SCENE ANALYSIS AND ORGANIZATION OF BEHAVIOR IN DRIVER ASSISTANCE SYSTEMS

W. von Seelen, C. Curio, J. Gayko, U. Handmann, and T. Kalinke

Institut für Neuroinformatik, Ruhr-Universität Bochum, institut@neuroinformatik.ruhr-uni-bochum.de

Abstract

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. Fully or partly autonomously guided vehicles, particularly for road traffic, pose high demands on the development of reliable algorithms. Principal problems are caused by having a moving observer in predominantly 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. In this paper we present a solution for a driver assistance system. We concentrate on the aspects of video-based scene analysis and organization of behavior.

1. INTRODUCTION

Systems for automated image analysis are useful for a variety of tasks and their importance is still growing due to technological advances and an increase of social acceptance. Especially in the field of driver assistance systems the progress in science has reached a level of high performance [4, 2, 7, 5]. The presented paper is focused on a vision-based scene analysis and a behavior organization.

2. SCENE ANALYSIS

We structure the scene analysis as a recognition process of scene elements. In our context of video-based driver assistance scene elements are all informations necessary to organize a reliable behavior for the selected driving task, e.g. intelligent cruise control. So, our aim is not the description of the whole world, but the extraction of object information that we call object attributes. Object attributes are the relative position, relative velocity of the objects and the object type. The three object attributes for scene characterization and behavior organization have to be extracted by different image processing modules. The relative position is determined by an initial detection process. Using tracking and classification algorithms these initial object hypotheses are verified and their relative velocity and their object type are determined. Additionally the road geometry and the free driving space are used [9].

2.1. Object Detection

The main motivation of using multiple methods is that the development of one basic method solving all conceivable scenarios seems to be impossible. Therefore in order to provide reliable results and to ensure a fast and robust processing a coupling of *specialists* is carried out. For the detection process three main methods are used: a fusion process of simple features, a symmetry-based detection and a model match. In the first instance when no information about the object positions is available a fusion of different image features is done using a neural network. Data fusion is one of the main goals to be achieved if a large amount of stability and reliability is necessary like in this application of driver assistance systems [5]. On one hand a gain in robustness is reached by creating higher redundancy so that poor or missing results of one data stream do not affect the overall result decisively. On the other hand the varying types of objects and background constellations demand a large spectrum of data to be processed to solve the given task. The high flexibility, the facility of expansion and the adaptive retraining processes have led to the choice of neural networks. Image features are contour informations based on the local orientation coding (LOC) and polygon approximation of contours, texture analysis based on local image entropy, local variance analysis, and local cooccurrence measures, and color analysis [5]. The aim of this method is to reduce the large amount of information. The areas where the result neuron indicates a reliable structure are analyzed (fig. 1).

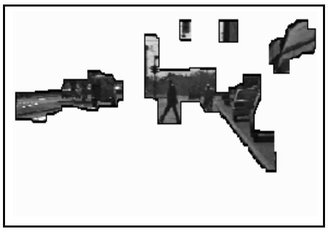

Fig. 1. Fusion of image preprocessing results.

After the reduction of the whole image amount two methods to extract the object positions are used in combination: at first a symmetry analysis and secondly a model matching. Both methods are applied to all object types with different parameters and models. The first method is used to detect the rear, front and side views of all object types by a measurement of the inherent vertical symmetric structure. The detection of symmetry is a challenging task under real condition. The method has to be robust against changes in illuminations and slight differences of the right and left part of the object. So, the symmetry detection is formulated as an optimization problem solved by a neural network (fig. 2 left,[9]). At last the model matching method is used to back up the fusion and symmetry results. Here, we use two different models for the first group (sedans, trucks and motorbikes) and the second group (the pedestrians) of objects. To detect trucks, sedans and motorbikes a shadow analysis in combination with a model matching based on the Hausdorff distance is used [5]. The detection of shadows is realized by thresholding the intensity image, some morphological processing and a region clustering stabilized over time. To increase the estimation goodness a template matching using a model representing the common downer shape of the vehicles shaped as an ”U” is realized. Figure 2 right shows the results of the object detection process. Beside some mismatches all solution relevant objects are detected. An object tracker and type classifier are applied for final decision making.

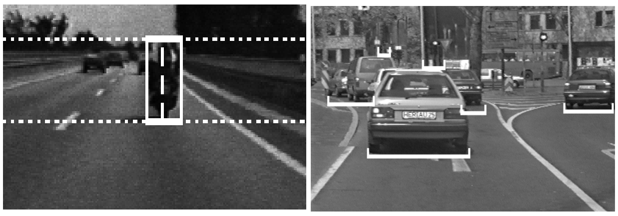

Fig. 2. Symmetry detection and ”U”-detection results.

The second group of objects to be detected, the pedestrians, are the most difficult to detect because their manifestation is of a higher variety. The model matching detection of pedestrians is performed similarly to the first group. Just the type of model is changed. We restrict the detection of pedestrians to the lower part of their body (i.e. hip and legs), since the torsos of the body typically shows huge variance in its form (e.g. due to the arm movements). Thus, the model resembles approximately a down-faced ”V”-form in different deformations corresponding to the various phases of a gait. Fig. 3 shows that from a subset of these discrete phases of the human walking model a detection of pedestrians via model matching is performed. In the left image all reliable models are displayed in the intensity image. The right image shows the finally detected regions.

Fig. 3. Detection of pedestrians (model matching).

2.2. Object Tracking

Algorithms for object tracking are the most important if a stabilization over time or a prediction of e.g. trajectories is demanded. Beside this, the relative velocity can be determined using the temporal change of the object size in the image. For rigid objects a tracking algorithm based on the Hausdorff distance is used as a measurement based on contour codes (LOC). This tracker has been tested successfully on a large set of different image sequences [5]. Supplementary for the pedestrians the texture based cross entropy tracking provides optimally results. The geometric comparison of shapes is a fundamental tool for model-based object recognition. The Hausdorff distance measures the divergence of a set of features with respect to a reference set of features [8]. These sets mostly describe object contours in our application. It is able to recognize vehicles and track them over a period of time. Two degrees of freedom were considered in our schema: translation and scaling of models to estimate the relative position and velocity [5]. One of the commonly used description of textures is obtained by the cooccurrence matrices. Especially non-rigid objects like pedestrians and two-wheeled vehicles that consist of a further rotational degree of freedom compared to other road-users can be tracked using the cross entropy. As described in [9] a matching process can be performed by comparison of two probability distributions based on the Kullback-Leibler divergence. In our application a model distribution D at time step $(t-1)$ is compared to several hypotheses at time t in the space of search (translation, scale, and deformation).

2.3. Classifiers for Traffic Participants

For the task of classification different methods are applied. Feature based and model based solutions have been developed [5]. For sedans, trucks and motorbikes both methods the LOC classifier (feature based) and the Hausdorff classifier (model based) can be applied efficiently. The LOC classifier is aimed at separating possible objects from the background. It is independent from the resolution of the objects due to a normalization in size. The model based Hausdorff classifier processes objects using the knowledge about their shape so that the region of interest detected and tracked can be enhanced in size and place. The LOC classifier is able to generalize from relative few training examples to the necessary features characterizing a vehicle by using a neural network for solving the classification task.

Furthermore, the geometric property of the Hausdorff distance leads to the idea of classifying various vehicles into separate classes according to the imposed dissimilarity measure. Because of the need of defining a reference contour for each class we deal with a model based approach. The design of accurate models (prototypes) is of great importance for our task. At a first step, the Hausdorff distance is used for the classification of sedans, trucks, motorbikes and pedestrians using different model databases. Each region is compared with all models of the data base. The contour features of the region and the models have been extracted using the LOC choosing different codes depending on the model under inspection. For more robust results the horizontal features are separated from the verticals, for both the region and the models. The Hausdorff distance is computed for each model over all possible translations inside the region, and a certain range of scales and deformations. The fractions of the features of the forward (model to image features) and the backward (image to model features) match that verify a given distance threshold constitute the criteria for its classification for each model. These values are learned by a neural network.

For the classification of pedestrians, we use a Hausdorff classifier with a set of phases of the gait and a feature based classifier analyzing the limb motion over time. Here, we use the periodic motion of the legs that is typical for detecting walking pedestrians are used additionally for the final classification to be most reliable in a large array of different scenarios [3]. The characteristic sequence of movement patterns evolving from the relative limb movement between realigned images enables one to classify walking pedestrians and to distinguish them from other tracked parts in the image not showing this characteristic motion pattern. A cross-correlation between one period ($T_p \approx 12$ frames) of model motion data and a window of the same size of real motion data is done. In order to provide unambiguous results a whole period of model motion data is necessary.

2.4. Object Attributes

As shown before, a scene analysis for driver assistance systems is realized by using a combination of different methods. The scene elements object positions, object trajectories including the object sizes over time and the object types are shown in fig. 4. With this information the object attributes relative position (distance d_{obj} and relative angle ψ_{obj}) and relative velocity Δv_{obj} are calculated. In fig. 4 the dark line indicates a lane change of the left sedan. The sedan in front has a relative velocity $\Delta v_{obj} \approx 0$. The classification results are also illustrated. Based on these informations an organization of behavior is realized.

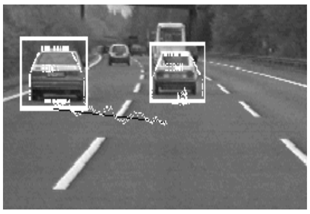

Fig. 4. Vision-based object detection, classification and tracking.

3. ORGANIZATION OF BEHAVIOR

The behavior control of a vehicle in real world traffic is a difficult problem to be solved, but finally the steering angle and the acceleration can only be influenced. In the presented driver assistance system the steering angle is calculated evaluating the dynamics of a one-dimensional neural field [6]. The acceleration is determined by a fuzzy controller optimized using an evolutionary algorithm [10]. The input of the behavior control results from the scene analysis. The distance d_{obj}, the relative angle ψ_{obj} and the relative velocity Δv_{obj} of other objects are used.

3.1. Neural Field

A neural field is a nonlinear dynamic system [1]. Originally it was introduced as a model of the neurophysiology of cortical processes. The field equation is given by

$$\begin{aligned} \tau \dot{u}(\psi, t) &= -u(\psi, t) + h + S(\psi, t) \\ &\quad + \int_{\Gamma} w(\psi - \psi') \varphi(u(\psi', t)) d\psi', \quad (1) \end{aligned}$$

where $u(\psi, t)$ is the field excitation at time t ($t \geq 0$) at the position $\psi \in \mathbb{R}$. The position ψ characterizes the field-site relative (steering angle) to a reference position $\psi = 0$. The excitation u varies with the time constant τ ($\tau \in \mathbb{R}^+$). A constant preactivation of the field is achieved by h. The stimulus $S(\psi, t) \in \mathbb{R}$ represents the input of the field. An interaction between u at position ψ and the excitation of the neighborhood is achieved by the convolution of an interaction kernel $w(\psi - \psi')$ over the set Γ and a nonlinear activation function $\varphi(u)$. The type of equilibrium solution depends on S, h and w. In the given task a peak solution is analyzed.

The main advantage of the neural field is the additive composition of the stimulus. The field can be stimulated starting with less information which can be additively broadened as more relevant information is obtained and is formulated in terms of the field-variable. The data for the field stimulus have to be coded adequately with respect to the effect they are supposed to have on the field activation.

The used field is influenced by the position and velocity of other traffic participants (especially the guiding vehicle), by information describing the free driving space and by lane information. According to this, S is determined by three stimuli-functions: The danger estimate $S_{\mathcal{O}}(\psi, t)$, street-course-factor $S_{\mathcal{L}}(\psi, t)$ and the leader-behavior $S_{\mathcal{D}}(\psi, t)$. The stimulus S of the field for the steering angle at time t is then determined by $S(\psi, t) = -S_{\mathcal{O}}(\psi, t) + S_{\mathcal{L}}(\psi, t) + S_{\mathcal{D}}(\psi, t)$. The field equation for the neural field controlling the steering angle is described by eqn. 1. The time constant τ is chosen according to the time scale on which the field is supposed to react on the stimulus. The preactivation is set to $h = -1$. The interaction kernel w is realized as Mexican Hat function (concentration). As nonlinearity $\varphi(u)$ a sigmoide function is used.

The evaluation of u is performed by determining the position of the maximum $\mathcal{N}(t) = \arg\max_{\psi} u(\psi, t)$ for the change in steering angle. To examine the behavior of the designed system different traffic scenes were tested by a simulation program. The required parameters are determined by evaluating the reaction of the system for a variety of scenes. The result of the field state for a defined scene at time t_0 is shown in fig. 5 left (solid line). According to the given data the stimuli of the field are shown as dashed lines.

The change of the steering angle over time according to field excitation is given in fig. 5 right. The vehicle drives through a right curve while keeping the lane and following

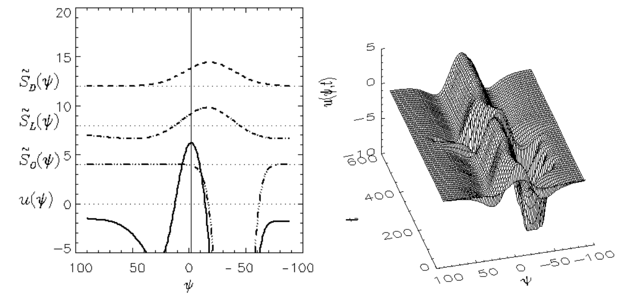

Fig. 5. Excitation $u(\psi, t)$ of the neural field (left). Additionally the stimuli according to objects, lane and leader are presented (shifted). Results over time (right).

a leader. The field-excitation shows negative values at the positions of objects to be avoided and positive values at angle positions to be favored. The maximum of u shifts to the left over time as long as the parking vehicles are detected, to keep a security distance. While controlling the steering angle oncoming traffic is taken into account (negative values at positive angle positions).

3.2. Fuzzy Controller

The task of the fuzzy controller is to control the distance to a vehicle driving in front. We are using the acceleration as a control variable. If there is no vehicle driving in front the system accelerates until a predefined speed is reached. Since the desired behavior of the controller could be described in linguistic rules, we decided to use a fuzzy rule controller. An initial controller was designed using human expert knowledge. Using an evolutionary algorithm we optimize the parameters and the structure of the fuzzy rule set to adapt the controller to human driving behavior. Further details of the optimization method and a comparison to other optimization methods are described in [10]. Fig. 6 shows the time-series of distance and speed of a scene with

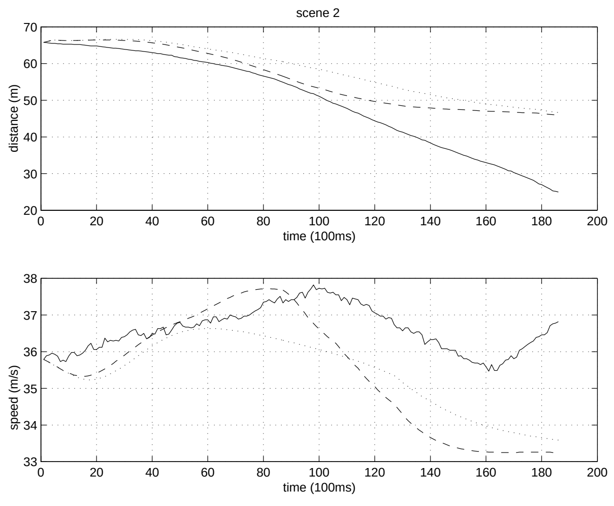

Fig. 6. Scene with safety constraint.

a car driving up to a car in front. The solid lines show the recorded values for the distance and the speed. The dotted lines show the simulation results using the initial controller before optimization. After the optimization the controller behaves as indicated by the dashed lines. The scene could be divided into two parts. From time-step 0 to time-step 100 it shows the adaption of the controller to the behavior of the human drivers. Distance and speed of the simulation with the optimized controller differ much less from the recorded data than the initial controller. After this point, the controller enforces a safe distance to the car in front and does not follow the (dangerous) behavior of the human driver. This example shows the influence of the problem dependent objective function.

4. CONCLUSION

In this paper a scene analysis and a behavior organization for a video-based driver assistance system is presented. In the first part an object detection, an object tracking and an object classification for rigid and non-rigid objects using multiple algorithms are introduced.

For behavior organization a neural field dynamic is chosen to adjust the steering angle of the vehicle, because the field dynamics is controlled by the object attributes easily. The acceleration of the vehicle is controlled by a fuzzy system. An evolutionary algorithm based optimization method is used to adapt an initial controller to human behavior.

5. REFERENCES

[1] S.I. Amari. Dynamics of pattern formation in lateral inhibition type neural fields. In *Biological Cybernetics*, volume 27, pages 77–87. Springer Verlag, 1977.

[2] 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*, volume 7(1), pages 62–81, 1998.

[3] C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakas, and W. von Seelen. Walking Pedestrian Recognition. In *ITSC'99*, pages 292 – 297, Tokyo, Japan, 1999.

[4] E.D. Dickmanns et al. Vehicles capable of dynamic vision. In *IJCAI'97*, pages 1–16, Nagoya, Japan, 1997.

[5] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W. von Seelen. An Image Processing System for Driver Assistance. *Image and Vision Computing (Elsevier)*, 18(5):367 – 376, 2000.

[6] U. Handmann, I. Leefken, and A. Steinhage. Behavior Planning for Driver Assistance using Neural Field Dynamics. In *NC 2000*, Berlin, Germany, 2000.

[7] U. Handmann, I. Leefken, C. Tzomakas, and W. von Seelen. A Flexible Architecture for Driver Assistance. In *Proceedings of SPIE Vol. 3838*, pages 2 – 11, 1999.

[8] D.P. Huttenlocher. Comparing Images Using the Hausdorff Distance. *IEEE Transactions on PAMI*, 15(9), September 1993.

[9] T. Kalinke. *Texturbasierte dynamische Erkennung veränderlicher Objekte*. PhD thesis, Ruhr-Universität Bochum, Germany, 1999.

[10] Q. Zhuang, J. Gayko, and M. Kreutz. Optimization of a Fuzzy Controller for a Driver Assistant System. In *Proceedings of the Fuzzy-Neuro Systems 98*, pages 376 – 382, München, Germany, 1998.