Computer Vision for Driver Assistance Systems

Uwe Handmann, Thomas Kalinke[a], Christos Tzomakas,
Martin Werner, and Werner von Seelen

[a]Institut für Neuroinformatik, Ruhr-Universiät Bochum,
D-44780 Bochum, Germany

ABSTRACT

Systems for automated image analysis are useful for a variety of tasks and their importance is still increasing 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. Fully or partly autonomously guided vehicles, particularly for road-based traffic, pose high demands on the development of 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 introduce a system which extracts the important information from an image taken by a CCD camera installed at the rear view mirror in a car. The approach consists of a sequential and a parallel sensor and information processing. Three main tasks namely the initial segmentation (object detection), the object tracking and the object classification are realized by integration in the sequential branch and by fusion in the parallel branch. The main gain of this approach is given by the integrative coupling of different algorithms providing partly redundant information.

Keywords: object detection, tracking, and classification, fusion of basic algorithms, integrative overall system

1. INTRODUCTION

Some systems presented in Ref. 1–3 show the principle feasibility of driver assistance systems based on computer vision. Although exclusively vision based systems and algorithms are not yet powerful enough to solve all driving relevant tasks, a large amount of different scenarios can be interpreted sufficiently. Additionally sensors like RADAR and LIDAR extent the contents of sensor information necessary for building a reliable system. The main focus of our system lies in combining various methods for the analysis and interpretation of images and in the fusion of a large spectrum of sensor information to extract most reliable information for the final planing and predicting behavior of the vehicle.

In contrary to other systems we do not want to restrict our system by using a defined camera geometry or assumptions about the environment (like flat roads) in our algorithms in order to be most flexible to the requirements of the cooperating partners of the car manufacturing industry. It is the task better solved by the companies to build up overall systems implementing constraints depending on the camera geometry and parameters and on the type of vehicle equipped. Nevertheless our system's performance will increase by this addition. So the main gain of our system lies in providing a high level of flexibility. Nevertheless the incorporation of a-priori knowledge like model prototypes is done. Another constraint motivated by the industry due to the high final expenditure is the application of only one camera for the front view and one for the rear view. But due to decreasing costs of hardware components this constraint will relax.

Principal problems are caused by having a moving observer in predominantly natural surroundings classifying moving objects with a task determined operating frequency and estimating their positions. In particular, it is attempted to isolate traffic participants from video images and to attribute the obtained object hypotheses (e.g., object class, distance, velocity, danger potential with respect to the planned self-movement trajectory, etc).

Other author information: (Send correspondence to Thomas Kalinke)
Thomas Kalinke: E-mail: ThomasKalinke@neuroinformatik.ruhr-uni-bochum.de

The great variety of different scenarios as well as the high degree of reliability necessary for the given task require an encompassing and flexible system architecture. The requirements concerning the reliability of the reached solution, the variety of geometric appearances of involved objects and that of environmental constraints of both deterministic as well as statistical nature necessitate a multitude of partial solutions based on different representations of the environment. Consequently, complexity and structure of the overall system have to be adaptable to increasing system complexity in order to allow for accommodation of additional modules without degeneration of already accomplished partial solutions. For this reason, even 'simple' applications are encumbered by considerations concerning the overall system architecture.

Basically, the overall system architecture can be divided into basic, fusion and integration algorithms. Basic methods are those providing specific partial solutions under given constraints. Results and application of the individual algorithms are not independent, resulting in an increase in redundancy making the overall system secure and reliable given a suitable coupling architecture. The necessary methods for fusion and integration ensure a flexible cooperation of the basic building blocks as well as the integrative derivation of results from them. In a similar vein, a sequentiell data transmission and system dynamics are necessary in order to build up an overall system with feedback stability solving complex tasks.

The fusion of different sensor information and preprocessing results increases the performance of the system. All basic algorithms themselves are *specialist* for a specific kind of sensor information. Figure 1 shows the different types of information principles depending on the spatial relationship to the vehicle. With respect to the requirements of various applications optimally adapted algorithms are built. In the area F1 contour based methods are chosen. On the one hand the sparse coding (edges) of the intensity information is sufficient enough due to the high resolution of the objects in the image and on the other hand it speeds up computation time for real time applications. Here, we mainly use a feature called local orientation coding (LOC).[4] In the field F2 we use motion detection algorithms to segment overtaking and overtaken vehicles. In contrary to other applications we use a pattern tracking based algorithm which ensures high stability. The long distance field F3 is analyzed by texture based methods. The low spatial resolution make an edge based processing infeasible. Nevertheless the integrative characteristics of texture analysis provides good results by separating the objects from the background by use of their texture.

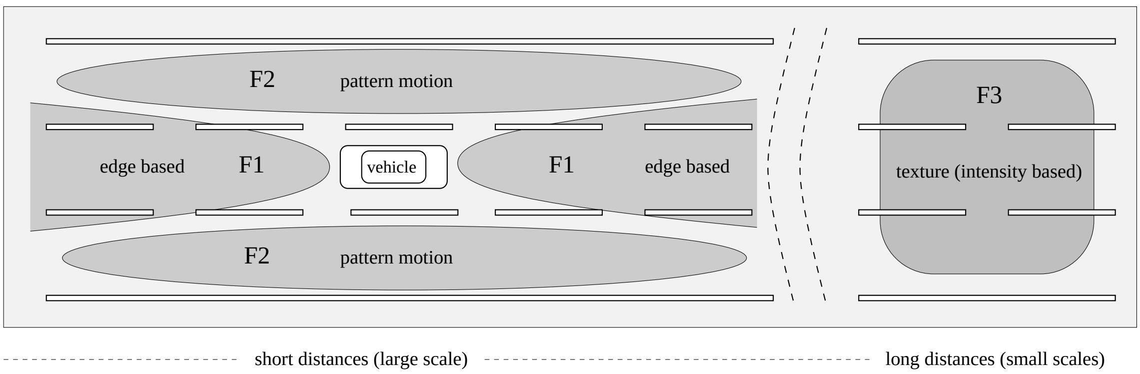

Figure 1. Separation of the road in fields F1, F2 and F3 in which different algorithms can be applied optimally.

In the area of preprocessing, a multitude of different methods for initial segmentation, object tracking, and object classification has been developed in the context of current research. A few inherent tendencies appear remarkable.

- Previous work often was based on the use of higher features, meaning the generation of a sequence of features beginning at the iconic (image-based) side and continuing to the symbolic side. There are two main reasons:

1. The historic rooting of image processing in material and surface inspection for quality control has lead to the existence of theoretically well-founded and practically tested algorithms.
2. Symbolic features are commonly used for compact coding purposes, so that processed data amounts can be largely reduced for accommodating limited processing resources.

The breath-taking evolution of processors has particularly alleviated the impact of this last constraint. In addition, it appears that particularly in the context of limited sensor resolution (i.e., in long distance regions) algorithms can be employed that rely on statistical measures of extensive 'early' (in the chain of processing) feature sets. These algorithms supplement the spectrum of methods explicit in more traditionally oriented algorithms.

- Often a formulation as an optimization problem can lead to implicitly robust solutions avoiding disadvantages of explicit methods (e.g., the correlation of model with image features, the correspondence problem). In this area as well the increase in available computational power has been aiding progress.
- Particularly in natural environments, flexible algorithms possessing a certain learning capability for input data driven adaption are preferably used.

At the *Institut für Neuroinformatik* algorithms providing partial solutions for object detection, tracking and classification have been incorporated in an driver assistance architecture. Namely the following enumeration gives an overview over the applied methods.
initial object detection: local orientation coding (LOC),[4] polygon approximation of contours,[5] use of local symmetry,[6] pattern motion analysis, texture analysis based on local image entropy and local cooccurrence measures,[7] shadow detection, and RADAR mapping.
object tracking: *Hausdorff* distance matching and cross entropy.[8]
object classification: local orientation classifier, *Hausdorff* distance classifier, and cooccurrence classifier.

A further task of high importance to solve is the lane detection. The solution of this is two fold. On the one hand the visibility of the road markings have to be ensured if a reliable lane detection has to be done over a large distance. For the lateral vehicle control a linear approximation of the first meters in front of the vehicle is sufficient enough. Solutions for this are developed many times. So our main interest lies in the lateral control of the vehicle. A solution of our work is given in Ref. 9,10 where the distance to a car in front is controlled by a fuzzy system using evolutionary algorithms and a RADAR sensor input. This system is running on a vehicle prototype. But in most of the typical driving scenarios on German *Autobahnen* the view in front is limited so that an object tracking algorithm is used to provide information about the lane as long as the tracked vehicle in front keeps in lane. Nevertheless two solutions for lane detection have been worked out to deliver additional information for initial object detection (segmentation). The first one is based on the road markings and operates on polygon approximations of contours and a road model.[11] The second one completes the road detection by using local image entropy[7] - in others words using local textures. Figure 2 shows the results of the different methods. Further more the data bases about the course of the road and actual traffic work are enhancing and getting more and more precise and will be available for every driver in the near future. The use of the GPS and GLONASS signals[12] provide a reliable determination of your position on the road. So informations about the course of the road and the detection of traffic signs seem to loose their relevance for computer vision solutions.

The next sections describe the functionalities of the basic algorithms first. Then the method of robust an flexible fusion is presented. Next the integrating system explained. The paper is concluded by a presentation of the results and a discussion including an outlook to future work.

2. THE BASIC ALGORITHMS

As explained in the introduction different basic algorithms are used to provide partial solutions for some spectrum of sensor input informations. All algorithms can be parted in methods working on

Figure 2. Lane extraction, determination of free space, and scaling estimates of vehicles using camera and environment geometry.

differential information (e. g. edges) and integral measurements (e. g. texture). For the application types, object initial detection, tracking and classification, a short description is given.

2.1. Initial object detection

The main motivation of using multiple simple methods is that the development of designing 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 "specialist" is carried out. A polygon approximation of gradient images (Sobel filtering) are performed by a standard hardware. The detection of shadows is realized by thresholding the intensity image, some morphological processing and a region clustering stabilized over time. The information of the RADAR signal is mapped into image coordinates. The others methods are describe shortly in the next sections.

2.1.1. Local Orientation Coding

The 'raw' gray scale (intensity) images are preprocessed by a method we call local orientation coding (LOC). The image features obtained by this preprocessing are bit strings each representing a binary code for the directional gray-level variation in a pixel neighborhood. In a more formal fashion the operator is defined as

$$b'(n,m) = \sum_{i,j} k(i,j) \cdot u(b(n,m) - b(n+i,m+j) - t(i,j)), \quad (i,j) \in \text{ neighborhood}$$

where $b(n,m)$ denotes the (gray scale) input image, $b'(n,m)$ the output representation, $\mathbf{k}$ a coefficient matrix, $\mathbf{t}$ a threshold matrix and $u(z)$ the unit step function. The output representation consists of labels, where each label corresponds to a specific orientation of the neighborhood.

An adaption mechanism for the parameters $\mathbf{t}$ of the coding algorithm yields a high level of flexibility with respect to lighting conditions.[4]

2.1.2. Symmetry

One powerful feature is the bilateral symmetry.[13] Our interest in vertical symmetry detection originates from the object detection task. Especially vehicle front and back views are strongly symmetric. Furthermore, we want to detect vehicles which are situated in a long-distance area as early as possible to increase the ability of scene analysis before the object is in the nearer surrounding. Under these circumstances the system has to cope with low pixel resolution, noise, partial occlusions of objects and the signal-to-noise ratio of the background to the object in given hypotheses (region of interest) may be low. We developed a neural network[6] that is capable of measuring the strength of symmetry. Furthermore a set of different positions of the symmetry axis are under inspection at a time. The lack of symmetry is plotted in figure 3 for different hypotheses of symmetry positions. The object covers only half of the image.

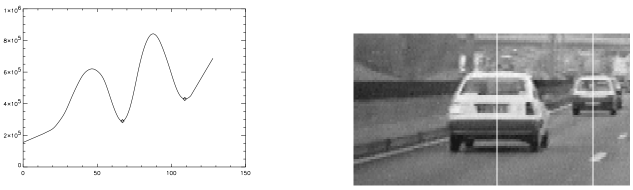

Figure 3. Plot of the lack of symmetry over possible horizontal positions in the image. Detected minima correlate with maxima of symmetry strength.

2.1.3. Texture

Besides operators like intensity derivation (gradients and the LOC) texture analysis as an integrating operator has been used successfully in image processing. The term *texture* is not explicitly defined. Globally texture is a description of image pixels or texture elements (groups of pixels) belonging to a specific texture class due to their spatial arrangement to other elements. Texture depends inherently on scaling. The spatial and intensity relationship between these elements define the kind of texture. Strong variation of intensity in a small area lead to fine textures and low variations produce coarse textures. Furthermore textures can be parted into properties weak and strong. Weak textures are described mostly by statistical methods. In strong textures the spatial interaction of elements are somewhat regular. Their recognition is usually accompanied by an exact definition of texture primitives (grammars). Actually two different methods for analyzing are commonly used: statistical and syntactic. In our applications we mainly work with statistical texture description. Every kind of texture is represented by a multi-dimensional feature vector in order to evaluate a statistical pattern recognition for every texture class based on suitable decision rules.

Local Image Entropy The *local image entropy (LIE)* has been developed at our institute.[7] In this method an estimation of the information contents of a pixel and its neighborhood is given. A saliency map is calculated so that a separation of objects and background can be evaluated. Figure 4 shows results of an initial segmentation process based on the LIE. The areas of objects and background are cut out by thresholding. A detection of road-users and the free driving space can be easily done.

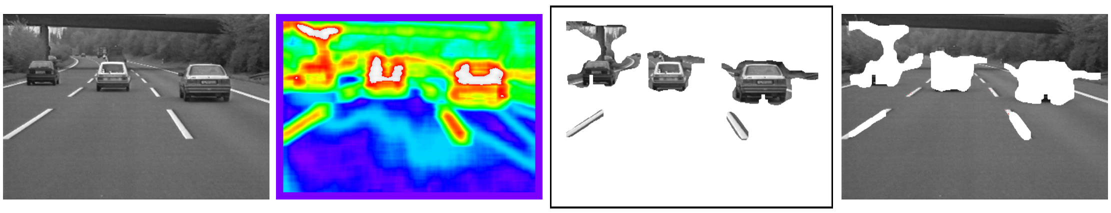

Figure 4. Image, local image entropy, image segmentation: objects and free driving space

Cooccurrence Matrix One of the fundamental tools in texture analysis, the cooccurrence matrices, were suggested by Ref. 14. In here, the probability of the cooccurrences of pixel pairs under predefined geometrical and intensity constraints are measured. These constraints are determined by the intensity ratio and the spatial relationship (angle and distance) of two image points. A definition of the cooccurrence matrix follows. In an image window $\mathbf{I}$ of size $M \times N$ and a maximum number of different gray values Q the cooccurrence matrix $\mathbf{P}$ is calculated under parameters angle α in a given distance d as follows

$$P_{d,\alpha}(i,j) = \frac{\text{number of pairs}_{((x,y),(x',y'))}\text{, verifying }_{(d,\alpha)}\text{ and }_{I(x,y)=i}\text{ and }_{I(x',y')=j}}{\text{number of all pairs in image window}_{((x,y),(x',y'))}}$$

A calculation of texture features is performed in most of the applications under four directions ($\alpha = 0, 45, 90,$ and 135) and different distances $d = 1, 2, \ldots$. A rotation-invariance can be obtained by accumulation of the matrices of the four directions. The amount of scaling variance can be reduced by calculating the matrices over different distances.

Julesz showed that the human perception of texture is based on cooccurrence statistics. Haralick, Shunmugan, and Dinstein suggested in[14] 14 different statistical features which can be obtained from the cooccurrence matrices.

In our field of research cooccurrence matrices are mainly applied to the initial segmentation. The matrices are calculated in overlapping windows. Features like energy, entropy, contrast, correlation and the highest cooccurrence of[14] are combined for the segmentation process.

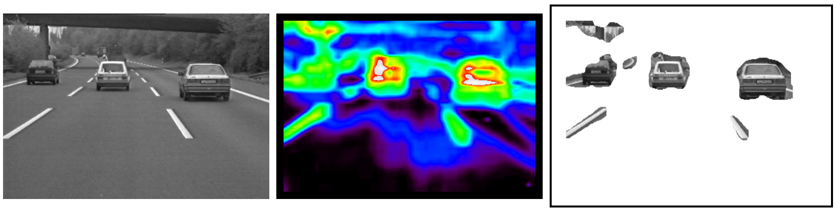

Figure 5. Image, Cooccurrence based saliency map and the thresholded and re-mapped segmentation result

2.2. Object Tracking

Algorithms for object tracking are the most important if a stabilization over time or a prediction e. g. trajectories should be given. Three methods for object tracking are introduced in the next sections. As it can be seen in figure 1 the tracking algorithms find their applications depending on the spatial resolution of the images. In the near field the Hausdorff distance or the order statistics are used as a measurement based on contour codes (LOC). Here we present the more stable method the Hausdorff distance tracker that has been tested successfully on a large set of different image sequences. For further details of the approach using order statistics see Ref..[15] Supplementary in the long distance field the texture based cross entropy provides optimally results.

2.3. Hausdorff Distance for Object Tracking and Classification

The geometric comparison of shapes is a fundamental tool for model-based object recognition. Most of the methods used in object recognition systems refer to a similarity measure between the model features and the image features.[16] The *Hausdorff distance* measures the divergence of a set of features with respect to a reference set of features.[17] These sets mostly describe object contours in our application. The comparison of similar object contours yields small distance values, whereby objects with different contours yield larger distances.

The directed Hausdorff distance h of one point set A against a point set B is the maximum of the minimum distances of each point of set A to the points of set B. The final Hausdorff distance H is simply the maximum of the two directed distances.

$$\begin{aligned} h(A,B) &= \max_{p\in A}\min_{q\in B}\|p-q\| \\ H(A,B) &= \max(h(A,B), h(B,A)) \end{aligned}$$

The partial Hausdorff distance performs a ranking of these minimum distances and considers a fraction of them instead of the maximum.

Unlike the classical correlation methods the Hausdorff distance uses Min-Max operations instead of multiplications, so it is more efficient in time. The partial Hausdorff distance is robust against partially occluded objects and outliers that may arise at the contours due to noise or insufficient feature extraction.

The partial Hausdorff distance can examine object hypotheses in a complex scene. This method was tested successfully with highway-traffic scenes. It was able to recognize vehicles on highways and track them over time. Two degrees of freedom were considered in our schema: translation and scaling of models.

2.3.1. Texture based Object Tracking - Cross Entropy

One of the simple description of textures is obtained by intensity histograms (first order statistics). Especially non-rigid objects like pedestrians and two-wheeled vehicles which consist of a further rotational degree of freedom compared to other road-users can be tracked using the cross entropy. As described in[8] a matching process can be performed by comparison of two probability distributions. In our application a model distribution at time step $(t-1)$ is compared to several hypotheses at time t. Figure 6 show tracking of pedestrians using intensity and edge probability distributions.

Figure 6. Tracking of pedestrians based on the cross entropy based on intensity distributions and LOC features

As an extension to the proposed method we use instead of statistics given by a histogram correlated statistics given by the cooccurrence matrices. The quality of the estimate of position and scale increases but the calculation time increases as well.

2.4. Neural Classifiers for Vehicles

As the effective application of different tracking schemes in the case of the classification task edge based and texture based solution have been developed in order to cover high and low object resolutions in the image. The LOC-Classifier is computational fast method used for a fast estimate of given ROI. It is aimed at separating possible objects from the background. It is independent from the resolution of the objects due to a normalization in size. Additionally two classifiers with higher computational costs perform a reliable classification. The Hausdorff distance classifier processes objects in the near field with high spatial resolution enhancing the ROI image coordinates. The cooccurrence classifier processes objects in the long distance field based in intensity texture analysis.

2.4.1. LOC-Classifier

With the given local orientation coding,[4] described in section 2.1.1), a classification of vehicles is realized. The classifier has to cope with partial occlusions, varying illumination conditions, tilt of an object, differently resolved structures depending on the distance of the object under consideration, noise and perturbations induced by the recording and processing equipment, different viewpoints and different kind of cars with different shapes and colors. Additionally, the classifier should be able to generalize from relative few training examples to the necessary features characterizing a car. Therefore, a neural network has been chosen for solving the classification task. It is a feed-forward neural network with one hidden layer trained by the error back-propagation algorithm.[18,19]

These networks are known to be universal approximators for any continuous valued function.[20] Furthermore, it is shown that these structures can, with some small modifications, approximate a-posteriori probabilities in the sense of a Bayesian classifier.[21]

The inputs for the classifier are certain subsets of the histograms. The output is the class of the region.
The complete system has been implemented and extensively tested on the Mercedes Benz VITA II test vehicle.[22] Different classes of vehicles have been trained. For a further evaluation of the system see Ref. 23.

2.4.2. Hausdorff Distance Classifier

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 here 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 cars and trucks. Due to the fact that rear views of cars differ significantly from rear views of trucks, one can expect that the design of generic models for each class can accomplish the separation of the objects of both classes.

The classification works according to the following scheme: Each region is compared with two models, i.e. a car model and a truck model. The features of the region and the models have been extracted using the Local Orientation Coding. For robuster results the horizontal features are separated from the verticals, for both the region and the models.

Figure 7. Each region is compared with two models

The Hausdorff distance is computed for each model over all the possible translations inside the region and a certain range of scales. The fractions of the features of the forward and the backward match that verify a given distance threshold constitute for each model the criteria for its classification. These values are learned by a multi-layer perceptron (MLP) network using the back-propagation algorithm.

2.4.3. Cooccurrence Classifier

In a lot of applications features of the cooccurrence matrices like energy, entropy, contrast, correlation and so on[14] were used for classification processes. But every reduction of the dimensionality implicates reduction of information as well. Therefore, the matrix itself is used for classification. It is the task of the neural network to extract the necessary feature by itself. Due to the fact that the matrix is symmetric only the non redundant part of the matrix is extracted to build the characteristic vector $S_1(i,j)$.

In order to estimate the statistics of even small objects a reduction of the grey-level dynamics is performed. The range of $Q \in \{0,\ldots,255\}$ is reduced to a 4-bit range $Q \in \{0,\ldots,15\}$ by a bit shift operation (extracting the highest 4 bits). Another method introduced in[24] implies a dynamic adaption to the actual grey-level distribution. It is the goal to keep the influence of the background to the grey-level reduction as small as possible. Therefore, the bit shift operation is selected to process all elements in the same manner without dependency on the background distribution.

The cooccurrence matrices were calculated from a classification database. A neural network was chosen as a classification structure. A multi-layer perceptron with one hidden layer using the quick-prop-learning algorithm was trained. The net contains 136 input neurons, 5 hidden neurons and 3 output neurons for the classes car, truck and background.

The input vectors are calculated from a classification database containing a wide spectrum of different objects. They were taken from different image sequences acquired by various camera under different points of view to cover large spectrum of objects. The training set contains about 680 examples: 300 cars, 200 trucks and 180 images of the background. All objects had different scalings starting from 40×20 pixel up to 400×200 pixels. Figure 8 shows classification results. A classification of the background is added. The performance of the classifier scales with the size of the ROI, thus if structural details get lost the differentiation of cars and trucks becomes more difficult. A stabilization over time improves the results.

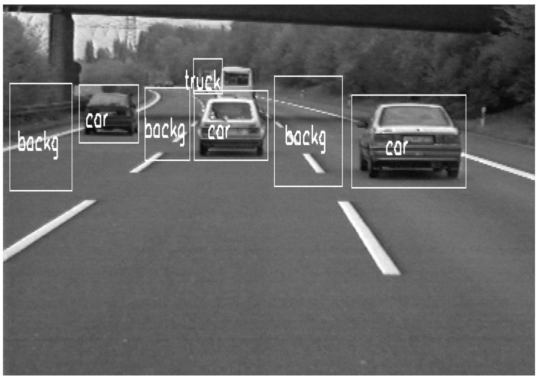

Figure 8. Classification results.

3. THE CONCEPT OF FUSION

Data fusion is one of the main gains one can get when a large amount of stability and reliability is necessary like in this application of driver assistance systems. On one hand a gain in robustness is achieved 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. Two different types of neural coupling mechanisms are introduced.[25] The high flexibility, the facility of expansion and the adaptive retraining processes have led to the choice of neural networks.

3.1. Fusion for Initial Segmentation

The aim of fusion in computer vision is to get an improvement of special solutions and single methods with a coupling net. Especially the modular coupling of single processing steps generates redundancy necessary for object recognition. Within this, greater flexibility and robustness of the image processing modules and higher adaptation of the modules regarding to the problems should be achieved.

3.1.1. Neural Coupling

A formulation of a robust and error tolerable segmentation is shown in figure 10. Computer vision modules, generating lines (polygon approximation of the contour),[5] local orientation coding,[4] local image entropy and cooccurrences,[7] symmetry, shadow detection, and Radar remapping are coupled in a neural network (figure 10). A feedback over time is possible, additional sensor information could be easily integrated at this level.

3.1.2. Segmentation by Fusion

Based on integrative (entropy) and differential (LOC, polygon approximations) representations a 12-dimensional input vector

$$\mathbf{u}(x, y)^T = (\mathbf{u_1}(x, y)^T, x, y)^T$$

for the coupling net is generated for each pixel. Here, the 10-dimensional vector $\mathbf{u_1}(x, y)$ refers to the contour and texture information and the values of x and y represent the pixel's coordinate position. Vector $\mathbf{u_1}(x, y)$ is defined by

$$\mathbf{u_1}(x, y) = \sum_{(i,j) \in R} \mathbf{v}(i, j)$$

where R is a local neighborhood (e.g. 9×9) of (x, y) and $\mathbf{v}(i, j)$ is a binary vector. $(v_1(x, y), \ldots, v_4(x, y))^T$ code a subset of the LOC results, $(v_5(x, y), \ldots, v_9(x, y))^T$ code the value of the entropy and $v_{10}(x, y) = 1$, if (x, y) belongs to a polygon.

A multi-layer perceptron (MLP)[26] with a 12-5-1-structure of the layers is used here as a coupling net. The feature vector is propagated through the twelve input neurons and the five hidden neurons to the output. The activity of the output neuron o describes the membership of each pixel to the initial segments, if $o > 0.5$.
The back-propagation algorithm[26] is used for training the net. With a hand-labeled database the error function between the received output activity after passing the net and the required output is calculated and propagated backwards. This structure is flexible to be enlarged and adaptable for new problems.

Figure 9 shows the robust segmentation results of cars in a sequence of traffic scenes. Only every 50th frame is shown in the figure.

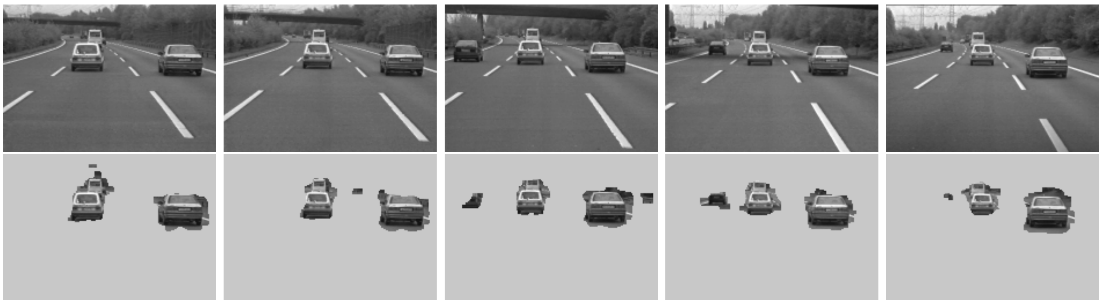

Figure 9. Initial segmentation with a MLP coupling net

3.2. Fusion of Classifier Results

Besides the task of object detection and tracking over time the process of classification is a valuable information. The large amount of different vehicle types and the interfering environmental conditions like weather conditions (shading, illumination, etc.) require high flexibility and robustness of the classifier. A-priori knowledge can be incorporated by decomposition of the classification task to subtasks. A modular neural network has been designed for vehicle classification in order to require a small input space for every module and to facilitate the analysis of the classifiers' performance. Every module is an expert in its small band application spectrum and is resulting in high performance by gaining from a suitable coupling structure. The Cooccurrence and the Hausdorff classifier are efficiently combined by a modular neural net using a multi-layer perceptron as a coupling network. To find a fixed threshold to decide which classifier to be taken is an difficult task. But the neural fusion network has learned during its training process to assess the results of the both classifiers by the sizes of the objects. It has been trained by distinguish sedans, trucks and varying parts of the background.

4. THE CONCEPT OF INTEGRATION

The overall system is shown is figure 10. The concept of integration single steps to a reliable working system is mainly based on feedback of single step results. As a sensor input the intensity image and a radar signals are used. The results of basic preprocessing algorithms are fed into a neural fusion architecture, the initial object detection, that provides hypotheses of possible vehicles. The very fast calculating LOC classifier reduces the set of hypothesis. An internal stabilization over time ensures further robustness. In order to confirm the hypothesis an object tracking is performed where the object size and type decides whether the Hausdorff tracking or the cross entropy tracking have to be used. The results, scale, position, and confidence are fed into the main stream and to the modular classifier. A neural network determines depending on the results of the object size, the Hausdorff and cooccurrence classifier what type of vehicle has been tracked. Concerning the calculation rates the object tracking has be performed for every time step. The initial object detection can process on a slower time axis. Finally the classification provides their results on larger time steps due to the fact that a tracked object with high confidence values will not change its class. So the object tracking is the most important task next to the detection process. To ensure a stable tracking over time a Kalman filter is implemented.

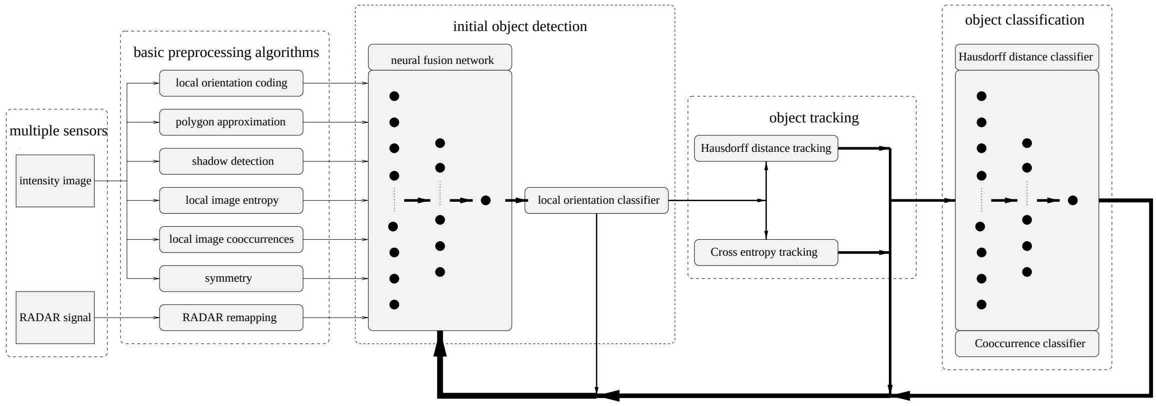

Figure 10. overall system

The main recoupling stream gathers all the results of the single tasks and is increasing over time. The type of information is changing from an iconic (preprocessing) to a symbolic (classification) description. A global data representation is built.

5. RESULTS AND DISCUSSION

We presented an overall system integrating the results and experiences of a long period of research in computer vision. Due to the increase of computational power and the development of reliable algorithms a fusion and integration of basic methods each solving specific problems can be performed to realize an overall stable system. The stability and robustness is largely increased. Because the overall computational time by using actual standard hardware is still quite long a spin off was realized. So if a real time operation system is the goal the whole processing has to be restricted to some algorithms due to limited computational power and time. In this application the initial object detection is restricted to a shadow analysis including a LOC classification. The objects are tracked by the Hausdorff tracker and classified by the Hausdorff classifier in order to use just one preprocessed feature map. On a standard DEC Alpha (500 MHz) the system uses 10 ms for the initial segmentation including a time stabilization, then the LOC classification needs 2 ms for every ROI, the tracking is performed in about 2 ms per object (we restrict the number of objects to five so that 10 ms for tracking is realistic) and finally the classification takes about 8-12 ms per object. As mentioned before the classification has not to be calculated for every frame. This system is quite capable of the obedience of real time requirements but the processing cannot cope with all different scenarios and we have restrict this application to extra-urban roads and motorways. Nevertheless if the performance of the hardware components will increase the the presented overall system is able to cope with most of the scenarios even in more complex situations.

ACKNOWLEDGMENTS

The authors like to thank all members of the Institut für Neuroinformatik who have aided in accomplishing the work presented.

REFERENCES

1. E. Dickmanns *et al.*, “The Seeing Passenger Car 'VaMoRs-P',” in *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, pp. 68–73, 1994. E.D. Dickmanns and R. Behringer and D. Dickmanns and T. Hildebrandt and M. Maurer and F. Thomanek and J. Schielen.
2. D. Pomerleau, “RALPH: Rapidly Adapting Lateral Position Handler,” in *Proceedings of the Intelligent Vehicles '95 Symposium, Detroit, USA*, pp. 506–511, 1995.

3. A. Broggi, “A Massively Parallel Approach to Real-Time Vision-Based Road Markings Detection,” in *Proceedings of the Intelligent Vehicles '95 Symposium, Detroit, USA*, pp. 84–85, 1995.
4. C. Goerick, D. Noll, and M. Werner, “Artificial Neural Networks in Real Time Car Detection and Tracking Applications,” *Pattern Recognition Letters* , 1996.
5. ELTEC Elektronik GmbH, Mainz, “THINEDGE-Processor for Contour Matching.” Hardware Manual, Revision 1A, 1991.
6. T. Kalinke and W. von Seelen, “A Neural Network for Symmetry-Based Object Detection and Tracking,” in *Mustererkennung 1996*, pp. 37–44, 1996.
7. T. Kalinke and W. von Seelen, “Entropie als Maß des lokalen Informationsgehalts in Bildern zur Realisierung einer Aufmersamkeitssteuerung,” in *Mustererkennung 1996*, pp. 627–634, Springer-Verlag, (Berlin, Heidelberg), 1996.
8. T. Kalinke and W. von Seelen, “Kullback-Libler Distanz als Maß zur Erkennung nicht rigider Objekte,” in *Mustererkennung 1997*, 1997.
9. Q. Zhuang, M. Kreutz, and J. Gayko, “Optimization of a fuzzy system using evolutionary algorithms,” in *Proceedings of the Fuzzy-Neuro Systems 98*, 1998.
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*, 1998.
11. T. Kalinke and C. Tzomakas, “Objekthypothesen in Verkehrsszenen unter Nutzung der Kamerageometrie,” *IR-INI 97-07, Institut f"ur Neuroinformatik, Ruhr-Universit"at Bochum* , 1997.
12. D. Walsh, S. Capaccioand, D. Lowe, P. Daly, P. Shardlow, and G. Johnston, “Real time differential GPS and GLONASS vehicle positioning in urban areas,” in *Proceedings of the Intelligent Vehicles Symposium, Boston*, p. 229, Digest 1997.
13. V. Bruce and P. Green, *Visual Perception Physiology, Psychology and Ecology*, Lawrence Erlbaum Associates, London, 1985.
14. R. Haralick, K. Shanmugan, and I. Dinstein, “Textual features for image classification,” *IEEE Transactions on Systems, Man and Cybernetics* **3**(6), 1973.
15. M. Werner and W. von Seelen, “Using Order Statistics for Object Tracking,” in *Proceedings of the Intelligent Vehicles, ITSC, Boston, USA*, IEEE, 1997. In this volume.
16. D. Paglieroni, “Distance Transforms: Properties and Machine Vision Applications,” *CVGIP* **54**, pp. 56–74, January 1991.
17. D. Huttenlocher, “Comparing Images Using the Hausdorff Distance,” *IEEE Transactions on PAMI* **15**, September 1993.
18. J. Hertz, A. Krogh, and R. Palmer, *Introduction to the Theory of Neural Computation*, Addison Wesley, 1991.
19. D. Rumelhart, J. McClelland, and P. R. Group, *Parallel Distributed Processing*, MIT Press, Cambridge, Massachusetts, USA, 1986.
20. K. Hornik, M. Stinchcombe, and H. White, “Multilayer Feedforward Networks are Universal Approximators,” *Neural Networks* **2**, pp. 359–366, 1989.
21. M. Finke and K.-R. Müller, “Estimating A-Posteriori Probabilities Using Stochastic Network Models,” in *Proceedings of the Summer School on Neural Networks, Bolder, Colorado*, pp. 276–281, 1993.
22. M. Brauckmann, C. Goerick, J. Groß, and T. Zielke, “Towards all around automatic visual obstacle sensing for cars,” in *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, pp. 79–84, 1994.
23. S. Bohrer, T. Zielke, and V. Freiburg, “An Integrated Obstacle Detection Framework for Intelligent Cruise Control on Motorways,” in *Proceedings of the Intelligent Vehicles Symposium, Detroit*, pp. 276–281, 1995.
24. P. Bhattacharya and Y.-K. Yan, “Iterativ Histogram Modification of Gray Images,” *IEEE Trans. on Systems, Man, and Cybernetics* **SMC-25**(3), pp. 521–523, 1995.
25. T. Kalinke and U. Handmann, “Fusion of Texture and Contour Based Methods for Object Recognition,” in *IEEE Conference on Intelligent Transportation Systems 1997*, 1997.
26. S. Haykin, *Neural Networks*, Macmillan College Publishing Company, Inc., 1994.