A Flexible Architecture for Intelligent Cruise Control

U. Handmann   I. Leefken   C. Tzomakas   W. v. Seelen
Institut für Neuroinformatik
Lehrstuhl für Theoretische Biologie
Ruhr-Universität Bochum
44780 Bochum, Germany

Abstract

In this paper we present a concept of a flexible and modular architecture for intelligent cruise control (ICC). The architecture can be subdivided into three different processing steps: the object-related analysis of sensor data, the behavior-based scene interpretation and the behavior planning. Each of it works on collected sensor information as well as on a knowledge base, which can be broadened by external knowledge like GPS and street maps. An implementation of the object-related analysis has been presented on the IV'98 in Stuttgart [1]. An intelligent car following system is described in this paper as a spin-off for behavior planning.

1 Introduction

The problems encountered building a driver assistance system are numerous. The collection of information about real environment by sensors is erroneous and incomplete. When the sensors are mounted on a moving observer it is difficult to find out whether a detected motion was caused by ego-motion or by an independent object moving. The collected data can be analyzed by several algorithms with different features designed for different tasks. To gain the demanded information their results have to be integrated and interpreted. In order to achieve an increase in reliability of information, stabilization over time and incorporation of knowledge about important features has to be applied.

Different solutions for driver assistance systems have been published. An approach proposed by Rossi et al. [2] showed an application for a security system. An application being tested on highways has been presented by Bertozzi and Broggi [3]. Dickmanns et al. presented a driving assistance system based on a 4D-approach [4]. Those systems were mainly designed for highway scenarios, while the software architecture presented by Franke and Görzig [5] has been tested in urban environment.

In contrast, the content of this paper concentrates on a flexible modular architecture of a driver assistance system working on evaluation and integration of the actual information gained from different sensors. The presented architecture is able to handle different tasks. New requirements to the system can be integrated easily. A spin-off is realized.

2 Architecture

The proposed architecture (fig. 1) is intended to produce different kinds of behavior according to given tasks. Information about the actual state of the environment is perceived by the system's sensors. The data collected by each sensor have to be processed and interpreted to gain the desired information for the actual task. This is done by the object-related analysis. It has to provide the scene interpretation with information. In the scene interpretation the partly redundant results have to be interpreted and integrated to achieve consistent information. The behavior relevant information has to be presented to the behavior planning. The behavior planning is the final element that has to evaluate which action should be taken to achieve the current task and which subtask has to be fulfilled based on the actual information from the scene interpretation and the actual knowledge. It also has to decide if the current decision or advice is reliable and can be proposed to the driver. The actual behavior planning should influence the scene interpretation to produce the optimal amount of information needed.

In the following sections the knowledge base, the object-related analysis, the scene interpretation, and the behavior planning are discussed in detail. As an example, an intelligent cruise control (ICC) is embedded in the given architecture. The ICC has to guide the driver according to a chosen object. This comprises advices for velocity adaptation to keep a secure distance, for changing lane and for choosing a new leading object, if the previous one is lost. In the shown example the only sensor applied is a visual sensor being mounted on the rear view mirror of the observing vehicle. Other sensors, like radar, could be integrated easily [6].

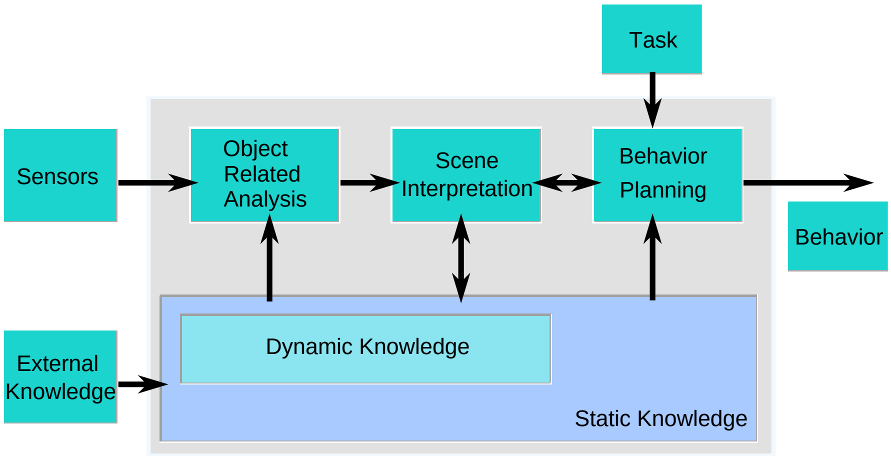

Figure 1: Architecture for a driver assistance system

3 Knowledge Base

In the knowledge base static and dynamic knowledge is represented. Static knowledge is known in advance independently of the scenery of movement (e.g. physical laws, traffic rules). Dynamic knowledge (e.g. actual traffic situation, scenery) is knowledge changing with the actual information or with the task to be performed (e.g. objects in front of the car). Dynamic knowledge can also be influenced by external knowledge like GPS-information.

4 Object-related Analysis

In the object-related analysis the sensor data are prepared for the scene interpretation. The structure is shown in fig. 2. This part of the architecture can be subdivided into a sensor information processing and a representational part. The sensor information processing is specialized for each sensor, while the representation performs the consistent integration of the processed sensor data in sensor coordinates over time. In the following an example of an object-related analysis for an ICC is described.

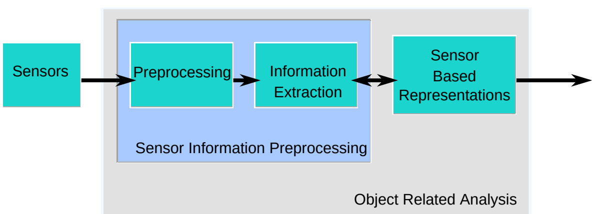

Figure 2: Structure of the object-related analysis

4.1 Sensor Information Processing

In the sensor information processing part the collected sensor data are preprocessed (e.g. segmentation, classification of regions of interest (ROI) or lane detection) and interpreted according to their capabilities. The processing can be performed for each sensor as well as information from different sensors can be fused. An implementation of an object-related analysis on vision data has been presented in [7, 1]. Objects are extracted by segmentation, classification and tracking. Results are shown in fig. 3. These object hypotheses are used to build movement sensitive representations to get detailed information about objects in front of the car.

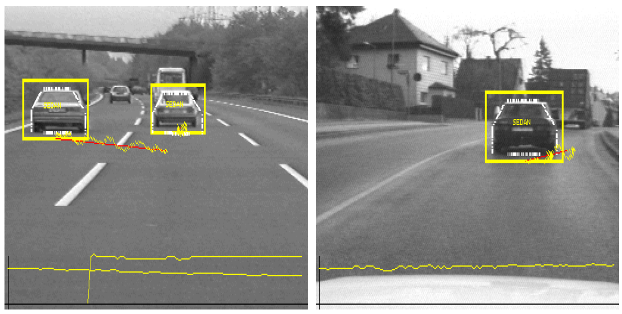

Figure 3: Vision-based object detection, object classification and object tracking

4.2 Sensor-based Representations

In the sensor-based representational part of the object-related analysis the data are combined consistently for each sensor. A representation in general can be subdivided into functional modules. It performs consistent integration of the processed sensor data over time. Each representation has a data integration and a knowledge integration module. An internal memory and internal dynamics have to be organized (fig. 4).

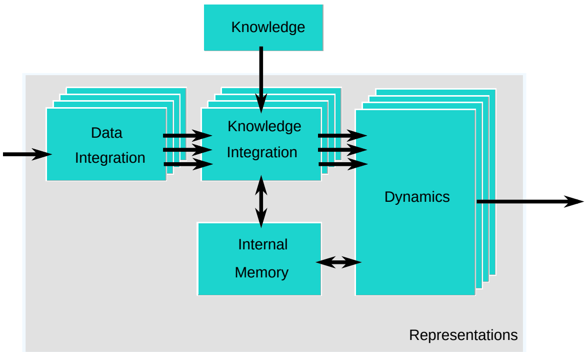

Figure 4: Structure of the representations

To be able to compare the results from different sensors each representation has to describe a common database on a comparable time scale (data integration). In this part the results of the information processing are evaluated in sensor coordinates according to consistency and discrepancy or ambiguity in information. The results of the sensor information processing are stabilized in movement sensitive representations by introducing a third dimension, the time dimension. In this sense a ROI is accepted as a valid hypothesis only if it has a consistent history. This can be implemented by temporal accumulation, based on the assumption that there is no abrupt change in the location of objects between subsequent frames (fig. 5).

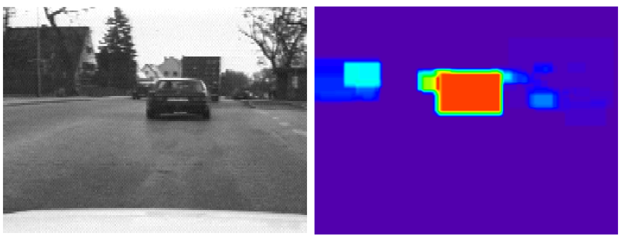

Figure 5: Image and representation

The accumulation results can be supported by a preactivation of a saliency map (fig. 6). This map is the result of a fusion process coupling texture and contour information by a multilayer perceptron [6].

This assumption could be verified for example for highway environments, where the relative velocity of ve-

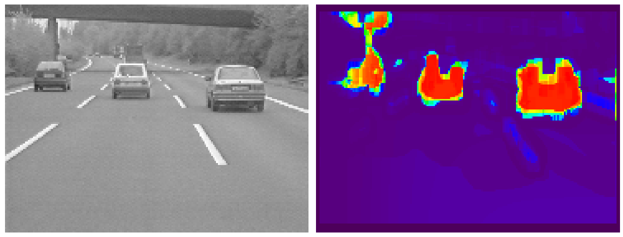

Figure 6: Image and saliency map

hicles is small. In the case of country road scenarios the locality assumption does not hold for all objects. Indeed, the high relative velocity of oncoming vehicles yields to abrupt movements between two subsequent scenes, especially when the vehicles approach the observer. The change in their position is a functional of their relative velocity and their distance to the observer. In order to apply a time stabilization to these regions and decide if they are valid or not, a prediction of their position in the knowledge integration part is realized.

Moreover this prediction can be useful for scene interpretation, since foregoing vehicles can be discriminated from oncoming ones. The prediction requires knowledge about the road trajectory. When the road boundaries can be localized in the image (e.g. either from GPS/road map information or by a vision-based approach like in [8]), then the trajectory of oncoming vehicles can be estimated, since it flows approximately parallel to the road boundaries. Indeed, assuming a trapezoid road model for small and middle distances (fig. 7), the non-parallel edges are defined by the line

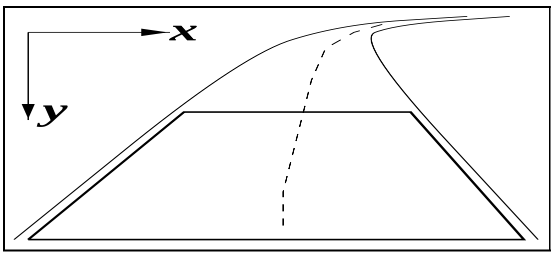

Figure 7: A simplified road model

equation $y = \lambda x + \xi$, where λ is the slope e.g. of the left boundary and ξ an offset parameter. Furthermore a detected object is assumed having a constant relative velocity v_r (regarding only translatory motion without rotational components) within the time interval of two successive frames. Within this interval Δt the running distance of the object is given by $s = v_r \Delta t$. Using the perspective geometry and assuming zero tilt for reasons of simplicity the motion of the object in the image plane from frame t_k to frame t_{k+1} is described by the following equation:

$$\Delta y'_k = \frac{-\Delta Y_k \cdot {y'_k}^2}{f_y \cdot H + \Delta Y_k \cdot y'_k} \quad \text{and} \quad \Delta x'_k = \frac{\lambda}{\Delta y'_k} \quad , \quad (1)$$

with $\Delta y'_k = y'_{k+1} - y'_k$, $\Delta Y_k = Y_{k+1} - Y_k$, and $\Delta x'_k = x'_{k+1} - x'_k$. For oncoming vehicles holds $Y_{k+1} < Y_k$ and $y'_{k+1} > y'_k$.

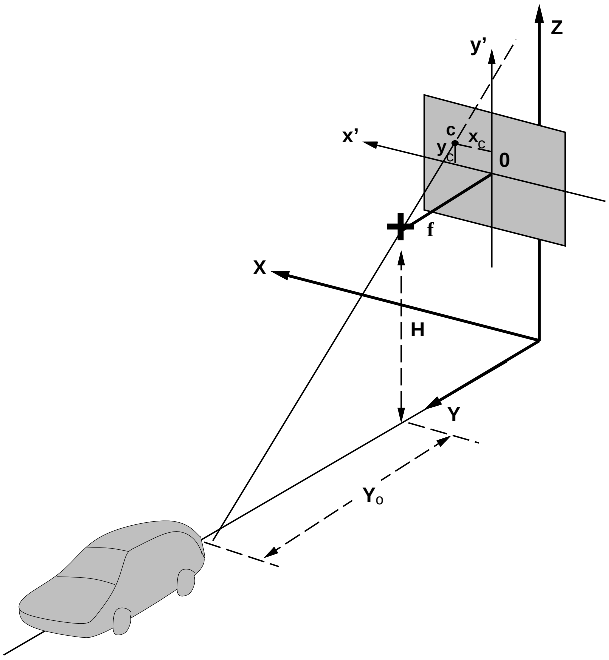

Figure 8: Camera geometry

H is the height of the camera, f_y the normalized focal length f to the camera parameters (fig. 8), x' and y' are the image coordinates, X, Y and Z are the world coordinates and k is the time index. Since the direction of movement is towards the observer it follows $s = -\Delta Y$ and equation (1) yields for the vertical translation of an oncoming vehicle:

$$\Delta y'_k = \frac{v_r \cdot \Delta t \cdot {y'_k}^2}{f_y \cdot H - v_r \cdot \Delta t \cdot y'_k} \quad . \quad (2)$$

That means, the translation in y-direction of an oncoming vehicle in the next frame depends on its translational velocity and its current position (height) in the image. Its lateral translation can be estimated as the ratio of the trajectory slope and the vertical translation. Thus, for each detected region its position is predicted for a finite number of frames and the predicted ROIs are registered in an accumulator field in a similar way as in the case of foregoing vehicles. The predicted ROIs are verified along with newer regions at each frame separately, and the total activity (volume) of a region within the time window is the criterion for the detection of oncoming objects (fig. 9).

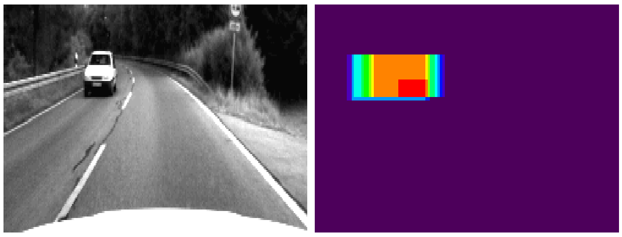

Figure 9: Prediction and object detection (oncoming)

The principle of the explained translation can be easily adapted for the prediction of oncoming vehicles on the right (overtaking task, fig. 10).

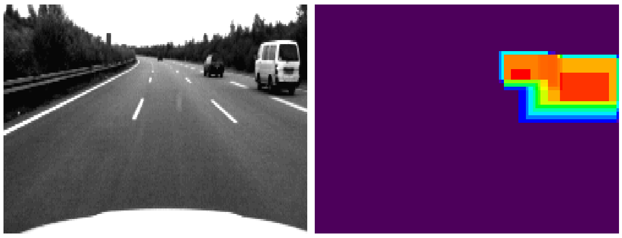

Figure 10: Prediction and object detection (overtaking)

5 Scene Interpretation

The scene interpretation interpretes and integrates the different sensor results to achieve consistent results. Behavior-relevant information is extracted. In [9] a perceptual architecture for scene interpretation based on a

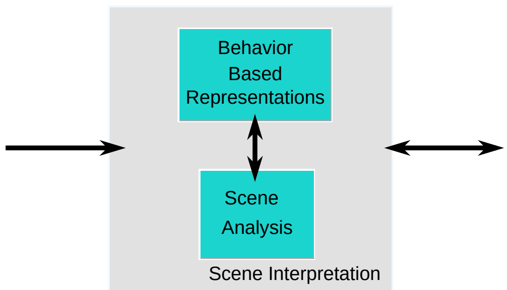

Figure 11: Scene Interpretation

4D-approach and Kalman-filtering is described. In contrast, in this paper the scene interpretation is behavior-based. It is subdivided into a behavior-based representational and a scene analysis part (fig. 11).

5.1 Behavior-based Representations

In the course of data integration of object hypotheses and lane information a transformation of the results to world coordinates with respect to the moving observer is realized. The positions of the detected objects are determined in a birds view perspective of the driving plane (fig. 12). The transformation rules follow the given position of the CCD camera in the car and the position of the car on the lane. The physical laws are given by the transformation equations for the camera (fig. 8) and physical considerations of the movement and position of potential objects. The transformation also depends on constant data (e.g. length of a vehicle according to its classification).

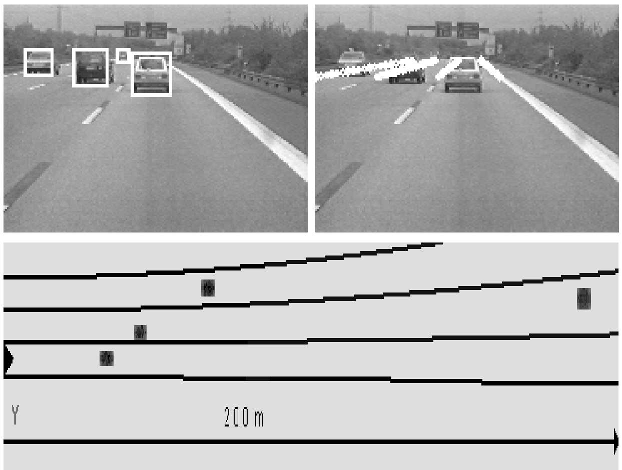

Figure 12: Image with objects and lanes and birds view perspective

The knowledge needed for the evaluation of the data and for information management is given by the efforts of the task of ICC, by physical laws and traffic rules. An improvement of the results can be achieved by information about the actual driving data like acceleration and steering angle, as well as GPS-information, street maps, and information about the current traffic conditions.

Regarding the task of an ICC a good performance is reached by gaining information about the actual object followed and about obstacles invading or occupying the actually needed driving space. The information about the leading vehicle comprises the distance to the vehicle, the lane information and its trajectory relative to the observer. This information is stored in terms of a list containing the properties of each detected object (internal memory). Those data (e.g. object class, position) can be used to show developments over time towards the actual situation, to work on different time scales, to give information about trajectories, to evaluate velocities of the objects, and to estimate the TTC (time to contact [10]).

5.2 Scene Analysis

The scene analysis sustains the ICC by evaluating the actual traffic condition as well as the scenery. According to the actual traffic condition and the planned behavior the risk-factor for actions is estimated. The determination of the traffic condition is performed by evaluating the information from scene interpretation. This is done by counting the objects, evaluating their relative speed and the movement according to their class. The scenario can be determined using GPS and street maps for investigating the kind of street, e.g. highway, country road or urban traffic. According to these scenarios different objectives have to be taken into consideration. The determined traffic condition as well as the classified scenario are proposed to the behavior planning.

6 Behavior Planning

The behavior planning depends on the given task and the scene interpretation. Different solutions for the planning task are possible. A rule-based fuzzy-logic approach is described in [11]. An expert system is shown in [12].

In the case of the ICC the first question is if there is a leading vehicle and if it has been tracked successfully and can be followed. The second one is the question if the observer can follow the tracked object safely.

Safety for the observer results in advices to the driver which are not only based on the intention to follow the leader but on regards concerning the safety of the own vehicle. This means that the object cannot be followed or might be lost in case of other object or obstacles endangering the observer. The signal behavior for the main tasks is determined by the flow diagram in fig. 13. This flow diagram represents a simplified version of a behavior planning.

At first it has to be evaluated if the leading object could be detected. If the leader could not be detected a new leader is searched automatically in the same lane. A preceding object is recognized as a leader if it has a consistent trajectory on the actual lane.

If no leader can be found the advices for deceleration or no change in action are given according to safety considerations. The distance to the preceding object is kept. If there is no preceding object below security distance and no object entering the actual lane below security distance no advice is given.

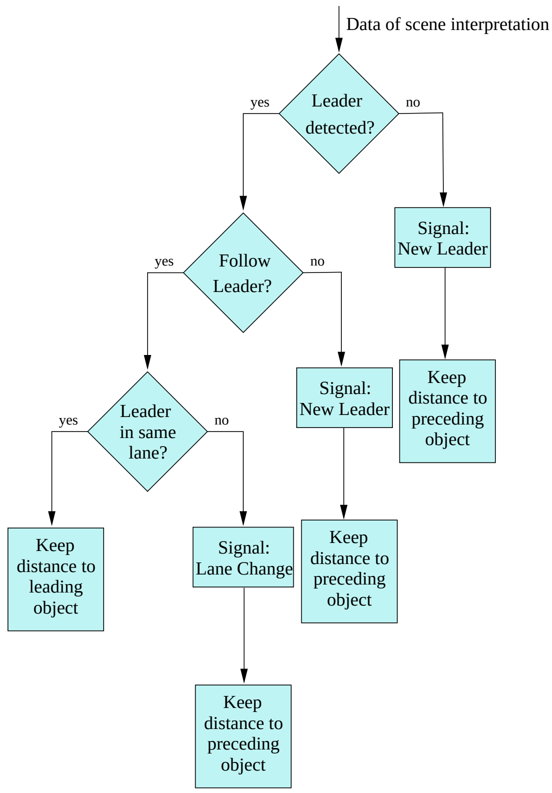

Figure 13: Flow diagram of signal behavior for ICC

In the case of a detection of the leader the question arises if the leaders' driving direction coincides with the intended one. If this is not given a new leader has to be chosen and the distances to the other objects have to be kept as mentioned before.

If the direction of the leader is corresponding to the intended one it is of interest if the leader can be located in the same lane. This decision is found by considering the lane occupancy and the trajectory of the leader. If the leader is detected in a different lane or if the trajectory of the leader points to a different lane an advice for lane change is given.

If no lane change can be performed the distance to the preceding object is kept. If the leader has been detected in the same lane the correct security distance has to be kept or reached by acceleration or deceleration.

7 Conclusion and Outlook

The proposed architecture shows flexibility for integrating different tasks in terms of a driver assistance system. Its applicability is demonstrated on the problem of ICC. The main advantages of the presented architecture are a flexible data integration structure, a task modularization, and multi-level (object-related and behavior-related) representations. Different sensors can be incorporated for the analysis and interpretation of the scene.

The part of the scene interpretation has to be broadened for a better specification of the actual sensed situation. The results will be incorporated into the decisions of the behavior planning. The behavior planning has to be expanded for several tasks and a method for choosing the optimal strategy will be developed to cope with different situations and goals. A dynamic adaptation is possible as well.

References

[1] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W. von Seelen, "An Image Processing System for Driver Assistance," in *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Stuttgart, Germany, 1998, pp. 481 – 486, IEEE.

[2] M. Rossi, M. Aste, R. Cattoni, and B. Caprile, "The IRST Driver's Assistance System," Technical Report 9611-01, Instituto per la Ricerca Scientificia e Technologica, Povo, Trento, Italy, 1996.

[3] M. Bertozzi and A. Broggi, "GOLD: a Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection," in *IEEE Transactions on Image Processing*, IEEE, Ed., 1997, vol. 4(2), pp. 114–136.

[4] E.D. Dickmanns et al., "Vehicles capable of dynamic vision," in *15th International Joint Conference on Artificial Intelligence (IJCAI)*, Nagoya, Japan, 1997, pp. 1–16.

[5] S. Goerzig and U. Franke, "ANTS - Intelligent Vision in Urban Traffic," in *IV'98, IEEE International Conference on Intelligent Vehicles 1998*. 1998, pp. 545–549, IEEE.

[6] U. Handmann, G. Lorenz, T. Schnitger, and W. von Seelen, "Fusion of Different Sensors and Algorithms for Segmentation," in *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Stuttgart, Germany, 1998, pp. 499 – 504, IEEE.

[7] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W. von Seelen, "Computer Vision for Driver Assistance Systems," in *Proceedings of SPIE Vol. 3364*, Orlando, 1998, pp. 136 – 147, SPIE, Session Enhanced and Synthetic Vision 1998.

[8] A. Broggi, "A Massively Parallel Approach to Real-Time Vision-Based Road Markings Detection," in *Proceedings of the Intelligent Vehicles '95 Symposium, Detroit, USA*, 1995, pp. 84–85.

[9] V. v. Holt and S. Baten, "Perceptual architecture for a vision system of autonomous vehicles," in *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Stuttgart, Germany, 1998, pp. 539 – 544, IEEE.

[10] D. Noll, M. Werner, and W. von Seelen, "Real-Time Vehicle Tracking and Classification," in *Proceedings of the Intelligent Vehicles '95 Symposium, Detroit, USA*, 1995, pp. 101–106.

[11] Qiang Zhuang, Martin Kreutz, and Jens Gayko, "Optimization of a Fuzzy System using Evolutionary Algorithms," in *Proceedings of the Fuzzy-Neuro Systems 98*, München, Germany, 1998, pp. 178 – 185.

[12] R. Sukthankar, *Situation Awareness for Tactical Driving*, Phd thesis, Carnegie Mellon University, Pittsburgh, PA, United States of America, 1997.