PROCEEDINGS

55. IWK

Internationales Wissenschaftliches Kolloquium
International Scientific Colloquium

13 - 17 September 2010

Crossing Borders within the ABC

Automation,

Biomedical Engineering and

Computer Science

Faculty of
Computer Science and Automation

www.tu-ilmenau.de

TECHNISCHE UNIVERSITÄT
ILMENAU

Home / Index:
http://www.db-thueringen.de/servlets/DocumentServlet?id=16739

Impressum
Published by

Publisher: Rector of the Ilmenau University of Technology
Univ.-Prof. Dr. rer. nat. habil. Dr. h. c. Prof. h. c. Peter Scharff

Editor: Marketing Department (Phone: +49 3677 69-2520)
Andrea Schneider (conferences@tu-ilmenau.de)

Faculty of Computer Science and Automation
(Phone: +49 3677 69-2860)
Univ.-Prof. Dr.-Ing. habil. Jens Haueisen

Editorial Deadline: 20. August 2010

Implementation: Ilmenau University of Technology
Felix Böckelmann
Philipp Schmidt

USB-Flash-Version.

Publishing House: Verlag ISLE, Betriebsstätte des ISLE e.V.
Werner-von-Siemens-Str. 16
98693 Ilmenau

Production: CDA Datenträger Albrechts GmbH, 98529 Suhl/Albrechts

Order trough: Marketing Department (+49 3677 69-2520)
Andrea Schneider (conferences@tu-ilmenau.de)

ISBN: 978-3-938843-53-6 (USB-Flash Version)

Online-Version:

Publisher: Universitätsbibliothek Ilmenau
ilmedia
Postfach 10 05 65
98684 Ilmenau

Home / Index:
http://www.db-thueringen.de/servlets/DocumentServlet?id=16739

INTELLIGENT BEHAVIOUR OF HUMANOID ASSISTIVE ROBOTS

Peter Nauth

Fachhochschule Frankfurt a.M.–University of Applied Sciences, 60318 Frankfurt, Germany

ABSTRACT

Robots operating under not standardized conditions need to behave intelligently. They must understand the goals a user wants to be achieved, recognize the environment in context with the goals, develop strategies to execute them and act autonomously. By means of intelligent sensors for speech recognition, proximity measurement, color measurement and image processing an intelligent humanoid robot has been developed. It understands goals by recognizing the name of an object a user has told the robot to take and searches for it by means of a smart camera and other sensors. After it has found and identified the object, it grabs it and brings it to the user. The speech recognition sensor and the smart camera can learn new words as well as shapes and colors of new objects in order to cope with new situations.

***Index Terms* -** Robot Control, Humanoid Robots, Goal Achievement, Goal Understanding, Image Processing, Speech Recognition, Sensor Fusion, Intelligent Behaviour

1. INTRODUCTION

Intelligent behaviour enables robots such as assistive or rescue robots to solve complex tasks which up to now only human beings can handle. They communicate with those people they are supposed to support in an intuitive way and act autonomously in nonrestrictive environments and changing boundary conditions. Key technologies for intelligent behaviour are Embedded Intelligent Systems [1] which analyze and fuse comprehensive sensor data and derive execution strategies in order to accomplish a goal.

In the Laboratory for Autonomous Systems und Intelligent Sensors at the Fachhochschule Frankfurt a.M., Germany, first a stationary intelligent robot with visual and auditive sensors has been developed. It understands spoken instructions and acts accordingly by differentiating objects with its camera, by grabbing the demanded object and giving it to a user [2]. This technology has been transferred to an autonomous humanoid robot.

Algorithms for robot control and navigation have been developed by several research groups [3]. This paper focuses on the understanding of goals, strategies to achieve these goals, learning methods, on sensing and navigating in a natural environment as well as on handling objects with respect to intelligent humanoid robots.

Another focus is on applying the algorithms to small robots. The advantages of small sized robots over other systems [4] are reasonable deployment costs and scalability. However, small robots alone cannot carry heavy or big objects or reach these lying in higher levels. A solution to this problem is the swarm robot approach where several robots co-operate as a team in order to solve a heavy task together.

In order to cope with situations where the robots are confronted with conflicting requirements such as a goal to be immediately achieved in the presence of a low battery status and the need to avoid a dangerous area, an approach for a self-generating will is introduced. The purpose is to equip the robots with a kind of cognitive intelligence which allows them to learn new behaviours by experiences.

2. SYSTEM ARCHITECTURE

An autonomous robot needs to know the goal to be accomplished, situation awareness and the ability to plan and perform actions depending on a situation. This requires the following functions [5]:

- Sensing by means of multiple sensors in order to acquire all necessary data about the environment. It includes getting to know the goal to be met, e.g. by understanding a spoken instruction.
- Fusion of the data acquired from intelligent sensors in order to assess the situation.
- Planning how to achieve the goal and
- Execution of the necessary steps by controlling the robot motors.

A distributed data base provides reference information e.g. for the pattern recognition algorithms in intelligent sensors, for strategies to fuse data or for setting an optimal execution plan. It is important that the data base can be adapted to new situations by methods such as learning algorithms. A robot embedding the features described above can be regarded as behaving intelligently because it can perform tasks depending on a goal to be met in a complex environment and can adapt to new situations by learning new words and objects. However, a higher level of intelligence can be achieved by robots with a self-generating will because they can develop new behaviours by themself depending on experiences they make.

The ability of a robot to recognize a situation where it needs the help of others is a feature of intelligent behaviour as well.

We use the Robonova 1 (Fig. 1) for humanoid robot development. Sensor fusion, the control of the actuators and the coordination of all components are executed on a 8-Bit Atmel ATMega 128L.

Fig. 1. Autonomous Humanoid Robot

Depending on the area of activity the robot is equipped with some or all of the following sensors (Fig. 2):

- Speech Recognition Sensor
- Proximity Sensor
- Smart Camera.

The signal processing for speech recognition is implemented on the DSP-based module Voice Direct 364 (Sensory Inc.). The image processing algorithms run on the smart color camera POB-Eye (POB Technologies) with an embedded 32-Bit controller ARM7TDMI. It is mounted on the neck of the humanoid robot and can be turned left and right as well as up and down. This enables the robot to look in various directions.

In order to cope with new situations the speech sensor and the vision sensor can learn new words as well as shapes and colors of new objects respectively [6]. The method used is supervised learning.

Two different proximity sensors (laser beam triangulation method) are used in order to measure the distance to an object in the near range from 4 to 24 cm and the far range from 15 to 80 cm. They are mounted on the moveable head of the robot in order to scan the environment in 2 dimensions.

Additionally, the proximity sensors are used to acquire distance functions of the objects which contain further information about the object's shapes supplementing the camera data.

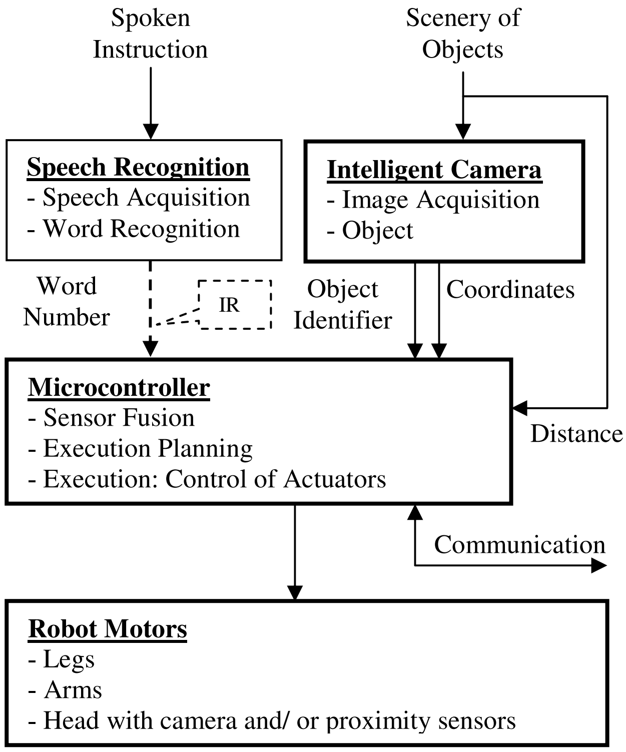

Fig. 2. System Architecture: Thick lines indicate parts mounted at the robot

3. SPEECH RECOGNITION

The intelligent speech recognition enables the robot to understand spoken instructions. These are either single words or a sequence of words which are spoken without breaks in between. After data acquisition, the algorithm divides the signal into segments and calculates the frequency spectra out of each segment. Next, frequency parameters are calculated and classified by means of a neural network.

As for the training phase, the user speaks a word and repeats it. If the frequency parameters from the first and the repeated word are similar, the word is accepted. The weighting factors of the classifier are being adapted to the word's frequency parameters and assigned to a word number. Then the training can be continued with the next word.

During the recognition phase the speech sensor asks by saying "Say a word" the user to speak the instruction. If the classifier detects a high similarity to one of the previously learned words, it sets a respective pin to High. An additional controller SAB 80535 turns the spoken word into a word number by monitoring the state of the pins and transmitting a bit sequence via an infrared (IR) LED to the robot. The sequence corresponds to the pin number set to high and therefore to the word number recognized by the speech module. This enables the user to command the robot remotely.

The robot controller receives the bit sequence via an infrared detector and decodes the word number. For each word number is assigned to an instruction, the robots now knows its goal, i.e. which object to search for.

4. OBJECT RECOGNITION

For the recognition of the demanded object and for obstacle avoidance during the search phase the intelligent camera and the proximity sensors are used.

The algorithms we have developed for the smart camera converts the acquired RGB – image into the HSL – space, segments the image [7] by means of an adaptive threshold algorithm (histogram analysis of hue) and extracts the form factor F

$$F = U^2 / A$$

from the area A and the circumference U as well as the mean hue – value H from each object detected. By means of a box classifier [8] each object is assigned to an object identifier which represents the class (Fig. 3). Given the extracted parameters, objects and obstacles can be differentiated regarding shape and color.

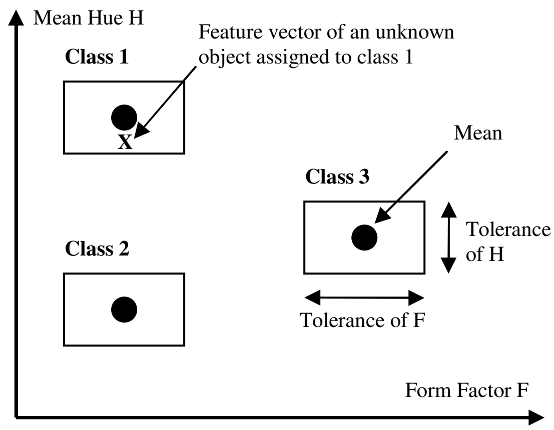

Fig. 3. Box Classifier: Mean values und tolerances result from the learning phase

Additionally, the coordinates of each object are calculated. The object identifier and the respective coordinates of all objects found are transmitted to the robots microcontroller.

New objects can be learned by a supervised learning algorithm: Typical examples of each object class are shown to the camera and the learning algorithm assigns the mean values of each parameter to the class these objects belong to. The tolerances which define the size of the classification box of each class equal 1.5 times the standard deviation calculated during the teach-in procedure.

Proximity sensors supplement the camera information by sensing the distance between the robot and the object. Additionally, proximity sensors can be used to provide shape parameters for object differentiation themselves, especially for obstacle detection [9]:

Given that the proximity sensors scan the environment in two dimensions, the distance z (α,ß) is a function of the horizontal angle α and the vertical angle ß. Appropriate parameters can be selected by modelling the distance function for typical object shapes. E.g., the distance function z (α,ß) of a round object with the radius r positioned in a lateral distance d can be modelled as

$$z(\alpha,\beta) = \frac{(d+r)\cos\alpha - \sqrt{(d+r)^2(\cos^2\alpha - 1) + r^2}}{\cos\beta}$$

whereas a wall has the distance function z (α,ß)

$$z(\alpha,\beta) = \frac{d}{\cos\alpha\cos\beta}$$

We use these equations to differentiate walls from objects for all the objects we use have round shapes.

At the edges of the steps of stairs the distance changes rapidly. Therefore, by turning the proximity sensors in the vertical direction ß, stairs can be differentiated from objects and walls if unsteady parts show up in the distance function z (α,ß).

5. SENSOR FUSION, PLANNING AND CONTROL

By fusing the auditive, visual and proximity data the robot knows all objects within its reach and their position as well as the goal it is advised to attain.

The fusion algorithm used is hierarchical and works as follows:

1. Auditive and visual data are fused by matching the word number (derived from the speech sensor data) with one of the object identifiers (derived from the camera data) by means of a table. The algorithm generates one of the following hypothesis:

- A negative match result (i.e. no object or the wrong object) leads to the hypothesis "object not found". This causes the robot to repeat the search.

- A match of one of the object identifiers with the word number results in the hypothesis “object found”.
- The hypothesis “wall” or “stairs” is derived if one of those obstacles has been classified regardless the spoken command.

2. The robot moves towards the object or obstacle until it is within the scanning range of the proximity sensor. In the next fusion step the hypothesis generated by the visual sensor is verified by the data acquired from the proximity sensor. If the class derived from the distance function z (α,ß) equals the hypothesis it is accepted. Otherwise the hypothesis is rejected.

This hierarchical approach results in a high specificity and a low sensitivity because in case of conflicting visual and proximity results they are rejected. We overcome this problem by repeating the search in this case. The robot moves to a different position before starting the renewed search as described below.

Currently we apply the fusion approach to differentiate 3 kinds of objects, a water bottle and bottles of 2 different kinds of soft drinks, as well as 2 different kinds of obstacles (wall and stairs).

In order to execute the appropriate steps for goal achievement the robot has to plan the next actions necessary to meet the goal. First, the robot assesses its state $s_j(\underline{x})$ which is a function of the sensory input $\underline{x}$. Next, the action

$$a_i\left(s_j(\underline{x}), g, a_{i-k}, \underline{x}\right)$$

is derived depending on the actual state $s_j(\underline{x})$, the goal g to be met, previous actions a_{i-k} and the sensory input $\underline{x}$. After having executed the action a_i the robot reaches the next state $s_{j+1}(\underline{x})$.

E.g. depending on the state being reached after the action “sensor fusion” has been performed the robot develops different execution plans and controls the robot motors accordingly:

- If the demanded object has been identified, the robot approaches it and grabs it in order to bring it to the user. During the movement towards the object its position relatively to the robot is tracked permanently.

- If no or the wrong object has been spotted or in case of conflicting results, the robot repeats the search by turning the camera head and the proximity sensors or by moving around in order to change the position.

- If an obstacle has been detected the robot develops an approach to overcome it, i.e. it climbs stairs or avoids colliding with walls.

The robot grabs the objects by pressing its arms from left and right at them. It stops the arm movements if a feedback signal indicates a resistance.

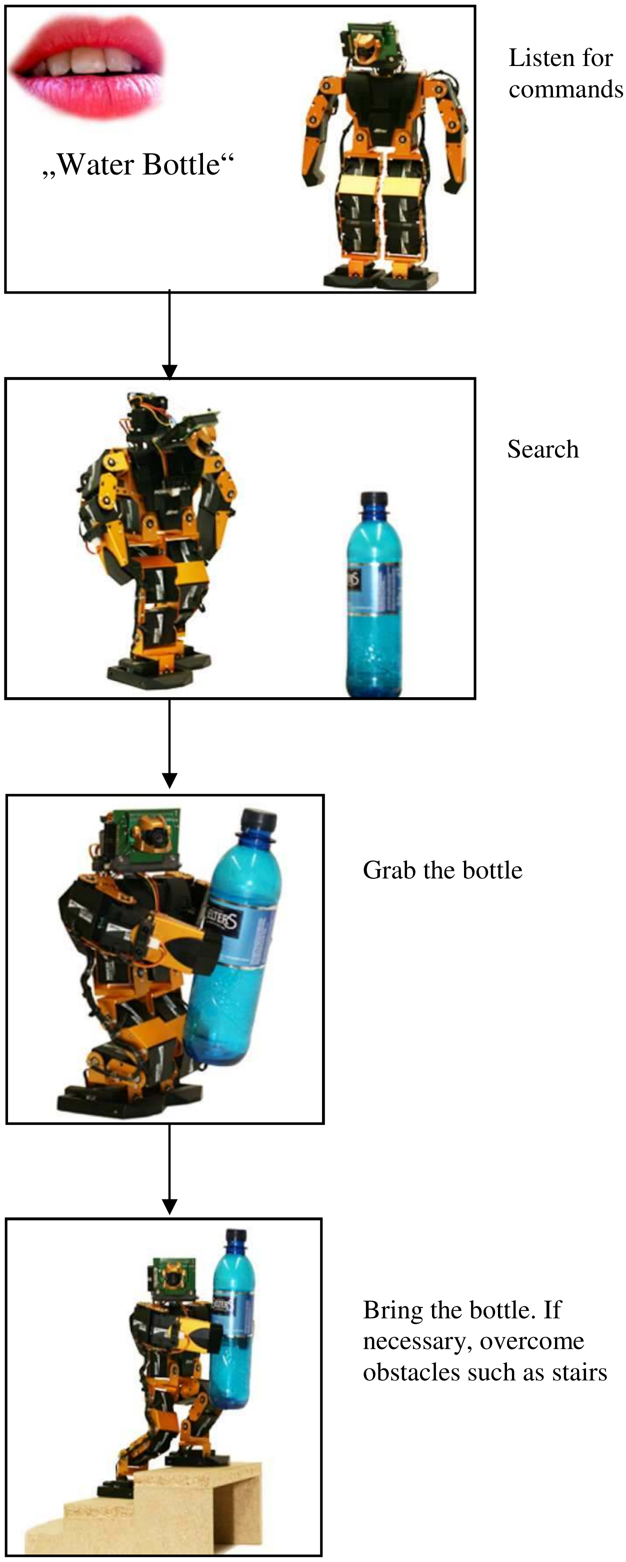

Fig. 4. Object Search and Fetch

6. APPLICATION SCENARIOS AND RESULTS

One typical example of the robot's performance is to search for objects and bring it to the user (Fig. 4). If the user says "Water Bottle", the robot understands its task and searches for the bottle. After detection, it grabs the bottle and brings it to the user. Stairs can be climbed by a co-ordinated arm and leg movement and the stabilization of the robot with sensors.

In the scenario mentioned above (1 water and 2 soft drink bottles, a wall and stairs) the right bottle has not been found in 2 out of 20 cases. In 6 cases the robot had to repeat the search at least once due to a mismatch or conflicting data during the fusion process. In 1 case stairs have been wrongly classified as a wall. The number of search repetitions was limited in all 20 cases to 10.

7. CO-OPERATIVE ROBOTS

If the robot sensors (feed-back signal of motor positions, tilt sensor) indicate that an object is too heavy to grab or to maintain a stable position, it is set to the state "need help". This triggers the communication with other robots via Bluetooth to come and assist in order to grab and carry the object together (Fig. 5) as swarm robots [10-11].

Fig. 5. Swarm Robots grabbing a bottle together

If a robot has found an object it is not instructed to grab, it informs other robots about the object type and position. A robot which has received the goal to fetch this particular object can go directly to it without the need for search.

Soccer robots which search for the ball and kick it to the goal or another robot of their team can benefit from the swarm robot approach, as well. A communication between the robots allows to co-ordinate their actions.

As for rescue robots, we focus on the scenario that an injured person (in this case a doll) lies on a stretcher. Two robots evacuate the injured person by jointly carrying the stretcher into a safe area. The algorithms we implemented synchronize their movements and exchange information about the direction to go. Additionally, the robot at the front end warns the other one about obstacles.

8. EXPERIENCE DRIVEN BEHAVIOUR

The robot we have developed obeys is intelligent in a sense that it can adapt by learning new object classes and words. However, it cannot cope with conflicting requirements and cannot use behavioural experiences it has made previously. In order to enable the robot to behave intelligently [12-13] when boundary conditions have changed and requirements contradict, we propose to equip the robot with a self-generating will based on the algorithms we use so far.

The limitation of the robot presented above is the dependence of the state $s_j(\underline{x})$ on the task-related sensory vector $\underline{x}$ derived from the visual and proximity data, only. Adding a sensory vector $\underline{d}$ for acquiring environmental data such as temperature as well as for drives such as hunger (low battery status) and the desire for praise by the user for having achieved a goal and "social contact", the robot can gather information about its well-being or dangerousness of situations. Combining $\underline{x}$ and $\underline{d}$ we get the sensory vector $\underline{y}$

$$\underline{y} = \left(\underline{x}^T, \underline{d}^T\right)^T$$

In order to implement a sort of feelings with respect to a given state the sensory data $\underline{y}$ are weighted by a weigthting vector $\underline{w}$ which has been trained by the experiences made during tasks have been performed previously. Hence, the states

$$s_j\left(\underline{y}, \underline{w}\right)$$

depend on the sensory input $\underline{y}$ and their "assessment" represented by $\underline{w}$. Experiences are made by recognizing e.g. that an action has led to an increasing battery voltage (positive), a word of acknowledgement spoken by the user (positive), falling down (negative) or a too high temperature (negative). This requires that some of the sensors such as battery voltage are directly linked to the classes of experience "positive" or "negative".

In order to describe the "feelings" about the current state, a quality of state function [14]

$$Q\left(s_j\left(\underline{y}, \underline{w}\right)\right) = \underline{w}^T \underline{y}$$

is introduced. A high quality value corresponds to well being. This will trigger the robot to continue the actions necessary to achieve its goal. A low quality value means the robot is "afraid". It would go back to the previous state and perform an alternative action in order to achieve the goal. In case no alternative action is available the robot would return to its (safe) start position. A more sophisticated approach would avoid going back and decide about the appropriate action in order to reach a state without fear and to continue the goal achievement. It requires to predict the quality of all the states which can be reached from the current state $s_j(\underline{y}, \underline{w})$ after the actions a_i have been executed. The prediction is calculated by a modified quality of state function

$$Q(s_j(\underline{y}, \underline{w}), a_i)$$

Hence, the generation of the robot's will to perform an action a_w in order to achieve the goal g

$$a_w(s_j(\underline{y}, \underline{w}), g, a_{i-k}, \underline{y})$$

is not rule based but the result of optimizing the quality of state function

$$Max\{Q(s_j(\underline{y}, \underline{w}), a_i)\} \rightarrow a_w$$

This operator predicts the quality of all states $s_j(\underline{y}, \underline{w})$ which can be attained from the current state with respect to the possible actions a_i that can be executed from the current state and selects the maximum. As a result the robot generates the will to perform an action a_w which maximizes the quality of state. The algorithm is currently implemented at our humanoid robot. We expect the robot to show some kind of independence and refuse tasks it has made bad experiences with, e.g. if it expects that the state "idle" will have a higher quality than the state "goal achieved", because the object to be grabbed has caused the robot to fall down in the past. This negative feeling might dominate over the expected praise of the user for delivering the object. Or it might seek co-operations more frequently if previous team work had been successful.

9. SUMMARY

A humanoid robot has been developed which behaves intelligently. It understands spoken commands and can act accordingly. If the user advises the robot to bring a specific object, the robot searches for the object by means of its smart camera and other sensors. After the robot has identified it, it grabs it and brings it to the user. In case the robot needs support, it can call other robots for assistance, e.g. in order to carry heavy objects together as swarm robots. The small size of the robots allow reasonable deployment costs. Furthermore an algorithm for implementing a self-generating will has been proposed.

10. ACKNOWLEDGMENT

The research was funded by the Fachhochschule Frankfurt a.M. and the Gesellschaft für technische Weiterbildung e.V.

11. REFERENCES

[1] Peter Nauth, Embedded Intelligent Systems, Oldenbourg Verlag , 2005

[2] Ella-Roxana Ivascu, Roboteransteuerung mit einem 32-Bit Mikrocontroller, Diplomarbeit, Polytechnische Universität Bukarest , 2007

[3] Tae-Seok Jin, Bong-Kee Lee and Jang Myung Lee, „AGV Navigation Using a Space and Time Sensor Fusion of an Active Camera", International Journal of Navigation and Port Research, Vol 27, No 3, 2003

[4] K. Hirai, M. Hirose, Y. Haikawa, T. Takenaka, The Development of Honda Humanoid Robot, IEEE Int. Conf. Robot. Autom.,pp.1321-1326, 1998

[5] Kristen Stubbs and David Wettergreen, "Anatomy and Common Ground in Human-Robot Interaction: A Field Study", IEEE Intellig. Systems, March 2007

[6] Christopher M. Bishop, Pattern Recognition and Machine Learning, Springer Verlag, 2008

[7] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Prentice Hall, 2008

[8] Desiree Toth, Objekterkennung von humanoiden Robotern mittels visueller Sensorik, Diplomarbeit, Fachhochschule Frankfurt a.M., 2009

[9] Azzedine El Haddaoui and Rateb Hamidi, Intelligente Sensorik zur Steuerung eines humanoiden Roboters, Diplomarbeit, Fachhochschule Frankfurt a.M., 2008

[10] V. Trianni, S. Nolfi, M. Dorigo, "Cooperative hole-avoidance in a swarm-bot", Robot. Autom. Syst. 54. 2006

[11] Bruno Siciliano and Oussama Khatib, Handbook of Robotics, Springer Verlag, 2008

[12] D. Dietrich, D. Bruckner, B. Müller, G. Zucker, A.Tmej, Psychoanalytical Model for Automation and Robotics, IEEE Africon, 2009

[13] G. Doeben-Henisch, Humanlike Computational Learning Theory. A Computational Semiotics Perspective, IEEE Africon, 2009

[14] Ron Sun, "Cognitive Social Simulation Incorporating Cognitive Architectures", IEEE Intelligent Systems, Sept 2007