An Insect Inspired Visual Sensor for the Autonomous Navigation of a Mobile Robot

Sjoerd van der Zwaan and José Santos-Victor*

Instituto de Sistemas e Robótica,
Instituto Superior Técnico,
Av. Rovisco Pais,
1049-001 Lisboa,
Portugal.

Abstract. This paper deals with the autonomous navigation of a mobile robot equipped with an insect-like visual sensor. The design of the sensor is based on models derived from the insect compound eyes that use optical flow information for navigation purposes. Restricting navigation to the ground-plane, we design a 1D panoramic sensor that performs parallel image processing and calculates local components of the optical flow. This information then is used to detect obstacles in the robot azimuth-plane so as to periodically update an obstacle map. With this map, we use a potential field approach to drive an obstacle avoidance algorithm. Closing the visual navigation loop then reveals a reflex-like behaviour. Results are described and discussed.

1 Introduction

The observation of some flying insects like bees or flies shows their amazing capability to navigate in a three dimensional space at a considerable speed. The insect's compound eyes, shown in Figure 1, play an important role in navigation and can be seen as the sensory input to a neural network that processes the images so as to generate wing control. Considering the dimensions of the insect's head, we may wonder how such good visual navigation can be embedded in such a small brain. Obviously, some efficient signal processing must take place and also the compound eye's lateral positions and pixel distribution seem to play an important role.

Fig. 1.: The two compound eyes of a bee

Inspired by the performance of some insects, recent work has been done so as to obtain insect-like behaviours in mobile robotics. Analysis of the housefly's visual system is done by Franceschini et. al. [1, 2, 8]. Horridge [4] has illustrated the principles of some bee-like behaviours and Santos-Victor [9] and Srinivasan [6]

* Email:sjoerd@isr.ist.utl.pt, jasv@isr.ist.utl.pt

showed how these behaviours can be embedded into mobile robots. From all the work, a common conclusion can be drawn: insect vision makes use of optical flow computation for navigation purposes.

At the front of the neural network of some insects, two compound eyes can be found that are considered to be the most important parts of their visual system. The visual system of the housefly has been well studied [1] and its relevant characteristics will be resumed here. The surface of the housefly's compound eye consists of a two-dimensional array containing about 3000 distributed pixels (ommatidia). On its retina, 24000 photosensors are defined so that every pixel contains a set of eight photosensors. The photosensors have coaxial receptive fields specified by their optics and sample the visual environment in a nonuniform way, with an acute zone of vision in the front part of the eye. The electrical outputs of the photosensors are read in parallel by a neural network consisting of about one million neurones. These neurones process the signals of the photosensors in different ways, depending on the function they serve. This way, the neural network realises a parallel and asynchronous signal processing to control the muscles that activate on the fly's wings.

An essential feature of this signal processing is that it computes components of the optical flow. The ability of humans, animals and insects to evaluate their distance relative to features in the environment by making use of ego-motion, was first introduced by Helmholtz and is called Motion Parallax. The principle of Motion Parallax relates angular velocity of projections on the spherical retina (similar geometry in [7]) with radial distance to the perceived features in the visual field, as illustrated in Figure 2. Under the assumption of pure translational locomotion, this relation is given by:

$$\Omega = \frac{V_0}{D} \cdot sin\varphi \quad [rad/s], \tag{1}$$

where Ω is the angular velocity of the feature on the retina, V_0 denotes the translational velocity of the observer, D is the radial distance to the feature and φ represents the angle of observation. The perceived optical flow can be obtained by multiplying the feature angular velocity by the compound eye radius, r_{eye}.

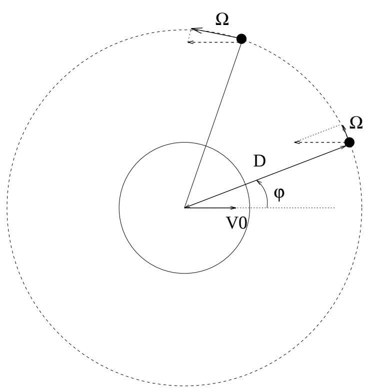

Fig. 2.: Motion Parallax: the angular velocity of a feature point as perceived on the spherical retina.

Just as stereo vision can be used to obtain metric measurements from images, the principle of Motion Parallax can be used to perceive radial distance to an observed point or to derive properties like time-to-crash or travelled distance. In insect vision, local signal processing neurones called Elementary Motion Detectors are believed to compute local components of the optical flow in such a way as to extract this kind of information.

This paper describes the design and implementation of an insect-like visual sensor for mobile robot navigation. The following Section 2 introduces the insect inspired visual sensor composed of an array of motion detectors. Section 3 presents a reflex-like navigation algorithm for a mobile robot, inspired by insect behaviours. Section 4 briefly looks at its hardware implementation and Section 5 shows some obtained results in simulation and real-time experiments. Finally, some conclusions are made and future work is indicated.

2 Panoramic Insect-like Visual Sensor

With the intention to design an artificial compound eye for mobile robot navigation, we have considered a horizontal section of the insect's two lateral compound eyes (Figure 1). This results in a disc that scans the world in the azimuth-plane with a 1D photosensor array. Defining a 1D panoramic sensor is sufficient for navigation purposes restricted to the ground-plane. The sensor's artificial retina is composed of an array of photosensors with local motion detectors realising parallel asynchronous image processing.

2.1 Elementary Motion Detector

An Elementary Motion Detector (EMD) computes the local component of the optical flow caused by a contrasted feature travelling in the visual field, using information of two neighbouring pixels. The signal processing within the EMD is based on the functional diagrams of the housefly's motion detecting neurones as described in [2, 8].

Upon translational locomotion, the local component of the optic flow is defined as the angular velocity of the projection of contrasted features on the local two-pixel retina, occurring in opposite direction to the locomotion:

$$\Omega = \frac{\Delta\varphi}{\Delta t} \quad [rad/s], \tag{2}$$

where Ω is the local component of the optical flow, $\Delta\varphi$ denotes the angle between the two motion detecting pixels and Δt represents the time interval of detection between the two pixels.

With the configuration given in Figure 3 and knowing the angles between the photosensors, measuring Δt is sufficient for an EMD to calculate the local component of the optical flow.

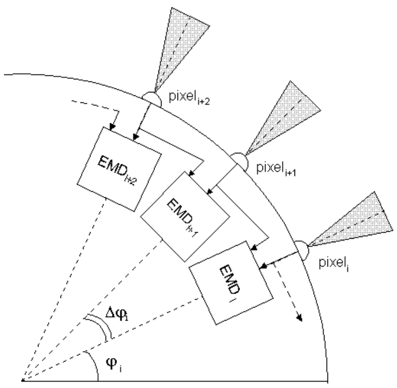

Fig. 3.: Organisation of the array of Elementary Motion Detectors with corresponding input pixels

The signal processing in each EMD consists of spatio-temporal filtering on its photosensor input-channels and can be divided into three stages: spatial filtering, temporal filtering and optical flow computation.

- Spatial filtering is performed at each EMD input channel, so as to perform adaptation to local light intensity levels. This adaptation can be found in all living eyes for which biology has many solutions.
- Temporal filtering performs a temporal edge detection on each channel, which emphasises fluctuations in pixel intensity over time. A threshold operation transforms the image into binary values over time and defines the EMD's sensitivity for edge detection.
- The impulses generated by the edge detection in each EMD channel then serve to measure the time Δt between edge detection in the two neighbouring pixels. Measuring this time is equivalent to knowing the local optical flow (see equation (2)).

The signal processing of an EMD is illustrated in Figure 4.

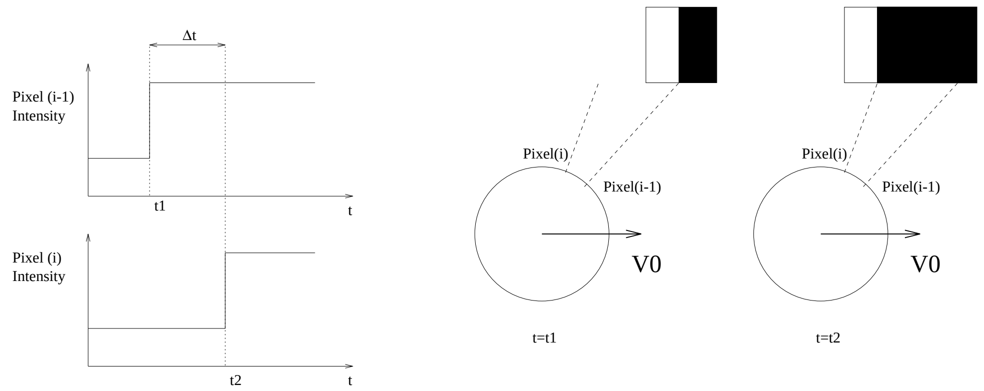

Fig. 4.: Signal processing within an EMD

2.2 Composition of the Artificial Retina

After establishing the basic principles of operation of the EMD, we need to design the layout of the compound eye, which will depend on the robot motion strategies and on the photosensors visibility radius, which will be analysed first.

Robot Locomotion

In general, if the robot translates and rotates, the perceived flow will have two components due to the two motion components. However, only the translational component depends on distance to obstacles and therefore contributes to the purpose of obstacle detection and avoidance. For the sake of simplicity, we assume that the robot moves with pure translational motion while performing optical flow computations. This assumption is justified by the fact that this motion pattern can also be observed from the housefly. In [10], van Hateren showed that the trajectory over time of a blowfly during flight consists of a sequence of translations with very quick turns, thus maximising the time available for analysing the surroundings. Although we do not consider rotational components of robot locomotion, one should realise that with quick turns, vision becomes blurred and the way in which 3D-structures are seen can be affected.

If we further constrain the robot motion to a **constant** linear velocity V_0, then we obtain the additional benefit that all EMD's can perform the same processing over time, since there is no need to adapt the temporal filters as the robot accelerates or slows down. We will therefore assume that the robot moves at constant speed and with fixed-length translational steps ΔL while measuring the optical flow.

Having defined this fixed translational step ΔL, we can drive the robot at some maximum speed and therefore fix a sampling period T for reading out the information processed in all EMD's.

Visibility Radius (r_{vision})

Since the compound eye photosensors have a limited sensitivity, another important parameter is the radius of visibility, r_{vision}, that specifies the maximum distance for which motion detection is possible. To assure anticipated detection, this radius has a lower bound:

$$r_{vision} \geq r_{eye} + \Delta L \tag{3}$$

Distribution of the Ommatidia

The distribution of the ommatidia in the compound eye must ensure that when a robot performs a translation ΔL, all edges within the visibility radius, r_{vision}, must cross at least two optical axes of some EMD, to enable detection.

In the frontal part of the eye, the perceived flow is quite small since the velocity tangential to the eye vanishes when $\varphi \rightarrow 0$ (where we have a singularity). To guarantee that a 3D point will cross 2 optical axes within one step and stimulate an EMD, we must impose an upper limit on the angular separation $\Delta\varphi_{min}$ between the photosensors. As the physical dimensions of the photosensors determine a lower bound for this angular separation, there will be a frontal area of the compound eye for which a different layout must be sought.

As we go towards the lateral visual field ($\varphi \rightarrow 90^0$), the perceived angular velocity increases. A minimal configuration can be obtained if we warrant that each point crosses *not more than* two optical axes within one step, as well. As a consequence the photosensors separation $\Delta\varphi$ must increase as $\varphi \rightarrow 90^0$. Another interesting comment is that this space variant resolution guarantees that the perceived optical flow will be approximately constant independently of the viewing direction. Hence we may have spatial and temporal filters tuned to the same scale all over the visual field, which is an important advantage.

This non-linear distribution of the ommatidia, with increasing resolution towards the central (foveal) region, is illustrated in Figure 5. The adaptation of the resolution across φ is given by equation (4).

$$\Delta\varphi_i = arctan\left(\frac{0.5 \cdot \Delta L \cdot sin\varphi_i}{r_{vision} + 0.5 \cdot \Delta L \cdot cos\varphi_i}\right) \tag{4}$$

$$\varphi_i = \varphi_{i-1} + \Delta\varphi_i \tag{5}$$

To assure that an obstacle crosses two optical axes within r_{vision}, the minimum angular separation on this radius has to be $0.5 * \Delta L$.

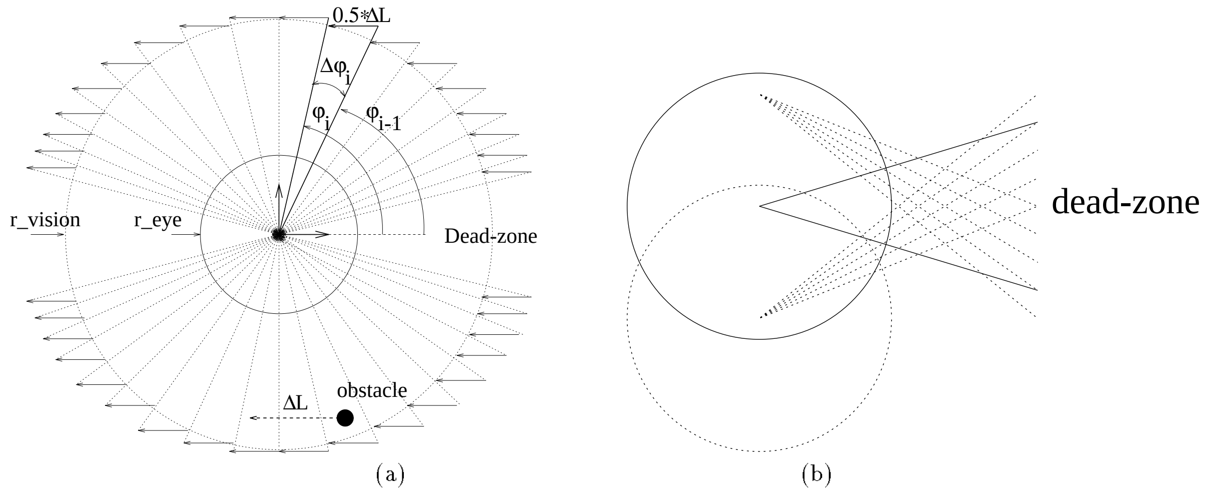

Fig. 5.: The pixel distribution in the compound eye; (a) lateral vision; (b) frontal vision.

The frontal part of the compound eye is designed based on sections of auxiliary compound eyes with the same radius and dislocated laterally, as shown in Figure 5. The non-uniform spacing between photosensors in this foveal region was determined to ensure obstacle detection within the visibility radius and assuming that all EMD's perform the same signal processing.

This design of the panoramic sensor is based on the model proposed in [2, 8]. A more recent approach is given by Mura et. al.[5], in which a scanning retina is proposed. With this retina, the photosensors are positioned on a vibrating strip and with separated fixed optic. Vibrating the array of photosensors will alternate the direction of the sensor's receptive field (optics are fixed in position), so each photosensor scans the world along two different visual axes. The main advantage of this approach is that one photosensor can be given as an input to the EMD. In this case, double resolution can be achieved with the same number of photosensors in the eye.

3 Reflex-like Robot Navigation

Equipping a mobile robot with the panoramic sensor allows visual servoing so as to achieve autonomous navigation restricted to the ground-plane. A locomotion with fixed size steps at constant speed allows to periodically evaluate the optic flow in the azimuth-plane. From this evaluation, a map of detected features in polar co-ordinates of the robot's frame can be periodically constructed and updated, assuming that the centre of the panoramic sensor coincides with the robot's reference frame. Using this map in a repulsive potential field approach allows to implement an obstacle avoidance algorithm and also reveals some embedded insect-like behaviours.

3.1 Constructing an obstacle map with the panoramic sensor

Whenever a contrasted feature within the robot's vision field crosses the receptive field of two photosensors, the corresponding EMD fires and outputs an estimate of the local optical flow. At the end of each translational step, any contrasted features inside the visual field will be detected by the array of EMD's. Identifying the stimulated EMD's then allows to locate these features and construct a map containing their polar co-ordinates in the robot reference frame. The polar angle is given by the angle of the second pixel corresponding to the input of the stimulated EMD. The radial distance is obtained by relating the amplitude of the local measured optical flow to the principle of Motion Parallax.

$$D_i = \frac{V_o \cdot sin\varphi_i}{\Omega_i} \tag{6}$$

where D_i is the radial distance to a feature detected by EMD_i, φ_i is the polar angle of EMD_i in the sensor reference frame and Ω_i is the local component of the optical flow measured by EMD_i.

This relation allows to calculate the radial distance under the assumption of an executed step with pure translation at constant velocity V_0, using equation (6).

3.2 Visual servoing

The robot heading direction is obtained by updating the obstacle map periodically. This is done by a potential field approach constructed from the polar obstacle map. Expressing the obstacle's polar co-ordinates as vectors in the robot reference frame, we calculate the resulting potential vector from the stimulated EMD's. The direction θ of this resultant force is given by:

$$\theta = arctan\left(\frac{\sum_i \frac{1}{D_i} \cdot sin\varphi_i}{\sum_i \frac{1}{D_i} \cdot cos\varphi_i}\right) \tag{7}$$

The inverse radial distance emphasises the obstacles near the robot. Hence, the obstacle avoidance manoeuvre is triggered by the robot ego-motion. Also note that the direction of this resultant force is not very sensitive to the effect of noise in single EMD measurements. Reminding that a single EMD calculates local components of the optical flow using only two pixels to estimate derivatives, one can discuss the effect of noise on these measurements. However, the use of the direction of the resulting repulsive force for visual servoing somewhat takes account for the effect of noise by averaging out these errors. Therefore using this repulsive force in visual servoing has a robust performance.

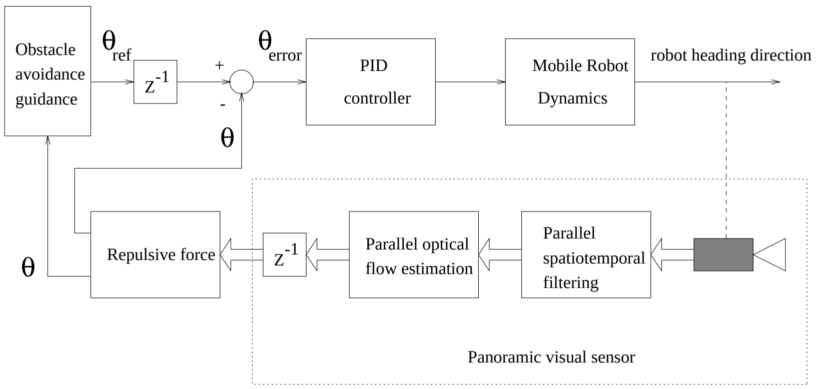

Fig. 6.: Navigation loop for controlling the robot's orientation. The time delay $T = \frac{\Delta L}{V_0}$ is introduced by the image processing. This delay is also the period for updating the polar map.

The obstacle avoidance system aims at orientating the robot perpendicularly to the repulsive force. Therefore, the reference heading direction, θ_{ref}, is given by:

$$\theta_{ref} = \begin{cases} 0^\circ & \text{, if } \theta = 0^\circ \vee \theta = 180^\circ \\ \theta - 90^\circ & \text{, if } 0^\circ < \theta < 180^\circ \\ \theta + 90^\circ & \text{, if } 180^\circ < \theta < 360^\circ \end{cases} \tag{8}$$

4 Implementation

The artificial compound eye consists of a spherical retina, with radius $r_{eye} = 8.63cm$, composed of an array of parallel processing and analogue EMD's. Photodiodes with build-in optics were used as EMD inputs. The EMD's are able to realise motion detection within a range of 30 cm. We tried to minimise the number of pixels in the eye so as to minimise the number of corresponding EMD's. The pixel distribution obtained is shown in Figure 7. The total number of pixels used is 30. The corresponding minimal step-size is $\Delta L = 5.75cm$, obtained from a minimal required radius of motion detection given by $r_{vision} = 14.38cm$. Note that this minimum required radius is within the range of motion detection. The robot translational speed was established at $V_0 = 35cm/s$.

The visual servoing is realised by an 8-bit micro-controller that periodically ($T = 163ms$) scans the outputs of the EMD's in the panoramic sensor and instantaneously updates the polar map. From the updated map, the robot heading direction is calculated. A digital PD-controller is used to control the robot's orientation, using a differential-drive configuration for the mobile robot.

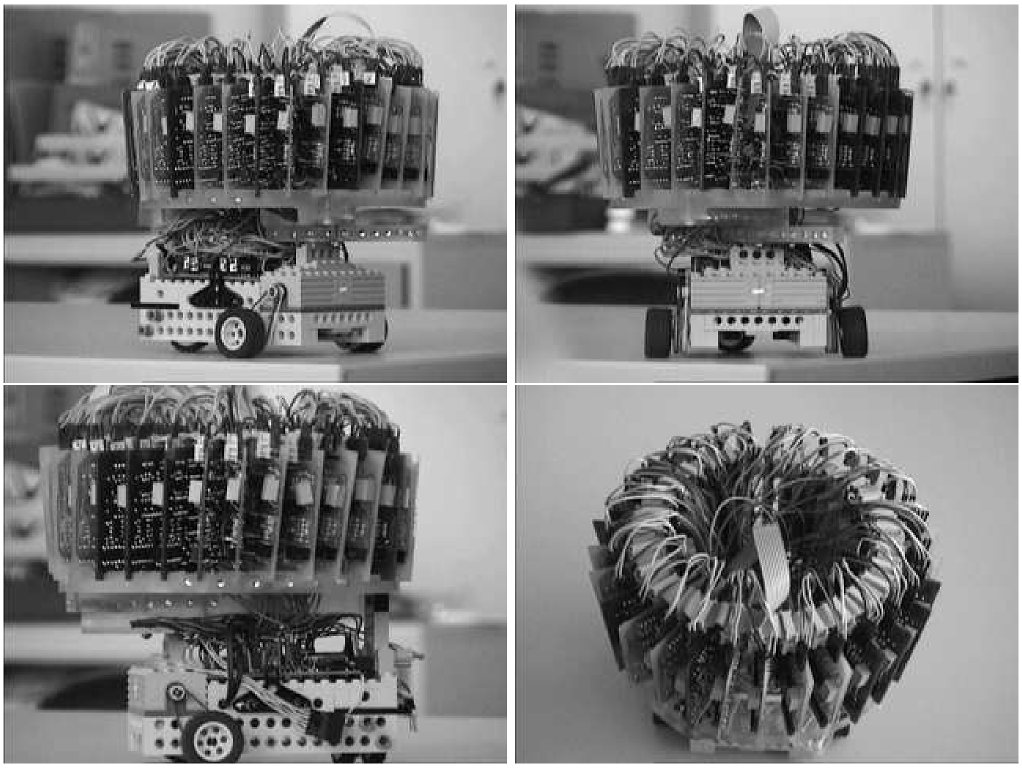

Fig. 7.: Hardware implementation

5 Results

The compound eye's image processing as well as the robot dynamics were simulated so as to test the navigation algorithm. Some simulation results are illustrated in Figure 8. The figures show the trajectory over time of the centre of the robot reference frame in configuration space [3]. The introduced point-obstacles are enlarged by the robot radius so as to indicate the restricted area in which the robot is not allowed to enter as it would collide with the obstacle.

From the upper left simulation in Figure 8 it becomes clear that obstacle avoidance reflexes are realised and the robot survives in an unknown scene. With the upper right simulation in Figure 8 one can see that the frontal vision detects obstacles at the robot's front and is able to realise obstacle avoidance in this extreme case. Observation of the lower left simulation in Figure 8 shows the centring behaviour as described by Santos-Victor [9] and Srinivasan [6]. An interesting fact is that no initially attempts were made so as to obtain this behaviour. It is embedded implicitly in the navigation algorithm, resulting from the calculations based on optical flow. In the lower right simulation of Figure 8 the robot is enclosed in an environment without any possibility to escape. The robot searches for the free space without success and wanders around in circles which is a safe attitude.

Similar experiments have been realised with real-time robot navigation. In Figure 9 the set-up is illustrated, from which one can see that a contrasted environment is required for realising motion detection. This dependency on highly textured environments is due to the established sensitivity for edge detection in each EMD, which is determined by the threshold value applied to the temporal high-pass filters used for edge detection. Proper tuning of the analogue filters requires careful frequency-analysis of the spatio-temporal signals, a topic that will be addressed to in future work.

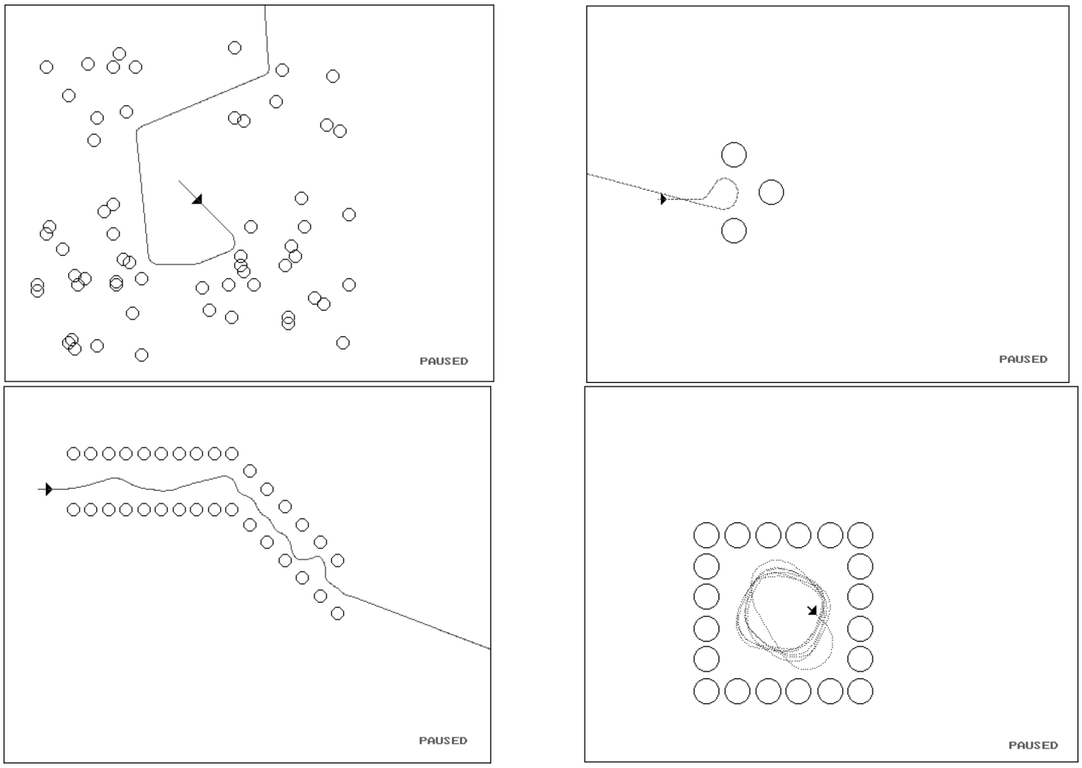

Fig. 8.: Simulation results

In each experiment, the robot's trajectory over time is drawn within the images so as to illustrate its reflex-like behaviour. Comparison with the simulation results shows the similarities. Another interesting feature is the robot's reflex to dynamic changes in its surrounding environment. A person crossing the robot's path causes an immediate reflex of the robot, thus trying to avoid collision.

6 Conclusions & Future Work

A panoramic visual sensor inspired by insect-vision was designed and implemented for real-time autonomous robot navigation. The robot equipped with the insect-like vision system realises motion detection from ego-motion, from which it can update its heading direction using a potential field approach. The implemented obstacle avoidance algorithm results in a reflex-like behaviour. An important fact is that the system does not need to have a notion about any world reference, neither does it have a priori knowledge of the surrounding environment. From our experiences we conclude that low-level parallel image processing can perform a

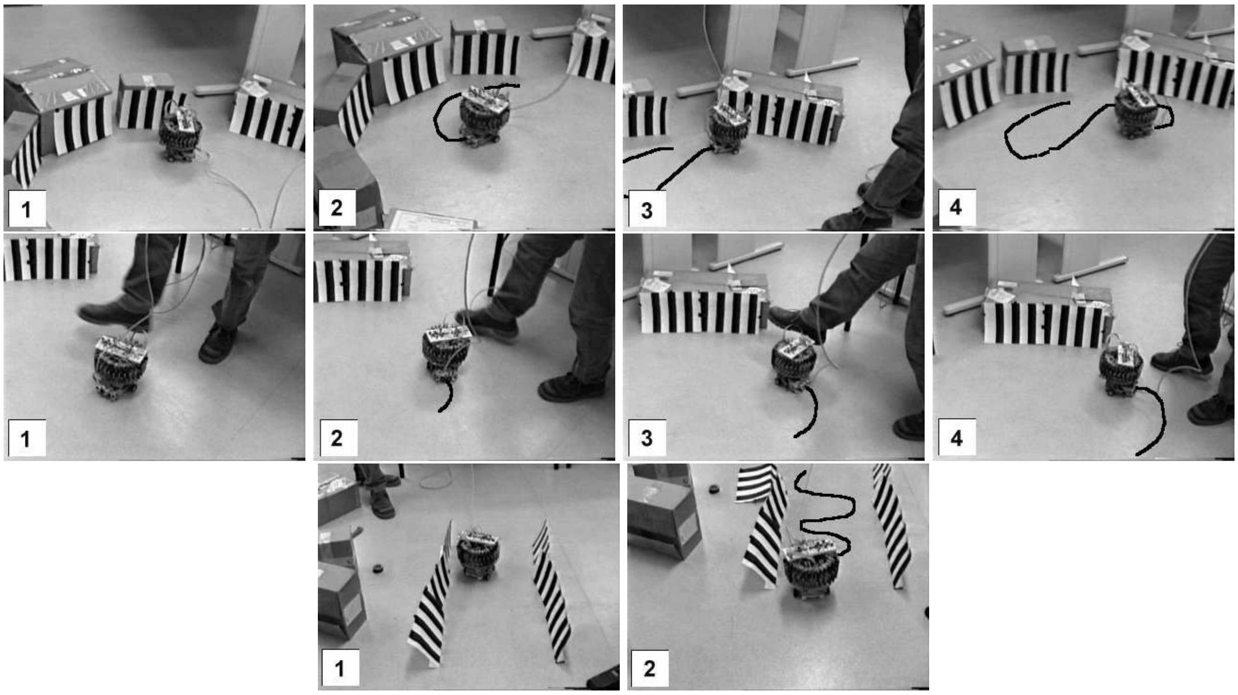

Fig. 9.: Real-time experiments: the first row illustrates a sequence in which the robot is able to detect a contrasted wall and navigate into safe directions; the second row shows the robot's reflex to dynamic features in its surrounding; the last row illustrates the passage of a contrasted corridor (oscillation is due to limited radius of vision)

natural adaptation to any contrasted environment, thus assuring the robot's survival. Although the system showed satisfactory reflex-like behaviour, no higher level goals are specified. It is interesting to give the robot an objective and integrate higher level controllers in such a way that they need not worry with the low-level obstacle avoidance that is realised by the insect-like visual sensor.

Future work will include some analysis and compensation of the neglected rotational component of the robot locomotion, which produces gaze shifts and therefore affects vision. The resulting rotational flow of optical signals interferes with the pattern of optic flow that normally reveals the 3D-structure of the surroundings during translation. We think that proper tuning of the analogue image-processing filters and threshold values can considerably reduce the error, introduced by the rotational component of locomotion, in local optic flow calculation. To do so, we think of a careful characterisation of the spatio-temporal signals in frequency-domain and realisation of some qualitative tests concerning motion detection. The frequency-analysis also will contribute to a better understanding of the parallel image processing. Also the design of a 2D insect-like visual sensor for navigation purposes that are not restricted to the ground plane, will be addressed.

We believe that in natural vision systems, optical flow computation plays an important role in navigation. Biological inspiration from insects is justified by the fact that their navigation skills are still not achievable in today's mobile robot applications.

7 Acknowledgements

We would like to thank Alexandre Miguel Lopes Barroso for his work and efforts in realising this project. This work has been partially funded by project PRAXIS 2/2.1/TPAR/2074/95, SIVA.

References

1. N. Franceschini. Sequence discriminating neural network in the eye of the fly. In Frank H. Eeckman, editor, *Analysis and Modeling of Neural Systems*. Kluwer Academic Publishers, 1992.

2. N. Franceschini, J.M. Pichon, and C. Blanes. From insect vision to robot vision. *Philosophical Transactions of the Royal Society, London*, B(337):283–294, 1992.
3. Jean-Claude Latombe. *Robot Motion Planning*. Kluwer Academic Publishers, 1991.
4. M. Lehrer, M.V. Srinivasan, S.W. Zhang, and G.A. Horridge. Motion cues provide the bee's visual world with a third dimension. *Nature*, Vol. 332(6162):356–357, 24 March 1988.
5. F. Mura, K. Hoshino, H. Morii, K. Suematsu, and I. Shimoyama. A small-sized panoramic scanning visual sensor inspired by the fly's compound eye. In *Proc. of the 1998 IEEE International Conference on Robotics and Automation*, May 1998.
6. M.V.Srinivasan, K. Weber, and S. Venkatesh. *From living eyes to seeing machines*. Oxford University Press, 1997.
7. R.C. Nelson and J. Aloimonos. Finding motion parameters from spherical motion fields (or the advantages of having eyes in the back of your head). *Biological Cybernetics*, (58):261–273, 1988.
8. Pichon, J.M. Blanes, and N. Franceschini. Visual guidance of a mobile robot equipped with a network of self-motion sensors. In *SPIE Vol. 1195 Mobile Robots IV*, 1989.
9. J. Santos-Victor, G. Sandini, F. Curotto, and S. Garibaldi. Divergent stereo in autonomous navigation : From bees to robots. *International Journal of computer Vision*, 14:159–177, 1995.
10. J.H. van Hateren and C. Schilstra. Stabilizing gaze in flying blowflies. *Nature*, Vol. 395:654, 15 October 1998.

This article was processed using the TEX macro package with SIRS99 style