Published in IEEE Trans.Neural Networks, vol.17, issue 3, May 2006, page 705-716, DOI: 10.1109/TNN.2006.873286

Collision Detection in Complex Dynamic Scenes Using a LGMD Based Visual Neural Network with Feature Enhancement[↓]

Shigang YUE, *Member, IEEE* and F. Claire RIND

***Abstract*—The lobula giant movement detector (LGMD) is an identified neuron in the locust brain that responds most strongly to the images of an approaching object such as a predator. Its computational model can cope with unpredictable environments without using specific object recognition algorithms. In this paper, a LGMD based neural network is proposed with a new feature enhancement mechanism to enhance the expanded edges of colliding objects via grouped excitation for collision detection with complex backgrounds. The isolated excitation caused by background detail will be filtered out by the new mechanism. Off-line tests demonstrated the advantages of the presented LGMD based neural network in complex backgrounds. Real time robotics experiments using the LGMD based neural network as the only sensory system showed that the system worked reliably in a wide range of conditions, in particular, the robot was able to navigate in arenas with structured surrounds and complex backgrounds.**

***Index Terms*—collision detection, dynamic visual scene, visual neural network, complex environment, mobile robot**

I. Introduction

The ability to avoid collisions is important for many mobile/intelligent machines. Mobile robots have used several kind of sensors, such as visual, ultrasound, infra-red, laser, and mini-radar, for object detection (for example, [10, 1, 38, 2, 21]). However, it is still very difficult for a robot to run autonomously without collision in complex, outdoor environments without human intervention. For intelligent machines one of the greatest challenges is to understand and cope with dynamic scenes [6]. Visual sensors have evolved as an important organ (eyes) for animals to exploit the plentiful cues in the real visual world and eyes play an important role in the survival of animals. However, artificial robot vision systems have not yet been able to quickly and cheaply extract the wealth information present in the visual environment [20, 31, 8].

For animals, such as insects, the ability to detect approaching objects is important, serving both to prevent collision as the animal moves and also to avoid capture by predators [28, 33]. Evolved over millions of years, the visual collision avoidance systems in insects are both efficient and reliable. The neural circuits processing visual information in insects are relatively simple compared to those in the human brain and are a perfect model for the optical collision avoidance sensors that should equip mobile intelligent machines [29]. As an example, the 'Elementary Motion Detectors' (EMDs) of the fly have been realised electronically and used to control a robot avoiding stationary objects or making a straight path as it localises and approaches a sound source (recent examples, [16, 19, 15, 37], reviewed by [12]). An identified interneuron in locust visual system, the lobula giant movement detector (LGMD), which responds vigorously to looming objects, is another neural model to be applied in autonomous robots [29].

The lobula giant movement detector (LGMD) is a large visual interneuron in the optic lobe of the locust [23] that responds most strongly to approaching objects [34, 26]. It is tightly tuned to respond to objects approaching the locust on a direct collision course [17], but produces little or no response to receding objects [26]. This makes the LGMD an ideal template to develop specialised sensors for automatic collision avoidance. Also, the electrophysiological knowledge about the LGMD neuron and its afferent pathway revealed in the last several decades (e.g., [23, 30, 34, 26] makes it possible to model it.

A functional neural network based on the LGMD's input circuitry was developed by Rind and Bramwell [27]. This neural network showed the same selectivity as the LGMD neuron for approaching rather than receding objects and responded best to objects approaching on collision rather than near-miss trajectories. The expanding edges of colliding objects and the lateral inhibition were the key features computed by the model. This neural network has also been used to mediate collision avoidance in a real-world environment by incorporating it into the control structure of a

Manuscript received May 11, 2005. This work was supported by the LOCUST project EU IST-2001-38097.

Shigang YUE is with Ridley Building, the School of Biology and Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne, NE1 7RU, UK. (corresponding author; tel: (0044) 0191 222 5720; fax: (0044) 0191 222 5229; e-mail: shigang.yue@ ncl.ac.uk).

F. Claire RIND, is with Ridley Building, the School of Biology and Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne, NE1 7RU, UK. (e-mail: claire.rind@ncl.ac.uk).

[↓] The paper in part was presented in the IEEE International Conference on Robotics and Automation, Spain, Barcelona, April, 2005.

miniature mobile robot [3, 4]. The collision avoidance controlled via the LGMD was successful on at least 69% of occasions, and for half of the speeds tested, it was successful on over 90% of occasions [4].

In real world challenges, robots have to deal with colliding objects against complex visual background. The previous LGMD based neural network either challenged only by pure computer generated visual stimuli [27], or tested in a simple, structured environment [4, 32]. The robustness of the neural network needs to be improved to increase the success rate in deal with colliding objects against a complex background.

This study further develops the LGMD based neural network to detect colliding objects in complex background. Based on the previous LGMD neural networks [27, 4, 32, 35], we propose a neural network to form the collision detection system with a new mechanism processing the excitations (refer to pixels with higher value hereafter) before the LGMD cell gathers excitations. The new mechanism favours grouped excitations by enhancing them and allowing them reach to next layer without decay; however the isolated excitations is not enhanced and is therefore subject to decay. This is consistent with the recent finding that the summation within the real LGMD's dendritic tree is highly non-linear and probably its dendritic tree structure and conductancees may play an important role in filtering incoming excitation and inhibition [14]. When integrated with a mobile robot, a special mechanism is used to regulate the response of the network by mediating the threshold. With an adaptable threshold, the system is able to work under extremely bright or dark conditions. Experiments are carried out to test the feasibility and advantages of the new collision detection system in complex environments with different conditions.

II. Formulation of the System

The LGMD based neural network proposed in this paper was based on previous studies described in [27, 4, 32, 35]. The LGMD neural network [27, 4] was composed of four groups of cells - photoreceptor, excitatory, inhibitory and summing, and two single cells - feed-forward inhibition and LGMD [27]. These groups of cells were also used as a basic for the modified neural network (Figure 1). To improve the robustness of the LGMD based neural network in situations where the background of a visual scene is complex, a new layer of grouping cells (G cells) were added to enhance the visual feature defining a colliding object and filter visual details irrelevant to the collision detection task. When integrated with robots, new cell feed-forward mediation (FFM), will be introduced to mediate the response of the LGMD cell by varying its threshold to cope with extreme luminance conditions. The proposed neural network (shown in Figure 1) used in the paper will be described in detail in the next part (please note that the G cells and FFM cell may not have exact counterparts in real locusts).

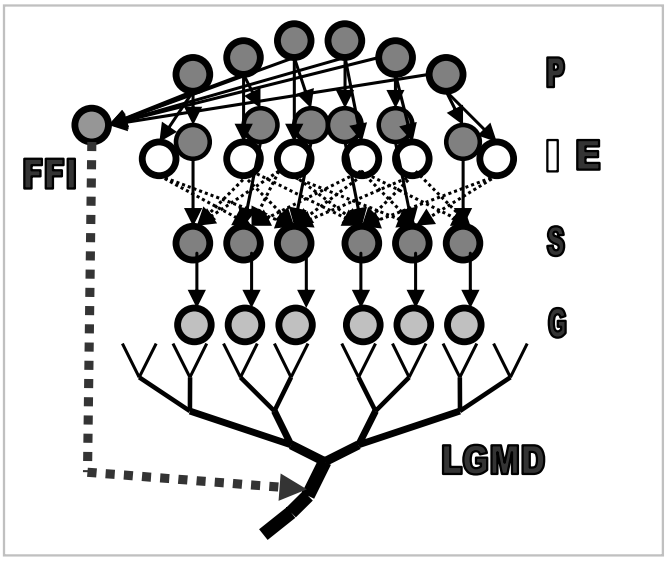


Fig.1 A schematic illustration of the LGMD based neural network for collision detection. There are five groups of cells and two single cells: photoreceptor cells (P); excitatory and inhibitory cells (E and I); summing cells (S); grouping cells (G); the LGMD cell and the feed forward inhibition cell (FFI). The input of the P cells is the luminance change. Lateral inhibition is indicated with dotted lines and has one frame delay. Excitation is indicated with black lines and has no delay. The FFI also has one frame delay. The input to FFI is luminance change from photoreceptor cells

A. P layer

The first layer of the neural network is the photoreceptor *P* cells arranged in a matrix; 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:

$$P_f(x,y) = \sum_i^{n_p} p_i P_{f-i}(x,y) + (L_f(x,y) - L_{f-1}(x,y)) \quad (1)$$

where $P_f(x,y)$ is the change of luminance corresponding to pixel (x,y) at frame f, x and y are the indices of the matrix, L_f and L_{f-1} are the luminance, subscript f denotes the current frame and $f-1$ denotes the previous frame, n_p defines the maximum number of frames (or time steps) the persistence of the luminance change can last, the persistence coefficient $p_i \in (0,1)$ and

$$p_i = (1 + e^{\mu i})^{-1} \quad (2)$$

where $\mu \in (+\infty, -\infty)$ and i indicates the previous i^{th} frame counted from the current frame f. Note that the LGMD neural network detects potential collision by responding expansion of the image edges, a strategy that needs computation rather than a strategy relying on object. If there is no difference between successive images, the *P* cells are not excited.

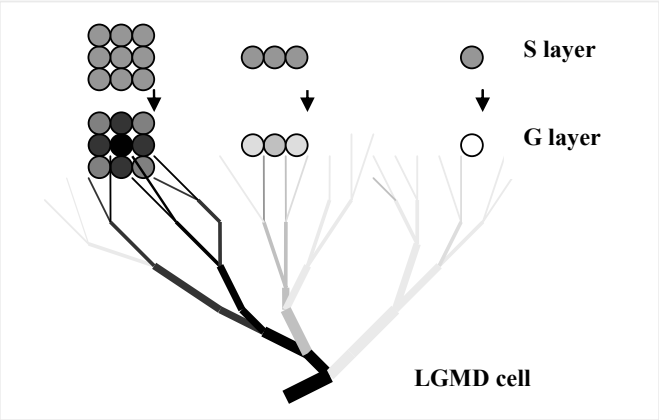


Fig.2 A schematic illustration of the grouped excitation processing mechanism. The cells(or pixels in an image) surrounded by strong excitations gain bigger passing coefficients and the isolated excitations get smaller passing coefficients and may be ruled out in the next layer G by threshold. Circles represent excitation in S and G layers. The strength of excitation in S layer, G layer and LGMD cell is indicated by grey levels where black represents the strongest excitation. S layer and G layer have one-to-one connection type.

B. I E layer

The output of the *P* cells forms the inputs to two separate cell types in the next layer. One type is called excitatory cells, through which excitation is passed directly to their retinotopic counterpart in the third layer, the *S* layer. The excitation *E(x, y)* in an *E* cell has the same value as that in the corresponding *P* cell. The second cell type are lateral inhibition cells, which pass inhibition, after 1 image frame delay, to their retinotopic counterpart's neighbouring cells in the *S* layer with one frame delay. The gathered strength of inhibition of a cell in this layer is given by:

$$I_f(x,y) = \sum_i \sum_j P_{f-1}(x+i, y+j) w_I(i,j), \ (if\ i = j,\ j \neq 0) \quad (3)$$

where $I_f(x,y)$ is the inhibition corresponds to pixel (x,y) at current frame f, $w_I(i, j)$ is the local inhibition weight. Please note i and j are not allowed to be equal to zero simultaneously. This means inhibition will only be allowed to spread out to its neighbouring cells in next layer rather than to its direct counterpart in the next layer.

C. S layer

The excitatory flow from the *E* cells and inhibition from the *I* cells is summed by the S cells using the following equation:

$$S_f(x,y) = E_f(x,y) - I_f(x,y) W_I \quad (4)$$

where W_I is the inhibition weight.

D. G layer

In the previous LGMD neural networks [27, 4] *S* cells connect directly with the LGMD and the LGMD sums input from all the *S* cells. In the LGMD based neural network, the expanded edges which are represented by clustered excitations, should be enhanced to extract colliding objects against complex backgrounds. Therefore, we have added a new layer, *G* cells, between the *S* cells and LGMD (Figure 1). A mechanism allows clusters of excitation in the *S* cells to easily reach the *G* layer and therefore provide a greater input to the membrane potential of the LGMD neuron compared with the input of a single *S* cell (as illustrated in Figure 2). To implement the new mechanism, the excitation in an *S* cell passed to the *G* layer is multiplied by a passing coefficient *Ce*. The coefficient is determined by the cell's surrounding neighbours, i.e., defined by a convolution process

$$Ce_f(x,y) = \sum_{i=-1}^{1} \sum_{j=-1}^{1} S_f(x+i, y+j) w_e(i,j) \quad (5)$$

where $w_e(i, j)$ represents the influence of its neighbours and this operation can be simplified as a convolution mask [7] and the passing coefficients can be computed in a matrix,

$$[w_e] = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (6a)$$

$$[Ce]_f = [S]_f \otimes [w_e] \quad (6b)$$

where $[w_e]$ is the convolution mask, $[C_e]_f$ is the passing coefficient matrix, $\otimes$ denotes the convolution operation and $[S]_f$ is the excitation matrix in the *S* layer.

When reaches the G layer, the excitation correspond to cell (x, y) becomes,

$$G_f(x,y) = (S_f(x,y)) Ce_f(x,y) \omega^{-1} \quad (7)$$

where ω is a scale and computed at every frame by

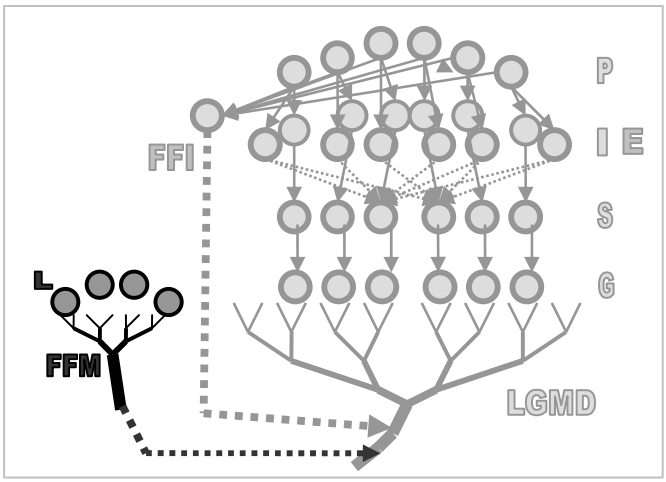


Fig.3 The feed forward mediation (FFM) cell is introduced to mediate LGMD cell's threshold when the images fed to the neural network are captured by a camera with white point calibration. The input to FFM is luminance through luminance cells L.

$$\omega = \Delta c + \max(abs[Ce]_f)C_w^{-1} \tag{8}$$

C_w is a constant, Δc is a small real number and $max(abs[Ce]_f)$ is the largest element in matrix $abs[Ce]_f$. As shown in Figure 2, the grouped excitations in the S layer (represent edges) become darker (stronger) when reaching the G layer and the isolated excitations become lighter (weaker). However, the previous unexcited cells will remain unexcited after the grouped excitation processing as indicated in equation (7).

From S layer to G layer, we set a threshold to filter decayed excitations,

$$\widetilde{G}_f(x,y) = \begin{cases} G_f(x,y) & if\ G_f(x,y)C_{de} \geq T_{de} \\ 0 & if\ G_f(x,y)C_{de} < T_{de} \end{cases} \tag{9}$$

where C_{de} is the decay coefficient and $C_{de} \in (0,1)$, T_{de} is the decay threshold. The presented grouped excitation (G) processing via equation (7) together with decay (D) processing via equation (9) can not only enhance the edges, but also filter out background detail caused excitations. The LGMD based neural network with grouped excitation and decay (GD) processing will be used and compared in the later experiments.

12.5%	25%	12.5%
25%		25%
12.5%	25%	12.5%

Fig.4 The local weights of inhibition spreading from the centre cell (in the IE layer) to neighbouring cells (in the S layer). The number in each cell represents the percentage of value it gains from the central pixel

E. LGMD cell

The membrane potential of the LGMD cell K_f at frame f is summed after G layer with a rectifying operation, which will turn the responses in negative values to positive before summing, as described by the following equation,

$$\mathrm{K}_f = \sum_x \sum_y abs(\widetilde{G}_f(x,y)) \tag{10}$$

The membrane potential of the $LGMD$ cell K_f is then transformed to sigmoid function as,

$$\kappa_f = (1 + e^{-\mathrm{K}_f n_{cell}^{-1}})^{-1} \tag{11}$$

where n_{cell} is the total number of the cells in G layer. Since K_f is greater than zero according to equation (10), the sigmoid membrane potential $\kappa_f \in (0.5\sim1)$.

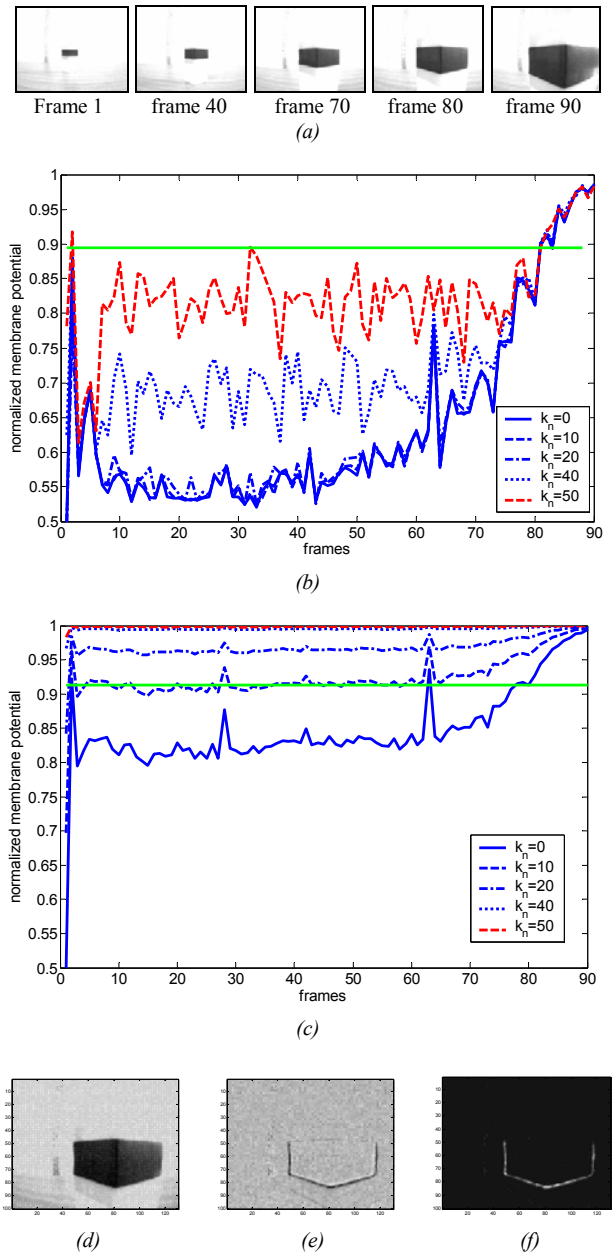


Fig.5 Off-line comparison when approaching an objecct. (a), the sample frames of the video clip to test the network. The video clip was taking with a robot approaching the block at speed 5.6cm/s and 24 frames per second. The collision happened at frame 91. (b), the sigmoid membrane potential at different noise level, with GD processing; collsions were detected at frame 84 when noise level is less than or equal 50. The straight horizontal line around 0.89 membrane potential is the threshold. (c), the sigmoid membrane potential at different noise level, without GD processing; will give false collision signal if noise slightly excedes 10. The thresholds have been reset to ensure both networks detect collision at frame 84, when no image noise presented. (d), the image with added noise at frame 84 with the presented noise level kn=50. (e), the image before GD processing at frame 84. (f), the image after GD processing.

F. FFM cell

The collision detection system consists of the LGMD based neural network and a CCD camera, which feeds the input

images to the network. The cameras of some robots have the ability to maintain a balanced image contrast (e.g., auto iris, K-team, Lausanne, Switzerland, http://www.k-team.com). Usually, a standard procedure (white point calibration) is used in the camera's hardware to normalize contrast within the image. This ability allows the robot to see objects under both dark and bright conditions and could be very important for feeding proper images to the LGMD based neural network when the robot is operating in dark or bright environments. However, this mechanism has a major shortcoming, for example, objects may have little contrast against the background if very bright objects or light sources get into the view field. The network may respond to the colliding objects too late in this case or too early in an opposite scenario [39]. To compensate for the drop of excitation in this situation, an adjustable threshold is needed. When integrated with a robot, threshold T_s can be used,

$$T_s = \alpha_{lt} T_{lt} + \alpha_{mp} T_{mp} \tag{12}$$

where T_{lt} is the adaptable part, T_{mp} is the constant part, α_{lt} and α_{mp} are the coefficients. $\alpha_{lt} > 0$ when integrated with the robot

TABLE 1. THE PARAMETERS OF THE LGMD BASED NEURAL NETWORK

name	value	name	value	name	value
p	0	α_L	1	n_{sp}	4
W_I	0.3	ΔT_{lt}	0.03	n_{ts}	4
C_{de}	0.5	Π_{lt}^{u}	230	$\alpha_{f\text{-}1}^{F}$	0
T_{de}	15	Π_{lt}^{l}	180	α_{ffi}	0.02
n_{cell}	13,000	T_{lto}	0	T_{FO}	10
α_{mp}	1	n_r	100	C_w	4
T_{mp}	0.86	n_c	130	Δc	0.01
n_p	1				

(the fed images are light compensated); otherwise, $\alpha_{lt} = 0$.

The feed forward mediation (FFM) (Figure 3) is introduced to adapt the threshold in response to low contrast colliding objects in situations when most parts of the input image are dim,

$$T_{lt} = T_{lto} + \alpha_L \begin{cases} \Delta T_{lt} & if \ L_d > \Pi_{lt}^{u} \\ 0 & otherwise \\ -\Delta T_{lt} & if \ L_d < \Pi_{lt}^{l} \end{cases} \tag{13}$$

where T_{ito} is the initial value of the T_{it}, α_L is a coefficient, ΔT_{it} is a constant, Π_{it}^{u} and Π_{it}^{l} are the upper and lower boundary respectively, the L_d is gathered as,

$$L_d = \left(\sum_{i=1}^{n_r} [\max L]_f\right) + \sum_{i=1}^{n_c} [\max L^T]_f \Big)(n_c + n_r)^{-1} \tag{14}$$

where n_c and n_r are the number of columns and rows in the luminance matrix $[L]_f$, $[maxL]_f$ is the row of the largest elements in each column of $[L]_f$ and $[maxL^T]_f$ is the row of the largest elements in each column of $[L]_f^T$. With the above mechanism, the LGMD based neural network is more likely to respond to colliding objects since threshold can be lowered.

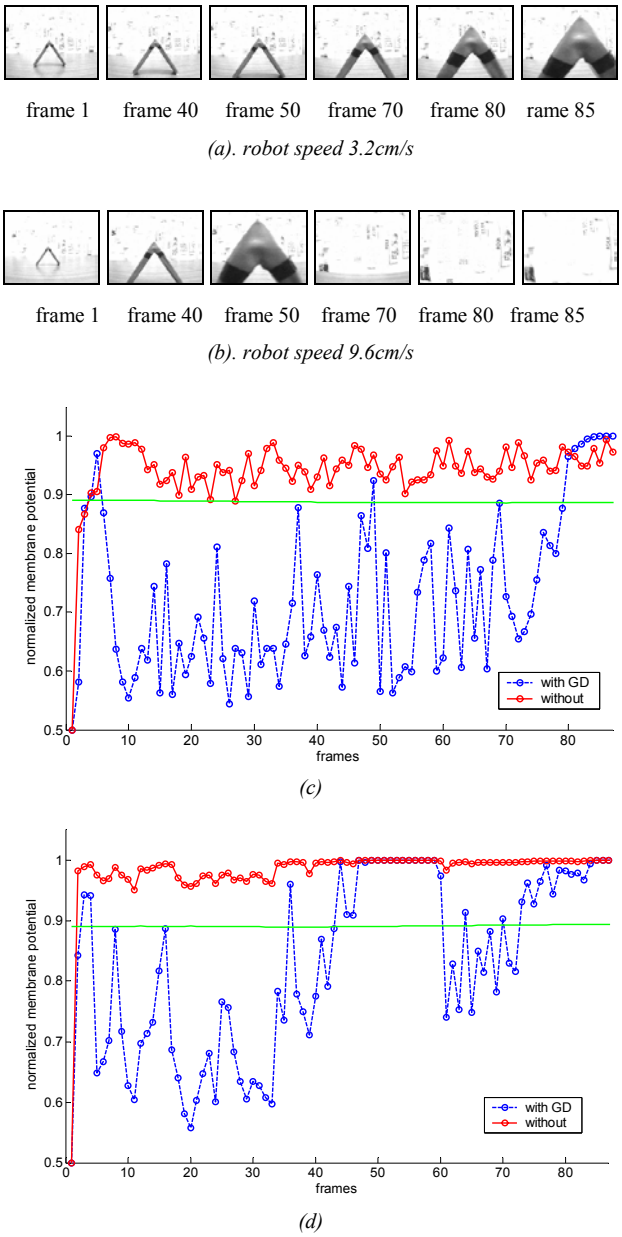


Fig.6 Comparison of the networks with/without GD processing at different approaching speeds. (a), the sample frames of the video clips, which was taking by a robot's camera when the robot was approaching an inverted V shaped object at speed 3.2cm/s) at 24 frames per second. Collisions happened at frame 87. (b), the sample frames of the video clips with robot speed at 9.6cm/s. Collision happened at frame 51. (c), the response of the membrane potential to the 3.2cm/s video. Collision detected at frame 83 with GD processing. (d), the response of the membrane potential to the 9.6cm/s video. Collision detected at frame 47 with GD processing. The horizontal straight line around 0.89 membrane potential value is the threshold for network with GD processing.

G. Spiking mechanism

The collision alarm is finally decided by a spiking mechanism. If the membrane potential κ_f exceeds the threshold T_s, a spike is produced,

$$S_f^{spike} = \begin{cases} 1 & if\ \kappa_f >= T_s \\ 0 & otherwise \end{cases} \quad (15)$$

where 1 represents a spike, 0 means no spike. A collision is detected if there are n_{sp} spikes in n_{ts} time steps (n_{sp}.<= n_{ts}) [35], i.e.,

$$C_{final} = \begin{cases} TRUE & if\ \sum_{f-n_{ts}}^{f} S_f^{spike} >= n_{sp} \\ FALSE & otherwise \end{cases} \quad (16)$$

where the value of C_{final} turns to be TRUE when collision is detected. The robot's avoidance behaviour is initiated once collision is detected. However, spikes may be suppressed by the FFI cell when the robot is turning.

H. The feed forward inhibition (FFI) cell

If it is not suppressed during turning, the network may produce spikes and even false collision alerts because of the sudden change in the visual scene. The feed forward inhibition and lateral inhibition work together to cope with such whole field movement [32]. The FFI excitation at current frame is gathered from the photoreceptor cells with one frame delay,

$$F_f = \sum_{j}^{n_a} \alpha_{f-j}^{F} F_{f-j} + \sum_{x=1}^{n_r} \sum_{y}^{n_c} abs(P_{f-1}(x,y)) n_{cell}^{-1} \quad (17)$$

where α^{F}_{f-j} is the persistence coefficient for FFI and $\alpha^{F}_{f-j} \in (0,1)$, n_a defines how many time steps the persistence can last.

Once F_f exceeds its threshold T_{FFI}, spikes in the LGMD are inhibited immediately. The threshold T_{FFI} is also adaptable,

$$T_{FFI} = T_{FO} + \alpha_{ffi} T_{FFI_{f-1}} \quad (18)$$

where T_{FO} is the initial value of the T_{FFI} , the adaptable threshold is decided by the previous T_{FFI} and α_{ffi} is a coefficient.

As described in the above subsections, the LGMD based collision detection system only involves low level image processing, such as excitation transferring and neighbouring operation; computationally expensive methods, such as object recognition or scene analysis, are not used. Because of this, the collision detection system is able to work in real time and is independent of object classification.

III. Test the Collision Detection System

Two kinds of experiments will be carried out to test the feasibility and robustness of the above collision detection system. One is the off-line test which mainly tests the effects of the added further excitation processing mechanism against complex backgrounds using recorded video clips. Then, the neural network is integrated with a Khepera robot (K-Team, Switzerland, http://www.k-team.com) to be tested in real time experiments.

A. Parameters and system setting

To obtain images, a K2D video turret (K-Team, Switzerland) with a CCD was mounted on top of the Khepera robot. The main properties of the camera are detailed in K2D video turret user manual, K-Team, Switzerland. We used the CCD camera, to sample images in real time. The camera was working at 25 frames per second in the experiment.

Parameters of the LGMD based collision detection system were set before the experiments. The input video images were 130 (in horizontal) by 100 (in vertical) pixels; images were grey scale ranging from 0 to 255 (parameter without unit, similar parameters hereafter will not be restated). The lateral inhibition spreads to its neighbour 1 layer away and with one frame delay. The local inhibition weight is set as shown in Figure 4. Other parameters are listed in Table 1. These parameters are tuned manually based on the early pilot experiments and will not be changed in the following experiments unless stated.

The LGMD based collision detection system is written in Matlab (the MathWorks, Inc., USA). The computer used in the experiments is a PC (Dell Precision 450) with one 2.40GHz CPU and Windows XP operating system. The communication between the computer and the robot is via serial port with Baud rate at 9600 bits/s. A USB frame grabber (Hauppauge Computer Works Ltd. UK) and video device access software Video for Matlab (VFM) (University of East Anglia, UK) are used to obtain live image input.

B. Off-line tests

The results of off-line tests can be fairly compared since the network can be challenged with the same visual images repeatedly. In the following off-line tests, we use recorded video clips to test and compare the collision detection systems.

1) Tests under simulated background

Background details but not colliding objects sometimes cause excitations. Since the level of a computer simulated background can be easily added into images and controlled in the experiment, we will use computer generated random dots to simulate the background in the off-line tests.

A video clip with 90 frames was recorded using a robot approaching a block at speed 5.6cm/s and frame rate at 24Hz,

example frames are shown in Figure 5 (a). Isolated excitations are generated using random values added to the input images. For example, at frame *f*, the random values are

$$L_f^n(x,y)=\begin{cases}k_n\, rand(1), & (x=1,3,5,\cdots 100)\ AND\ (y=1,3,5,\cdots,130)\\ 0, & otherwise\end{cases} \quad (19)$$

where κ_n is a scale representing the level of background caused image noise (shorted to noise hereafter) and will be set from 0 to 60 units with increase step 10, function *rand(1)* will generate a uniform distribution value between (0,1). Noises at different pixels and the successive frames are independent. An element at (x,y) in the input image to the LGMD neural network will be,

$$\widetilde{L}_f^{\,n}(x,y)=L_f(x,y)+L_f^n(x,y) \quad (20)$$

The averaged excitation (averaged from 100 frames) in each frame distributed to each isolated *L* cell (or pixel) caused by these simulated noise is around (3.32, 6.60, 9.88, 13.20, 16.52, 19.80) for noise levels of (10, 20, 30, 40, 50, 60) respectively.

The LGMD based neural networks were challenged with the video clip and results are shown and compared in Figure 5. The network with GD processing worked well when κ_n was less or equal to 50; the collisions were consistently detected at frame 84 (Figure 5, b). The curves climb up sharply when collision is imminent. As a comparison, a similar network but without GD processing was also tested with the same video sequence and the results shown in Figure 5 c; the threshold has been reset to ensure both networks detect collision at frame 84, when no image noise presented. However, with this threshold the network without GD processing will inevitably fail when noise, i.e. complex background, is presented as shown in the figure. The GD processing has significantly

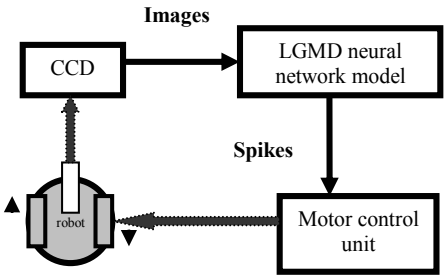


Fig.7 A diagram of the Khepera robot's (K-team, Switzerland) visual collision detection system and its connection with the motor control unit. The images are captured by a CCD camera mounted on the robot and are fed to the LGMD based neural network. Spikes are generated by the neural network and are passed to a motor control unit which has only one pattern of reaction. The motor control unit can be activated by four successive spikes. It controls the two wheels of the Khepera robot and makes the robot turn in one direction, clockwise. The turning angles are generated randomly within lower and upper limits. In the real time experiments, further background of the environment was a typical lab not excluded.

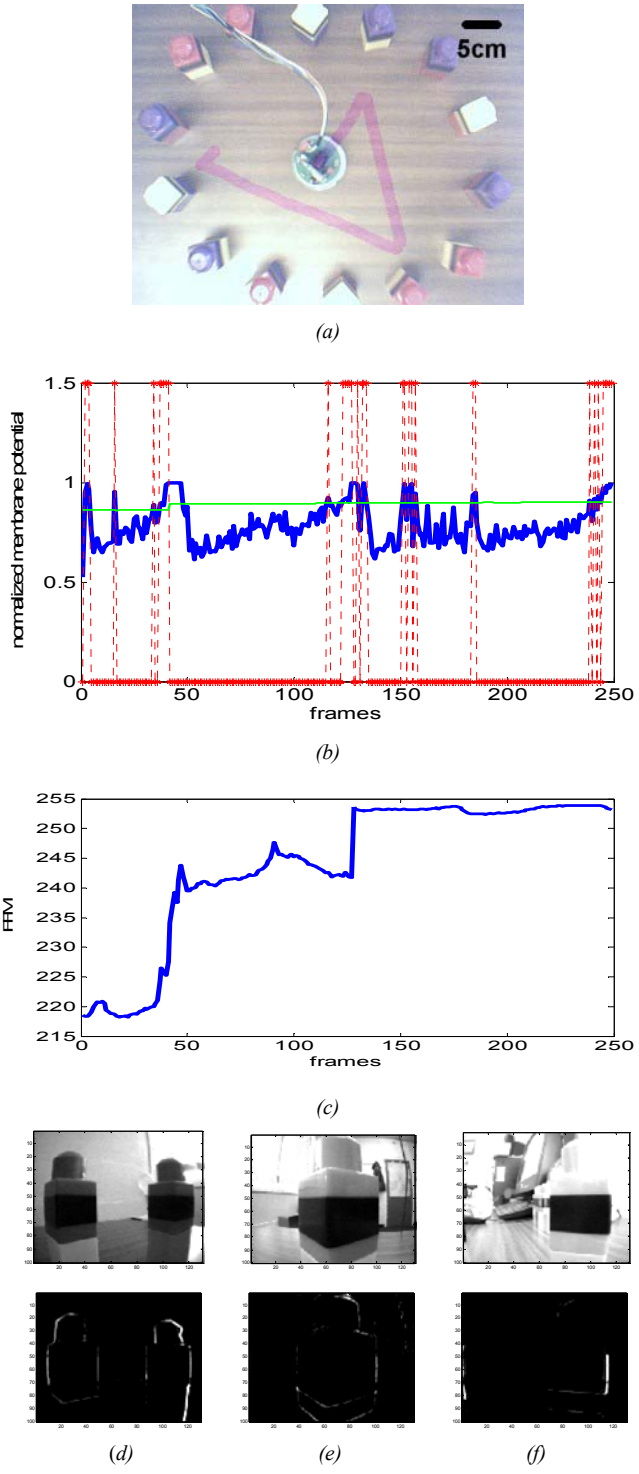


Fig.8 The robot has moved for 15 seconds within an arena at speed 4.8cm/s. (a) The robot trajectory within the arena; the trajectory is indicated by bold lines. Three times of collision were detected by the LGMD based collision detection system and three turns were conducted. (b) The LGMD sigmoid membrane potential (bold red solid line), spikes (dashed line with stars at the peaks) and the threshold (solid horizontal straight line). The three collisions detected were indicated with 4 successive spikes. (c).The FFM during the movement. It reflects the changed contrast intensity after each turn. (d). The scene of the first turn and the image after GD processing. (e) The scene of the second turn and the image after GD processing. (f) The scene of the third turn and the image after GD processing.

improved the network's robustness in these cases.

2) Test under real background generated image noise

The network was also challenged with clips captured by robot approaching an inverted V shaped object before a complex background (Figure 6 a, b) at different speeds. The V shaped object was placed on a table 50cm away from the wall which was decorated with textured slow drifting stripes of paper. The object is 7.5cm high. Sample frames of the two video clips are shown in Figure 6 (a) and (b), the approaching speeds were 3.2cm/s and 9.6cm/s respectively. Again, the networks with GD and without GD processing were compared. The results are shown in Figure 6 (c-d).

With GD processing in G layer, the neural network responded most differently to non-colliding scenes and imminent colliding objects; it detected the colliding object at different approaching speeds, i.e., at frame 83 for the lower speed video and at frame 47 for the high speed video. However, without GD processing, the network was unable to detect collision at all as shown in the Figure 6 (c) and (d) since there is no room to set a proper threshold. These off-line tests may suggest that with edge enhancement, the network can detect imminent collision robustly, especially on an open, complex background.

C. Real Time Tests

In the off-line tests, the objects and scenes were known in advance or remaining unchanged for each new test. Moving autonomously within arenas might cause the robot to face new and unpredictable situations. To test if the collision detection system works reliably, the best way is to challenge it with real world situations in real time.

1) Environmental set up

In the real time experiments, the LGMD based neural network (with GD processing if not indicated otherwise in the following parts) together with FFM is integrated with a Khepera robot (K-team, Switzerland) (Figure 7). The robot is controlled by a motor control unit, which can be triggered by several (4 in this study) successive spikes from the LGMD cell and outputs two commands to the left and right wheels to control the robot's turning behaviour. The luminance intensity is deliberately not controlled; however, the light from above the arenas was measured in the experiments from 86μw/cm^2 to 130μw/cm^2 if not stated differently.

As mentioned in the introduction, previous experiment [4] demonstrated the ability of a LGMD neural network to avoid collision in an arena with a simple background. Since the presented LGMD based collision detection system is aiming to tackle challenges in a more complex situation, all the arenas used in the experiments will have complex backgrounds.

In the autonomous navigating experiments, the robot was allowed to move at a speed within an arena for a period of time, for example, 15 seconds. Once it detected an imminent collision, it stopped and turned before continuing its straight line path. The turning speed is set to be 3.2cm/s for the left wheel and -3.2cm/s for the right. The turning angle was controlled by time t_α and

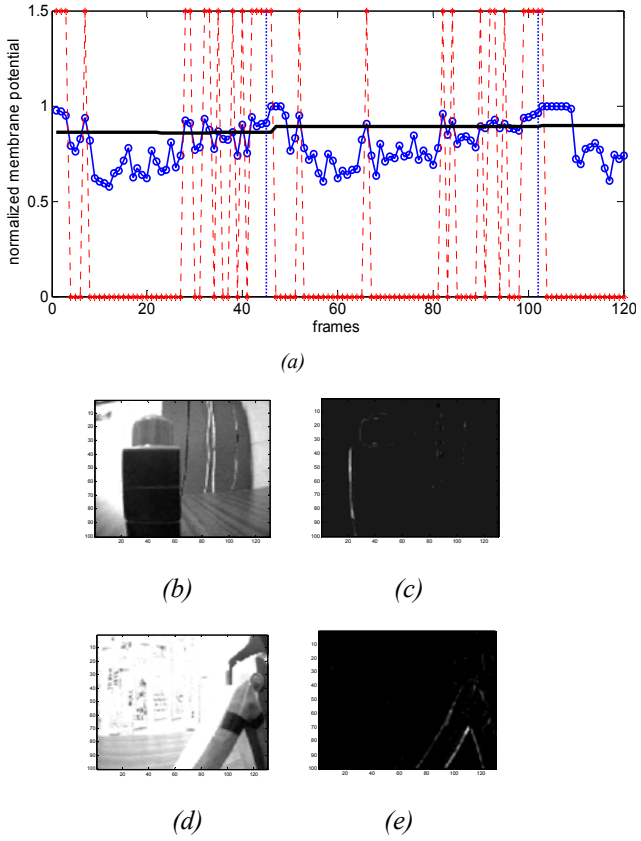


Fig.9 (a), the effects of the FFM regulated threshold. The threshold is in solid straight line, sigmoid membrane potential is indicated by circles for each frame and spikes represented by upper stars. Collisions were detected at frame 45 and frame 102, and indicated by dotted vertical lines. (b), the image at frame 45. (c), the image after GD processing at frame 45. (d), the image at frame 102. (e), the image after GD processing at frame 102. In the images, white represented the highest value and black represented the lowest.

$$t_\alpha = t_{con} + \lambda\, rand(1) \tag{21}$$

where t_{con} is a constant and set to 0.7 second in the experiment, λ is a scale and set to 0.25 and *rand(1)* is a function to generate uniform random numbers between 0 to 1. Therefore, the period of time for the robot to turn is between 0.7~0.95 seconds.

In the following experiments, unless otherwise stated, the visual input to the LGMD neural network will be shut during turning; the LGMD based neural network will be the one with GD processing. The trajectories of the robot are recorded with a webcam (Trust 380 USB 2.0 SPACEC@M, http://www.trust.com) hanging above the arena and are

extracted via an off-line trace extracting programme written in Matlab. In the trace extracting programme, template matching method [9] is used to locate the robot position in each frame with updated templates and calibrating points.

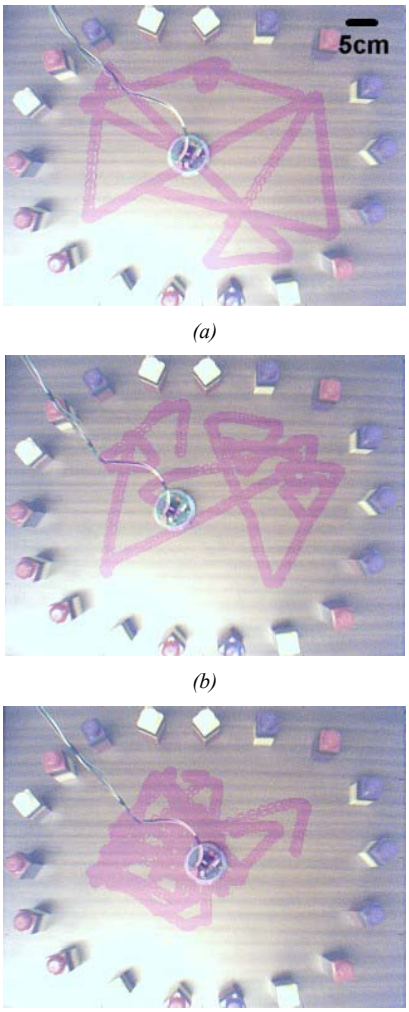


Fig.10 The robot moves autonomously within an arena at different speeds for 60 seconds. (a) robot speed is 6.4cm/s; (b) robot speed is 9.6cm/s; (c) robot speed is 12.8cm/s.

2) Experiments and results

A short robot movement (15 seconds, speed at 4.8cm/s) was conducted to test and show the mechanism of the collision detection system. Results are shown in Figure 8. Three imminent collisions were successfully detected (Figure 8 a) at frame no.44, no.127 and no.250 respectively (Figure 8, b). The membrane potential threshold jumped to a higher level when the robot turned and faced a new scene because of the change in FFM (Figure 8, b). The three scenes and processed images, just when the collisions were detected, were also shown in the figure. The stripes on the table, wires and other small objects were filtered out by the GD processing. Only the expanding edges, the feature which is used by the detection system, remained. This made the neural network more robust as it only concentrated on colliding objects.

The FFM sometimes can affect the detection moment in several frames, as shown in the results (Figure 9) from an experiment in which two collisions were detected in time; the collision avoidance behaviour would be either several frames later for the first colliding object or several frames earlier for the second one if without the mediation by FFM.

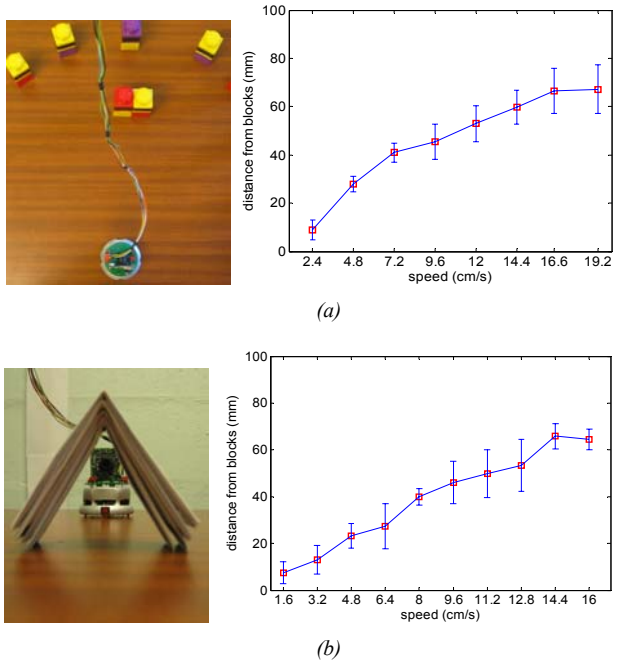
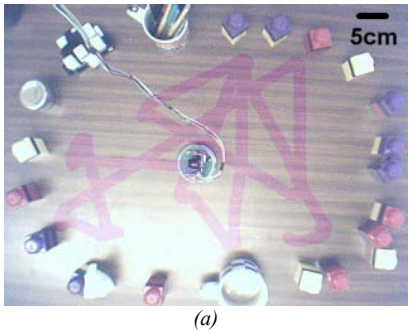


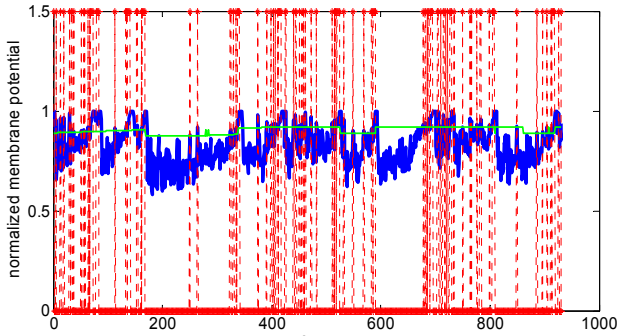
Fig.11 The detection distance versus robot speeds. (a), the robot approaching two blocks with several blocks scattered on the background; the data is collected with robot's speed changed in the order of, 2.4-4.8-7.2-9.6-12-14.4-16.6-19.2, 19.2-16.6-14.4-12-9.6-7.2-4.8-2.4, 2.4-4.8… till 5 groups of data were collected. (b), the robot approaching an A shaped object with different speeds; data is collect in the same way as that in (a). Data (mean and standard deviation) at each speed level is averaged from five times of the experiments. Robot was placed on the mark 25cm away and directed to the colliding objects before a trial started. The moving direction of the robot was calibrated at the start position for each trial using a TV monitor. The detection distance refers to distance between robot and the colliding objects.

Further experiments were carried out to show how the robot behaved within an arena at different speeds (Figure 10). With the speed of 6.4cm/s, the robot sometimes was quite close to the object before it detected collision and turned (Figure 10, a) as less excitation occurred due to smaller changes in successive images. It was found that at higher speeds (9.6cm/s and 12.8cm/s) the robot increased its distance from the blocks (Figure 10, b, c). This was because more excitation was caused by higher speeds as the change in image between successive frames was larger.

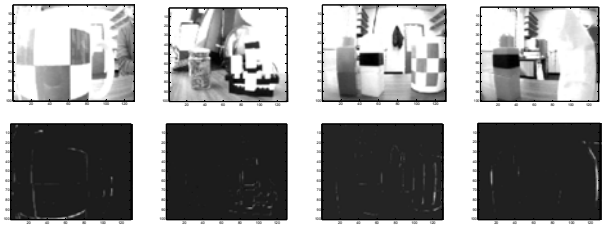
The collision detection distance was found to increase reliably with the increase in approach speed in two



(a)



(b)



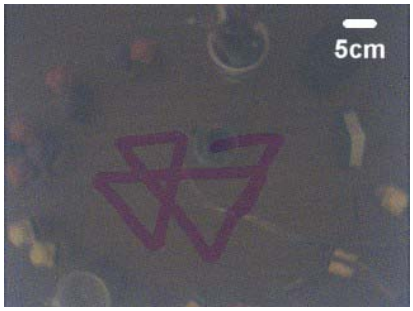
(c) (d) (e) (f)

Fig.12 The robot has moved for 60 seconds within an arena surrounded by different (by shape, colour and material) objects at speed 6.4cm/s. (a) The robot trajectory within the arena. The trajectory is indicated by bold lines. 12 times of collision were detected by the LGMD based collision detection system and 12 turns were conducted. (b) The LGMD sigmoid membrane potential (bold solid line), spikes (upper stars) and the threshold (solid straight line). (c)~(f). Some of the scenes when imminent collisions were detected and their corresponded images after GD processing.

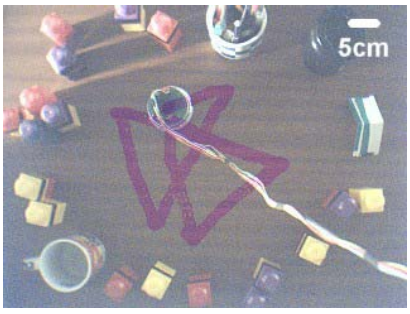
experiments that have been done with different colliding objects, as shown in Figure 11. The collision detection distance here refers to the distance between the robot and the colliding object when an imminent collision is detected. In the experiments, the robot was approaching the two blocks (Figure 11 a) and the inverted V shaped object (Figure 11 b) respectively with different speeds and stopped once an imminent collision was detected.

When some of the blocks were changed to other types of objects, the neural network also worked quite well; navigating a course about 60 seconds, 12 collisions were detected and avoided as indicated in the figures (Figure 12, a and b). Mugs, strange shaped blocks and curved paper were all successfully detected by the system (Figure 12, c~f).

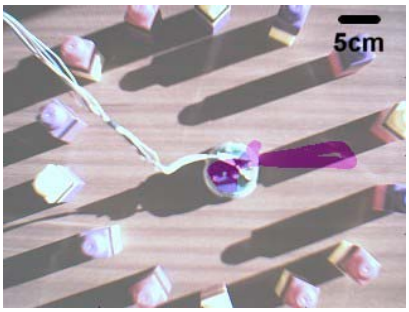
We want to see if the system works in extreme conditions. As shown in Figure 13, three experiments have been done to test the system in three scenarios: (1) very dark with luminance intensity at 17μw/cm^2 (Ealing Electro-Optics, Holliston, MA, USA), (2) partly in sunlight (917~1,274μw/cm^2), and (3) very bright sunlight (3,240μw/cm^2) with long shadows. The experiments showed that the robot can still detect collision without any difficulty in these situations. Interestingly, it detected collision very early



(a)



(b)



(c)

Fig.13 The performance of the robot with the LGMD based collision detection system in extreme conditions. (a) extremely dark (17μw/cm^2), robot speed is 6.4cm/s; (b) partly in mild sunlight (917~1,274μw/cm^2), robot speed is 6.4cm/s; (c) in bright sunlight (3,240μw/cm^2) with long shadows, robot speed is 4.8cm/s.

when facing long shadows (Figure 13, c) because the system detecting objects relying on the contrast.

D. Further discussions

In the above sections, the presented LGMD based collision detection system has been tested using both off-line and real time tests. GD processing enhances the key features of a colliding object. The collision detection system demonstrated a reliable ability to detect collision in different situations with complex backgrounds, regardless of the shape, material or colour of the colliding objects. The system allows the robot to navigate in an unstructured, complex environment without intensive computing cost.

For many species of animal, vision plays a key role in their survival. Different visual based navigation methods have also been proposed towards autonomous robots (for example, [22, 11]). Nowadays, visual sensors are becoming very cheap and reliable. This makes it possible for many mobile machines (e.g., mobile robots, cars, boats, planes and some toys) equipped with visual sensors and visual-based navigation systems to avoid unwanted collision automatically in the real world.

However, the locust LGMD is only one example of a visual feature detector that has evolved within the visual system of insects. In the insects' brain, there are numerous interneurons related to vision working together to extract the plentiful visual cues simultaneously. To separate/extract other visual cues from the dynamic scenes simultaneously, other specialized neurons need to be integrated into the system in the future. For example, directional selective neurons (e.g., [24, 25] in locust, [5] in fly, [36] in rabbit) may be used to detect high speed big translating objects which project big areas on the retina. The further ongoing investigations of the LGMD and its postsynaptic interneuron in locust also provide new ideas and alternative ways to further modify the system (for example, [13, 32, 33]).

Although the presented LGMD based collision detection system's robustness was demonstrated in the experiments, it still needs to be noted that the detection system is entirely reliant on an objects' contrast against background. The system does fail to detect a colliding object if the object has no contrast to its background as seen by the robot.

IV. Conclusions

In this paper, we proposed a modified LGMD based neural network as a real time robust collision detector with edge enhancement, especially for collision detection against complex backgrounds. The excitations (represents expanded, moving edges) extracted in the input images are further enhanced with GD processing. Experiments showed that the presented LGMD based neural network worked reliably in the situations with complex backgrounds. Integrated with a robot, the collision detection system demonstrated its reliable ability in detecting imminent collisions in wide range of robot speeds and situations; therefore enabling the robot to autonomously avoid collision within arenas with only visual input.

In the future, other neurons, directional selective neurons may be integrated into the system to extract other visual cues simultaneously from the same sequential images to cope with fast translating objects.

Acknowledgment

We thank the comments raised in the internal review on the manuscript. We also thank the anonymous reviewers for their valuable comments. We thank Mr. Mark Bendall for reading this paper.

References

[1] M. D. Adams, Sensor modelling, design and data processing for autonomous navigation. River Edge, NJ, World Scientific, 1998.

[2] G. Benet, F. Blanes, J.E. Simo, and P. Perez, "Using infrared sensors for distance measurement in mobile robots," Robotics and Autonomous Systems, vol.40, pp.255-266, 2002.

[3] M. Blanchard, P. F. M. J. Verschure, & F. C. Rind, "Using a mobile robot to study locust collision avoidance responses," International Journal of Neural Systems, vol. 9, pp.405-410, 1999.

[4] M. Blanchard, F.C. Rind and P.F.M.J. Verschure, "Collision avoidance using a model of the locust LGMD neuron," Robotics and Automonous Systems, vol.30, pp.17-38, 2000.

[5] A. Borst, and J. Haag, "Neural networks in the cockpit of the fly," J. Comp. Physiology, vol.188, pp.419-437, 2002.

[6] H. Buxton, "Learning and understanding dynamic scene activities: a review," Image and Vision Computing, vol.21, pp.125-136, 2003.

[7] E.R. Davies, Machine vision: theory, algorithms, practicalities, Academic Press, Inc., San Diego, USA, 1997.

[8] G.N. DeSouza and A.C. Kak, "Vision for mobile robot navigation: a survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24(2), pp.237-67, 2002.

[9] R.O. Duda and P.E. Hart, Pattern classification and scene analysis, John Willy & Sons Inc. New York, 1973.

[10] H.R. Everett, Sensors for Mobile Robots: Theory and Application. AK Peters, Wellesley, MA, 1995.

[11] M. Fiala and A. Basu, "Robot navigation using panoramic tracking," Pattern Recognition, vol.37, pp.2195-2215, 2004.

[12] N. Franceschini, "Visual guidance based on optic flow: a biorobotic approach," Journal of Physiology Paris, vol.98, pp281-292, 2004.

[13] F. Gabbiani, H.G. Krapp, C.Koch and G. Laurent, "Multiplicative computation in a visual neuron sensitive to looming," Nature, vol.420(21), pp.320-324, 2002.

[14] F. Gabbiani, H.G. Krapp, N. Hatsopoulos, C-H Mo, C. Koch and G. Laurent, "Multiplication and stimulus invariance in a looming-sensitive neuron," Journal of Physiology- Paris, vol.98, pp.19-34, 2004.

[15] R.R. Harrison and C. Koch, "A silicon implementation of the fly's optomotor control system," Neural Computation, vol.12, pp.2291-2304, 2000.

[16] S.A. Hubber, M.O. Franz and H.H. Buelthoff, "On robots and flies: modelling the visual orientatin behaviour of flies," Robotics and Autonomous Systems, vol.29, pp.227-242, 1999.

[17] 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.

[18] Khepera User Manual, version 1.1, K-team, Switzerland, March, 2002.

[19] F. Iida, "Biologically Inspired Visual Odometer for Navigation of a Flying Robot," Robotics and Autonomous Systems, vol.44/3-4, pp. 201-208, 2003.

[20] G. Indiveri, and R. Douglas, "Neuromorphic vision sensors," Science, vol.288, pp.1189-1190, 2000.

[21] R. Manduchi, A. Castano, A. Talukder and L. Matthies, "Obstacle detection and terrain classification for autonomous off-road navigation," Autonomous Robots, vol.18, pp.81-102, 2005.
[22] C.F. Olson, L.H. Matthies, M. Schoppers and M.W. Maimone, "Rover naviation using stereo ego-motion," Robotics and Autonomous Systems, vol.43, pp.215-229, 2003.
[23] M. O'Shea, C.H.F. Rowell, J.L.D. Williams, "The anatomy of a locust visual interneurone: The descending contralateral movement detector," Journal of Experimental Biology, vol.60, pp.1–12, 1974.
[24] F.C. Rind, "A directionally selective motion-detecting neurone in the brain of the locust: physiological and morphological characterization," J. Exp. Biol., vol.149, pp.1-19, 1990a.
[25] F.C. Rind, "Identification of directionally selective motion-detecting neurones in the locust lobula and their synaptic connections with an identified descending neurone," J. Exp. Biol. Vol.149, pp.21-43, 1990b.
[26] F.C. Rind, P.J. Simmons, "Orthopteran DCMD neuron: A reevaluation of responses to moving objects. I. Selective responses to approaching objects," Journal of Neurophysiology, vol.68, pp.1654–1666, 1992.
[27] F.C. Rind, and D.I. Bramwell, "Neural network based on the input organization of an identified neuron signaling impending collision," Journal of Neurophysiology, vol.75, pp.967– 985, 1996.
[28] F.C. Rind, P.J. Simmons, "Seeing what is coming: Building collision sensitive neurons," Trends in Neurosciences, vol.22, pp.215-220, 1999.
[29] F.C. Rind, Roger D. Santer, J. Mark Blanchard and Paul F.M.J. Verschure, "Locust's looming detectors for robot sensors," Sensors and Sensing in Biology and Engineering, FG Barth, JAC Humphrey, and TW Secomb (Eds.), Spinger-Verlag, Wien, New York, 2003.
[30] C.H.F. Rowell, M. O'Shea and J.L.D. Williams, "The neuronal basis of a sensory analyser, the acridid movement detector system," J. Exp. Biol., vol.68, pp.157-185, 1977.
[31] A. Rosenfeld, "From image analysis to computer vision: an annotated bibliography, 1955-1979," Computer Vision and Image Understanding, vol.84, 298-324, 2001.
[32] 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.R. Soc. Lond. Interface, vol.1, pp.61-77, 2004.
[33] R. D. Santer, P. J. Simmons, F. C. Rind, "Gliding behaviour elicited by lateral looming stimuli in flying locusts,'' J Comp Physiol A, vol.191, pp.61-73, 2005.
[34] G.R. Schlotterer, "Response of the locust descending contralateral movement detector neuron to rapidly approaching and withdrawing visual stimuli," Canadian Journal of Zoology, vol.55, pp.1372–1376, 1977.
[35] Tecnical Report D11: Biological Model Report, Project IST-2001-38097, LOCUST: Life-Like Object Detection for Collision-Avoidance Using Spatio-Temporal Image Processing, February 2004. . http://www.imse.cnm.es/locust/main.html
[36] S.F. Stasheff, and R.H. Masland, "Functional inhibition in direction-selective retinal ganglion cells: spatiotemporal extent and intralaminar interactions," J. Neurophysiology, vol.88, pp.1026-1039, 2002.
[37] B. Webb, and R. Reeve, "Reafferent or redundant: integration of phohotaxis and optomotor behaviour in crickets and robots," Adaptive behaviour, vol.11 (3), pp.137-158. 2003.
[38] G. Wichert, "Can robots learn to see," Control Engineering Practice, vol.7, pp.783-795, 1999.
[39] S. Yue and F.C. Rind, "A Collision detection system for a mobile robot inspired by locust visual system," IEEE Int. Conference on Robotics and Automation, Spain, Barcelona, Apr.18-22, pp.2843-3848, 2005.

Shigang YUE (M'05) received the B.E. degree from Qingdao Technological University in 1988, received the M.E. and Ph.D. degree from Beijing University of Technology (BUT), Beijing, China, in 1993 and 1996 respectively.

He is currently with the School of Biology and Psychology, University of Newcastle upon Tyne, UK. He worked at University College London (2002-2003) and University of Kaiserslautern, Germany as an Alexander von Humboldt Research Fellow (2000-2001). He worked at the City University of Hong Kong (1998-1999) as a Senior Research Assistant and BUT (1996-1999) as a Lecturer/Associate Professor. His current research interests include biological/artificial vision systems and applications, computer vision, neural information processing, machine learning, neural systems co-evolution, robotics, and computational intelligence.

F. Claire RIND was awarded a BSc from the University of Canterbury, Christchurch New Zealand and received a PhD in Zoology from Cambridge University in 1982.

She is currently a Reader in Invertebrate Neuroscience in the School of Biology, University of Newcastle, Newcastle upon Tyne, UK. Previously she held a Royal Society University Reasearch Fellowship and a BBSRC Advanced Reasearch Fellowship. Her current reasearch interests include sensory processing by the insect brain, neuronal pathways for collision avoidance in locusts, bio-inspired robotics and ASICS for visual tasks. Dr Rind is a member of the Society for Experimental Biology, The Physiological Society and the International Society for Neuroethology