

ANALYSIS OF DYNAMIC SCENES

Uwe HANDMANN, Iris LEEFKEN, and Christian GOERICK

Institut für Neuroinformatik, Lehrstuhl für Theoretische Biologie

Ruhr-Universität Bochum, 44780 Bochum, Germany

E-mail: Uwe.Handmann@neuroinformatik.ruhr-uni-bochum.de

KEY WORDS: Dynamic Scenes, Computer Vision, Driver Assistance

ABSTRACT

In this paper the proposed architecture for a dynamic scene analysis is illustrated by a driver assistance system. To reduce the number of traffic accidents and to increase the drivers comfort, the thought of designing driver assistance systems rose in the past years. Principal problems are caused by having a moving observer (ego motion) in predominantly natural surroundings. In this paper we present a solution for a flexible architecture for a driver assistance system. The architecture can be subdivided into four different parts: the object-related analysis, the knowledge base, the behavior-based scene interpretation, and the behavior planning unit. The object-related analysis is fed with data by the sensors (vision, radar). The sensor data are preprocessed (flexible sensor fusion) and evaluated (saliency map) searching for object-related information (positions, types of objects, etc.). The knowledge base is represented by static and dynamic knowledge. It consists of a set of rules (traffic rules, physical laws), additional information (GPS, lane-information) and it is implicitly used by algorithms in the system. The scene interpretation combines the information extracted by the object-related analysis and inspects the information for contradictions. It is strongly connected to the behavior planning using only information needed for the actual task. In the scene interpretation consistent representations (i.e., bird's eye view) are organized and interpreted as well as a scene analysis is performed. The results of the scene interpretation are used for decision making in behavior planning, which is controlled by the actual task.

1 INTRODUCTION

Systems for automated image analysis are useful for a variety of tasks. Their importance is still growing due to technological advances and increased social acceptance. Especially driver assistance systems have reached a high level of sophistication. Fully or partly autonomously guided vehicles, particularly for road traffic, require highly reliable algorithms due to the conditions imposed by natural environments. At the *Institut für Neuroinformatik*, methods for analyzing driving relevant scenes by computer vision are developed in cooperation with several partners from the automobile industry. We present a system extracting important information from an image taken by a CCD camera installed at the rear-view mirror in a car.

The problems encountered in building a driver assistance system are numerous. The collection of information about real environments by sensors is error-prone 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, a stabilization over time and knowledge about important features have to be applied. Different solutions for driver assistance systems have been published. An approach published by (Rossi et al., 1996) showed an application for a warning system for security distance and lane-keeping. An application being tested on highways and being based on inverse perspective mapping has been presented by (Bertozzi and Broggi, 1998). (Dickmanns et al., 1997) presented a driving assistance system based on a 4D-approach. Those systems were mainly designed for highway scenarios, while the architecture presented by (Goerzig and Franke, 1998) has been tested in urban environment.

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 modules of the architecture are represented by the object-related analysis, the scene interpretation and the behavior planning. The accumulated knowledge is organized in the knowledge base. The presented architecture is able to handle different tasks. New requirements to the system can be integrated easily.

2 SYSTEM ARCHITECTURE

The proposed architecture is intended to produce different kinds of behavior according to given tasks. The structure is shown in fig. 1. 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 behavior planning. The behavior planning unit is the final element that has to evaluate which action should be taken to achieve the current task and which subtask has to be completed based on the actual information from the scene interpretation and the actual knowledge. It also has to decide whether the current decision or advice is reliable and can suggest the driver. The actual behavior planning should induce the scene interpretation to produce the optimal amount of information needed.

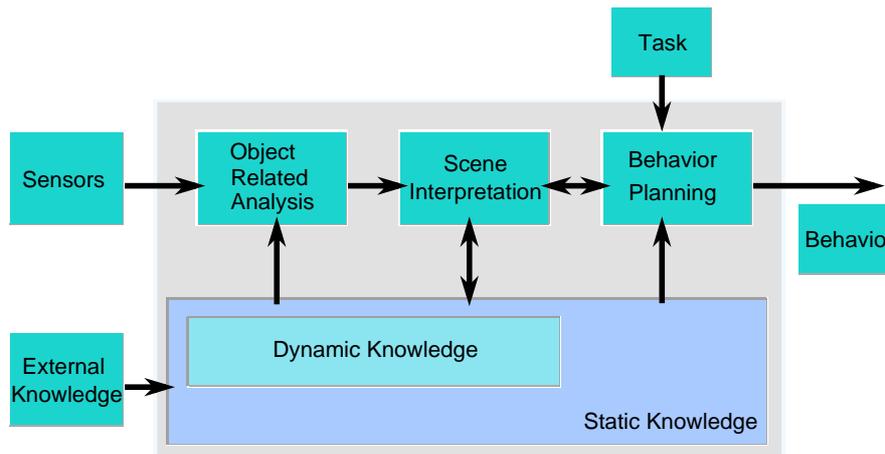


Figure 1: Architecture for a driver assistance system.

In the following sections, the object-related analysis, the knowledge base, the scene interpretation, and 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 advice for velocity adaptation so as to keep a secure distance, for changing lanes 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, as shown by (Handmann et al., 1998c).

3 OBJECT RELATED ANALYSIS

The object-related analysis can be subdivided into a sensor information processing and a representational part. The structure is shown in fig. 2. The sensor information processing is specialized for each sensor. In a preprocessing step, the sensor input is analyzed and relevant features for the sensor information processing are extracted. The representation performs the consistent integration of the processed sensor data in sensor coordinates over time.

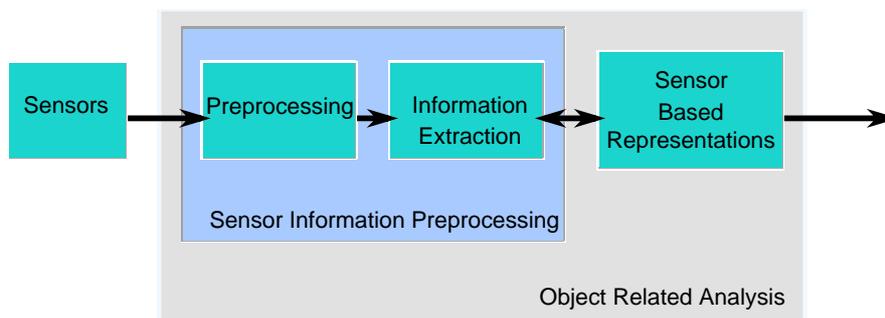


Figure 2: Structure of the object-related analysis.

3.1 Sensor Information Processing

In the sensor information processing part, the collected sensor data are preprocessed. Texture-, contour- and flow-field-analysis are realized which was proposed by (Handmann et al., 1998b). Relevant features for the sensor information processing (e.g., segmentation, classification of **Regions Of Interest (ROI)** or lane detection) are extracted and interpreted

according to their capabilities. The processing can be performed for each sensor, also information from different sensors may be fused as shown by (Handmann et al., 1998c). Objects are extracted by segmentation, classification and tracking (fig. 3).

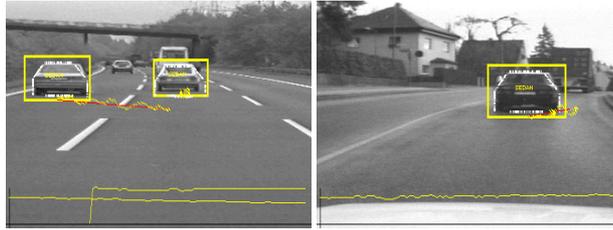


Figure 3: Vision-based object detection, object classification and object tracking.

3.2 Sensor-based Representations

The results of the sensor information processing stage are stabilized in movement-sensitive representations by introducing the time dimension, which was presented by (Handmann et al., 1999). In this sense, a ROI is accepted as a valid hypothesis only if it has a consistent history. This is implemented by spatio-temporal accumulation using different representations with predefined sensitivities. The sensitivities are functions of the objects' supposed relative velocity and of the distance to the observer (fig. 4). In order to apply a time stabilization to these regions and to decide whether they are valid or not, a prediction of their position in the knowledge integration part is realized. A competition between the different

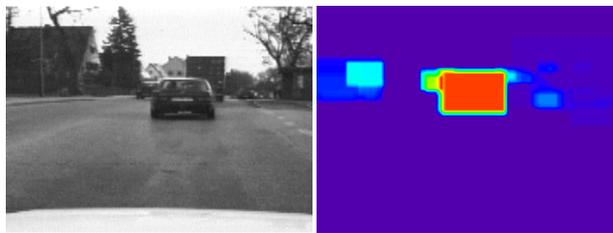


Figure 4: Image and representation. The value in the representation indicates the activation for an ROI.

representations and a winner-takes-all mechanism ensures reliable object detection. An implementation of an object-related analysis on vision data has been presented by (Handmann et al., 1998a, Handmann et al., 2000a). The results are passed to the scene interpretation.

4 KNOWLEDGE BASE

The knowledge needed for the evaluation of the data and for information management is determined by the specific task of driver assistance, physical laws and traffic rules. An improvement of the results can be achieved by the information from the 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, lane-information) is knowledge changing with the actual information or with the task to be performed (e.g., objects in front of the car). It can also be influenced by external knowledge like GPS-information. The accumulation results of the movement-sensitive representations in the object-related analysis, i.e., can be additionally supported by a preactivation depending on the dynamic knowledge (lane-information, fig. 5).

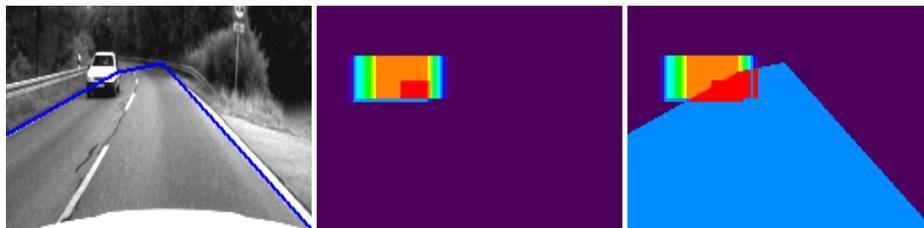


Figure 5: Oncoming traffic (image with lane-information, movement-sensitive representation without and with preactivation). The preactivated region is determined by the lane information. The value indicates the grade of activation.

5 SCENE INTERPRETATION

The scene interpretation interprets and integrates the different sensor results in order to extract consistent, behavior-based information. The scene interpretation is subdivided into a behavior-based representational and a scene analysis part (fig. 6).

5.1 Behavior-based representations

The data prepared by object-related analysis have to be integrated in order to detect and evaluate inconsistencies and discrepancies. The incoming data (in sensor coordinates) are transformed to a common description base in the data integration. Actually object- and lane-information are transformed to world coordinates with respect to the moving observer. The positions of the detected objects are determined in a bird's-eye view of the driving plane. The transformation rules follow the given position and direction of each sensor (e.g., the position and pitch angle of CCD-cameras in the car) and the position of the car in the lane. The sensor parameters are determined from the knowledge base (e.g., transformation equations for CCD-Cameras using a pinhole model presented by (Brauckmann, 1994)). Physical considerations concerning the movement and the position of potential objects are incorporated as well as constant parameters (e.g., length of a vehicle according to its classification). Explicit information from the knowledge base (e.g., physical rules and traffic rules, sensor evaluation factors) is applied in the knowledge integration. In this module, the synchronized information from the different sensors is coupled using knowledge. In this part e.g., ROI detected above the horizon are eliminated and lane positions are determined relative to the moving observer according to lane-information. The properly organized information is presented to the behavior-based representations of the situation as well as to an

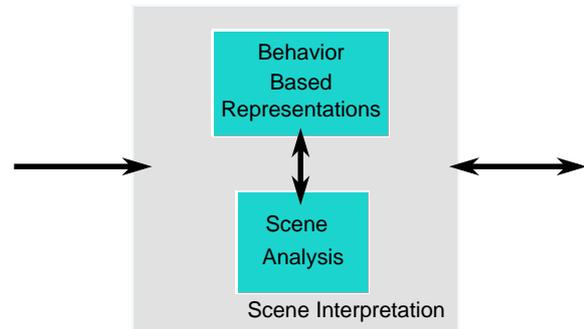


Figure 6: Structure of the scene interpretation.

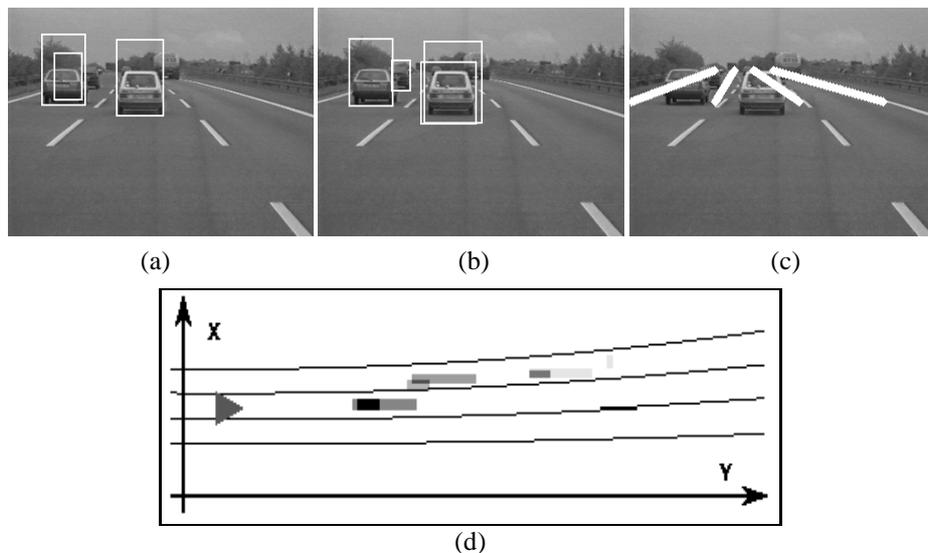


Figure 7: Image with segmentation results (a), tracking results (b), lane information (c) and bird's-eye view (d).

internal memory implemented as an object list. The data representations comprise a bird's-eye view representation, a representation containing free driving space and trajectory information as well as object-related information if the actual task requires it. Those representations are organized dynamically for stabilization of data and for performance of planning tasks. An example for a dynamically organized bird's-eye view representation is shown in fig. 7(d). The results of the object-related analysis (segmentation, object tracking) and the lane-information are integrated to build the bird's-eye view representation.

The internal memory is fed by the results of knowledge integration as well as by results of the dynamics of the representations. The object list in the internal memory contains information of previously detected objects to enable time-stabilized object detection and to determine object-related results. According to the data of the objects accumulated over time and sustained by the representation-dynamics, an object-related **Time To Contact Value** (TTC mentioned by (Noll et al., 1995)). The object and the observer are supposed to collide if they would occupy the same space at the same time. The space is determined by the intersection of the estimated trajectories of the object edges.

The prepared data (e.g., bird's-eye view, TTC-estimation, object-list) are passed to the scene analysis for determination of global measures and to behavior planning.

5.2 Scene Analysis

The scene analysis supports the driver assistance by evaluating the actual traffic condition as well as the scenery. The traffic condition is assessed by evaluating the information from scene interpretation. According to the actual traffic condition and the planned behavior a risk-factor

$$r = \sum_{i=1}^n \frac{|\vec{v}_{rel,i}|}{d_i} \quad (1)$$

for actions is estimated. The risk factor evaluates the sum over all relative velocities $|\vec{v}_{rel,i}|$ between observer and objects normalized by the distances d_i to each of the n objects detected in the scene. The scenario can be determined using GPS and street maps for investigating the kind of street, e.g., highway, country road or urban traffic. Depending on these scenarios, different objectives have to be taken into consideration. The determined risk factor as well as the classified scenario are passed to behavior planning. The scene analysis does not have to be performed for every image, as e.g., the kind of road does not change abruptly, so the evaluated data are stored in the knowledge base so as to remain available to behavior planning.

6 BEHAVIOR PLANNING

Behavior planning depends on the given task and on the scene interpretation. Different solutions for the planning task are possible. A rule-based fuzzy logic approach is described by (Zhuang et al., 1998). An expert system is shown by (Sukthakar, 1997). At the *Institut für Neuroinformatik* a system with complex behavior by means of dynamical systems for an anthropomorphic robot was realized by (Bergener et al., 1999). A behavior control for autonomous vehicles controlling the steering angle and the velocity by neural field dynamics (Amari, 1983) was presented by (Handmann et al., 2000b). In the present system, a flow diagram controlled behavior planning realizing an ICC was integrated.

Behavior planning for the observer results in advice to the driver not only based on the intention of following the leader but also taking into account the safety of the own vehicle. This means that the object cannot be followed or might be lost in case of other objects or obstacles endangering the observer. The signal behavior for the main tasks is described by the flow diagram shown in fig. 8.

At first it has to be determined whether the leading object could be detected. If not, a new one 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, is positioned in a reasonable (detectable) distance and if it could be detected at least once in a certain number of previous time-steps. If no leader can be found, advice for deceleration or no change in action is given according to safety considerations. A safe and comfortable driving velocity has to be achieved by taking the security distance, the TTC to other objects and possible collision positions as well as a general risk factor (eq. 1) into account. In the case of leader detection it is of interest whether the leader can be located in the same lane.

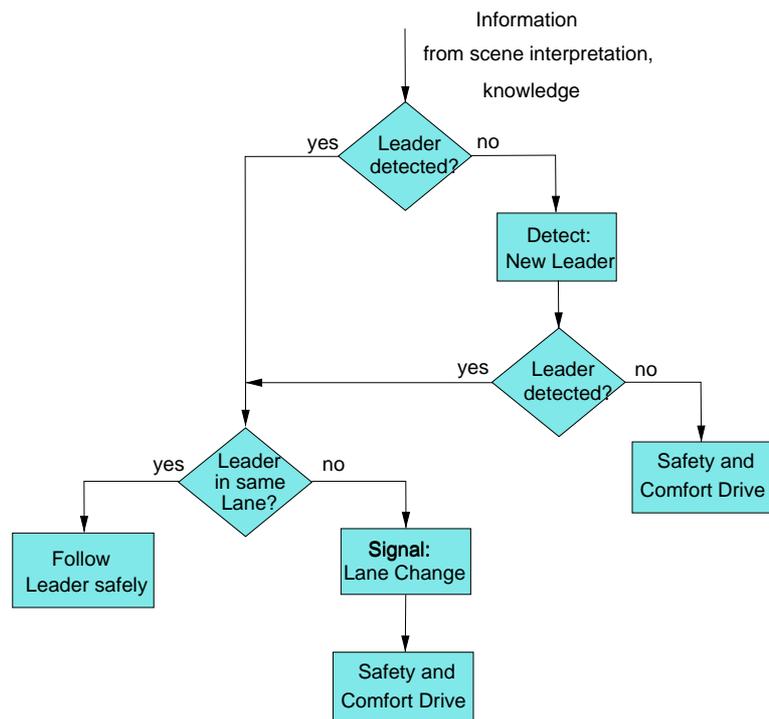


Figure 8: Flow diagram of signal behavior for ICC.

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, advices for a safe and comfortable ride are given. If the leader has been detected in the same lane, the correct security distance has to be kept or reached by acceleration or deceleration. In this case safety and comfort also have to be considered.

7 SIMULATION ENVIRONMENT

To test different assistance systems with different sensor constellations in different traffic conditions a simulation environment was realized. In fig. 9 results of test environment are shown. The bird's eye view of the sensor constellation on a vehicle (black) with three cameras (white lines) and two radar sensors (black lines) is shown in fig. 9 (a). The appropriate sensor output is shown in fig. 9 (b)-(f). For the ICC the camera and the radar sensor, both mounted in the front of the car, are used. Fig. 9 (b) shows the output of the camera (*aperture* = 28.0724°). Fig. 9 (e) illustrates the relative velocity [$\frac{m}{s}$] of the object hypotheses of the radar sensor (x-coordinate: distance [m], y-coordinate: angle [deg]).

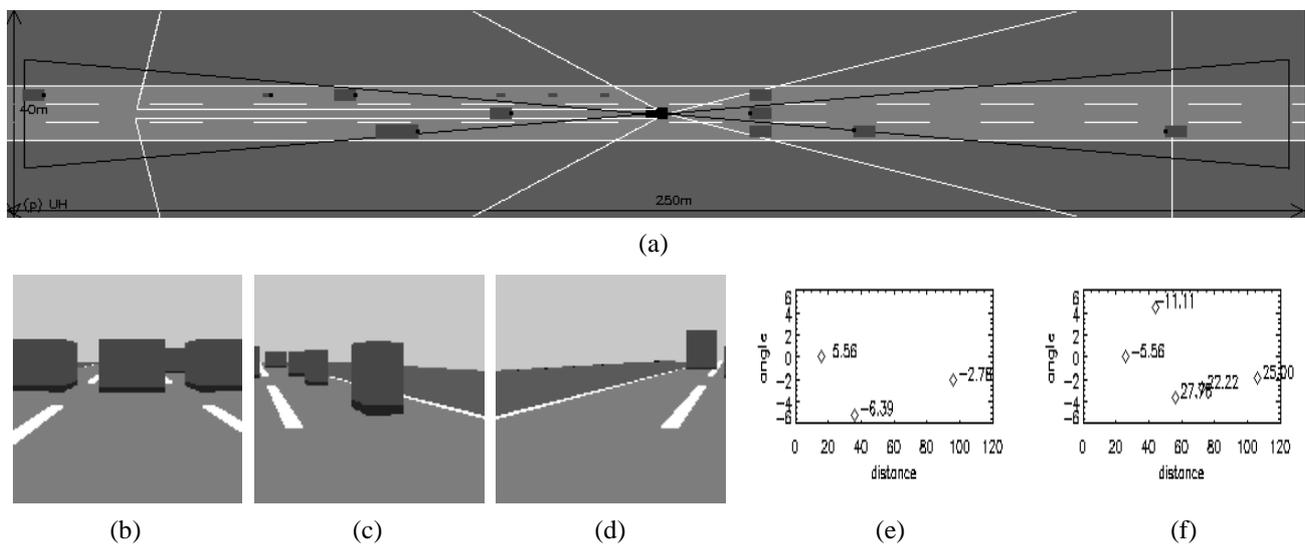


Figure 9: Sensor simulation: (a) bird's-eye view with mounted sensors on the vehicle; (b)-(d) camera output of the three cameras; (e)-(f) plotted results of the radar sensor output.

Other assistance systems, e.g., blind-spot observer, can be realized with different combinations of the mounted sensors (fig. 9 (a)-(f)) in the vehicle.

8 RESULTS

The proposed results have been gained from a visual sensor mounted on the rear view mirror of a car. The sensor data were collected on a German highway. In fig. 10 the results of a scene with 1000 frames are shown. Four frames showing special situations have been chosen to present the main characteristics of the system.

The first row (a)-(d) contains the segmentation results of the object-related analysis. In this part of the system ROI are extracted. The second row (e)-(h) shows the tracking results of the object-related analysis. In image (e) the necessity of the tracking module becomes obvious. An object not being found by the segmentation module (a) is detected by the tracking module (e) because of its history. In row three (i)-(l) the lane information provided by the knowledge base is shown. The three lanes of the highway are mapped from image coordinates to world coordinates. The bird's-eye view representations of the scene interpretation are shown in (m)-(p). The observing vehicle is symbolized by a dark triangle. The dimensions of the representation correspond to $[x, y] = [32m, 110m]$ in world coordinates. The observer is located at the point $[16.0m, 10.0m]$. The observed objects are mapped to world coordinates according to the results of the object-related analysis. The results of the behavior planning module are presented in (q)-(t). The white dot in the images represents the actually chosen leader. If the leader has been chosen, the object is followed and a lane change is advised in case of a lane change of the leader (s).

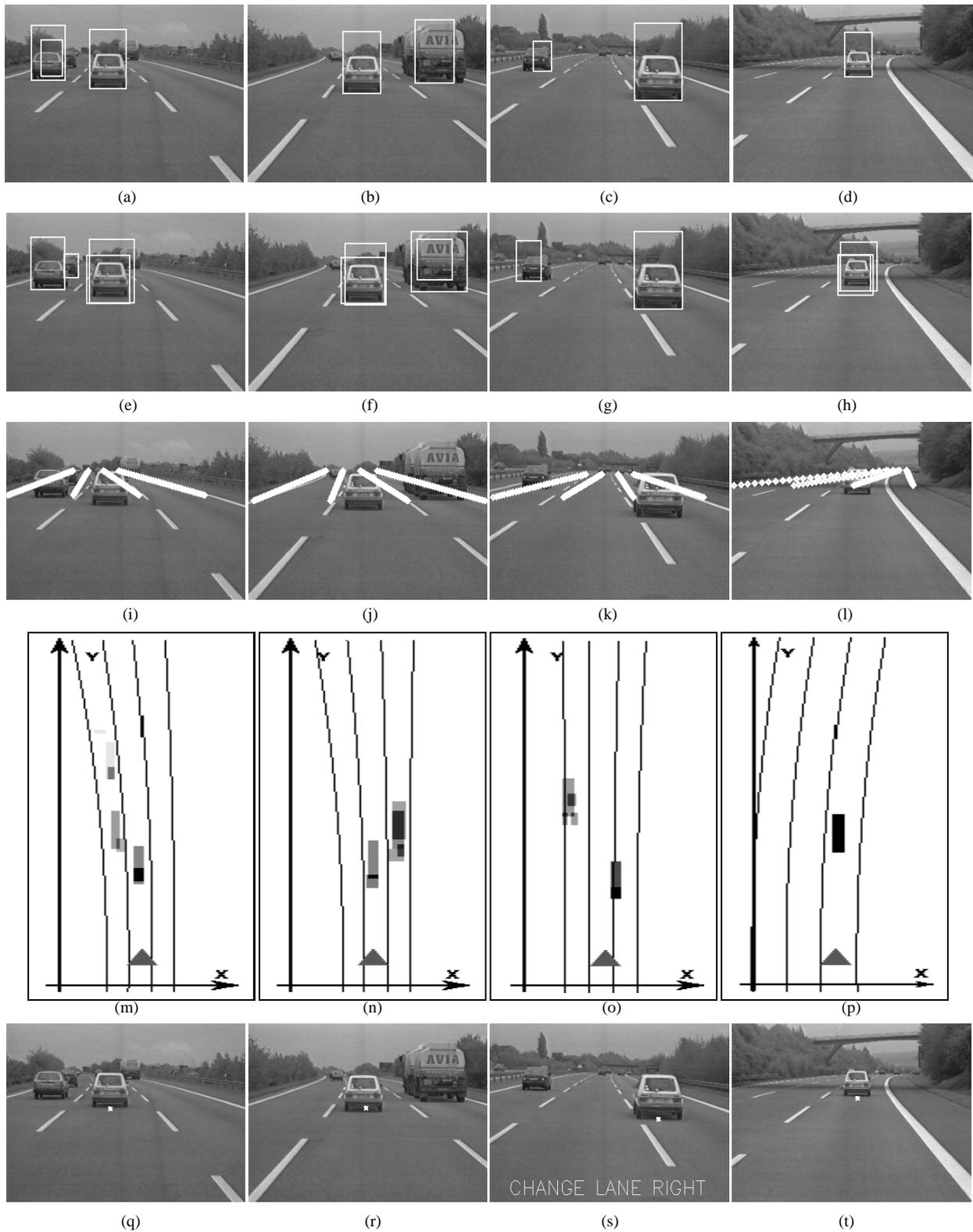


Figure 10: System results for the task ICC. Object-related analysis: segmentation results (a)-(d), tracking results (e)-(h); Knowledge base: lane information (i)-(l); Scene interpretation: bird's eye view (m)-(p); Behavior planning: (q)-(t).

9 CONCLUSIONS

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, task modularization, and multi-level (object-related and behavior-related) representations. Different

sensors (e.g., Radar, Lidar and image processing) can be incorporated for the analysis and interpretation of the scene to increase the quality of the system. In the part of the scene interpretation, neural field dynamics (Amari, 1983) will be used to stabilize the given object hypotheses with an adequate interaction kernel. For behavior planning a dynamical system is supposed to be realized as well. A successful solution to a behavior planning using this method was presented by (Handmann et al., 2000b).

REFERENCES

- Amari, S.-I., 1983. Field Theory of Self-Organizing Neural Nets. IEEE Transactions on Systems, Man, and Cybernetics.
- Bergener, T., Bruckhoff, C., Dahm, P., Janßen, H., Joublin, F., Menzner, R., Steinhage, A. and von Seelen, W., 1999. Complex Behavior by means of Dynamical Systems for an Anthropomorphic Robot. special issue of Neural Networks.
- Bertozzi, M. and Broggi, A., 1998. GOLD: a Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection. In: IEEE (ed.), IEEE Transactions on Image Processing, Vol. 7(1), pp. 62–81.
- Brauckmann, M. E., 1994. Visuelle Automobilerkennung zur Fahrzeugführung im Straßenverkehr. Fortschrittberichte, Reihe 12, VDI Verlag.
- Dickmanns, E. et al., 1997. Vehicles capable of dynamic vision. In: 15th International Joint Conference on Artificial Intelligence (IJCAI), Nagoya, Japan, pp. 1–16.
- Goerzig, S. and Franke, U., 1998. ANTS - Intelligent Vision in Urban Traffic. In: IV'98, IEEE International Conference on Intelligent Vehicles 1998, IEEE, Stuttgart, Germany, pp. 545–549.
- Handmann, U., Kalinke, T., Tzomakas, C., Werner, M. and von Seelen, W., 1998a. An Image Processing System for Driver Assistance. In: IV'98, IEEE International Conference on Intelligent Vehicles 1998, IEEE, Stuttgart, Germany, pp. 481 – 486.
- Handmann, U., Kalinke, T., Tzomakas, C., Werner, M. and von Seelen, W., 1998b. Computer Vision for Driver Assistance Systems. In: Proceedings of SPIE Vol. 3364, SPIE, Orlando, pp. 136 – 147. Session Enhanced and Synthetic Vision 1998.
- Handmann, U., Kalinke, T., Tzomakas, C., Werner, M. and von Seelen, W., 2000a. An Image Processing System for Driver Assistance. Image and Vision Computing (Elsevier) 18(5), pp. 367 – 376.
- Handmann, U., Leefken, I., Steinhage, A. and v.Seelen, W., 2000b. Behavior Planning for Driver Assistance using Neural Field Dynamics. In: NC 2000, second international symposium 'Neural Computation', Berlin, Germany.
- Handmann, U., Leefken, I., Tzomakas, C. and von Seelen, W., 1999. A Flexible Architecture for Intelligent Cruise Control. In: ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999, IEEE, Tokyo, Japan, pp. 959 – 963.
- Handmann, U., Lorenz, G., Schnitger, T. and von Seelen, W., 1998c. Fusion of Different Sensors and Algorithms for Segmentation. In: IV'98, IEEE International Conference on Intelligent Vehicles 1998, IEEE, Stuttgart, Germany, pp. 499 – 504.
- Noll, D., Werner, M. and von Seelen, W., 1995. Real-Time Vehicle Tracking and Classification. In: Proceedings of the Intelligent Vehicles '95 Symposium, Detroit, USA, pp. 101–106.
- Rossi, M., Aste, M., Cattoni, R. and Caprile, B., 1996. The IRST Driver's Assistance System. Technical Report 9611-01, Istituto per la Ricerca Scientifica e Tecnologica, Povo, Trento, Italy.
- Sukthankar, R., 1997. Situation Awareness for Tactical Driving. Phd thesis, Carnegie Mellon University, Pittsburgh, PA, United States of America.
- Zhuang, Q., Gayko, J. and Kreutz, M., 1998. Optimization of a Fuzzy Controller for a Driver Assistant System. In: Proceedings of the Fuzzy-Neuro Systems 98, München, Germany, pp. 376 – 382.