MOTION DRIVEN SEGMENTATION SCHEME FOR CAR OVERTAKING SEQUENCES

S. Mota[1], E. Ros[1], J. Díaz[1], G. Botella[1], F. Vargas-Martín[2] and A. Prieto[1]

[1] Departamento de Arquitectura y Tecnología de Computadores, Universidad de Granada, Granada, Spain
[2] Departamento de Física, Universidad de Murcia, Murcia, SPAIN

The rear-view mirror is unhelpful when an overtaking car is at the blind spot. This paper describes a simple algorithm that watches overtaking scenarios from the rear-view mirror point of view. Although motion extraction requires high computational resources and normally produces very noisy patterns in real sequences, if an overtaking vehicle approaches our car we detect it as forward moving features, while the rest of the patterns in the rear mirror visual field move backwards due to the ego-motion of our vehicle. Therefore motion provides useful cues to achieve an efficient segmentation in this application framework. In this paper we use the Reichardt motion detector to extract forward moving objects and we apply a rigid body motion rule to filter features that could belong to an overtaking vehicle. This scheme is used to efficiently segment overtaking cars, using the rear mirror visual field. This system alerts the driver of the host car when an overtaking car (approaching trajectory) is detected.

1. INTRODUCTION

One of the most dangerous operations in driving is to overtake another vehicle. The driver's attention is on his way forwards, and sometimes does not use the rear-view mirror or it is unhelpful because of the blind spot.

The automobile industry is very interested in introducing systems applied to driver assistance, (Franke, 2000; Handmann, 1998). Artificial vision systems would be very effective; however, current bio-inspired vision models (based on vertebrates' visual systems) are limited and require high computational cost. Simpler models based on insects' motion detection are being developed.

Flies are capable of exploiting optical flow by calculating the local image motion through Elementary Motion Detectors (EMDs) and integrate these signals (Krapp, 2000). Reichardt et al. developed a correlation based model of motion detection that captured the functionality of these neural circuits (Reichardt (1961)).

This paper describes the software implementation of an algorithm, based on EMDs, that watches overtaking scenarios: it detects the vehicle behind us, discriminates whether it is approaching or not and alerts us about its presence if necessary. Figure 1 illustrates the posed problem. The area (1) corresponds to the direct driver vision area; of course, the driver must head leftwards to access this angle, losing the visual track of his way. The area (2) can be seen by the driver through the rear-view mirror. The area (3) is the one covered by the camera. A vehicle (a) could be seen by the camera and by the driver through the rear-view mirror. On the other hand, the vehicle (b) is positioned in the blind spot and therefore only the camera could track it. This blind spot area is the one we want to be monitored by the system.

This work is framed in the European Project ECOVISION (Ecovision, 2003). 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 (provided by Hella[1], a partner in ECOVISION) that have been taken with a camera placed on the driver's rear-view mirror. The sequences are composed 500 frames (20 seconds of recording) with a resolution of 288x384 pixels per frame and 256 grey levels. These sequences include different overtaking processes: slow overtaking, rapid overtaking, inverse overtaking, etc; and represent an initial test bed for this kind of application.

The detection algorithm follows different stages. The original sequence is pre-processed. We extract the edges in each frame. This pre-processed sequence is the input for the Reichardt motion detector, where we detect motion through a set of EMDs tuned to different velocities. At this point, each pixel has been detected moving with a velocity (module and direction). This saliency map from the Reichardt step is a noisy pattern. Therefore, dynamic filters modify the saliency map applying rigid body motion rules to characterize the motion of the vehicle. We also apply rules related with perspective correction and temporal coherency to finally segment the vehicle. All these stages will be explained in detail in the next sections.

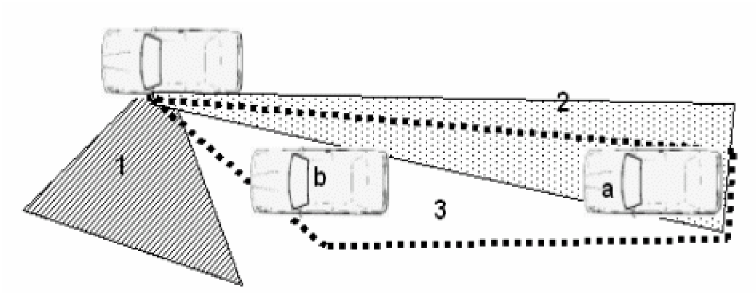


Figure 1. Problem description.

2. PRE-PROCESSING

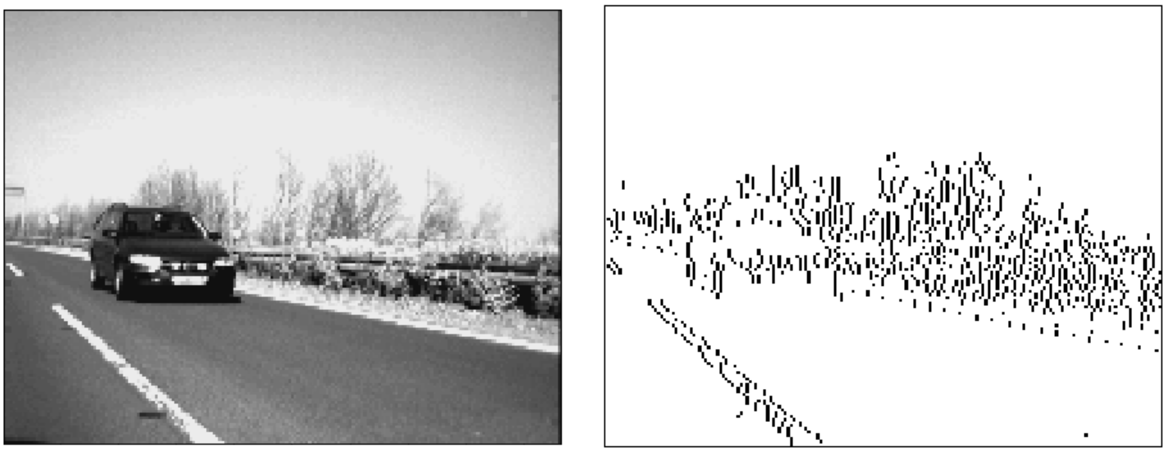
It has been pointed out, through neurophysiology experiments, that object borders are the most important cues to extract the scene structure (Marr, 1982). Because of this, we start obtaining the spatial image edges. We have chosen a simple edge detector, Sobel gradient detector (Gonzalez, 1992), because we want to implement the proposed algorithm through specific hardware to be used in embedded systems, and this detector is able to provide acceptable real-time outputs to Reichardt detectors.

The obtained edges are processed to get them thinner. On the other hand, it will be enough to detect vertical edges because the scenes' dominant movements are sidewards and with vertical edges these motion patterns become more explicit.

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

The output of this pre-processing stage is a sparse map composed of pixels with intensity different from zero when an edge is detected.

Figure 2a shows a real image from an overtaking sequence. After the pre-processing (vertical edges detection) we obtain a sparse map of pixels represented in Figure 2b.



(a) (b)

Figure 2. (a) Real image from an overtaking sequence; **(b)** Vertical edge detected in (a).

3. REICHARDT DETECTOR AND VELOCITY TUNING

Figure 3 shows a simple Reichardt detector. When a motion pattern is detected, it is seen as a stimulus that reaches the two detector inputs (a and b) with a certain delay.

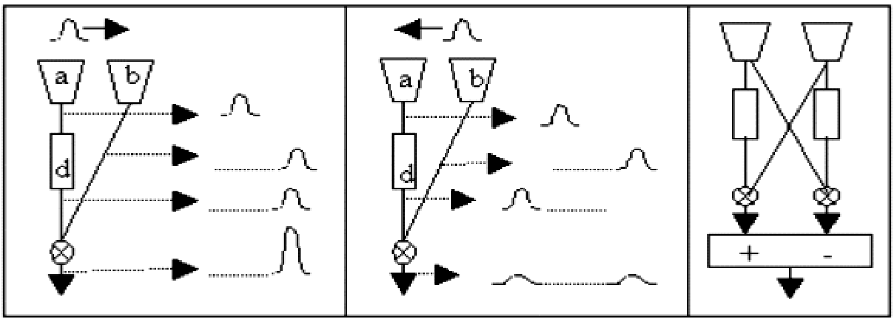
When the pattern moves in the preferred detector direction (Figure 3a) the temporal lag of the signals in both input channels is compensated by a delay (d). The output of both channels will be simultaneous when the delay is appropriated. In this way the two stimuli will be perfectly correlated. When the pattern moves in the opposite direction (Figure 3b) the delay increases the temporal lag between the two channel outputs. The output signals are less correlated and therefore the detector response will be weak.

Two sub-units, as the one described above, form the complete EMD. The EMD output will be the subtraction between the two sub-unit outputs; a positive final output indicates motion to the right, while a negative total output indicates motion to the left (Figure 3c).

A single EMD gives us information, mainly, about the motion direction of the detected features moving with its characteristic velocity (temporal delay associated to the EMD). However, if we apply a set of EMDs, each of them tuned to a different velocity, we can also obtain information about velocity module using a competitive system among the EMDs, i.e., we apply the set of EMDs to the same pixels and we extract its motion. The EMD that maximized the correlation is the one that detects the correct velocity (module and direction) of the studied pixel. We repeat the process for all the pixels along each frame.

The set of velocities that can be tuned is enough for the purpose of the application, i.e., we do not have problems related with the blindness of the system to some velocities. This is the case of the *Australian tiger beetle* (cincindela hudsoni) that is the most rapid running insect in the world (10 km/h). However, its motion detector cells respond only to slower speeds. In consequence, the *Australian tiger beetle* is blind when it runs, and to adjust its trajectory to the prey movement, it needs to stop from time to time (Gilbert, 1997).

When we apply the Reichardt detector to the vertical edges we obtain the representation in Figure 4a (features moving to the right). The output of the Reichardt layer is composed of a cloud of points; each pixel is associated to a different velocity module, but the same direction (rightward direction). Figure 4a is very noisy; it shows also incorrect points, i.e., those points that belong to the landscape, and we know they must move leftward because of the ego-motion of our vehicle. In order to reduce the noise and effectively segment the overtaking car we can use the velocity channels responses (Figure 4c) to dynamically filter these feature maps.



(a) (b) (c)

Figure 3. Reichardt detector; **(a)** an edge moves in the preferred detector direction; **(b)** an edge moves in the opposite direction; **(c)** Complete Elementary Motion Detector.

4. VELOCITY CHANNELS AS DYNAMIC FILTERS

We have a number of points that move at their detector characteristic velocity (Figure 4a). But how to localize the overtaking vehicle from this diffuse cloud of points?.

We apply a Rigid Body Motion (RBM) rule to segment the vehicle. If we only consider lateral transactions in a rigid body all points move at the same speed (module and direction). Therefore, if we detect a population of pixels that have associated the same speed, and all of them are in a limited area of the image, they track the vehicle or other rigid body.

However, this is not true in our sequences due to the perspective correction. In this case, points in the distant part of a rigid body seem to move more slowly than closer points. Because of this, it will be necessary to consider different velocities to synchronize both parts of a car. The back of the vehicle is far away and it seems to move slowly, therefore slow velocity detectors will respond to it, while the front of the vehicle is closer and it seems to move faster. Therefore, it is detected through rapid speed detectors. We need to cluster the velocities of the pixels into a range of velocities.

We introduce a new concept, the velocity channels, that allow us to apply the RBM rule.

We divide the set of detected velocities into groups of neighbouring velocities; and we also divide the image into a grid (Figure 4a illustrates the grid). The next step is to calculate the velocity channels.

The velocity channel V_i is the plot of the number of points tuned with velocities in a range of V_i for each square of the grid (receptive field). We will have as many velocity channels as velocity sub-sets we consider.

The central plot of Figure 4c represents the velocity channel V_i. This 3D figure shows the grid in the x-y plane, and z direction represents the number of points synchronized with velocity V_i for all the receptive fields in the grid. These velocity channels work as band pass dynamic filters. Only the points of the saliency map tuned to the cluster velocities that produce the local maximum in each receptive field of the velocity channel plot are maintained active. The maximum corresponds to points where a rigid body motion induces coherent feature motion. Hence, the receptive field **B** (in the upper part of the Figure 4c) does not produce an output, i.e., the points belonging to this receptive field are filtered because the features of this receptive field moving leftwards (not shown in this plots) are dominant. However, the points tuned by V_i in receptive field **A** appear in the final output layer.

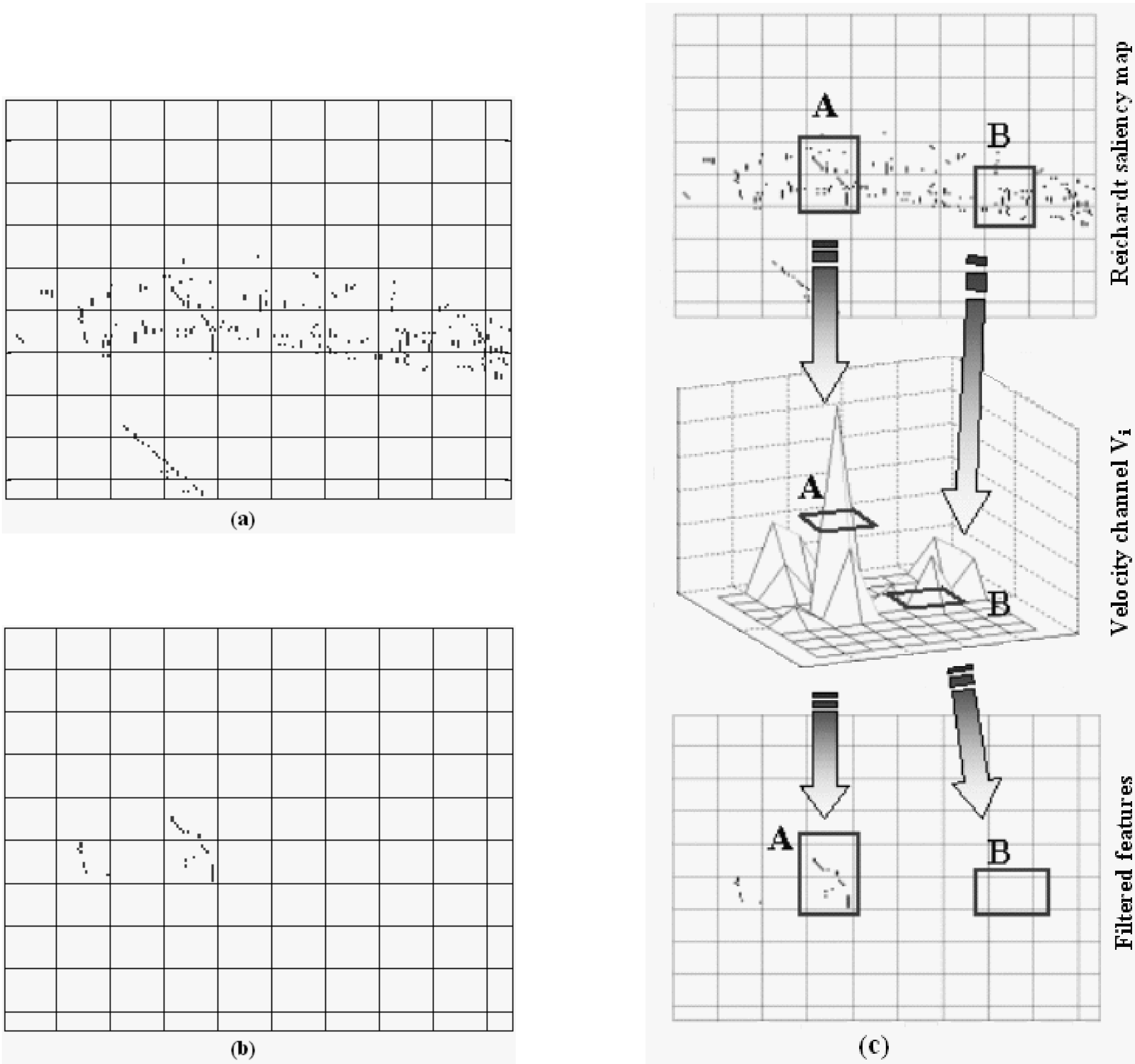


Figure 4. (a) Rightward motion features detected by Reichardt stage; **(b)** Filtered rightward movement with dynamic filters; **(c)** Velocity channel V_i used as dynamic filter to segment the overtaking vehicle.

Finally, we obtain the results in Figure 4b when the full set of velocity channels are applied to the input map.

Moreover, if the detectors indicate that the vehicle moves rightwards, then it is approaching us, and this is the alerting situation we would like to distinguish. On the other hand, if the detectors indicate that the vehicle moves leftwards, then it is moving away, and it will not disturb us. We must remember that these figures represent only the rightward moving features but we have also obtained similar results with ego-motion features (the leftward moving ones), and moving features in opposite directions compete in the saliency map.

5. RESULTS

The results shown are very promising, but are they really good?. Now we are going to evaluate if the previous results really allow us to detect the overtaking vehicle and to warn the driver if necessary.

We have rounded the overtaking car in a square along the sequence. This process has been done manually, so we know the limits of the square in each frame (Figure 5 shows the result of this operation in a frame). The centre of the square for each frame is also known.

We calculate the centre of mass of all the points that are moving rightward in the dynamic filter output (the correspondent centre of mass have been represented in Figure 5 with a cross).

We also calculate the distance between this centre of mass and the centre of the square (this distance is normalized dividing by the size of the square in every frame). If the centre of mass is into the square this Quality Measure (QM) is below 1 and we are detecting the overtaking vehicle properly. In other cases the QM is higher than 1 and the noisy pattern is dominant and lead to a wrong estimation.

Figure 5. Frame 110 from the overtaking car sequence. From this distance on the detection algorithm does not lose the trajectory of the overtaking vehicle; the overtaking car has manually rounded with a square. The centre of mass of rightward moving features is shown with a cross.

Taking as an example an overtaking sequence in which the car drives at a velocity of 90 Km/h and the overtaking car at a contact velocity of 110 Km/h, Figure 6a shows the QM along the sequence. When the overtaking vehicle is small (it is far away) the detection has errors because of the sparse number of points, i.e., we detect few points from the car and the noisy features turn aside the centre of mass. But from a medium distance (from frame 110 on), the vehicle is big enough (the features detected belong to the overtaking car) and the detection is done properly. In fact, accurate detection occurs when the overtaking vehicle begins to be dangerous for us.

Figure 6b shows the variance of the QM along the sequence. We have 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 forward the algorithm detects the overtaking vehicle without errors, and the system is able to warn us about the presence of this overtaking car, if necessary. In Figure 6a shows that approximately from frame 110 on (shown in Figure 5) the detecting 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 only advantage of block matching is that the QM becomes below 1 earlier, and the convergence to zero is earlier too.

We have also tested the detection algorithm in adverse weather conditions (in a foggy and rainy day) where we have low contrast images and reflection artefacts (such as water drops). We have obtained similar results.

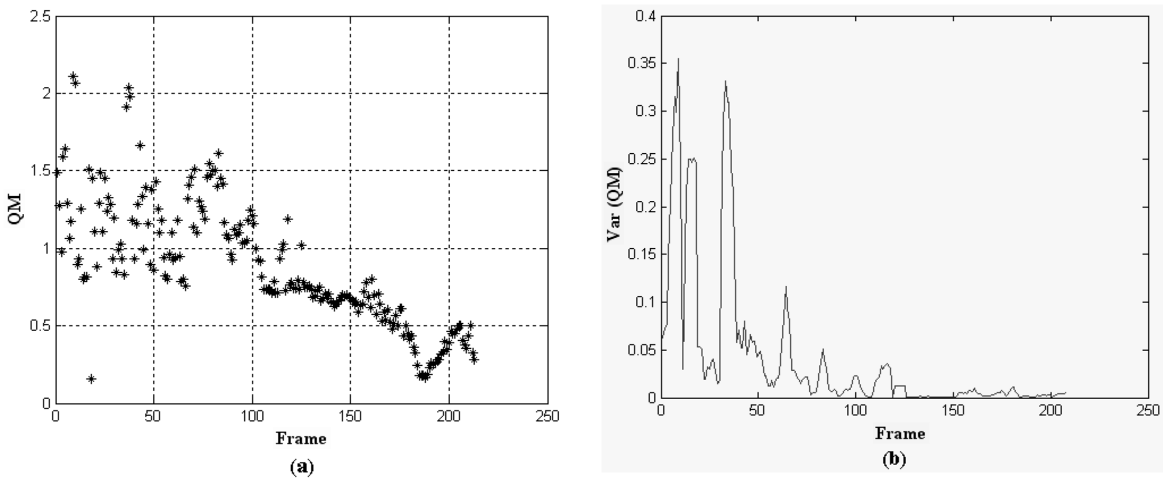


Figure 6. (a) QM of the detection along the slow overtaking sequence; **(b)** Variance of the QM along the sequence.

6. 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 uses 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 the overtaking vehicle (if present). This filtering technique is a robust scheme 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, and these motion patterns compete with opposite direction motion features. In this way, the moving features are processed in a competitive 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 being tested with sequences taken from a high dynamic range camera in different weather light conditions to evaluate its robustness to low contrast and reflection artefacts (such as water drops).

7. ACKNOWLEDGEMENTS

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

REFERENCES

Gavrila, D. M., Gern, A., Görzig, S., Janssen, R., Paetzold, F. and Wöhler, C. (2000) From door to door- Principles and Application on Computer Vision for driver assistant systems. IN: Proceeding of Intelligent Vehicle Technologies: Theory and Applications, Arnold.

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

Gonzalez R. & Woods R. (1992) Digital Image Processing. Addison-Wesley.

Handmann, U., Kalinke, T., Tzomakas, C., Werner, M., v. Seelen, W. (1998) Computer Vision for Driver Assistance Systems. IN: Proceeding of SPIE, vol. 3364, Session Enhance and Synthetic Vision, pp. 136-147.

Krapp H. G. (2000) Neuronal Matched filters for optic flow processing in flying insects. International review of neurobiology, 44, pp. 93-120.

Marr D. (1982) Vision. New York, W. H. Freeman and Company.

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

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