuron 36, $V2_R$-neuron 38, $V1_{UR}$-neuron 29, $V2_{UR}$-neuron 29, $V1_U$-neuron 22, $V2_U$-neuron 21, $V1_{UL}$-neuron 21, $V2_{UL}$-neuron 13, $V1_L$-neuron 23, $V2_L$-neuron 15, $V1_{LL}$-neuron 24, $V2_{LL}$-neuron 17, $V1_D$-neuron 27, $V2_D$-neuron 32, $V1_{LR}$-neuron 35, and $V2_{LR}$-neuron 64. Since the $V2_{LR}$-neuron get the maximum value of activation strength, the global direction of this 64-pixels object is Lower Rightward and the velocity of it is 2.

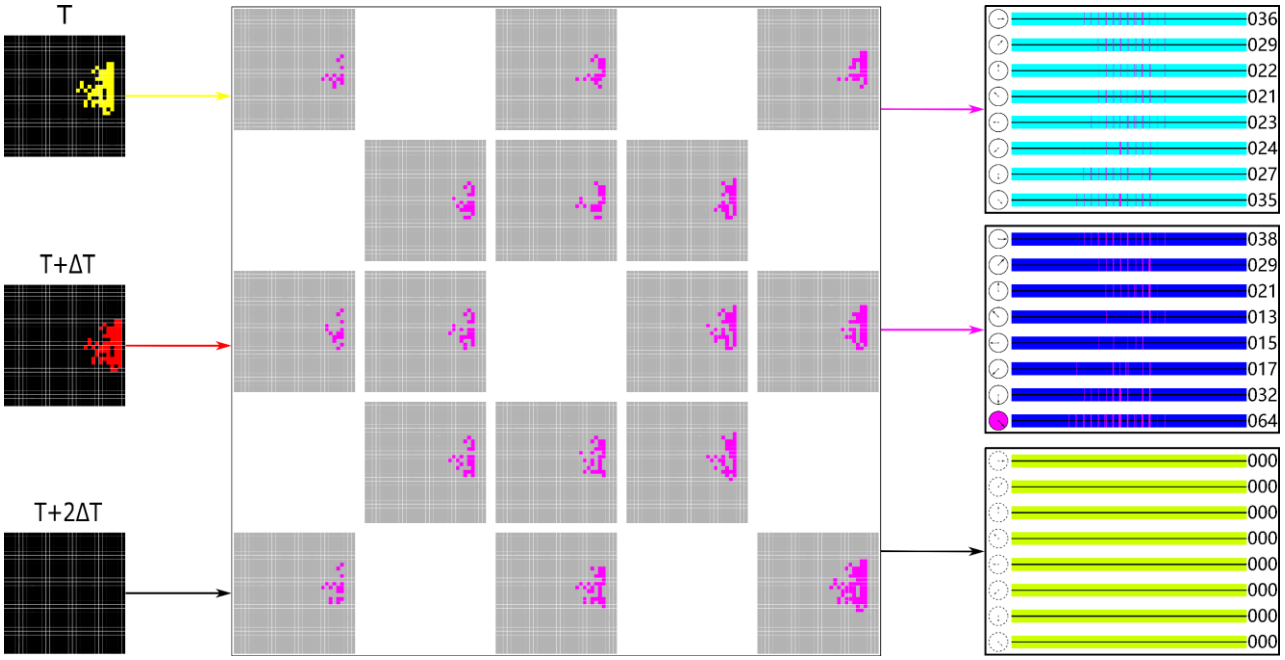

Figure 4.6 The activation plots of the AVS for 64-pixels object motion direction detection within the background with no noise (Dataset V2).

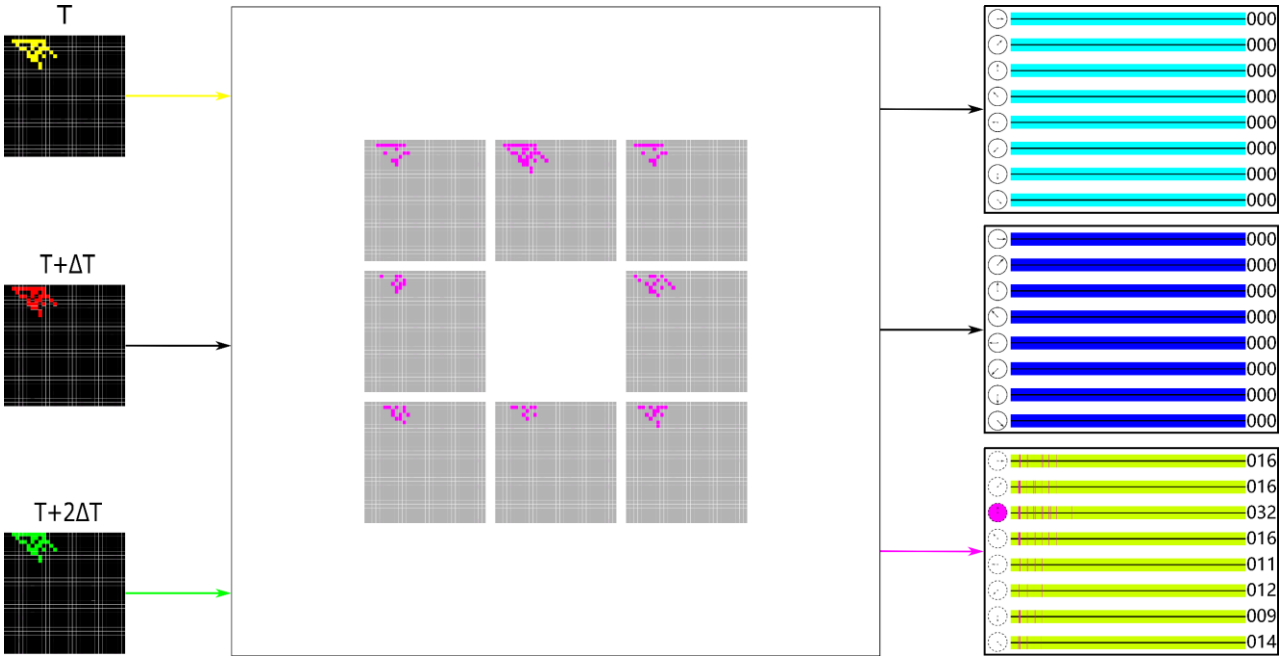

Figure 4.7 The activation plots of the AVS for 32-pixels object motion direction detection within the background with no noise (Dataset V1/2).

The activation plots of the AVS for a 32-pixels object motion direction detection within the background with no noise (Dataset V1/2) have been shown in Fig 4.7. As the detection occurs at time T+2ΔT, 8 local velocity-sensitive directionally detective neurons are applied to scan over every point in the visual field. The activation strength of each neuron is: $V1/2_R$-neuron 16, $V1/2_{UR}$-neuron 16, $V1/2_U$-neuron 32, $V1/2_{UL}$-neuron 16, $V1/2_L$-neuron 11, $V1/2_{LL}$-neuron 12, $V1/2_D$-neuron 9 and $V1/2_{LR}$-neuron 14. Since the $V1/2_U$-neuron get the maximum value of activation strength, the global direction of this 32-pixels object is Upward and the velocity of it is 1/2.

In the second series of experiments, we test our AVS within the background with 1% to 10% separated noises. As we mentioned above, we designed our local velocity-sensitive directionally detective neurons based on the core computation of the HRC model, so the effect of stationary stimulus should be considered during the global motion detection. In this research, we defined the light spot which does not have spatial movement as noise. The dataset has totally 240,000 images (each size of object has 10,000 images and each type of velocity dataset has totally 80,000 images) and the detection results are presented in Table 4.2.

The activation plots of the AVS for a 32-pixels object motion direction detection within the background with 10% separated noises (Dataset V1) have been shown in Fig 4.8. As the detection occurs at time T+ΔT, 16 local velocity-sensitive directionally detective neurons are applied to scan over every point in the visual field. The activation strength of each neuron is: $V1_R$-neuron 11, $V2_R$-neuron 19, $V1_{UR}$-neuron 12, $V2_{UR}$-neuron 17, $V1_U$-neuron 16, $V2_U$-neuron 27, $V1_{UL}$-neuron 32, $V2_{UL}$-neuron 27, $V1_L$-neuron 18, $V2_L$-neuron 23, $V1_{LL}$-neuron 17, $V2_{LL}$-neuron 17, $V1_D$-neuron 12, $V2_D$-neuron 22, $V1_{LR}$-neuron 14, and $V2_{LR}$-neuron 20. Since the $V1_{UL}$-neuron get the maximum value of activation strength, the global direction of this 32-pixels object is Upper Leftward and the velocity of it is 1.

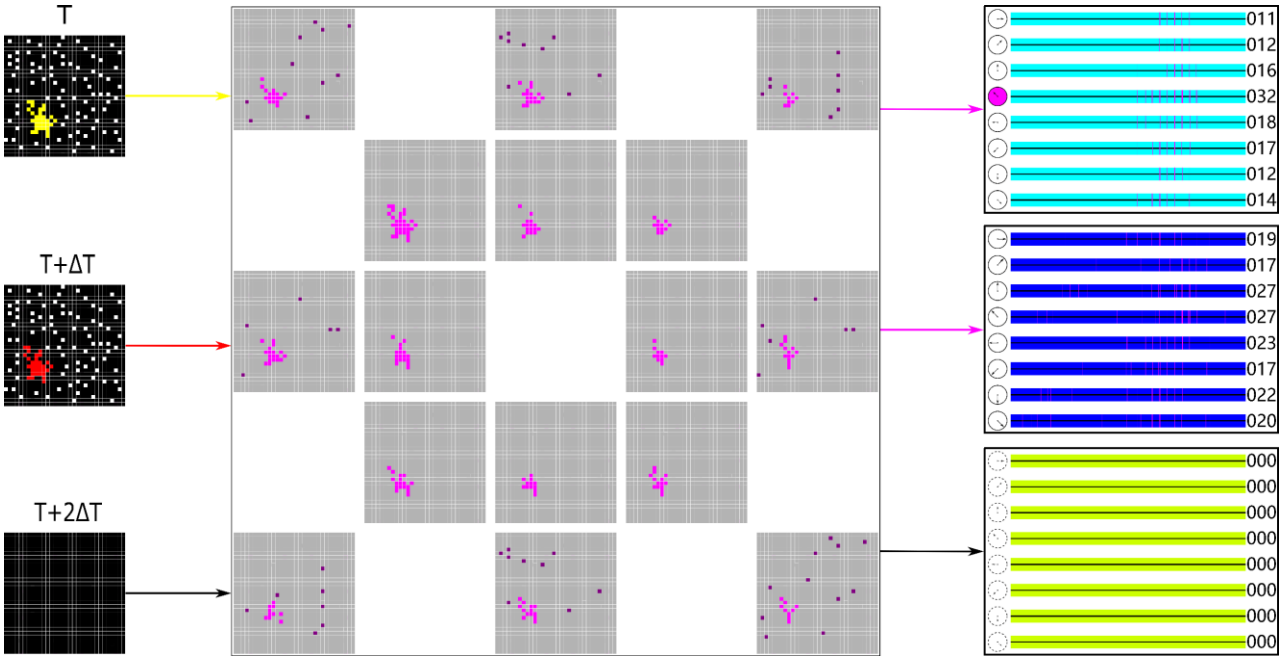

Figure 4.8 The activation plots of the AVS for 32-pixels object motion direction detection within the background with 10% separated noises (Dataset V1).

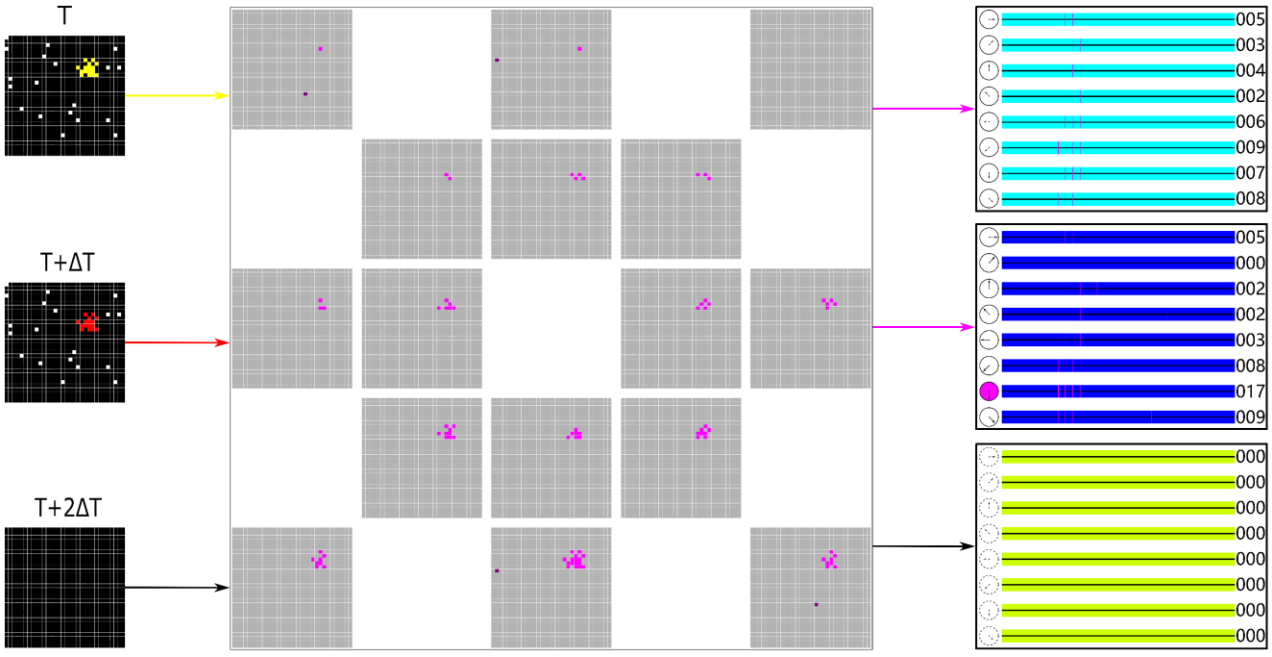

Figure 4.9 The activation plots of the AVS for 16-pixels object motion direction detection within the background with 2% separated noises (Dataset V2).

The activation plots of the AVS for a 16-pixels object motion direction detection within the background with 2% separated noises (Dataset V2) have been shown in Fig 4.9. As the detection occurs at time T+ΔT, 16 local velocity-sensitive directionally detective neurons are applied to scan over every point in the visual field. The activation strength of each neuron is: $V1_R$-neuron 5, $V2_R$-neuron 5, $V1_{UR}$-neuron 3, $V2_{UR}$-neuron 0, $V1_U$-neuron 4, $V2_U$-neuron 2, $V1_{UL}$-neuron 2, $V2_{UL}$-neuron 2, $V1_L$-neuron 6, $V2_L$-neuron 3, $V1_{LL}$-neuron 9, $V2_{LL}$-neuron 8, $V1_D$-neuron 7, $V2_D$-neuron 17, $V1_{LR}$-neuron 8, and $V2_{LR}$-neuron 9. Since the $V2_D$-neuron get the maximum value of activation strength, the global direction of this 32-pixels object is Downward and the velocity of it is 2.

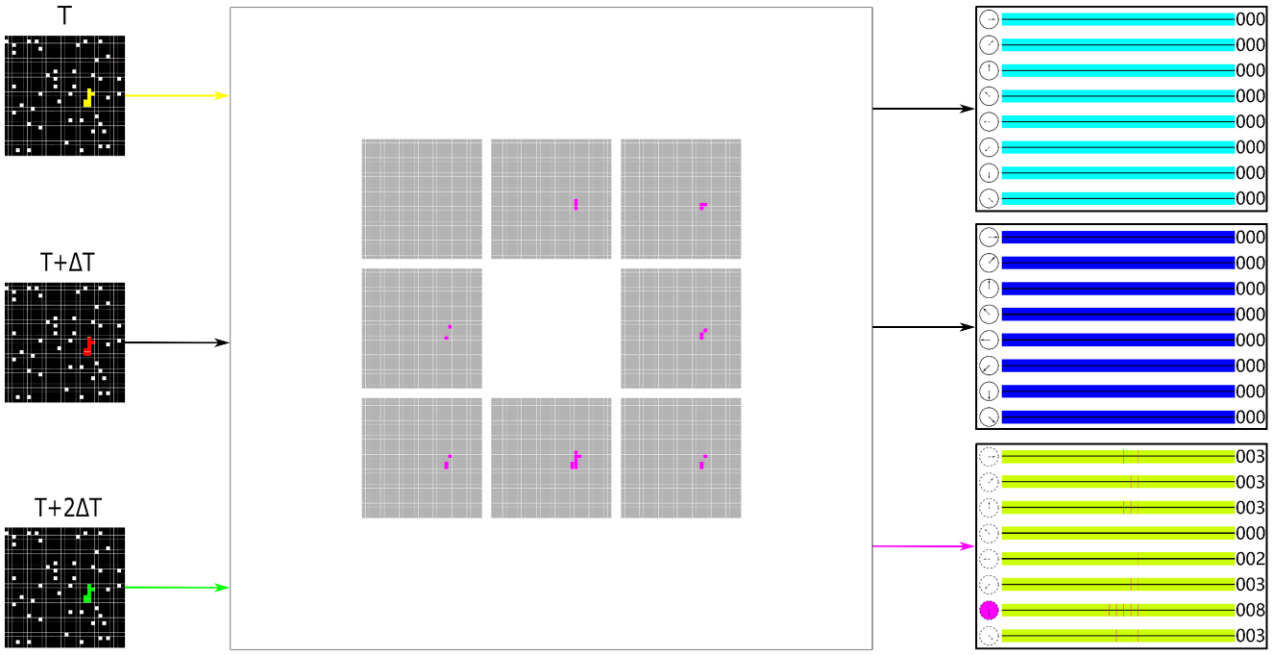

Figure 4.10 The activation plots of the AVS for 8-pixels object motion direction detection within the background with 5% separated noises (Dataset V1/2).

The activation plots of the AVS for an 8-pixels object motion direction detection within the background with no noise (Dataset V1/2) have been shown in Fig 4.10. As the detection occurs at time T+2ΔT, 8 local velocity-sensitive directionally detective neurons are applied to scan over every point in the visual field. The activation strength of each neuron is: $V1/2_R$-neuron 3, $V1/2_{UR}$-neuron

3, $V1/2_U$-neuron 3, $V1/2_{UL}$-neuron 0, $V1/2_L$-neuron 2, $V1/2_{LL}$-neuron 3, $V1/2_D$-neuron 8 and $V1/2_{LR}$-neuron 3. Since the $V1/2_D$-neuron get the maximum value of activation strength, the global direction of this 8-pixels object is Downward and the velocity of it is 1/2.

In the third series of experiments, we test our AVS within the background with 1% to 10% connected noises. The dataset has totally 240,000 images (each size of object has 10,000 images and each type of velocity dataset has totally 80,000 images) and the detection results are presented in Table 4.2.

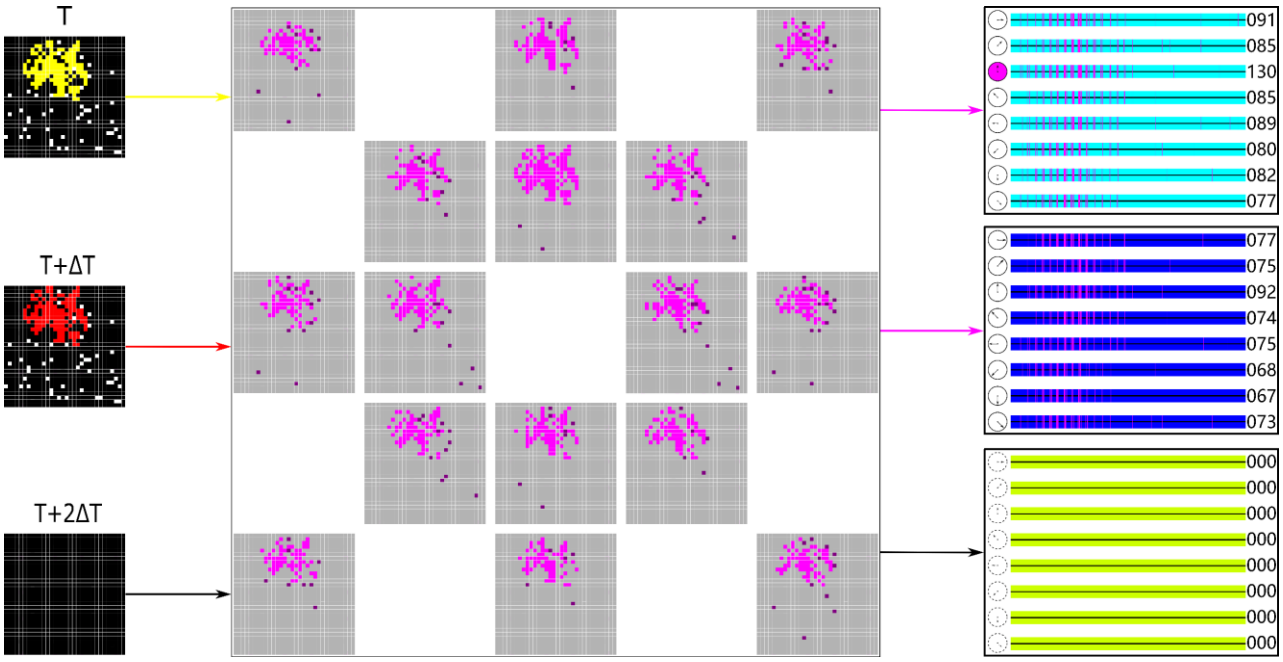

Figure 4.11 The activation plots of the AVS for 128-pixels object motion direction detection within the background with 5% connected noises (Dataset V1).

The activation plots of the AVS for a 128-pixels object motion direction detection within the background with 5% connected noises (Dataset V1) have been shown in Fig 4.11. As the detection occurs at time T+ΔT, 16 local velocity-sensitive directionally detective neurons are applied to scan over every point in the visual field. The activation strength of each neuron is: $V1_R$-neuron 91, $V2_R$-neuron 77, $V1_{UR}$-neuron 85, $V2_{UR}$-neuron 75, $V1_U$-neuron 130, $V2_U$-neuron 92, $V1_{UL}$-neuron 85,

$V2_{UL}$-neuron 74, $V1_{L}$-neuron 89, $V2_{L}$-neuron 75, $V1_{LL}$-neuron 80, $V2_{LL}$-neuron 68, $V1_{D}$-neuron 82, $V2_{D}$-neuron 67, $V1_{LR}$-neuron 77, and $V2_{LR}$-neuron 73. Since the $V1_{U}$-neuron get the maximum value of activation strength, the global direction of this 128-pixels object is Upward and the velocity of it is 1.

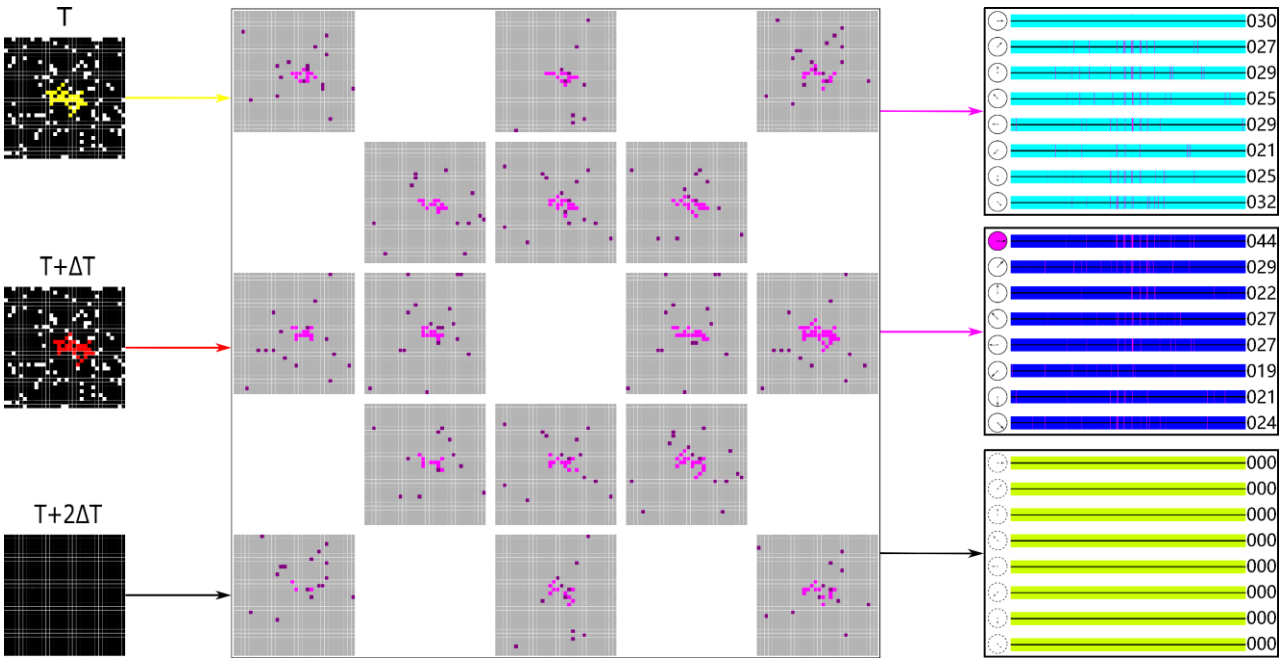

Figure 4.12 The activation plots of the AVS for 32-pixels object motion direction detection within the background with 10% connected noises (Dataset V2).

The activation plots of the AVS for a 32-pixels object motion direction detection within the background with 10% connected noises (Dataset V2) have been shown in Fig 4.12. As the detection occurs at time T+ΔT, 16 local velocity-sensitive directionally detective neurons are applied to scan over every point in the visual field. The activation strength of each neuron is: $V1_{R}$-neuron 30, $V2_{R}$-neuron 44, $V1_{UR}$-neuron 27, $V2_{UR}$-neuron 29, $V1_{U}$-neuron 29, $V2_{U}$-neuron 22, $V1_{UL}$-neuron 25, $V2_{UL}$-neuron 27, $V1_{L}$-neuron 29, $V2_{L}$-neuron 27, $V1_{LL}$-neuron 21, $V2_{LL}$-neuron 19, $V1_{D}$-neuron 25, $V2_{D}$-neuron 21, $V1_{LR}$-neuron 32, and $V2_{LR}$-neuron 24. Since the $V2_{R}$-neuron get the maximum value of activation strength, the global direction of this 32-pixels object is Rightward and the velocity

of it is 2.

The activation plots of the AVS for a 16-pixels object motion direction detection within the background with no noise (Dataset V1/2) have been shown in Fig 4.13. As the detection occurs at time T+2ΔT, 8 local velocity-sensitive directionally detective neurons are applied to scan over every point in the visual field. The activation strength of each neuron is: $V1/2_R$-neuron 16, $V1/2_{UR}$-neuron 7, $V1/2_U$-neuron 7, $V1/2_{UL}$-neuron 8, $V1/2_L$-neuron 6, $V1/2_{LL}$-neuron 3, $V1/2_D$-neuron 6 and $V1/2_{LR}$-neuron 8. Since the $V1/2_R$-neuron get the maximum value of activation strength, the global direction of this 16-pixels object is Rightward and the velocity of it is 1/2.

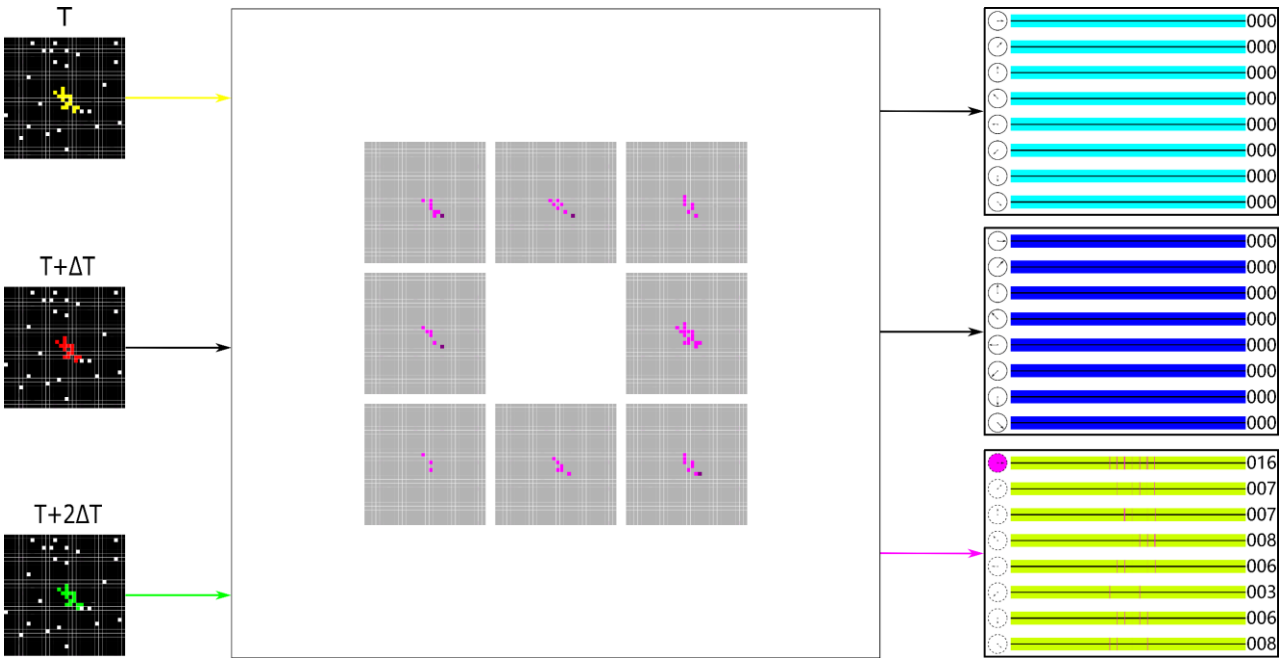

Figure 4.13 The activation plots of the AVS for 16-pixels object motion direction detection within the background with 2% connected noises (Dataset V1/2).

In the last series of experiments, we compare the performances of the AVS with a 2-channels CNN. The structure of the 2-channels CNN is shown in Figure 4.14. We hypothesize the CNN output signals at time T+ΔT, so we design the CNN contains 16 sets of convolutional filters with size of 5 × 5. By summing up the results of filters, 16 feature maps are obtained and then a 2 × 2 average

pooling is conducted before they come to Full-connect layer. Finally, the CNN obtain 16 outputs. In this research, we use the dataset with no noise in the background (Dataset V1 and Dataset V2) and randomly picked 75% of them to train the 2-channel CNN. When training accuracy reached 100%, we saved the model and test it with the rest 25% images. The detection results in Dataset V1 and Dataset V2 are both 99.99%. Moreover, we test the 2-channels CNN in detecting the Dataset V1 (separated noises and connected noises) and the Dataset V2 (separated noises and connected noises), the detection results are presented in Table 4.3.

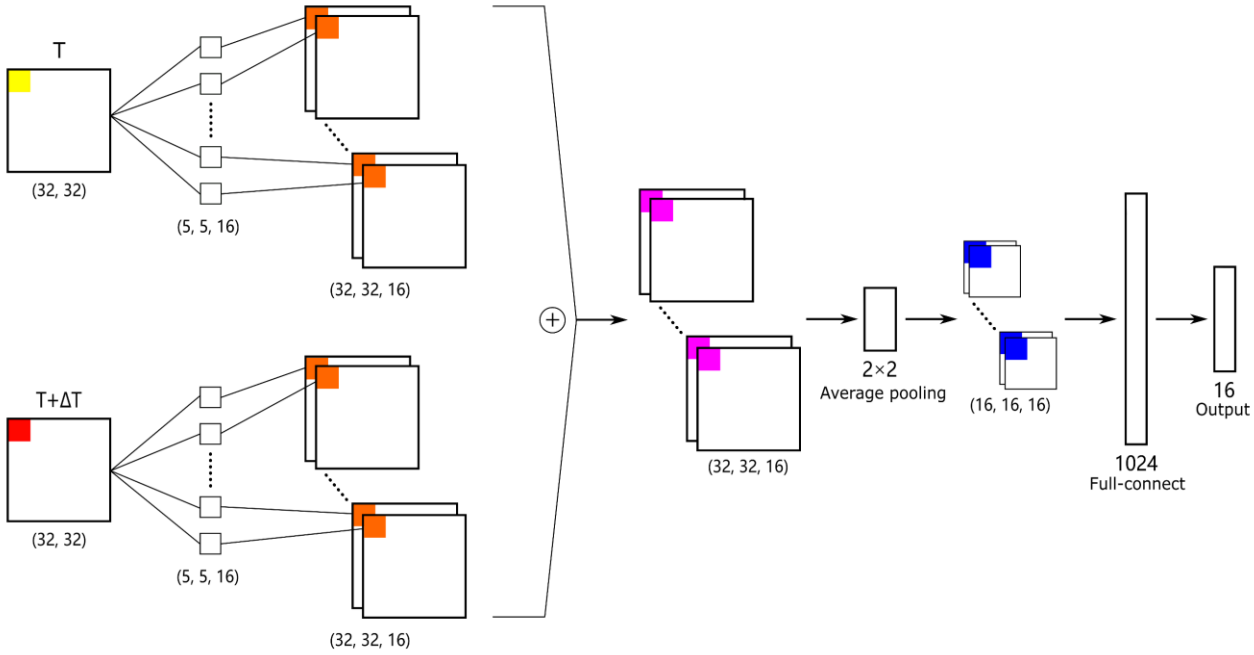

Figure 4.14 Schematic of the 2-channels CNN.

Table 4.2 Detection results of AVS within the background with separated noises and connected noises.

Background Noises	Object Sizes	Models / Noise Types / Accuracy (%)	AVS					
			Separated (Dataset V1)	Connected (Dataset V1)	Separated (Dataset V2)	Connected (Dataset V2)	Separated (Dataset V1/2)	Connected (Dataset V1/2)
1%	1		71.80	55.75	80.97	65.06	100	79.92
	2		96.84	91.71	99.06	94.95	100	95.37
	4		99.95	99.38	99.98	99.70	100	99.67
	8		100	100	100	100	100	100
	16		100	100	100	100	100	100
	32		100	100	100	100	100	100
	64		100	100	100	100	100	100
	128		100	100	100	100	100	100
2%	1		23.98	29.85	53.62	37.83	100	54.23
	2		73.03	66.64	89.20	74.69	100	81.52
	4		97.93	95.34	99.60	97.46	100	97.16
	8		99.99	99.84	100	99.96	100	99.97
	16		100	100	100	100	100	100
	32		100	100	100	100	100	100
	64		100	100	100	100	100	100
	128		100	100	100	100	100	100
5%	1		0.02	18.89	34.76	16.69	100	33.78
	2		0.84	32.31	54.64	33.50	100	49.03
	4		22.61	58.56	83.40	62.61	100	72.12
	8		83.70	89.94	98.95	93.49	100	93.29
	16		99.73	99.64	99.98	99.84	100	99.80
	32		99.99	100	100	100	100	99.90
	64		100	100	100	100	100	100
	128		100	100	100	100	100	100
10%	1		0	15.87	25.62	12.61	100	26.01
	2		0	22.61	36.12	18.72	100	34.95
	4		0	33.62	54.00	31.70	100	48.57
	8		0	57.09	78.44	58.48	100	68.17
	16		7.28	87.19	96.91	90.26	100	91.56
	32		81.17	99.29	99.96	99.73	100	99.38
	64		99.82	100	100	100	100	100
	128		100	100	100	100	100	100

Table 4.3 Detection results of the 2-channels CNN within the background with separated noises and connected noises.

Models / Noise Types / Accuracy (%) — Background Noises	Object Sizes	2-channels CNN			
		Separated (Dataset V1)	Connected (Dataset V1)	Separated (Dataset V2)	Connected (Dataset V2)
1%	1	29.91	27.40	0.02	0
	2	50.54	46.20	2.52	0.95
	4	79.60	72.98	31.53	22.20
	8	93.96	91.32	80.75	73.40
	16	98.12	97.21	96.71	94.83
	32	99.41	99.36	99.50	99.17
	64	99.89	99.84	99.88	99.81
	128	99.97	99.98	99.98	100
2%	1	20.77	20.56	0	0
	2	25.95	25.40	0	0
	4	39.25	34.90	1.27	0.36
	8	64.25	57.41	18.17	11.33
	16	84.08	78.08	61.57	49.22
	32	94.16	90.81	91.02	84.95
	64	98.74	97.36	98.33	97.05
	128	99.78	99.48	99.77	99.35
5%	1	16.05	15.78	0	0
	2	19.32	19.57	0	0
	4	22.95	22.56	0	0
	8	26.07	25.57	0.03	0
	16	37.88	34.26	3.77	0.35
	32	55.41	46.13	20.99	3.86
	64	71.34	54.42	59.20	21.43
	128	87.73	66.60	88.12	54.85
10%	1	13.75	14.40	0	0
	2	15.97	16.15	0	0
	4	19.84	18.65	0	0
	8	26.07	21.67	0	0
	16	27.60	25.84	0.03	0
	32	38.83	35.28	1.12	0
	64	49.11	46.07	11.08	0.02
	128	55.90	49.80	47.56	0.74

4.4 Summary

In this research, we validate the reliability of our AVS with a series of experiments and the high accuracy experimental results have proved that the Temporal-based multi-neurons scheme motion detection mechanism is suitable for global motion speed perception in a two-dimensional view. As we cited the biological motion detector to receive visual signal directly from the visual field, the detection results are impacted when the detecting condition is small size object and many stationary light stimuli. However, this effect will decrease as the size of object increases. Moreover, we compare the performance of our AVS with a 2-channels CNN under the same experimental conditions. The comparison results further proved that our AVS is not only capable of global motion speed perception tasks, but also has excellent performance in noise resistance.

Chapter 5 Conclusions

In this dissertation, we proposed two potential motion detection mechanisms that could be used for global motion direction detection and global motion speed perception in a two-dimensional view. We cited the core computation of the HRC model, employed the local motion-sensitive directionally detective neurons to gather the direction of local motion. With reference to the concept of simple cells, we designed our neurons with 3 × 3 local receptive field and eight neurons are employed for multi-directions detection. Moreover, we based on the characteristic of single unidirectional motion detectors, extend our motion-sensitive neurons with different sampling base and temporal delay to local velocity-sensitive directionally detective neurons for local motion speed perception. Considering the ON motion pathway is sufficient to drive the optomotor response at the high pattern contrast, we assigned value 1 to the visual signals and value 0 to the background. Through a series of experiments, we validate the reliability of the Full-neurons scheme motion detection mechanism and the Temporal-based multi-neurons scheme motion detection mechanism. The experimental results have proved our mechanisms is not only suitable for global motion detection tasks, but also has an excellent performance in noise resistance.

In this research, we only consider the excitatory inputs enhancement in the preferred direction and the simplest structure of a classical motion detector: Hassenstein-Reichardt Correlator model. Thus, the limitation of our research is the motion can only be detected in a binary dataset. However, this problem can be solved by extending the artificial visual system with more neural architectures. For example, with the application of horizontal cells, the AVS can be used for detecting the global motion in grayscale images. Furthermore, with the extension of binocular vision, the global motion can be detected in a three-dimensional view.

Visual systems have been the focus of research for past decades, however, our understanding

of it is far from complete. With the development of technology, interdisciplinary integration is increasingly valued, such as the scanning mechanism which can both be applied to the field of computer vision and the field of electron microscopy (EM). We hope our research can encourage the biologists to quantify the global motion detection from the perspective of fundamental units and contribute to the field of both neuroscience and computer vision.

References

1. Moor J. The Dartmouth College artificial intelligence conference: The next fifty years[J]. Ai Magazine, 2006, 27(4): 87-87.
2. Buchanan B G. A (very) brief history of artificial intelligence[J]. Ai Magazine, 2005, 26(4): 53-53.
3. Minar M R, Naher J. Recent advances in deep learning: An overview[J]. arXiv preprint arXiv:1807.08169, 2018.
4. DeWeerdt S. How to map the brain[J]. Nature, 2019, 571(7766): S6-S6.
5. Todo Y, Tang Z, Todo H, et al. Neurons with multiplicative interactions of nonlinear synapses[J]. International journal of neural systems, 2019, 29(08): 1950012.
6. London M, Häusser M. Dendritic computation[J]. Annu. Rev. Neurosci., 2005, 28: 503-532.
7. Medina J. Brain rules: 12 principles for surviving and thriving at work, home, and school[M]. ReadHowYouWant. com, 2011.
8. Livingstone M S. Art, illusion and the visual system[J]. Scientific American, 1988, 258(1): 78-85.
9. Nakayama K. Biological image motion processing: a review[J]. Vision research, 1985, 25(5): 625-660.
10. Exner S. Entwurf zu einer physiologischen Erklärung der psychischen Erscheinungen[M]. F. Deuticke, 1894.
11. Wertheimer M. Experimentelle studien uber das sehen von bewegung[J]. Zeitschrift fur psychologie, 1912, 61.
12. Hartline H K, Graham C H. Nerve impulses from single receptors in the eye[J]. Journal of Cellular & Comparative Physiology, 1932.

13. Hartline H K. The response of single optic nerve fibers of the vertebrate eye to illumination of the retina[J]. American Journal of Physiology-Legacy Content, 1938, 121(2): 400-415.
14. Kuffler S W. Discharge patterns and functional organization of mammalian retina[J]. Journal of neurophysiology, 1953, 16(1): 37-68.
15. Barlow H B, Hill R M. Selective sensitivity to direction of movement in ganglion cells of the rabbit retina[J]. Science, 1963, 139(3553): 412-414.
16. Dvorak D R, Bishop L G, Eckert H E. On the identification of movement detectors in the fly optic lobe[J]. Journal of comparative physiology, 1975, 100(1): 5-23.
17. Sanes J R, Zipursky S L. Design principles of insect and vertebrate visual systems[J]. Neuron, 2010, 66(1): 15-36.
18. Mauss A S, Vlasits A, Borst A, et al. Visual circuits for direction selectivity[J]. Annual review of neuroscience, 2017, 40: 211.
19. Borst A, Helmstaedter M. Common circuit design in fly and mammalian motion vision[J]. Nature neuroscience, 2015, 18(8): 1067-1076.
20. Vlasits A L, Bos R, Morrie R D, et al. Visual stimulation switches the polarity of excitatory input to starburst amacrine cells[J]. Neuron, 2014, 83(5): 1172-1184.
21. Poleg-Polsky A, Diamond J S. Retinal circuitry balances contrast tuning of excitation and inhibition to enable reliable computation of direction selectivity[J]. Journal of Neuroscience, 2016, 36(21): 5861-5876.
22. Leonhardt A, Ammer G, Meier M, et al. Asymmetry of Drosophila ON and OFF motion detectors enhances real-world velocity estimation[J]. Nature neuroscience, 2016, 19(5): 706-715.

23. Mauss A S, Meier M, Serbe E, et al. Optogenetic and pharmacologic dissection of feedforward inhibition in Drosophila motion vision[J]. Journal of Neuroscience, 2014, 34(6): 2254-2263.
24. Takemura S, Nern A, Chklovskii D B, et al. The comprehensive connectome of a neural substrate for ‘ON’motion detection in Drosophila[J]. Elife, 2017, 6: e24394.
25. Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. The Journal of physiology, 1962, 160(1): 106.
26. Fu Q, Wang H, Hu C, et al. Towards computational models and applications of insect visual systems for motion perception: A review[J]. Artificial life, 2019, 25(3): 263-311.
27. Clifford C W G, Ibbotson M R. Fundamental mechanisms of visual motion detection: models, cells and functions[J]. Progress in neurobiology, 2002, 68(6): 409-437.
28. Borst A, Egelhaaf M. Principles of visual motion detection[J]. Trends in neurosciences, 1989, 12(8): 297-306.
29. Hassenstein B, Reichardt W. Systemtheoretische analyse der zeit-, reihenfolgen-und vorzeichenauswertung bei der bewegungsperzeption des rüsselkäfers chlorophanus[J]. Zeitschrift für Naturforschung B, 1956, 11(9-10): 513-524.
30. Frye M. Elementary motion detectors[J]. Current Biology, 2015, 25(6): R215-R217.
31. Barlow H B, Levick W R. The mechanism of directionally selective units in rabbit's retina[J]. The Journal of physiology, 1965, 178(3): 477.
32. Adelson E H, Bergen J R. Spatiotemporal energy models for the perception of motion[J]. Josa a, 1985, 2(2): 284-299.
33. Yang H H, Clandinin T R. Elementary motion detection in Drosophila: algorithms and mechanisms[J]. Annual Review of Vision Science, 2018, 4: 143.

34. Borst A, Euler T. Seeing things in motion: models, circuits, and mechanisms[J]. Neuron, 2011, 71(6): 974-994.
35. Cruz-Martín A, El-Danaf R N, Osakada F, et al. A dedicated circuit links direction-selective retinal ganglion cells to the primary visual cortex[J]. Nature, 2014, 507(7492): 358-361.
36. Fukushima K, Miyake S. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition[M]//Competition and cooperation in neural nets. Springer, Berlin, Heidelberg, 1982: 267-285.
37. Borst A, Haag J, Reiff D F. Fly motion vision[J]. Annual review of neuroscience, 2010, 33: 49-70.
38. Maisak M S, Haag J, Ammer G, et al. A directional tuning map of Drosophila elementary motion detectors[J]. Nature, 2013, 500(7461): 212-216.
39. Joesch M, Schnell B, Raghu S V, et al. ON and OFF pathways in Drosophila motion vision[J]. Nature, 2010, 468(7321): 300-304.
40. Borst A, Haag J, Mauss A S. How fly neurons compute the direction of visual motion[J]. Journal of Comparative Physiology A, 2020, 206(2): 109-124.
41. Takemura S Y, Lu Z, Meinertzhagen I A. Synaptic circuits of the Drosophila optic lobe: the input terminals to the medulla[J]. Journal of Comparative Neurology, 2008, 509(5): 493-513.
42. Shinomiya K, Huang G, Lu Z, et al. Comparisons between the ON-and OFF-edge motion pathways in the Drosophila brain[J]. Elife, 2019, 8: e40025.
43. Groschner L N, Malis J G, Zuidinga B, et al. A biophysical account of multiplication by a single neuron[J]. Nature, 2022, 603(7899): 119-123.
44. Rister J, Pauls D, Schnell B, et al. Dissection of the peripheral motion channel in the visual system of Drosophila melanogaster[J]. Neuron, 2007, 56(1): 155-170.

45. Schnell B, Raghu S V, Nern A, et al. Columnar cells necessary for motion responses of wide-field visual interneurons in Drosophila[J]. Journal of Comparative Physiology A, 2012, 198(5): 389-395.

46. Takemura S, Bharioke A, Lu Z, et al. A visual motion detection circuit suggested by Drosophila connectomics[J]. Nature, 2013, 500(7461): 175-181.

47. Ammer G, Leonhardt A, Bahl A, et al. Functional specialization of neural input elements to the Drosophila ON motion detector[J]. Current Biology, 2015, 25(17): 2247-2253.

48. Strother J A, Wu S T, Wong A M, et al. The emergence of directional selectivity in the visual motion pathway of Drosophila[J]. Neuron, 2017, 94(1): 168-182. e10.

49. Haag J, Arenz A, Serbe E, et al. Complementary mechanisms create direction selectivity in the fly[J]. Elife, 2016, 5: e17421.

50. Egelhaaf M, Borst A, Reichardt W. Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system[J]. JOSA A, 1989, 6(7): 1070-1087.

51. Glünder H. Correlative velocity estimation: visual motion analysis, independent of object form, in arrays of velocity-tuned bilocal detectors[J]. JOSA A, 1990, 7(2): 255-263.

52. Zanker J M, Srinivasan M V, Egelhaaf M. Speed tuning in elementary motion detectors of the correlation type[J]. Biological cybernetics, 1999, 80(2): 109-116.

53. Bai S, Kolter J Z, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[J]. arXiv preprint arXiv:1803.01271, 2018.

54. Srinivasan M, Zhang S, Lehrer M, et al. Honeybee navigation en route to the goal: visual flight control and odometry[J]. The Journal of experimental biology, 1996, 199(1): 237-244.

55. Barlow H B, Hill R M, Levick W R. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit[J]. The Journal of physiology, 1964, 173(3): 377.
56. Movshon J A, Blakemore C. Functional reinnervation in kitten visual cortex[J]. Nature, 1974, 251(5475): 504-505.
57. Orban G A, Kennedy H, Maes H. Response to movement of neurons in areas 17 and 18 of the cat: velocity sensitivity[J]. Journal of neurophysiology, 1981, 45(6): 1043-1058.
58. Tynan P, Sekuler R. Simultaneous motion contrast: velocity, sensitivity and depth response[J]. Vision research, 1975.
59. Levinson E, Sekuler R. The independence of channels in human vision selective for direction of movement[J]. The Journal of Physiology, 1975, 250(2): 347-366.
60. Levinson E, Sekuler R. Inhibition and disinhibition of direction-specific mechanisms in human vision[J]. Nature, 1975, 254(5502): 692-694.
61. Juusola M, Dau A, Song Z, et al. Microsaccadic sampling of moving image information provides Drosophila hyperacute vision[J]. Elife, 2017, 6: e26117.
62. Schneider J, Murali N, Taylor G W, et al. Can Drosophila melanogaster tell who's who?[J]. PloS one, 2018, 13(10): e0205043.