

A bio-inspired model reliably predicts the collision of approaching objects under different light conditions

Ana Carolina Silva, and Cristina Peixoto dos Santos

Industrial Electronic Department, University of Minho, Portugal

Abstract. In this paper, we present a model of the Lobula Giant Movement Detector, which is a part of a visual pathway responsible for triggering collision avoidance manouvres in the locust *Locusta Migratoria*. Also based on locust neural adaptation to transitions in light intensities, the model proposed here integrates a mechanism for light adaptation. The tests performed with the model demonstrate its ability to reproduce several characteristic properties of the LGMD response, including the firing rate profile for different visual stimuli. Additionally, results obtained for different light conditions show that the increase in the LGMD model efficiency is provided by the new mechanism of light adaptation.

Keywords: Bio-inspired model, Lobula Giant Movement Detector neuron, artificial neural networks, collision avoidance, spatiotemporal summation

1 Introduction

The choice of a model system for study is often difficult, as the complexity of behaviour is quite often related to the complexity of an animal's nervous system. Nevertheless, some animals have been pressured by natural selection to obtain very complex behaviours, while maintaining a relatively simple nervous system. Insects are a perfect example. One advantage of studying the nervous system of insects is based on the fact that they have fewer neurons than other animals such as vertebrates. In addition, the properties of a single identified neuron can often yield general properties and mechanisms that are applicable to other systems [1].

One insect in particular, the locust, has evolved a dedicated and well-studied collision avoidance neural pathway that is responsible for generating collision avoidance behaviours to avoid predation and continual in-flight collisions with conspecifics [2]. The visual system of the locust is paramount to its survival, and acts as a great model system for study.

The eyes of the *Locusta Migratoria* are apposition compound eyes, highly tuned to activity during day light. However, it is known that locusts perform migration flights at low level intensities [3]. After a deep research it was found that, in environments with low level intensities, locusts improve the photon capture neurally by summing the outputs of neighbouring visual channels (spatial

summation) and/or by increasing the length of time a sample of photons is counted by the compound eye (temporal summation). So, using spatiotemporal summation, locusts are able to extend their vision into dim light [4].

Locusts have a large neuron in their brain called the Lobula Giant Movement Detector (LGMD), that is tightly tuned to respond to objects approaching on a direct collision course [5,6,7]. The selectivity of the LGMD for looming stimuli is established by circuits among neurons in the locust optic lobe. The LGMD receives excitation from a large number of afferents, which excitation level depends on the subtended angle of the object on the locust eye. Thus, LGMD excitation increases as a looming object gets closer. On the other hand, non-colliding objects do not show the same increase in excitation and, by this reason, are unlikely to trigger avoidance reactions [2,8,9,10,11]. LGMD spikes are transferred to the Descending Contralateral Movement Detector (DCMD) which, subsequently, is connected to flight interneurons and motorneurons within the thoracic ganglia [2,10]. Therefore looming responses in this pathway may have consequences for collision avoidance behaviours.

Based on the relatively simple encoding strategy of the LGMD, we concluded that it was a good candidate as a bio-inspired model for detecting collisions. According to literature, the first physiological and anatomical bio-inspired model for the LGMD neuron was developed by Bramwell [12]. The model continued to evolve [13,14,15,16] and it was used in mobile robots and even in automobiles for collision detection. However, further work is needed to develop more robust models that can account for complex aspects of visual motion [10], as well as be able to work at different light intensities [4]. In this article, we are interested in integrating two recent LGMD models, [14] and [16], in order to take the advantage of noise immunity proposed in the first and direction sensitivity proposed in the second. Based on the principles of the locust visual system previously referred, we subsequently add a capability of light adaptation to the developed model.

In this article the proposed model is verified when submitted to artificial visual stimuli of approaching and receding objects, that include both different light intensities and noise and a real video captured by a camera. The results show that the system avoids approaching obstacles in an effective way despite these disturbances.

2 The model

The biologically inspired neural network here proposed is based on previous models described on [13,14,15,16]. The modified neural network is shown on figure 1. In this model, we integrated an adaptative spatiotemporal summation system, which has the capability to adapt the neural network to different light intensity scenarios.

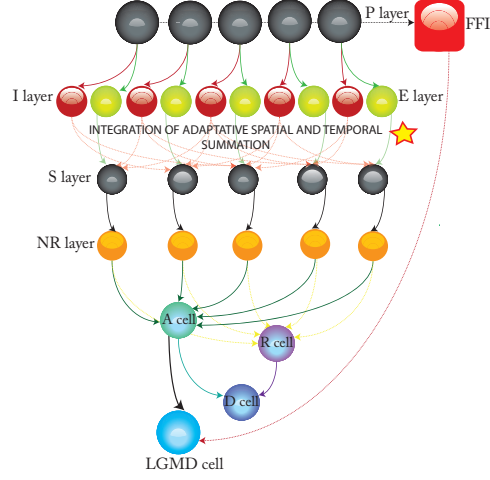


Fig. 1. Schematic illustration of the proposed LGMD model.

A grayscale image of the camera current field of view, represented has a matrix of values (from 0 to 255), is the input to a matrix of photoreceptor units (P layer).

This layer calculates the absolute difference between the luminance of the current and of the previous input images, that is:

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

where $P_f(x, y)$ is the output relative to the cell in the (x, y) position at frame f , $L_f(x, y)$ and $L_{f-1}(x, y)$ are the captured luminance at position (x, y) for frames f and $f - 1$, respectively. This layer allows us to detect the object edges. The output of the P layer is the input of two different layers: the excitatory (E) and the inhibitory (I) layer. To the excitatory cells of the E layer, the excitation that comes from the P layer is passed directly to the retinotopic counterpart at the S layer. And the inhibition layer (or I layer) receives the output of the P layer and applies a blur effect on it, using:

$$I_f(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 P_{f-1}(x+i, y+j) \cdot W_l(i, j), \quad i, j \neq 0, \quad (2)$$

where $I_f(x, y)$ is the inhibition relative to the cell in the (x, y) position at frame f , $W_l(i, j)$, an empirically set kernel, represents the local inhibition weight. The inhibition from the (x, y) cell only spreads to the nearest neighbors and does not inhibits itself. This process is strongly supported by the biological nervous systems. In the biology, an excited neuron do not inhibits himself, it inhibits the neighboring neurons, with a temporal delay associated to the inhibitory synapses (and this is the reason why we use the P cell excitement corresponding to the previous time-step, $f - 1$). Relative to the definition of the values holding by this

kernel, the inhibition value of a particular cell is given by the distance at which a neighboring cell is located. The use of such kernel is also based on biological systems, since distant neurons inhibit a particular neuron with less intensity than those that are closest to a neuron, due to the decrement of the neuronal signal with increasing distance. Finally, the excitatory flux from the E cells and the inhibition that comes from the I cells are summed by the S cells (summing cells), using the following equation:

$$S_f(x, y) = E_f(x, y) - w_i \cdot I_f(x, y), \quad E_f(x, y) = P_f(x, y), \quad (3)$$

where w_i (a scalar) represents the inhibition strength. Based on [14], a new mechanism for the LGMD neural network was added to filter background noise. This mechanism, implemented in the NR layer, takes clusters of excitation in the S units to calculate the input to the LGMD membrane potential. These clusters provide higher individual inputs than the ones of isolated S units. The excitation that comes from the S layer is then multiplied by a passing coefficient C_{ef} , which value depends on the surrounding neighbours of each pixel, calculated as follows:

$$C_{ef}(x, y) = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 S_f(x+i, y+j) \quad (4)$$

The final excitation level of each cell in the NR (Noise-Reduction) layer, at frame f (NR_f), is given by:

$$NR_f(x, y) = |S_f(x, y) \cdot C_{ef}(x, y) \cdot w^{-1}| \quad (5)$$

$$w = \max(|C_{ef}|) C_w^{-1} + \Delta c \quad (6)$$

C_w is set to 4, Δc is a small number (0.01) to prevent w from being zero, and $\max(|C_{ef}|)$ is the largest element in matrix $|C_{ef}|$. Within the NR layer, a threshold filters the decayed excitations (isolated excitations), as:

$$\tilde{NR}_f(x, y) = \begin{cases} NR_f(x, y), & \text{if } NR_f(x, y) \cdot C_{de} \geq T_{de} \\ 0, & \text{if } NR_f(x, y) \cdot C_{de} < T_{de} \end{cases}, \quad (7)$$

where $C_{de} \in [0, 1]$ is the decay coefficient and T_{de} is the decay threshold (set to 20). The decay threshold here used was experimentally determined. The NR layer is able to filter out the background detail that may cause excitation. Hence, only the main object in the captured scene will cause excitation. The $LGMD$ potential membrane K_f , at frame f , is summed after the NR layer,

$$LGMD_f = K_f = \sum_{x=1}^n \sum_{y=1}^m (\tilde{NR}_f(x, y)), \quad (8)$$

where n is the number of rows and m is the number of columns of the matrix representing the captured image. The A (Approaching) and R (Receding) cells (adapted from [16]) are two grouping cells for depth movement direction recognition. The A cell holds the mean of three samples of the $LGMD$ cell.

The R cell shares the same structure as the A cell but with a temporal difference, having one frame delay from A . According to the previously described, it can be concluded that if the object is approaching $A_f > R_f$ and if the object is receding, $R_f > A_f$. The D cell or Direction cell ($\in \{-1, 0, 1\}$ in case of receding, no movement and approaching object, respectively) is used to calculate the direction of movement. This cell exploits the movement direction in depth. It is based on the fact that a looming object (approaching) gets larger whereas a receding object gets smaller. In a way to distinguish the movement direction detected by the D cell, it was added a threshold mechanism, $T_D(0.05 \times n \times m)$, which was experimentally determined.

$$D_f = \begin{cases} 1, & \text{if } |A_f| - |R_f| \geq T_D \\ 0, & \text{if } T_D < |A_f| - |R_f| < T_D \\ -1, & \text{if } |A_f| - |R_f| \leq T_D \end{cases} \quad (9)$$

The LGMD membrane potential K_f is then transformed to a spiking output $k_f \in [0.5, 1]$ using a sigmoid transformation,

$$k_f = (1 + e^{-K_f \cdot n_{cell}^{-1}})^{-1}, \quad (10)$$

where n_{cell} is the total number of cells in the NR layer. The collision alarm is decided by the spiking of the $LGMD$ cell.

A spiking mechanism was implemented using an adaptable threshold. This threshold starts with a value experimentally determined, T_s (0.88) and it is updated at each frame, through the following process,

$$T_s = \begin{cases} T_s + \Delta t, & \text{if } s_{av} > \Pi \text{ and } (T_s + \Delta t) \in [T_l, T_u] \\ T_s - \Delta t, & \text{if } s_{av} < \Pi \text{ and } (T_s - \Delta t) \in [T_l, T_u], \\ T_s, & \text{others} \end{cases} \quad (11)$$

where $[T_l, T_u]$ defines the lower and upper limits for adaptation (T_l is 0.80 and T_u is 0.90), $\Delta t = 0.01$ is the increasing step, $\Pi = 0.72$ is a threshold that limits the averaged spiking output s_{av} , between frame $f - 5$ to frame $f - 2$,

$$s_{av} = 0.25 \sum_{i=2}^5 s_{f-i}. \quad (12)$$

If the sigmoid membrane potential k_f exceeds the threshold T_s a spike is produced, as follows:

$$s_f = \begin{cases} 1, & \text{if } k_f \geq T_s \\ 0, & \text{others} \end{cases} \quad (13)$$

Finally, a collision is detected when there are n_{sp} spikes in n_{ts} time steps ($n_{sp} \leq n_{ts}$), where n_{sp} is 4 and n_{ts} is 5 (values experimentally determined).

$$C_f = \begin{cases} 1, & \text{if } \sum_{f-n_{ts}}^f s_f \geq n_{sp} \\ 0, & \text{others} \end{cases} \quad (14)$$

The escape behavior is initialized when a collision is detected. Besides that, the spikes can be suppressed by the *FFI* cell when whole field movement occurs. If it is not suppressed during the tuning of the robot, for example, the network may produce spikes and even false collision alerts due to sudden changes in the visual scenario.

The *FFI* cell is a cell which is very similar to the *LGMD* cell but receives the output from the *P* layer (and not from the *S* layer), as follows:

$$\text{FFI}_f = \frac{\sum_{x=1}^m \sum_{y=1}^n |P_{f-1}(x, y)|}{\text{n cell}}, \quad (15)$$

where P_{f-1} is the output of the *P* layer at frame $f - 1$. If FFI_f exceeds a threshold T_{FFI} (experimentally set to 25), the spikes produced by the *LGMD* cell are automatically inhibited.

We have introduced a light adaptative system within the *S* layer, that comprises three different states depending on the mean gray scale value in the captured image: STATE 1 is activated in the presence of a very low light intensity scenario (mean gray scale value lower than 50); STATE 2 is activated in the presence of a medium light intensity scenario (mean gray scale value lower than 150); and STATE 3 is activated in the presence of a high light intensity scenario (mean gray scale value greater than 150).

In STATE 1 and STATE 2, a Gaussian spatial kernel (K_k) is chosen, and the following processment is done within the *S* layer:

$$S_f \text{ spatial}(x, y) = \sum_{i=a}^b \sum_{j=a}^b S_f(x+i, y+j) \cdot K_k, \quad (16)$$

where K_k ($k = \text{State1}, \text{State2}$) is a kernel that allows the spatial summation between neighboring pixels within the *S* layer; and a (b) is -2 (2) or -1 (1) depending whether on State 1 or 2, respectively. The K_{State1} has a gaussian distribution, with a standard deviation of 1, and a strenght of 20 units. The K_{State2} has a gaussian distribution, with a standard deviation of 0.5, and a strenght of 4 units. This kernel allows the spatial summation between neighboring pixels within the *S* layer.

Besides this spatial summation, when any of these states is activated, the model will activate a temporal summation mechanism, that sets the final value for each pixel of the *S* layer, as follows:

$$S_f(x, y) = \sum_{f=c}^f S_f \text{ spatial}(x, y), \quad (17)$$

where $c = -3$ or -1 whether on State 1 or 2, respectively.

In case STATE 3, due to the high intensity level, the spatial and temporal summation are not activated.

3 Experiments and Results

In order to assess the effectiveness of the proposed LGMD model, we develop a simulation environment in MATLAB, using a Laptop (Toshiba Portegé R830-10R) with 4 GHz CPU and Windows 7 operating system. To test the efficiency of the LGMD model, different data sets were used. The first experiment was made on a simulated data set, showing a square approaching at different velocities, with a high noise level (500 pixels of random noise, corresponding to 5 percent of the image) and with different light conditions. The second experiment was made with receding stimuli, showing the same characteristics previously described for the approaching stimuli.

As previously mentioned, the input of the LGMD model are images, which represent a 2 dimensional information of light intensity. Lower intensity is represented by a lower gray scale value within an image. Artificial visual stimuli, with different mean of background intensity: A)10; B)30; C)50; D)70; E)90; F)120; G)200; H)255, were used to stimulate the LGMD model.

In our first experiment, we evaluated the LGMD model, by using a looming stimulus consisting of a solid square, for two different relations between the object half size (represented by l) and object velocity (represented by v). This ratio determines the time course of the angular size of the looming stimulus. We selected these two relations between l and $|v|$ due to their importance on the stimulation of the locust visual system [17][5]. Figure 2 shows the output from the LGMD model, when using background H. In this figure, at each time step we can observe the result of different mathematical processing, corresponding to the layers of the proposed model, executed sequentially, necessary to detect an imminent collision. Observing the top panels of figure 2, in which it is represented the *Spike Rate* of the LGMD model (obtained by dividing the A_f value for the number of pixels in the captured image) we observe that the waveform obtained is consistent with the biological data reported to the same visual stimuli [10]. The analysis of these results showed that the LGMD neural network detected a collision at time -0.07 seconds (for $l/v=25$ ms) and -0.12 seconds ($l/v=50$ ms), being the simulated square located at 14 cm and 12 cm, respectively.

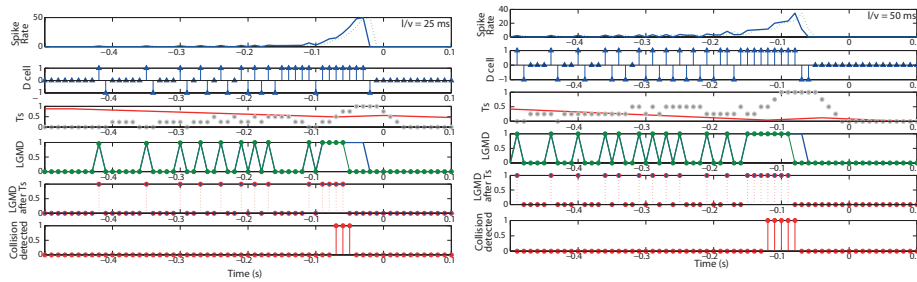


Fig. 2. LGMD model response to an approaching object with l/v set at 25 milliseconds (left panel) and l/v set at 50 milliseconds (right panel). In all these graphs, the zero value corresponds to the time of collision.

After this test, we decrease the background light intensity, following the order from G to A. For these tests, we obtained the results resumed in figure 3.

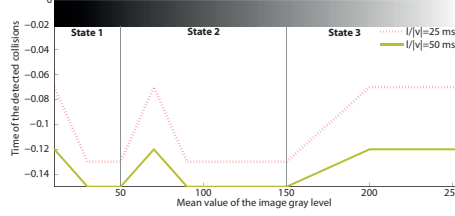


Fig. 3. LGMD model final output to $l/v = 25$ (dots) and 50 milliseconds (line), when subjected to visual stimulus with different light intensities.

At different mean values of the image gray level and for each l/v relation, the output of the LGMD model relative to the detected collision (figure 3), did not always happen at the same time instant (ranging from -0.07 seconds (or 14cm of distance) to -0.12 seconds (or 26 cm) for $l/v = 25$ ms and ranging from -0.12 (or 12 cm) to -0.15 seconds (or 15 cm) for $l/v = 50$ ms). Through these results we verify that the distance at which the LGMD model detected a collision almost doubled in some situations. This happened due to the fact that we only have 3 different states of light adaptation. So, for the highest mean values of the image gray level within a state, the intensity of the image will be already too high for the size of the kernel used in the spatial summation processing, as well as for the temporal integration used. Consequently, the excitation level of the LGMD cell will strongly increase, leading to the generation of premature “collision detected” spikes. However, for $l/v = 50$ ms, the collisions were detected when the object was located at 12 or 15 cm, depending on the mean value of the image gray level. Based on these results we can conclude that for high l/v values, spatial and temporal summation has a lower effect on the distance at which the collisions were detected. After this first set of tests, we repeated the same stimuli but using a receding trajectory. In all the situations tested no collisions were detected, as expected.

In addition to these simulated visual stimuli, and in order to test the LGMD model in a real environment, we recorded a real video sequence, using a Sony Cyber shot digital camera 7.2 megapixels to obtain the video clip (figure 4, left panel).

The resolution of the video images is 640 by 480 pixels, with 30 frames per second of acquisition frequency. After the video recording and using a movie editor, we created two new video recordings, one showing an environment with high light intensity and other with very low intensity (see figure 4). By using these three video sequences, we were able to test three different stages of the LGMD model light adaptation. Observing the left panel in figure 4, we observe that, for video sequences with low and medium intensity level, the collisions were detected at -0.08 seconds, ie, when the ball was at 26 cm relatively to the camera. However, for high intensity level, the collision was detected later

(-0.04 seconds), ie, when the ball was located at 16 cm to the camera. Despite this difference, consequence of the distinct processing made in each state, the obtained results were very satisfactory, since the LGMD model here proposed is able to: remove the background noise, detect the direction of stimuli movement as well as auto-adapt to different light intensity scenarios, in a high effective way.

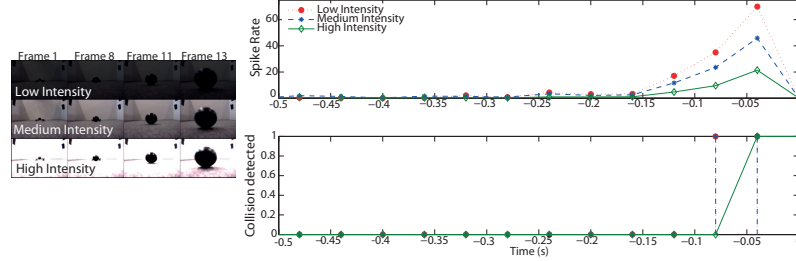


Fig. 4. LGMD model response to a real video sequence under different light conditions.

Based on our results, we conclude that the LGMD model proposed here reliably detects collisions of objects of different sizes, different trajectories and under different light intensities. Our bioinspired model reproduces the response of the locust LGMD neuron and is able to account for spatiotemporal summation adaptation in different light conditions.

4 Conclusions

In this paper, we have presented a model of the collision avoidance behaviour of the *Locusta Migratoria*. The model reproduces the behaviour of the locust LGMD neuron and is able to account for spatiotemporal summation adaptation in different light conditions. It was shown that a neural network based model could describe the behaviour of the locust in a wide range of stimulus conditions. This model has a mechanism favouring grouped excitation, as well as two cells with a particular behaviour that provide additional information on the depth movement direction. For applications as collision detectors for robots, the model proposed is able to remove the noise captured by the camera, as well as enhance its ability to recognize the direction of the object movement and, by this way, remove the false collision alarms produced by the previous models when a nearby object is moving away. Furthermore, the model is able to autonomously adjust to different light conditions.

Acknowledgements

Ana Silva is supported by PhD Grant SFRH/BD/70396/2010, granted by the Portuguese Science Foundation. The authors would like to thank Glyn McMillan for his revisions and suggestions to improve the manuscript.

References

1. Zupanc, G. K. H.: Behavioral Neurobiology: An Integrative Approach. Oxford University Press, (2010)
2. O'Shea, M., Williams, J. L. D.: The anatomy and output connection of a locust visual interneurone; the lobular giant movement detector (LGMD) neuron. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 257–266 (1974).
3. Chapman, R. F.: The insects: Structure and Function. Hodder and Stoughton, London (1980)
4. Warrant, E. J.: Seeing better at night: light style, eye design and the optimum strategy of spatial and temporal summation. *Vision Res* 39, 1611–1630 (1999)
5. Gabbiani, F., Krapp, H., Laurent, G.: Computation of object approach by a wide-field motion-sensitive neuron. *J. Neurosci.* 19, 1122–1141 (1999)
6. Gabbiani, F., Mo, C., Laurent, G.: Invariance of Angular Threshold Computation in a Wide-Field Looming-Sensitive Neuron. *The Journal of Neuroscience* 21(1), 314–329 (2001)
7. Gabbiani, F., Krapp, H. G., Koch, C., Laurent, G.: Multiplicative computation in a visual neuron sensitive to looming. *Nature* 420, 320–324 (2002)
8. Gray, J. R., Lee, J. K., Robertson, R. M.: Activity of descending contralateral movement detector neurons and collision avoidance behaviour in response to head-on visual stimuli in locusts. *Journal of Comparative Physiology A*, 115–129 (2001)
9. Rind, F. C.: Non-directional, movement sensitive neurones of the locust optic lobe. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 477–494 (1987)
10. Guest, B. B., Gray, J. R.: Responses of a looming-sensitive neuron to compound and paired object approaches. *Journal of Neurophysiology* 95(3), 1428–1441 (2006)
11. Gray, J. R., Blinow, E., Robertson, R.: A pair motion-sensitive neurons in the locust encode approaches of a looming object. *Journal of Comparative Physiology A* 196(12), 927–938 (2010)
12. Rind, F. C., Bramwell, D. I.: Neural Network Based on the Input Organization of an Identified Neuron Signaling Impeding Collision. *Journal of Neurophysiology* 75(3), 967–985 (1996)
13. Blanchard, M., Rind, F. C., Verschure, P. F. M. J.: Collision avoidance using a model of the locust LGMD neuron. *Robotics and Autonomous Systems* 30(1), 17–37 (2000)
14. Yue, S., Rind, F. C.: Collision detection in complex dynamic scenes using an LGMD-based visual neural network with feature enhancement. *IEEE transactions on neural networks* 17 (3), 705–716 (2006)
15. Stafford, R., Santer, R. D., Rind, F. C.: A bio-inspired visual collision detection mechanism for cars: combining insect inspired neurons to create a robust system. *BioSystems* 87, 164–171 (2007)
16. Meng, H., Yue, S., Hunter, A. A., K., Hobden, M. P., N., Hobden, P., Pettit, C.: A modified neural network model for the Lobula Giant Movement Detector with additional depth movement feature. In: *Proceedings of International Joint Conference on Neural Networks*, pp. 14–19. Atlanta, Georgia (2009)
17. Badia, S. B. i., Bernardet, U., Verschure, P. F. M. J.: Non-Linear Neuronal Responses as an Emergent Property of Afferent Networks: A Case Study of the Locust Lobula Giant Movement Detector. *PLoS Comput Biol* 6(3), e1000701 (2010)