BIO-INSPIRED MOTION DETECTION FOR A BLIND SPOT OVERTAKING MONITOR

S. Mota, E. Ros, E. M. Ortigosa, F. J. Pelayo

Departamento de Arquitectura y Tecnología de Computadores, E. T. S. I. Informática, Universidad de Granada, Periodista Daniel Saucedo Aranda s/n, 18071 Granada, Spain.
{smota, eros, emartinez}@atc.ugr.es ; fpelayo@ugr.es

Abstract
The rear-view mirror is unhelpful when an overtaking car is in the blind quadrants (blind spot). In this contribution we describe the software implementation of an algorithm to monitor vehicle overtaking processes. This algorithm detects the vehicle to the rear, and discriminates whether it is approaching or not, and if approaching, it alerts us of its presence. The proposed system is based on the Reichardt correlator model [1]. The approach presented uses the saliency of motion features in a competition scheme to filter noise patterns. In this way features corresponding to rigid body motion self-emerge from the background. Real overtaking sequences have been to develop this monitoring system.

Key Words: Reichardt correlator model, motion detection, segmentation, blind spot.

1. Introduction

Human mistakes or distractions are the cause of many traffic accidents. This motivates research efforts in different areas (vision, smart control, multi-modal fusion, etc) applied to the automotive technology, mainly related to vehicle security and driver assistance [2, 3, 4]. In 10 years time, vehicles will most likely have built-in devices based on computer vision to aid the driver. The medium-long term goal is to implement devices based on the vertebrate visual system, due to their astonishing efficiency in analysing dynamic scenes. However, current vision models are limited and of high computational cost.

Flies are capable of exploiting optical flow by calculating the local image motion through Elementary Motion Detectors (EMDs) and integrate these signals. This processing takes place in neurons with large dendritic trees and whose receptive fields match certain optic flow fields [5]. Reichardt and other researchers developed a correlation based model of motion detection that captures the functionality of these neural circuits [1].

This paper describes the software implementation of an algorithm, based on EMDs, to analyse overtaking scenarios. The algorithm detects the vehicle to the rear, and discriminates whether it is approaching or not. The whole system, using the algorithm, can alert the driver of a potential hazard. It could prevent an accident when the driver plans to change lanes when another vehicle is close to the blind spot of his rear-view mirror. Fig. 1 illustrates the problem. Only area (1) is directly visible when the driver turns his head left, therefore, loosing his visual track. Area (2) can be seen by the driver through the rear-view mirror. Area (3) is the one covered by the camera. A vehicle (A) can be seen by the camera and by the driver through the rear-view mirror.

The vehicle in (B) is positioned in the blind spot and therefore only the camera can track it. This blind spot area is the one to be monitored by the system.

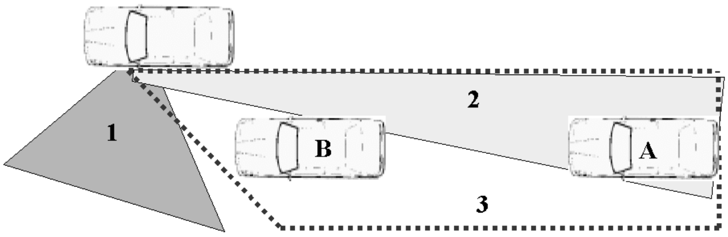

Figure 1. Problem description.

This work is part of the European Project ECOVISION [6]. Its goal is to employ basic knowledge about biological vision systems to design a hybrid software-hardware system to address the posed problem.

We have used real overtaking sequences to develop this monitoring system. The images have been taken with a camera fitted to the driver's rear-view mirror. They have been provided by Hella [7] (partner of ECOVISION). The sequences comprise 500 frames (20 seconds of recording) with a resolution of 288x384 pixels per frame and 256 grey levels. The sequences include various overtaking processes: slow overtaking, rapid overtaking, inverse overtaking, etc; and represent the initial stages of an ideal test bed for this type of application.

The overtaking database includes sequences in very different weather conditions taken with a High Dynamic Range Camera. This has shown to be more robust for outdoor applications. The proposed algorithm has been tested with different weather conditions and has produced very promising results as shown in Fig. 9.

The approach described in this paper is being implemented in digital real-time hardware, specifically in the FPGA device. A preliminary design has already been completed using a low cost 200K equivalent gates device appropriate for the parameters shown above.

2. The Detection Algorithm

The original sequence is pre-processed using the Sobel detector for edge extraction. This pre-processed sequence is the input for a Reichardt block, where we correlate each pair of pixels in the sequence through a set of EMDs tuned to different velocities. Finally, we assign a velocity to each pixel at every frame. This velocity map from the Reichardt detector block is processed in the next stage, where dynamic filters modify the velocity map applying rigid body motion rules to characterize moving objects. We also apply rules related to perspective correction and temporal coherency to finally segment the vehicle. All these stages will be explained in detail in the next sections.

2.1 Pre-Processing

David Marr's visual perception theory [8] pointed out, according to the results obtained by neurophysiology experiments, that object borders are the most important cues to extract scene structure. For this reason, we start by obtaining the spatial image edges and, in the second step, these edges are processed to detect motion.

Eventually, we want to implement the proposed algorithm through specific hardware to be used in embedded systems. For this reason, we have chosen a simple edge detector able to provide acceptable real-time outputs to Reichardt detectors. Consequently, we have used a Sobel gradient detector [9] to extract the edges. However, thinner edges are desirable. So, we only use those edges which are local maximums in a chosen gradient direction.

We only detect the vertical edges. Due to the aperture problem, this approach is unable to obtain non-horizontal movement components. However, this is not so important for the posed problem since the scene' dominant movements are sideward. Furthermore, the horizontal velocity component is enough to determine the approaching trajectory and estimate the time to contact or, at least, the relative velocity with respect to our vehicle. In overtaking sequences on highways with more than two lanes, the lane change process is detected in a higher-level stage (following the sparse map evolution).

Horizontal edges can be also computed but this increases the noise in the next processing stage (Reichardt detectors).

The output of the pre-processing stage is a sparse map composed of pixels with non zero intensities indicating an edge. Consequently, a system based on this approach has a low computational cost.

2.2 Reichardt Detector

Both predators and preys have visual systems specialized in motion detection. They have provided inspiration for many motion detection algorithms proposed in the literature [10, 11, 12, 13, 14]. The system described in this paper is based on the Reichardt correlator model [1] and, in particular, on the fly motion detection model.

Although the Reichardt model was originally developed to explain insect motion detection mechanisms, it has been also used to explain motion detection in humans, cats and birds [15, 16, 17]. Moreover, most spatio-temporal energy models are mathematically equivalent to the correlator model [1].

Fig. 2 shows a simple Reichardt detector. A stimulus sequentially reaches the two detector inputs (a and b).

When the pattern moves in the preferred direction (Fig. 2.a.) the temporal separation of the signals from both input channels is compensated by a delay (d). The output of both channels will be simultaneously active when the delay is appropriate. In this way the two outputs will be perfectly correlated. When the pattern moves in the opposite direction (Fig. 2.b.) the delay increases the temporal separation between the two channel outputs. The output signals have less correlation and, consequently, the detector response will be weak.

Two sub-units, as the one described above, form the complete elementary detector. One sub-unit detects motion to the right and the other sub-unit detects motion to the left (Fig. 2.c.). The complete detector output will be the subtraction of the two sub-unit output signals; a positive final output signal indicates motion to the right, while a negative total output indicates motion to the left.

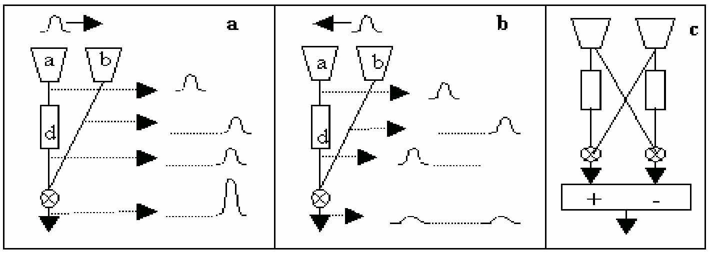

Figure 2. Reichardt detector; a) an edge moves in the preferred detector direction; b) an edge moves in the opposite direction; c) Complete Elementary Detector.

2.3 Velocity Tuning

From the above description, it should be noted that precise velocity tuning adjustment of the Reichardt detector delays is very important. If the input pattern speed and cell delay are different the detector will not respond significantly. The biological example of this is the *Australian tiger beetle* (cincindela hudsoni) [18] which is the fastest running insect in the world (10 km/h). However, its motion detector cells respond only to slower speeds. Consequently, the *Australian tiger beetle* is blind when it runs, and from time to time, it needs to stop in order to adjust its trajectory to the prey movement.

Although biological visual systems appear to be capable of estimating image speed, the basic Reichardt detector does not function as velocity estimator. It only shows motion direction. Some authors have proposed that animals capable of estimating image speed have a collection of detectors tuned to different velocities, or have alternative motion detector systems [19]. In this paper, we have used such a multi-velocity-detector composed of a group of single-velocity-detector cells as the one shown in Fig. 2.c. Each single-velocity-detector cell in the set is tuned to a different velocity. The set of velocities that can be tuned is enough for the purpose of the application, i. e., difficulties arising from blindness of the system to some velocities are not a problem.

We use the set of multi-velocity-detectors to correlate pairs of pixels. As a result, we get a set of velocities associated with each pixel-pair, one from each EMD in the set. The velocity finally associated with a pixel-pair is the one obtained from that particular EMD which maximizes the correlation.

The saliency map of this stage shows the velocity (amplitude and direction) associated with each pixel in all the frames in a sequence.

All the EMDs tuned to a particular velocity compose what we call *velocity channel*. We integrate the EMDs responses of a local area of the input image as shown in Fig. 3, obtaining linear plots that represent the population activity of EMDs of a particular *velocity channel* along the x-axis of the image.

2.4 Rigid Body Motion and Perspective Correction

The output of the detector layer is a cloud of points. Each point (pixel) is associated to a velocity and the velocities in a neighbourhood of pixels can be different in amplitude and/or direction. Consequently, we have a number of points which move at their detector characteristic speed. But how to isolate the overtaking vehicle from this diffuse cloud of points?.

To do this we apply a rigid body motion rule to segment the vehicle. In a rigid body all points move at the same speed (amplitude and direction). Therefore, if we can detect such a population of pixels in a limited area of the image, they are indicative of the presence of the vehicle or another rigid body.

However, due to the perspective distortion this is not true in our sequences. In this case, the points in the distant part of the rigid body seem to move more slowly than closer points. The rear of the vehicle is farther away and appears to move slowly, so the slow velocity detectors will respond to it. The front of the vehicle is closer and it seems to be moving faster and rapid speed detectors detect it. For this reason, different velocities in the sequence must be taken into account in order to synchronize front and rear of the car.

The perspective deformation effect explained above applies to the complete visual field of the rear-view mirror. Hence, a moving object (overtaking car) is expected to move slowly when it is localized in the very left side of the image and its speed will apparently increase as it moves (overtakes) rightwards through the visual field due to the perspective effect in the rear-view mirror.

We have introduced some rules to deal with this perspective deformation. These rules allow us to filter the output signals from our Reichardt multi-velocity-detector.

See Fig. 3 for an example of how this is done. In the upper part of the figure we can see the velocity map of the Reichardt stage. Each point in the figure is associated to a velocity. The medium part of the figure represents a single velocity channel, the plot of the number of points in each restricted area within the figure, with velocities that belong to cluster V_i. Locations **Pa** and **Pb** of the velocity channel V_i gather the points of the area **A** and **B** in the image, respectively. This process is computed in identical windows along the *x*-axis, producing the velocity channel response. **Pb** receives only poor contribution of its window (**B**) while **Pa** is a local maximum; in other words, there is an important population of points tuned to the cluster characteristic velocity V_i. These velocity channels work as band pass filters. Only those points of the velocity map tuned to the cluster velocities that produce the local maximum of the velocity channel plot are retained. The maximum corresponds to points were a rigid body motion induces coherent feature motion. Hence, the window **B** does not produce an output (the points belonging to this receptive field are filtered); however, the points tuned by V_i in window **A** appear in the final output layer.

In other words, the sum of the components detected in a local region of the frame enhances the detection capability of solid objects. This approach helps to eliminate patterns such as those produced by the wind on the vegetation and others.

Moreover, if the detectors indicate that the vehicle moves rightwards, then it is approaching us. This is the situation we would need to be alerted of. On the other hand, if the detectors indicate that the vehicle moves leftwards, then it is moving away, and it will not disturb us.

So far, we have demonstrated the application of the dynamic filters by counting points in areas only along x axis (by columns) in order to understand the method, but our results have been obtained by applying a grid; in other words, we have considered windows along both *x* and *y*-axes.

3. Results

Fig. 4.a shows a real camera image. The representation in Fig. 4.b shows the results of pre-processing (vertical edges detection).

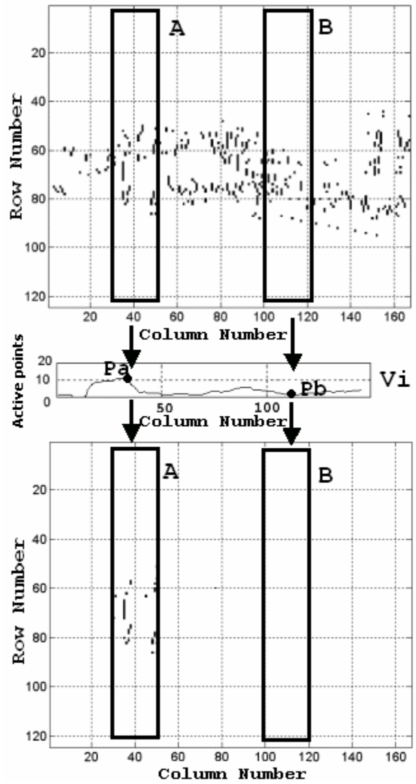

Figure 3. Using velocity channels as dynamic filters

When we apply the Reichardt detector to the vertical edges we obtain the representation in Fig. 4.c (features moving to the right) and Fig. 4.d (features moving to the left).

Figs. 4.c and 4.d are very noisy; they also contain erroneous points (leftward moving points in the area occupied by the car and rightward moving points in the road area). In order to reduce the noise and effectively segment the overtaking car we can use the velocity channel responses (Fig. 5) to dynamically filter these velocity maps.

Each plot presents the number of points synchronized with a speed in the range of V_i (the characteristic velocity of each of the 10 channels or motion detector populations). Its windows cover 20 columns from the input image (Figs. 6 and 7). In this way, the points of Figs. 4.c and 4.d have been grouped into different velocity ranges around certain spatial areas. This is what is represented in Fig. 5, the left column shows rightward velocity channels and the right column shows leftward velocity channels.

These velocity channel responses are used as dynamic filters to produce the representation shown in Figs. 6.a and 6.b that contain less erroneous points. It can be seen that Fig. 6.a represents the filtered rightward moving features (only present in regions filtered by the maximums indicated in Fig. 5). All these points belong, in fact, to the overtaking car (rigid body motion), so we can take the estimated centre of the overtaking car as the centre of mass for these filtered features moving rightward. On the other hand, the points plotted in Fig. 6.b represent the leftward moving features filtered by the maximums indicated in Fig. 5. All of them correspond to the ego-motion pattern.

The results shown are very promising, but are they really accurate? To evaluate this, we manually marked the overtaking car with a rectangle in every frame in the sequence (Fig. 7 shows this).

We calculated the distance between the centre of mass (in Fig. 7 it is represented by a cross) and the centre of the rectangle. This distance is normalized dividing it by the radius of the minimum circle containing the rectangle in each frame. This distance is what we call *Quality Measure* (QM). If the centre of mass falls into this circle this QM is below 1 and we are detecting the overtaking vehicle. In other cases higher than 1 noisy pattern are dominant and lead to incorrect estimations. Fig. 8.a shows the QM along the sequence. When the overtaking vehicle is small (far away) the detection has errors (left image shown in Fig. 7). However, at a medium distance (around frame 110), the vehicle is big enough and the vehicle is correctly detected. In fact, accurate detection occurs when the overtaking vehicle becomes a potential hazard.

Fig. 8.b shows the variance of the QM along the sequence. We calculated the variance of the QM in every 5 estimations. We can see a convergence to zero, in other words, from a point in the overtaking process we detect the overtaking vehicle without errors, and the system is able to warn of the presence of the overtaking car. In Fig. 8.a can be seen that from approximately frame 110 (shown in the centre image of Fig. 7), the detection is done accurately.

During the Reichardt stage we have used the correlation between single pixels. We have also repeated the process using a correlation between blocks of pixels (block matching). The advantage of block matching is that the Quality Measures falls below 1 sooner, and the convergence to zero is also faster.

4. Conclusions

The present contribution describes a motion processing system to be used as a blind spot monitor. It is intended to detect overtaking cars. The front-end of the system comprises Reichardt motion detectors. We define dynamic filters based on motion patterns of the image that seem to correspond to moving objects. These dynamic filters effectively clean noisy patterns and help to segment on overtaking vehicle. This filtering technique is robust because it is only based on a rigid body motion rule. It detects areas within a population of features moving coherently (with the same velocity and direction), being good candidates for a moving rigid body. The moving features are processed in a competitive a manner, only patterns that activate a whole population of detectors with a similar velocity become salient and pass through this dynamic filter stage. The system has been tested on real overtaking sequences. The system is currently being tested with sequences taken from a high dynamic range camera in different weather and light conditions to evaluate its robustness to low contrast and reflective mediums (such as water drops).

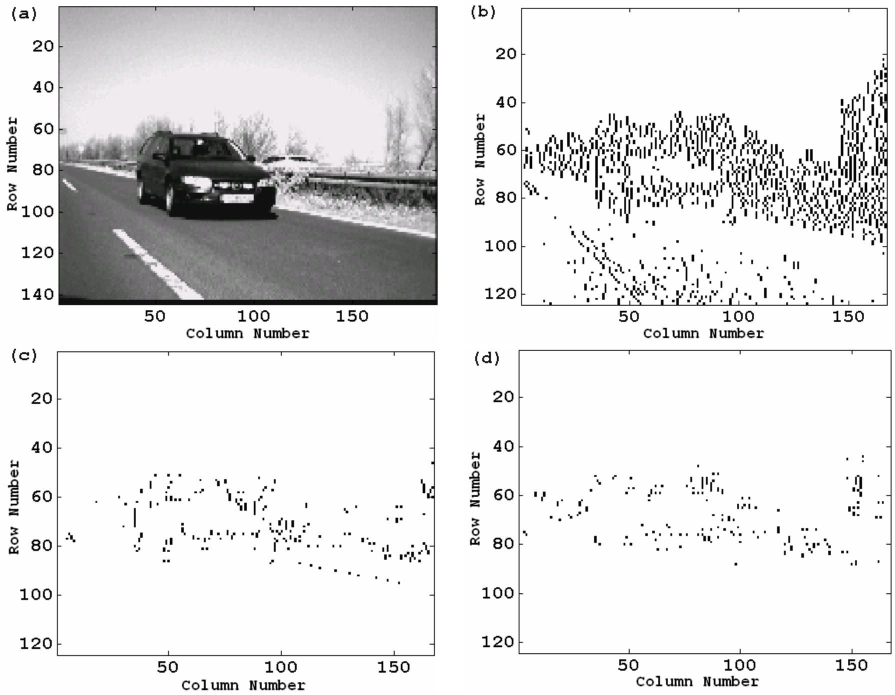

Figure 4. Motion extraction: **(a)** Real Image, **(b)** Vertical edges of the image, **(c)** Rightward moving detected features, **(d)** Leftward moving detected features.

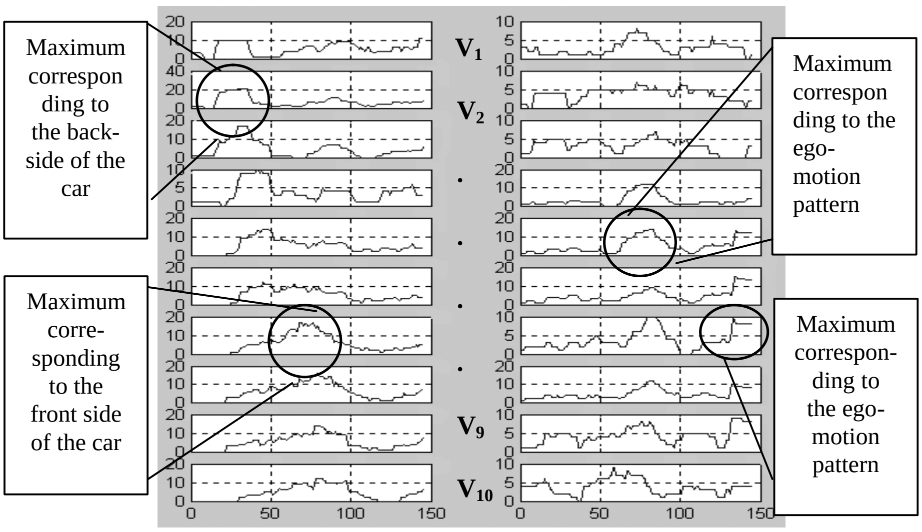

Figure 5. Velocity channels as dynamic filters. The left column shows rightward motion dynamic filters, and the right column shows leftward motion dynamic filters.

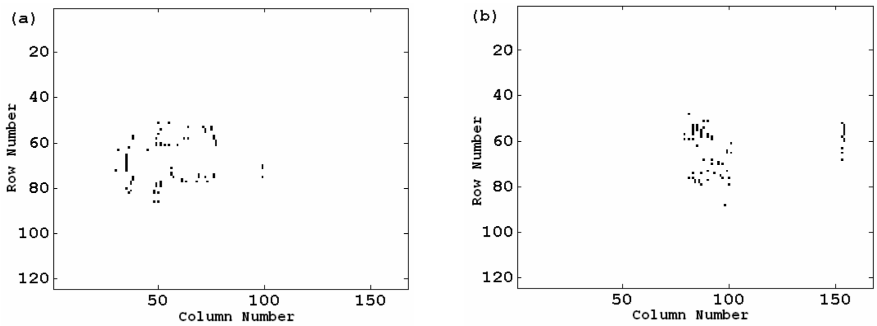

Figure 6. Filtered motion features: **(a)** rightward movement, **(b)** leftward movement.

Figure 7. Overtaking car manually marked with a trust rectangle. In the first image the car is too far away to be correctly detected and the system fails.

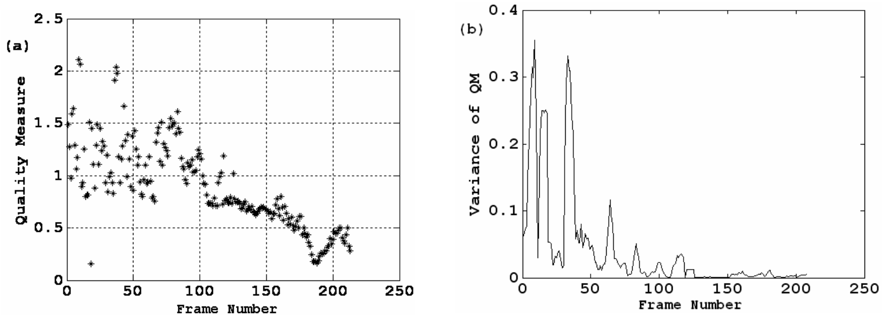

Figure 8. (a) QM of the detection along the sequence of Fig. 7; **(b)** Variance of the QM along the sequence of Fig. 7**.**

Figure 9. Example of the system performance under adverse weather conditions (foggy and rainy). The sequence had been taken with a High Dynamic Range Camera.

5. Acknowledgements

This work has been supported by the 5th EU research framework funds through the European Projects ECOVISION (IST-2001-32114) [6] and CORTIVIS (QLK6-CT-2001-00279). We would like also to acknowledge Hella [7] for providing us with the overtaking sequences.

References

[1] W. Reichardt: "Autocorrelation, a principle for the evaluation of sensory information by central nervous system" in Sensory Communication, W. A. Rosenblith ed. 303-317 (1961).

[2] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner & W. von Seelen, Computer Vision for Driver Assistance Systems, Proc. *Session Enhance and Synthetic Vision*, 1998, 136-147.

[3] S. Görzig & U. Franke, ANTS-Intelligent Vision in urban Traffic, Proc. *IEEE Conference on Intelligent Transportation Systems*, 1998.

[4] U. Franke, D. Gavrila, A. Gern, S. Görzig, R. Janssen, F. Paetzold & C. Wöhler, From door to door- Principles and Application on Computer Vision for driver assistant systems, Proc. *Intelligent Vehicle Technologies: Theory and Applications*, 2000.

[5] H. G. Krapp, Neuronal Matched filters for optic flow processing in flying insects, *International review of neurobiology, 44*, 2000, 93-120.

[6] ECOVISION, http://www.pspc.dibe.unige.it/~ecovision/

[7] Dept. of predevelopment EE-11, Hella KG Hueck & Co., Lippstadt, Germany, www.hella.de

[8] D. Marr, Vision (New York: W. H. Freeman and Company, 1982)

[9] R. Gonzalez & R. Woods, Digital Image Processing, (Addison Wesley, 1992)

[10] A. Johnston, P. W. McOwan & C. Benton, Robust velocity computation from a biologically motivated model of motion perception, *Proc. Royal Soc. Lond. B266*, 1999, 509-518.

[11] A. Watson & A. J. Ahumada, Model of human visual-motion sensing, *J. Opt. Soc. Am. 2(2)*, 1985, 322-341.

[12] K. Nakayama, Biological image motion processing: a review, *Vision Res. 25*, 1985, 625-660.

[13] P. R. Schrater, D. C. Knill & E. P. Simoncelli, Mechanism of visual motion detection, *Nature America 3(1)*, 2000, 64-68.

[14] E. H. Adelson & J. Bergen, Spatiotemporal energy models for the perception of motion, *J. Opt. Soc. Am. A2*, 1985, 284-299.

[15] J. P. H. van Santen & G. Sperling, Elaborated Reichardt detectors, *J. Opt. Soc. Am. A2*, 1985, 300-321.

[16] F. Wolf-Oberhollenzer & K. Kirschfeld, Motion sensitivity in the nucleus of the basal optic root of the pigeons, *J. Neurophysiol. 71*, 1994, 1559-1573.

[17] R. C. Emerson, M. C. Citron, W. J. Vaughn & S. A. Klein, Nonlinear directionally selective subunits in complex cells of cats striate cortex, *J. Neurophysiol. 58*, 1987, 33-65.

[18] C. Gilbert, Visual control of cursorial prey pursuit by tiger beetles (Cicindelidae), *Journal of Comparative Physiology A181*, 1997, 217-230.

[19] D. J. Field, Relations between the statistic of natural images and the response properties of cortical cells, *J. Opt. Soc. Am. A4*, 1987, 2379-2394.