Multi-sensor Data Fusion in Automotive Applications

Thomas Herpel[1], Christoph Lauer[1], Reinhard German[1] and Johannes Salzberger[2]
[1]Department of Computer Science 7 – Computer Networks and Communication Systems,
Friedrich-Alexander University, Erlangen, Germany
[2]Department of Safety Electronics,
Audi AG, Ingolstadt, Germany
{herpel, christoph.lauer, german}@informatik.uni-erlangen.de, johannes.salzberger@audi.de

Abstract

The application of environment sensor systems in modern – often called "intelligent" – cars is regarded as a promising instrument for increasing road traffic safety. Based on a context perception enabled by well-known technologies such as radar, laser or video, these cars are able to detect threats on the road, anticipate emerging dangerous driving situations and take proactive actions for collision avoidance. Besides the combination of sensors towards an automotive multi-sensor system, complex signal processing and sensor data fusion strategies are of remarkable importance for the availability and robustness of the overall system. In this paper, we consider data fusion approaches on near-raw sensor data (low-level) and on pre-processed measuring points (high-level). We model sensor phenomena, road traffic scenarios, data fusion paradigms and signal processing algorithms and investigate the impact of combining sensor data on different levels of abstraction on the performance of the multi-sensor system by means of discrete event simulation.

Keywords: multi-sensor data fusion, simulation, intelligent cars, environment perception, automotive

1 Introduction

Increasing road traffic safety and at the same time reducing the number of fatal car accidents is one of the most challenging future tasks for both car manufacturers and research institutions worldwide. Besides intelligent roadside infrastructures, advanced traffic routing and information services considerable effort is spend on enhancing the intelligence of individual vehicles within the traffic flow. Presently, sensor technologies well-known from other application areas like military or civil aviation are employed. Radar, laser, ultrasonic or video devices perceive information about the environment and possible threats around the vehicle either actively or passively. This significantly enhances the car's ability to anticipate dangerous driving situations and to act early and effectively in order to avoid a collision or at least mitigate the accident severity by proactive activation of adequate protection means. The quality of context perception by a set of environment sensors is of utmost importance for the so-called Advanced Driver Assistance Systems (ADAS) which rely on the sensor data. Important sensor properties that influence the quality of the environment perception include range, field-of-view (FOV), weather robustness, power consumption or placement constraints. Single sensor systems often have undesired weaknesses that suggest the use of multi-sensor systems. However, the sensor signal processing and fusion of sensor data from multiple devices is a sophisticated process including important design decisions regarding system performance and dependability. Various algorithms have been investigated for tasks of clustering measurement points, association of data with real world objects and filtering of sensor information [1]. Depending on the system's fusion paradigm the data integration takes place at a specific level of data abstraction. In a low-level data fusion near-raw data from various devices is combined at a very early stage of signal processing and the algorithms are applied to the conglomerate of measurement points. A high-level data fusion strategy pre-processes the data of the single sensors individually - i.e. each sensor is capable of dedicated clustering, association and filtering - and fuses the edited information, often represented as lists of detected objects. Both approaches are expected to have certain advantages and disadvantages in terms of entropy of information, computational complexity and adaptivity. In this paper, we use the concept of discrete event simulation for the analysis of a model consisting of various multi-sensor systems, sensor phenomena like reflection of radar or laser beams, road traffic scenarios and sensor data fusion strategies. The simulation results allow for a comparison on what data fusion paradigm, low-level or high-level, performs best in which scenario and is preferable with respect to maximum detection performance, robustness and reliability of proactive ADAS applications. The paper is organised as follows: Chapter 2 presents related work, chapter 3 introduces to the design of multi-sensor data fusion architectures and important techniques for context perception. Chapter 4 describes the implemented generic fusion system model. The simulation results are presented in chapter 5. Finally, chapter 6 concludes this paper and presents some areas of future work.

2 Related Work

Both industrial and scientific researchers spent considerable efforts on design and implementation of automotive environment sensor systems. Most work focuses on installation of sensor devices into real test cars and hot testing of various aspects, e.g. radar image acquisition [2], multi-target tracking [3], sensor-based cruise control [4], pedestrian protection [5] and devices for ADAS functions [6]. However, the benefits of modelling and simulating various aspects of sensors and sensor data processing in terms of saving costs, time and manpower are up to now exploited rarely, e.g. in [7] or in this paper.

3 Architectural Design

Safety-critical ADAS comprise of devices for context perception, data procession and actuators. This section briefly describes some of the key technologies for environment perception, related techniques for signal processing and approaches for obtaining more valuable information by multi-sensor data fusion.

3.1 Sensor Hardware

Technologies for context perception have been widely used for surveillance tasks in military, aviation and other fields. Some of these technologies (e.g. radar) are now being adopted for automotive applications. Other approaches (e.g. laser-based, vision-based or satellite-supported) extend the repository of available technologies and provide different advantages and disadvantages. The devices currently under investigation can be divided into three classes: active sensors, passive sensors and data dissemination techniques.

3.1.1 Active Sensor Systems

Sensor devices actively probing the environment, like radar devices (long-range radar (LRR) and short-range radar (SRR) sensors) and laser-based devices (e.g. lidar, laser-scanner (LS), photonic mixer devices (PMD) or closing velocity (CV) sensors), are able to detect targets based on reflections of an emitted signal. The target's distance can be derived directly by measuring the time-of-flight between pulse emission and pulse response. An intrinsic disadvantage of active sensors is their incapability to distinguish between targets that are relevant for the system and clutter targets that are, by definition, unimportant to the system. Clutter targets can be physically present, but nevertheless uninteresting, targets, like trash cans or rail roads, or ghost objects which are caused by signal interference. Some sensor types, like laser-scanners, produce enough measurements to allow a pattern matching approach to exclude objects with a geometry that suggests them to be clutter. Radar devices suffer particularly from clutter objects but provide the intrinsic possibility to perceive the target's relative velocity. Laser-based sensors are not able to detect this information directly but have to derive velocity information via the target's range rate. Moreover, laser-based devices are more vulnerable to bad weather conditions. Acoustic devices (e.g. ultra-sonic sensors), which have been widely used in automotive applications like parking aids, suffer from very short sensing ranges and negative influences of loud driving noise and therefore play a minor role in safety-critical ADAS.

3.1.2 Passive Sensor Systems

Contrary to active sensors, passive sensor devices do not emit any probing signal but passively perceive the environment. Especially vision-based devices (e.g. mono, stereo, night-vision and infra-red cameras) receive a large amount of attention from researchers and engineers. Vision-based devices offer excellent means for target classification but provide inaccurate distance information and are also affected negatively by bad weather conditions.

3.1.3 Data Dissemination

The third class of context perception devices utilizes the data dissemination via a wireless communication channel (e.g. Car-2-Car, Car-2-Infrastructure) or on non-volatile storage devices (digitally stored maps). It is obvious that such devices can offer a wealth of information for vehicle occupants but are of minor interest due to hard real-time communication constraints of safety-critical ADAS.

3.2 Signal Processing

The signal processing of fusion systems incorporates the refinement of raw sensor data to an abstract context description as well as the actual fusion step. Depending on the fusion paradigm sensor data can be fused on a low level of abstraction, high level of abstraction or on multiple levels of abstraction (hybrid approach) [8]. Figure 1 illustrates the signal processing chain for low-level and high-level fusion architectures.

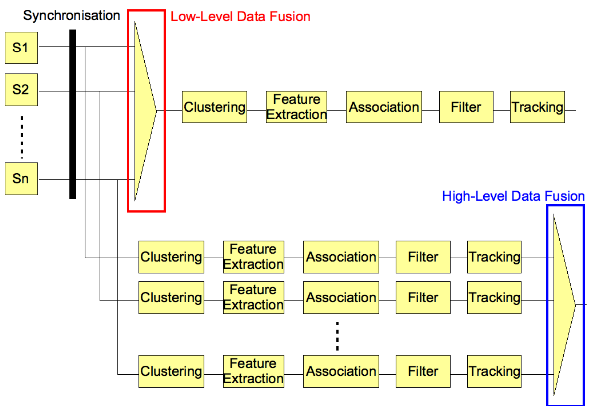

Figure 1: Low-level and high-level sensor data fusion

Before data integration can take place a common temporal and spatial base of the sensor data needs to be established through data synchronization. Objective of the clustering step is the assembly of groups of data points that are believed to be originating from the same real world entity. The data clusters are now analysed by the feature extractor to derive relevant attributes of the clustered entities, e.g. distance, velocity, acceleration, width, height, length, weight, yaw rate and others. In order to establish a timely correlated record of the objects, the association of objects from the last time step with the currently present objects is necessary. To lessen the effect of sensor inaccuracy and signal noise a filter algorithm is executed which also provides the state predictions for the next iteration's association step. High-level fusion architectures integrate context data based on associated and filtered sensor data whereas low-level fusion architectures integrate sensor data at an early stage of signal processing.

3.2.1 Algorithms for Context Perception

Most research concerning multi-sensor data fusion for target detection and tracking has been conducted in the fields of track association and signal filtering. However, these topics are still the most controversially discussed. The de-facto standard for automotive filtering applications is the Kalman-filter or variations of it [9]. In an iterative process the filter calculates a weighted state update $\overline{x}_k$ of the state prediction $\overline{x}_k^-$ from the last iteration, based on the actual measurement $\overline{z}_k$ and the Kalman gain matrix K_k (equation 1). The Kalman gain is derived from the signal covariance values. Other emerging filter approaches include, for example, the particle filter.

$$\overline{x}_k = \overline{x}_k^- + K_k(\overline{z}_k - C\overline{x}_k^-) \tag{1}$$

The association of objects from two different time steps can be done using greedy approaches (Nearest Neighbor (NN) and variants), Joint Probabilistic Data Association (JPDA) or Multiple-Hypothesis Testing (MHT). These algorithms differ in terms of computational load and ability to find local (iterative NN) or global optimal (JPDA, MHT) solutions. Moreover, JPDA and MHT take multiple possible object updates into consideration whereas NN approaches assume that a track can only be associated with exactly one object update.

3.2.2 Fusion Paradigms

As mentioned before, different approaches of fusing data from multiple sources are possible. The level of abstraction on which the data integration takes place is what we call fusion paradigm. Three main fusion paradigms can be observed in current research proceedings: *Low-level fusion* of raw sensor data, near-raw sensor data or sensor data that has been processed little, *high-level fusion* of sensor data that was processed independently to a high level of abstraction and *hybrid fusion* of sensor data on multiple levels of abstraction.

Some suspect that the potential benefit that can be obtained by data fusion is higher using low-level approaches [10]. From the perspective of information theory, the information content of the sensor data repository is lower on high levels of abstraction, thus reducing the synergetic opportunities for a sensor data fusion on this higher level. More obvious properties of low-level fusion architectures include higher computational load, higher communicational load, standardisation issues and reduced modularity and scalability compared to high-level architectures. Hybrid architectures promise to provide the advantages of both low-level and high-level paradigms. Note that not every sensor-set can be fused using any of the above mentioned paradigms. For example, a fusion of raw video data and low-level radar data is physically counterproductive. Lately many publications provided insight into various architectures, fusion paradigms and case studies [2-6, 8, 10]. However, analyses of the actual performance of the different systems is difficult as no common test strategy exists and different algorithms, sensor technologies and fusion paradigms have been used. Another problem emerges from the experimental research methodology itself. Although reference systems are under development it is not a trivial task to compare the outcomes of the sensor data fusion with the real world situation (ground truth).

3.2.3 Fusion Benefit

In [11] the definition of sensor data fusion includes a definition of 'greater quality'. In the automotive context there are two major benefits data from multiple sensors is fused for:

Enhanced coverage: Data from sensors with (partially) disjoint fields-of-view are fused in order to achieve a larger overall coverage. This co-operative (or complementary) fusion is comparable to a 'logical-OR' operation on sensor data and is often necessary to enable sufficient coverage for long and short range applications.

Increased confidence: Data from sensors with (partially) joint fields-of-view is used to validate object detections of the other sensor(s). This approach is also called competitive fusion, comparable to a 'logical-AND' operation on individual observations.

Safety-critical applications require highly dependable system behaviour and therefore themselves rely on dependable subsystems. Attributes of dependable systems include availability, safety and integrity, hence, both of the above mentioned fusion benefits are required to provide dependable system behaviour.

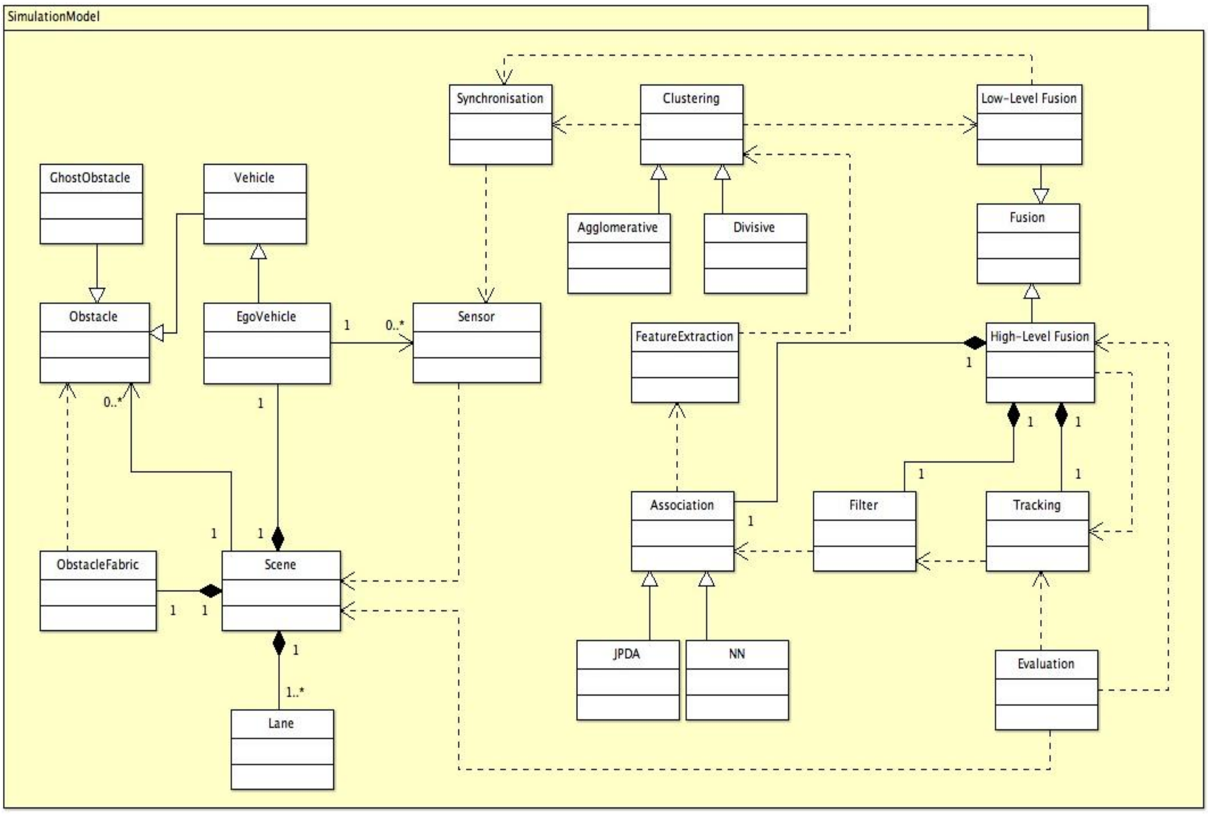

Figure 2: Basic UML-diagram of the simulation model

4 Simulation

Although some studies that investigate the performance of separate subsystems (like filter-algorithms) using simulation approaches have been conducted, most of the research in automotive multi-sensor data fusion relies on experimental methods. As authentic and valuable this approach is, it lacks flexibility in investigating fusion paradigms and sensor-sets. In this paper we use discrete event simulation to evaluate the performance of fusion architectures. The simulation model was created using the multi-method simulation tool AnyLogic 5.5.4[1] enabling real-time UML behavioural descriptions and full Java support.

4.1 Simulation Model

The fusion system model was created in a fashion that allows a thoroughly analysis of the signal processing elements and fusion paradigms (cf. figure 1), different sensor-sets and different traffic situations. The UML-diagram in figure 2 depicts the simulation model design. The central components of the simulation include the scene generator, the signal processing and fusion classes and the evaluation class. The scene generator is responsible for spawning objects (relevant and clutter) into the context of the ego-vehicle and for the creation of reflection points which are generated taking into account the sensor-set's individual sensor properties (FOV, range, lateral resolution, sensor variance) and physical measuring principle(s). The signal processing elements and the fusion elements perform environment perception based on these reflection points. Note that different implementation of the data refinement steps as well as different fusion paradigms are available for the analysis. The evaluation of both high-level and low-level fusion architectures is performed simultaneously and relies on ground truth information provided by the scene generator. The complexity and modularity of the model permits diverse simulation scenarios with different sensor-sets, traffic scenarios and fusion paradigms and provides excellent means for performance analyses of the complete fusion architecture.

4.2 Simulation Settings

We conducted several simulation experiments to evaluate the impact of the sensor data fusion paradigm on the multi-sensor system behaviour. Table 1 shows the scenario settings for urban and motorway traffic. Objects are generated uniformly according to the object spawn rate with a specific fraction of ghost or clutter objects. Note that the settings result in a denser and more clutterous scene for urban traffic scenarios.

[1] http://www.xjtek.com/anylogic/

Table 1: Parameter selection for simulation scenario

Parameter	Urban	Motorway
Object spawn period	uniform ($\sigma = 0.3$)	uniform ($\sigma = 0.12$)
Ghost fraction	0.45	0.35
Max. objects in scene	5	6
Number of lanes	2	3
Lane width	3 m	3.5 m
Scene horizon	50 m	80 m
Ego-vehicle speed	40 km/h	110 km/h
Lane speed difference	10 km/h	20 km/h

In table 2, the parameter settings for the sensor devices are summarised. The range denotes the upper limit of longitudinal detection ability and azimuth denotes the angle of the beam. The σ-values refer to the setting of the Gaussian probability distribution with mean zero that models the sensor variance, a higher value results in more inaccurate sensing.

Table 2: Parameter selection for sensor devices

Sensor	Range	Azimuth	σ_x	σ_y	σ_{vrel}
LRR	120 m	9°/10°	0.45	0.15	0.3
SRR	40 m	60°	0.5	0.2	0.5
Lidar	120 m	10°	0.15	0.15	--
LS	100 m	100°	0.11	0.11	--
PMD	40 m	55°	0.25	0.25	--
CV	20 m	45°	0	0.15	--

Out of these devices, we arranged four multi-sensor sets according to table 3:

Table 3: Multi-sensor sets for simulation

	Long-range	Short-range
Sensor Set 1	2 x LRR[2]	LS
Sensor Set 2	Lidar	SRR
Sensor Set 3	LRR	PMD
Sensor Set 4	LRR	CV

In order to increase the detection ability, each set consists of devices with different physical measuring principles. Sensor data is processed and fused according to the strategies described in section 3.2, using agglomerative clustering, generic feature extraction, JPDA and Kalman filtering. For both low-level and high-level the enhanced coverage strategy is applied.

5 Results

The simulation outcomes are presented and discussed in the following. The results were obtained by at least ten independent replications of single simulation experiments, applying a simulation control mechanism according to [12], requiring the stop criterion in equation 2 to be met before the simulation is aborted.

$$\frac{CI(x_k)}{mean(x_k)} \leq \frac{\varepsilon}{1+\varepsilon} \qquad (2)$$

[2] The LRR devices are installed with a horizontal displacement in the left and right half of the bumper.

CI(x_k) is the 95% confidence interval, i.e. the interval in which 95% of the measurements of the k-th experiment reside. Note that the individual measurements themselves are mean values of a particular variable. *mean(x_k)* is the mean value of the measurements of the k-th experiment and represents a normalisation of the confidence interval. The right part of equation 2 describes the desired maximum relative error ε, which was set to 10%.

The performance measures we obtained are the mean overall error in detected object position (in metres) and the mean false positive detection ratio. The former one characterizes the accuracy of the system. It is computed according to formula 3 as the mean value of all Euclidian Distances between actual coordinates (r_x, r_y) of objects in the scene and the corresponding positioning information (o_x, o_y) provided by the multi-sensor system, i.e. the instances of the mapping relation ◊.

$$Error = \frac{\sum_{\forall o,r|o\Diamond r} \sqrt{(o_x - r_x)^2 + (o_y - r_y)^2}}{size(\Diamond)} \qquad (3)$$

The mean false positive detection ratio resembles the robustness of the system against misinterpretation of the scene and thus propagation of irrelevant clutter objects, e.g. manhole covers, or creation of ghost objects by inappropriate signal processing.

The outcomes are shown as box-whisker plots in figure 3 and 4 and numerically in tables 4 and 5, respectively. Each figure is subdivided into parts for urban traffic (columns 1-4) and motorway traffic (columns 5-8) with each column providing the results of both low-level (LL) and high-level (HL) data fusion for a specific multi-sensor system (SET 1-4).

Obviously, the low-level fusion paradigm performs consistently better with regard to both accuracy of positioning and detection reliability. This corroborates the assumption that fusing the sensor data early (at a low-level of abstraction) is preferable to a high-level fusion of individually pre-processed sensor data.

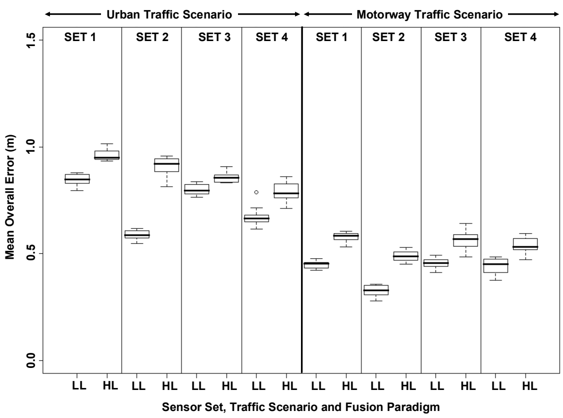

Figure 3: Mean overall error in object position

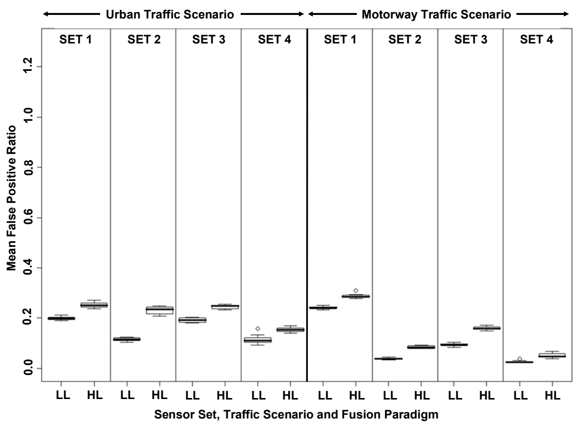

Figure 4: Mean false positive detection ratio

Table 4: Mean overall error (in metres)

	Urban Traffic				Motorway Traffic			
Set	1	2	3	4	1	2	3	4
LL	0.84	0.58	0.79	0.66	0.45	0.32	0.45	0.45
HL	0.95	0.92	0.85	0.78	0.58	0.48	0.56	0.53

Table 5: Mean false positive detection ratio

	Urban Traffic				Motorway Traffic			
Set	1	2	3	4	1	2	3	4
LL	0.19	0.11	0.19	0.11	0.24	0.03	0.09	0.02
HL	0.25	0.23	0.24	0.15	0.28	0.08	0.15	0.04

In most cases, the dense urban traffic leads to increased positioning errors and false detection rates compared to the motorway scenario, which is likely to happen with more objects in the scene and in turn smaller distances and diffuse boundaries in-between.

6 Conclusions and Future Work

In this paper we show that both high-level and low-level fusion paradigms can be used for reliable context perception for ADAS. Using discrete event simulation we analysed various sensor-sets, traffic scenarios and fusion paradigms. The simulation model includes implementation of state-of-the-art techniques for sensor data fusion and implicit knowledge of ground truth information through traffic scene generation to support analyses of different architecture designs. The simulation output analysis shows advantages of the low-level fusion paradigm in both overall tracking accuracy and false positive detections. More clutterous and denser traffic scenarios result in higher tracking inaccuracies for both fusion paradigms, while the false positive detection rate remains almost unaffected.

The presented simulation model allows for analyses of fusion paradigms which are difficult to perform using experimental research methods. More sensor-sets, impact of specific algorithms and fusion benefits (extended coverage, increased confidence) can be evaluated using this model. An implementation of context perception algorithms for passive sensor devices would increase the extent to which fusion architectures can be investigated. Progress in ground truth reference measurements could enable experimental validation of the simulated system characteristics. Furthermore, new concepts of low-level and early sensor data fusion could be investigated and help automotive engineers in understanding the opportunities of multi-sensor data fusion and related design decisions.

7 References

[1] D.L. Hall and J. Llinas. *Handbook of multisensor data fusion*. Boca Raton, FL: CRC Press, 2001.

[2] U. Meis and R. Schneider, “Radar Image Acquisition and Interpretation for Automotive Applications”, *Proc. of the IEEE Intelligent Vehicle Symposium*, Columbus, Ohio, USA, 2003, pp. 328-332.

[3] R. Mobus and B. Kolbe, “Multi-Target Multi-Object Tracking, Sensor Fusion of Radar and Infrared”, *Proc. of IEEE Intelligent Vehicle Symposium*, Parma, Italy, 2004, pp.732-737.

[4] G. Widmann, M. Daniels, L. Hamilton, L. Humm, B. Riley, J. Schiffmann, D. Schnelker and W. Wishon, “Comparison of Lidar-Based and Radar-Based Adaptive Cruise Control Systems“, *Proc. of SAE World Congress*, Detroit, Michigan, USA, 2001, pp. 1-14.

[5] M. Meinecke, M. A. Obojski, M. Töns, R. Dörfler, P. Marchal, L. Letellier, D. Gavrilla and R. Morris, “Approach for Protection of Vulnerable Road Users Using Sensor Fusion Techniques“, *Proc. of International Radar Symposium*, Dresden, Germany, 2003.

[6] J. Langheim, A. Buchanan, U. Lages and M. Wahl, “CARSENSE – New Environment Sensing for Advanced Driver Assistance Systems“, *Proc. of IEEE Intelligent Vehicle Symposium*, Tokyo, Japan, 2001, pp. 89-94

[7] M. Bühren and B. Yang, “Simulation of Automotive Radar Target Lists Considering Clutter and Limited Resolution”, *Proc. of International Radar Symposium*, Cologne, Germany, 2007, pp. 195-200.

[8] A. Amditis, Ed., *PReVENT Fusion Forum e-Journal*, vol. 2, 2008.

[9] A. Gelb, *Applied Optimal Estimation*, MIT Press, 1974.

[10] T. Tatschke, “Early Fusion,” *PReVENT Fusion Forum e-Journal*, vol. 1, pp. 6–7, 2006.

[11] L. Wald, “Definitions and terms of reference in data fusion,” *International Archives of Photogrammetry and Remote Sensing*, vol. 32, no. Part 7, pp. 651–654, 1999.

[12] A.M. Law and W.D. Kelton, *Simulation Modeling and Analysis*, McGraw-Hill Higher Education, 1997.