A Motion-Direction-Detecting Model for Gray-Scale Images Based on the Hassenstein–Reichardt Model

Zhiyu Qiu 1, Yuki Todo 2,*, Chenyang Yan 1,* and Zheng Tang 2

1 Faculty of Electrical and Computer Engineering, Kanazawa University, Kakuma-Machi, Kanazawa 920-1192, Japan; qiuzy1916@stu.kanazawa-u.ac.jp
2 Department of Intelligence Information Systems, University of Toyama, 3190 Gofuku, Toyama 930-8555, Japan; ztang@eng.u-toyama.ac.jp
* Correspondence: yktodo@se.kanazawa-u.ac.jp (Y.T.); yy772013879@gmail.com (C.Y.)

Abstract: The visual system of sighted animals plays a critical role in providing information about the environment, including motion details necessary for survival. Over the past few years, numerous studies have explored the mechanism of motion direction detection in the visual system for binary images, including the Hassenstein–Reichardt model (HRC model) and the HRC-based artificial visual system (AVS). In this paper, we introduced a contrast-response system based on previous research on amacrine cells in the visual system of Drosophila and other species. We combined this system with the HRC-based AVS to construct a motion-direction-detection system for gray-scale images. Our experiments verified the effectiveness of our model in detecting the motion direction in gray-scale images, achieving at least 99% accuracy in all experiments and a remarkable 100% accuracy in several circumstances. Furthermore, we developed two convolutional neural networks (CNNs) for comparison to demonstrate the practicality of our model.

Keywords: artificial visual system; amacrine cells; mechanism; neuron; motion detection; gray-scale

1. Introduction

The ability to receive information about movement is crucial for the survival of living things. Vision is an important method of gaining motion information. The study of motion vision has been a popular topic in recent years, as it not only facilitates the development of image-processing techniques but also contributes to a deeper understanding of how the brain works. In 1956, Hassenstein and Reichardt proposed a motion-direction-detecting model, which is known as the Hassenstein–Reichardt Model (HRC model), by analyzing the steering tendencies of green leaf beetles [1]. This model has had a great influence on studies in the motion vision field [2]. Subsequently Goetz, Joesch, Schnell et al. respectively demonstrated the validity of the HRC model at different levels of the visual system of Drosophila [3–6]. Since some properties of Drosophila were quite suitable for genetic experiments, scientists conducted a series of studies on Drosophila to unravel the mechanism of visual circuits. It has been shown that the HRC model could reasonably explain some mechanisms of motion direction detection in the fly visual system [7]. In our previous work, we have proposed several models for motion direction detection and object orientation detection based on biological models and theories [8–14]. Yan et al. proposed an artificial visual system (AVS) based on the HRC model [15]. This model can be used for motion direction detection in a binary environment (where the value of each pixel is limited to 0 or 1) and this model also proposed a global motion-detection mechanism. However, this model is not applicable to motion detection in gray-scale images. In the case of motion detection in gray-scale images, Tang et al. proposed a mechanism for motion direction detection in gray-scale images based on the function of bipolar cells and horizontal cells, which called the on-off response [8,16]. However, there is no horizontal cell in the visual system of Drosophila [17]. Therefore, in order to propose a motion-direction-detecting...
model for gray-scale environments based on the visual system of Drosophila, we needed to understand the mechanism of the visual system of Drosophila and the function of each cell. It is known that the neural circuit for motion detection of Drosophila is split into two parallel motion circuits specialized to detect the motion of luminance increments and decrements separately [18–23]. A study showed that these two circuits seemed to be in line with the HRC model [24]. In recent years, Bahl et al. found that the contrast response in the Drosophila visual system proceeds in a visual pathway independent of motion detection through genetic experiments, and there was a global contrast-response mechanism that shared some elements with the motion-direction-detecting pathway [25]. Sanes et al. compared the structure of vertebrates and flies and hypothesized that amacrine cells in the fly visual system may perform a similar function to horizontal cells in the vertebrate visual system [17]. Takemura et al. discovered a kind of large amacrine cell in the Drosophila visual system and some research showed that this cell has a role in the motion-direction-detecting pathway; however, it is still not clear exactly how it acts in the motion-direction-detecting pathway [26–29].

In this paper, we hypothesized that such large amacrine cells in the Drosophila visual system perform a similar function to horizontal cells in the vertebrate visual system. Incorporating these speculations, we constructed a motion-direction-detecting system for gray-scale images based on the HRC model. Compared to previous motion-direction-detecting models [8,11,14,15], our innovative approach sets our model apart. Our model uses three frames of images as input, resulting in a substantial improvement in stability, especially for small objects and messy backgrounds. Moreover, theoretically, each pixel in the input image can take any numerical value in our model. In this paper, we have restricted the pixel values to a range of 0 to 255 for the sake of convenience.

Our initial step was constructing the core detector, which utilizes the HRC model to detect motion in a single direction. According to the HRC model, direction-selective neurons receive signals from two separate photoreceptors to detect the direction of motion. We then constructed a contrast-response system that receives input from the same photoreceptors and inhibits motion-direction-detecting neurons according to the contrast information of the input signals. Furthermore, we extended the model to two-dimensional planes in order to detect eight movement directions. In the two-dimensional model, the contrast-response system would receive input from a number of surrounding photoreceptors and output an inhibitory signal to the motion-direction-detecting neurons based on the contrast information from photoreceptors. Finally, we constructed a global motion-direction-detecting model based on the biological theories and the work of Yan [15]. We tested the model using a group of images with a time lag. The results showed that our model could detect the motion direction in gray-scale images perfectly under various situations. In addition, we used convolutional neural networks (CNNs) as a comparison and the results showed that our model had better performance than CNNs in motion direction detection. We firmly believe that our proposed model presents a promising solution for motion-detection tasks in gray-scale images.

The remaining sections of this paper are structured as follows: Section 2 introduces the HRC model and motion vision of Drosophila and how we construct the motion-direction-detecting model for gray-scale image based on these theories. Section 3 present the experimental results of our model, comparing the practicality with the convolutional neural networks (CNNs). Section 4 makes conclusions for our study.

2. Methods

In recent years, the exact nature of the motion-direction-detecting process has been extensively studied in the visual system of Drosophila. In addition, significant strides have been made in determining the neural circuits that generate directional motion information [23]. In this section, we discuss how we construct the motion-direction-detecting model for gray-scale image based on the HRC model and motion vision of Drosophila.
2.1. Hassenstein–Reichardt Correlator Model

The Hassenstein–Reichardt correlator (HRC) model was proposed by Hassenstein and Reichardt after analyzing the steering tendencies of green leaf beetles [1]. This model has great value in the field of motion vision [2]. The HRC model consists of a set of mirror circuits (Figure 1A). Borst et al. demonstrated that each subunit could detect motion direction independently [30]. Studies have shown that the HRC model can explain the mechanism of motion detection in Drosophila [23,25–30]. In the study of motion-direction-detecting neurons, the direction to which each neuron responds most strongly was defined as the preferred direction (PD) [15]. The opposite direction, in which the neuron responds less strongly or not at all, was known as the null direction (ND). This distinction is important for understanding how the motion-direction-detecting neuron responds to specific motion directions. As an example, we will focus our discussion on one of the subunits in the HRC model (Figure 1B). There are two inputs in this subunit and one of the branches exhibits a delay of $\Delta t$. The signals from these two inputs are combined in a multiplier, which produces the final output signal. As the light point moves from left to right, suppose it passes photoreceptor $A$ at time $T$. This signal will be transmitted to the multiplier after a delay of $\Delta t$. At time $T + \Delta t$, the light spot passes the adjacent photoreceptor $B$, and the signal from photoreceptor $B$ is combined with the delayed signal from photoreceptor $A$ in the multiplier. Finally, the multiplier will export the signal, which means that the neuron will respond. However, if the light moves from right to left, the signal from photoreceptor $A$ will only reach the multiplier after the signal from photoreceptor $B$ due to the delay circuit. As a result, the neuron will not respond to the light stimulus when it moves in this direction. The function of a subunit in the HRC model Figure 1B can be defined as:

$$X = A(T)B(T + \Delta t)$$  \hspace{1cm} (1)

$X$ is the output, and $A(t)$ and $B(T + \Delta t)$ are two inputs of the HRC model with a fixed delay.

2.2. Local Motion-Direction-Detecting Neuron for Gray-Scale Images

In the fly’s visual system, visual processing starts with the optic lobe, where photoreceptors are present to receive external light stimuli. The optic lobe sends the signal to the laminae via two pathways that detect light increments (on) and decrements (off), respectively. These two pathways then respectively connect to the lobule and lobule plate and finally converge on the lobule plate. Studies have revealed that T4 and T5 cells in the lobules and lobule plates are the first cells in the fly visual system that show directional selectivity [18–23]. T4 is located in the L1 pathway and responds to the on edge,
while T5 is located in the L2 pathway and responds to the off edge. Both types of cells have four subtypes and each subtype forms a neural circuit that is specialized to detect a specific direction of movement. Previous research has provided ample evidence that the HRC model can reasonably explain the mechanism behind motion direction detection in both pathways [17,24]. Therefore, we further speculate that it might be possible to merge these two pathways into a single HRC model mathematically that can respond to both on and off signals. Bahl et al. made a discovery that there is a contrast-response mechanism in the visual system of Drosophila. Studies have shown that it might be a global contrast-response mechanism [25]. Takemura et al. discovered a large amacrine cell [26] that receives input from a wide range of L1 and L2 pathways [31]. This cell acts as an inhibitor of direction-selective neurons, T4 and T5 cells [23,27–29]. This function is quite similar to that of horizontal cells in the vertebrate visual system [16]. It has been suggested that amacrine cells in the fly visual system may serve a similar function to horizontal cells in vertebrates [17]. In addition, this kind of amacrine cell has been reported to have an important function in the motion-detecting system [28]. A recent study has shown that they released the inhibitory neurotransmitter GABA when there was a change in the local contrast [29]. Based on the findings from these prior studies, we hypothesized that amacrine cells could have an inhibitory effect on T4 and T5 cells in response to changes in contrast. Based on these prior studies, we constructed a local motion-direction-detecting system that is applicable to gray-scale images using the HRC model. First, we constructed local motion-direction-detecting neurons that can identify a single moving direction. These neurons respond to changes in luminance and detect the moving direction based on the mechanism of the HRC model. Mathematically, we utilized integers from 0 to 255 to represent the various strengths of the light signals received by the photoreceptors. We assumed that the photoreceptor would receive the information of luminance; however, two branches under the photoreceptors would transmit the signal only when the luminance changed. Additionally, we assumed that the luminance intensity of an object remains constant as it moves. Here, we consider the neuron that detects rightward motion first (Figure 2A). The photoreceptor located in the center of the receptive field was defined as $A(x, y)$, and the photoreceptor located to its right was defined as $B(x + 1, y)$. At time $T - \Delta t$, the object was located on the left of photoreceptor $A$, and the luminance of the signals on both photoreceptors was 0. At moment $T$, the object passed by photoreceptor $A$ and it received the light signal $S_1(x, y, T)$. Since the luminance of the signal on photoreceptor $A$ changed, the left branch of the HRC model transmitted a signal $X_A$ downstream. At moment $T + \Delta t$, the object passed by photoreceptor $B$ and it received the light signal $S_2(x + 1, y, T + \Delta t)$. As the luminance of the signal on photoreceptor $B$ also changed, the right branch of the HRC model transmitted a signal $X_B$ downstream. The function of $X_A$ and $X_B$ could be formulated as:

$$X_A(x, y) = \begin{cases} 1 & |S(A, T) - S(A, T - \Delta t)| \geq \epsilon \\ 0 & |S(A, T) - S(A, T - \Delta t)| < \epsilon \end{cases}$$

(2)

$$X_B(x + 1, y) = \begin{cases} 1 & |S(B, T + \Delta t) - S(B, T)| \geq \epsilon \\ 0 & |S(B, T + \Delta t) - S(B, T)| < \epsilon \end{cases}$$

(3)

where $\epsilon$ is a positive number approaching 0. According to the mechanism of the HRC model, the activation function of a local direction-detecting neuron that detects rightward movement can be represented as:

$$Q'_{x,y}(x, y) = X_A(x, y)X_B(x + 1, y)$$

(4)

$Q'_{x,y}$ is the output of the motion-direction-detecting neuron on $(x, y)$ detecting rightward movement. We can observe that the local direction-detecting neuron that detects rightward movement is activated only when the luminance of light on photoreceptor $A$ changed at moment $T$ and the luminance of light on photoreceptor $B$ changed at moment $T + \Delta t$. 

$X_A(x, y)$ and $X_B(x + 1, y)$ represent the light signal on photoreceptor $A$ and $B$ at moment $T$. The change in luminance intensity is given by $\Delta X = X_T - X_{T-\Delta t}$, where $X_T$ is the luminance intensity at moment $T$ and $X_{T-\Delta t}$ is the luminance intensity at moment $T - \Delta t$. The output of the motion-direction-detecting neuron is then calculated as:

$$Q'_{x,y}(x, y) = X_A(x, y)X_B(x + 1, y)$$

(4)
Subsequently, we extended our one-direction-detecting model to detect motion in 8 directions on a two-dimensional plane by constructing a $3 \times 3$ local receptive field. Based on the concept of local receptive fields, we used 8 local motion-direction-detecting neurons to detect motion in each of the 8 directions. Figure 2B shows the scheme of the motion-direction-detecting model for 8 directions. Each of these neurons was responsible for detecting motion in one of 8 directions—upper left (UL), upper (U), upper right (UR), left (L), right (R), lower left (LL), lower (Lo) and lower right (LR). The activation functions for all kinds of neuron except for the one detecting rightward motion are shown below:

$$Q'_{UL}(x, y) = X_A(x, y)X_B(x - 1, y + 1)$$

$$Q'_U(x, y) = X_A(x, y)X_B(x, y + 1)$$

$$Q'_{UR}(x, y) = X_A(x, y)X_B(x + 1, y + 1)$$

$$Q'_L(x, y) = X_A(x, y)X_B(x - 1, y)$$

$$Q'_{LL}(x, y) = X_A(x, y)X_B(x - 1, y - 1)$$

$$Q'_Lo(x, y) = X_A(x, y)X_B(x, y - 1)$$

$$Q'_{LR}(x, y) = X_A(x, y)X_B(x + 1, y - 1)$$

Then we constructed a contrast-response neuron according to the function of an amacrine cell. This neuron receives all light signals within the receptive field of each local motion-direction-detecting neuron. It then compares the light intensity received by the photoreceptors and outputs an inhibitory signal to the specific local motion-direction-detecting neuron based on the result of the comparison. Essentially, the contrast-response neuron helps to refine the detection of motion direction by inhibiting signals that are not relevant to the moving object.

Figure 3 shows the structure of this neuron. Specifically, we will discuss the detailed mechanism of this neuron using the example of detecting rightward movement. Photoreceptor $A(x, y)$ receives the light signal $S_1(x, y, T)$ at time $T$, while photoreceptor $B(x + 1, y)$ receives the light signal $S_2(x + 1, y, T + \Delta t)$ at time $T + \Delta t$. If the absolute difference in intensity between $S_1$ and $S_2$ exceeds the threshold $\alpha$ (a positive number close to 0), the
neuron will export an inhibitory signal (0) to the motion-direction-detecting neuron that detects rightward movement. If the absolute difference in intensity is smaller than the threshold $\alpha$, the neuron will export an activating signal (1) to the motion-direction-detecting neuron. The activation function for this process can be expressed by the following equation:

$$Z_R = \begin{cases} 1 & |S_1(x, y, T) - S_2(x + 1, y, T + \Delta t)| \leq \alpha \\ 0 & |S_1(x, y, T) - S_2(x + 1, y, T + \Delta t)| > \alpha \end{cases}$$ (12)

![Figure 3. Structure of contrast-response system.](image)

The activation function of our system for detecting rightward motion after the addition of the contrast-response neuron can be expressed by the following equation:

$$Q_R(x, y) = Q'_R(x, y) \cdot Z_R$$ (13)

It is obvious that the system will only be activated when the strength of the light signals is the same before and after moving after adding the contrast-response neuron. By incorporating the contrast-response neuron, our system is able to detect rightward motion with greater precision and accuracy. The activation function of the contrast-response neuron for detecting the remaining 7 directions of motion can be expressed by the following equations:

$$Z_{UL} = \begin{cases} 1 & |S_1(x, y, T) - S_2(x - 1, y + 1, T + \Delta t)| \leq \alpha \\ 0 & |S_1(x, y, T) - S_2(x - 1, y + 1, T + \Delta t)| > \alpha \end{cases}$$ (14)

$$Z_U = \begin{cases} 1 & |S_1(x, y, T) - S_2(x, y + 1, T + \Delta t)| \leq \alpha \\ 0 & |S_1(x, y, T) - S_2(x, y + 1, T + \Delta t)| > \alpha \end{cases}$$ (15)

$$Z_{UR} = \begin{cases} 1 & |S_1(x, y, T) - S_2(x + 1, y + 1, T + \Delta t)| \leq \alpha \\ 0 & |S_1(x, y, T) - S_2(x + 1, y + 1, T + \Delta t)| > \alpha \end{cases}$$ (16)

$$Z_{l} = \begin{cases} 1 & |S_1(x, y, T) - S_2(x - 1, y, T + \Delta t)| \leq \alpha \\ 0 & |S_1(x, y, T) - S_2(x - 1, y, T + \Delta t)| > \alpha \end{cases}$$ (17)
The activation function of our model to detect the remaining 7 directions after adding the contrast-response neurons can be expressed by the following equation:

$$Q_d(x,y) = Q'_d(x,y) \cdot Z_d$$

(21)

‘d’ indicates one of the eight directions (UL, U, UR, R, L, LL, Lo, LR).

Now, a local motion-direction-detecting system for grey-scale images is established. The structure of the model for detecting rightward movement is shown in Figure 4 as an example.

2.3. Global Motion-Direction-Detecting System

Studies have shown that a single dendritic arbor of each T4 and T5 cell in the Drosophila visual system can sample from different locations in the visual field. Additionally, these cells send signals to the lobule plate tangential cells, which sum these signals to produce a wide-field motion response [27]. Inspired by this theory and biophysical studies in Drosophila, we constructed a global motion-direction-detecting system. The basic idea behind our model is that each light spot in the visual field can be received by different local motion-direction-
detecting neurons. Specifically, we assume that the signals received by each photoreceptor transmits to different local motion-direction-detecting neurons. Each photoreceptor is connected to a local motion-direction-detecting neuron that detects 8 directions. At the same time, these light signals will also be transmitted to the contrast-response neurons. Each local motion-direction-detecting neuron will be activated when the strength of signals on two photoreceptors changes at time $T$ and $T + \Delta T$ and the contrast-response neuron did not detect light intensity changes exceeding the threshold. The amount of activated neurons with the same preferred direction is then summed up to determine the activation strength in that direction. The global motion-detecting neuron will give a detection result based on the maximum value of the activation strength of the eight directions. Therefore, the final output can be expressed by the following equation:

$$r_d = \sum Q_d$$  \hspace{1cm} (22)$$

$$R = \max(r_d)$$  \hspace{1cm} (23)$$

$Q_d$ is the output of a single neuron in a specific direction, ‘$d$’ indicates one of the eight directions (UL, U, UR, R, L, LL, Lo, LR) and $r_d$ is the sum of the outputs of all neurons in a given direction, which is the activation strength in that direction. A flowchart of the global motion-direction-detecting process is shown in Figure 5.

![Figure 5](image_url)

Figure 5. The flowchart of global motion-direction-detecting process. $Q_d$ represents the output of each local motion-direction-detecting neuron. The global motion-direction-detecting neuron counts the amount of $Q_d$ for each direction and then judges the global movement direction by determining the maximum value of the sum of $Q_d$ in each direction, which is represented as $r_d$.

The overall structure of our model is shown in Figure 6. The local motion-detecting neurons (as discussed in Section 2.2) gather movement information from each pixel with the assistance of amacrine cells. Then the global motion-detecting neuron uses this information to determine the global motion direction, as explained earlier.
3. Results

To demonstrate how our model works, we will begin with a simple example. We consider a $5 \times 5$ region with a photoreceptor under each pixel. Each pixel is connected to a contrast-response neuron and a corresponding motion-direction-detecting neuron. Suppose there is a 4-pixel object that has moved to the lower right. According to the previously discussed theory, the activation of local motion-detecting neurons occurs when an object moves by a single pixel between two consecutive images. Therefore, we have set the velocity of the object to 1 pixel per $\Delta t$ in our experiments. Figure 7A shows images of the object at time $T - \Delta t$, $T$ and $T + \Delta t$. The number in each pixel represents the intensity of light received by the photoreceptor (0–255). For clarity, we have used colored pixels to indicate the object.

In order to determine the direction of motion, we need to count the output of all kinds of motion-direction-detecting neurons. Theoretically, all motion-direction-detecting neurons produce their outputs at the same time in our model; however, since it is difficult to count the output of every neuron simultaneously, we check each pixel one by one. In our model, there is a set of opposite motion-direction-detecting neurons between every two neighboring pixels, as shown in Figure 7B. Based on the detecting mechanism of the HRC.
model, the direction-selective neuron responds when the light intensity on a photoreceptor changes at moment $T$ and the intensity on the neighboring photoreceptor changes at moment $T + \Delta t$. Therefore, first, we need to identify the regions where the light intensity changed at moment $T$ and $T + \Delta t$. Then we check each pixel in that region in the image of time $T$. At the same position in the image of time $T + \Delta t$, we check all the neighboring pixels where the light intensity changed. According to the function of the contrast-response neuron and the motion-direction-detecting neuron, the motion-direction-detecting neuron in one direction responds when the light intensity of a pixel at moment $T$ and the intensity of a surrounding pixel at moment $T + \Delta t$ were equal and both changed. The detecting process is shown in Figure 8. In addition, we use spike plots to represent the activation of the 8 classes of motion-direction-detecting neurons, as shown in Figure 9. The horizontal axis represents the position and the vertical line represents an activation at that position. The direction with the most activated neurons indicates the moving direction of the object. We specified rightward as $0^\circ$ and increased the angle counterclockwise. In this example, we determined that the object moved lower rightward.

![Figure 8](image1)

**Figure 8.** (A) The regions where the light intensity changed at moment $T$ and $T + \Delta t$. (B) Schematic diagram of motion-direction-detecting process.

![Figure 9](image2)

**Figure 9.** Spike plot of motion-direction-detecting result discussed above.

We designed and executed experiments to test our model’s performance, generating eight datasets with different combinations of object types and backgrounds. These included a constant object with a black background, a constant object with a constant background, a constant object with a random background, a random object with a black background, a random object with a constant background, a random object with a random background, a black object with a constant background and a black object with a random background. Each dataset had object sizes of 1, 2, 4, 8, 16, 32, 64 and 128 pixels. We collected 10,000 sets of data for each dataset, with each set including images of time $T - \Delta t$, $T$ and $T + \Delta t$ and a
label. The image size was $32 \times 32$ pixels. Table 1 shows the test results for each dataset. Our model achieved a remarkable accuracy rate of 100% across all datasets, regardless of the object and background brightness and patterns. However, for the random-background cases, the accuracy did not reach 100%, but remained at least 99%. This was likely due to some of background pixels having the same light intensity as the object pixel, which resulted in the motion-detecting neurons not being activated since there was no change in light intensity when the object passed that pixel. This phenomenon would affect the total amount of activated neurons and lead to a reduction in accuracy. However, the influence of this phenomenon would be reduced when the object became larger due to the increment of the total amount of activated neurons.

Table 1. Test results of our model on 8 kinds of datasets.

<table>
<thead>
<tr>
<th>Size</th>
<th>Type</th>
<th>C0</th>
<th>CC</th>
<th>CR</th>
<th>R0</th>
<th>RC</th>
<th>RR</th>
<th>0C</th>
<th>0R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>99.38%</td>
<td>100%</td>
<td>100%</td>
<td>99.26%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>99.5%</td>
<td>100%</td>
<td>100%</td>
<td>99.99%</td>
<td>100%</td>
<td>99.82%</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>99.8%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>99.95%</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>64</td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>128</td>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

C0: constant object with black background; CC: constant object with constant background; CR: constant object with random background; R0: random object with black background; RC: random object with constant background; RR: random object with random background; 0C: black object with constant background; 0R: black object with random background.

We conducted a detailed analysis of how our model detects various types of images using spike graphs. Specifically, we focused on three cases: a random object with a black background, a black object with a constant background and a constant object with a random background. As shown in Figures 10–12, spike plots demonstrate that our model works properly in all cases, regardless of the type of object and background. Notably, even in the case of a random background, which can be regarded as 100% statistical background noise, our model showed high accuracy. This indicates the robustness of our model in detecting motion signals in complex environments.

![Figure 10. (A) Sample data of random object with black background case, object moving upward (90°). The images of time $T - \Delta t$, $T$ and $T + \Delta t$ are from left to right. The color of pixel represents the brightness, as shown in the bar on the right. (B) Spike plot of motion-direction-detecting result.](image-url)
Figure 11. (A) Sample data of random object with black background case, object moving leftward (180°). The images of time $T - \Delta t$, $T$ and $T + \Delta t$ are from left to right. (B) Spike plot of motion-direction-detecting result.

Figure 12. (A) Sample data of random object with black background case, object moving lower rightward (315°). The images of time $T - \Delta t$, $T$ and $T + \Delta t$ are from left to right. Since it is difficult to accurately determine the boundaries of the object under these circumstances, we have marked the approximate position of the object with a square marker. (B) Spike plot of motion-direction-detecting result.

To further validate the feasibility of our model, we conducted a comparative study using two convolutional neural networks (CNNs). We designed CNN1 to have a similar structure to our model, with a convolutional layer for detecting the local motion direction and a fully connected layer for detecting the global motion direction. In contrast, CNN2 had a more general architecture with four convolutional layers. The structures of the two kinds of CNN are shown in Figure 13. We randomly selected 2500 sets of data from each dataset as the test set. As for the remaining 7500 sets of data, we mixed data with same the object and background type and made a dataset containing 60,000 sets of data as the training set. We used the Adam optimizer and trained each convolutional neural network (CNN) for 30 epochs with a batch size of 100. For the 4-layer CNN, we set the Maxpooling strides to (1,1). We carried out 10 trials, each starting with randomly selected data, and averaged
the results to obtain the final accuracy. The final results of the two CNNs are presented in Tables 2 and 3. Although the CNNs had high accuracy, they faced challenges in achieving 100% accuracy and did not perform well in a random background environment. In contrast, our model achieved a higher accuracy than the CNNs and had a simpler structure.

Figure 13. (A) Structure of 1-layer CNN. (B) Structure of 4-layer CNN.

Table 2. Test results of 1-layer CNN on 8 kinds of datasets.

<table>
<thead>
<tr>
<th>Size</th>
<th>C0</th>
<th>CC</th>
<th>CR</th>
<th>R0</th>
<th>RC</th>
<th>RR</th>
<th>0C</th>
<th>0R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>98.57%</td>
<td>59.33%</td>
<td>12.95%</td>
<td>98.94%</td>
<td>66.52%</td>
<td>12.38%</td>
<td>96.58%</td>
<td>21.32%</td>
</tr>
<tr>
<td>2</td>
<td>99.29%</td>
<td>51.04%</td>
<td>12.64%</td>
<td>99.47%</td>
<td>69.2%</td>
<td>12.19%</td>
<td>96.5%</td>
<td>38.49%</td>
</tr>
<tr>
<td>4</td>
<td>99.53%</td>
<td>41.61%</td>
<td>13.06%</td>
<td>99.46%</td>
<td>72.32%</td>
<td>12.64%</td>
<td>88.23%</td>
<td>36.96%</td>
</tr>
<tr>
<td>8</td>
<td>99.72%</td>
<td>67.94%</td>
<td>13.93%</td>
<td>99.64%</td>
<td>89.45%</td>
<td>21.05%</td>
<td>89.66%</td>
<td>26.73%</td>
</tr>
<tr>
<td>16</td>
<td>99.93%</td>
<td>74.12%</td>
<td>14.23%</td>
<td>99.72%</td>
<td>92.54%</td>
<td>12.08%</td>
<td>90.58%</td>
<td>36.07%</td>
</tr>
<tr>
<td>32</td>
<td>99.95%</td>
<td>80.24%</td>
<td>14.3%</td>
<td>99.99%</td>
<td>98.11%</td>
<td>51.58%</td>
<td>99.62%</td>
<td>29.01%</td>
</tr>
<tr>
<td>64</td>
<td>99.99%</td>
<td>80.11%</td>
<td>13.56%</td>
<td>100%</td>
<td>99.29%</td>
<td>28.57%</td>
<td>91.12%</td>
<td>38.43%</td>
</tr>
<tr>
<td>128</td>
<td>100%</td>
<td>69.95%</td>
<td>17.24%</td>
<td>100%</td>
<td>99.38%</td>
<td>19.32%</td>
<td>91.16%</td>
<td>47.26%</td>
</tr>
</tbody>
</table>
C0: constant object with black background; CC: constant object with constant background; CR: constant object with random background; R0: random object with black background; RC: random object with constant background; RR: random object with random background; 0C: black object with constant background; 0R: black object with random background.

Table 3. Test results of 4-layer CNN on 8 kinds of datasets.

<table>
<thead>
<tr>
<th>Size</th>
<th>C0</th>
<th>CC</th>
<th>CR</th>
<th>R0</th>
<th>RC</th>
<th>RR</th>
<th>0C</th>
<th>0R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>99.98%</td>
<td>96.1%</td>
<td>60.08%</td>
<td>99.87%</td>
<td>88.36%</td>
<td>43.7%</td>
<td>99.37%</td>
<td>92.41%</td>
</tr>
<tr>
<td>2</td>
<td>99.93%</td>
<td>96.69%</td>
<td>47.58%</td>
<td>99.98%</td>
<td>98.11%</td>
<td>51.58%</td>
<td>99.62%</td>
<td>96.96%</td>
</tr>
<tr>
<td>4</td>
<td>99.84%</td>
<td>96.99%</td>
<td>74.69%</td>
<td>99.98%</td>
<td>99.02%</td>
<td>45.37%</td>
<td>99.33%</td>
<td>97.03%</td>
</tr>
<tr>
<td>8</td>
<td>99.99%</td>
<td>95.55%</td>
<td>21.71%</td>
<td>99.95%</td>
<td>99.51%</td>
<td>70.99%</td>
<td>99.47%</td>
<td>98.04%</td>
</tr>
<tr>
<td>16</td>
<td>99.93%</td>
<td>93.59%</td>
<td>68.68%</td>
<td>99.91%</td>
<td>99.53%</td>
<td>65.18%</td>
<td>99.86%</td>
<td>97.96%</td>
</tr>
<tr>
<td>32</td>
<td>99.88%</td>
<td>95.4%</td>
<td>68.98%</td>
<td>99.89%</td>
<td>99.66%</td>
<td>58.05%</td>
<td>99.54%</td>
<td>98.46%</td>
</tr>
<tr>
<td>64</td>
<td>99.86%</td>
<td>96.38%</td>
<td>69.44%</td>
<td>99.87%</td>
<td>99.71%</td>
<td>49.72%</td>
<td>99.87%</td>
<td>99.18%</td>
</tr>
<tr>
<td>128</td>
<td>99.88%</td>
<td>92.22%</td>
<td>67.39%</td>
<td>99.61%</td>
<td>99.71%</td>
<td>50.46%</td>
<td>99.76%</td>
<td>98.75%</td>
</tr>
</tbody>
</table>
C0: constant object with black background; CC: constant object with constant background; CR: constant object with random background; R0: random object with black background; RC: random object with constant background; RR: random object with random background; 0C: black object with constant background; 0R: black object with random background.

4. Conclusions

In this paper, we present a motion-direction-detecting model for gray-scale images that is inspired by the Drosophila visual system and the HRC model. Specifically, we used the HRC model to construct a basic structure for a unidirectional motion-direction-detecting neuron. Based on existing biological theories, we hypothesized the role of amacrine cells in motion direction detection and integrated it with the HRC model to develop a motion-direction-detecting model that can be applied to gray-scale images. Our test results
demonstrated the robustness of our model to a variety of object and background scenarios and provided evidence of its potential for practical applications. The accuracy achieved by our model suggests that it could be used in a wide range of motion-detection tasks, including but not limited to object tracking and recognition. Moreover, the experimental results are consistent with existing biological theories to a certain extent. To further verify the feasibility of our model, we compared it with two types of CNNs through experiments. The results indicated that our model not only has a simpler structure but also performs better than the CNNs in terms of accuracy and noise immunity. Overall, our proposed motion-direction-detection system offers a promising solution for motion-detection tasks in grayscale images. We hope that our model can make a small contribution to the study of the motion vision of Drosophila and to the field of machine vision and image processing.

Author Contributions: Conceptualization, Z.T.; Methodology, Z.Q.; Software, Z.Q.; Validation, C.Y. and Z.T.; Formal analysis, Z.Q.; Resources, C.Y.; Writing—original draft, Z.Q.; Writing—review & editing, Y.T. and C.Y.; Supervision, Y.T., C.Y. and Z.T.; Project administration, Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by JSPS KAKENHI Grant No. 23K11261.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

9. Han, M.; Todo, Y.; Tang, Z. Mechanism of Motion Direction Detection Based on Barlow’s Retina Inhibitory Scheme in Direction-Selective Ganglion Cells. Electronics 2021, 10, 1663. [CrossRef]
15. Yan, C.; Todo, Y.; Kobayashi, Y.; Tang, Z.; Li, B. An Artificial Visual System for Motion Direction Detection Based on the Hassenstein–Reichardt Correlator Model. Electronics 2022, 11, 1423. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.