

Visualization of Spatial Sensor Data in the Context of Automotive Environment Perception Systems

Marcus Tönnis*
Technische Universität München
Fakultät für Informatik
Boltzmannstraße 3
85748 Garching b. München, Germany
Tel/Fax: +49 89 289 17083 /-17059

Rudi Lindl†, Leonhard Walchshäusl‡
BMW Forschung & Technik GmbH
Hanauer Straße 46
80992 München
Tel.: +49 (0) 89 382 -13243 / -13242
Fax: +49 (0) 89 382 44988

Gudrun Klinker§
Technische Universität München
Fakultät für Informatik
Boltzmannstraße 3
85748 Garching b. München, Germany
Tel/Fax: +49 89 289 18215 /-17059

ABSTRACT

Spatial sensor systems in cars are gaining more and more importance. Such sensor systems are the foundation of future safety systems, such as automatic emergency brakes, as well as for interactive driver assistance systems. We have developed a system that can visualize such spatial sensor data. Two environments are supported: A laboratory setup for off-line experience and a car setup that enables live experience of spatially aligned laser scanner and video data in real traffic. We have used two visualization devices, a video see-through LCD Flat Panel (TFT) and an optical see-through Head-Mounted Display (HMD) in both setups. For the laboratory setup, a back-projection table has been integrated as well. To present data in correct spatial alignment, we have installed tracking systems in both environments. Visualization schemes for spatial sensor data and for geometric models that outline recognized objects have been developed. We report on our system and discuss experiences from the development and realization phases.

The system is not intended to be used as a component of real driver assistance systems. Rather, it can bridge the gap between Human Machine Interface (HMI) designers and sensing engineers during the development phase. Furthermore, it can be both a debugging tool for the realization of environmental perception systems and an experimental platform for the design of presentation schemes for upcoming driver assistance systems.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Ergonomics; H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces—Evaluation/Methodology; H.1.2 [Models and Principles]: User/Machine Systems—Human Factors

1 INTRODUCTION

Modern cars are equipped with an increasing number of sensors perceiving the environment – especially towards the area in front of a vehicle. Fusing such sensor data and further analysis to detect other traffic participants is expected to help driver assistance systems increase driver safety.

For the development of such multi-sensor systems and driver assistance systems, it is necessary to visualize representations of all levels of such data, starting with raw data from each single sensor up to fused data and interpreted contextual data. Such visualization is necessary for debugging purposes during the development process of perception systems. They will also become invaluable as

cars with increasing sensoric functionality are introduced into market and (re)calibration becomes part of the daily production and maintenance routine since the correct operation of the sensors has to be evaluated or maintained on a regular basis. Visualization of sensor data also can bridge the gap between researchers in sensorics and in HMI presentation concepts, thus leading to new, preferably visual interaction schemes in safety assistance systems.

Such future assistance systems might use large scale Head-Up Display (HUD) technology to place information in the driver's field of view. HUDs and sensory systems combined allow for development of presentation schemes based on the fact that spatially embedded information does not require the driver to look off-road, for instance onto the dashboard. The focus of analysis then relies in finding minimally distractive presentation schemes. Such can be warning symbols superimposed on for example, obstacles, as well as awareness guidance systems for objects that are not in the driver's view [12]. Augmented Reality (AR) could allow for ergonomic testing of those presentation schemes before large scale Head-Up Displays are available for real in-car environments. Thus an accelerated technology transfer is obtained, since parallel research on technology and human factors is enabled. Human-computer interaction researchers can then experience and evaluate any kind of presentation scheme long before it is fully technically mature. Their results can be used as input for research in technological issues by communicating where to focus research.

The paper at hand reports on our presentation setup on our system that successfully displays live sensor data from a equipped car. We have first developed the system in our laboratory to visualize recorded sensor data. Presentation takes place in a Head-Mounted Display (HMD), on a back-projection workbench and on a portable TFT. As a next step, we have ported the system to a test vehicle, such that actually available. In a car, real sensory data can be presented directly where and while it is perceived. The HMD allows for a personal experience, while the TFT enables collaborative experience of the visualization schemes. We report on our experiences with calibration of such setups, system internal data-flows and tracking issues in the test vehicle.

2 RELATED WORK

Bock et al. [4] equipped a sedan with an optical see-through HMD for a vehicle in the loop application. While driving in a real world environment simulated traffic is displayed to the driver. Based on this simulation data advanced driver assistance systems can be evaluated economically and without endangering the test vehicle or other road users. A similar concept is sketched by Regenbrecht et al. [10] who developed a driver safety training system with overlaid virtual content. Both, the trainee and the trainer wear head-mounted displays with video see-through capabilities which are tracked by a combination of fiducial and inertial systems. In a carefully controlled environment, the trainee learns to cope with adverse road and weather conditions and simulated accident situations

*e-mail: toennis@in.tum.de

†e-mail: Rudi.Lindl@bmw.de

‡e-mail: Leonhard.Walchshaeusl@bmw.de

§e-mail: klinker@in.tum.de

rarely found in everyday life.

Another Augmented Reality (AR) application within an automotive environment is navigation. Nartz et al. [8] for example present advanced navigation information to the driver by superimposing the recommended route or a virtual guiding vehicle with the image of a fixed camera.

Nevertheless, none of the approaches developed so far apply augmented reality to visualize live sensor- and tracking data to the co-driver.

3 HARDWARE SETUP

To present the sensor data, a tracking system has been integrated into a test vehicle. This tracker delivers orientation and position data of the visualization devices to the rendering system.

3.1 Lab Setup

The first system has been carried out in the laboratory, enabling us to focus on the tracking issues in a convenient development environment. The perception sensor data provided by pre-recorded scenes is rendered in a down-scaled form to both portable visualization devices (HMD and portable TFT) and to a back projection workbench (cf. figure ??). The table provides an overview of the scenery, whereas the HMD and the portable TFT offer an in-detail visualization of arbitrary sections of the scene (cf. figure 1).

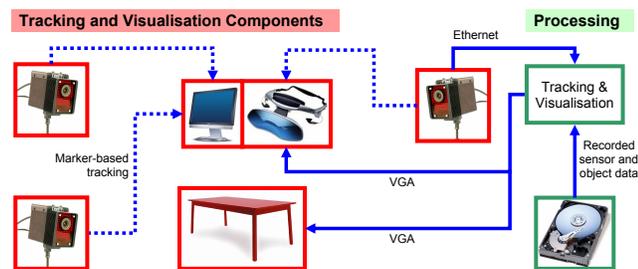


Figure 1: Overview of the hardware components of the lab setup

3.2 In-car Setup

A test vehicle has been equipped with a multi-sensor perception system for road user detection. In the following this system is called *iFuse*. For a more detailed description see [7, 15]. The system consists of a laser scanner utilized for object localization and an automotive video camera used for object classification (cf. figure 2). Furthermore, two infrared tracking cameras (ART smARTrack) mounted on a rigid carrier have been installed at the rear seat behind the driver surveying the area of the co-driver (cf. figure 10 for more details). Any distraction of the driver or the co-driver can be avoided by this mounting position. Most of the communication paths are based on 100 Mbit ethernet links. A schematic overview of the setup is shown in figure 3.

3.3 Visualization Devices

The visualizations are shown in a monocular Head-mounted Display (HMD), on a portable video see-through TFT or on a back-projection workbench. The HMD allows for a individual experience, while the TFT and the table enable collaborative experience.

HMD We use a Sony Glasstron PLM-S700. Although it is equipped with two displays capable of a 800x600 resolution, this HMD is limited to monocular vision. Six optical markers have been attached at the upper front part of the HMD to ensure robust and precise tracking of the position and orientation (cf. figure 14(a)).

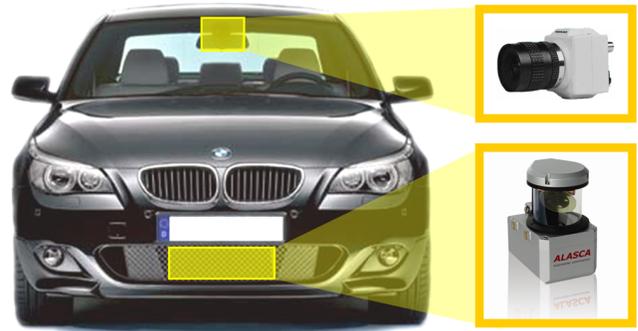


Figure 2: Mounting points of the two perception sensors. The automotive video camera is located behind the windshield and the laser scanner beneath the license plate

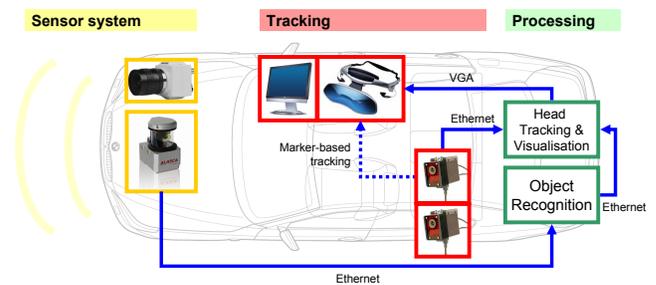


Figure 3: Overview of the hardware components, the communication channels and the rough respective installation locations in the test vehicle

Portable Video See-Through-TFT We obtained a video see-through display by mounting a FireWire camera (resolution 1024x768 at 30Hz) at the middle of the back side of a common 19" TFT display (cf. figure 14(b)). Five optical markers at the top of the display allow for robust tracking.

Back-projection Workbench For collaborative work and discussion with many people we use a workbench [5]. A mirror mounted beneath the tabletop reflects the image of a projector so that it is shown on the workbench. Using back-projection, visualization is not disturbed or occluded by interacting participants.

4 SOFTWARE ASPECTS

The real world vehicle surroundings and the sensor configuration are reflected by a virtual environment, modeled as a hierarchical scene-graph structure [3], ensuring centralized data access and efficient spatial dependency processing. A vector-quaternion-scalar (VQS) [11] representation has been chosen to achieve coordinate system transformations between the entities of the scene-graph.

4.1 Spatial Relationship Graph

The system works on continuous data streams relating 3D positions and orientations (poses) of various sensors, objects and display devices to one another. Figure 5 illustrates these relations in a spatial relationship graph (SRG) [9]. Relationships that are of special relevance to this paper are the edges with labels A through L.

The left part of the SRG describes the internal setup of the *iFuse* system, consisting of the a laser scanner and a camera in the car. Their pose is described relative to the center of the front bumper. Both sensors measure their environment. Figure 6(d) shows the



(a) HMD (Sony Glasstron PLM-S700) used for visualization worn by an artificial head (b) Backside of the 19" TFT video see-through display with an attached FireWire camera at the center

Figure 4: Utilized visualization devices with attached optical markers

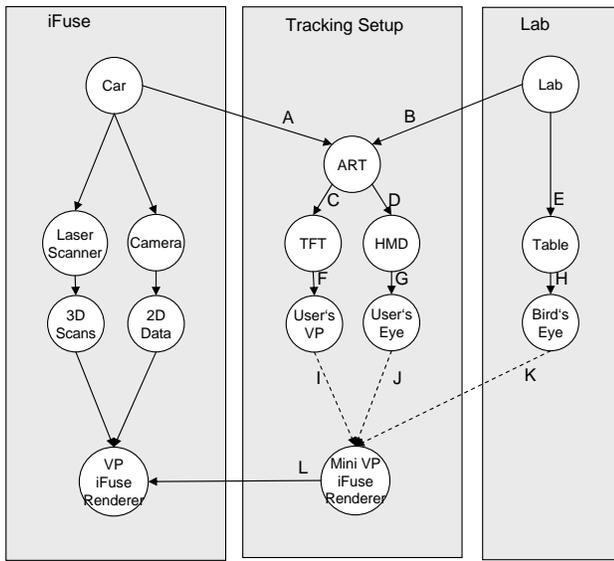


Figure 5: The spatial relationship graph of the presentation system

results rendered as 3D objects relative to the sensors by the *iFuse* renderer [7, 15]. In the SRG, the *iFuse* node represents the virtual viewpoint according to which the scene is rendered. In the bare *iFuse* system, the viewpoint can be controlled via mouse input – as typical for VR scenes.

The center part of the SRG shows the tracking setup that is required to determine the mobile pose of the TFT and the HMD in the car or in the lab. Tracking is provided by an outside-in optical tracker from ART. Depending on whether the system is set up in the car or in the lab, the origin of the tracker is related either to the car or to the lab (edges A, B). It tracks optical markers on both display devices (edges C, D). The user's viewing position is provided relative to the TFT or HMD (edges F, G). The viewing positions control the virtual viewpoint of the *iFuse* renderer (edges I, J). If the system is set up in the lab, the presentation is scaled down by a factor of 10 (edge L).

The right part of the SRG represents the pose of the projection table within the lab (edge E) and the placement of a bird's eye viewpoint relative to the table (edge H). This viewpoint controls how the *iFuse* system renders the sensor data on the table (edge K).

4.2 Calibration of Sensor and Tracking Coordinate Frames

The origin of the ART tracking system is set manually when the tracking system is initialized and calibrated. To this end, among other things, the system coordinator has to place a special calibration object (a large, L-shaped optical marker) at a position in the scene where it is well visible by all cameras of the tracking system. The pose of this marker defines the pose of the tracking coordinate system.

For the lab setup, we place the L-shaped marker on the projection table. That way, the sensor data of the *iFuse* system is automatically aligned with the tracker, the lab and the table. These nodes in the SRG represent the same pose, and the edges B and E are unit matrices.

For the car setup, we cannot place the L-shaped marker on the bumper plate since that area is not within the visible range of the tracker. We thus have to compute a transformation between the origin of the car and the origin of the ART tracker (edge A). The origin of the tracker is located near the co-driver's seat. The origin of the sensor coordinate frame is placed in the middle of the front bumper. We register the tracker relative to the bumper by computing its absolute orientation [6]. To provide the matching points, we touch the required number of points on the car surface with a tracked pointing device. We measure the same points relative to the *iFuse* reference frame by mounting the test vehicle on a measurement base plate. A large scale measuring sensor can measure the transformation between the origin of the sensor system and the target points in the test vehicle. This setup is part of a well-established routine for registering various sensors for *iFuse*. We have also implemented a GUI to manually readjust parameters of the absolute orientation calibration that exhibited minor errors. This way, minor errors that occurred during the manual procedure of selecting matching calibration points can be fixed without requiring to mount the test vehicle on the base plate again.

4.3 Calibration and Tracking of Visualization Devices

Device Poses The poses of the TFT and the HMD (edges C, D) are provided in real-time by the ART tracker. The projection table is placed at a fixed position in the lab (edge E). It could easily be equipped with an optical marker like the TFT and the HMD, such that it can be wheeled to different positions. However, this has not been necessary for the current setup.

User Viewpoints To show *iFuse* visualizations on the TFT, HMD or the projection table, the view point of the user needs to be described relative to these display devices (edges F, G, H). Since the tracking system places the origin of tracked markers at a position within the target that cannot be easily determined in a real setup, calibration routines are needed to relate the marker poses to a specified point of each display, such as the center of the display.

For the TFT (edge F), we describe the static transformation from the pose of the marker to a fixed viewpoint at the center of the display (edge F). To this end, we use a pointing device (similar to the one used for the absolute orientation algorithm) to mark all four corners of the presentation area and compute the center. For the lab setup, we then move the viewpoint along the optical axis 40 cm in front of the TFT since, if a person holds the display right in front of his eyes, the average distance of the display is about 40 cm to the user's head. Thus if he looks perpendicularly at the display surface, a field of view that is correctly aligned to the real world is covered. If the display is seen from an oblique angle, e.g., when viewed by a second person, the augmented video picture does not perfectly match up with reality. The resulting transformation is written to a configuration file. This file is parsed by the *iFuse* system on startup and the transformation matrix is placed in the rendering system's data flow.

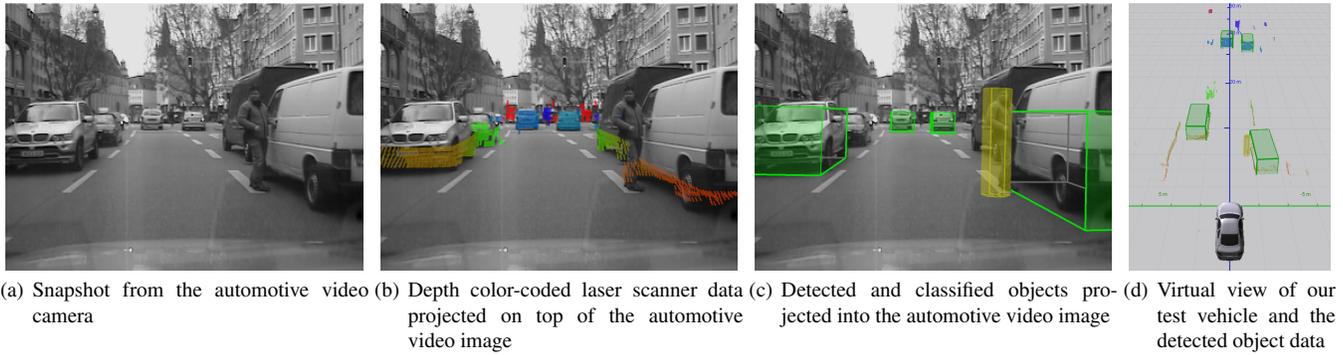


Figure 6: A snapshot from a city-scene recorded with our test vehicle. The depicted presentation schemes are at the moment used with our offline perception system development toolkit

For the HMD (edge G), we use the Single Point Active Alignment Method (SPAAM from [14]). This method requires a user to align a cross-hair on the 2D screen of the HMD with several points in the 3D space. The algorithm computes a projection matrix for a perfectly aligned view and also can incorporate the transformation of the marker-target to the viewpoint. After the HMD calibration, the resulting projection matrix is written to the *iFuse* configuration file. In contrast to the startup of the system in HMD mode, not only the transformation is set, but the projection matrix including the transformation replaces the standard projection of the rendering system.

For the projection table (edge H), we set the view point to a fixed position above the surface.

5 PRESENTATION SCHEMES

We visualize sensor data as well as perception data in real-time. In the see-through HMD case only the laser scanner data and the perception data is rendered (cf. figure 7). In case of the portable TFT, the laser scanner data is rendered on top of the video images of the automotive camera (cf. figure 6(b)).

Our laser scanner has a horizontal aperture angle of 160 degrees with four vertically spanned layers with an inter-layer distance of 0.8 degrees and a horizontal single-scan sampling of 0.5 degrees. The lowest layer points forward in exactly horizontal direction. The topmost (fourth) layer aims 2.4 degrees upward. Layer one and two are taken at angles of 0.8 degrees and 1.6 degrees, respectively. The total sampling rate is 10Hz . Thus we visualize $(160 \div 0.5 \cdot 4 \cdot 10) = 12800$ distance measurements per second. The video signal is in NTSC format with 29.97 interlaced frames of video per second. To avoid flickering of the laser scanner data due to the reduced sampling rate of the scanner compared to the TFT-mounted camera, data from old scans is repainted in a fading-out manner until new scan data arrives. Each individual measurement from the laser scanner is drawn as solid rectangle, depicting the intersecting plane of the laser beam at the given distance. The rectangles are color-coded with respect to their distance (cf. figure 6(b) and 6(d)) to preserve the depth perception with both the monocular HMD and the portable TFT.

The perception system of *iFuse* detects and tracks vehicles and pedestrians in the data from the laser scanner and the automotive video camera. The recognition results are visualized as bounding boxes which enclose the respective object. The colors of the bounding boxes encode the object type (green for vehicles and yellow for pedestrians, cf. figure 6(c)).

6 ISSUES DURING SYSTEM DEPLOYMENT

The system was primarily developed in a laboratory and tested with an offline version of *iFuse*. This off-line version is capable of re-

playing sensor data recorded earlier. When the tracked HMD and TFT were integrated, the system was transferred into the automotive environment. For both environments, we had to cast a number of trade-offs, to be described and discussed next.

6.1 Lab Setup

To visualize prerecorded sensor data in a virtual laboratory environment, some extensions to the original *iFuse* system had to be made.

Scaled Presentation To have a suitable test environment, the original sensor data had to be scaled such that all presentations fitted into the laboratory's space. For the see-through HMD, users would have been irritated had they perceived the real walls of the laboratory to be closer than parts of the virtual traffic scene.

Furthermore, it is easier to examine a scaled environmental scene from different vantage points around or inside the test vehicle. Users can walk around the scaled-down test vehicle and even dive into the car to get a view from the driver's perspective. The entire scenery was scaled down by a factor of 10.

Location-Fixed Car Sensor data is delivered relative to the car coordinate frame.

In principle, the car is moving relative to the real world. However, it does not make sense to account for such motion in the laboratory-based visualization since the car would then quickly leave the lab setup on its test rides. Instead, the coordinate system of the (moving) car was fixed to the origin of the tracking coordinate system in the lab (edges B and E in figure 5), and the visualizations are shown relative to the car coordinate system. As a consequence, the road, as well as world-fixed objects can be seen as mobile objects floating by the car in reverse direction. The car is standing still in the laboratory, while the sensor data appears to move toward it, passing by and leaving the sensed area.

Absolute Placement The laboratory constitutes a fully virtual environment. No real object such as the real test vehicle or obstacles are present. For a better understanding of the virtual setup, a user has to know where to perceive the virtual scenery in the lab. To ease scene discovery and joint discussions among researchers, we showed a stationary visualization of the sensor data from a bird's eye view on a back-projection workbench. The origin of the presentation on the table is aligned with the origin of the tracking system, enabling all users to get an impression of what is going on in the virtual world (cf. figure 8).

Users holding the TFT or wearing the HMD can get another perspective view of the same scenery.

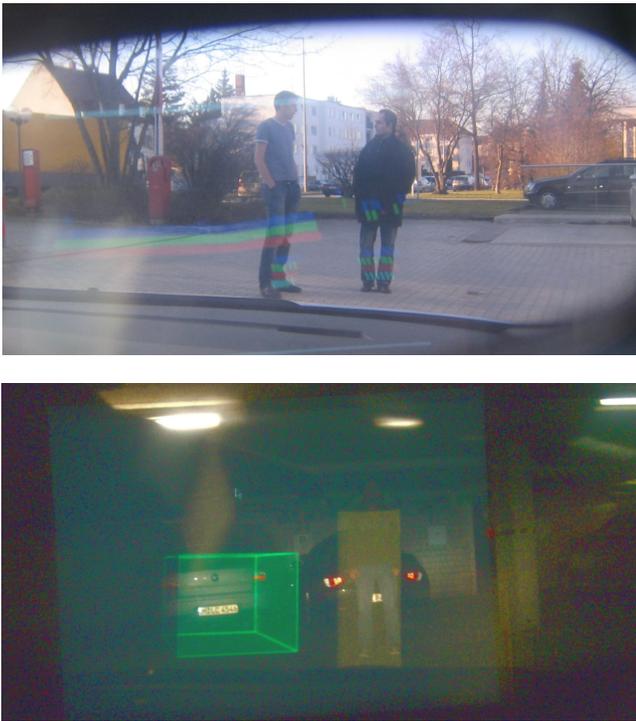


Figure 7: Two handmade photos taken through the HMD. Both are not perfectly aligned (=calibrated), because no mount for the camera and the HMD was available. The image on top shows sensor data from the automotive laser scanner. The persons' legs and the curbstone can be seen superimposed with sensor data. The image on the bottom has been taken in a basement garage and shows the detection results of the perception system (pedestrian: yellow box, vehicle: green box)

6.2 In-Car Environment

To deploy the system in the test vehicle, the major issue was to find a suitable mounting point for the tracking system. Certain concerns were related to tracking, others to presentation issues and a sufficient throughput of the system:

- Non-occluded tracking volume to track the visualization devices.
- Stable vibration-free tracking of the visualization devices.
- Minimal perturbation of in-car systems to infrared tracking cameras.
- Brightness and contrast of visualization devices
- Superimposing reality.
- Minimal time difference between environment perception, in-car tracking and rendering.

In-Car Tracking The cameras in the car environment must be able to track the following region: The two presentation variants for the *iFuse* system consume an area from up to 40 cm in vertical, 80 cm in horizontal and 60 cm in depth from a seated person. A wide angled tracking camera must be used to supervise this volume. To receive valuable position and orientation data for all degrees of freedom, each camera's orientation to one another should have a certain angle, not facing along the same axis and should have a certain distance in between [1].

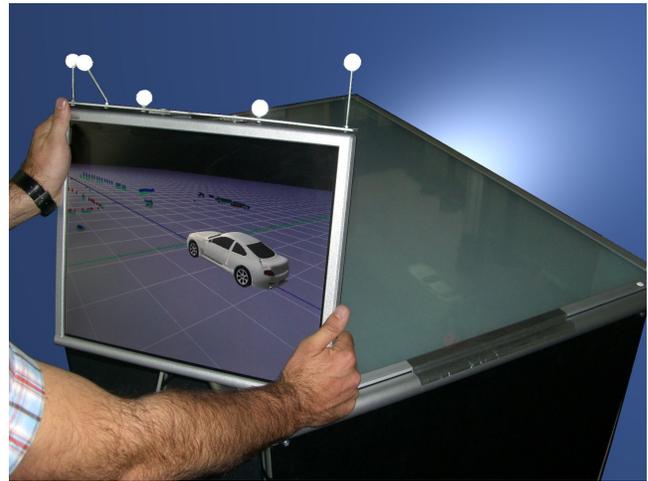


Figure 8: The portable TFT held into the tracking volume at the laboratory visualizes the virtual test vehicle and the perceived laser scanner data. The same scene is rendered from a birds-eye view on the back-projection workbench below

The ART smARTrack cameras provide such tracking quality – aside from the fact that their size is not really suitable for in-car use. Both cameras are mounted on a carrier in a suitable distance (about 50 cm). Their focal axes cross in a distance of 160 cm to the middle of the carrier and the tracking volume starts in a distance of about 30 cm, where the borders of the field of view cross almost perpendicularly. With this fixed setup, the resolution of the cameras for a standard target ranges from 0.25 mm to 0.5 mm depending on the distance to the cameras (about 50 – 100 cm) and axis relative to the lines of sight of the cameras [2].

The carrier bar had to be mounted such that both cameras can track movements of the HMD and of the TFT. First thoughts of placing the cameras in front of the co-driver, above the right part of the dashboard failed due to the required minimum distance of the tracking volume. Moving the cameras further away, to the drivers side also failed because the car still has to be maneuverable without any handicap.

Thus the cameras had to be moved behind the front seats. The attempt to mount the cameras directly behind the co-driver's head-rest prohibited tracking the visualization devices, because the user's head often occluded the line of sight of one or both cameras. From the three remaining options to mount the cameras near the seat behind the driver or on the back-shelf, the two horizontal approaches failed. Placing the camera bar horizontal on the back-shelf failed, because the co-driver's head-rest occluded too much of the tracking volume. Also the placement directly behind the driver, the bar turned about 45 degree to the right, facing toward the tracking volume did not prove suitable, because here, both head-rests occluded some parts of the tracking volume.

Finally, the approach to place the bar vertically above the right knee of a left backside passenger enabled a large enough view of the required tracking volume. The photo in figure 9(a) illustrates the extent of the freely visible tracking volume.

When a visualization device is brought into this volume, a passenger has taken seat. To illustrate what the tracking cameras record: Figure 11 shows the TFT as it can be held in the tracking volume. Both cameras can recognize the markers. Figure 9(b) shows the HMD as it can be worn. Especially the upper tracking camera sees all marker balls. Yet, the lower one also has a good vantage point.



(a) Snapshot of the two tracking cameras from behind and the covered tracking volume (b) The HMD worn in the tracking volume. The tracking cameras (back right) can see the markers

Figure 9: The tracking cameras and the tracking volume inside the test vehicle

The volume that can get occluded is characterized by the following coincides: First, the driver's right elbow, when moved backwards or the co-driver's left arm can occlude parts of the lower volume. Second, when a tracked object (TFT or HMD) is moved or turned far to the right side, the co-driver's left arm or left shoulder can occlude the TFT's markers, while the head still can occlude the marker-balls of the TFT.

After a suitable location was found for the camera carrier, it had to be ensured, that no (infrared) light source from the interior of the car impaired the spot detection algorithm of the tracking system. Detection checks of the tracking system showed that a transponder of the car mounted near the driver's back-mirror placed a unintended spot in the video image. Right next to this spot, depending on its adjustment, the back-mirror reflected the infrared pulses of the tracking system, generating a second dead spot.

A final issue in incorporating the tracking system into the car concerned finding a stable, vibration-free mounting of the tracking system. The two-camera system is rigid in itself. It correctly maintains the calibration between the two cameras themselves. To ensure that the origin of the full tracking system does not alter due to vibrations of the driving car, a stable connection between the car and the tracking system had to be built. A construction, applying the two handrests near the roof of the car above the rear car doors (cf. figure 10(b)) and the mount point of a child seat at the lower end of the backrest gave enough stability against longitudinal (front-to-back) and vertical (top-down) vibrations. Lateral movements (left-to-right) were reduced by an additional aluminum profile in the passenger's footwell. The complete construction is depicted in figure 10.

With all these issues solved, a working tracking system was mounted successfully in the car.

Presentation Issues While driving on the road, various lighting conditions can occur. From almost full darkness on rural roads during night, brightness can range to full sunlight and reflection spots from other objects, especially from the windshields of other cars in certain daytime conditions. The human eye is capable of adjusting its iris to such conditions. If a visualization device does not produce an image bright enough to be comparable with outside lighting, the perception of the AR presentation schemes is reduced. Under large changes in brightness, focal adaption between different brightness levels takes longer. Perception of a standard TFT display can thus be more exhausting for the user. Furthermore, a direct overlay on an optical display can have such a low brightness,



(a) Overview of the mount for the tracking system (b) Left side handhold mount for the tracking system

Figure 10: Construction of the mount for the tracking system

that it is no longer visible in bright daylight conditions. Fortunately our car had slightly toned windows, so that incoming light is reduced in the in-car environment.

For the TFT, the camera mounted on its backside had a lens with adjustable iris. In addition, the software driver automatically adjusted the image's brightness. Hence, overlaid sensor data always appeared in a good perceivable relationship. The TFT had a screen diagonal of 48 cm and a brightness of $300 \frac{cd}{m^2}$. Presenting the generated image to several people was judged satisfactory on a sunny summer day with hardly any clouds. Figure 13 and figure 11 show photos of the TFT while presenting a scenery on that day. In contrast to the digital camera used to take these pictures, the differences in brightness are handled easily by the eyes of a human being.



Figure 11: The TFT held into the tracking volume while driving. One can see the laser-scanner augmentation on the vehicles and the guardrail

For the HMD, the real environment is directly augmented with virtual objects. Here brightness of the HMD (Sony Glasstron) is a more critical issue. A scenery viewed in full daylight is too bright to still perceive any augmentation. For normal sunny daylight with no direct solar radiation (cf. figure 7), a respectable amount of the superimposed scenery remains perceivable for the user.

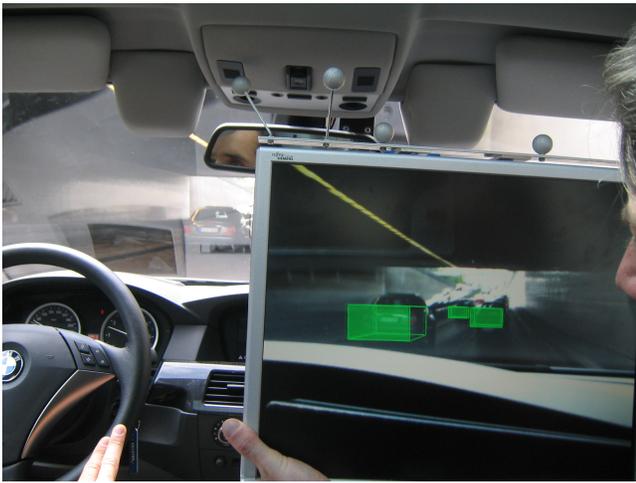


Figure 12: The TFT held into the tracking volume while driving. One can see the green augmentation of the detected vehicles

System Performance An immediate system response is crucial for augmented reality systems. Therefore we put some effort into achieving real-time performance. The system response times are composed of the timings from the tracking system, the TFT-mounted camera, the perception system, the rendering and the communication paths. The communication times are negligible. The scene rendering has been implemented in OpenGL and is accelerated by a 3D graphic adapter, resulting in a processing time of about $20ms$. The laser scanner frequency is $10Hz$ which results in a worst case processing time of $100ms$. Altogether the processing time sums to about $100ms$ in average.

Visualization on the Road Once all issues of installation and presentation were addressed, the system was ready to be tested.

First tests on a parking lot quickly revealed, that the construction of the TFT introduced several problems: First, the rather large TFT did not fit easily into a sedan. Moving and turning were possible, but the user had to be careful, not to disturb the driver's area around the gearstick. Furthermore, the TFT was heavy and thus uncomfortable to hold for a longer time. Third, due to its size, the TFT display had to be held very close to the ceiling of the car. In such position, the marker-balls were tracked well. Yet, the user had to watch out, that the markers on top of the display were not deformed. Handling the TFT was possible, but more complicated than necessary. The fourth observation concerns the mounting point of the digital camera on the backside of the display. It was mounted on the opposite of the center of the screen, to give an intuitive understanding to the user, that he is observing the scene from inside TFT. Due to the layout of car interior, the height between the dashboard and the roof allowed for the camera only to look over the dashboard, when the TFT was held so high, that it almost touched the ceiling. In that case, about two thirds of the environment shown on the TFT was outside area, while the remaining part showed the dashboard.

These first tests also gave feedback about the quality of the marker-tree construction. The setup of the TFT display required that all markers be attached along the upper edge of the display. Otherwise, single marker-balls would have had too large a distance to the TFT, complicating handling even more. The same applies to the HMD. All marker-balls are attached on top of HMD to improve its trackability. The marker-balls on the HMD define a wide, but not that deep and high volume. So, for both displays, rotations like nodding (i.e.: pitch) produce the high tracking imprecision. Thus the largest misalignments in rendering attribute to verti-



Figure 13: Snapshot of the See-Through TFT. One can see the superimposed sensor data on the garbage container and the Vespa scooter

cal (non)conformity.

The depth based alignment between the real environment and the superimposed sensor data in the TFT was good when we used a camera model with a fixed focal length to initialize the corresponding projection matrix in the rendering system. For the HMD, the projection matrix was generated by the SPAAM algorithm - a tricky task to achieve in a car since the algorithm requires the alignment of a calibration device with a crosshair, and both the calibration device and the HMD must be tracked accurately. The extremely limited space to calibrate the HMD in the car are the reason, why the HMD's frustum tends to be slightly too wide - as can be observed for objects at large distances.

After the first tests in the parking lot, the system was tested under driving conditions on the road. Here the stability of the mount of the tracking system was tested. The vibration of the car, when driving on normal concrete streets generated perceivable jitter. Augmented objects, sensor data and boxed objects reached misalignments of 0.5 cm to 1 cm in average when driving over a bump. To determine the effect of the not quite perfectly rigid mounting, we mounted a marker-target in a fixed position at the maximum distance (1.02 m) to the tracking cameras. This target, being light in comparison to the cameras should not swing very much, even if the car drives over a bump. Then, tracking data was recorded while standing and while driving. Recorded tracking data was analyzed and results showed that, for a standing car position, tracking precision varied below a value of 0.1 mm and orientation below a value of 0.2 degrees. When driving with the car, positional changes reached about 4 mm in all three axes. Rotational changes remained below 1 degree. The increased positional imprecision contributes to up to 0.6 degree in altered orientation. Thus the cumulated error of the vibrating tracking system can sum up to 3 cm for a maximum amplitude in both position and orientation. In general, on even concrete, calculated with 68% of the maximum amplitude, the error is about 1 cm .

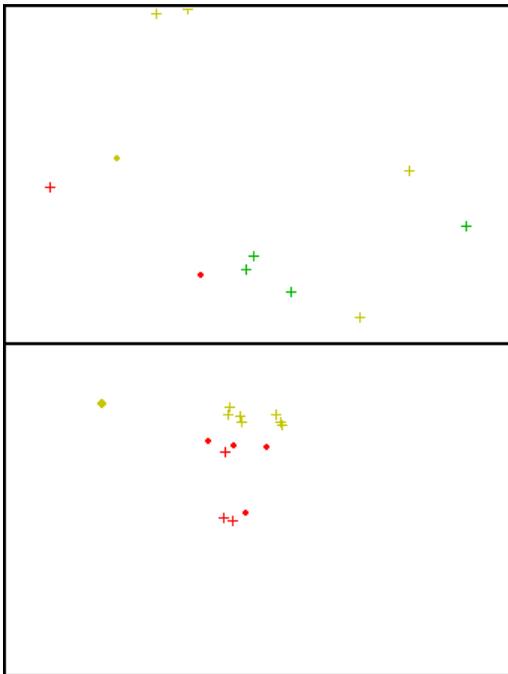
In summary, misalignments can be attributed to four major factors. First, they are caused by tracking errors due to suboptimal marker-target design, as described above. Second, system lag can sum to 100 ms , causing significant misalignments for a moving car. A third cause of misalignments are blinding spots from bright outside objects and in-car light sources. Fourth, a non-rigid mounting of the smARTrack camera carrier disturbs the tracking data.

We noted another issue during our test drives, pertaining to the

rendered data from laser scanner during. When driving through, e.g., a tunnel, sensor data was perceived from the walls, too. Looking at the graphical representation gave the impression of a miscalibrated frustum. This effect vanished when the setup of the laser scanner was recaptured. The lowest layer of the scanner aims horizontally forward, while every layer above aims another 0.8 degrees more upwards. Therefore the graphical representation can appear as not scaling down fast enough with increasing distance.



(a) A photography taken from the center of the smARTrack bar: Sunny day and a reflecting windshield



(b) The two smARTrack images. The upper part shows the upper camera and the lower part shows the lower camera

Figure 14: Two images showing the centered view out of the windshield and the tracked points, the tracking software receives from the cameras – both pictures were taken at the same point in time

Signal-Noise Ratio of Infrared Tracking Automotive environments are no usual environment for infrared trackers. Natural sunlight, reflection spots and lighting rays can flood in through the windows. In-car bulbs and infrared interfaces like central locking systems senders generate noise that reduces tracking quality. Static spots like most of in-car light sources can easily be neglected by disabling tracking of these spots in the corresponding camera im-

ages. Thus tracking the brightest spots, which then are the marker-spheres again, is enabled again. For car-external light sources, the situation is different. As the car can move through the environment, light-source locations can change and no blind spots can be placed in the camera images. Adjusting the infrared flash intensity can increase the signal to noise ratio. Users have to stand occurring inconsistencies brought up by occurring wrong spots.

We tested tracking in different environments under various lighting conditions, from bright sunlight to just i-car lighting during night. In the first tests we realized, that we have to adjust the camera flash intensity to the upper end so that the marker-spheres' reflection is stronger than unintended lighting. Then, surprisingly, tracking worked well under most conditions. Fig. 14 shows a scene taken on a bright sunny day and the tracking software's screenshot showing the tracked points. One can estimate the location of the marker target's three spheres (green crosses) in the lower part of the upper tracking software screenshot. The other spots are false positives, but are colored in red and yellow, indicating not-as-good hits. This figure is a good example to illustrate the good overall tracking.

Extreme light conditions still bring in high noise and thus tracking fails, when, e.g., the sun is directly in front of the cameras. The slightly darkened glasses surely facilitate tracking, but can not cope with the strength of natural light.

7 CONCLUSION AND FUTURE WORK

We have built a presentation system for spatial sensor data. The idea for such presentation of sensor data is motivated by the need to debug sensor and perception data in real setting and by the need of user interface designers to get a better understanding of what sensor measurements can principally be provided to novel driver assistance applications. The system is not intended to as the final presentation scheme for in-vehicle driver assistance systems.

7.1 Checking Sensor Functionality

Both setups, the laboratory setup and the car setup can be used for a visual check, whether a sensor works properly.

The laboratory setup can be used for visualization of sensor data, which for instance, had been recorded directly after assembly of each sensor. Experts can compare object recognition schemes by looking at recognized objects from different perspectives in a down-scaled setup.

The car setup allows for direct inspection of the perception system. While our system has been developed for in-car use, the tracking system surely could also be mounted around a (stationary) car such that an inspector can walk around it and investigate the quality of sensor calibrations and of sensor operation in a much larger space.

New prototypes of sensors, analysis strategies and object detection methods could now be more easily experienced, evaluated and discussed by the sensoric community.

7.2 Designing Visual Driver Assistance Systems

In addition to the inspection of the sensor's functionality, car embedded visualization of presentation schemes can support development of upcoming visual driver assistance systems. Although large scale HUDs are not available yet, research in usability and applicability of such embedded presentation schemes is enabled by our presentation system (e.g., as used by [13]). Furthermore, new warning schemes for road hazard warnings or traffic dependent navigation can easily be developed and experienced.

The laboratory setup enables fully joint examination of such presentation schemes. The visualization on the table enables all participants of a research group to get an overview of a proposed scheme, while the direct neighbors of the person holding the TFT can examine the presentation scheme from a certain point of view. A fully

personalized view can be obtained in the HMD, where the user can choose an arbitrary perspective.

The car setup enables in-place examination of such driver assistance schemes. Perception based driver assistance systems that have been declared valuable for evaluation can be integrated directly into the car setup and thus become AR presentation schemes. Test subjects can experience these augmentations in real traffic.

7.3 Future Work

There are a number of open issues that need to be solved in the future. Most of them have been identified after the system was set up and tested.

Extensions to the System The usability of the portable TFT and the HMD for in vehicle applications could be further increased: Additional markers would enhance tracking performance. For the TFT, the markers are currently placed on top of the display, to be visible all the time. Those markers are nearly linearly aligned along the horizontal axis. Rotations around the horizontal axis thus cannot be tracked with high precision. At least one additional marker, placed more toward the lower end of the TFT, can fix this issue. For the HMD, the long marker-sticks on top of the HMD make it difficult for large persons to move their head freely, because those sticks can hit the roof. Using OLED technology will enable more flexible, lightweight and thus more ergonomic displays.

Moreover, further testing should evaluate the potential of video see-through, high contrast or stereo vision HMDs regarding user acceptance and usability.

Moving the video see-through camera to the top of the TFT would increase the variety of viewpoints that show image data that is not partially occluded by the vehicle's dashboard. This in turn requires changes of the TFT calibration algorithm.

Bridging the Gap to HMI Further research is needed to evaluate the suitability of the presented setup and visualization schemes regarding the design and assessment of advanced driver assistance systems for both HMI and application developers.

ACKNOWLEDGEMENTS

The multi-sensor perception system *iFuse* for road user detection utilized in this paper is part of the results achieved by BMW in the *COMPOSE*-project, an application-driven subproject of the *PREVENT* Integrated Project – an automotive initiative co-funded by the European Commission's Sixth Framework Programme.

We would like to thank the students Troels Frimor and Philipp Kemmeter for their programming support.

REFERENCES

- [1] M. Bauer, M. Schlegel, D. Pustka, N. Navab, and G. Klinker. Predicting and estimating the accuracy of vision-based optical tracking systems. In *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR'06)*, Santa Barbara (CA), USA, October 2006.
- [2] M. A. Bauer. *Tracking Errors in Augmented Reality (to appear)*. PhD thesis, Technische Universität München, 2007.
- [3] B. A. G. Bishop and G. Welch. Tracking: Beyond 15 minutes of thought: Siggraph 2001 course 11. Technical report, University of North Carolina at Chapel Hill, 1995.
- [4] T. Bock and M. Maurer. Vehicle in the loop simulation. In *4. Workshop Fahrerassistenzsysteme*, pages 125–134, Löwenstein / Höllinsülz, October 2006.
- [5] F. Ehtler, M. Huber, and G. Klinker. Hand tracking for enhanced gesture recognition on interactive multi-touch surfaces. *submitted to Horizontal Interactive Human-Computer Systems*, 2007.
- [6] B. Horn et al. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.
- [7] R. Lindl and L. Walchshäusl. Three-Level Early Fusion for Road User Detection. *PREVENT Fusion Forum e-Journal*, (1):19–24, 2006.

- [8] W. Narzt, G. Pomberger, A. Ferscha, D. Kolb, R. Müller, J. Wiegardt, H. Hörtnner, and C. Lindinger. Pervasive information acquisition for mobile ar-navigation systems. In *Fifth IEEE Workshop on Mobile Computing Systems and Applications*, 2003.
- [9] D. Pustka, M. Huber, M. Bauer, and G. Klinker. Spatial relationship patterns: Elements of reusable tracking and calibration systems. In *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR'06)*, October 2006.
- [10] H. Regenbrecht, G. Baratoff, and W. Wilke. Augmented reality projects in the automotive and aerospace industries. *IEEE Computer Graphics and Applications*, 25(6):48–56, 2005.
- [11] W. Robinett and R. Holloway. The visual display transformation for virtual reality. Technical Report TR94-031, University of North Carolina at Chapel Hill, Department of Computer Science, Chapel Hill, NC, USA., 10 1994.
- [12] M. Tönnis and G. Klinker. Effective Control of a Car Drivers Attention for Visual and Acoustic Guidance towards the Direction of Imminent Dangers. In *Proceedings of the 5th International Symposium on Mixed and Augmented Reality (ISMAR)*, Oct. 2006.
- [13] M. M. Trivedi and S. Y. Cheng. Holistic Sensing and Adaptive Displays for Intelligent Driver Support Systems. *IEEE Computer*, 40(5):60 – 68, 2007.
- [14] M. Tuceryan and N. Navab. Single point active alignment method (SPAAM) for optical see-through HMD calibration for AR. *Augmented Reality, 2000.(ISAR 2000). Proceedings. IEEE and ACM International Symposium on Augmented Reality*, pages 149–158, 2000.
- [15] L. Walchshäusl and R. Lindl. Multi-sensor classification using a boosted cascade detector. In *IEEE Intelligent Vehicles Symposium 2007*, June 2007.