CoTeSys White Paper
CA4 - Task Force Nonverbal Communication

Dr. Frank Wallhoff [1], Prof. Dr. Gerhard Rigoll [1], Prof. Dr. Berthold Färber [2]

[1] Human-Machine Communication (MMK), Fakultät für Elektrotechnik und Informationstechnik, TU München, Tel: 089/289-285{52,41}, E-Mail: {wallhoff,rigoll}@tum.de

[2] Institut für Arbeitswissenschaft (IfA), Fakultät für Luft- und Raumfahrttechnik, Universität der Bundeswehr München, Tel.: +49-89-6004-3250, E-Mail: berthold.faerber@unibw.de

Summary— **This white paper reports about the activities, aims and goals of the Task Force Nonverbal Communication, which brings together all CoTeSys researchers that are affected by this topic. By synchronizing relevant activities inside the cluster projects, e.g. emotion and intention recognition, this instrument can bundle synergy effects for the entire cluster with interdisciplinary domain knowledge, which will improve state-of-the-art research and results in innovative and excellent technologies that can be ported to the cluster-wide demonstration scenarios.**

I. Introduction

The Task Force Nonverbal Communication originally started to find synergy effects between a few projects that had the topic *Intention Recognition* in their research portfolio. However, after the first meeting it has been identified, that this topic must not be seen as an isolated one, but has a high overlap with similar topics, for example emotion recognition. Thus, after an additional advice from the Executive Board the initial topic was expanded to all relevant aspects of *Nonverbal Communication*. In fact, a task force with this expanded topic can act more effective by covering interdisciplinary researchers from Neuroscience, Psychology, Engineering, Computer Science, and even Biomechanics and Sports Science.

During the workshop-internal discussions the participants realized that this task force should not result in additional workload, but should be an instrument to bring together different and interdisciplinary researches within the cluster. The task force is further understood to link common problems in different sub-projects and to act as a platform for knowledge exchange. As such, this task force strengthens the CoTeSys community and can provide demonstrator oriented solutions without resulting in an additional bottleneck.

Since most members are in different sub-projects, first a manifold set of interests arose in the beginning. During the discussions it has turned out, that some partners are rather on the implementation side, while others are acting more like a consumer providing the community with detailed problem specifications. All these interests were collected to form a common basis for the future activities of the task force. From an application driven point of view the following cross-activities within the cluster projects were identified:

- Emotion recognition
- Intention recognition
- Emotion synthesis and induction
- Gaze detection

However, the task force NVC shall not to be interpreted as in isolated community. Due to the nature of the above listed topics, input from other task forces, i.e. CA-1 Image-Processing and Understanding, CA-2 Cognitive Architectures, and CA-5 Learning and Planning will be taken into account wherever indicated. By addressing several factory-related projects, output for CA-7 Cognitive Factory will be generated.

II. Addressed Problems

Nonverbal communication is usually understood as the process of communication through sending and receiving wordless messages. Such messages can be communicated through gesture; body language or posture; facial expression and eye contact; object communication such as clothing, hairstyles or even architecture; symbols and infographics; prosodic features of speech such as intonation and stress and other paralinguistic features of speech such as voice quality, emotion and speaking style [1]. Since four applications have been identified it becomes obvious to concentrate just on them, which again can be divided into several subtasks, which are introduced below.

Emotion recognition

In a first step, relevant emotions for the specific task have to be defined with respect to

- a technical feasible way to measure and
- a physiologically reliable way to induce these emotions.

For example in the context of the Cognitive Factory this can be stress or tiredness. Secondly, in an experimental phase the properties of

- emotion induction/elicitation and
- domain and context knowledge

have to be evaluated. For the automated emotion classification, following multidimensional technical measurements have been considered so far:

- Feature/Object identification
- Face recognition
- Speech recognition
- Motion recognition
- Heart rate
- Body temperature

The first experimental setup will consist of a valid measurement procedure also consisting of body mounted sensors. In the second step the available sensors are restricted to re-

mote measurements only. It is planned to provide a common software toolkit for the different sub-project related tasks.
So far several task force members have shown their willingness to take care for the topics Computer Vision, Physiological Experiments, and Pattern Recognition.

Intention recognition

The recognition of a human's intention can be solved in an analogous way as the emotion recognition problem. However, due to the new focus additional tasks and measurements are added:

- Analyzing human interactions by observation
- Gait recognition
- Face and Eye Gaze detection

As result of first experiments gaze detection is an important factor to predict human's intentions. Thus the input of the Gaze Detection partners in the task force is urgently needed.

Emotion synthesis and induction

Besides the recognition of emotions, it is further planned to synthesize emotions with the purpose of equipping our desired cognitive technical system demonstrators with the ability to actively and naturally provide the human partner with feedback. This can be done over the visual or acoustical cue. Further specifications will follow in the future.

Gaze Detection

Due to its complexity, the fourth goal of the Nonverbal Communication Task Force will be the eye gaze detection with three lines of activities in CoTeSys:

1) The first line aims at equipping both the Cognitive Factory robots and the humanoids with the ability to actively direct their camera-based vision systems to task-relevant locations in space. This will help the human cooperator to make conclusions about robot actions that are to be expected in the immediate future. The related CoTeSys activities are mainly located in the current projects #106 and #125 and will possibly be extended in the new projects #315, #316, and #334.
2) The second line of activities aims at equipping all CoTeSys robot demonstrators with the ability to perceive the direction of gaze of human cooperators. This ability is mandatory if robots are intended to seamlessly (co-) operate in human-machine interaction tasks. By analyzing task-relevant targets of gaze a robot will then be able to make conclusions about the intention of humans. This is the main goal of project #121.
3) The third line of activities aims at providing standardized but unprecedented gaze detection capabilities to those CoTeSys projects that conduct fundamental studies on human gaze behavior. These activities are mainly located in projects #106 and #121 and the results of these have already been requested and used in projects #147, #159, and #127. One of the major goals of project #106 is to combine a low-latency eye tracker with pivotable head cameras that can be oriented to the user's target of gaze. It is also planed to put the developed camera orientation devices on JOHNNIE's head and to drive his eyes (see the first line of activities).

III. Associated Projects

With currently nine nominated projects the NVC group covers nearly 20 percent of the overall volume. In the future the relations within this task force will further help to cluster projects and merge synergies in the subsequent funding period. This section lists the CoTeSys projects associated with the task force NVC:

- *Action recognition from look-ahead fixations of objects in space with self-calibrating gaze control of head-mounted stereo cameras (#106).* The expertise of project #106 in eye movements and gaze detection technology has been communicated to the participants of the Non-verbal Communication Task Force by way of an overview talk on the occasion of a Workshop and by a Wiki-based review of the relevant literature. This coordinated effort helped to trigger new cooperations with projects #147, #159, and #127.
- *Eyetracker-Based Driver State and Intention Recognition for Realtime Control and Configuration of Human-Machine Interfaces in Vehicles (#121).* Gaze has been identified as a central modality of non-verbal communication in human-human as well as in human-machine interaction scenarios. Gaze direction is an experimentally accessible indicator of future action since it correlates with spatial attention and intention.
- *Recognizing and modeling the action intention of humans (#123).* This project deals with the automatic recognition and interpretation of human action intentions. Since we started off with the traffic context and the recognition of pedestrian intentions we concentrate on the nonverbal communication from a distance (the only form of communication there is between driver and pedestrian). So far we worked on image processing from an infrared camera (that could also be made available to other projects if needed) to reliably detect humans and distinguish them from other objects. In an experiment first parameters were identified that help humans to perceive the intention of others. This experimental approach could also be transferred to other scenarios for intention recognition. The experiments showed the importance of movements of the human as well as his visual behavior for the intention detection. Thus cooperation with #106 and #121 is obvious. We are now working on detecting the identified parameters by technical means and on finding the critical values for those parameters.
  In cooperation with the other partners of NVC we will provide a platform for intention recognition in different application areas.
- *Emotional Control of Human - Cognitive System Interaction (#126).* Project #126 is concerned with emotional aspects of human-robot interaction, which are a crucial component of natural and efficient communication processes even between man and machine (see e.g. Picard, 2000 [2]; Breazeal et al., 2005 [3]). One main focus is the development of an emotion recognition system based on physiological data (ECG, SC, BVP, EMG). Contrary to common emotion recognition techniques by face or speech analysis, these involuntary and continuous signals allow for emotion detection even in situations without spoken words or in case of non-extreme emotions which are

typical for working scenarios (e.g. frustration, overchallenge, and irritation). Another focus of this project is the display of emotional expressions by the robotic head EDDIE (Sosnowski et al., 2006 [4]). The integration of a combined emotion recognition and display system into a humanoid robot will in the long run facilitate a systematic control of the emotional state of the human cooperation partner for a more natural and efficient human-machine interaction and manufacturing process.
Beyond doubt, the ability to detect and react to emotions occurring during HMI is an extremely important aspect of nonverbal communication. Besides the very valuable exchange of general methodological knowledge and experiences, we profit from other projects' support concerning methods of image processing that we might use in the future for facial expression and gesture analysis, or different emotion induction procedures, or experimental settings, as well as from the easy access to an appropriate demonstrator that is capable of executing our planned experiment.

- *ACIPE - Adaptive Cognitive Interaction in Production Environments (#159)*. The project ACIPE concentrates on the integration of cognitive systems in manual assembly environments by creating an assistive system which helps the human worker to optimize his assembly task. Several models have been defined during the first project year. One of them is the cognitive human model, which represents the worker with his mental capabilities and mental state. The assistive system described above has to be adapted to the current situation of the worker and the workpiece. One major challenge is to present the appropriate information to the worker in the most optimal way, minimizing distraction and transcoding effort. Therefore, we intent to prepare the information matched to the worker's skills and his mental state in the current working situation. Additionally the information has to be presented at the right time, which means, we need to know where the worker's focus is and what his intentions are. In order to generate cognitive human models that actually correspond to real human cognitive processes, results from RA-A are essential and the base for a successful estimation of the worker's mental state.
- *JAHIR - Joint Action for Humans and Industrial Robots (#207)*. This project aims to integrate industrial robots in a human-dominated working area, so that humans and robots can cooperate on a true peer level regarding safety aspects. Therefore JAHIR will focus on the information gained from the intuitive non-verbal communication channels based on advanced computer vision algorithms including model-based visual tracking. Within the project a representative assembly cell will be created where industrial robots, dedicated robot platforms and human workers work jointly on complex assembly tasks. In this way, the project provides an ideal experimental stage for an integrated exploration of research results from research Areas A-E, with focus on D and E. In future phases of the cluster, this project will also serve as an integration platform for embedding humanoid action in the cognitive factory.
- *MOHMIP - Mental Model based Human-Machine-Interaction in Production Processes (#194)*. MOHMIP covers the analysis, modeling and implementation of human cognitive processes in the context of the cognitive factory. The goal is to enable a collaboration of human and machine resembling human-human interaction. This is achieved by the formalization of human mental models and cognitive processes using the cognitive architecture ACT-R. The resulting models will then be implemented in technical systems allowing for intelligent cognitive assistants as well as an adaptive allocation of work among workers and machines.
- *Recognition of states and traits by whole-body movement analysis (#151)*. Goal of this project is the kinematic and kinetic analysis of the human gait under seven different induced emotions. Goal of the project is the development of algorithms which will allow the humanoid robot, vehicle, and factory demonstrators to estimate precisely the emotion and personality of an approaching human. Additionally, to that methodological fundaments will be investigated to allow humanoid robots to walk in an emotionally expressive way in a consecutive project planned for a later stage in CoTeSys. According to this, in a first step every trajectory of all main joints during gait will be tracked while different emotions are induced (exploration phase). In a second step emotionally induced gait patterns will be distinguished on the basis of dynamical classification methods (analysis phase). The found algorithms will be tested at a new sample in a third step (confirmation phase). At the end the emotionally depended gait patterns will be tested in a simulation study. This study will explore if human observers also classify the extracted gait patterns in the same emotional classes.
- *Cognitive Factory and Humanoid Robots: enabling multi robot-human joint action through multimodal cue fusion (#147)*. The main focus of the project is on the research on learning by demonstration in a scenario of n robots and m humans. Together they should accomplish a complex task which can typically be found in a manufacturing domain. In order to standardize specifications, the assembly of building brick structures was chosen as manufacturing task. Despite this simplification, the complexity of the task still requires a group of at least two robots working together with a human expert. The principal role of the human within the group is to plan the task, assign subtasks to other group members and to demonstrate previously unknown procedures. In this context the need for communication within the group arises. The first step towards this development is the observation of communication about a task within a team of humans. Beside linguistic formulations, this will also comprise modalities like eye movement as an indicator for attentiveness and pointing gestures which are able to affect the attentiveness of an apprentice. Experiences made with the human team should be transferred step by step on a mixed team of humans and robots.

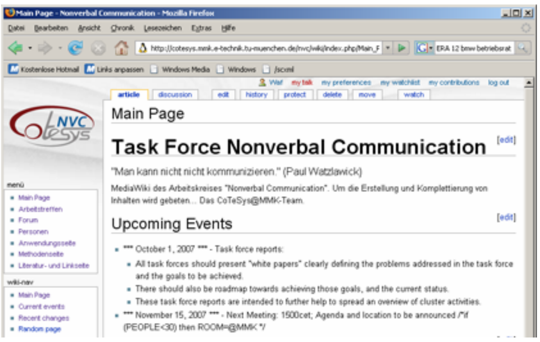
IV. Information Sharing

Providing state-of-the-art information exchange instruments, two common tools have been set up on a server at the Human-Machine Communication institute:

- A cluster-internal mediaWiKi:
  http://cotesys.mmk.e-technik.tu-muenchen.de/nvc/wiki

- A discussion group: http://cotesys.mmk.e-technik.tu-muenchen.de/forum/

To restrict unwanted access from outside, both platforms are protected by a username and password: *nvc*, *cotesys*. To add wiki content and to place or answer questions in the forum, each CoTeSys member is invited to create his own account. In order to preserve the intellectual rights of the authors that offer information all users are strongly asked to respect the common copyright rules.



V. Meetings and Activities

During the first year three meetings have taken place until now. As mentioned before the central idea is to have periodical meetings with a distance of approximately 2 months.

Exploratory Meeting

The first meeting took place on June 9th 2007 at *iwb* in Garching and was a first attempt to identify common interests between the participants' projects, i.e. ACIPE (#159), JAHIR(#207) and Intention Recognition (#123). The discussion showed that the topic intention recognition is a common question throughout these projects. The participants were Alois Knoll (i6, TUM), Berthold Färber (ifa, UniBW), Sabrina Schmidt (ifa, unibw), Ana Perez Grassi (VMS, TUM), Hendrik Neumann (IFA, UniBW), Marc Halbrüge (ifa, UniBW), Christian Stössel (NCP, LMU), Florian Engstler (LFE, TUM), Florian Friesdorf (LFE, TUM), Claus Lenz (i6, TUM), Manuel Giuliani (i6, TUM), Markus Rickert (i6, TUM), Alexander Bannat (MMK, TUM), Jürgen Gast (MMK, TUM), Frank Wallhoff (MMK, TUM), Wolfgang Vogl (iwb, TUM).

1st Nonverbal Communication Worksop

After the recommendation of the CoTeSys Executive Board the original goal intention recognition was expanded to the more general topic nonverbal communication aiming at covering more projects within the cluster. Thus this first get together at the institute MMK on July 12th 2007 concentrated on getting in touch with each other by presenting the treated projects. The participants: Michael Popp (IFA, UniBW), Cornelia Wendt (IFA, UniBW), Sabrina Schmidt (IFA, UniBW), Ana Perez Grassi (VMS, TUM), Anna Schubö (NCP, LMU), Christian Stössel (NCP, LMU), Bernd Spanfelner (SSE, TUM), Christian Leuxner (SSE, TUM), Fernando Puente Leon (VMS, TUM), Frank Wallhoff (MMK, TUM), Jürgen Gast (MMK, TUM), Alexander Bannat (MMK, TUM), Gerhard Rigoll (MMK, TUM), Dirk Wollherr (LSR, TUM), Kolja Kühnlenz (LSR, TUM), Wolfgang Rösel (IWB, TUM), Claus Lenz (Info VI, TUM), Manuel Giuliani (Info VI, TUM), Thomas Villgrattner (AM, TUM), Erich Schneider (nefo, LMU), Stefan Kohlbecher (nefo, LMU), Stanislaus Bardins (nefo, LMU), Aleksandra Kupferberg (nefo, LMU), Stefan Glasauer (nefo, TUM).

2nd Nonverbal Communication Worksop

In the second workshop the common interests were in the focus of the presentations. Therefore three of the main topics were on the agenda: Physiological Emotion Recognition (C. Wendt), Object Detection (G. Panin) and Eyetracking (E. Schneider). Each topic was presented and questions have been extensively discussed afterwards. Following participants were at the workshop at MMK: Zahid Riaz (i9, TUM), Jan Bandouch (i9, TUM), Florian Friesdorf (LfE, TUM), Stanislaus Bardins (nefo, LMU), Matthias Rungger (LSR, TUM), Heike Valery (LSR, TUM), Giorgio Panin (i6, TUM), Erich Schneider (nefo, LMU), Claus Lenz (i6, TUM), Wolfgang Rösel (iwb, TUM); Susanne Ihsen (lte, TUM), Kolja Kühnlenz (LSR, TUM), Michelle Karg (LSR, TUM), Cornelia Wendt (IFA, UniBW), Ana Perez Grassi (VMS, TUM), Sabrina Schmidt (IFA, UniBW), Manuel Giuliani (i6, TUM), Markus Huber (nefo, LMU), Thomas Villgrattner (AM, TUM), Klaus Bartl (nefo, TUM), Stefan Kohlbecher (nefo, LMU), Ferdinand Tusker (sport, TUM), Wolfgang Seiberl (sport, TUM), Christian Stößel (NCP, LMU), Anna Schubö (NCP, LMU), Stefan Schwärzler (MMK, TUM), Jürgen Gast (MMK, TUM), Michael Popp (LRTM, UniBW), Berthold Färber (IFA, UniBW), Frank Wallhoff (MMK, TUM).

VI. Roadmap

To continue reaching the above mentioned goals, further bimonthly meetings will be arranged. The next meeting is scheduled for November 15th 1500-1800CET. Without actual deadline constraints following work packages are in the scope of the discussion:

- Emotion Recognition Software Module
- Intention Recognition Software Module
- Gaze Detection Module
- Probably Satellite workshop with major conference

Further input is expected after the task force presentation on October 11th and 12th 2007 at the Grande CoTeSys Meeting.

VII. References

[1] http://en.wikipedia.org/wiki/Nonverbal_communication

[2] Picard, R. W. (2000). Affective computing. Cambridge: MIT Press.

[3] Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. October 9-15, Beijing, China.

[4] Sosnowski, S., Bittermann, A., Kühnlenz, K., & Buss, M. (2006). Design and evaluation of emotion-display EDDIE. Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. October 9-15, Beijing, China.