in: Proc. IEEE/RJS Int. Conf. on Intelligent Robots and Systems (IROS 2011), San Francisco, USA, pp. 2430-2437, IEEE Omnipress 2011

Progress in Developing a Socially Assistive Mobile Home Robot Companion for the Elderly with Mild Cognitive Impairment

H.-M. Gross, Ch. Schroeter, S. Mueller, M. Volkhardt, E. Einhorn, A. Bley, Ch. Martin, T. Langner, M. Merten

***Abstract*— The paper is addressing several aspects of our work as part of the European FP7 project "CompanionAble" and gives an overview of the progress in developing a socially assistive home robot companion for elderly people with mild cognitive impairment (MCI) living alone at home. The spectrum of required assistive functionalities and services that have been specified by the different end-user target groups of such a robot companion (the elderly, relatives, caregivers) is manifold. It reaches from situation-specific, intelligent reminding (e.g. taking medication or drinking) and cognitive stimulation, via mobile videophony with relatives or caregivers, up to the autonomous detection of dangerous situations, like falls, and their evaluation by authorized persons via mobile telepresence. From the beginning, our approach has been focused on long-term and everyday suitability and low-cost producibility as important prerequisites for the marketability of the robot companion. Against this background, the paper presents the main system requirements derived from user studies, the consequences for the hardware design and functionality of the robot companion, its system architecture, a key technology for HRI in home environments - the autonomous user tracking and searching, up to the results of already conducted and ongoing functionality tests and upcoming user studies.**

I. INTRODUCTION

The European FP7 project CompanionAble [1] running from 2008-2012 addresses the issues of social inclusion and home assistance of elderly people suffering from mild cognitive impairment (MCI) and living alone at home. The objective of the project is to allow those people to remain in their accustomed home environment for as long as possible by using assistive technologies - smart home technologies as well as a socially assistive, mobile robot companion. The fact that a mobile robot is able to navigate autonomously or, if required, by means of tele-operation in the elderly's home has proved as an convincing argument in accepting such technologies by the different end-user target groups (the elderly, their relatives, and caregivers). For them it is well comprehensible that mobility of an assistive system significantly simplifies the interaction with the care recipient and makes a spectrum of assistive functions possible at all, as for example greeting the user upon his/her returning home, facilitating contact (by videophony) with relatives or caregivers, or detecting dangerous situations, like falls anywhere in the home, and evaluating them by authorized persons via mobile telepresence. To this end, within the CompanionAble project a socially assistive, mobile robot companion has been developed (see Fig. 1), which aims at assisting the elderly in their daily life. The specifics of socially assistive robotics is that it is focused on helping human users through social rather than physical interaction [3]. So, manipulation skills are explicitly not subject of this project. Nonetheless, the spectrum of assistive functionalities that can be provided by the robot companion is manifold and reaches from situation-specific, intelligent reminding (e.g. taking medication or drinking) and cognitive stimulation by means of tailored training exercises, up to the aforementioned detection and evaluation of dangerous situations.

Fig. 1. 89 years old lady in interaction with the mobile home robot companion (pre-final version) developed in the CompanionAble project [1].

The paper is giving a summarizing overview of the progress in developing such a home robot companion. For this purpose, it is addressing several aspects beginning with the system requirements and the required functionalities (Sec. II), continuing with a brief overview of related work in the field of socially assistive robotics for domestic use (Sec. III), the derived consequences for the hardware design and the functionality of the robot companion (Sec. IV), its system and control architecture (Sec. V), a key technology for human-robot interaction in home environments - the user tracking and searching (Sec. VI), up to the results of

---

This work has received funding from the European Community's 7th Framework Program (FP7/2007-2013) under grant agreement no. 216487.

H.-M. Gross, Ch. Schroeter, S. Mueller, M. Volkhardt, and E. Einhorn are with the Neuroinformatics and Cognitive Robotics Lab, Ilmenau University of Technology, 98684 Ilmenau, Germany. `Horst-Michael.Gross@tu-ilmenau.de`

A. Bley, Ch. Martin, T. Langner, and M. Merten are with MetraLabs Robotics GmbH Ilmenau, 98693 Ilmenau, Germany. `www.metralabs.com`

already conducted and ongoing functionality testings with a particular focus on experimental results of the "Search user" behavior (Sec. VII). Section VIII closes with an outlook on the upcoming usability evaluations of the robot companion.

II. REQUIREMENTS SPECIFICATION

The specification of the requirements on the home robot companion is based on a set of surveys and studies the CompanionAble consortium carried out in the first year of the project, as for example that presented in [2]. In the different surveys, more than 250 end-users (care recipients (CR), their relatives, caregivers, and care professionals) in Austria, Belgium, Great Britain, France, Holland and Spain have been interviewed about their individual needs and priorities in assistive smart home technologies and services that could be provided by home robot companions. As a result of these end-user studies the following categories of system requirements could be identified:

- *Communication:* any function that allows keeping in touch with the CR's family, friends, or care professionals, e.g. by videophony and telepresence
- *Safety:* any function that directly increases the safety of the CR, e.g. reminding taking medication, user monitoring and fall detection including emergency call functionality and remote-control of the robot for evaluating critical situations anywhere in the home
- *Services and assistive functions:* any function that improves the CR's daily life, as e.g. daytime management (daily routine, agenda updating), suggestion of clothing based on weather forecasts, keeping and dispensing personal items of the CR, entertainment and interactive media access
- *Therapy:* all functions that are intended to improve the CR's state of health, e.g. cognitive training exercises or stimulating games
- *Smart situation awareness:* information the robot companion can collect to allow for adapting its behavior to the CR's specifics, e.g. her typical daily activity course, or her preferences in communication with the robot.

To fulfill these requirements, a great number of different functionalities need to be ensured for navigation, HRI, and the assistive services. For autonomous *navigation* this includes: a) to build a detailed map of the apartment while the robot is manually driven around, b) to robustly self-localize within the apartment, c) to drive to any target position in the home, d) to robustly avoid collisions with known and unknown obstacles, e) to efficiently pass through narrow doors or gaps between furniture, f) to autonomously drive and dock to the charging station, and g) to pass over thresholds and carpets.

For *HRI*, the following requirements were defined as mandatory: h) to robustly detect and keep track of a moving/static person, i) to orient towards a user or drive in a position facing the user, j) to follow a user through the room and the apartment, k) to autonomously search for a user in the apartment, l) to express simple emotions by using its facial capabilities, m) to notice when personal items of the CR are placed in the robot's tray, n) to understand a defined set of command words/phrases, and o) to recognize a defined set of critical sounds, e.g. glass shattering.

Regarding the *overall services*, the following use cases were specified: p) initiating a videophone call to relatives or a care service, q) reminding the user to do a cognitive stimulation exercise, r) reminding of a forgotten event (e.g. taking medication), s) greeting the user at the entrance of the apartment when returning home including t) storing personal items in the robot's tray, and u) detecting a fallen user anywhere in the home and evaluating the situation by tele-operating the robot.

The user studies have also shown that the existence and the design of a robot head play a crucial role in the context of enabling and stimulating social interaction. Only a robot head can really stimulate and motivate people to listen to the robot and to engage in dialog with it showing adequate robotic emotions during interaction. Moreover, the robot should be able to keep and identify personal items with a tendency to be misplaced by the elderly. These and some more requirements have directly influenced the design and the functionality of the robot companion (see Sec. IV and V).

III. RELATED WORK

In recent years, socially assistive robotics for domestic use has been a rapidly increasing field of research and development [3]. Therefore, this section can only give a fragmentary overview of the state-of-the-art in this field with a focus on socially assistive home robot companions. Developments on mobile robots with manipulation skills, as for example the well-known MOVAID, Hermes, ARMAR, or the probably most advanced wheel-based robotic home assistant Care-O-Bot, are not covered here. A very good overview of this branch of assistive robotics is given in [4].

The CMU Pittsburgh *NurseBot* project [5] (1998-2005) was one of the first research endeavors that dealt with robotic assistance for the elderly. The robotic platform was used to test several ideas for such a companion, including intelligent reminding functions, telepresence applications, surveillance, simple social interaction, and help for physically impaired people. The European FP6 project *COGNIRON* [6] (2004-2007) was focused on researching enabling technologies for socially assistive robots whose ultimate task was to serve humans (not primarily elderly or cognitively impaired persons) as a companion in their daily life. To this end, the focus was on studying the perceptual, representational, reasoning and learning capabilities of embodied robots in human-centered environments. The European FP6 project *Robots@Home* [7] (2007-2010) aimed at creating an open mobile platform that should pave the way for introducing robots in homes. This involved not only the development of the hardware, but also the creation of an embedded perception system for learning rooms and maps which should facilitate safe and robust navigation in homes. The still running FP7 project *LIREC* (Living with Robots and Interactive Companions) [8] (2008-2012) addresses a number of key scientific and technical areas required for long-term companion relationships with a focus on perception, memory, emotions, communication, and learning. This also includes an empirical evaluation in real social settings to validate the theory developed.

The approaches mentioned here are either research-oriented systems or proof-of-concept demonstrators not yet ready for operation in real home environments with ordinary end-users. Typically they show limited functionality regarding a completely autonomous navigation behavior and require instructed users and the presence of roboticists during their operation to handle unexpected situations. An important step on the way to introduce social robotics technologies in real homes has been done with the Swedish mobile telepresence system *Giraff* [9], which has been available on the market since 2010. The *Giraff*, however, is completely operated by remote-control by authorized relatives and caregivers. The communication occurs by means of a tele-operated, tiltable display with camera, and a speaker and microphone mounted on a mobile base.

Having true autonomy in mind, none of these systems can already be considered as home robot companion suitable for everyday use. Moreover, most of them have not been involved in assistive tasks over longer periods of time, or were subject of long-term field trials in real homes. Therefore, the work in CompanionAble is aiming at both aspects. In this paper, however, we are focusing on the technical functionalities and design aspects of a home robot companion that fulfils the user requirements, is suitable for everyday and long-term use, and allows for a low-cost producibility as important prerequisite for the marketability of such a robot.

IV. DESIGN OF THE ROBOT COMPANION

To ensure a high acceptance of the robot companion by the different end-user target groups, their requirements regarding the functionality, usability, and appearance of the robot (see Sec. II) have been taken seriously from the beginning of the design process. With a height of 120 cm, the robot companion is comparable to the size of an 8 years old child. Its size is optimized for a friendly appearance and an ergonomic operability by a standing or sitting user. Fig. 2 depicts the pre-final (left) and the final, still to be completed version (right) of the robot companion. As usual for socially assistive robots, our companion does not have any manipulators. In both versions, the drive system consists of a differential drive and a castor on the rear. This gives it a good maneuverability and stability in spite of its height and weight of less than 50 kg. For navigation, user monitoring, HRI and safety, both versions of the robot are equipped with various interaction devices and sensor systems subsequently described.

Tiltable touch-Screen: The touch-screen is the central unit for graphical and touch-based communication with the robot. It is designed as vertically mounted 15,4 inch wide screen display. To ensure the best usability, the display unit can be tilted up and down in order to allow for an optimal interaction with a standing and sitting user (see Fig. 1). To avoid trapping of fingers or hands between the display and the robot casing, for safety reasons a force sensor is integrated in the tiltable display to measure the pushing force of the display motor.

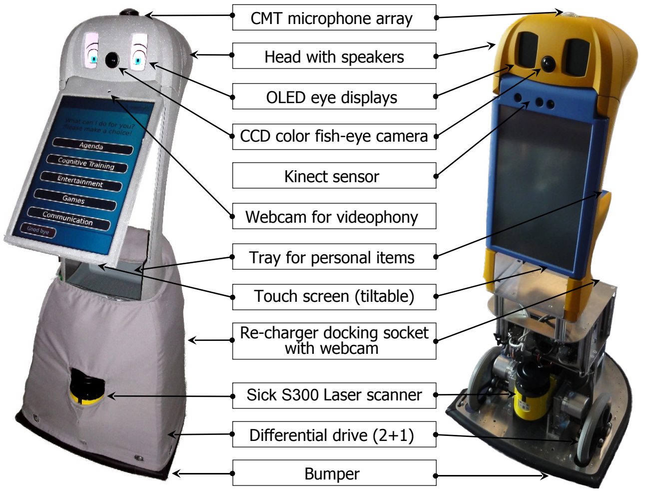

Fig. 2. Design stages of the robot companion developed in the CompanionAble project [1] with its main equipment for environment perception, navigation, and HRI. The left figure shows the pre-final version of the robot which has been and is being used for first functionality tests and diverse user trials in 2010 and 2011. The upper part of the robot's enclosure is made of ABS plastic mixed with aluminium. As interim solution, the lower part is made of textile fabric, which will be replaced by a whole-body ABS plastic enclosure in the final version shown on the right (lower part omitted).

Robot head and face: As additional interaction channel and imitating a pleasant technical face, the robot has a head with two OLED displays as eyes, which can be used for expressing a spectrum of robotic emotions or following a moving user with the eyes. These features are very helpful for intuitive HRI and for getting and staying in contact with the user.

Tray for personal items: As requested by all end-user target groups, the robot companion is equipped with a tray, where the elderly person can place personal items, which s/he tends to mislay more or less often, like glasses, wallet, keys, mobile phone, etc. As result of usability studies, the tray has been rearranged from a position below the display in the pre-final version to a higher position behind the touch-screen for the purpose of a better reachability when sitting and standing. As an important feature, the tray uses RFID technology to detect and automatically register personal items equipped with RFID tags. This tagging is not a limitation anymore, because very small tags with a size of 2x2 mm are available meanwhile. Using the knowledge about the items put down in its tray, the robot is able to remind the user to move something in or out, if needed.

Docking concept for battery re-charging: For completely autonomous re-charging of the battery, a new docking concept had to be developed. One of the main challenges was to find a technological solution that allowed for a direct connection of the robot to the line voltage of the home, because it should be possible to connect and disconnect the charging connector by the robot platform itself without expansive actuators or the assistance by a human. Inspired by the connector concept known from water boilers used in millions of households, a safe charging connector has been developed with a socket on the rear side of the robot (see Fig. 2) and a spring loaded plug for the wall. Above the socket, there is a low-cost webcam mounted to detect synthetic markers that are used for the autonomous docking to the socket. In functional testings, we achieved a very good and robustly repeatable docking behavior which is an important prerequisite for autonomous re-charging and, therewith, the long-term applicability of the robot companion.

For HRI and navigation, both versions of the robot are equipped with diverse sensor systems (see Fig. 2). Here only a brief overview is given, while in [10] more details are described. One high resolution camera (1600 × 1200 pixels) with a 180^o fish-eye lens mounted in the middle of the robot's face as a kind of snub nose is used for people tracking and obstacle detection. In the pre-final version, a further webcam is placed within the tiltable display immediately above the touch screen. This way, the camera can continuously face the user during interaction with the robot which is a prerequisite for such functionalities like videophony or facial expression analysis during interaction. This webcam has been replaced by a Microsoft Kinect sensor in the final version of the robot. The depth camera of the Kinect provides a dense depth image that can be used to build dense 3D point clouds of the environment in front of the robot. Therefore, in the final version the depth camera will be used as additional sensor for more robust obstacle detection as well as person detection and tracking. The RGB camera of the Kinect is to substitute the webcam for videophony and facial expressions analysis during interaction. Moreover, the robot is equipped with a microphone array using the coincidence microphone technology (CMT) from project partner AKG in Austria (www.akg.com). The CMT array placed at the top of the robot is utilized for localizing a talking person and sound signal enhancement based on beam-forming which is helpful for videophony and speech-based people tracking.

For obstacle detection, map building, localization, and person tracking, the robot is additionally equipped with a SICK S300 2D-laser range finder. The protective field areas of this sensor are used to detect obstacles close to the robot. This allows for the lowering of the moving speed if a collision is imminent. A closed rubber-based security collision sensor (bumper) mounted around the robot's base plate is used to detect collisions of the robot with obstacles that are invisible for the other sensor systems or moving faster than the reaction time of the robot. The sensor information of the laser range finder and the collision sensor is directly connected to the motor controller. This guarantees a rapid and reliable system response and stopping brake independent of software control. All safety-relevant sensors are integrated in accordance with technological standards and will be certified by the German Technical Inspection Agency in the near future.

V. SYSTEM ARCHITECTURE

To guarantee the main requirements to a modern robot control architecture, like modularity, extensibility, efficiency, customizability, reusability, and rapid application development, in continuation of our approach presented in [11] we have separated the robot-specific methods and skills from the application itself resulting in a flexible layered system architecture (Fig. 3). In this architecture, the bottommost layer, the *Hardware Layer*, encloses the robot hardware (sensors and actuators), the operating system, and the low-level interfaces to the hardware. The low-level sensor information is processed in the next higher level, the *Skill Layer*, which provides a set of necessary, robotic-specific basic skills that are executed in the Hardware Layer. These skills are only weakly coupled and do not need any coordination due to non-conflicting hardware resources. The *Skill Layer* covers the whole spectrum of navigation skills including map building and localization. Furthermore, specific perception skills, like person detection and tracking as well as face analysis, are implemented in that layer. Speech synthesis and the GUI are also placed in this layer and provide their services to the higher application levels. Due to the modularity of this layer, existing open source software solutions, e.g. for navigation and user detection, can be and have already been integrated here (see Sec. VI).

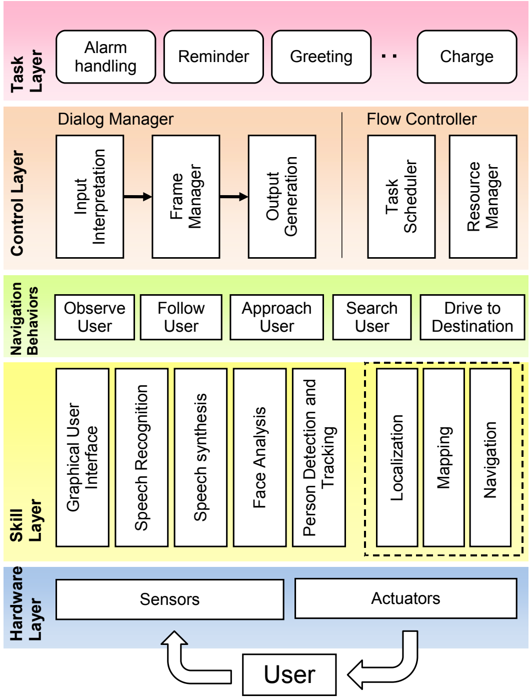

Fig. 3. Main components of the robot's layered system architecture.

Above the independent skills there are diverse modules representing the application control which make use of the basic features provided by the skills. The simplest controllers using the navigation and people detection skills are the *Navigation Behaviors*. These are exclusive units each representing an individual control loop for accomplishing the different task-oriented navigation behaviors of the robot. Here, for example, a "Search user" behavior is realized that uses the history of person detections and a set of navigation points, which are checked one after the other in order to cover the whole apartment, while searching for the user using the ”Person detection and tracking” and ”Navigation” skills. Other behaviors are ”Approach user” and ”Follow user” which are necessary for direct interaction as well as passive user observation, which is active most of the time when no direct interaction takes place. When this navigation behavior is active, the robot looks for an unobtrusive position, from where it has a direct view to the user in order to recognize critical situations, like fall events.

The exclusive navigation behaviors as well as the input and output skills are utilized by the modules of the *Control layer*, which form the ”Dialog Manager” and the ”Global flow control”. The ”Dialog Manager” consisting of an input interpreter, a frame manager, and an output generation module is responsible for organizing the direct interaction with the user. Here, the infrastructure for multi-modal input fusion, coordination of the dialog progress, and generation of multimodal outputs using speech, display, and facial expressions is provided. The ”Dialog Manager” follows a frame-based approach, where the content of the dialog is defined in a set of scripts residing in the highest layer, the *Task Layer* where the content of the application is defined.

Since the ”Dialog Manager” is more a reactive subsystem handling interaction with human once the dialog has been initiated, the proactive part consists of a ”Task Scheduler” and a ”Resource Manager”, which are responsible for coordination of the partially competing task controllers implemented in the *Task layer*. These controllers are realizing different primitive control loops, as e.g. adjusting the display to the user position or coordinating the charging of the robot, which incorporates a docking to the charging station as well as an undocking if any other task needs to drive around. Also hierarchical dependencies among these tasks can occur.

By means of this decomposition, we could avoid instantiating one complex state machine for application control and preserved extensibility and reusability. Coordination among these controllers of the *Task layer* is realized by a set of system resources, which are granted according to a dynamic prioritization by means of the ”Resource Manager”. This prioritization is also represented in the ”Task Scheduler”, which is responsible for the activation of event- and time-triggered tasks, like reminding to take medication, giving advice in daily activities, or offering the cognitive training exercises. All these tasks have a priority value allowing for a coordination with the user’s current activity, in order to be situation-adaptive and not disturbing.

Example for activation flow: For better explaining the internal flow of activation, a typical use case has been selected - the delivery of a reminder (e.g. to take medication), which is defined by a period in time when it has to be delivered, and a priority to prevent an immediate interruption of other ongoing and possibly more important activities. In that case, the robot normally is in the passive observation behavior, and the ”Person detection and tracking” skill tries to recognize the user’s current activity, which is written to a user state model. The ”Task Scheduler” now waits for an adequate moment in time when the user has finished the current activity and is available for a reminder dialog. In this case, the ”Task Scheduler” will indicate the need for interaction to a controller in the *Task layer*, which will request the drive resource in order to activate the ”Approach user” behavior. Hence, the other controllers have to reach a safe state and have to release the drive resource first. Once the ”Approach user” behavior has been activated, the ”Navigation” skills are utilized for driving to the position determined by the ”Person Detection” skill. In case of a user found in interaction distance to the robot, the ”Task Scheduler” finally will execute the task and trigger the respective frame in the ”Dialog Manager”. This reminder frame will use the ”Dialog output” generation in order to express the message verbally as well as on the screen. A possible answer given by the user will be recognized by the ”Speech recognition” skill or the touch display. Also a head nodding is possible and will be interpreted by the input interpretation module as a confirmation that the message has been understood by the user. In case of a missing confirmation or a request for presenting the reminder later, the frame can reschedule the reminder at the task scheduler. If the message has been delivered successfully, the frame will be finished and the dialog will be finalized.

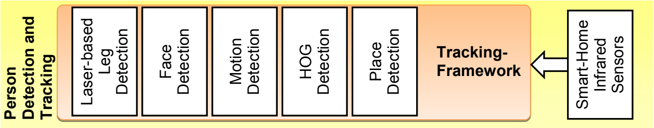

Fig. 4. Overview of the tracking framework. In this work, we used 4 different observation cues to feed the tracker. The innovation is a module to detect people lounging at resting places. Furthermore, optional IR motion sensors can guide the behavior of the robot.

VI. USER TRACKING AND SEARCHING

To offer the different service functionalities to the care recipient (CR), the robot system employs several autonomous behaviors (see Fig. 3, middle layer). First, observing the user in a non-intrusive way allows to facilitate services that require interaction or to react on critical situations. A second behavior is following and approaching the CR if interaction is desired. Third, the robot must seek for the CR if a reminder has to be delivered or a video call comes in, and the CR is not in direct proximity of the robot. In that case, the robot should also be aware if the CR is not at home at all. A prerequisite to all these behaviors is the robust detection and tracking of the user in the apartment. In contrast to other interaction applications in public environments, people in home environments often do not face the robot in an up-right pose but sit on chairs or lie on sofas. Therefore, our system combines state-of-the-art methods for up-right pose people detection with a module to detect users independently of their pose at places, where they usually rest.

A. User Detection and Tracking

Our person detection and tracking system comprises a multi-modal, multi-cue tracking framework based on the

Kalman Filter update regime. The advanced system handles a set of independent 3D position hypotheses of people, which are modeled by Gaussian probability distributions. Adding the velocity results in a six dimensional state space $\mathbf{s} = (x, y, z, v_x, v_y, v_z)$ for each hypothesis. We use the head of the user as the reference for alignment. Therefore, z denotes the height of the user's head. The tracking system is designed in a framework-like fashion to incorporate the detections of arbitrary observation modules, also from existing open source solutions (see below). New position observations are transformed to the 3D representation of the tracking system. When using range-based detection modules, a Gaussian is created at the x, y position of the range measurement with a height value set to the common size of a person $z = 1.70$. In case of visual detection modules, we transform the bounding box of the user into a 3D Gaussian by using the parameters of the calibrated camera and estimating the distance my means of the size of the bounding box. The detection quality of the respective sensor is incorporated into the covariance of the Gaussian distribution, i.e. laser-based detection results in low variance in distance and direction but in large variance in height while visual detections have a high variance in distance estimation. Each resulting detection is associated with the closest hypothesis in the system. If a distance threshold is exceeded, a new hypothesis is introduced. Once the system knows the associated source for that observation, the position of that hypothesis is updated using the Kalman filter technique. In this work, we apply laser-based leg detection and multiple visual detection modules (Fig. 4). The first module is based on the boosted classifier approach of [12] and discovers legs in laser-range data. By searching for paired legs, the system produces hypotheses of the user's position. The face detection system utilizes the well-known face detector of Viola & Jones [13]. The motion detection cue of [14] is only activated when the robot is standing still and uses a fast and simple image difference approach. Furthermore, we apply a combination of a full-body HOG detector [15] and an upper body detector [16]. The system described so-far is able to detect and track upright standing (mainly through legs and HOG) and sitting in frontal-view people (mainly through face and upper-body HOG) in the surroundings of the robot.

B. User Detection at Common Resting Places

When a person is watching TV, reading newspaper, making phone calls, working or sleeping, s/he is typically not doing this in an upright sitting or standing pose, but is sitting or lounging on chairs or lying on sofas. In these cases the cues typically fail. Therefore, we developed a method that first learns the appearance of places in the apartment where the user usually rests. Afterwards, the deviation of occupied places from the respective models and the similarity to an up-to-date user model are used for detection.

Definition of places: We define places as positions in the apartment where the user is usually encountered, e.g. chairs, sofas, working desk. Each place P is represented by a 3D box. Figure 5 (left) shows an exemplary place position in the world centered occupancy map used for navigation, and the 2D-projection of the place box into the current camera image of the robot. Naturally, the content of the place-boxes looks completely different in the camera image, if observed from different positions. Since the system is learning the appearance of a number of places in the apartment, we need to restrict the number of poses from which the robot is observing them. Therefore, each place is assigned n observation poses $\mathbf{O} = (\mathbf{o}_1, \ldots, \mathbf{o}_n)$, where $\mathbf{o} = (x, y, \phi)$ with x, y representing the world coordinates of the robot's position and ϕ denoting the heading of the robot. The restriction of the number of observation positions ensures that the variance of the place appearance is limited. Additionally, some kind of feature description model $\mathcal{M}$ of the place is added. In this work, we use a contextual color histogram as place model.

Fig. 5. Place definition. (Left) Bounding box (red rectangle) of one place (the couch) in the occupancy map of the test apartment. The robot is in its observation position (red circle). (Right) The place's bounding box projected into the camera image shows the couch with the user resting on it.

Contextual color-based place model: The color-based place model comprises the appearance of each place in multi-modal histograms (RGB color space with 8 bins in each dimension). Each place is observed from different, but pre-defined view-points given different illumination conditions, e.g. ambient day-light and electric lighting in the evening. For an efficient representation, we use a multi-modal color place model augmented by a discrete context distribution capturing the circumstances of the histogram's acquisition (e.g. the origin of the histogram, like viewpoint and day-time). In the process of learning, the model maintains unique and distinctive representations of a place, but merges similar descriptions. An important condition is that the user is not occupying these places during the learning phase.

User model: The color model of the user is similar to the color model of places, but without the context distribution. Model learning is done by first creating a Gaussian Mixture background model [17], when the robot is standing still and no hypothesis is in front of the robot's camera (given by the tracker output). This background model is used for background subtraction. Then, the user is asked to walk in front of the robot's camera to learn the 3D color histogram of the user in the foreground. To remove shadows and to refine the segmentation, we apply the GrabCut algorithm [18].

Recognition of the user: Once the place models and the user model have been trained, the system is able to detect the user in arbitrary poses at the learned places. For that purpose, the robot drives to the predefined observation positions and checks each place. By comparing the current appearance to

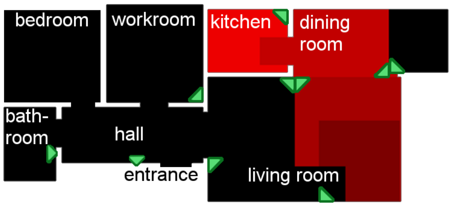

Fig. 6. User position likelihood map of the apartment: triangles indicate positions of IR-sensors. Red color codes the time elapsed since the respective sensor was activated for the last time (the brighter the more recent).

the place and user model, the system can decide if the place is occupied by the user. If the user is present, this results in low similarity to the place model, because the appearance of the place is partially covered, and a high similarity to the user model. If the user is not present, the results are vice versa. To finally decide if a place is occupied, we trained a single linear Support Vector Machine (SVM) [19] on data of multiple labeled runs with empty and occupied places. The resulting SVM then decides for each place if the user is present given the similarities to the place and user model. A more detailed description of our approach is given in [20].

C. Integration of Smart Home Sensors

Every time the user needs to be searched in the apartment, because s/he got out of the range of the robot's sensors, a proper search strategy has to be activated in the "Search User" behavior. Then the robot checks each of the aforementioned observation positions for the user's presence using the tracking framework including the place-detection module. The tracking system also detects a standing user while driving from one observation point to another. Generally, if no data is available from the motion sensors, the robot starts with the observation position closest to the last known position of the user. By incorporating the detection capabilities of IR motion sensors installed in the smart home environment, a more sophisticated search strategy can be applied. The achievable spatial accuracy of the sensors is not sufficient for fine user detection suitable for interaction, but more than enough to decide where to search first. The user position likelihood map shown in Fig. 6 is built up by the robot using the history of activations of the IR motion sensors. In the given example, the user was sitting on the couch in the living room, before s/he moved through the dining room to the kitchen, where s/he is resting now. On a given "Search User" task, the robot now can start its search with the place with the highest (most recent) activation (the kitchen), before it begins to check the other places with weaker activations (dining room, living room). If nobody is detected after all places have been visited, the user is assumed to be not at home. Using the sensor information decreases seeking time enormously, because the robot usually drives directly to an observation position close to the user.

VII. FUNCTIONALITY TESTINGS

As described in Section II, we defined a number of system requirements with respect to the navigation and HRI capabilities of the robot companion to assess the performance of the different subsystems and the robot companion as a whole before usability studies can take place. With respect to the *navigation functionality*, most of the defined tests (a)-(g) have already been completed successfully from qualitative point of view, only test case (f), the autonomous driving and docking to the charging station, still needs to be finalized until the upcoming user studies. A detailed quantitative analysis of the navigation capabilities regarding localization and positioning errors, velocity profile and smoothness of the movement trajectories, distances to obstacles, etc. is still pending, but in preparation.

The defined requirements on *HRI* (h)-(o) are also currently being tested and evaluated. Most of these tests could be finalized successfully from qualitative point of view, however, the functionalities (j), (l), (n), and (o) still need to be made more robust. First quantitative studies could be finalized quite recently, and we expect the others to be finished by fall 2011.

Of particular importance for the robot's service functionality is the requirement (k), the search user behavior, because the robot should robustly find its current user as quickly as possible if s/he has been lost from view and reminders and incoming video calls have to be delivered. For the quantitative study of this behavior, we played more than 100 hide-and-seek "games" (test runs) and determined the average search time and the success rate (see [20], too). In these studies, all detection modules of the tracking system shown in Fig. 4 were activated, and the smart home IR-sensors were utilized to sequence the observation positions as described in Sec.VI-C. Each game started with the robot situated on a fixed starting position. The user then "hid" somewhere in the apartment by resting on one of the learned places. Occasionally, the user did not occupy any of these places, but stood somewhere in the apartment. Furthermore, in a few games the user was not present in the apartment at all. The ground truth of the user's position was labeled manually. The robot then started searching for the user by driving to each observation position and checking for the user's presence. If the user was not found on a specific place, the robot went on checking the other places. If the robot found the user lounging at a resting place or standing somewhere in the apartment, it logged the detection position and the time when the detection was made and returned to the starting position to end the game. If the user was not present in the apartment or the robot failed to detect her or him on the specific place, the robot returned to the starting position after checking all places, logging the moment of arrival.

Table I shows the results of these experiments. We regarded a test run as successful if the distance between the robot's user detection and the ground truth was below 1 meter, or if the robot returned to the starting position after checking all places, if the user was not present. Table I (a) depicts games in which the user was mostly lounging at different places and occasionally just standing. The success rate was rather high with 74%. Errors mostly occurred when the user was in an unfavorable position and proper color

TABLE I
RESULTS OF DIFFERENT HIDE-AND-SEEK STUDIES.

	test runs	successful runs	success rate	average duration
(a) user lounging	73	54	74%	27.8 s
(b) user standing	15	13	87%	32.2 s
(c) w/o IR-sensors	19	8	44%	37.2 s
(d) user not at home	18	3	15%	25.4 s

histogram extraction was impossible. The average time over all successful and unsuccessful games was 27.8 seconds with a minimum of 5 and a maximum of 122 seconds. We also tested the performance of the system in additional test runs, in which the user was always standing somewhere in the apartment (Tab. I (b)). Compared to the aforementioned games, where the user usually lounged at a place, the success rate increased to 87%. This is because the range- and HOG-based detection modules are particularly dedicated detecting standing people. The average time to find the user increased a little, because the user was not on one of the predefined places but had to be found by the robot when driving from one observation point to another. When disabling the IR-sensors (Tab. I (c)), the success rate dropped down to 44%, and the average time to finish a game increased from 32 to 37 seconds. Without the initial hint of the IR-sensors, the robot had to check each place, increasing the average search time and the chance of false positives. The more places the robots needs to check before reaching the place occupied by the user, the lower the probability of a successful game. This becomes extreme, if the user is not at home, so that the robot has to check all defined places and return to the starting location. Given the seven places used in the test apartment and a mean classification rate of 0.8 for each place, the expected probability of a successful run can be only 21%. This explains the low success rate of only 15% when no person was present in the apartment (Tab. I (d)).

VIII. OUTLOOK ON USER STUDIES

The focus of the still pending *usability studies* will be put on the evaluation of the user-robot interaction, and the way the system is perceived by the users as meeting their needs. The assessment of the usability is going to be started in August 2011 and will include the observation of the test participants by recording usability problems, and *real end-users* - volunteers suffering from MCI. To guarantee a strong scientific approach, the studies are prepared and conducted by domain experts, i.e. experts for usability engineering and ethical evaluation, and partners with strong background in geriatric medicine, who belong to the CompanionAble consortium [1]. The evaluation will take place in two different settings. A first setting will be provided with the robot in a fully equipped smart house of the Dutch project partner Smart-Homes (www.smart-homes.nl) in Eindhoven, the Netherlands, which serves as a test house in general. In this case, the test users will temporarily stay in the house and perform the specified test cases. Because it is important for persons with MCI to stay in their own home environment, a subset of the usability test will be conducted with the robot companion and the volunteers in their own apartments.

In the future work, particular attention has to be paid to quantifying the long-term effectiveness of socially assistive robot companions on health and well-being of the care recipients. As suggested by [21], this could be done by measuring (i) the level of stress (e.g. by stress hormones in urine), (ii) the positive mood (evaluating the facial expressions and using questionnaires), or (iii) the communication activity between the elderly or with their families (measurable by the frequency of contact between them). However, for these investigations it is necessary to expand the length of the studies from days to several weeks or even months, what requires really autonomous robot companions suitable for long-term use in real homes.

REFERENCES

[1] CompanionAble project homepage: www.companionable.net

[2] A. Badii, I. Etxeberria, C. Huijnen, A. Maseda, S. Dittenberger, A. Hochgatterer, D. Thiemert, A-S. Rigaud. ”CompanionAble: Graceful integration of mobile robot companion with a smart home environment”, *Gerontechnology*, 8 (2009) 3, 181

[3] A. Tapus, M. Mataric, B. Scassellati, ”Socially Assistive Robotics”, *IEEE Robotics & Automation Magazine*, 14 (2007) 1, 35-42

[4] B. Graf, H. Staab, ”Service Robots and Automation for the Disabled/Limited”, *Springer Handbook of Automation*, 1485-1502, Springer 2009

[5] M. Montemerlo, J. Pineau, N. Roy, S. Thrun, and V. Verma, ”Experiences with a Mobile Robotic Guide for the Elderly”, in *Proc. AAAI Nat. Conf. on AI 2002*

[6] Project homepage: www.cogniron.org

[7] Project homepage: www.robots-at-home.org

[8] Project homepage: www.lirec.org

[9] Project homepage: www.giraff.org

[10] H.-M. Gross, Ch. Schroeter, S. Mueller, M. Volkhardt, E. Einhorn, A. Bley, Ch. Martin, T. Langner, M. Merten, ”I’ll keep an Eye on You: Home Robot Companion for Elderly People with Cognitive Impairment”, in *Proc. IEEE-SMC’11*, to appear

[11] H.-M. Gross, H.-J. Boehme, C. Schroeter, S. Mueller, A. Koenig, E. Einhorn, Ch. Martin, M. Merten, A. Bley, ”TOOMAS: Interactive Shopping Guide Robots in Everyday Use - Final Implem. and Experiences from Long-term Field Trials”, in *Proc. IROS 2009*, 2005-12

[12] K.O. Arras, Ó.M. Mozos and W. Burgard, ”Using Boosted Features for the Detection of People in 2D Range Data”, in *Proc. ICRA 2007*, 3402–3407.

[13] P. Viola, M. Jones, ”Rapid Object Detection using a Boosted Cascade of Simple Features”, in *Proc. CVPR 2001*

[14] Chr. Martin, E. Schaffernicht, A. Scheidig, H.-M. Gross, ”Sensor Fusion using a Probabilistic Aggregation Scheme for People Detection and Tracking”, *Robotics and Auton. Systems*, 54 (2006) 9, 721–728.

[15] N. Dalal, B. Triggs, ”Histograms of Oriented Gradients for Human Detection”, in *Proc. CVPR 2005*, 886–893.

[16] V. Ferrari, M. Marin-Jimenez, A. Zisserman, ”Progressive search space reduction for human pose estimation”, in *Proc. CVPR 2008*, 1–8

[17] C. Stauffer, W.E.L. Grimson, ”Adaptive Background Mixture Models for Real-Time Tracking”, in *Proc. CVPR 1999*, 2246–2253.

[18] C. Rother, K. Vladimir, A. Blake, ”Grabcut: Interactive foreground extraction using iterated graph cuts”, in *Proc. ACM SIGGRAPH 2004*, 309–314.

[19] C. Lin, Chih-Jen, ”LIBSVM: a library for support vector machines.” 2001. Available: www.csie.ntu.edu.tw/~cjlin/libsvm

[20] M. Volkhardt, St. Mueller, Ch. Schroeter, H.-M. Gross. ”Playing Hide and Seek with a Mobile Companion Robot”, submitted to *11th IEEE-RAS Int. Conf. on Humanoid Robots 2011* (unpublished)

[21] J. Broekens, M. Heerink, H. Rosendal, ”Assistive Robots in elderly care: a review”, Gerontechnology 8 (2009) 2, 94-103