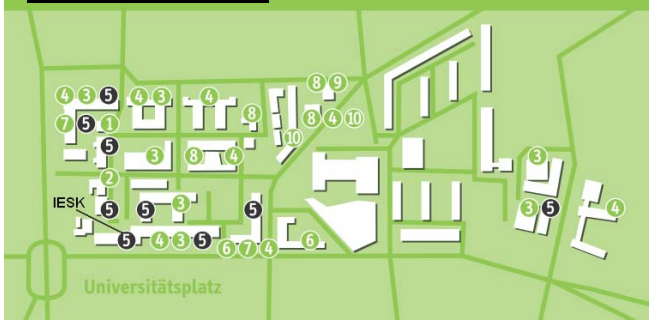
DER UNIVERSITÄTSCAMPUS

1 Rektorat
2 Dezernat für Studienangelegenheiten
3 Fakultät für Maschinenbau
4 Fakultät für Verfahrens- und Systemtechnik
5 Fakultät für Elektrotechnik und Informationstechnik
6 Fakultät für Informatik
7 Fakultät für Mathematik
8 Fakultät für Naturwissenschaften
9 Fakultät Geistes-, Sozial- und Erziehungswissenschaften
10 Fakultät für Wirtschaftswissenschaft



University:
Otto-von-Guericke-Universität Magdeburg
Postfach 4120, D-39016 Magdeburg

Tel.: +49 391 / 67 - 01
Fax: +49 391 / 67 - 11156

Institute:
Otto-von-Guericke-Universität Magdeburg
Werner-von-Siemens-Gebäude
Institut für Informations- und Kommunikationstechnik (IIKT)
Lehrstuhl für Technische Informatik
Universitätsplatz 2
D-39106 Magdeburg
Tel.: +49 391 / 67 - 18709 Ayoub Al-Hamadi
- 18861 Office
- ☎ Group members
Fax: +49 391 / 67 11 231

Email: ayoub.al-hamadi@ovgu.de
http://www.iikt.ovgu.de/NIT

09-2013

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG
EIT FAKULTÄT FÜR ELEKTROTECHNIK UND INFORMATIONSTECHNIK

Institute for Information Technology and Communications (IIKT)

Neuro-Information Technology

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

Brief research profile

The research group is working on the areas of digital image processing, artificial neural networks, fuzzy systems and processor architectures for real-time processing. Its topics of interest span the basic research and applications in automation, information engineering, medicine and biology.

Lectures and other teaching activities

- Information Engineering
- Electronic Foundations of Computer Science
- Artificial Neural Networks
- Image Processing
- Pattern Recognition

Focal Points of Research

Transregional Collaborative Research Centre SFB/TRR 62 – Companion-Technology for Cognitive Technical Systems

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 following subprojects:

C1 – Environment perception: A companion system must perform its surroundings, interpreting, to recognize the user groups and communicate with him, involving the facial expressions, language and gesture. In this sub-project, the important tasks of this area detection and environment modeling that based on gesture recognition are solved. To capture the environment, methods for multi-sensor fusion, information fusion and temporal filtering are investigated based on the finite set theory, and further allow the simultaneous estimation of object and existence of the object state. The detection, tracking and classification of persons and other objects made using multi-sensory data and the classification of user gestures is purely image based on the use of hidden Markov models (HMMs). For a robust gestures classification, static and dynamic features characteristics that are appropriate for their extraction as a result of recent research can be used, also color information. The segmentation using color information and 3D information provides a high degree of robustness to disturbances, occlusions, brightness variations

and background perturbations.

The goal of this project is to create a situation model for the area under surveillance, which localizes, relates and classifies the persons and objects, as well as undertaking an assessment of the emotion and intent of the user. The created situation model is the basis for planning and decision-making levels of sub-projects of the areas of planning and decision. The basis for the creation of the situation model are different sensor data, which are preprocessed in the other C-projects. Besides the geometrical description of the scene, also the emotion of the acting person is to be recorded. This requires the development of a multi-modal emotion recognition from speech, gestures and facial expressions, and psycho-biological data of the user. The individual modules will be developed in the projects C1-C4. (S. Handrich ☎-491, A. Al-Hamadi -18709) (S. Handrich ☎-491, A. Al-Hamadi -18709)

C3 – Mechanisms of nonverbal communication: Mimic emotion recognition and analysis of head and body posture: Effective companion systems require robust methods for detection of faces and facial mimics as well as situations of non-verbal person approach, i.e. via body posture and gaze. Currently, under real world conditions these tasks cannot be carried out at a satisfying degree of quality. Thus, to achieve robust analysis of faces and body posture, new mono- and binocular image processing approaches will be investigated. For vision based emotion recognition from faces, two methods will be studied, which combine static and dynamic features in temporal and spatial image sequences. In addition, neural computational mechanisms will be developed which aim at replicating key functionalities of processing in visual cortex. Here, the segregated processing along the motion and form pathway and their fusion will be studied to analyze sensory signals in visual communication and vision-based social interaction via head and body poses. This serves to derive a deeper understanding of the details in early and mid-level visual processing in a biologically inspired architecture and their use for advanced human-computer interaction. (R. Niese ☎-11483, A. Al-Hamadi ☎-18709)

Affective State Recognition in HMI

One example is the estimation of the affective state of the user in a Human Machine Interaction (HMI) which could be important 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 1, 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 good way. Experiments have been performed with subjects of different age, gender and educational background. (G. Krell ☎-11476)

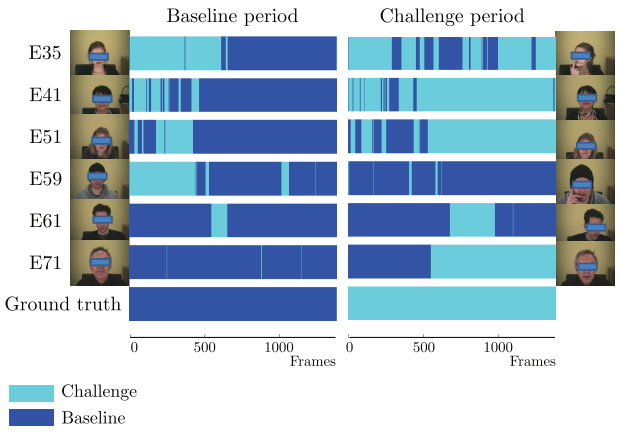


Figure 1: Classification of the affective states Baseline and Challenge for selected (anonymized) subjects in Baseline (blue) and Challenge (cyan) period based on a temporal window of facial features.

Automatic Pain Recognition based on Facial Expression and Psychobiological Parameters

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 Fig. 2). 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. In this project, we collaborate with the Emotion Lab of the University of Ulm. (Ph. Werner ☎-11491, A. Al-Hamadi ☎-18709)

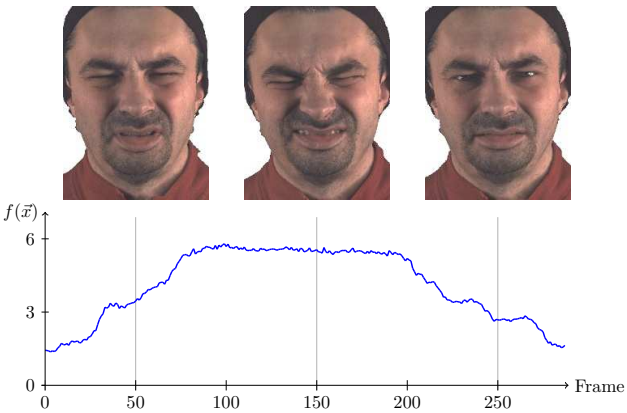


Figure 2: Pain expression intensity for a video sequence with sample frames.

Dynamic height measurement for adaptive focused image capturing of large volume objects

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. In general, special imaging systems are required, which are often very costly. The speed of image acquisition is limited primarily by the height measurement and the mechanical focusing.

Currently we are working on a new method which allows a significant performance increase compared to previous approaches. In contrast to existing systems, we avoid in our approach a mechanical focus. The focus is to work purely electronically without moving parts like mirrors and lenses, etc. The expected latency is thus comparatively very low, if the height values are calculated in real time. This is one of the challenges in this topic.

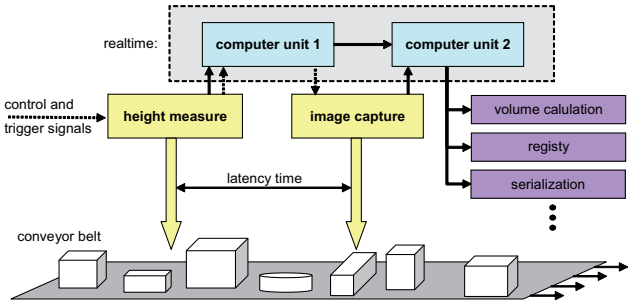


Figure 3: processing chain on a conveyor belt

The dynamic height measurement should not only provide a single value to focus the optical system. We also want to capture the objects in their spatial surface shape with a high resolution. This allows us to implement features such as volume calculation, detection of deformations, object registration and serialization. (E. Lilienblum ☎-11126, A. Al-Hamadi ☎-18709)

3-D 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 3-d 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. (E. Lilienblum ☎-11126, C. Bendicks ☎-11473, A. Al-Hamadi ☎-18709)

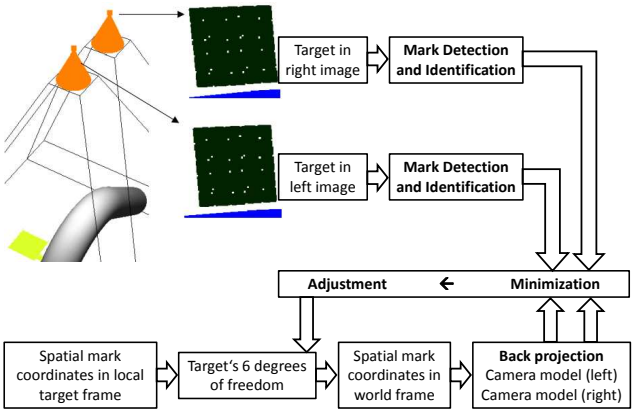


Figure 4: Target localization

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

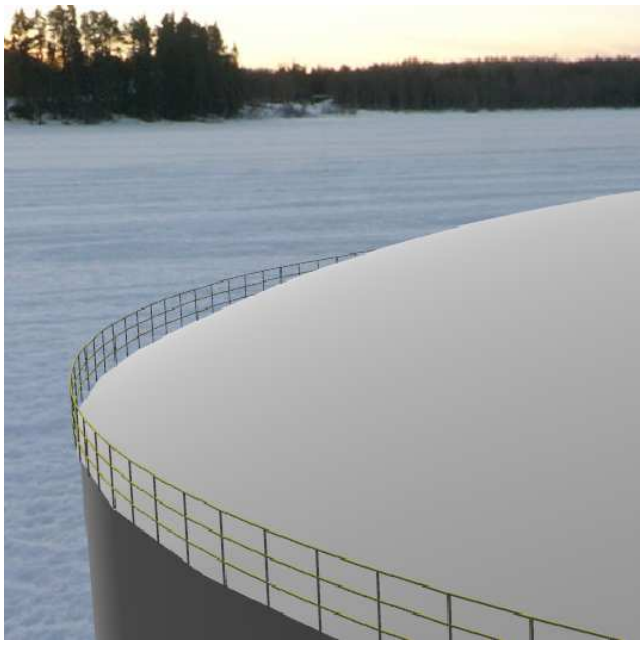


Figure 5: Simulation of the industrial environment.

research is to locate and navigate the robot in such an industrial environment. Therefore a stereo camera system is used to find landmarks and calculate 3-d point clouds of them (see fig. 6). 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,

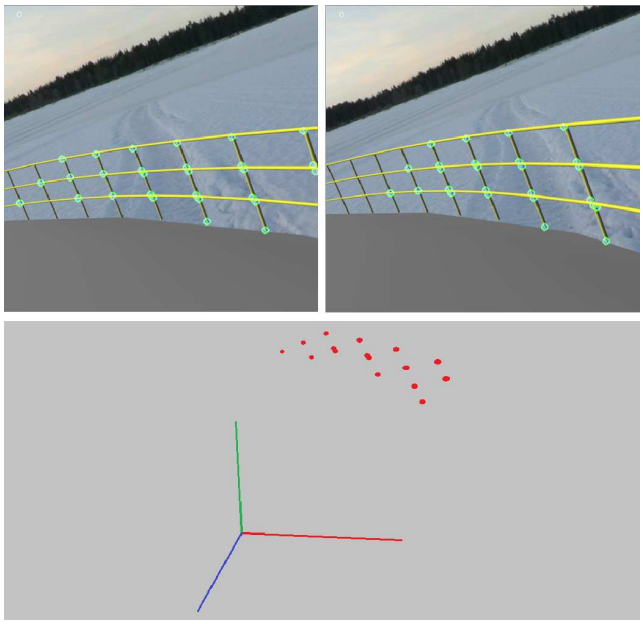


Figure 6: Results of the landmark extraction.

A. Al-Hamadi ☎-18709)

Vehicle Environment Perception in Advanced Driver Assistance System

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 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

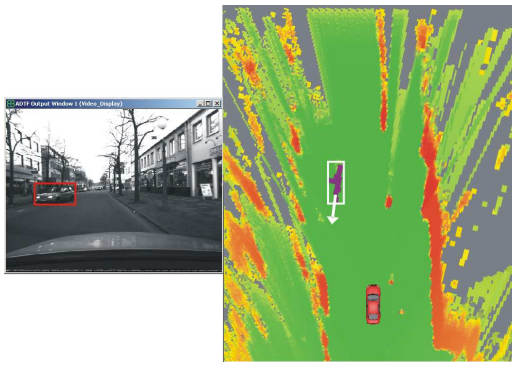


Figure 7: Grid-based environment perception in combination with object tracking

concepts. Compared to Kalman filter approaches of object tracking grid-based techniques are of high potential for these scenarios. The idea of using the advantages of the compact description in object-based tracking and the advantages of grid-based environment description to yield to a hybrid object-/ grid-based approach is the focus of this research. One disadvantage of the classic grid-based fusion is the capability of moving objects. Hence a main challenge at this point is to develop an extended grid-based fusion which is responsive to dynamic information but still maintain the robustness against noise. This prepares a decisive basic for effective combination with the object-based method. The goal of this hybrid approach is to provide a better recognition of other traffic participants, drivable areas as well as static obstacles. This is a basic for predictive and effective action concepts to enable more safety and comfort for the driver. (* M. Tornow ☎-11481, A. Al-Hamadi ☎-18709)

3-d 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

*In cooperation with Volkswagen AG

sensing, such as laser scanning, radar, and ultrasound etc., which can be implemented in combinations to balance their respective weaknesses. The majority of these techniques is based on active processes. Photogrammetry is a passive position measurement technique, in which images from several cameras are analyzed. Optical sensors perform very well for an exact recognition of the object position and give a lot of extra information for further processing.

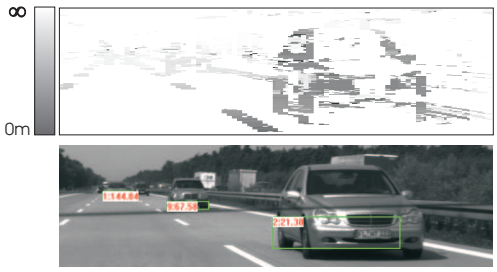


Figure 8: Depthmap (close objects are dark) and result of the clustering

The goal of this project is a robust, real-time 3-d 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.

A fast area correlation of an image pair taken by two cameras is used for the 3-d measurement. Which is mainly a matching algorithm to find the correspondences in the two images and calculate the disparity. 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. The coordinates are combined to “Clusters”. Each one of these clusters indicate an image region of a raised object, which possibly rep-

resents a vehicle. A 3-d coordinate is calculated for the center of every cluster.

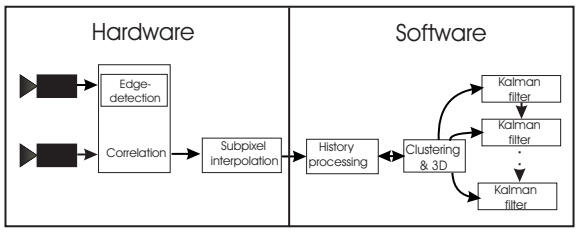


Figure 9: System concept

The results of the optical position sensor are continuous available after a determined processing time because of the use of a combination of programmable hardware (FPGA) and embedded software. (M. Tornow ☎-11481*, A. Al-Hamadi ☎-18709)

Surface inspection of deformable workpieces

Optical methods for 3d-surface measurement are widely used in industrial applications as e.g. quality control of car body parts. Next to increasing accuracy of optical sensors, research effort has to be put into methods for processing 3d data.

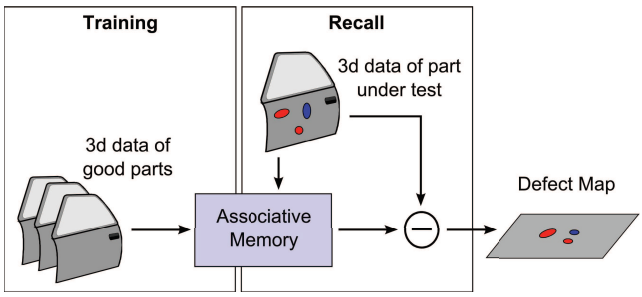


Figure 10: Operation of the Associative Memory during training with reference parts and recall with measured parts

*In cooperation with Volkswagen AG

A modified Associative Memory developed at our research group allows for detection of surface defects close to optical resolution and noise limits, even for deformable workpieces. With this approach, dents within the micrometer-range in depth can be detected within a single measurement of a 1m x 1m surface area.

The Associative Memory is based on an Artificical Neural Network which is trained with good parts, thus containing the allowed part variance in the weights of the network. After training, a reference part is obtained by recalling the network with the currently measured part. The ANN internally performs a series expansion with adapted base functions (Karhunen-Loève transform). This reference part is then used for comparison with the measurement, leading to a difference map which can be further processed for classification and projected onto the surface for localisation of defects.

Current research focuses on several issues that allow higher accuracy and easier setup and training of the Associative Memory for the evaluation of complex parts. Methods for virtual training with parts derived from CAD data is researched in a joint project with the Fraunhofer Institute Magdeburg (IFF) and the INB Vision AG. (S. v. Enzberg ☎-11126, T. Lilienblum +49 391 6117305, E. Lilienblum ☎-11126, A. Al-Hamadi ☎-18709)

3-d Highspeed Particle Tracking Using Coloured Tracers

3-d particle tracking velocimetry (PTV) is an established technique in the field of fluid mechanics to obtain 3-d velocity fields up to large Lagrangian trajectories (Fig. 11). 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 3-d 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.

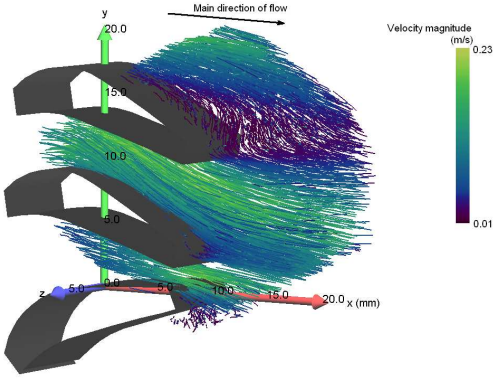


Figure 11: Reconstructed particle trajectories in 3-d

This point is known as a major issue for PTV, since the correspondence analysis between the different camera views becomes extremely difficult. To establish the spatial correspondence to determine 3d positions (by photogrammetric methods), each tracer has to be found in all pictures of the involved cameras during a timestep. This central problem can be partially solved for the correspondence both from a temporal and from a spatial point of view using coloured tracer particles. In that case, classifying in all images the initial dense cloud regarding to particle colours results in different particle subsets. Analysing these subsets for correspondences separately is much easier than analysing all particles as a whole because the number of ambiguities is reduced dramatically. Using colour classes diminishes the apparent particle density without inducing any a-priori restriction for the measuring accuracy. Only the tracer particles belonging to a specified colour group have to be recognized at successive timesteps. (C. Bendicks ☎-11473, A. Al-Hamadi ☎-18709)

Robot based Optical Test Engineering

Optical methods in the area of 3d measurement for test engineering become more and more important in the industrial production process. The performance of test

engineering systems is characterized with velocity, accuracy and sufficiency. Because the application should work also very adaptable on extensive workpieces we use as an additional technical basis industrial robots.

Generally, robot based 3d measurement systems are known and already used in industrial applications. But they are normally based on standard methods like phase shifting or laser scanning. However, this signifies a so called stop-and-go system being awkward in case of extensive workpieces. On each stop the robot needs some time to die.

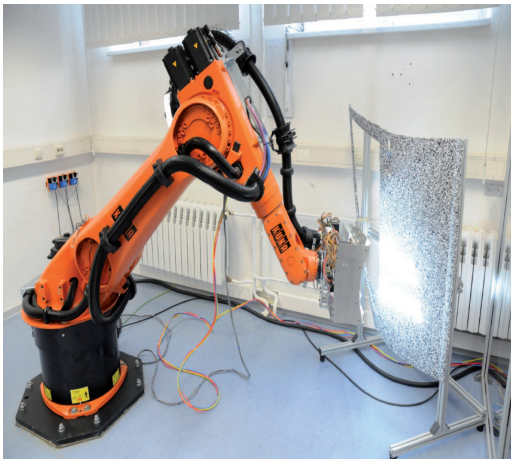


Figure 12: Robot with optical sensor scanning a car door

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.

The 3d surface reconstruction is the essential basis for test engineering. We develop new algorithms to find surface defects like dents, dints, pimples and scratches. Here we also aim continuous methods getting results

in real time with short delays. Due to high resolution and high line frequency the measurement data volume is tremendous. So, we have to implement our developed methods on parallel hardware like FPGAs and GPUs. (E. Lilienblum ☎-11126, S. v. Enzberg ☎-11126, A. Al-Hamadi ☎-18709)

Digitalisation of Warped Documents

Digital archiving of hard-back literature becomes increasingly an essential part of the work of libraries and museums. Although for this purpose the modern computer technology accomplishes already major premises, the status quo of the scanner technology is not satisfying. In particular it is hardly possible to get distortion-free copies from thick books without damaging them. This represents currently for digitalisation of valuable historical books a large problem. In addition, a full automated character recognition in the area of book crease is often impossible. Removing the projective distortion in a two-dimensional image of a three-dimensional warped surface is a very common problem. There are many different approaches. But either in case of pure software solutions the exactness and reliability are not particulary high or the techniques are costly concerning the additional hardware and provide only low resolutions in the copy of documents due to the exclusive use of matrix cameras.

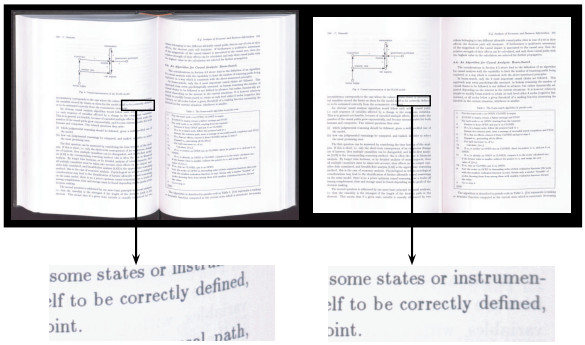


Figure 13: Dewarping a document

Our contribution to handling the problem of dewarp-

ing is the development of a hardware extension for a top view scanner. This kind of book 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 geometry of the scanner and the warped surface of the page. Using an additional camera we capture during the scanning process an image sequence of the moving stripe lighting of the scanner. From this image sequence we calculate a 3-d surface reconstruction of the warped page through a new method which is a further development 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, A. Al-Hamadi ☎-18709)

Smart Line Sensor System for Fast 3d Surface Measurement

With line-scan-cameras very high local resolution with thousands of pixels per line and high clock frequencies are possible. Therefore, they are used in 3d remote sensing applications for processing of aerial images and in close range applications for 2d inspection at the industrial production line. For 3d shape acquisition today normally matrix-cameras are used, but with comprehensive low spatial- and time-resolution. Our

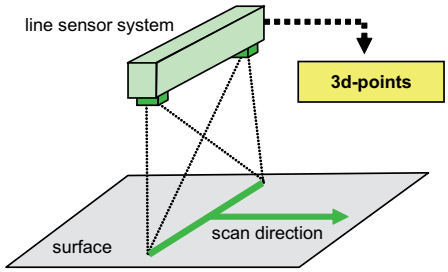


Figure 14: Setup for line based 3d surface measurement

new system (Fig. 14) combines the advantages of the

both approaches. It uses two high resolution line-scan-cameras and captures 3d-shape and 2d-texture with very high local and spatial resolution simultaneously. If the measurement object is well contrasted, no active light pattern projection unit is needed. But in low contrast regions accuracy and robustness may be improved by projecting static or dynamic illumination patterns.

The measurement system supports powerfull algorithms for offline-calibration and realtime-image-processing. The current implementation at two graphics-processing-units (2xGTX285) produces an array of 10000 by 10000 3d-points in just one second. Therfore the system allows the fast 3d-surface and 2d-texture inspection at the production line. Another possible application could be a passive 3d-panoramic scanning device for environmental scanning: here, both cameras would be mounted on top of each other moved by vertical-rotation-unit.

At the current stage of the project a shortening of the latencies between image capturing and point computation is aimed, to get a continuous scanning device and enabling closed loop distance control for the capturing process. This project furthermore serves as basic for robot guided scanning.

All the stages of development are guided by "Virtual Engineering", which refers to the simulation of the complete system chain: measurement objects, cameras and motion-units. This enables the development of image-processing algorithms at the IESK in Magdeburg without need of real images. Simultaneously, the camera-hardware is developed in parallel at the Chromasens GmbH in Constance. In this case "Virtual Engineering" allows a comprehensive short run-time of the whole project, by breaking typical dependencies. (E. Lilienblum ☎-11126, A. Al-Hamadi ☎-18709)

Electronic image correction for image acquisition and reproduction systems and for document processing

Blur, noise, geometric distortion and vignetting are typical errors of digital imaging devices. In color image processing, additional errors, such as chromatic aber-

rations, occur. This kind of irregularities - that often occur spatially variant - should be corrected for a better visualization and in order to avoid problems in further processing of the images.

The application of artificial neural networks to the correction of optical and electrical irregularities of image forming devices is one subject of research. The transfer properties of image acquisition or reproduction systems are compensated by trained correction systems that take the influence of noise and local dependencies of parameters into consideration. In order to meet the

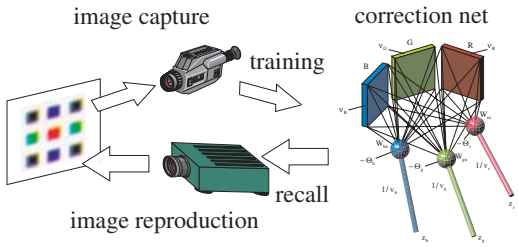


Figure 15: Image correction

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

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.

The main focus is the processing and restoration of colored documents and their correct reproduction. A learning algorithm reduces the influence of often unknown effects of intermediate image representations in the cycle between document capture and reproduction. Transfer properties of the system parts including image storage, different input and output media and the influence of compression algorithms are considered. The investigations result in generalized solutions for real-time processing of colored documents including dedicated hardware solutions where appropriate. (G. Krell ☎-11476, A. Al-Hamadi ☎-18709)

Simulation and application of neurobiological networks and design of biologically inspired self-organizing systems

In this area, bioinformatics and some of the biological background are investigated computationally within the field of neuroscience. The established researching points of interest in Artifical Neural Networks and Image Processing are concerned in cooperation with neurobiology with the understanding and applications of biologically plausible neural networks. As in the connectivity structures of early mammalian brains, the activation patterns and macroscopic phenomena developing thereby and their meaning for information processing of cognitive behaviour at an operational point of view are considered. For instance, biochemical processes of systemic changes in the brain, produced by biological learning, which are auto-adaptive and self-organizing systemic descriptions of the impact of changes on cellular level and connectivity on the systemic behaviour, are abstracted as controlling operations for applications, as associative memories from spiking neurons and self-organizing decentralized adaptive systems, which are to be contemplated.

By computational simulated phenomena of structural emergence and dynamic information overlay in the cortex, biologically detailed simulations are used to investigate universal mechanisms of stability and regularization principles of nervous systems, and are brought to transfer neural behaviour of massive data streams in real brains towards numerical paradigms of data processing in machine operations and applications for real-world-problem-solving. Detailed biologically plausible simulations, which are undertaken to observe behaviour and validate biological hypothesises from the neurobiological detailed point of view, and their complexity of modelling is accompanied by an exponential increase of the requirements of the numeric processing capabilities, so the systems are abstracted and simplified by high level parameters. (S. Handrich ☎-11491, G. Krell ☎-11476, A. Al-Hamadi ☎-18709)

Multi-Objects Tracking in Colour Image Sequences

Multi-objects tracking and data association have received considerable attention in the field of computer vision, mobile robotic, autonomous systems and intelligent transportation system (ITS). It has to deal with difficulties existing in single object tracking, such as changing appearances, non-rigid motion, dynamic illumination and occlusion, as well as the problems related to multiple objects tracking including inter-object occlusion, multiple object confusion. To be effective, any proposed method for multi-objects tracking has to meet several stringent requirements:

- automatic segmentation of each moving object, from the background, and from other objects, so that all objects are detected,
- robust operation under a wide range of real-world conditions, i.e. congestion, partial occlusions of objects,
- robust operation in a wide variety of lighting conditions, shadow, sunny, etcetera.

Even though a number of tracking methods have been introduced in the literature, many of these criteria still cannot be met. In our research work, we describe a novel object tracking technique in color video sequences, with application to multi-object tracking in crowded scenes. The proposed paradigm integrates object detection into the object tracking process and provides a robust tracking framework under ambiguity conditions. In order to reduce the computational complexity and to increase the robustness, we use a trisectional structure. I.e., firstly it distinguishes between real world objects, secondly extracts image features like motion blobs and colour patches and thirdly abstracts objects like meta-objects that shall denote real world objects.

Through such a tight integration of the motion blobs and colour patches, as well as the global optimization of object trajectories, we have accomplished not only robust and efficient multi-object tracking, but also the ability to deal with merging/splitting of objects, irreg-

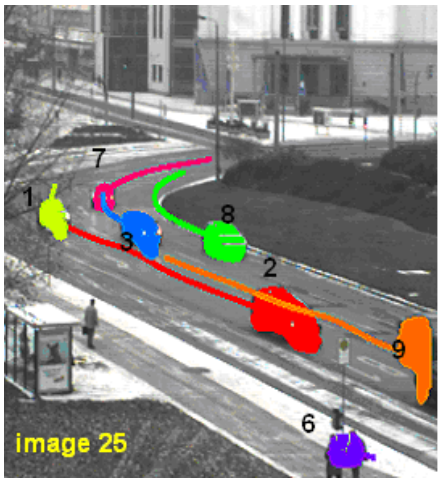


Figure 16: Tracking of multiple objects.

ular object motions, changing appearances, etc. which are the challenging problems for the most traditional tracking methods. For solving the problems of the fluctuation detection and dealings with object interactions, a data association step is suggested in further step with data exclusion, data allocation and data administration. The efficiency of the suggested technique for multi-objects detection and tracking is demonstrated and published in several papers on the basis of analysis of strongly disturbed real image sequences. (R. Niese ☎-11483, A. Al-Hamadi ☎-18709)

Analysis of Facial Expressions

Automatic analysis of the human face is a vivid research topic in computer vision. There are numerous potential applications involving face recognition for security and law enforcement as well as facial expression analysis in human computer interaction (HCI). Also automatic pain recognition from face presents a potential application.

In general, robust facial analysis demands for an accurate localisation and tracking of the face and its features or feature points. A number of known techniques like Deformable Templates, Statistical Models, Active Ap-

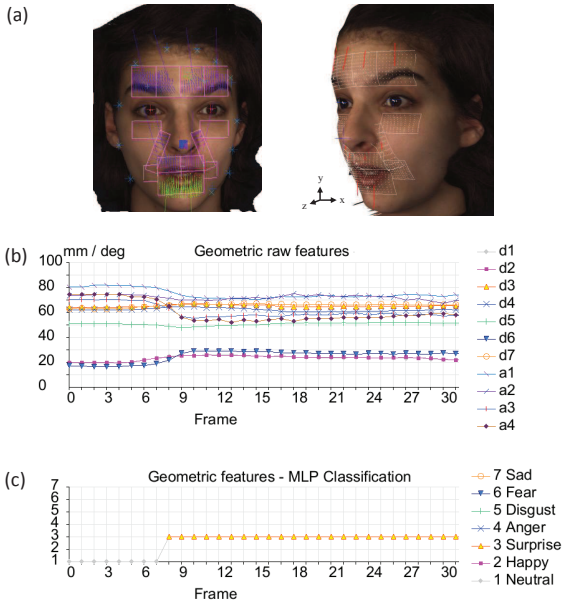


Figure 17: Example, a) feature extraction and 3D feature representation, b) raw features and c) MLP classification

pearance Models and combined Shape Models address this task. Edges, intensity maxima and other features guide the matching process of such models. However, many of them do not incorporate stereo and colour information or prior knowledge, such as calibration data of the cameras and subject specific model data. Integrating such information, not only the recognition can greatly be improved, but also the demand for robustness under varying poses and lighting conditions can be satisfied. Common techniques often assume that the person observed is cooperative. In many applications it is not feasible or possible to constrain the user in order to always acquire frontal images of the face. The work at IESK addresses this issue.

At our institute new colour, monocular and stereo based methods for common HCI and also clinical applications are being developed, which are able to automatically detect facial expressions of emotion and pain. In particular, by using combined 2D-3D feature extrac-

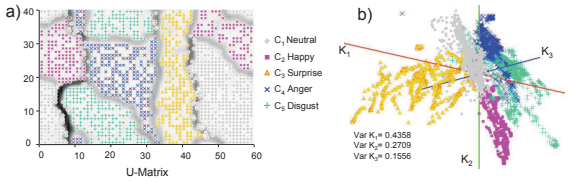


Figure 18: Feature space; a) U-Matrix of a Self Organizing Map with added class labels, b) Distribution in K1, K2, K3 space after PCA dimension reduction

tion methods, we achieve great invariance with respect to the so-called pose problem. With respect to 2D, our approach uses colour, gradient and optical flow information to extract the facial features. In the 3D part, it includes camera models and initial registration, in which the system automatically builds person specific face models from stereo. Photogrammetric techniques are applied to determine 3D geometric measures as features. Feature normalization is carried out and Artificial Neural Network (ANN) and Support Vector Machine (SVM) based classifiers are trained and applied. This leads to minimal mixing between different facial expression classes. Our framework achieves robust and superior classification results across a variety of head poses with resulting perspective foreshortening and changing face size (Fig. 17). Analysis of the feature space demonstrates the good separation of the classes (Fig. 18). (R. Niese ☎-11483, A. Panning, A. Al-Hamadi ☎-18709)

Gesture Recognition using Hidden Markov Models

Sign language recognition from hand motion or hand posture is an active area in gesture recognition research for Human Computer Interaction (HCI). Hand gesture recognition has many applications such as: Sign Language Recognition, Communication in Video Conference, using a finger as a pointer for selecting options from menu, interacting with a computer by easy way for children. Over the last few years, many methods for hand gesture recognition have been proposed. These

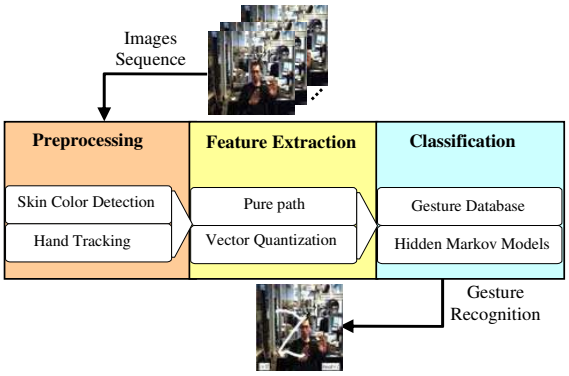


Figure 19: Suggested method for alphabets gesture recognition using HMM.

methods differ from one another in their models; Neural Network, Syntactical Analysis and Hidden Markov Models (HMM). Since HMM are used widely in handwriting and speech recognition, we develop a method to recognize the alphabets from a single hand motion using Hidden Markov Models. The gesture recognition for alphabets is based on three main stages; preprocessing, feature extraction and classification. In preprocessing stage, colour and depth information are used to detect both hands and face in connection with morphological operation. After the detection of the hand, the tracking will take place in a further step in order to determine the motion trajectory; so-called gesture path. The second stage, feature extraction, enhances the gesture path which gives us a pure path and also determines the orientation between the center of gravity and each point in a pure path. Thereby, the orientation is quantized to give a discrete vector that is used as input to HMM. In the final stage, the gesture of alphabets is recognized by using Left-Right Banded model (LRB), in conjunction with Baum-Welch algorithm (BW) for training the parameters of HMM. Therefore, the best path is obtained by Viterbi algorithm using a gesture database. In our experiment, 520 trained gestures are used for training and also 260 tested gestures for testing. Our method recognizes the alphabets from A to Z and achieves an average recognition rate of 92.3%. (M. Elmezain, S. Handrich ☎-11491, A. Al-Hamadi ☎-

18709)

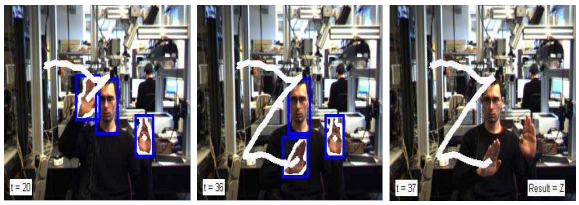


Figure 20: System output for alphabet Z.

Content-based Image Retrieval

Content-based image retrieval (CBIR) has become a very popular research area in recent years. The ultimate goal of CBIR is to develop an image search engine, not by using the text annotated to the image by an end user (as traditional image search engines), but using the visual contents available in the images themselves. A CBIR system works by getting an image from the user as a query. Then, it searches the whole database in order to find the most similar images to the query image.

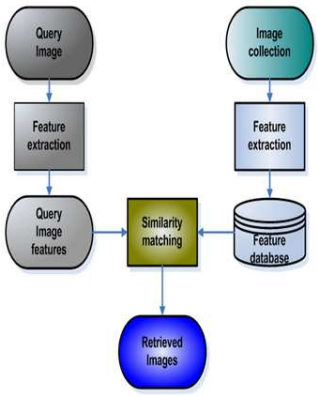


Figure 21: Process of CBIR

The key issues in developing a CBIR system are:

- Which features should be derived to describe the images better?

- Which data structure should be used to store the feature vectors?
- Which learning methods should be used in order to make the CBIR more efficient?

Therefore, we are currently working on these issues: image segmentation, colour and texture feature derivation and colour indexing. The final goal of our research is to develop new approaches for CBIR which should be efficient in terms of precision and recall compared to the existing ones. (A. Al-Hamadi ☎-18709)

Authentication of Images and Video

The wide use of digital images on Internet and in electronic commerce demand grantees against malicious manipulations and for copyright protection. Thus, digital watermarking technology has been developed to protect the copyright and integrity of multimedia contents. Digital watermarking is the process of inserting a piece of information (watermark) into multimedia contents which can later be detected and verified for many purposes, including image authentication or copyright protection. For image authentication purposes, it is essential that the watermarking algorithm is blind, secure and so sensitive to trivial changes in image contents and can localize the alteration detection.

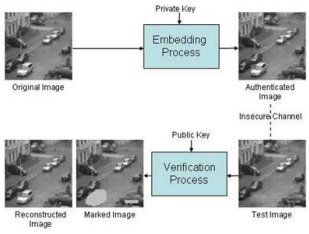
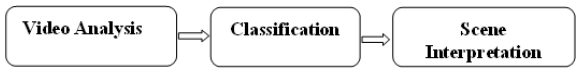


Figure 22: Block diagram of the image authentication concept

We are developing so-called self-embedding algorithms to not only localize the alteration, but also to restore the missing image contents. (G. Krell ☎-11476, 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. (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). OCR's systems can be of great help in a wide

spectrum of industries, e.g. Banking, Postal, Education, Libraries and Archives. It can be also an integral part of systems for machine translation, text-to-speech and text mining.

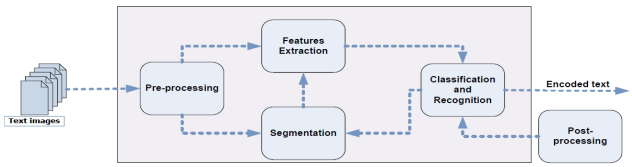


Figure 23: Model for Optical Character Recognition (OCR) System.

Our research in IESK 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-AHR), in which limitations of the only one available database, are avoided. To avoid the lack of handwriting database needed for the training of OCR systems we also research modeling of handwriting's variations by using deformable models.

Handwriting specific pre-processing is a very crucial prerequisite for the subsequent phases in an OCR system. Hence, we have developed an approach for the Arabic words base-line estimation: we combined two well known techniques, namely Hough's transform based and local minima regression based on contours, in a way that they compensate the limitations of each other. The slant correction problem is also solved by a projection profile based technique calculated upon the horizontal gradient image.

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.

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 meth-

Figure 24: Source text image (*left*) and segmented text image (*right*)

ods, particularly the restricted lexicons. In addition, it bypasses the segmentation phase, which tends to be expensive and error-prone. (M. Elzobi ☎-11065, A. Al-Hamadi ☎-18709)

Currently funded projects

- 3 Projects in SFB/Transregio 62 "A Companion-Technology for Cognitive Technical Systems" (DFG)
  - Environment perception, C1
  - Mechanisms of nonverbal communication, C3
  - Information fusion, C5
- Automatic Pain Recognition based on Facial Expression and Psychobiological Parameters (in cooperation with Section of Medical Psychology at the Ulm University Hospital, DFG)
- Dynamic height measurement for adaptive focused image capturing of large volume objects (in cooperation Chromasens GmbH, ZIM)
- 3-d Ultrasonic measuring method for corrosion inspection in piping system segments (in cooperation with MBQ GmbH, ZIM)
- Advanced radar tracking and classification for enhanced road safety (in cooperation with the Volkswagen AG, EU)
- Driver Assistance Systems (in cooperation with the Volkswagen AG)
- Surface inspection of deformable workpieces (in cooperation with Fraunhofer Institute Magdeburg and INB Vision AG, LSA/EU)
- Advanced methodical Developments for IBR-2M Spectrometers Complex (AMD)
- 3-d Highspeed Particle Tracking using coloured Tracers (in cooperation with the Institute of Fluid Dynamics and Thermodynamics, DFG)
- 2 Projects in the Bernstein Group "Components of Cognition: Small Networks to flexible Rules" (BMBF)
- Engineering and Computational Science for Oncology Network (ECSON, UK Engineering and Physical Sciences Research Council)

Technical Equipment

The Neuro-Information Technology (NIT) group is equipped with optical measuring devices for the acquisition of 3-d 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 á 2.6 GHz
- Spherical camera system
- Stereo camera systems with hardware processing
- Time of flight camera system
- 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 multi-senor system (SFB-TRR62)