Cooperative Cognitive Automobiles

Christoph Stiller, Georg Färber, and Sören Kammel

Abstract—Safety requirements are among the most ambitious challenges for autonomous guidance and control of automobiles. A human-like understanding of the surrounding traffic scene is a key element to fulfill these requirements, but is a still missing capability of today's intelligent vehicles. Few recent proposals for driver assistance systems approach this issue with methods from the AI research to allow for a reasonable situation evaluation and behavior generation. While the methods proposed in this contribution are lend from cognition in order to mimic human capabilities, we argue that in the long term automated cooperation among traffic participants bears the potential to improve traffic efficiency and safety beyond the level attainable by human drivers. Both issues are major objectives of the Transregional Collaborative Research Centre 28 *'Cognitive Automobiles,'* TCRC28 that is outlined in the paper. Within this project the partners focus on systematic and interdisciplinary research on machine cognition of mobile systems as the basis for a scientific theory of automated machine behavior.

I. INTRODUCTION

Among the most fascinating capabilities of intelligent beings is the seamless perception and interaction with their environment. Guidance and control of automobiles comprises a comprehensive example for these capabilities. A human driver needs to perceive and understand the automobile's environment. Based on the understanding of the scene he plans, initiates, supervises, and controls suitable behavior.

Driver assistance systems aim to project those capabilities onto artificial systems. Longitudinal control is supported by *adaptive cruise control* (ACC) systems that have been introduced in several vehicle models around the turn of the century. While those systems originally were restricted to comfort enhancement, i.e. to an operational speed range of about $50 - 150$ km/h with a potential acceleration range of about $-0.25 - 0.1$ g, operational function currently migrates to a full speed range as well as to intervention in safety-critical situations with extended acceleration setting amplitudes (see e.g. [1]). Likewise, assistance systems for lateral control have been introduced into the market that emerge from lane departure warning functions, as e.g. in the Mercedes Actros truck, to active heading support as, e.g. in Honda's HIDS-system [2]. Night vision enhancement systems introduced in the Lincoln Navigator and Mercedes S-Class are examples of functions that still rest at a pure information and warning level.

In research, ambitious additional driver assistance functions have been demonstrated in experimental vehicles under supervision and on tracks that exclude public traffic. Some prominent examples are the autonomous 'VaMoRs-P', 'FhG-Codriver' or 'Navlab' vehicles [3], [4], [5]. Recently, the Darpa Grand Challenge competitions in 2004 and 2005 gathered wide attention, when more than a dozen unmanned vehicles traveled a longer distance through the Mojave Dessert and five vehicles even accomplished the complete course of some 150 miles [6][1].

Beyond these encouraging successes, unsupervised autonomous driving in public traffic is still a far fetched vision. One of the major limiting shortcomings of driver assistance systems is their lack to reliably identify those situations in that sufficient performance cannot be guaranteed. From the point of view of models for human behavior [8] (Fig. 1), one might find that state-of-the-art autonomous automobiles are able to conduct skill based behavior to a large extend, i.e. they master stabilization tasks such as distance or lane keeping in simple situations. Even though integration of some rule based behavior has been successful[2], an exhaustive set of rules for autonomous driving has not even yet been formulated. The extension of rule based and the implementation of knowledge based capabilities require the implementation of cognitive capabilities to understand traffic scenes.
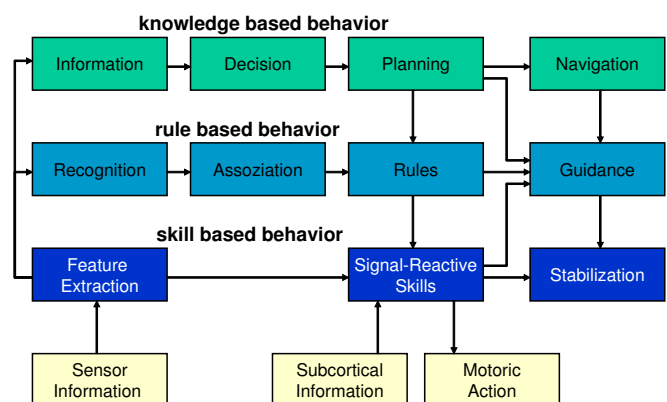

Fig. 1. Model for human behavior (cf. [8])

While cognition aims to mimic human capabilities, automated cooperation among traffic participants bears the

[1]The team *'Dessert Buckeyes'* that gathered partners from Ohio State University and Universität Karlsruhe became 10th of 195 participants [7].

[2]A well known example for rule based behavior is to neglect of closely cutting-in vehicles with positive relative velocity in longitudinal control.

potential to improve traffic efficiency and safety beyond the level attainable by human drivers. Fig. 2 illustrates this potential for a highway scenario in mixed traffic. It is worth noting that participants benefit from cooperative perception as well as from cooperative behavior.
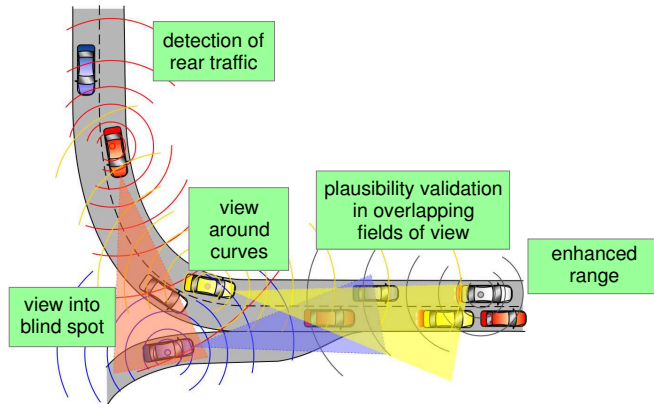

Fig. 2.   Cooperation of vehicles in mixed traffic

Both procedures are major objectives of the Transregional Collaborative Research Centre 28 *'Cognitive Automobiles,'* TCRC28 founded in January 2006. Within this project the partners

- Universität Karlsruhe (TH),
- Fraunhofer Institut IITB Karlsruhe,
- Forschungszentrum Karlsruhe,
- Technische Universität München, and
- Universität der Bundeswehr München

focus on systematic and interdisciplinary research on machine cognition of mobile systems as the basis for a scientific theory of automated machine behavior. The potential of cooperative perception and behavior is examined. Analytic research is accompanied by closed-loop simulations. Experimental autonomous vehicles build an important platform for the TCRC that allows demonstration and validation of the theoretical findings.

The same partners also contribute to the team *'AnnieWAY'* in the Urban Challenge 2007 competition [9]. As compared to the Grand Challenge 2005, this competition will pose additional challenges such as compliancy with selected traffic rules, passing, and merging with moving traffic [10]. The team expects that some basic principles developed in TCRC28 may be simplified under the restricted scenarios of the Urban Challenge to meet the real time requirements.

The remainder of this paper is organized as follows: Chapter II gives a brief overview of the system architecture used in the cognitive automobiles. Subsequently, in Chapter III, the machine cognition principles applied to traffic scenes are described. Chapter IV outlines the usage of this information for a cooperation among traffic participants and finally an outlook over the future work of the TCRC is given based on the conclusions of this paper in Chapter V.

II. SYSTEM ARCHITECTURE

Figure 3 depicts the block diagram of a cooperative cognitive automobile. It is worth noting the following two distinctions as compared to most other autonomous automobiles: First, the representation of knowledge is explicit to allow for knowledge based behavior. This representation comprises geometric and conceptual description of the dynamic vehicle environment and traffic situation as well as a formulation of fundamental goals and skills of the vehicle that lead to the current mission plan. Second, the stand-alone models of both perception and behavior generation are augmented by information gathered through the cooperation with other vehicles.
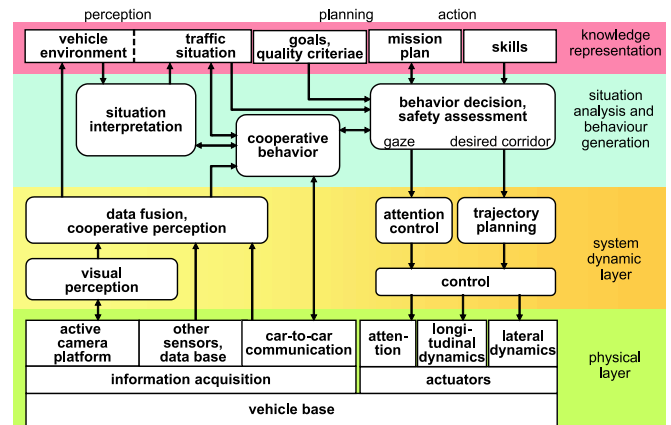

Fig. 3.   Block diagram of a cooperative cognitive automobile

The functional system structure is mapped on a hardware architecture as shown in Fig. 4: The modules corresponding to the upper three cognitive layers are implemented on a common-off-the-shelf AMD Opteron multiprocessor PC system as outlined in [11]. It delivers a computing power comparable to a small cluster, yet offers low latencies and high bandwidth for module interprocess communication. The unified hardware architecture ensures an active interchange of information among participating researchers. The control of the active vision platform is handled by an embedded system, providing the necessary response times for an inertial stabilization of the telephoto camera. We also dedicate a dSpace AutoBox to drive the vehicles actuators in order to meet our safety requirements. The chosen hardware architecture is supported by a realtime-capable software architecture as proposed in [11]: It consists of a central database (KogMo-RTDB), giving all cognitive modules a unique view of all knowledge available to the cooperative cognitive automobile. An easy-to-learn programming interface allows fast development and integration of new components.

The consortium has procured three Audi Q7 and a Volkswagen Passat for the project. Furthermore, a Smart roadster and a Volkswagen Touareg are used in the project context. All vehicles are designed to conduct full autonomous behavior[3]. Thus the consortium operates a fleet of in total six
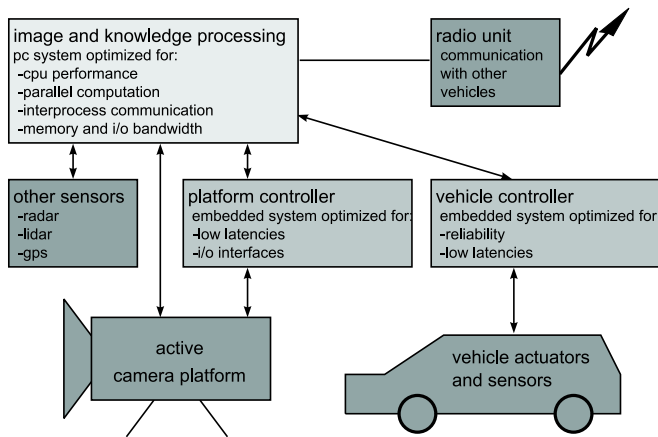
Fig. 4. Overview of the hardware architecture

Fig. 5. Active camera platform a) in the vehicle; b) schematic sketch [12]

vehicles which allows assessment of cooperative perception and behavior in mixed traffic as developed in the project. The modular software and hardware architecture with specified interfaces enables exchange and fusion of hardware and software modules among the partners.

Due to its rich information content, particular emphasis of the sensor system is laid on vision. Figure 5 depicts our active camera platform that includes three cameras. The yaw directions of the outer two cameras with wide angle lenses are independently steerable. The cameras can be steered to yield disjunct fields of view to monoscopically survey a wide field. Alternatively, the cameras may be steered in the same direction to cover some 70° instantaneous field of view stereoscopically. Any mixture or any dynamic transition between the two is possible as well. Through smooth or saccadic panning each camera can survey an almost 180° field of view. A strength of the active camera system is its capability for dynamic self-calibration [12]. This allows for 180° field of view 3D stereoscopic scene perception which, to the best knowledge of the authors, is a unique qualification. The third camera is a telephoto camera with steerable yaw and tilt direction that allows high resolution tracking of distant objects.

III. COGNITION: PERCEPTION, REASONING, AND INFERENCE

A central issue for any driver assistance function relevant to safety is its ability to assess the perception and decision performance under current conditions. In a typical perception process as sketched in Fig. 6, information emerges from the signal level (sensor raw data) over several processing steps via a geometrical-symbolic representation of the current traffic environment to the generation and control of suitable behavior.

Robustness is an important factor in this process: One successful method is to fuse data from different sensors. This may happen on the pixel level so that combined images from video- and IR-sensors provide information less sensitive to illumination; the fusion may happen also on a subsymbolic or symbolic level: Sensors that differ in nature like video, Lidar
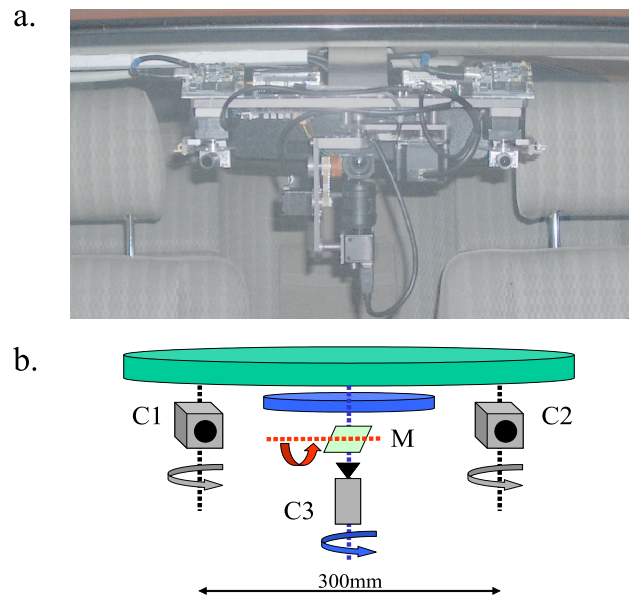
and Radar may detect objects, and combining these proposals generates more robust hypotheses [13], [14]. Tracking over time enhances reliability of object hypotheses further, and reduces computational effort since only regions of interest have to be processed. Object tracking with Extended Kalman Filters (EKF) or particle filters and model based object detection has been reported as an important element for the early success of approaches to autonomous driving, as e. g. for the 4D-approach [15].

It is crucial to not only propagate knowledge through the cognition scheme but to augment this knowledge with confidence measures. These measures are consistently processed at each step of the cognition chain, considering the confidence of previous processing steps along with additional noise introduced by sensors and the uncertainty introduced by the individual algorithms. At the top level the procedure quantifies the confidence of the set behavior considering the uncertainties of all previous levels.

In reverse direction, selective enquiries are conducted to resolve ambiguities at lower levels [16]. This scheme exhibits interesting parallels to biological image processing schemes, see e. g. [17]. Expectation based image processing that focusses the processing power to appropriate regions and features of interest not only reduces computational effort may also resolve ambiguities and uncertainties at decision level. As a concise example, the feature level may specifically search for some basic features and initiate feature detection at this level with a decreased decision threshold when a hypothesis at the object detection layer predicts a complex feature in a specific image region [18].

As outlined before, the partners in the TCRC28 research group from Karlsruhe and München are cooperating. Nevertheless they are also competitive in some aspects. In the field of perception, one group elaborates stereo algorithms to acquire depth information from motion and disparity keys,

whereas the other group focuses on biologically inspired principles: They use *a priori* knowledge about object classes and dimensions to estimate the distance from the object sizes in the images; the 4D-algorithm again stabilizes the results over a short image sequence; first results being very promising [19]. The different methods will be evaluated through competitive benchmarks. Through joint analysis of the performance of both groups it will be decided which method or which diverse combination of methods will further be pursued.
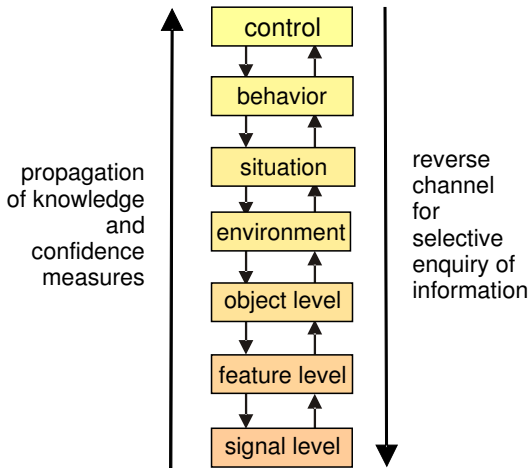

Fig. 6. Propagation of uncertainty, airiness and ambiguity through the cognition chain

Information may be purely quantitative at the lower levels of the cognition chain. E. g. at signal level it may be composed of a set of RGB intensities for all pixels and at control level by the setting amplitudes. In order to generate knowledge based behavior, however, conceptual information is required at intermediate levels. E. g. a situation may include assumed intentions of other traffic participants, such as *'driver intends a lane change'*. We employ a probabilistic inference process described in the sequel [20].

Let $\mathbf{s} = (s_1, s_2, \ldots, s_m)$ denote a situation, i. e. the set of all parameters that are relevant for driving. We model a situation as a sample from a random variable $\mathbf{S}$. Let further $\mathbf{g} = (g_1, g_2, \ldots, g_n)$ denote the set of the available evidence, i. e. pixel intensities, features derived thereof, rules, or prior knowledge. As before, $\mathbf{g}$ is considered as a sample of random variable $\mathbf{G}$. Likewise, behavior is denoted by $\mathbf{b}$. Possible behavior is constrained by the current skills of the autonomous system. With each behavior $\mathbf{b}$ applied in situation $\mathbf{s}$ we imply a cost functional $c(\mathbf{b}, \mathbf{s})$ that reflects our goals, values and quality criteria, such as e. g. some measure for safety. For any given evidence $\mathbf{g}$ one can then associate the Bayesian expected cost functional with every possible behavior $\mathbf{b}$

$$k(\mathbf{b}; \mathbf{g}) = \int c(\mathbf{b}, \mathbf{s}) p(\mathbf{S} = \mathbf{s} | \mathbf{G} = \mathbf{g}) d\mathbf{s}$$

and the behavior minimizing k may be generated. Formulation of the second factor in this integral is, in general,

a difficult task and dimensions m and n of the situation and evidence may be large. In order to decompose this *a posteriori* distribution, we impose a Markov model or, more precisely, we construct a Markov random field (Markov network). The Markov network, which can be represented by an undirected graph, allows us to incorporate *a priori* knowledge about the traffic scene in its structure. Because only a subset of variables in a traffic situation is directly dependent this reduces the complexity of the optimization problem: A Markov network comprises the variables in its neighborhood definition $\mathcal{G} = \{\mathcal{G}_1, \ldots, \mathcal{G}_m\}$, where $\mathcal{G}_i$ denotes the set of all neighbors to i, i. e. those variables that are conditionally dependent. Then the Markov property holds

$$p(S_i | S_j, \ \forall j \neq i) = p(S_i | S_j, \ \forall j \in \mathcal{G}_i).$$

Due to the Hammersley-Clifford theorem, Markov random fields may always be represented by a Gibbs distribution [21]

$$p(\mathbf{S} = \mathbf{s}) = \frac{1}{Z} \exp \left\{ -\sum_{c \in \mathcal{C}} u(\mathbf{s}_c) \right\},$$

where the sum is taken over all sets of variables (cliques) c whose any pair are neighbors, $\mathbf{s}_c$ denotes a vector composed of the variables in c, u denotes a clique potential, and Z denotes a scalar partition function that normalizes the probability distribution. Since this representation decomposes the distribution into a product whose factors are each determined by a small set (clique) of variables only, it is used during the probability maximization procedure. One disadvantage of Markov random fields is that its manual design is tedious and error prone especially for complex knowledge bases as are required for the evaluation of traffic scenes. To overcome this problem, a special variant of Markov networks with only binary random variables is used. This *Markov logic network* is defined as a set of pairs (F, w) where the first component F is a formula in first-order logic involving only variables in one clique and the second component w is a real number. Loosely speaking F are weak rules, e. g. *'Vehicle x will most likely conduct a lane change manoeuvre if its own lane is blocked and the adjacent lane has a suitable gap.'*, and w quantifies the belief in this rule. Hence Markov networks allow the explicit formulation of rules and can cope with sporadic violations of these rules without becoming inconsistent. Inference machines that solve for desired probabilities for a given grounding in Markov logic networks are available (cf. [22]).

Figure 7 shows a graph illustration of a simple Markov logic network, where the binary random variables (validity of first order formulas) form the nodes and weighted neighborhood relationships are marked as edges. Markov logic networks are related to situation graph trees as proposed in [23] and generalize Bayesian belief networks in some aspects which can be considered as random fields on directed graphs. They have successfully been applied to lane change prediction [24]. The combination of a probabilistic reasoning framework like Markov networks with logic expressions has several advantages: Compared to classical logic, uncertainties
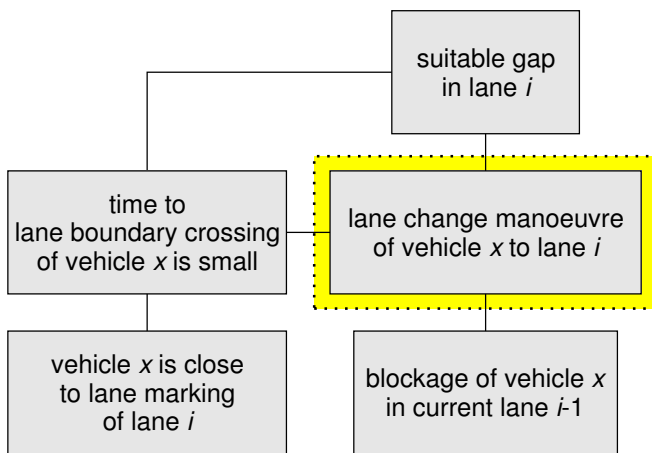
Fig. 7.   Simplified Markov network for a lane change manoeuvre

in measurements and contradictory rules can be resolved. Compared to Markov random fields on the other hand, the representation of rules using a logic notation makes the system comprehensible to humans: common tools for ontology engineering like *Protégé* [4] become applicable and simplify the design and maintenance of the knowledge base.

IV. COOPERATIVE PERCEPTION AND BEHAVIOR

As the equipment rate of vehicles with capabilities for environmental sensing increases, it becomes likely that a vehicle within a group possesses information about the environment that is relevant to others. Hence through exchange of information via car-to-car communication individual vehicles may enhance their field of view as well as the degree of accuracy and plausibility of the sensed information. Furthermore, vehicles can augment their scene perception by intentions communicated by cooperating traffic participants. It is worth noting from Fig. 2 that cooperative sensing does not require a 100% equipment rate, but provides benefit even at moderate rates. Preliminary experiments with cooperative perception between vehicles have recently been reported [25]. An important issue in this context is the spatio-temporal registration of data transmitted in the coordinate system of other vehicles. Since the uncertainty of the spatio-temporal alignment adds itself to the intrinsic uncertainty of the sensor information, this alignment must be conducted with high precision. It is shown that an alignment strategy that combines the coarse localization information of a GPS system with the sensor output of the video sensor itself yields good results for the envisaged application.

Cooperative perception and behavior generation impose significant requirements on communication. Due to the lack of a fixed infrastructure, communication relations between vehicles have to be set up 'ad hoc'. Demands such as high data rates, minimum delivery ratios and guaranteed maximum delays, commonly denoted as Quality-of-Service, have to be met. Scalability and QoS in self-organizing networks (see [26], [27]) are current research issues that have

not yet been generally answered. Especially in the context of cognitive vehicles, deterministic real-time behavior that assures observance of deadlines is an important prerequisite to enable distributed control. Also, aging of messages due to the ongoing perception has to be considered.

V. CONCLUSIONS AND OUTLOOK

Within the Transregional Collaborative Research Centre 28, the partners focus on systematic and interdisciplinary research on machine cognition of mobile systems as the basis for a scientific theory of automated machine behavior.

A software and hardware architecture that enables exchange of individual modules has been developed and implemented with of-the-shelf components. Emphasis has been laid upon the active camera platform that allows for 180° field of view 3D stereoscopic scene perception. The consortium operates in total six autonomous vehicles to validate and demonstrate cooperation perception and behavior in mixed traffic.

For the sake of rule-based and knowledge based cognition methods from artificial intelligence have been adopted. The combination of a probabilistic reasoning framework with a formal logic language enables a cognitive automobile to handle uncertainties in measurements and contradictory rules. Using ontological concepts for a detailed description of traffic scenes, this complex knowledge base stays comprehensible and maintainable.

Once groups of traffic have reached agreement on the perceived situation, they may negotiate to adapt their behavior cooperatively to the benefit of all. Emerging from successful experiments with cooperative city cars reported in [28], we are currently building dynamically self-organizing cooperative groups for cooperative passing and emergency brake manoeuvres.

REFERENCES

[1] J. Gayko, "CMS - Honda's collision mitigation system," in *Proc. IEEE Intelligent Vehicles Symposium*, Las Vegas (NV), USA, June 2005.
[2] A. Takahashi and N. Asanuma, "Introduction of Honda ASV-2 (advanced safety vehicle - phase 2)," in *Proc. IEEE Intelligent Vehicles Symposium*, Dearborn (MI), USA, June 2000.
[3] E. Dickmanns, R. Behringer, D. Dickmanns, T. Hildebrandt, and M. Maurer, "The seeing passenger car 'VaMoRs-P'," in *Proc. IEEE Intelligent Vehicles Symposium*, Paris, France, Oct. 1994, pp. 68 – 73.
[4] H.-H. Nagel, W. Enkelmann, and G. Struck, "FhG-co-driver: From map-guided automatic driving by machine vision to a cooperative driver support," *Math. and Computer Modeling*, vol. 22, pp. 101 – 108, 1995.
[5] C. Thorpe, *Vision and navigation - The Carnegie Mellon Navlab*. Kluwer Academic Publishers, 1990.
[6] Darpa, "Grand challenge 2005 official website," http://www.darpa.mil/grandchallenge05/index.html.
[7] Ü. Özgüner, C. Stiller, and K. Redmill, "Systems for safety and autonomous behavior in cars: The DARPA Grand Challenge experience," *IEEE Proceedings*, vol. 95, no. 2, pp. 397–412, Feb. 2007.
[8] J. Rasmussen, "Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models," *IEEE Trans. Systems, Man, and Cybernetics*, vol. SMC-13, no. 3, pp. 257–266, Mai/Juni 1983.
[9] Universität Karlsruhe - Institute for Metrology and Control Theory, "Team AnnieWAY Urban Challenge 2007 website," http://annieway.mrt.uni-karlsruhe.de/.
[10] Darpa, "Urban challenge 2007 official website," http://www.darpa.mil/grandchallenge/index.asp.

---

[4] http://protege.stanford.edu

[11] M. Goebl and G. Färber, "A realtime-capable hard- and software architecture for joint image and knowledge processing in cognitive automobiles," in *Proc. IEEE International Conference on Intelligent Vehicles*, 2007.

[12] T. Dang, C. Hoffmann, and C. Stiller, "Self-calibration for active automotive stereo vision," in *Proc. IEEE Intelligent Vehicles Symposium*, Tokyo, Japan, June 2006, pp. 364 – 369.

[13] H. Ruser and F. Puente León, "Informationsfusion - eine Übersicht," *Technisches Messen*, vol. 74, no. 3, 2007.

[14] J. Beyerer, F. Puente León, and K.-D. Sommer, Eds., *Informationsfusion in der Mess- und Sensortechnik*. Universitätsverlag Karlsruhe, 2006. [Online]. Available: http://www.uvka.de/univerlag/volltexte/2006/159/

[15] E. D. Dickmanns, "The development of machine vision for road vehicles in the last decade," in *Proc. IEEE Intelligent Vehicles Symposium*, vol. 1, 2002, pp. 268–281 vol.1.

[16] C. Stiller, "Cooperative environment perception," in *Proc. 1st Autocom Workshop on Preventive and Active Safety Systems for Road Vehicles*, Istanbul, Turkey, Sept. 2005, pp. 39 – 41.

[17] S. E. Palmer, *Vision Science: Photons to Phenomenology*, 3rd ed. The MIT Press, Cambridge, (MA), USA, 1999.

[18] G. Färber, "Biological Aspects in Technical Sensor Systems," in *AMAA05*, Berlin, Mar. 2005.

[19] S. Neumaier, P. Harms, and G. Färber, "Videobasierte Umfelderfassung zur Fahrerassistenz," in *4. Workshop Fahrerassistenzsysteme FAS2006*, Löwenstein, Oct. 2006.

[20] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 2nd ed. Morgan Kaufmann Publishers, San Francisco, (CA), USA, 1988.

[21] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Aanalysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.

[22] M. Richardson and P. Domingos, "Markov logic networks," *Machine Learning*, vol. 62, pp. 107 – 136, 2006.

[23] H.-H. Nagel, "Steps toward a cognitive vision system," *AI-Magazine*, vol. 25, no. 2, pp. 31 – 50, 2004.

[24] I. Dagli, M. Brost, and G. Breuel, "Action recognition and prediction for driver assistance systems using dynamic belief networks," in *Conf. Agent Technologies, Infrastructures, Tools, and Applications for E-Services*, 2002, pp. 179 – 194.

[25] K. Tischler, M. Clauss, Y. Guenter, N. Kaempchen, R. Schreier, and M. Stiegeler, "Net-worked environment description for advanced driver assistance systems," in *Proc. IEEE Intelligent Transportation Systems Conference*, Vienna, Austria, Sept. 2005.

[26] S. Chakrabarti and A. Mishra, "Qos issues in ad hoc wireless networks," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 142–148, Feb. 2001.

[27] H. Xiaoyan, X. Kaixin, and M. Gerla, "Scalable Routing Protocols for Mobile Ad Hoc Networks," *IEEE Network*, vol. 16, no. 4, pp. 11–21, July/August 2002.

[28] J. Baber, J. Kolodko, T. Noel, M. Parent, and L. Vlacic, "Cooperative autonomous driving - Intelligent vehicles sharing city roads," *IEEE Robotics and Automation Magazine*, vol. 12, no. 1, 2005.