LGMD and DSNs Neural Networks Integration for Collision Predication

Guopeng Zhang School of Computer Science University of Lincoln Lincoln, UK LN6-7TS gzhang@lincoln.ac.uk Chun Zhang
Institute of Microelectronics
Tsinghua University
Beijing, China 100-084
zhangchun@tsinghua.edu.cn

Shigang Yue School of Computer Science University of Lincoln Lincoln, UK LN6-7TS syue@lincoln.ac.uk

Abstract-An ability to predict collisions is essential for current vehicles and autonomous robots. In this paper, an integrated collision predication system is proposed based on neural subsystems inspired from Lobula giant movement detector (LGMD) and directional selective neurons (DSNs) which focus on different part of the visual field separately. The two type of neurons found in the visual pathways of insects respond most strongly to moving objects with preferred motion patterns, i.e., the LGMD prefers looming stimuli and DSNs prefer specific lateral movements. We fuse the extracted information by each type of neurons to make final decision. By dividing the whole field of view into four regions for each subsystem to process, the proposed approaches can detect hazardous situations that had been difficult for single subsystem only. Our experiments show that the integrated system works in most of the hazardous scenarios.

Keywords—LGMD, DSNs, looming, lateral movements, visual neural networks, collision detection, expanding edge enhancement, integration

I. INTRODUCTION

Predicting collision is a fundamental but vital task for autonomous robots and intelligent vehicles. Although a large amount of researches have been carried out for decades, no complete uniform solution satisfy the current application demands [1]. On autonomous vehicles, typical methods that combine several sensors, such as ultrasound, infra-red, laser and radar, have demonstrated the ability of driving autonomously in normal or restricted driving scenarios [2]. However, the performance of these systems is not always reliable enough and they are not properly challenged with complex collision scenes without human intervention. Moreover, the costs of these techniques are too high because of heavy computing loads and high power consumption.

A cheap and reliable solution to extract the wealth information in the visual scenes is possible, as most animals do. For example, insects, such as locusts or flies, are able to use visual systems to exploit the plentiful visual cues to avoid collision in complex dynamic scenes. In the visual pathways of insects, there are identified specialized neurons which have been known for several decades. The properties of these neurons revealed so far can be used to build unique computational models for visual based collision recognition.

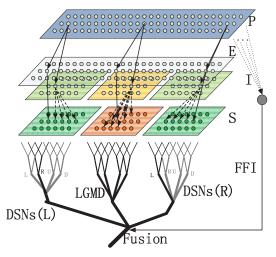


Fig. 1. A schematic illustration of the proposed system. The system consists of one LGMD network, two DSNs networks and a fusion component, which includes a FFI neuron and some other structures. Decisions are made by the fusion unit which connects to all the outputs of neural systems. The LGMD and DSNs networks have analogical structures: photoreceptor cells (P); excitatory and inhibitory cells (E and I); summing cells (S); but have different inhibitory cells connections. of a figure caption

Lobula Giant Movement Detector (LGMD) and Direction Selective Neurons (DSNs) are two classical types of identified neurons found in locusts, each of which has unique characteristic could contribute to collision prediction [3][4].

LGMD is a wide-filed visual neuron located in the Lobula layer of the Locust nervous system [3]. The LGMD increases its firing rate in response to both the velocity of the approaching object and its proximity. The responses are very quickly to looming stimuli and can trigger avoidance reactions when a rapidly approaching object is detected. It is tightly tuned to respond to objects approaching on a direct collision course. Early implementations have been applied in minirobots [5][6] and more extra-features implements can be found in FPGA [7] or ASIC [8].

Directional selective neurons are another type of specialized visual neurons with the features to detect certain directional motion cues, which have also been found in animals for decades [4]. Different directional visual motion

cues are extracted by DSNs through inhibiting a particular direction. Organized in an asymmetrical layered network, these DSNs can also be organized as a neural network specialized for collision recognition, as demonstrated in [9]. For the collision recognition role, it has been demonstrated that the LGMD has been able to build up its ability for collision recognition quickly and robustly in both robotics and driving environments, reducing the chance of other types of neural networks, such as DSNs, to play the same role [10].

Although the LGMD networks could be an ideal model to be realized for efficient collision prediction, it is still difficult to pick up specific collision cues caused solely by lateral visual movements. For example, a pedestrian walks into vehicle lanes suddenly could lead to an imminent collision. However, an integrated system based on LGMD and DSNs can be more practical. Furthermore, the direction and the location is critical – for example, a pedestrian runs into driving lane may lead to an accident while walks out of the same lane may not. These direction and location cues could be easily perceived by DSNs connected to that part field of view.

In this study, we propose an integrated prediction system based on LGMD and DSNs to address these related scenarios. We first implement the LGMD and DSNs models respectively, and then propose methods to connect these neural subsystems to specific regions of the whole field of view depending on their functionalities. The visual cues extracted by these visual neural subsystems are fused – with the outputs of LGMD boosted by the outputs of DSNs in a particular event, to form efficient collision prediction alarms. Experiments are carried out to test the effectiveness of the proposed approach.

The rest of the paper is organized as follows. Section II describes the integrated system in details and section III illustrates the experiments. In section IV we discuss the experimental results against typical driving scenarios and finally draw conclusions in section V.

II. THE INTEGRATED SYSTEM

The proposed system consists of several subsystems - one LGMD network, two DSNs networks and a compact fusion component, as shown in Fig.1. Sharing the same visual inputs, these neural subsystems process visual information with their own unique structures. These subsystems are re-integrated post-synoptically – with the LGMD outputs enhanced by the DSNs for critical scenarios such as hazardous pedestrian crossing. The models of LGMD and DSNs proposed in this paper are based on the models in [5][6][9][11] with some modifications to improve special features in driving scenarios, as detailed in the sections below.

A. The LGMD Neural Subsystem

The LGMD model illustrated in the middle of Fig. 1 is based on the previous model described in [5][6], and to cooperate with DSNs networks, the summing method is changed.

The LGMD model is composed of four groups of cellsphotoreceptor P, excitatory E, inhibitory I and summing S

and a single cell – LGMD. Different from the classical model, here we move the feed-forward inhibition (FFI) to the fusion part.

The first layer of the neural network is the photoreceptor P cells. The luminance L_f of each pixel in the input image is captured by each photoreceptor cell, the change of luminance P_f between frames of the image sequence is calculated and forms the output of this layer. The output of a cell in this layer is defined by equation:

$$P_f(x, y) = |L_f(x, y) - L_{f-1}(x, y)|$$
 (1)

where $P_f(x,y)$ is the change of luminance corresponds pixel (x,y) at frame f, x and y are the pixel coordinates, L_f and L_{f-1} are the luminance, subscript f denotes the current frame and f-1 denotes the previous frame.

Cells in the E and I layers are arranged in matrix forms. The output of the P cells forms the inputs to two separate cell types in the next layer. One is excitation and the other is inhibition. The excitation of a P cell passes to its counterpart in the E layer and the inhibition from a P cell passes to its counterpart in the I layer directly.

The S layer cells are also arranged in a matrix form. The excitation of a E cell passes to its counterpart in the S layer directly. The inhibition from a I cell passes to its retinotopical counterpart's neighboring cells in the S layer with one image frame delay. The gathered strength of inhibition to a cell in this S layer is:

$$S_f(x, y) = E_f(x, y) - W_I \times I_f(x, y) \otimes w_I$$
 (2)

where W_I is the global inhibition weight; w_I is the inhibition template, which in the proposed system is

$$w_{I} = \begin{bmatrix} 1/12 & 1/6 & 1/12 \\ 1/6 & 0 & 1/6 \\ 1/12 & 1/6 & 1/12 \end{bmatrix}$$
 (3)

The symbol \otimes means convolution operation. S_f is the output of a threshold of the sigmoid function which translates the sum of S_f into the range from 0.5 to 1.

$$\tilde{S}_f = \begin{cases}
1, & \text{if } (1 + e^{-\frac{\sum S_f(x,y)}{N_{field}}})^{-1} > th_{LGMD} \\
0, & \text{otherwise}
\end{cases} (4)$$

where N_{field} is the number of the whole vision field and th_{LMGD} is a threshold.

The output of the LGMD cell is alarm signal:

$$LGMD_{f} = \begin{cases} 1, & \text{if } \sum_{i=0}^{len-1} \tilde{S}_{f-i} = len \\ 0, & \text{otherwise} \end{cases}$$
 (5)

Here len is the default number of the spiking. When the continuous spiking number of \tilde{S}_f exceeds len, the

LGMD cell will generate the alarm by setting the value to 1.

B. The DSNs Neural Subsystem

The DSNs models are based on the previous model described in [9][12]. The part of the subsystems is similar to the LGMD module. In the DSNs network, the layers of photoreceptor cells, excitatory cells, inhibitory cells and summing cells also exist while the postsynaptic networks are different. In Fig. 2, the postsynaptic network consists of four direction selective neurons representing the direction of left, right, up and down. For this reason, there are four kinds of inhibition templates corresponding to the four directions, for example, the proposed right-direction inhibition template is

$$w_{I_{_R}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$
 (6)

and W_I is 1/15.

Through comparing the value of SL_f , SR_f , SU_f and SD_f , the output DSN_f is located within five status (unknown, right, left, up, down) represented by (0, 1, 2, 3, 4), which is decided by one of the four DSNs with minimum membrane potential.

Similar to that of the LGMD cell (Equation 5), a filtering process is conducted to increase reliability:

$$DSNo_{f} = \begin{cases} DSN_{f}, & \text{if } DSN_{f} = DSN_{f-1} = \dots = DSN_{f-len+1} \\ 0, & \text{otherwise} \end{cases}$$
 (7)

where DNS_{of} is the output of the filtering process at frame f. Equation 7 checks if the motion direction is stable or not.

C. The Fusion Model and Auxiliary Structure

In driving scenarios, it is important to recognise and respond to hazardous lateral movements caused by, for example, a pedestrian step down to road from pavement or a car slide down to main road at T-junction. However, a normal LGMD will not respond to these scenarios. This is because a pedestrian may be in a moderate distance that only stimulates weak excitation rather than an alarm. Another general problem is false alerts maybe triggered by fast moving objects while the vehicle stay at a T-junction.

To overcome those problems, we propose an integrated system based on integration of LGMD and DSNs – with the outputs of LGMD boosted by the outputs of DSNs in a particular event, to form efficient collision prediction alarms. The deployment layout of the divided field of view is illustrated in Fig. 2.

1) Fusion of the LGMD and DSNs subsystem

In particular events, visual motion from one side to the centre of the field of view, which may be caused by for example pedestrian crossing or vehicle moving into main driving lane in T-junction, could develop to imminent collisions. This type of abnormal visual movements should be

picked up by the specialized subsystems such as DSNs and fed to the system to form a collision alarm. According to this principle, the cues of lateral moving objects acquired by the involved DSNs can be combined with the gain of the membrane potential of LGMD subsystem. We use S_{fu} to represent the cooperated result.

$$S_{fu} = S_{\lg md_L} \times (1 + \alpha_{fu} \times DSNL_f) + S_{\lg md_R} \times (1 + \alpha_{fu} \times DSNR_f)$$
 (8)

where α_{fu} is the interaction strength of two subsystems; the fusion component imports the membrane potential of the LGMD subsystem in S layer, and amplifies the weights of the central regions, then splits S layer to left and right parts from the middle symmetrically, called $S_{lgmd\ L}$ for the left region and

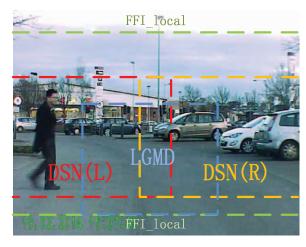


Fig. 2. A deployment of the separated field of view of the proposed system. The area enclosed by a blue box is connected to the LGMD subsystem while in the red and yellow box, DSNs (L) and DSNs (R) are deployed respectively. A FFI neuron captures the outputs of whole field photoreceptor cells and to extra regions at the top and bottom of the field vision are connected to a FFI local component for distinguishing between moving and stop of vehicles.

 $S_{lgmd,R}$ for the right region. The **location** and moving **direction** are further emphasized $-DSNL_f$ is defined in the (9) and $DSNR_f$ has the analogous definition.

$$DSNL_f = \begin{cases} 1, & \text{if } DSNo_f = right \\ 0, & \text{otherswise} \end{cases}$$
 (9)

With Equation (8) and (9), the contribution of the left or right LGMD will be enhanced if there is specific DNS excitation in that part of specific field of view.

2) Suppression of vision field global movement

Without suppression during turning, the whole system could be excessively exciting and even false collision alerts maybe be produced. The feed forward inhibition and lateral inhibition work together to cope with such whole field movement [13]. The FFI excitation at the current frame is gathered from the photoreceptor cells.

$$FFI_{-}all_{f} = \sum_{x} \sum_{y} P_{f}(x, y)$$
 (10)

In typical LGMD models, once FFI_all_f exceeds its threshold th_{FFI} , spikes in the proposed system are inhibited immediately. But it is difficult to distinguish turning scenes from approaching scenes only depending on the threshold.

An important characteristic is the growth trend of FFI_all_f when it exceeds th_{FFI} . In a turning scene, the trend is more flat than in approaching scenes. So in such situation, FFI needs to be set to 1 and a decision to reject false alerts is made through a sample threshold function.

3) Recognition of self-moving

For automobile application, another general principle is that if a car stalls, the alarms should be immunized. Whereas typical LGMD model could be excessively sensitive to particular situations such as other cars crossing the road before the still vision in a near distance, as a result of undesirable alert raise.

In this work, a ratio method of FFI outputs is introduced:

$$R_{\rm f} = FFI \quad | \log a |_{f} / FFI \quad | a |_{f}$$
 (11)

where FFI_local_f is the sum of the outputs from P layer in the upper and lower regions which are showed in green boxes in Fig. 2. If the vehicle is stationary, in many situations the value of FFI_local_f generally remains unchanged; nevertheless FFI_all_f increases when there are objects moving in the vision field. Contrasted to those self-moving scenes, R_f is much lower if no vehicle movement is involved. A factor β is imported to compute the membrane potential of the fusion neuron:

$$\beta = (1 + e^{-\frac{R_f}{C_{FFI}}})^{-1} \tag{12}$$

where C_{FFI} is a constant depending on the area ratio of green box region and the whole vision field of other regions of motion.

4) The output of the fusion neuron

The fusion neuron also has a membrane potential to predict the collision, which is effected by the states of all subsystems:

$$S_{fusion} = S_{fu} \times \beta \tag{13}$$

The same procedures in the LGMD subsystem are conducted with S_{fusion} to generate $S_{O fusion}$ for collision warnings.

The final output of the fusion system is:

$$FUSION = \begin{cases} 1, & \text{if } So_{fusion} = 1 \text{ and } FFI = 0 \\ 0, & \text{otherwise} \end{cases}$$
 (14)

where So_{fusion} is the output of the fusion neuron. *FUSION* equals 1 representing that the collision is imminent.

III. EXPERIMENT SET-UP

We choose several daily driving scenarios to demonstrate the performance of the proposed system. Fig. 3 illustrates the details about the test video sequences. The frame sizes of all the video sequences are 640x480 and the frame rates are 30 frames per second. The original video sources are colour and are converted to grey scale from 0 to 1.

The proposed system is written in Matlab and the computer used in the experiments is a Dell OPTIPLEX 7010 with Intel Core i5-3470.



Fig. 3. Samples from video sequences represent driving environments. The car with the camera is moving forward at different speeds in all the above sequences but sequence 3. The arrows in the images are added to indicate visual motion directions. Video sequence 1 is a car collision scene at about 45mph; video sequence 2 is a left-turning scene at low speed (about 10 mph, the same to the rest if not state differently); video sequence 3, the car with camera with no movements while other cars passing from right to left; video sequence 4 is a left walking pedestrian from the left side of road to the centre; video sequence 5 is a right walking pedestrian from the right side of road to the centre; video sequence 6 is also a left walking pedestrian from the left side of road, but he is far away from the camera; video sequence 7 is a pedestrian walking from the centre of the road to the left side; video sequence 8 is a car driving in the right lane from distance; video sequence 9 is some pedestrians walking along the both sides of the road. Among the test videos, sequence 1, 4 and 5 are dangerous as they are developing quickly to imminent collision and need to trigger alarm. Sequence 2 is a difficult scene. For the proposed system, the FFI neuron will suppress the output of the LGMD subsystem and give a signal representing the global turning

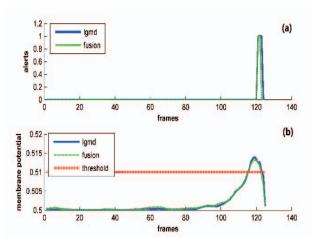


Fig. 4. Comparison of the proposed collision prediction system and a classical LGMD system - response to a looming car (Fig. 3, sequence 1). (a)The blue response comes from the LGMD subsystem while the green from the whole integrated system. The upper figure shows the collision alerts generated by two systems where 1 represents a collision alarm and 0 is for no alarms. The curves demonstrate the successful collision prediction from both systems. (b)The similar curves in the lower figure are the membrane potentials in the LGMD neuron and the fusion neuron, which illustrate the two systems are equivalent for looming collision prediction.

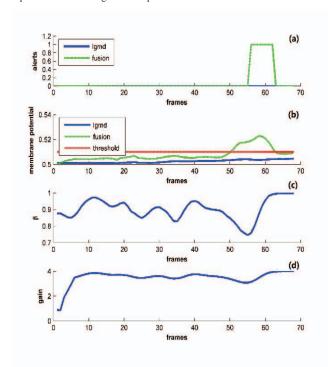


Fig. 5. Comparison of the proposed collision prediction system and a classical LGMD system - response to a left walking pedestrian from the left side of road to the centre (sequence 4). Figure (a), (b) is the same contents to the Fig. 4 while (c) shows the factor β and (d) shows the gain factor affecting the membrane potential in the fusion neuron. In the experiment, the LGMD subsystem fail to predict the crossing pedestrian but the fusion system successfully gives the alerts when the pedestrian is near to the centre of the road.

In the experiments, we challenge the proposed system with the video streams as sampled and explained in Fig. 3. The proposed system passes all the test video sequences listed above. It generates alerts in sequence 1, 4 and 5, and suppresses extra excitation in sequence 2, 3, 7, as expected. The results mean the system overcomes the weaknesses of classical LGMD systems in picking up critical collision cues of pedestrian crossing (Fig. 3, sequence 4 and 5) and demonstrate its ability to predict various collision patterns in hazardous situations.

Further details of the tests are shown in the following Fig. 4 to Fig. 7. Fig. 4 shows the result of video sequence 1 of a typical rear end collision scene. The outputs of classical LGMD model and the proposed system are similar, which means the proposed system inherits the advantages of classical LGMD model in extracting imminent collision cues.

In sequence 4 and 5, a pedestrian is walking into the driving lane, which could develop into a hazard rapidly. The proposed system predicts the lateral moving objects in a moderate distance around 2 meters. As shown in Fig. 5, however, the LGMD model neglects the hazardous pedestrian while our proposed system successfully predicted it as hazard scenes and generates alerts when objects is close to the centre of the field.

Through re-integration post-synoptically with the LGMD outputs, the DSNs, which perceive the specific lateral movements naturally, help the fusion system to augment the outputs of the membrane potential — which eventually resulting in collision prediction alarm. As a comparison, the contribution of DSNs in sequence 7 is completely different. The leftward lateral movement cue in the left of view has not been considered as critical cues in the proposed system because the pedestrian walking from the middle of the road to the left side exert no harm to the pedestrian or vehicles — there is no need for the DSNs modules to enhance the excitation of the LGMD subsystem.

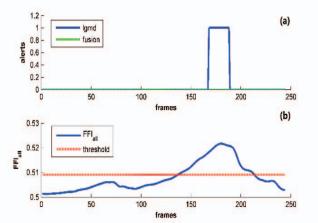


Fig. 6. Comparison of the proposed collision prediction system and a classical LGMD system - response to a left-turning scene (sequence 2). (a) shows the system alerts and (b) is the FFI_all response. Through the threshold the system suppress the false alters caused by vision turning. Here the LGMD subsystem fails because the FFI module in original models has been moved to the fusion component.

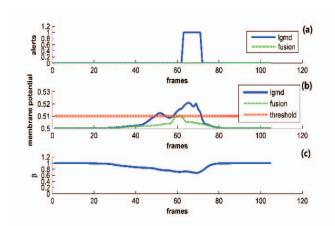


Fig. 7. Comparison of the proposed collision prediction system and a classical LGMD system - response to sequence 3. In this sequence, other cars are moving along the road in front of the still car. (a) is the system alerts, (b) is the membrane potentials in the LGMD neuron and the fusion neuron and (c) is the factor β to suppress the excitations generated by crossing cars. The figures show that the LGMD subsystem is extremely sensitive to the scene and generate the false alerts. Benefit by the factor β , the membrane potential in the fusion neuron descends and the false alters are inhibited.

Similarly, the responses to the sequence 8 and 9 also show that proposed system pays little attention to the harmless objects moving in the left and right edges, which are rational for driving environment.

Fig. 6 and Fig. 7 show the ability of suppressing false alarm from LGMD caused by near range lateral movements (Fig. 3, sequence 2, and 3). As shown in Fig. 7, the excitation caused by global turning is well reduced by the proposed FFI module. In sequence 3, the car stalls facing the road and some other cars drive along the road quickly exhibiting fast lateral visual movements to the system. Because the distance is short, the classical LGMD model gives a false alarm. Our proposed system responds to this scenario correctly with the help of the FFI_side module, which generates a lower factor and perfectly suppresses the output of LGMD subsystem.

Although the proposed integrated collision prediction system, as shown in the above, can predict most of the critical collision cues in similar driving scenarios, it does not mean the system can deal with all the complex collision cues in other different situations with large variances in background complexity, illumination conditions, moving speeds and sizes of objects, and so on. Considering the complexity of typical driving scenes, we will address the robustness and efficiency of the proposed collision prediction system in future work.

V. CONCLUSIONS

In the paper, we propose a bio-inspired collision prediction system based on the fusion of a LGMD subsystem, the two DSNs subsystems and other auxiliary modules. We modify these network models to suit for easy integration and applications for moving machines such as vehicles and autonomous robots. Benefiting from the robust collision

prediction ability inherited with LGMD subsystem and the enhancement by the DSNs in particular events, the experiment results demonstrate that our proposed approach can pick up critical lateral movement cues lead to hazard situations which should trigger imminent collision alarm.

Future work will include finer integration of LGMD network with DSNs modules to further enhance the selectivity for the visual challenges and implementing the system onto FPGA chips.

Acknowledgment

This work is partially supported by EU FP7 projects LIVCODE (295151), HAZCEPT (318907), Horizon 2020 projects ENRICHME (643691) and STEP2DYNA (691154).

References

- A. Vahidi, A. Eskandarian, Research advances in intelligent collision avoidance and adaptive cruise control, IEEE Trans. Intell. Transp. Syst. 4 (3) (2003) 143–153.
- [2] E. Guizzo, How Google's self-driving car works, IEEE Spectrum, Robotics Blog 18 (2011).
- [3] S.J. Judge, F.C. Rind, "The locust DCMD, a movement detecting neurone tightly tuned to collision trajectories," Journal of Experimental Biology, vol.200, pp.2209–2216, 1997.
- [4] F. C. Rind, "Identification of directionally selective motion-detecting neurones in the locust lobula and their synaptic connections with an identified descending neurone," J. Exper. Biol., vol. 149, pp. 21–43, 1990b.
- [5] S. Yue, F.C. Rind, A collision detection system for a mobile robot inspired by the locust visual system, in: ICRA, 2005, pp. 3832–3837.
- [6] S. Yue, F.C. Rind, Collision detection in complex dynamic scenes using a LGMD based visual neural network with feature enhancement. IEEE Transactions on Neural Networks, 17(3), pp.705-716,2006
- [7] H. Meng, K. Appiah, S. Yue, A. Hunter, M. Hobden, N. Priestley, P. Hobden, and C. Pettit, "Modified model for the lobula giantmovement detector and its FPGA implementation," Comput. Vis. Image Understand., vol. 114, no. 11, pp. 1238–1247, 2010.
- [8] G. Linan-Cembrano, L. Carranza, C. Rind, A. Zarandy, M. Soininen, A. Rodriguez-Vazquez, "Insect-vision inspired collision warning vision processor for automobiles", Circuits and Systems Magazine, IEEE, Volume 8, Issue 2, 2008, pp. 6-24.
- [9] S. Yue and F. C. Rind, "Postsynaptic organization of directional selective visual neural networks for collision detection," Neurocomputing, vol. 103, pp. 50–62, 2013.
- [10] S. Yue and F. C. Rind, "Redundant Neural Vision Systems-Competing for Collision Recognition Roles," IEEE Transactions on Autonomous Mental Development, Volume 5, Issue 2, 2013, pp. 173-186.
- [11] M. Blanchard, F. C. Rind, and P. F. M. J. Verschure, "Collision avoidance using a model of the locust LGMD neuron," Robot. Autonom. Syst., vol. 30, pp. 17–38, 2000.
- [12] S. Yue and F. C. Rind, "A synthetic vision system using directionally selective motion detectors to recognize collision," Artif. Life, vol. 13, no. 2, pp. 93–122, 2007.
- [13] R. D. Santer, R. Stafford, and F. C. Rind, "Retinally-generated saccadic suppression of a locust looming detector neuron: Investigations using a robot locust," J. Roy. Soc. London Interface, vol. 1, pp. 61–77, 200