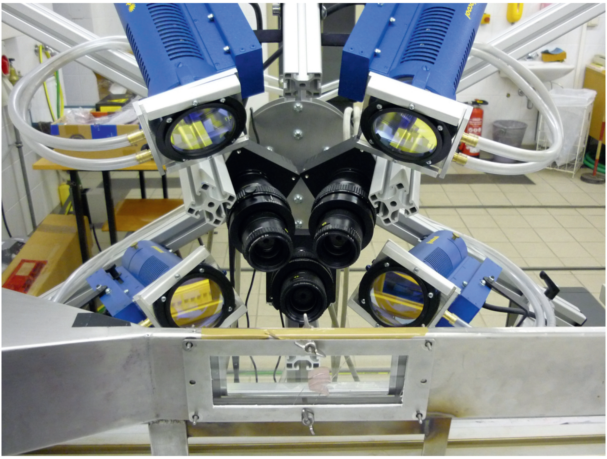
OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

EIT FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY

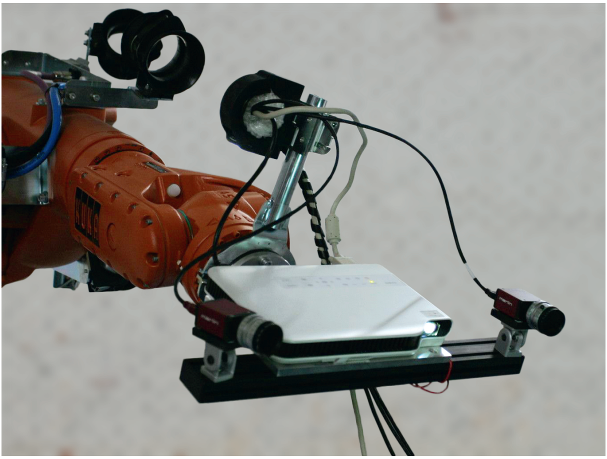
Institute for Information Technology and Communications (IIKT)

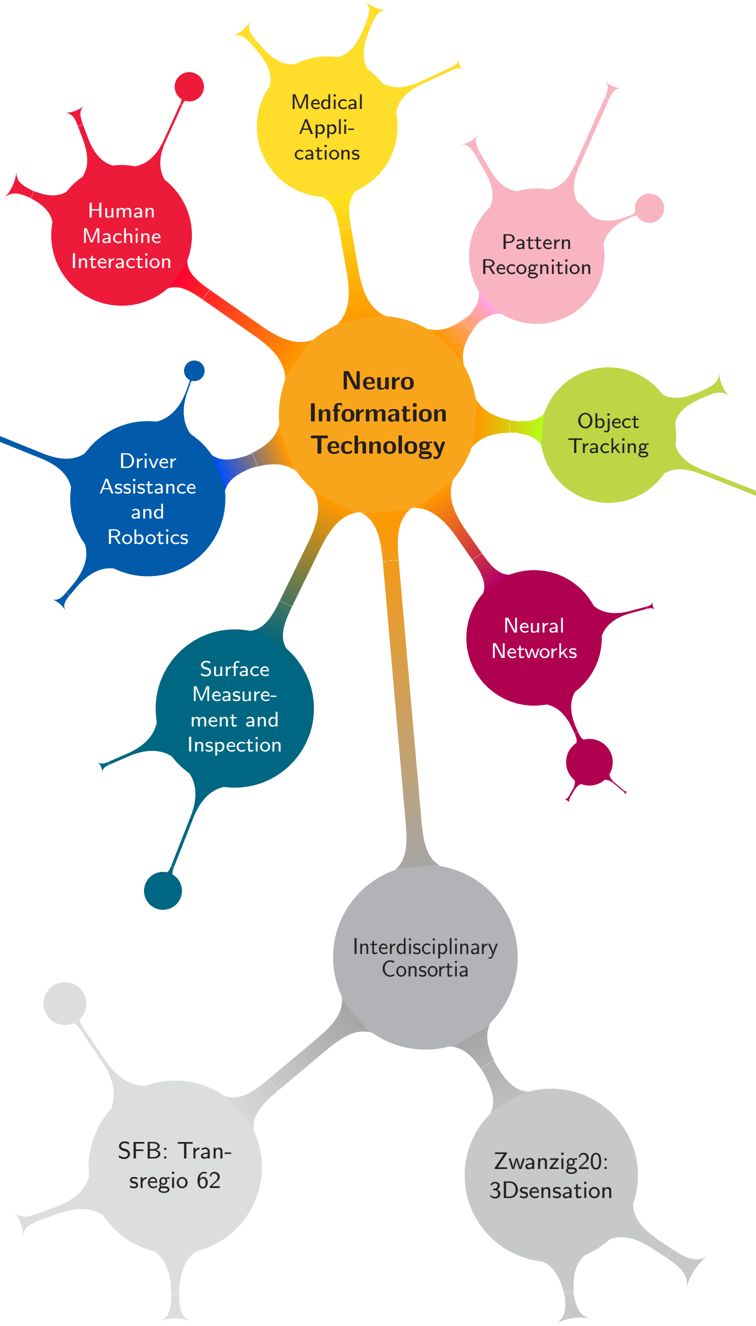
Neuro-Information Technology

apl. Prof. Dr.-Ing. habil. Ayoub Al-Hamadi



Research Profile




Neuro Information Technology
Medical Appli-cations
Human Machine Interaction
Pattern Recognition
Object Tracking
Driver Assistance and Robotics
Neural Networks
Surface Measure-ment and Inspection
Interdisciplinary Consortia
SFB: Tran-sregio 62
Zwanzig20: 3Dsensation

Neuro-Information Technology

The professorship for Neuro-Information Technology (NIT) is located at the intersection of the research fields multisensory information gathering and processing, human-machine interaction (HMI), and information perception. This involves at first the use of modern methods of information technology and pattern recognition for signal, image and video-based industrial and medical applications.

Our research field is the analysis of signal, image and video-based data in a variety of domains, mainly:

- Surface Measurement and Inspection, e.g. for automated quality control in industrial production.
- Driver Assistance and Robotics, including driver monitoring, autonomous driving, and environmental perception.
- Human Machine Interaction, to communicate contact-free using gestures and/or facial expression.
- Medical Applications, such as pain recognition, video based heart rate estimation.
- Pattern Recognition, e.g. arabic handwriting recognition.
- Object Tracking in microscopic and macroscopic video data.
- Neural Networks for image correction and self-organizing systems.

The following pages provide a brief overview of our work in the above mentioned domains.

Ayoub Al-Hamadi

SFB: Transregio 62

SFB/TRR 62 is an interdisciplinary research activity to investigate the communication between technical systems and human users. It is particularly focused on the consideration of so-called Companion-features - properties like individuality, adaptivity, accessibility, cooperativity, trustworthyness, and the ability to react to the user's emotions appropriately and individually. The research program comprises the theoretical and experimental investigation as well as the practical implementation of advanced cognitive processes in order to achieve Companion-like behavior of technical systems. With that, it will lay the foundations for a technology which opens a completely new dimension of interaction between man and technical systems. The Neuro-Information Technology (NIT) group contributes to subprojects Environment Perception (C1), Mechanisms of Nonverbal Communication (C3) and Companion-Technology for Augmented Reality Based Worker Assistance in Automotive Applications (T1).

3Dsensation

3Dsensation is a consortium founded by the BMBF program "Twenty20-Partnership for Innovation". The consortium consists of currently 20 research institutes and more then 40 industrial enterprises. It aims at the potentials of 3D image processing for human machine interaction in the pioneering fields of production, health, mobility and security. The technological challenges that arises should be solved by various research projects through the interdisciplinary combination of competencies of at least four partners. Our group contributes various projects in the fields of human machine interaction and 3D surface measurement.

High-Resolution Surface Quality Inspection for Large Industrial Surfaces

Manufacturing of premium quality surfaces is a major challenge, especially for high-priced goods as car body panels for the automobile industry. As of today, 3D surface measurement systems are either limited by field of view or resolution. The acquisition of data for large scale objects at consistent precision is not possible within a single measurement.

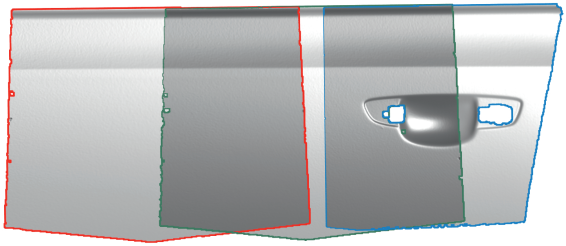


Figure 1: Partly overlapping measurements of an unpainted sheet metal car door panel.

A reasonable solution is the registration of individual measurements into one combined measurement. However, most industrial surfaces are feature-poor, and thus not well suited for common registration methods. Also a naive combination of measurements does not meet the requirements for precision and resolution necessary for surface quality inspection. Here, surface defects as dents and bumps of very small height (starting at 20 micrometers) have to be found.

Our research is focused on new methods for large-area, high-resolution 3D surface measurement suitable for industrial surface quality applications. This is achieved by combining our know-how of 2D and 3D data and image processing algorithms and high-precision 3D measurement with pattern projection. Our current approach combines 2D texture-based image registration with a 3D algorithm derived from Iterative Closest Point (ICP).

(S. v. Enzberg ☎-20113, L. Dinges ☎-11487)

3D Surface Inspection of Deformable Workpieces

Research effort is not only put in increasing optical sensor accuracy, but also in new methods for 3D data processing. In surface quality inspection, the objective is to detect defects as dents or bumps with heights in the micrometer-range. Surface as sheet metal or plastic panels are easily deformable before assembly and thus vary in shape. For curved surfaces with complex details, which are commonly used in the automobile industry, surface modelling thus is a demanding task.

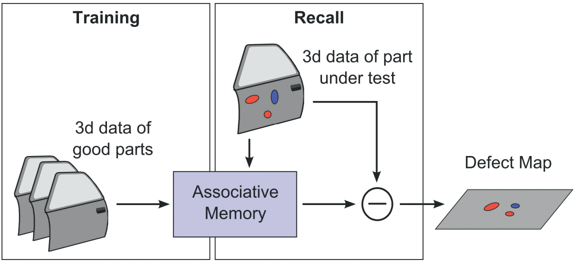


Figure 2: Functional principle of the Associative Memory

We developed an Associative Memory based on B-Spline surface descriptions, which allows for detection of surface defects close to optical resolution and noise limits, even for deformable workpieces. The system combines reverse engineering methods used in Computer Aided Design with Machine Learning methods to form an implicit surface model based on training data. This allows detection of defects which are hard to quantify and avoids the need to explicitly design a surface model that includes possible tolerance ranges and deformations.

Current research focuses on generation of virtual training parts, automatic defect classification and easier setup and training methods for evaluation of complex parts. The joint project with INB Vision AG and Fraunhofer Institute Magdeburg project was awarded third place of the 2014 innovation price of Saxony-Anhalt (Hugo-Junkers-Preis).

(S. v. Enzberg ☎-20113, E. Lilienblum ☎-11126)

Robot Based Optical Test Engineering

Test engineering systems based on the 3D surface reconstruction of large and complex workpieces often work with industrial robots. Generally, such robot based 3D measurement systems are known and already used in industrial applications. They are based on optical standard methods like phase shifting or laser scanning. However, this signifies a so called stop-and-go system being awkward because on each stop the robot needs some time to die.

To avoid stop-and-go systems we aim a continuous 3D scanning method based on line cameras and special illuminations. Line cameras distinguish themselves by high resolution and fast data acquisition. Both properties are essential requirements for high velocity and high accuracy. However, the high line frequency and the short exposure time allow an image acquisition during the movement of the robot. So the developed methods are suitable for continuous scanning processes getting a high resolution 3D surface reconstruction. (E. Lilienblum ☎-11126, S. v. Enzberg ☎-20113)

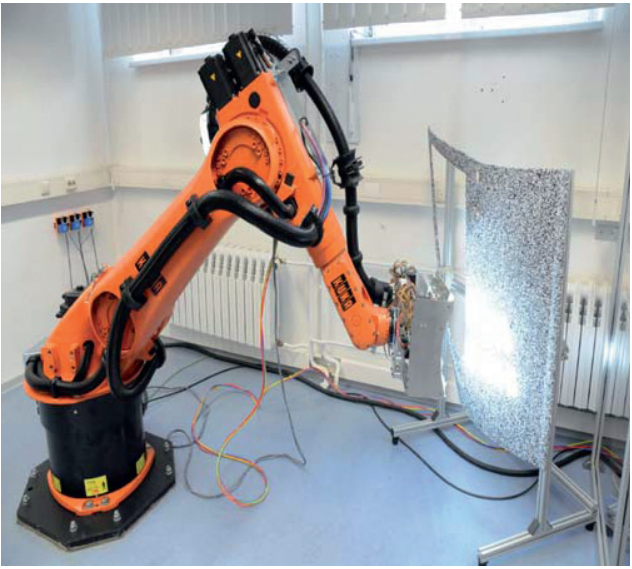


Figure 3: Robot with optical sensor scanning a car door

3D Surface Measurement with Active and Passive Line-Scan Systems

The measurement of workpieces for 3D surface inspection within a given clock rate require both high precision and high speed. There are a lot of applications, where the well-established 3D methods fails on that demands. One of the main reason for that is the use of matrix cameras. Either we do not have enough light or we do not have a sufficient resolution or we do not have the respective frame rate.

We try to solve those conflicts by using line-scan cameras instead of matrix cameras. The first approach is based on the correlation of texture on the surface. Of course, the lateral resolution of this approach depends on the correlation window size, whereby an increasing window size results in a decreasing lateral resolution. For some 3D applications the approach is still purposeful. But for other 3D applications, the decreasing resolution and the dependency from surface texture is an exclusion criterion.

In matrix camera systems there is the same conflict. It led to active approaches using coded light. But coded light approaches cannot be directly transferred to line-scan cameras. In general, the light intensity of the available projection technology is insufficient and all coded light approaches imply a static imaging setup. Thus, we pursue two goals. We research on new lighting techniques which generates light patterns with high intensity on a line. And we research on new photogrammetric techniques which apply coded light methods on continuously captured line-scan images. (E. Lilienblum ☎-11126, S. v. Enzberg ☎-20113).

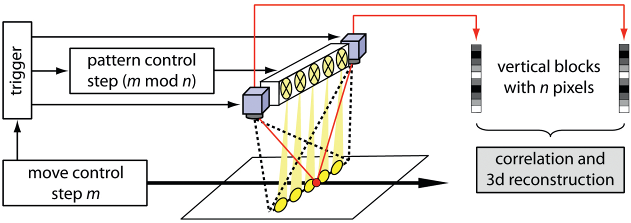


Figure 4: A line-scan system using coded light

Dynamic Height Measurement for Adaptive Focused Image Capturing

The fast imaging of large volume objects on conveyor belt systems for the distribution of packages or luggage is not possible with simple camera arrangements. The main problem is focusing the optical system at different levels of objects. The speed of image acquisition is limited primarily by the height measurement and the mechanical focusing.

To avoid this limitation we relinquish in our approach on a mechanical focus, which allow a significant performance increase. By using the Scheimpflug principle the focus works purely electronically without moving parts like mirrors and lenses, etc. Thus, the latency between height measurement and focusing the image acquisition can be very short.

Of course, a presumption for a short latency is also is a dynamic and fast height measurement, which is the second challenges in this topic. We use a single line scan camera with roughly structured light working on an appropriate height range. The method provides not only single hight values to focus the optical system but a lines of hight values. Considering the conveyor belt movement we get a surface measurement of the captured objects which allows us to implement features such as volume calculation, detection of deformations, object registration and serialization. (E. Lilienblum ☎-11126)

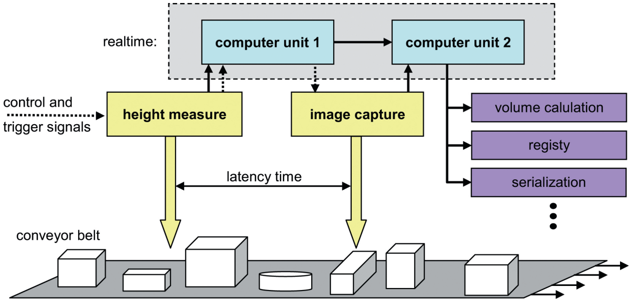


Figure 5: Processing chain on a conveyor belt

Digitalization of Warped Documents

Digital archiving of literature becomes increasingly an essential part of the work of libraries and museums. But on normal scanner technologies it is hardly possible to get distortion-free copies from hardback books without damaging them. This represents especially for digitalization of valuable historical books a large problem. There are many different approaches to remove the projective distortion, but in case of pure software solutions the exactness and reliability are not particularly high.

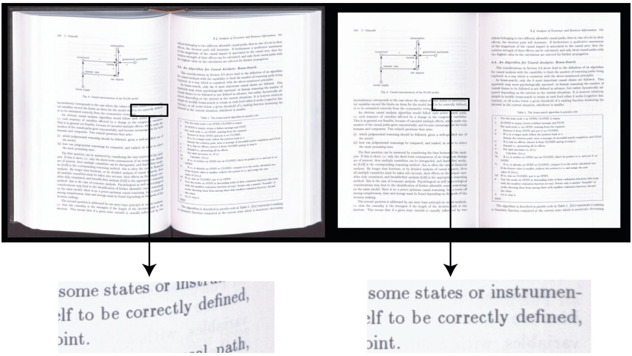


Figure 6: Dewarping a document

Our contribution to handling the problem is a hardware extension for a top view scanner. A top view scanner captures a copy of the book page from above and from a certain distance. Through the use of a line camera with an appropriate stripe lighting it is possible to receive an evenly sharp and well illuminated two-dimensional image. However, we get inevitably a distorted copy in consequence of the projective capturing and the warped page. Using an additional camera we take during the scanning process an image sequence of the moving stripe lighting of the scanner. From this image sequence we calculate a 3D surface reconstruction of the warped page through a special kind of light sectioning. By combining the surface reconstruction and the original scanner image we can calculate a nearly distortion-free copy of the book page. (E. Lilienblum ☎-11126)

3D Reconstruction of Hydrogen-Induced Cracks

Hydrogen-induced Cracking (HIC) results from atomic hydrogen being absorbed by solid bodies made of steel. There are cracks that connect adjacent hydrogen blisters on different planes in the material. In consequence, the life time of steel tanks containing chemicals for instance will be reduced and cause risks to man and environment.

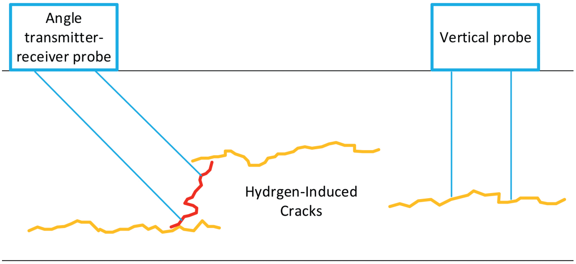


Figure 7: Crack inspection with different orientations

Usually HIC errors are detected subjectively by an examiner using a hand-guided ultrasonic sensor. Our approach to support an area-wide inspection is based on spatial tracking of sensor probes with various incident angles using a stereo camera system. Finally, measurements of different probes are fused together. (C. Freye ☎-11492, C. Bendicks ☎-11473)

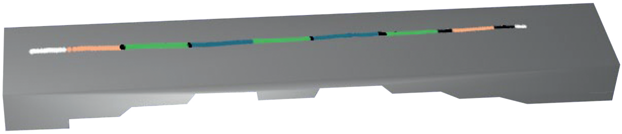


Figure 8: Color-coded track of a straight beam probe

3D Wall Thickness Measurement on Pipe Segments

A new system is being developed with which the corrosion course can be measured in pipes with ultrasound and be shown three-dimensional. This work is done in close cooperation with the MBQ GmbH. The sphere of activity is the nondestructive inspection of wall thickness in pipeworks. At the moment corrosion investigations are carried out by means of radiological examination or conventional ultrasonics. At this, only a fraction of the surface is inspected.

The aim of the research project is to track an ultrasonic sensing head in 3D by a stereo camera sensor while it is moved over the pipe surface. The position measurement should take place through the evaluation of optical fiducial marks. A new system originates which allows a three-dimensional representation, interpretation, and storage of ultrasonic data. In this way, the whole pipe surface in the effective range of the stereo camera sensor can be captured without gap to ensure a nearly one hundred percent evaluation of wall thickness. (C. Bendicks ☎-11473, C. Freye ☎-11492)

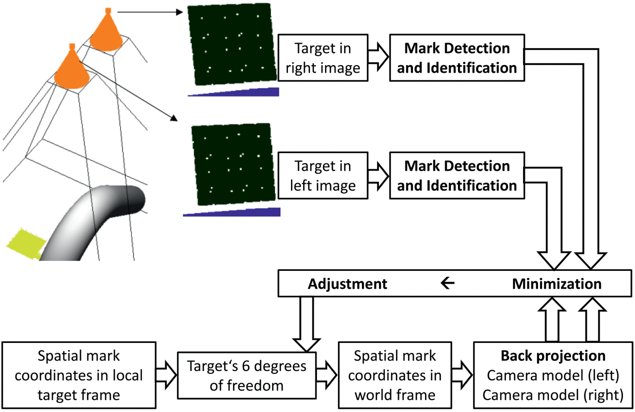


Figure 9: Target localization

Vehicle Environment Perception

Different advanced driver assistance systems (ADAS) are available in today's serial vehicles and provide important contribution in comfort-oriented as well as safety-oriented functions. In 1999 the first ACC (adaptive cruise control) system was launched on serial passenger cars. Since then the variety of ADAS was raising steadily. Many systems have shown great effort in highway application such as Side Assist, Lane Keeping Assist, etc. Nowadays research

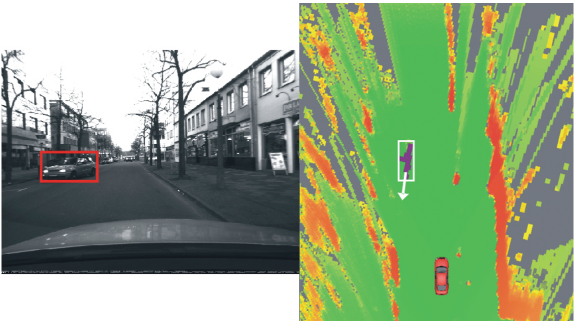


Figure 10: Grid-based environment perception in combination with object tracking

activities are addressing assistance for inner-city scenarios as well. A robust environment perception with sensors like radar, vision and laser is required, which is able to cope with the unstructured environment in urban scenarios. These and other demands require alternative and innovative sensor fusion concepts. Grid-based techniques are of high potential for these scenarios. We focus on using the advantages of the compact description in object-based tracking with the advantages of grid-based environment description to yield a hybrid object-/ grid-based approach. A main challenge is to develop an extended grid-based fusion which is responsive to dynamic information but still maintains the robustness against noise. The goal of this hybrid approach is to provide a better recognition of other traffic participants, drivable areas as well as static obstacles. (M. Heuer[1], A. Al-Hamadi ☎-18709)

[1] In cooperation with Volkswagen AG

3D Vision for Autonomous Driving

Environmental sensing is important for driver assistance systems, autonomous robots and vehicles interaction with their surroundings. There are a multitude of sensor techniques that are available for environment sensing, such as laser scanning, radar, and ultrasound etc., which can be implemented in combinations to balance their respective weaknesses. The goal of this project is a robust, real-time 3D object recognition, measurement and tracking system which uses a continuous data stream of a stereo camera system. The measurement range can be adapted to the application and has its maximum at 150 m. A depth map is determined from the stereo image, whose data is then fed further processing for object recognition and position determination.

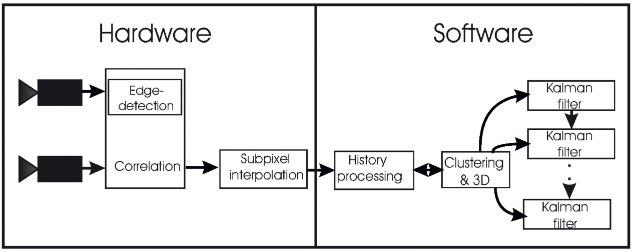


Figure 11: System concept

A fast area correlation of an image pair taken by two cameras is used for the 3D measurement. As this is the most intensive computation section, it is realized using full parallel hardware structures with massive pipelining in a FPGA. The depth map is passed to the processor where the more sequential sections are running. Statistical cluster methods are used to detect regions of a certain height and similar local coordinates. Each one of these clusters indicate an image region of a raised object, which possibly represents a vehicle. A 3D coordinate is calculated for the center of every cluster. (M. Tornow[1] ☎-11481, A. Al-Hamadi ☎-18709)

[1]In cooperation with Volkswagen AG

Automated Tank Roof Inspection

In close cooperation with MBQ GmbH, the concept for an autonomous robot system is being developed. The robot will be equipped with an ultrasonic sensor to measure the wall thickness of a tank roof. To guarantee the technical security at the operation of refineries, a monitoring of the corrosion infestation is prescribed. At the moment the corrosion is investigated manually and random check-like. This system has the ability to secure the safety of the inspector and give a nearly complete overview of the corrosion infestation.

The aim of the research is to locate and navigate the robot in such an industrial environment. Therefore a stereo camera system is used to find landmarks (see figure) and calculate 3D point clouds of them. These points are then used to locate and navigate the robot on the tank roof, while collecting ultra sonic measurements of the surface. Because of the limited access to real testing areas, the SLAM algorithms will be tested on a self developed simulator first. (C. Bendicks ☎-11473, C. Freye ☎-11492, A. Al-Hamadi ☎-18709)

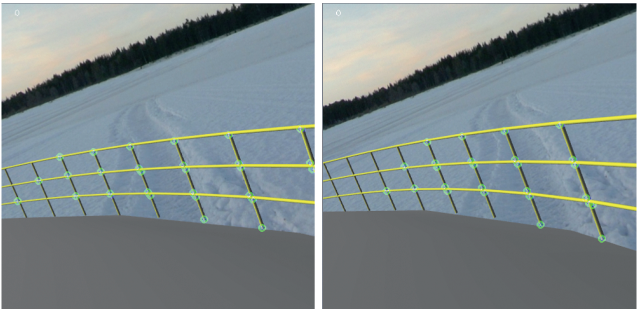


Figure 12: Results of the landmark extraction.

3D Simultaneous Localization and Mapping

Due to advances in sensing technology, more and more high quality depth cameras are available at low cost. If the camera is mounted on mobile platform like a robot, tablet, or smart-phone, a common task is the simultaneous localization and mapping (SLAM), i. e. the estimation of the camera movement and (simultaneously) the creation of a 3D surface model of the environment.

A field of research in our group is the analysis and advancement of methods for real-time 3D SLAM with depth cameras. We work on benchmark datasets for comparing methods and assessing the accuracy of the visual odometry and surface reconstruction. Further, a focus is on cooperative robotics, i. e. on mapping the environment with multiple robots cooperatively. (D. Werner ☎-11671, Ph. Werner ☎-11231, A. Al-Hamadi ☎-18709)

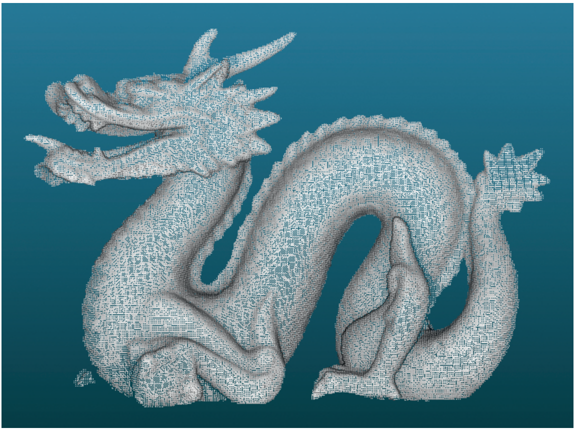


Figure 13: Object surface reconstructed from a series of depth maps with a 3D SLAM approach.

Affective State Recognition in HMI

In a Human Machine Interaction (HMI) scenario it is important to estimate the affective state of the user. This allows to control the strategy in a dialog in order to adapt to the skills of the user. Different modalities, such as audio, video and physiological signals can be used to obtain features in order to derive an estimation of the affective state of the user which could be relevant for his disposition. In figure 14, a time window of features, generated out of facial features (action units), prosodic inputs and gestures is used to estimate whether the test persons are in a relaxed (baseline) or stressed (challenge) affective state. Investigations showed that quite simple classifier architectures, such as linear filter classifiers, provide good results if the input information is supplied in a suitable way. Experiments have been performed with subjects of different age, gender and educational background. (G. Krell ☎-11476)

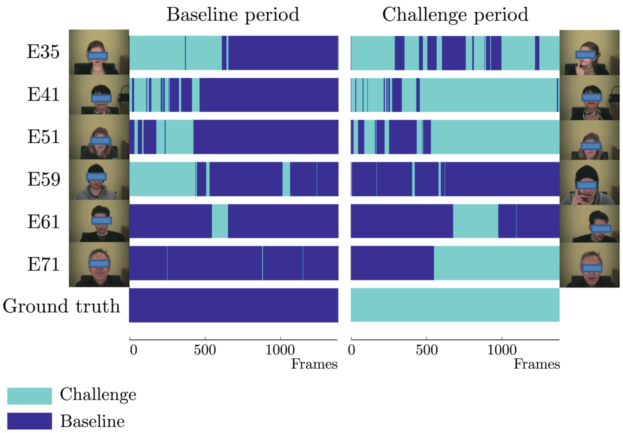


Figure 14: Classification of the affective states Baseline and Challenge for selected (anonymized) subjects in Baseline (blue) and Challenge (cyan) period based on a temporal window of facial features.

Analysis of Facial Expressions

Automatic analysis of the human face is a vivid research topic with numerous potential applications involving face recognition for security and law enforcement as well as facial expression analysis in HCI scenarios. Also automatic pain recognition presents a potential application. In general, robust facial analysis demands for an accurate localisation and tracking of the face. A number of known techniques like Deformable Templates, Statistical Models, and Active Appearance Models address this task. However, many of them do not incorporate stereo and color information or prior knowledge, such as calibration data of the cameras and subject specific model data. Integrating such information, the recognition can greatly be improved. Common techniques also assume that the person observed is cooperative but in many applications it is not feasible to constrain the user in order to always acquire frontal images of the face. The work at NIT addresses this issue. In particular, by using combined 2D/3D features like color, gradient, optical flow information and automatically generated person specific face models we achieve great invariance with respect to the so-called pose problem. Our framework achieves robust and superior classification results across a variety of head poses with resulting perspective foreshortening and changing face size (R. Niese ☎-11483, A. Al-Hamadi ☎-18709).

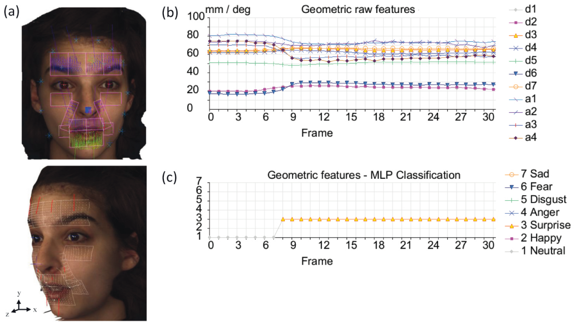


Figure 15: Example, a) 2D/3D feature extraction, b) raw features and c) MLP classification

Gesture Recognition

The recognition of gestures is an active research area and has many applications such as sign language recognition, augmented realities or the gesture-based control of technical systems. Moreover, gestures are in addition to speech and facial expressions a key modality, when it comes to determine the affective state of an user in an HMI environment. Gesture-recognition systems rely on a robust body part detection. Since the human body is capable of an enormous range of poses, this is a challenging task. In the NIT group, new technologies for human pose estimation are developed: Depth information is used to generate a graph-based surface description of the human body. Geodesic paths on the surface are extracted and classified to identify the distinct body parts. Their positions, in combination with derived parameters like velocities, form the feature vectors for the subsequent gesture classification. Among others, Hidden Markov Models and Conditional Random Fields are used as classifiers. We could demonstrate the accuracy and robustness of our system in recent experiments: In a database of more than 10k depth images, the mean joint position errors were below 4cm. In a gesture recognition experiment with 260 tested gestures, our system could recognize 92.3% of all gestures correctly (S. Handrich ☎-11150, A. Al-Hamadi ☎-18709).

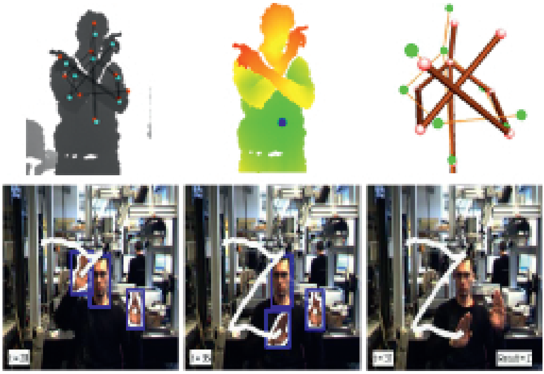


Figure 16: Top: Depth image, geodesic distances and final skeleton. Bottom: Gesture system output.

Automatic Pain Recognition

The assessment of acute pain is one of the basic tasks in clinics. To this day, the common practice is to rely on the utterance of the patient. For mentally affected patients this is little reliable and valid. For non-vigilant people or newborns it cannot be used at all. However, there are several characteristics that indicate pain. These include specific changes in the facial expression and in psychobiological parameters like heart rate, skin conductance or electrical activity of skeletal muscles.

We are working towards an automatic system, which can distinguish whether a patient feels pain or not, and can assess the intensity of the pain. Based on experiences in facial expression recognition our system can already distinguish facial expressions of pain from others and rate the intensity of the expression (see Figure below). In the current comprehensive study, we investigate the relations between pain, the facial expressions and the psychobiological feedback. The results are used to improve the robustness, reliability and validity of our system. Further, the data recorded in the study, which is named BioVid Heat Pain Database, is available to the scientific community. In the project, we collaborate with the Emotion Lab of the University of Ulm. (Ph. Werner ☎-11491, A. Al-Hamadi ☎-18709)

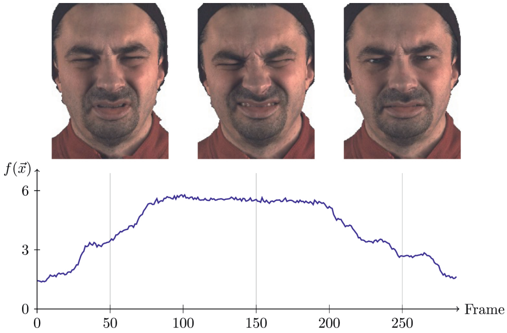


Figure 17: Pain expression intensity estimated on a video sequence with three exemplary frames.

Visual Measurement of Vital Signs

Vital signs like pulse, respiration, or body temperature are important parameters in medicine, affective computing, and related areas. Several vital signs can be measured with video-based methods, which are more convenient than classical contact-based methods.

One research focus of our group is visual heart rate measurement from the face. Current methods can only be applied in very constrained scenarios. In contrast, we work towards practical applicability by reducing measurement time and improving the robustness regarding head movement, facial expression, and illumination changes. (M. Rapczynski, Ph. Werner ☎-11231)

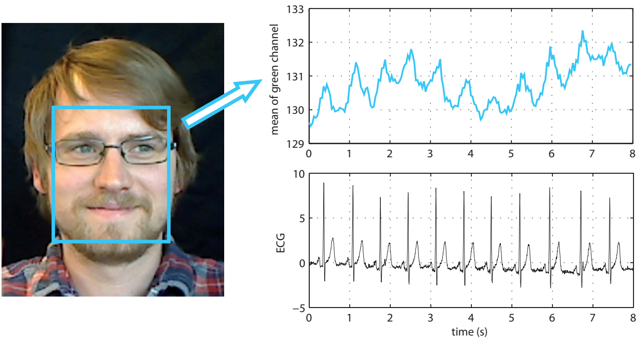


Figure 18: The heart rate measurement utilizes subtle changes of the skin color. If there are no motion or lighting artifacts, these changes reflect the blood volume pulse and occur in the rhythm of the heart beat (compare with ECG).

3D Measurements for a Virtual Training Simulator used in Laparoscopic Surgery

A laparoscopy is an operation used to look inside the abdomen. A thin instrument called a laparoscope (similar to an endoscope) is inserted through a tiny cut to help the doctor look, examine and operate in the abdomen without making large cuts. In laparoscopic surgery doctors need wide practical experience. This motivates the development of a virtual training simulator for educational purposes. Beside the modelling of preferably realistic visual effects, modelling of mechanical properties of organs during an operation is the goal of this project. Therefore, a force-feedback system will be intergrated into the simulator. Scientists of engineering mechanics needs data from measurements of force and torque, which operates at the instrument, and their interaction with the surface of inner organs in abdomen. Ex vivo 3D measurements of surfaces of animal organs under the influence of instruments has been carried out. A 3D measurement system based on an endoscope will be developed to enable in-vivo measurements to simulate more realistic visual and mechanical properties. (N. Riefenstahl, A. Al-Hamadi ☎-18709)

Sensor Fusion in Medicine

In medical therapy, different kinds of imaging devices are used for diagnosis and guidance. In cancer treatment, for example, the correct and reproducible positioning of patients at the treatment machine (e.g. linear accelerator) as well as monitoring and control of the correct position during the treatment are important. In contrast to the diagnosis and planning phase, the possibilities to gather information on position and motion of inner organs and bones during irradiation sessions are limited.

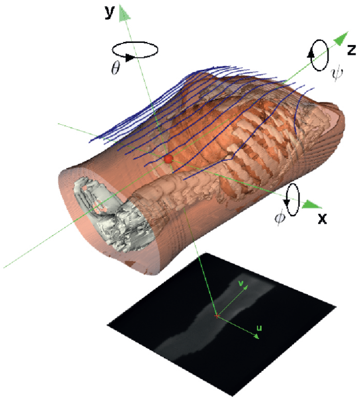


Figure 19: Fusion of surface, projection and volume data in radio therapy

Actually, optical 3D surface sensors are developed and installed at the treatment machine enabling a fast acquisition of surface of patient's body.

Optical 3D sensors can detect motion in the direction vertical to the irradiation plane which is otherwise impossible. This motivates the fusion of surface data obtained from the optical 3D sensor with other existing imaging devices, such as electronic portal images and CT data to improve positioning accuracy.

(G. Krell ☎-11476, A. Al-Hamadi ☎-18709)

Unconstrained off-line Arabic Alphabet based Handwriting Recognition

People nowadays expect that modern as well as historical human knowledge and cultural resources are digitally available as electronic text (e.g. Unicode, ASCII, and, etc.), which can be fast, efficiently and easily accessed. The technical means for converting images of typewritten, handwritten or printed text into a digital form, is what so called Optical Character Recognition (OCR).

Our research in NIT is focusing on OCR issues related to the Arabic alphabet based scripts, e.g. Arabic, Persian, Urdu, Ottoman, and etc. We started by creating our own handwriting database (IESK-arDB), in which limitations of the only one available database, are avoided.

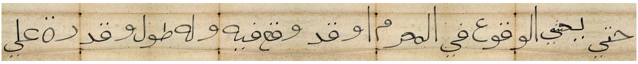
Given the vital importance of segmentation in the OCR process, a segmentation approach that makes use of topological and geometrical features is proposed, in order to identify the character borders within a word. Results were very satisfactory and to our knowledge outperforming literature available results so far.

Figure 20: Source text image (*left*) and segmented text image (*right*)

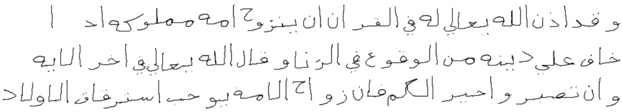
Future works will investigate issues like cyclic segmentation-recognition. Furthermore, a holistic sub-word based approach will be researched as an alternative that avoids the drawbacks of complete holistic based methods, particularly the restricted lexicons. In addition, it bypasses the segmentation phase, which tends to be expensive and error-prone. (M. Elzobi ☎-11065, L. Dinges ☎-11487, A. Al-Hamadi ☎-18709)

Synthesis of Arabic Handwriting

Comprehensive handwriting databases are crucial to train and test script recognition systems. However their generation is expensive in sense of manpower and time. As a result there is a lack of such databases which impedes research and development. This is especially true in case of holistic word recognition, since various samples must be available for each entry of the underlying vocabulary. To bypass this problem for Arabic, we present an efficient system that automatically generates images of synthetic handwritten words or one column text pages from unicode. Detailed Ground Truth files are generated automatically, enabling training and validation of diverse methods.



(a)



(b)

Figure 21: Example of a synthesized text-line (a) and -paragraph (b), using different rendering techniques.

A total of 28046 online samples of multiple writers are created to compute Active Shape Models (ASM) for over hundred letter classes. ASMs are used to generate unique letter representations for each synthesis. Subsequently these representations are modified by affine transformations, smoothed by B-Spline interpolation and composed to text. Finally the text is rendered and saved. In this way our system produces off-line pseudo handwritten samples, as shown in fig. 21, with variations in shape and texture. We compare samples of the IFN/ENIT database with corresponding syntheses to show that these can be used to surrogate real samples. (L. Dinges ☎-11487, M. Elzobi ☎-11065, A. Al-Hamadi ☎-18709)

Video-based Event Recognition

Scene understanding includes: state recognition, event detection and recognition, and situation interpretation. We focused on the event detection and recognition which bridges the gap between the raw data and the high level description of identities and activities. However, many pre-requisites are required before addressing the "video understanding". The research is divided into three phases:

In video analysis, the research is focused on tracking multiple objects considering the real-time challenges. Many techniques have been proposed to address these crucial issues but still no general solution exists. For that, we have proposed two novel approaches to handle the multi-object tracking under confusion based on Correlation-Weighted Histogram Intersection (CWHI) and the second approach multiple features of moving object are fused along with CWHI. Further, we are interested in integrating the inferential knowledge interpretation framework to recognize the state of moving object under inter-object occlusion and separation. After tracking, classification of detected objects will be carried out to obtain the individual objects in the scene. These individual objects are then interpreted separately to infer their activities.

(M. Elmezain, A. Al-Hamadi ☎-18709)

3D PTV using Colored Tracers

3D particle tracking velocimetry (PTV) is an established technique in the field of fluid mechanics to obtain 3D velocity fields up to large Lagrangian trajectories (see figure). Such information helps to understand the development of flow instabilities in turbomachines. After examining turbulences (e.g. rotating stall) in the liquid phase, it is now important to understand the birth of instabilities in the gas phase. For this purpose a 3D PTV system with a high temporal as well as spatial resolution is being developed based on a three-camera setup. This ensures the ability to investigate gas flows at relatively high speeds and involving small eddies. In that case it is clear that PTV can only be successful when there is a high concentration of seeded tracer particles. To simplify the correspondence problem due to particle projections in multiple views and in successive images the approach is based on the application of different colored particles. It diminishes the apparent particle density without inducing any a-priori restriction for the measuring accuracy.

(C. Bendicks ☎-11473, A. Al-Hamadi ☎-18709)

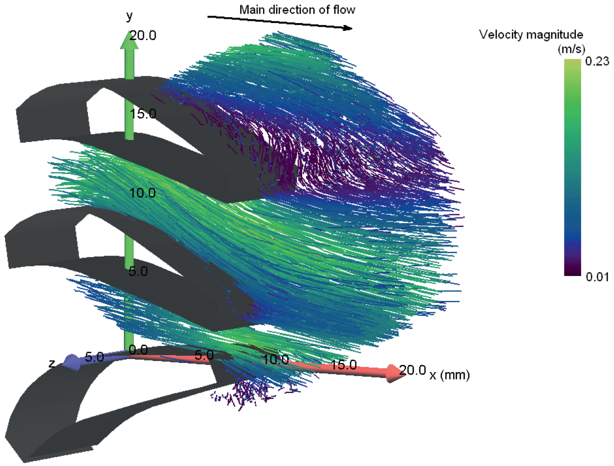


Figure 22: Reconstructed particle trajectories in 3D

Multi-Objects Tracking in Color Image Sequences

Multi-objects tracking and data association have to deal with difficulties existing in single object tracking, such as changing appearances, non-rigid motion, dynamic illumination, occlusion, and problems related to multiple objects tracking including inter-object occlusion and multiple object confusion. Even though a number of tracking methods have been introduced in the literature, many criteria still cannot be met. Our proposed paradigm integrates object detection into the object tracking process and provides a robust tracking framework under ambiguity conditions. I.e., firstly it distinguishes between real world objects, secondly extracts image features like motion blobs and color patches and thirdly abstracts objects like meta-objects that shall denote real world objects.

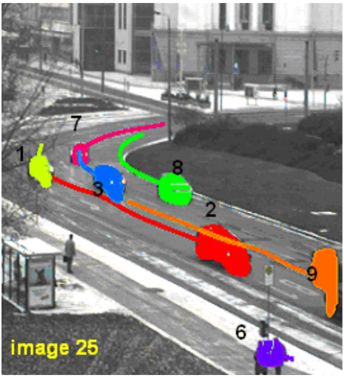


Figure 23: Tracking of multiple objects.

Through such a tight integration of the motion blobs and color patches, as well as the global optimization of object trajectories, we have accomplished robust and efficient multi-object tracking, the ability to deal with merging/splitting of objects, irregular object motions, changing appearances. This are challenging problems for the most traditional tracking methods. (R. Niese ☎-11483, T. Ljouad ☎-11922, A. Al-Hamadi ☎-18709)

Electronic Image Correction for Image Acquisition, Reproduction and Document Processing

Artificial neural networks are applied to correct typical errors in digital imaging devices, such as blur, noise, geometric distortion, chromatic aberrations and vignetting. The transfer properties of image acquisition or reproduction systems are compensated by trained correction net that takes the influence of noise and local dependencies of parameters into consideration. In order to meet the

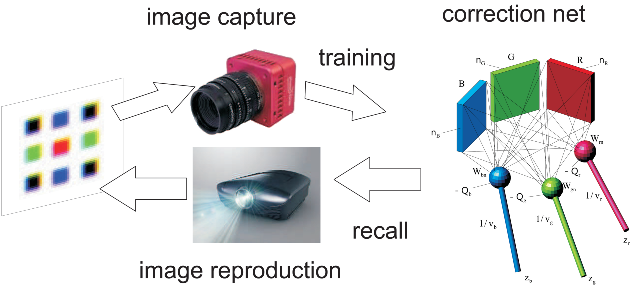


Figure 24: Neural correction net for imaging system

speed requirements of video applications hardware solutions, such as GPU or FPGA implementations, for real-time video processing are considered.

Automatic analysis of documents like forms and cheques and the evaluation of the containing data is an important challenge for information managing systems. One focal point of research is the development of document processing systems that can be adapted to a wide range of tasks using learning capabilities.

(G. Krell ☎-11476, A. Al-Hamadi ☎-18709)

Design and Analysis of Biologically Inspired Neural Networks

The world's most powerful *computer* in the field of computer vision is the human brain. The successful application of classical artificial neural networks to typical computer vision tasks, such as image understanding, object recognition and image restoration proves this already and has led to an attempt to duplicate the abilities of human vision. In neurobiology, the components of biological visual systems and the involved neural structures have been studied extensively. Several aspects of the information encoding and processing, however, are not yet fully understood. In cooperation with the Institute of Biology and the Leibniz Institute for Neurobiology, the NIT group contributed to this research with the design and simulation of biologically inspired neural networks. This included the modeling and analysis of self-organizing neural structures, the simulation of dopamine-modulated spike-time-dynamic-plasticity (STDP) learning mechanisms and the design of activity dependent axon connection strategies. Further research points were the automated segmentation of axons and the reconstruction of dendritic spines from microscopy images (S. Handrich ☎-11491, G. Krell ☎-11476, A. Al-Hamadi ☎-18709).

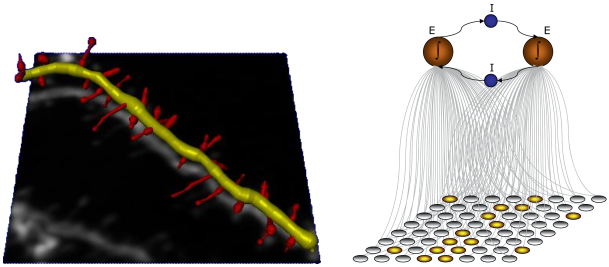


Figure 25: Left: 3d visualisation of a dendritic spine reconstructed from the underlying microscopy image. Right: Biological plausible Winner-takes-all (WTA) architecture. The stimulated input layer activates two excitatory neurons (E) that inhibit each other via inhibitory neurons (I).

Technical Equipment

The Neuro-Information Technology (NIT) group is equipped with optical measuring devices for the acquisition of 3D and motion parameters. In particular, besides high speed camera systems, active and passive stereo and multi-camera sensors are available as well as suitable hardware for processing. A key element for simulation of artificial neural networks and simulations is the Beowulf computer cluster, which also receives requests beyond the NIT group. A current list of available hardware and software at the NIT group is the following:

- Beowulf computer cluster (capacity: 14 TB, working memory: 1 TB) consisting of 50 nodes which corresponds to 272 kernels 2.6 GHz
- Spherical camera system
- Stereo camera systems with hardware processing
- Time of flight camera system
- Infrared camera
- Active vision scanner with flash bulb
- Software for camera calibration, ANNs, numerical computations and simulations
- FPGA development system
- Lab for human-computer interaction with a multisenor system (SFB-TRR62)
- Test vehicle for driver assisting systems (VW AG)
- Multi-robot system
- Set-up for pain analysis with multi-camera system and Nexus biodata gathering
- Mobile assistance system with augmented reality (Vuzix Star 1200)
- Highspeed camera system (3 cameras with 4000 frames/sec)

Contact Information

Mail: Otto von Guericke University Magdeburg
Institute for Information Technology
and Communications (IIKT)
Neuro-Information Technology
Universitätsplatz 2
D-39106 Magdeburg

Phone:

☎ Prof. Al-Hamadi	+49 391 67 - 18709
☎ Group members	+49 391 67 - (see text)
Fax:	+49 391 67 - 11231
E-Mail:	Ayoub.Al-Hamadi@ovgu.de
Web:	www.iikt.ovgu.de/NIT