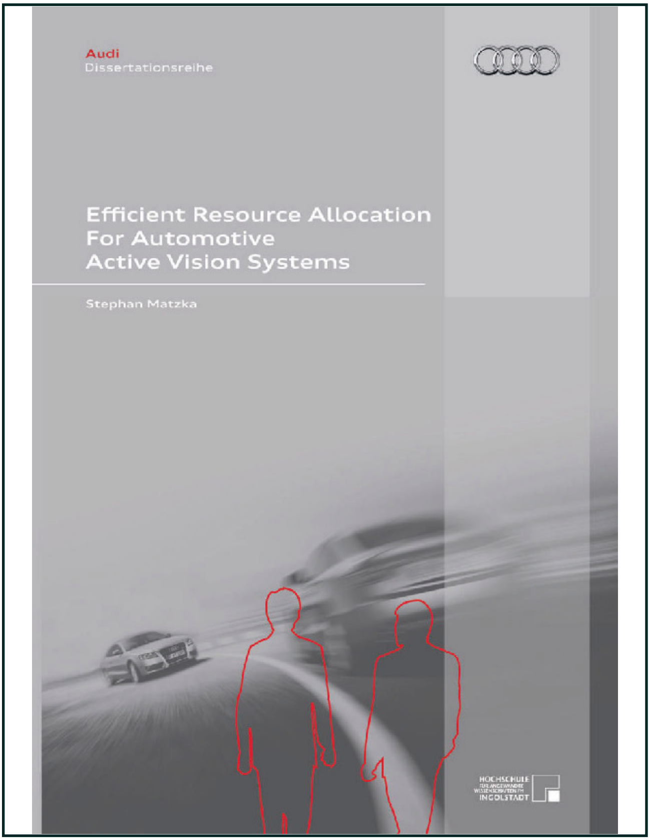
Stephan Matzka (Autor)

Efficient Resource Allocation for Automotive Active Vision Systems



https://cuvillier.de/de/shop/publications/852

1.2.3 Sensor Configuration of the Test-Vehicle

The proposed system obtains its sensor-level data from a multi-modal sensor system mounted on a test-vehicle provided by Audi AG consisting of

- two high-resolution video cameras
    - ○ one pan-tilt-zoom camera
    - ○ one fixed camera
- one 3-D camera (photonic mixer device, PMD)
- one time-of-flight laser scanner
- two short-range radars (SRR)
- one long-range radar (LRR)
- eight ultrasonic sensors (US)
- one differential global positioning system (DGPS)

In Fig. 1.3a the test vehicle provided by Audi AG is shown, in Fig. 1.3b the maximum detection distances and aperture angles of the different sensors is illustrated.

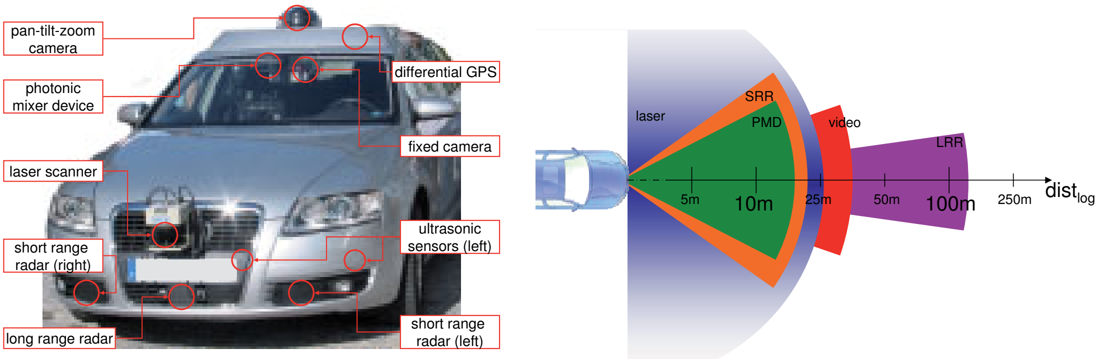


Figure 1.3: In a) the test-vehicle used to acquire road traffic scenes used for evaluation is shown. Figure b) the distance and aperture angles of the sensor array are drawn on a logarithmic scale.

Our selection of sensors from the test-vehicle's multi-modal sensor configuration is presented in section 1.3 below.

1.3 System Overview

In this thesis an effective resource allocation for an automotive active vision system is presented. A literature review shows that a variety of active vision systems exist. Despite the differences in the considered systems, all approaches encounter the same fundamental questions on system design:

- What is the overall system architecture?
- Which objectives are considered during optimisation?
- How are objectives determined for every candidate region?
- Which decision making concept is chosen to perform multi-objective optimisation?
- How is the complexity of the system reduced or, if this is not possible, handled?
- Is the system required and capable to fulfil real-time constraints?

Our active vision system processes data over various stages, beginning at sensor level and increasing both in level of abstraction and in significance towards allocation level (cf. Fig. 1.4). The system is organised using five levels of abstraction from sensor level towards allocation level:

1. Sensor level, containing raw sensor data representations.
2. Data level, containing the results of low-level sensor data processing.
3. Semantic level, containing semantic data resulting from high-level data interpretation
4. Reasoning level, containing combined semantic data to be used for. reasoning
5. Allocation level, containing the system's current resource allocation.

An increasing level of abstraction is highly desirable to maximise the system's efficiency, yet requires a set of serial processing steps. In order to mitigate the latency associated with serial processing, data processing tasks are run in parallel for every level of abstraction. Parallel processing requires the independence of the executed processes, which necessitates the use of individual data representation objects (drawn as parallelograms in Fig. 1.4) in every level of abstraction. Each representation object is updated by a single or multiple processes, providing data for subsequent processing steps.

The sensors used in our proposed system are a selection from the multi-modal sensor system mounted on a test-vehicle provided by Audi AG (cf. section 1.2.3). Our presented system performs computationally expensive tasks such as object detection and object

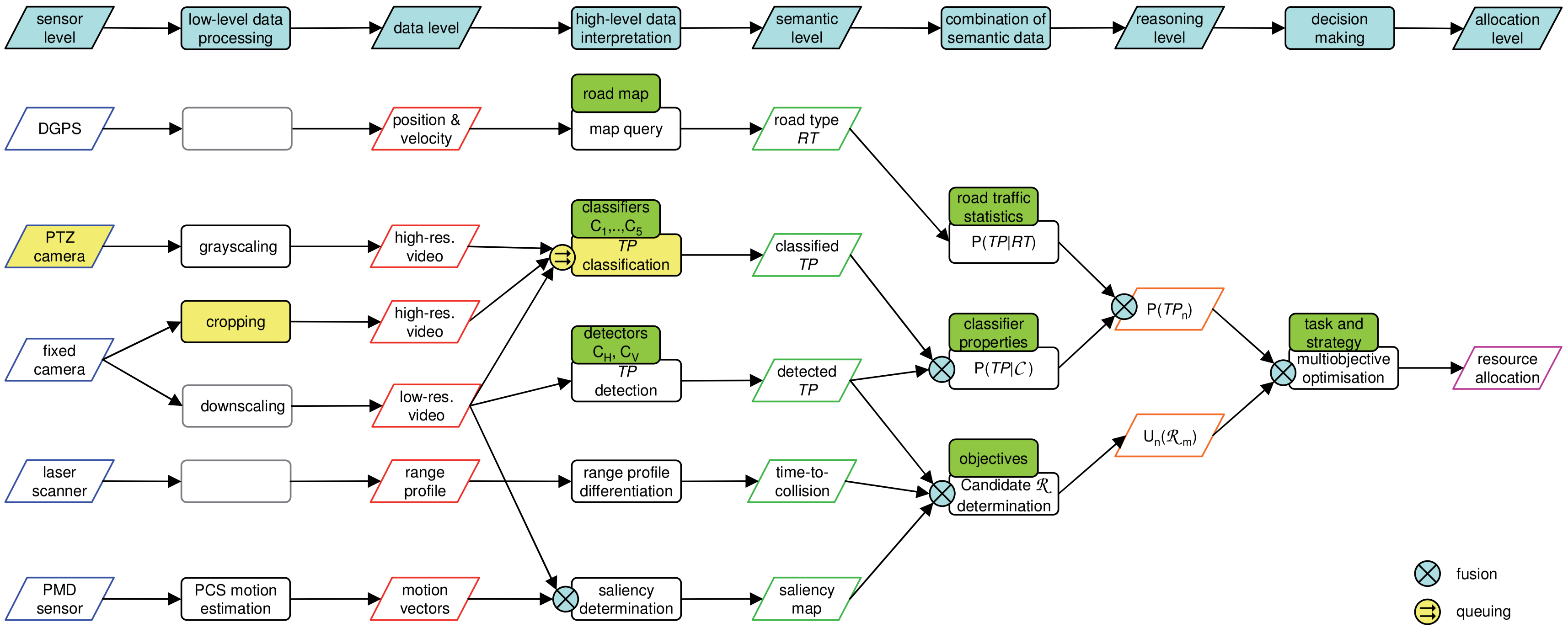


Figure 1.4: Block diagram of proposed system. Parallelograms indicate a data representation, whereas rounded boxes represent processing steps between two data representations (empty gray boxes indicate that no processing is necessary). Green containers indicate prior knowledge used in our proposed system. Circles represent pre-processing steps such as fusion (blue) and queuing (yellow). A yellow background colour indicates that the respective sensor or process is directly influenced by our resource allocation system, whereas all non-yellow sensors and processing steps work independently.

classification only on low-resolution data or on focused regions of high-resolution data. Controllable sensors acquire high-resolution data only for regions determined by the sensor-resource allocation concept, reducing the amount of sensor data in the system. This is in contrast to related systems discussed in section 2.4, where a high-resolution representation of the entire environment is required to determine regions of interest.

All processes inside the system are designed to fulfil soft real-time constraints, in the sense that all processes are designed to terminate within a given cycle time at which the proposed system operates. If a process exceeds the current cycle's deadline, the subsequent data representation object is not updated. Dependent subsequent processing steps will then pause until the update of the outdated data representation object is performed. This architecture also ensures that the most current data representation is made available to subsequent steps. Computationally inexpensive processes such as bottom-up saliency determination increase the system's reactivity to changes in the environment even if traffic participant detection or classification processes fail to terminate prior to the current cycle's deadline.

The proposed system is designed to avoid the problem of single points of failure. If any data representation object is outdated, the system is able to continue operation, albeit at the expense of decision making quality. Two system failures can be identified as most problematic, however. First, failure of traffic participant classification is critical because driver assistance systems are no longer provided with updated traffic participant positions and classes. Second, failure of the resource allocation system itself presents a problem, which is partly mitigated by a graceful degradation of both sensors and computational resources towards a static operation mode using a predefined resource allocation scheme.

The system design itself, two data processing methods, and the resource allocation process present original contributions and are discussed in section 1.4 below.

1.4 Contribution

The contribution of this thesis is divided into three aspects: the proposal of a novel system design for automotive vision system, extensions to sensor data processing algorithms, and the formalisation and evaluation of decision making in the resource allocation process.

System Design

Our proposed system is organised in five levels of abstraction, extending the four-layered architecture presented in Matzka *et al.* [12]. This architecture ensures that the amount of processed and transferred data decreases as the level of abstraction increases. The reduction of processed data lowers the computational demands on the vehicle's electronic control units and the reduction of transferred data reduces the load of the vehicle's bus system. In order to counter the latency caused by serial processing over multiple levels, processes within the same levels are run in parallel. In addition semantic information is made available to driver assistance systems in the third out of five levels, with both sensor level and data level processes designed to be computationally inexpensive.

Sensor Data Processing

We present two extensions to existing sensor data processing algorithms. First, an extension of the PMVFAST method to estimate 2-D motion vectors towards the PCS method is published in Matzka *et al.* [13] that efficiently estimates 3-D motion vectors in range maps is presented in section 4.5.2. Second, the use of a sparse input of single scanlines to be used in 3-D spin image object classification. The generation of suitable sparse scanlines is described and evaluated in Matzka *et al.* [14] and is presented in section 5.3. Beyond this, the fusion of correlated pre-filtered radar tracks is investigated in Matzka and Altendorfer [15, 16].

Formalisation and Evaluation of Resource Allocation

The central contribution of this thesis is the formalisation and evaluation of the decision making process required for resource allocation first presented in Matzka *et al.* [12], extending existing active vision systems discussed in section 2.4. Our proposed system is novel in the respect that it combines a formal, Pareto efficient decision making method with bottom-up and top-down information acquired using low-resolution data. This is in contrast to methods presented in the literature selecting regions of interest from high-resolution data. An optimum decision making strategy is determined and the problem of decision making complexity is solved by presenting efficient search heuristics to determine the allocation with the highest estimated utility.

1.5 Thesis Outline

This thesis is organised in chapters corresponding to the levels of abstraction of our proposed system as shown in Fig. 1.5.

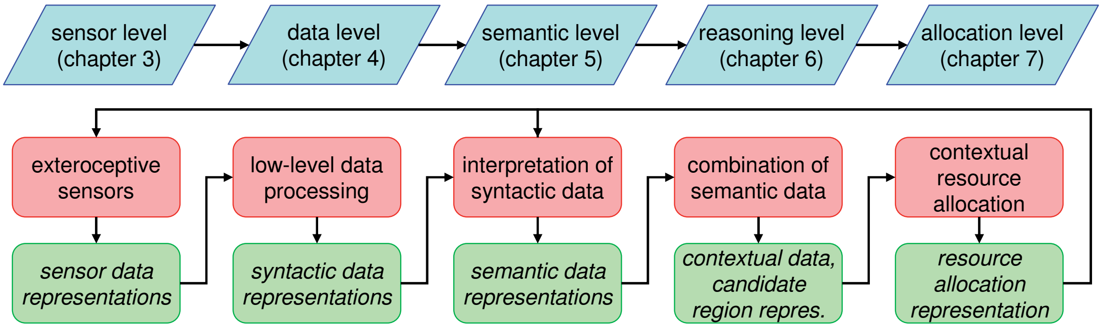


Figure 1.5: Organisation of the thesis corresponding to the levels of abstraction in the system overview given in Fig. 1.4. Red boxes show processing steps in the system, green boxes point out the resulting data representations.

After a review of integral parts and existing concepts for active vision systems in chapter 2, sensors and their data representations are covered in chapter 3. Low-level data processing steps towards a set of syntactical environment descriptions are given in chapter 4 on data level modules. Chapter 5 describes the bridging of the semantic gap, resulting in a set of relevant semantic data representations. In chapter 6 the combination of semantic data into a contextual data representation and our candidate region determination method is described and evaluated. The contextual resource allocation concept is described, evaluated, and discussed in chapter 7. Our conclusions and an outlook on future work are given in chapter 8.