DETECTION OF ROAD USERS IN FUSED SENSOR DATA STREAMS FOR COLLISION MITIGATION

L. Walchshäusl, R. Lindl, K. Vogel

BMW Group Research and Technology
Hanauerstr. 46, 80992 Munich, Germany
{leonhard.walchshaeusl, rudi.lindl,
katrin.vogel}@bmw.de

T. Tatschke

FORWISS, University of Passau
Institute for Software Systems in Technical
Applications of Computer Science
Innstr. 43, 94032 Passau, Germany
tatschke@forwiss.uni-passau.de

ABSTRACT

This paper deals with a novel sensor fusion approach to detect and track cars and pedestrians to facilitate a collision mitigation application for vehicles. Robust collision mitigation requires a perception performance of an unprecedented degree of reliability, since an erroneous application of emergency braking caused by false alarms would greatly impede road safety improvement not lastly due to the major setback such an incident would represent for driver acceptance. However, current off-the-shelf single sensor approaches can hardly fulfil the challenging demands. Accordingly, we develop a multi-sensor recognition system. It is composed of a far infrared imaging device, a laser scanner and several radar sensors, which operate integrated into a BMW sedan.

1. INTRODUCTION

Statistic evidence of the European Union shows that accidents resulting in fatalities or serious injuries are caused to the highest percentage by collisions of cars with vulnerable road users. This fact points to the urgent need for active and passive automotive safety systems as a significant contribution to the overall road safety.

Focusing on a novel approach for environmental perception based on a multi-sensor system this paper offers a collision mitigation application for cars by means of autonomous braking. To meet the application's requirements regarding accuracy and reliability of the perception result, we propose a fusion processing scheme, which operates only on slightly pre-processed sensor data. This "early fusion" approach uses the synergetic effect of a common and consistent data processing as well as an interpretation of sensor low-level data to tap almost the full sensor potential. To this end it strongly relies on a combined modelling of the environment, which contains object assumptions as well as a-priori knowledge.

1.1. Related Work

Many research groups have contributed significantly in the field of sensor fusion. With regard to an automotive environment Kämpchen et al. [KFD04] propose an adaptive cruise control system (ACC) based on a combination of an early and a track-based fusion approach. A laser scanning sensor and an imaging camera are used to detect vehicles. Schweiger et al. [SNR05] utilize a radar and a monocular imaging sensor to build an ACC system. A collision warning and vision enhancement system is proposed by Polychronopoulos et al. [PST04]. Vulnerable road users and vehicles are identified by a far infrared camera and a radar sensor.

In this paper the multi-sensor perception system is composed of four radar sensors, a laser scanning device and a far infrared camera to detect both vulnerable road users and other vehicles utilizing a novel early fusion approach.

1.2. Overview

This publication focuses on a novel fusion approach based on only slightly pre-processed sensor data. The

given sensor platform and its configuration is discussed in chapter 2. The envisaged safety application on top of the perception system is presented in chapter 3. Chapter 4 is dedicated to the perception system. After a short motivation with respect to the early fusion concept (4.1), section 4.2 is giving an overview on the fusion cycle and the system's main components. In the following sections (4.2.1 to 4.2.3) the data acquisition as well as the time prediction of our system is introduced, and the generation of predicted measurements is explained in detail. Sections 4.2.4 and 4.2.5 illustrate the data association process and the generation of new object hypotheses within the proposed fusion framework before the filtering of the object information is addressed (section 4.2.6). Finally, the last section gives an overview on the system architecture and implementation details of the fusion system.

2. SENSOR CONFIGURATION

BMW has set up an experimental car with the following sensor configuration depicted in figure 1.
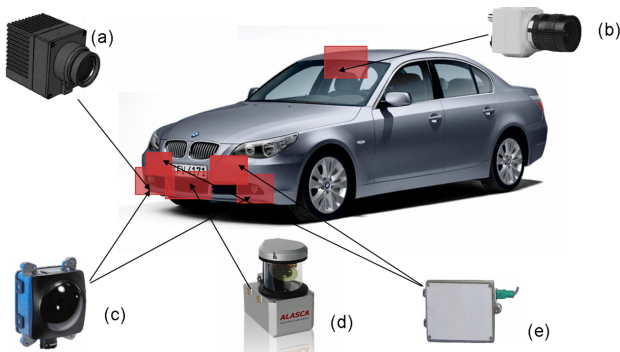
Fig. 1. BMW experimental car equipped with the following sensor configuration: (a) SAGEM far infrared camera, (b) Grey-scale camera, (c) Bosch SGU long range radar, (d) IBEO laser scanner and (e) MA/COM short range radar.

Concentrating on the surveillance of the area in front of the vehicle, these cooperative sensors, which operate on the basis of distinct physical principles, complement each other both in effective range and spatial accuracy.

The usage of a far infrared (FIR) sensor guarantees both perception at bad lighting conditions and straightforward vehicle and pedestrian detection (see figure
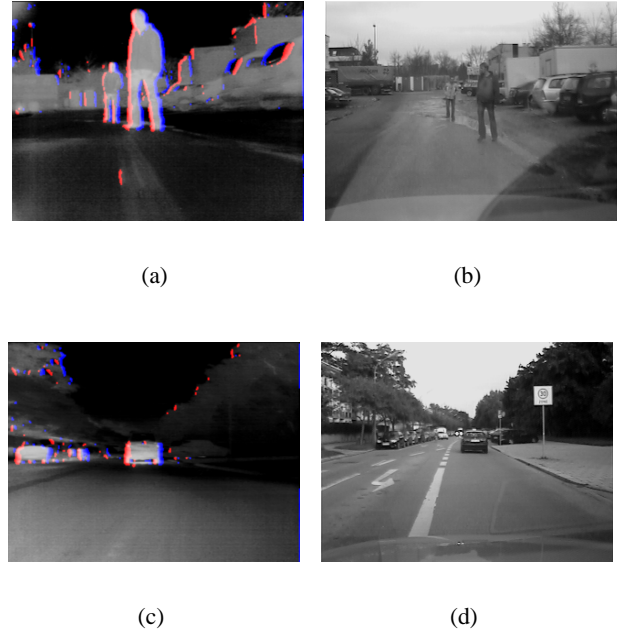
Fig. 2. (a) Pedestrian (FIR camera). (b) Pedestrian (grey-scale camera). (c) Vehicle (FIR camera). (d) Vehicle (grey-scale camera).

2). As most pedestrian-scenarios covered by the experimental vehicle, are situated in the area to the right of the road, this sensor is mounted at the right of the frontal bumper. Long and short range radar sensors are surveying the environment ahead providing a seamless transition in distance and field of view resolution. Moreover, a laser scanning (lidar) device is mounted beneath the number plate to enhance the detection and tracking quality for both pedestrians and vehicles. The visual grey-scale cameras are used for supervising and controlling purposes only.

3. COLLISION MITIGATION APPLICATION

The target application of the demonstration system is collision mitigation by means of autonomous braking. Accident statistics tell us that most of the accidents with severely injured persons happen through collisions of cars with vulnerable road users – often in urban areas on straight roads. These accidents can be attenuated or even prevented by our collision mitigation system. The second crash scenario addressed in our system deals with rear-end collisions.

The basis for any intra-system decision is a situation assessment. Taking into account the geometric and kinematic data of object-models provided by the perception system as well as probabilistic attributes and physical limits the collision risk is estimated. Only in case of an inevitable collision, the system engages the brakes autonomously. It uses the fact that technical systems are capable of reacting much faster than human beings.

Emergency braking caused by false alarms would greatly impede road safety improvement not lastly due to the major setback such an incident would represent for driver acceptance. Thus such an active autonomous intervention in the process of driving requires an outstanding degree of perception performance, particularly with regard to accuracy, availability and robustness. Therefore the attention is especially concentrated on the construction and design of the environment perception system.

4. PERCEPTION SYSTEM

The central challenge for advanced driver assistance systems is an adequate perception of the vehicle's environment and the generation of an environment description.

A multi-sensor system containing a set of sensors based on distinct physical principles establishes a basis for an accurate and robust environment perception – notably if they complement each other in their sensing capabilities. Key ingredient of the perception system is the way how the diverse and sometimes conflicting measurement data from different sensors are combined in order to increase the information content on the one hand and to reduce the amount of data on the other hand.

In track-based fusion approaches several sensor data streams are processed independently from each other until the level of object data is reached. Fusion on the object level runs the risk that useful data is discarded during early processing steps (e.g. data reduction or feature extraction) because it doesn't seem to be significant enough from one sensor's point of view. That way contradictions could arise on object level, that are difficult to be resolved due to the lack of lower level information.

Therefore it is this paper's standpoint that the "early fusion concept" is the more promising approach to exploit the synergies of the different sensor data.

4.1. Early Fusion Concept

The term "early" means to combine data provided by multiple and even diverse sensors at an early stage of the data processing chain. These input data can be slightly pre-processed – limited to untracked and raw sensor data.

During the subsequent fusion algorithm, data from one sensor is assessed with regard to the relevance of its information, always in the light of data provided by other sensors.

The early fusion approach ensures consistency of models in the whole processing chain. In particular, one common environment model is used for describing the same aspect of reality seen from different sensors. Thus the whole sensor data contributes to one global environmental description, i.e. one outgoing result. With this early fusion approach it is expected to achieve a robust and reliable output of the environment perception system.

4.2. Fusion Cycle

Generally speaking an abstract inference problem is composed of three circular steps namely time prediction, data association (data matching) and measurement update (correction) [FP02]. On top of this basic pattern we added further steps to come up with multi-sensor and multi-object demands. The following subsections describe the fundamental structure of the implemented fusion cycle, also condensed in figure 3.

4.2.1. Data Acquisition

As most of the used sensors are working on different clock rates and time is crucial in collision mitigation applications, we preserve a high time resolution by a semi-asynchronous data acquisition. The fastest sensor (master sensor) with respect to the refresh rate is used to trigger this step. The actual data acquisition is done by polling every sensor for new data. This raw data is stored in a sensor specific coordinate system, which is relative to the sensor's mounting position. Measurements of sensors, which work asynchro-
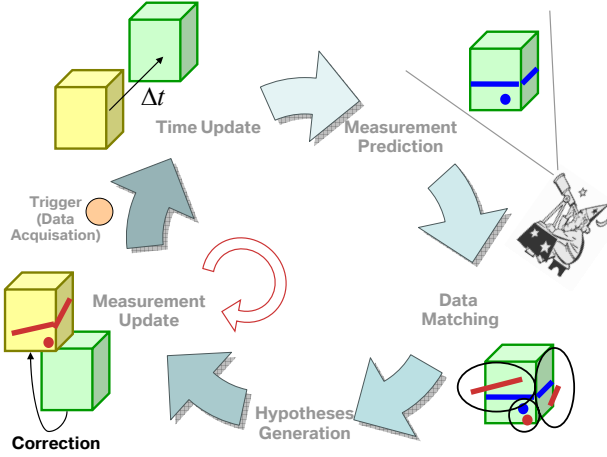
Fig. 3. Overview of Fusion Cycle. The cycle start is at the red circle. Yellow boxes symbolize tracked object-models. Green boxes and blue lines/circles represent predicted object-models respectively predicted measurements. Red lines/circles show real measurements.

nously to the master sensor, obtain a time delay as these measurements occur between the last and the current cycle. Assuming a maximum object velocity of $15m/s$ in urban environments, this error is negligible compared to the sensor specific errors. In addition, a slight pre-processing of the raw data, i.e. noise reduction or edge extraction, is performed. As the sensors have probably moved along with the own-car, a coordinate transformation from sensor specific to world coordinates can not be applied in this step. Nevertheless, this conversion is handled in the subsequent "time prediction" step.

4.2.2. Time Prediction

According to every objects' state (position, orientation, velocity ...) at the last cycle, these states have to be predicted to the current time. With the nomenclature from [BW95] the object-states $\hat{x}_k$ and their prediction error covariances P_k are projected one time step ahead via

$$\hat{x}_{k+1}^- = f(\hat{x}_k, u_k) \quad (1)$$
$$P_{k+1}^- = A_{k+1} P_k A_{k+1}^T + Q_k \quad (2)$$

on the basis of the objects' underlying dynamic models $f(\cdot)$ and their derivative A_k.

A single track dynamic model [Hie04] utilizing own-car sensor data like current velocity, steering angle and lateral acceleration is used to predict the own-car's position. As all other sensors are mounted to the own-vehicle, their position is directly deducible and a world coordinate transformation is performed. In combination with a global coordinate system, this simplifies and standardizes the subsequent time prediction step of tracked vehicles and pedestrians. Currently for these object-models a linear dynamic model is applied.

4.2.3. Predicted Measurement Generation

In the previous step for every object an updated representation (state) with regard to the current time is generated. These predicted states are the basis for the following step, which estimates what each sensor would measure under the assumption that every objects' state was correctly predicted. With the notation of [BW95], this process reflects the non-linear function $h(\hat{x}_k^-)$ where $\hat{x}_k^-$ represents the result of the "time prediction" step. In the following, a representation of the general task of the "predicted measurement generation" is given. The basic principle of this task is also shown in figure 4. Currently all object models are composed of simple polygons.
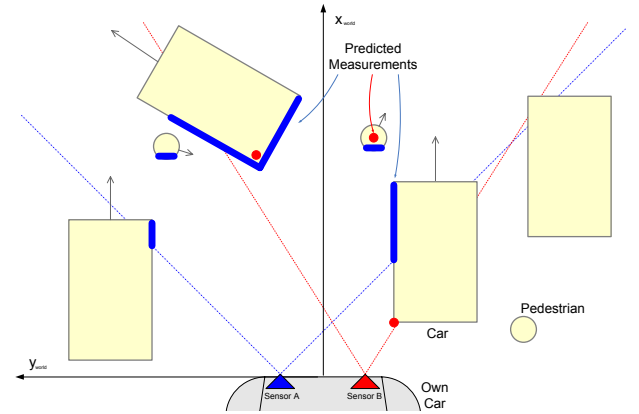
Fig. 4. Showcase for the predicted measurement generation. Using two different physical sensor principles predicted measurement generation is illustrated for both vehicles and pedestrian object-models, considering the sensors' view-port as well as partial occlusions. (Exemplified scene is in birds-eye-view and not in scale)

4

For all valid sensors within the current cycle the following steps are performed:

Sensor Transformation: Transform all known object-models into the sensor coordinate system.

Back-face Culling: Temporarily remove the objects' parts which are not "visible" to the sensor, because they are occluded by other parts of the object.

Clipping: All objects residing outside the view-port of the sensor are ignored.

**Occlusion Test:** Discard all objects which are "invisible" because they are occluded by other objects.

Measurement Generation: For the remaining objects respectively object parts the object specific predicted measurements are computed. This requires, that all object models have their own sensor specific representation.

Most of the necessary steps, like sensor coordinate transformation, clipping or occlusion testing, are strongly related to common computer graphics tasks. Therefore, we use a scene-graph representation for all object-models, which allow for an easy adaptation of these algorithms to the specific sensor characteristics.

4.2.4. Data Matching

The next step within the aforementioned fusion cycle is the data association that extracts and assigns corresponding pairs of real and predicted measurements. Due to the large data amount compared to most track-based fusion systems and the resulting complexity to determine the matching pairs, gating mechanisms are essential to support the fast finding of data correspondences. Thus, sensor data specific rectangular gates (derived from the object's shape model; see figure 5) are computed from the object's predicted estimation error covariance P_k^-, the Jacobian H_k of partial derivatives of the state-to-measurement function with respect to states and the sensor's measurement noise covariance R_k. Therefore

$$\epsilon = \gamma \sqrt{s_i} \tag{3}$$
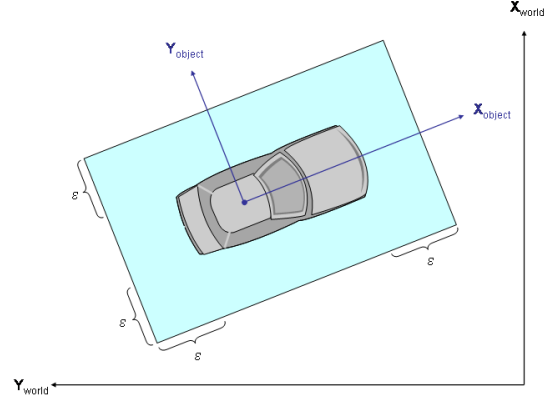
Fig. 5. Gate calculation of an object

is added to the object's length and width respectively on each side, where $\gamma \in (0; \infty)$ and s_i denotes the respective element in the innovation covariance matrix

$$S_k = H_k P_k^- H_k^T + R_k. \tag{4}$$

Now the real measurement data is tested for being inside one or more gates. Next, the data within the gating area of an object-model is assigned sensor-specifically to the object's predicted measurements (as calculated in 4.2.3) on the basis of a Global Nearest Neighbour approach.

4.2.5. Hypothesis Generation

A priority goal of the hypotheses generation is a direct and complete detection of all so far untracked and possibly relevant objects in the sensors' ranges. Thereto a high error of second kind is consciously taken into account. Usually a succeeding classification procedure as well as an observation of the objects over time can select and eliminate irrelevant assumptions. To limit the cost of computation, the hypothesis generation focuses to salient and unmatched sensor data in the detection range.

Currently the unmatched salient points, where new assumptions are placed, are radar responses, lidar segments within certain dimensions and vertical image edges from the far infrared imaging device. To limit the amount of assumptions a first coarse pre-classification step rejects impractical assumptions and a second aggregation step tries to combine overlapping hypotheses.

4.2.6. *Extended Kalman Filter*

A conventional Extended Kalman Filter (EKF) (see [AM79] for instance) has been chosen. It handles the nonlinearities of this application quite well. For every assigned pair of real and predicted measurement, which has been calculated before, a measurement update on the underlying object is performed. This procedure propagates the measurement information into the states of the respective objects. In doing so, the information of several measurements enhance the states by updating the objects' state values and furthermore, lowering the estimation error covariances. Thereby, for each assigned sensor data a measurement update step is conducted before the next cycle starts with the object's state prediction in time. With the notation of [BW95] the equations at time step k of the EKF's measurement update can be written as

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \qquad (5)$$
$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - h(\hat{x}_k^-)) \qquad (6)$$
$$P_k = (I - K_k H_k) P_k^-. \qquad (7)$$

Thereby the specific term $h(\hat{x}_k^-)$ has already been evaluated during the calculation process of predicted measurements and thus equation (6) can be written as

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - y_k^-) \qquad (8)$$

for every pair (y_k, y_k^-) of measurement and predicted measurement, matched by the data association process. As all sensor data is projected into the 3D global world coordinate system, the entries of the Jacobian H_k can be easily deduced from the underlying object-model without any further complex and time consuming calculations.

4.3. Implementation Details

A cyclic top-down architecture has been implemented to facilitate the detection, classification and tracking of relevant road users over time. The real world vehicle surroundings and the sensor configuration are reflected by a virtual environment, which is modelled as a hierarchical scene-graph structure [BW95], ensuring centralized data access and efficient spatial dependency processing (see section 4.2.3). To allow an efficient graph traversal as well as a decoupling of algorithm and data portions, the so called Visitor Design Pattern [BMRS96] has been used extensively.

5. CONCLUSION

This paper proposed a novel sensor fusion approach to detect and track cars and pedestrians in real-time to facilitate a collision mitigation application for vehicles. The system is composed of a far infrared imaging device, a laser scanner and several radar sensors which operate integrated into a BMW sedan. The proposed fusion framework in combination with the consistent use of global world coordinates for measurements, matching and the predicted measurement generation provides a high level of abstraction. Thus, more sensors at different view-ports up to a 360 degree sensor configuration could easily be implemented.

6. FURTHER WORK

There are still ample possibilities for improving the system. The most important ones to be tackled in future are to extend the system by a classification unit, to develop an auto-calibration of sensors and to apply alternative filtering approaches. In addition to these improvements, an extensive evaluation of the system performance is planned.

7. ACKNOWLEDGEMENT

8. REFERENCES

[AM79]   B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice Hall, Eaglewood Cliffs, NJ, 1979.

[BMRS96] Frank Buschmann, Regine Meunier, Hans Rohnert, and Peter Sommerlad. *Pattern-Oriented Software Architecture: A System*

of Patterns, volume 1. John Wiley and Sons Ltd, 1996.

[BW95]   B.D. Allen Gary Bishop and Greg Welch. Tracking: Beyond 15 minutes of thought: Siggraph 2001 course 11. Technical report, University of North Carolina at Chapel Hill, 1995.

[FP02]   David A. Forsyth and Jean Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, 2002. FOR d 02:1 1.Ex.

[Hie04]   Marcus Hiemer. *Model Based Detection and Reconstruction of Road Traffic Accidents*. PhD thesis, Universitätsverlag Karlsruhe, 2004.

[KFD04]   Nico Kämpchen, Kay Ch. Fürstenberg, and Klaus C.J. Dietmayer. Ein Sensorfusionssystem für automotive Sicherheits- und Komfortapplikationen. *Aktive Sicherheit durch Fahrerassistenz*, 2004.

[PST04]   Aris Polychronopoulos, Ullrich Scheunert, and Fabio Tango. Centralized data fusion for obstacle and road borders tracking in a collision warning system. *Seventh International Conference on Information Fusion*, 2004.

[SNR05]   Roland Schweiger, Heiko Neumann, and Werner Ritter. Multiple-cue data fusion with particle filters for vehicle detection in night view automotive applications. *IEEE Intelligent Vehicle Symposium*, 2005.