



Hobbit, a care robot supporting independent living at home: First prototype and lessons learned



David Fischinger^{a,*}, Peter Einramhof^a, Konstantinos Papoutsakis^b,
Walter Wohlkinger^a, Peter Mayer^c, Paul Panek^c, Stefan Hofmann^e, Tobias Koertner^d,
Astrid Weiss^a, Antonis Argyros^b, Markus Vincze^a

^a Automation and Control Institute (ACIN), Vienna University of Technology, 1040 Vienna, Austria

^b Institute of Computer Science, FORTH, Heraklion, Crete, Greece

^c Fortec group, Vienna University of Technology, 1040 Vienna, Austria

^d Academy for Aging Research, 1160 Vienna, Austria

^e Hella Automation GmbH, 9913 Abfalterback, Austria

HIGHLIGHTS

- We present a care robot for aging in place by means of fall prevention/detection.
- Detailed description of sensor set-up, hardware, and the multimodal user interface.
- Detailed description of major software components and implemented robot tasks.
- Proof-of-concept user study (49 user) on usability, acceptance, and affordability.

ARTICLE INFO

Article history:

Available online 7 October 2014

Keywords:

Social robotics

Robots for elderly

Care robot for independent living

ABSTRACT

One option to address the challenge of demographic transition is to build robots that enable aging in place. Falling has been identified as the most relevant factor to cause a move to a care facility. The Hobbit project combines research from robotics, gerontology, and human–robot interaction to develop a care robot which is capable of fall prevention and detection as well as emergency detection and handling. Moreover, to enable daily interaction with the robot, other functions are added, such as bringing objects, offering reminders, and entertainment. The interaction with the user is based on a multimodal user interface including automatic speech recognition, text-to-speech, gesture recognition, and a graphical touch-based user interface. We performed controlled laboratory user studies with a total of 49 participants (aged 70 plus) in three EU countries (Austria, Greece, and Sweden). The collected user responses on perceived usability, acceptance, and affordability of the robot demonstrate a positive reception of the robot from its target user group. This article describes the principles and system components for navigation and manipulation in domestic environments, the interaction paradigm and its implementation in a multimodal user interface, the core robot tasks, as well as the results from the user studies, which are also reflected in terms of lessons we learned and we believe are useful to fellow researchers.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Because of the fact that older adults prefer to independently live on their own at home as long as possible [1], the necessity

* Corresponding author. Tel.: +43 158801376622.

E-mail addresses: peter.einramhof@gmail.com (P. Einramhof), walter.wohlkinger@gmail.com (W. Wohlkinger), s.hofmann@inno-cube.at (S. Hofmann), Tobias.Koertner@altersforschung.ac.at (T. Koertner).

<http://dx.doi.org/10.1016/j.robot.2014.09.029>

0921-8890/© 2014 Elsevier B.V. All rights reserved.

of developing assistive technology is increasing. However, older adults themselves experience challenges in maintaining their home [2] and the need of assistive technology can be perceived as stigmatization [3].

Care robots are considered as one option to support independent aging in place. In recent years the development of this type of robot assistance quickly increased. Several research projects were focusing on the development of care robots for older adults, which can reduce loneliness, support in household tasks, and



Fig. 1. The “naked” Hobbit robot (left) and the Hobbit robot (prototype 1) used for the first round of user trials (right) in Austria, Greece, and Sweden.

connect users to the outside world (e.g. KSERA [4], DOME0 [5], Cogniron [6], Companionable [7], SRS [8], Care-O-Bot [9], Accompany [10], HERB [11]). A large number of studies on the impact of care robots on older adults (a thorough overview can be found in [12]) demonstrate their positive effects and it also seems that the elderly in the Western society are open for this kind of technology.

However, most of the developed robotic platforms are still research platforms, which did not enter private households as robotic products so far. Only pet-like social companion robots with limited (but essential) care functionality, such as the seal-like robot Paro [13], are available in the market for end users. To our conviction, one of the biggest challenges is still the development of a multifunctional care robot for independent living, which is affordable for end users.

Therefore our aim is to develop an affordable multifunctional care robot that sustainably enables independent aging in place. According to gerontological studies, falls are the leading cause of injuries among elders. EU-wide 30% of people over 65 and 50% of those over 80 years fall each year and falls are the main reason for moving to a care facility. Subsequently, our main goal is to develop a care robot which prevents falls by picking up objects from the floor, detecting falls by patrolling through the apartment, and handling emergencies by calling relatives or in last instance the ambulance.

Many questions about how a care robot can be developed that offers high usability and user acceptance, but is affordable for its intended target group are still unanswered. What is the right compromise between user expectations and system capabilities? How can existing hardware and software be used to develop a robotic product and not another research platform? Which interaction paradigms are really beneficial for older adults? To address these challenges, we designed and developed the Hobbit robot.

In this article we present results from the development of the first Hobbit robot prototype (subsequently called PT1, see Fig. 1) and the first set of user trials in a controlled laboratory setting in order to explore the reception of Hobbit from its target user group. Section 3 describes the overall system, including the mobile platform, the sensor system, the arm and gripper, and the multimodal user interface. The components are described in Section 4 expanding on navigation, human detection and tracking, gesture recognition, grasping, and object learning and recognition. Next,

the robot tasks are presented in detail in Section 5 followed by a description of the user study and its results on perceived usability, acceptance, and affordability from the perspective of potential end users. Throughout the article lessons learned from the PT1 development are presented for all sub domains. Thereby we want to share our experiences with fellow researchers and make our knowledge available in the research community as a stepping stone towards affordable robotic products for private users.

2. Motivation and contribution

State-of-the-art robots which should increase the quality of life for older adults can be divided into three main categories: (1) social companion robots, (2) household service robots, and (3) telepresence systems. Social companion robots for elderly should decrease the feeling of loneliness and are often designed in a way to substitute real pets. The most prominent example is Paro. Paro [13] is a seal type mental commitment robot. It has been developed for those who cannot take care of real animals and those who live in places where pet-animals are forbidden. Paro is designed to provide three types of effects: psychological, such as relaxation and motivation, physiological, such as improvement in vital signs, and social effects such as instigating communication among inpatients and caregivers. Paro is affordable and publicly available, but obviously limited in support functionalities when it comes to household tasks.

Household service robots for older adults are those that take over household tasks in order to enable independent aging in place. One of the most popular examples is the *Care-O-Bot* research platform [14], developed at the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA). Care-O-Bot is designed as general purpose robotic butler which can fetch and carry objects and also detect emergency situations (e.g. a fallen person) and contact help, however, it is an expensive research platform not intended for the end user market. Similarly, MobiNa, a small (vacuum-sized) robot was developed by Fraunhofer, specifically aiming at performing fallen person detection and video calls in emergency. Another prominent example is the robot Hector developed within the EU project CompanionAble [15]. Hector was designed as a robotic assistant for older adults integrated in a smart home environment and a remote control center to provide the most comprehensive and cost efficient support for older people living at home. However, also Hector is only available as research platform.

Telepresence systems for older adults or homebound people are intended to enable medical and social assistance for communication with parents, nurses, doctors, and patients, such as InTouch Health's RP-VITA remote presence robot, VGO Communications' post-op pediatric at-home robots, Double Robotics Double telepresence robot, and the Giraff telepresence robot. Most of these systems are affordable and are available also for the end consumer market, however, as social companions they only offer very limited support besides social connectedness.

The basic motivation for the development of Hobbit was to combine the three main aspects of the above-mentioned categories (decreasing loneliness, support in household tasks, medical and social assistance through remote communication) in one affordable robotic product (meaning around 15 000 Euro in costs for purchase) for aging in place. Thereby our main goal is that the robot provides older adults with a feeling of safety and being supported in everyday tasks. Consequently, the main functionalities the robot should provide are emergency detection (the robot should autonomously patrol through the apartment after three hours without any user activity and check if the user is ok), handling emergencies (automatically calling relatives or an ambulance and calming the user), as well as fall prevention measures (clearing the

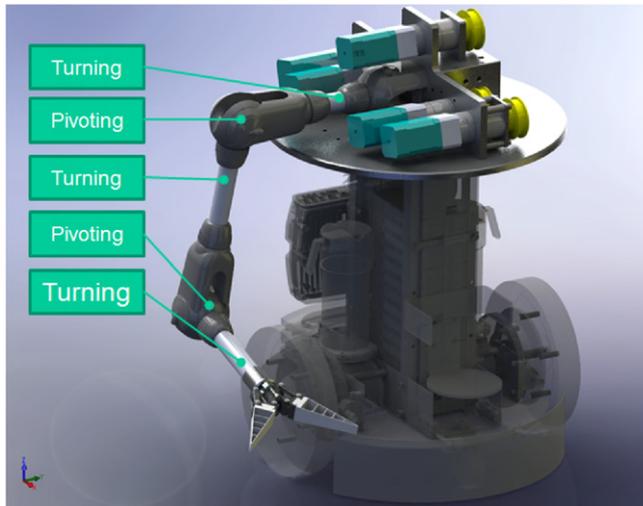


Fig. 2. Platform with 5-DOF IGUS Robolink arm and Fin Ray gripper.

floor twice a day from clutter, searching and bringing known objects to the user). Besides that, functionalities were added to increase Hobbit's abilities to serve as a companion, such as general video telephony, reminding the user of medication, drinking, and appointments, entertainment, such as games, videos, and audio books.

The contribution presented here is the PT1 robot system that sets out to close the gap of existing care robots: rather than focusing on specific care aspects of aging in place, we offer a combined solution that is affordable and provides a satisfying usability and acceptance. We present the contributions in this paper in the form of "lessons learned", where we discuss the challenges we had to face starting from building a complete system and stepping through the key robot components and capabilities. The lessons learned shall help other researchers and engineers to profit from our results to solve challenges complying with many, often conflicting requirements in terms of user requirements, costs, and achievable functionality. The article starts with the development choices regarding hardware components and related dependencies, which we will outline in the next section.

3. System and hardware

3.1. Platform

The lower part of the Hobbit system is a mobile platform (see Fig. 2) with differential drive kinematics. It has a circular cross-section with a diameter of about 45 cm. This combination allows the robot to turn on the spot within its footprint, which is important when navigating in narrow and cluttered domestic environments. The platform houses the batteries (24 V, 18 Ah) that power all electrical components of the robot and currently allow for an average autonomy time of three hours. An onboard PC ("XPC") runs the high-level control software of the robot. An additional controller board provides the low-level motion control for the platform, which can execute translational and rotational speed commands as well as fine positioning commands.

Lessons learned

Regarding the hardware design of the mobile platform it was most challenging to harmonize technical requirements, user requirements, and the goal of a low cost robot. An example is the user requirement for a small robot: Clearly a care robot should not be too big in order not to be threatening for a sitting older adult

and, moreover, the domestic environment also poses restrictions, such as narrow hallways and doorways. Subsequently, the limitations in size (PT1 requirement studies revealed that the maximum size should be 130 cm in height) make it difficult to place sensors and also the arm in a way that objects on table tops or shelves can be detected, reached, and respectively grasped.

Another aspect is the power management to facilitate not only long autonomy times of the robot but also safe operation. The state of the batteries needs to be tracked to know how much autonomy time is left to prevent the batteries from failing and the robot endangering the user (e.g. by blocking the user's way or not being able to execute a complete emergency handling scenario).

3.2. Sensor system

For being able to move safely and in a meaningful way through its environment and to interact with it, Hobbit requires an appropriate perception system for the following tasks:

- Map building and self-localization,
- Safe navigation (obstacle detection and avoidance),
- Human-robot interaction (user and gesture detection),
- Object detection and subsequent grasping.

For *map building and self-localization* seeing larger vertical planar structures such as walls or the faces of closets that are further away is desired. The "classical" approach is to mount a 2D laser range finder in the front of the robot that is scanning parallel to the floor. Since such laser scanners are currently still quite expensive, a more cost-effective solution is to use a depth camera facing parallel to the ground instead.

Safe navigation in domestic environments requires detecting obstacles up to the robot's height, and holes such as stairs leading downwards. A depth camera – when facing downwards – can be used to cover the space directly in front of the mobile platform and also left and right of it. Furthermore, some auxiliary range sensors need to cover the back of the robot to detect obstacles when moving backwards.

For *human-robot interaction* a Microsoft Kinect or ASUS Xtion Pro Live RGB-D camera is required to detect the user and to allow gestures as input modality. This hardware selection is based on technical requirements, user requirements, and the need for a low-cost solution. The recommended mounting height ranges from 60 to 180 cm, and the optical axis should be approximately parallel to the ground plane for detecting standing persons as well as their gestures.

Object detection requires seeing objects at heights of up to 90 cm (kitchen counter) and also on the ground. Table tops of "standard" tables (75 cm) and of couch tables (40 cm) as well as lower shelves are covered by this height range, too. Object detection requires RGB-D data. Support planes for target objects (e.g. table tops) need to be viewed from above to see at least parts of the horizontal plane. An appropriate mounting height for an RGB-D camera was identified as around 130 cm, which coincides with the user-preferred maximum height of the robot.

Taking the requirements listed above into account, the sensor system of the PT1 Hobbit was set up as the following: In the front of the mobile platform, at a height of 40 cm, there is a floor-parallel depth camera (ASUS Xtion Pro, see Fig. 3 top left). In the "head" of the robot, on a pan-tilt unit mounted at a height of 130 cm, there is an RGB-D camera (Microsoft Kinect, see Fig. 3 top right). The former camera is used for self-localization, the latter is used for obstacle detection (see Fig. 4), object detection, and grasping as well as human-robot interaction, that is based on depth-based human body observation and analysis (see Section 4.2). To be more compact, the Kinect was stripped off its original housing. An array of eight infrared and eight ultrasound distance sensors in the back

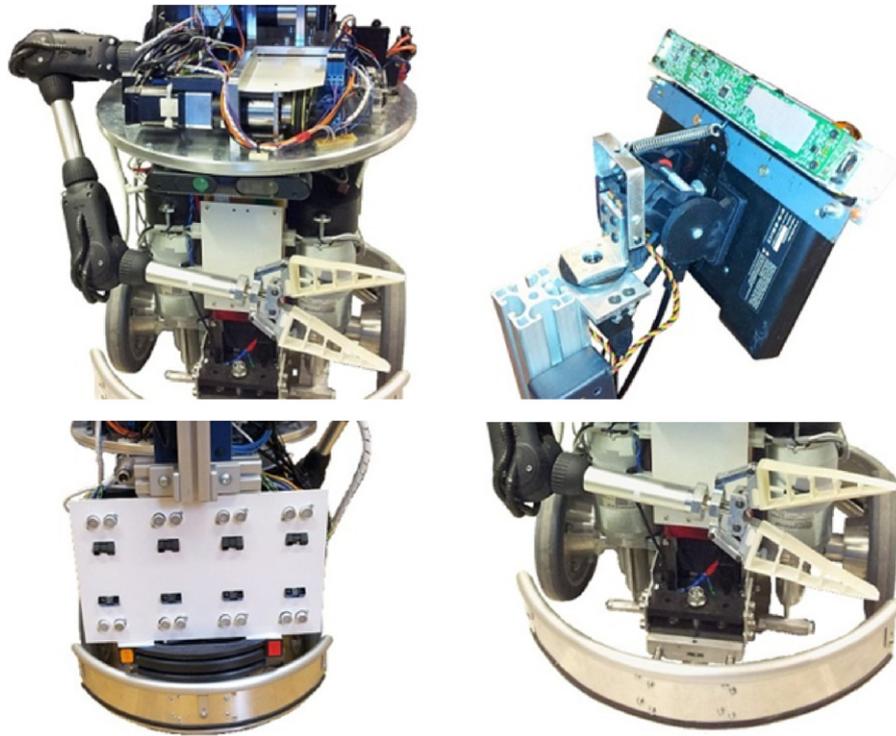


Fig. 3. The sensor setup of Hobbit: head RGB-D camera on a pan-tilt unit (top left), floor-parallel body RGB-D camera (top right), sensor array in the back (bottom left), and bumpers in the front and back (bottom right).

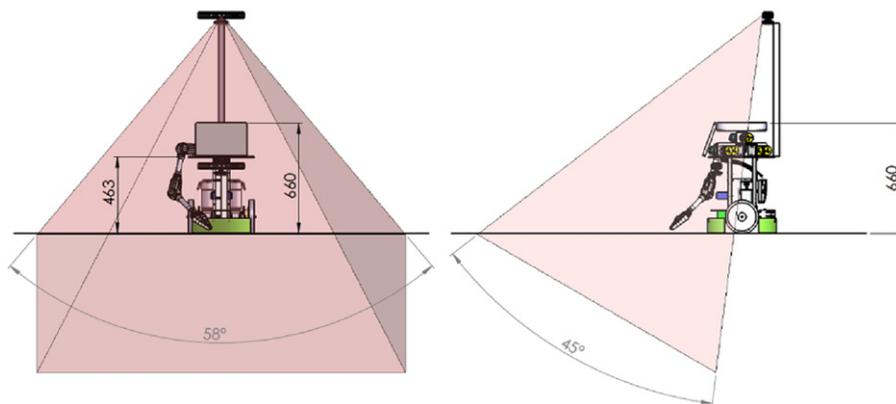


Fig. 4. Field of view of the top RGB-D camera when tilted downwards for obstacle detection.

of the platform allows for obstacle detection when backing up (see Fig. 3 bottom left). Additionally, and as last resort, there is one bumper in the back and another in the front of the mobile platform (see Fig. 3 bottom right). Finally, incremental encoders (odometry) on the two shafts of the drive motors allow measuring motion of the 20 cm drive wheels with a resolution of 70 μm per encoder tick. The encoder information serves as input to the speed control of the drive motors; it is also the basis for fine positioning.

Lessons learned

For a socially assistive robot that should autonomously work as a care-giver at home, it is of utmost importance that the individual components run robustly and are failure-tolerant, above all as a human is in the loop. Considering a failure probability of 1% per day for each of let's say 30 of the components, the robot would only run stable a whole day with a probability of $(1 - 0.01)^{30}$ equals 74%. For a whole week this probability is 12%, and for 3

weeks, the intended duration per user trial and user with the next prototype, the probability results to 0.18%. One solution to avoid a fast abrasion of hardware (which can subsequently lead to system failures) was considered for the design of the “head camera” of Hobbit. For PT1 we used a spring to relieve servos that were under constant load. The head design for the next prototype enables the servos to move the head with minimum moment based on an improved mechanical design balancing the weight of the head.

Furthermore, as with the overall size of the robot the optimal positioning of the head camera posed a challenge due to different requirements and constraints: (1) The more forward the camera is positioned, the better for obstacle avoidance, (2) the more the camera is in the back, the better for user recognition, (3) the higher the camera is positioned, the better for the overall perception, (4) the lower the camera is mounted, the better is the resolution for grasping on the floor. For PT1 we decided in favor of option two, as a robust user recognition lies in the focus for our care robot and failures in this area are crucial in terms of acceptance.

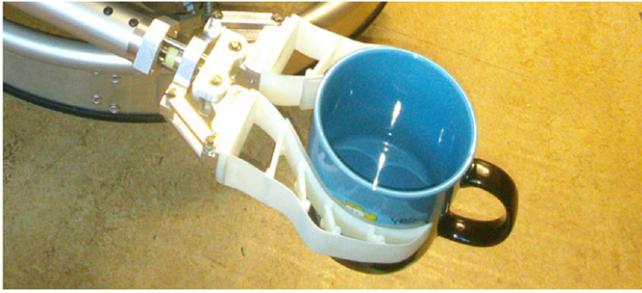


Fig. 5. Fin Ray Effect: gripper automatically wraps around object without additional actuation.

3.3. Arm

The design goal for the arm was to use an affordable, lightweight component with a human-like design. The so-called “IGUS Robolink” [16] has a freely configurable arm length and up to 5 degrees of freedom. Due to its modular design it is used to fulfill these requirements. The arm has a weight of 1.5 kg, payload is 500 g additionally to the gripper, and each joint is driven by tendons. This has the advantage that the motor drives can be mounted on the Hobbit platform. The control of the arm system is done by the XPC using TCP/IP commands which are received by the motor controller.

Lessons learned

During the PT1 user studies it became apparent that the arm reachability was too limited due to the 5 degrees of freedom. Especially when grasping objects from the floor, the platform had to be positioned very accurately to enable the arm to grasp the object, which was time-consuming and boring for the user. Using a 6 degree of freedom arm for the next prototype (PT2) will increase the reachability and the speed of the grasping from floor process, due to the fact that the fine positioning of the platform does not have to be as accurate when grasping an object from floor.

3.4. Gripper

The manipulator consists of a gripping system based on the FESTO “Fin Ray Effect” [17]. More specifically, the fingers mechanically wrap around any object shape without additional actuation (see Fig. 5). The assembled fingers on the manipulator can adjust themselves to the object by means of the “Fin Ray Effect”. In combination with a simple open/close mechanism, a variety of objects with different shapes (like mugs, keys, pens, etc.) can be grasped. Due to the slip-proof materials used for the fingers, it is possible to reliably grasp objects.

Lessons learned

Robustness wins over specific adaptations to all use cases. If the main functionality is reached, a well-tested solution (Fin Ray gripper from the arm manufacturer) is cheaper, easier to access, and more reliable than a self-developed and 3D-plotted gripper skeleton with Fin Ray fingers from Festo.

3.5. Multimodal user interface

The multimodal user interface (MMUI) consists of a Graphical User Interface (GUI, see Fig. 6) with touch, Automatic Speech Recognition (ASR), Text to Speech (TTS), and Gesture Recognition Interface (GRI). It provides web services (e.g. weather, news, RSS feed), video phone service (based on previous successful projects [18]), serious games, control of a manipulator, access to an Ambient Assisted Living (AAL) environment, and emergency call features.

Hobbit makes use of the MMUI to combine the advantages of the various user interaction modalities. The touch screen has strengths

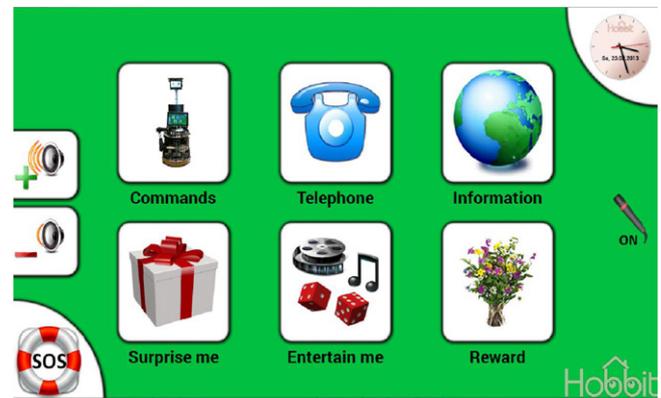


Fig. 6. Example of GUI used for Hobbit PT1.

such as intuitiveness, reliability, and flexibility for different persons and different sitting positions, but requires a rather narrow distance between user and robot. ASR allows a wider distance and can also be used when no free hands are available, but it has the disadvantage of being influenced by the ambient noise level, which may reduce recognition performance significantly. GRI allows a wider distance between the robot and user as well and also works in noisy environments, but it only operates when the user is in the field of view of the robot.

The touch screen is mounted in an approximately 45 degrees angle in a slightly protruding position which is a design compromise to avoid complex mechanics for the tilting. The MMUI is mounted on a mechanical slider so that it can be pulled towards the user for the most ergonomic position. Hobbit also provides a second small display on its “head” in order to present facial expressions (emotions). Additionally, we aim at presenting affective states of the robot towards the user, e.g. by different ways of navigating the robot (approach trajectory and speed or moving the robot slowly when recharging of its battery is needed).

The GUI is structured into several thematic menus with big clearly spaced icons taking into account the needs of older users and the operation from free standing. Immediate multimodal feedback (written text and text-to-speech) is provided for every command activation, which can be done by any of the input modalities.

The GUI was developed following an established interface paradigm taking into account standards of user interface design for older adults [18], therefore explicit usability testing was not done before the first user trial, but it was an integral part of the study presented in Section 6, as users had the option to freely choose between the modalities and were afterwards asked to rank the possibilities with respect to perceived usability. The interaction with Hobbit is always initiated by calling the robot, which can be done with three different input modalities, which are differently suitable depending on the distance between the user and Hobbit. It can be done either by

- a wireless call button (far, from other rooms),
- ASR and GRI (2–3 m),
- touchscreen (arm length).

The speaker-independent ASR and TTS are offered in four languages: English, German, Swedish, and Greek. Contemporary ASR systems work well for different applications, as long as the microphone is not moved far from the speaker’s mouth. The latter case is called distant or far-field ASR and shows a significant drop in performance, which is mainly due to three different types of distortion [19]: (a) background noise, (b) echo and reverberation, and (c) other types of distortions, e.g. room modes or the orientation of the speaker’s head. For distant ASR currently no off-the-shelf solution exists, but acceptable error rates can be achieved for distances up to 3 m by careful tuning of the audio components and the ASR engine [20].

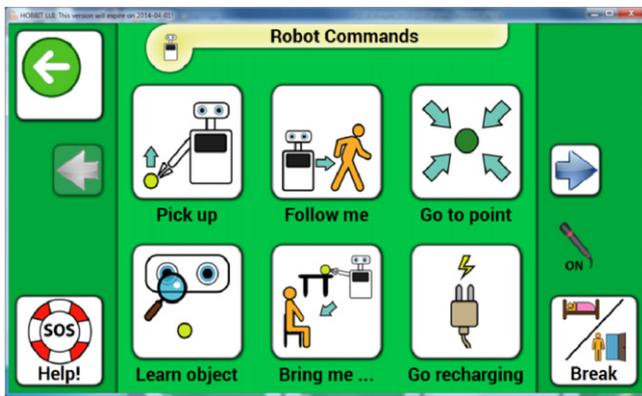


Fig. 7. Reworked GUI for the next Hobbit prototype.

Lessons learned

During the PT1 user studies it could be observed that the round corner icons of the GUI (SOS and clock in Fig. 6) were not always identified as buttons by the users and therefore were changed to a rectangular design comparable to that of the other buttons. In Fig. 7 the new icons as designed for the next prototype are depicted, including icons for new robot tasks that should also be integrated, such as sending the robot to its charging station. Moreover, it turned out that the option of extending the MMUI in a comfortable ergonomic position for the user, was hardly ever used by participants, even though they were reminded of this option. As a consequence, the mounting of the touchscreen for the next prototype will be changed to a fixed, protruding position. Furthermore, while initially the user was approached from the front, what can be considered as natural for standing persons, it can also block the way in case the user is seated. Hence, it is preferable to approach the user from the right or left side while seated, which is more positively experienced by the user [21]. Additionally, this position offers the advantage that the robot is close enough for the user to interact via the touchscreen, but it does not invade the personal space of the user (limiting movement space or restricting other activities such as watching TV).

4. Components

In order to fulfill its tasks as a care robot, Hobbit must be able to safely navigate in a domestic environment, detect and track humans, recognize gestures, and grasp objects. In this section we describe the major software components of Hobbit and the algorithms used to achieve the required functionality.

4.1. Navigation

To enable safe navigation in domestic environments, Hobbit must be able to generate a map of the environment, localize itself, detect obstacles, and find a drivable path (including local navigation and fine positioning). For the overall envisioned Hobbit robot the following navigation strategies need to be considered:

- The robot can be called (see Section 5.2) by the user and then autonomously navigate to this pre-defined location.
- The robot should be able to follow the user autonomously into a specific room, after the user indicated the room (in every room a pre-defined location is specified).
- The robot should autonomously patrol through the flat after three hours without any user activity to check if the user is ok (after three hours of an undetected fall, the risk of hydrating and severe follow-up health threats increases dramatically).

- The robot should patrol the apartment to search for clutter on the floor on two pre-defined times a day.
- The robot should not follow the user all the time and should not enter bathrooms and toilets due to privacy reasons; here the users will have additional emergency buttons in case of a fall.

This section describes the approaches used for SLAM-based map-building of the environment and subsequently for self-localization based on AMCL.

Map building

Many processes in Hobbit depend on the estimated pose of the mobile platform in relation to its environment. A (metric grid) map of the environment serves as basis for self-localization. In the first prototype of Hobbit we refrained from using the full 2.5D information computed from the depth images of both RGB-D cameras for mapping and self-localization. Instead, we reduced the 2.5D data of only the bottom RGB-D camera to a ground-parallel “virtual 2D laser scan” along the optical axis of the camera. This allows to use standard algorithms initially developed for 2D laser range finders with an RGB-D camera. Such algorithms are available and ready to use in ROS, and thus enable immediate practical testing. Furthermore, working with reduced amounts of data allows fast processing to meet real-time constraints even on low-power PCs.

From the floor-parallel depth camera we compute a virtual 2D laser scan. The horizontal field of view is that of the depth camera (about 60°) and the maximum range is 5 m. The map is generated in the traditional SLAM fashion: The robot moves through an environment and incorporates the measurements of the virtual laser scanner and of odometry.

We use the “gmapping” algorithm proposed by Grisetti et al. [22] for mapping, since it is able to cope with uncertain odometry data and short-range depth data. Gmapping uses Rao-Blackwellized particle filters for map generation. Each particle represents a hypothesis of the map itself. The particles are only updated when the robot has moved a specific distance, in our case 0.25 m. Due to the rapid decrease of depth resolution of RGB-D cameras with increasing depth, only virtual laser scans up to 4 m are used. Scans that report a higher distance will be only used for maintaining free space information. The scans are aligned under consideration of the non-holonomic constraints with a simple ICP approach [23] for each particle while the robot is moving. We use a minimum number of 500 particles and a maximum number of 2000.

During the setup phase of the Hobbit system, an expert will execute mapping due to the technical nature of this process. The expert has to take care that the map is consistent. It is necessary that all movable objects are removed from the map in a manual post-processing step, since those objects can easily change position and thus must not be used during self-localization. We are aware that this poses a limitation for the PT1 platform, as it cannot autonomously explore new rooms, however, in the case of a service robot for elderly we consider it relevant that an expert generates the map case to guarantee safe and reliable navigation.

Our experiments have shown that it is necessary to remove artifacts caused by the mapping process, e.g. single-standing cells that are occupied. Those artifacts can prevent the path planner from finding a suitable path later on. Fig. 8 illustrates the result of mapping an office environment; the resolution of the map is 5 cm × 5 cm per pixel. Although mapping must be done only once prior to acquainting the user with Hobbit, we will investigate possible approaches to automate many of the steps that currently require an expert.

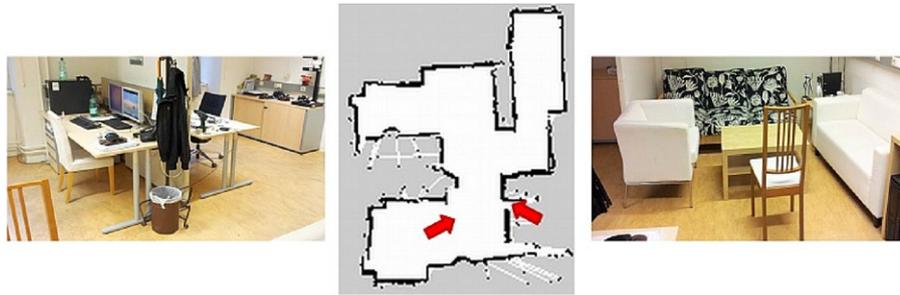


Fig. 8. Center: map of an office built from the virtual 2D laser scans using SLAM. Left and Right: Views when standing in the office at the positions of the red arrows and looking in the respective direction.

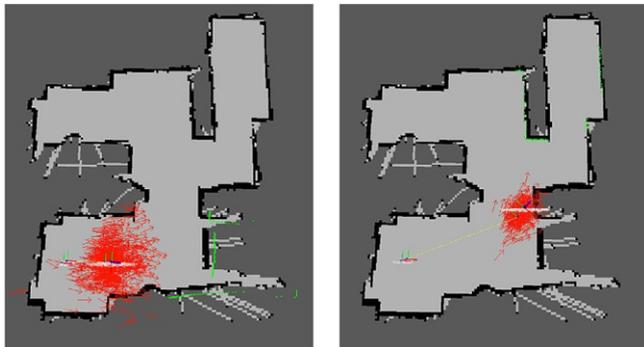


Fig. 9. Self-localization using AMCL with virtual 2D laser scans and the SLAMed map. The initial large uncertainty of the pose (left) grows smaller as the robot moves through the environment and updates its pose estimates.

Self-localization

Self-localization of the mobile platform is done using the traditional “Adaptive Monte Carlo Localization” method, short AMCL, originally proposed by Thrun et al. [24]. The robot pose is represented as a set of multiple hypotheses with respect to an a-priori known map. AMCL incorporates sensor data from the virtual 2D laser scanner and from the odometry of the mobile platform. It allows both pose tracking and initial localization to cope with the “kidnapped robot problem”.

Fig. 9 shows two stages of self-localization using AMCL in ROS, the SLAMed map (using gmapping), the virtual 2D laser scan and odometry data. The red arrows show hypotheses for the platform pose, and the green dots represent the virtual 2D laser scan. The platform was initially only roughly positioned on the map origin so that the scan points do not match the map very well. After moving a few meters, platform pose hypotheses form a denser cluster and the scan points match the map reasonably well.

Obstacle detection

When the top RGB-D camera is tilted downwards, it covers the space in front of and aside the robot. We apply a “v-disparity”-based approach [25] to the disparity images provided by the camera in order to detect and remove the floor. The remaining disparity data represents obstacles from which we compute a second virtual 2D laser scan with a horizontal field of view of about 150° and a maximum range of 2 m.

Fig. 10 shows an example result of the approach.

Path planning

The objective of the path planner is to seek for a possible path from the current position of the mobile platform to a given (task-related) destination. It is assumed that the environment and platform pose are known at any time through self-localization. We use the search-based planning (SBPL) algorithm for robot path planning. Proposed by Phillips and Likhachev, SBPL [26] differs

from traditional A* methods. Originally developed for robotic arms, it can also be applied for mobile robot motion planning. Instead of planning a path with the shortest Euclidean distance, SBPL uses predefined motion primitives that are kinematically feasible. Planning is done in x , y , and θ dimensions, resulting in smooth paths that take the orientation of the robot into account, which is especially important if the robot has non-holonomic constraints. Plans are found using the AD* planner [27], a variant of the A* algorithm. SBPL runs in real-time and needs approximately 300 ms to find a path, depending on the length of the path.

Lessons learned

In narrow passages of cluttered domestic environments – due to the small field of view of the depth camera – it is not always possible to extract useful features for self-localization. In order to bridge such a period without good features without losing self-localization, good odometry is required. Moreover, to provide an additional source for estimating the motion of the robot, an IMU can be used. The decision of a depth camera instead of a laser was primarily made to keep the robot affordable. However, it turned out to be beneficial to use 3D data to generate 2D data for virtual laser data. Using this data generation, we can handle protruding table tops and other objects sticking out at any height while 2D lasers would fail. With the two-camera solution we can also assure that we see below tables or chairs to walls, which is helpful for localization, while the top camera guarantees that the immediate front of the robot is supervised. This considerably adds to the safety of the robot navigation. For PT1 autonomous recharging was not implemented, however, as this is clearly a crucial functionality for a domestic service robot, it will be integrated in PT2.

4.2. Human detection and tracking

Vision-based human observation [28] encompasses a set of fundamental perceptual mechanisms that socially assistive robots should support. The first approach of the corresponding framework for Hobbit and the developed perceptual competences are presented in more detail in [29].

Based on recent advancements in the field of computer vision for 3D human observation and the availability of low-cost depth-aware sensors, like MS Kinect [30], algorithmic techniques for human pose estimation, recognition and tracking using depth visual data (e.g. [31,32]) have become computationally cheap and readily available at real-time performance on conventional computers. We exploit the opportunity to set Hobbit capable of supporting a rich set of vision-assisted competences regarding both full-scale observation of a human (full 3D human body detection, localization, tracking) and partial, close-up observation (hand/arm/face detection and tracking). Moreover, additional vision-based competences rely on this module of the platform, such as gesture and activity recognition, vision-based emergency (fall) detection (see Section 5.7), etc. To achieve these goals, we rely on RGB-D visual data acquired by the “head” RGB-D camera of the robot.

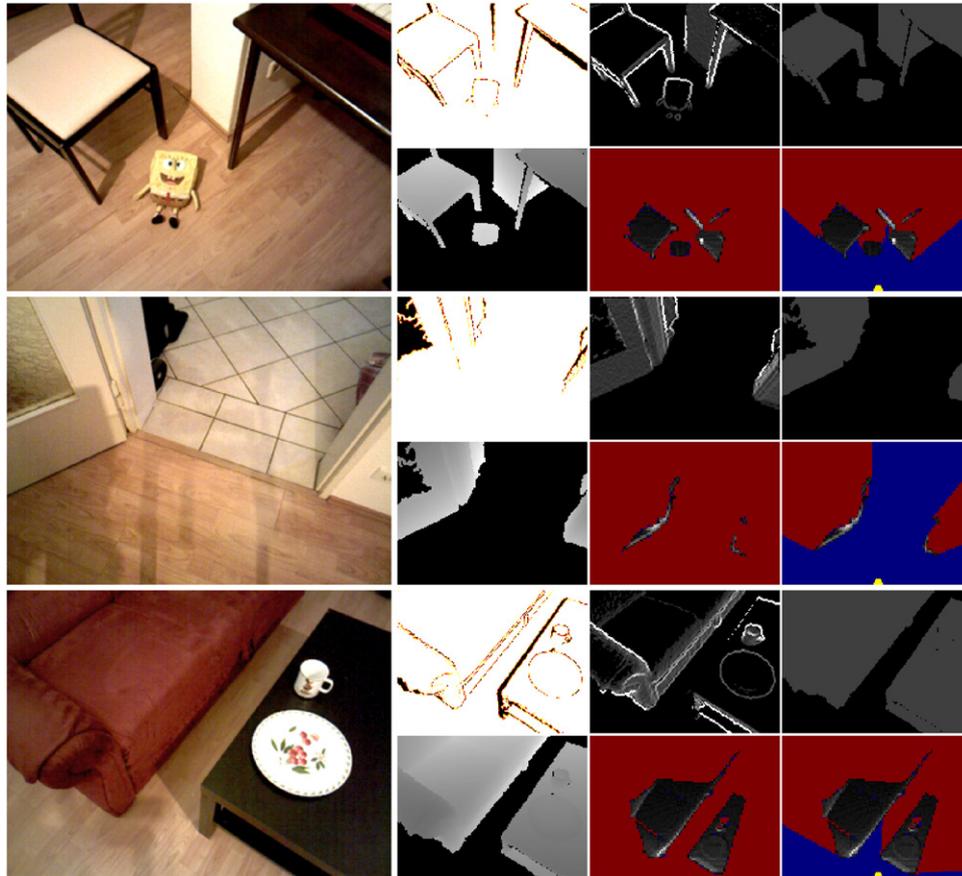


Fig. 10. V-disparity-based floor detection and removal, and generation of a virtual 2D laser scan for obstacle detection. Three example scenes are depicted. The left column shows the RGB images of the scenes. The processing results for each scene—top rows: confidence map of the reduced-resolution disparity image, biased gradient magnitude, mask for disparity values outside the calibrated range for floor points; bottom rows: reduced-resolution disparity image without floor points, cell grid (red means free, gray to white means occupied), and virtual 2D laser scan (blue indicates the laser beams emitted from the virtual laser scanner indicated by the yellow dot). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

On a technical level, 3D scene segmentation and foreground detection is initially performed for each acquired depth frame, while the robot is moving or operating in place. Vision-based extracted information regarding the scene background, foreground as well as 3D floor-plane estimation are computed. Subsequently, user detection and segmentation are performed to identify human bodies among the detected foreground depth-based objects in each frame and track them in the scene across frames providing a label map and unique persistent user IDs for each pixel. The latter process is closely related to 3D human body pose estimation and skeletal tracking that are also applied as higher level processes towards human body observation. Body pose estimation relies on a 3D skeletal model that is comprised of 15 main body joints and 10 body limbs, as reported in [33,34]. For each frame the detected depth-based pixels assigned to a human body are fed to the body pose estimator to fit the 3D skeletal model, see Fig. 11. Moreover, 3D skeletal tracking is performed to obtain seamless fitting of skeletal joint-related information across frames. Practically, a readjustment of the skeletal body model is performed in order to track the 3D positions/orientations of basic body limbs and joints across frames.

Hobbit is capable of detecting and tracking both a standing (moving or still) or a sitting user. In the first case, a full skeletal model is employed as described above, whereas a sitting user is detected and tracked based on a truncated upper body version of the described skeletal model (see Fig. 12(a)).

Moreover, face detection and 3D head pose estimation [35] are supported in order to enrich the vision-based extracted information provided by the system, as illustrated in Fig. 12(b). The face detector performs as a stand alone module providing



Fig. 11. Vision-based 3D human observation. 3D human body detection is performed identifying a human body (green color-left depth image) and segmenting it from the background in the observed depth scene. Subsequently, 3D pose estimation is applied on the segmented body data to fit a 3D skeletal model of the human body (right color image) and estimate the 3D position and orientation of 15 body joints (white dots) and infer the body limbs (red/green lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

reliable information to user detection and segmentation modules while it can also be bootstrapped by the latter in case of strong detection confidence of human body, eliminating false positives, in case multiple or no face detection results are obtained.

Lessons learned

The performance of 3D user detection and tracking during the task performance of participants was challenging. In many cases, the performance of human body detection and pose estimation for a sitting user was deteriorated due to occlusion by the chair, table or couch for specific poses of the user. In such cases, face detection

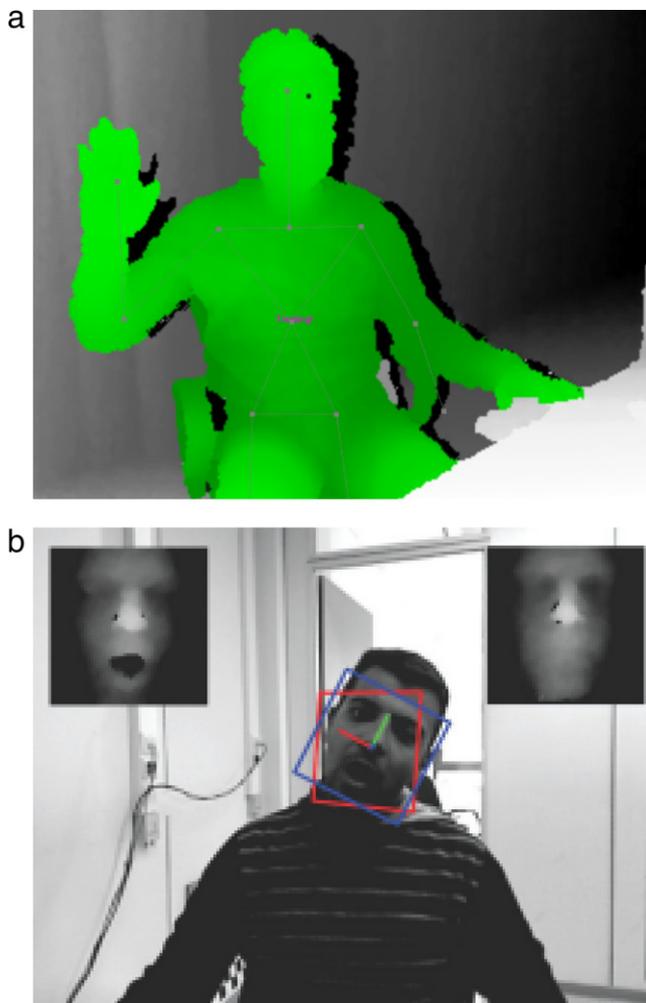


Fig. 12. In (a) 3D human body detection and tracking is performed for a sitting user (green color) segmenting the relevant pixels from the background in the observed depth scene. 3D pose estimation is applied to the data to fit the 3D skeletal model of human body (gray lines representing the main body parts of the skeletal model). In (b), face detection and 3D head pose estimation [35] are demonstrated for a sitting user, based on RGB-D data acquired by the upper (head) sensor of the robot. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

served as a fall-back to localize the user and act according to the executed task. Moreover, the performance of the face detector is deteriorated in cases when a user wears glasses or a hat, which are known issues for face detectors. Our intention is to further improve user detection, tracking, and the face detector based on the noticed failures and false negatives during the user trials.

4.3. Gesture recognition

A gesture recognition interface (GRI) has been developed as part of the MMUI of Hobbit (see 3.5), to provide an additional input modality based on the interpretation of physical arm/hand gestures to robot commands. This type of interaction provides an intuitive control modality for human–robot interaction that aspires to facilitate the communication of older adults with the robot.

In order to realize this type of interaction, a number of predefined gestures are supported by the GRI, as a physical action-based vocabulary. Gestures are defined as a series of postures performed using the upper body parts within a time window of configurable length. The supported gestures can be described as actions consisting of intermediate predefined postures of upper body parts. During interviews conducted with elderly prior to the user trials, their

preferences, intuition and physical convenience were recorded and evaluated in order to consider the predefined gestural vocabulary and the correspondences to robot commands. The following physical actions were validated as appropriate for usage in the GRI.

Each of the gestures consists of two or three primitives, as composite actions. The “Raise hand” primitive is always preceding any of the following combinations. It corresponds to the physical movement of rising each of the hands at the height of the chest or the shoulders with open palm towards the camera (see Fig. 13(a)). A hand tracking method is initiated in the background each time any of the user hands is raised as described. Subsequently, hand trajectories are recorded towards supporting gesture recognition. The list of gestures includes the following actions: (a) “Push towards the robot”, (b) “Keep palm steady and Swipe up or down or left or right”, (c) “Move cyclic”, (d) “Raise both hands and Cross wrists” and (e) “Keep palm steady and Extend the other arm to point” (Pointing gesture).

Given that the user is within the field of view of the “head” robot camera, he/she can perform any of the following gestures to intuitively initiate a specific robot command/task to be executed by the robot upon successful recognition. “Help-the-user” robot command is triggered after a “Cross hands” gesture is performed by the user and recognized by the robot, see Fig. 13(a) “Pick-up-object” command is also supported by performing the pointing gesture (extending the arm to point any location-object in 3D space) in order for the robot to pick up an unknown object from the floor. An illustration of the “Pointing” gesture is provided in Fig. 13(b). Moreover, answering “Yes/No” in human–robot dialogs is also feasible using GRI by mapping any of the Swipe up/down and Swipe left/right to affirmative and negative answers, respectively.

The hand tracking and gesture recognition algorithms used in our implementation of the described functions relies on the open source OpenNI framework [34] and the middle-ware library NITE [36].

Lessons learned

The user studies revealed that many participants found it difficult to adapt and perform the designed gestures, despite the selection of intuitive physical actions as gestures, even though appropriate training by demonstration took place on site during the user trials. Moreover, in many cases participants did not recall the set of gestures during the interaction with the robot.

Regarding the GRI, a new methodology will be introduced in order to enhance detection and tracking of hands, but most important is to extend its functionality to the fingers of the users. Thus, a new set of finger-based hand postures and gestures will be designed to replace the available robot commands. Moreover, a learning mechanism will be introduced aspiring to further explore adaptability and customizability of actions performed by the users, loose the required fidelity of execution for an action to be recognized, and therefore enhance the recognition performance. In other words, the user will need to only approximately perform any of the predefined gestures, while an online learning procedure will customize the recognition algorithm accordingly to adapt to the specific way the individual performs those. In addition, an updated system will also incorporate the ability for online definition, configuration, and learning of new gestures and postures that the user may desire to introduce to the interface and assign them to any of the existing robot commands. Therefore, the user may adapt the interface according to personal, daily habits, physical capabilities, and cultural differences in using body language.

4.4. Grasping

For grasping unknown (Section 5.4) and known (Section 5.6) objects, first the dominant horizontal plane, e.g. floor or table surface is detected and the corresponding points are eliminated. Clustering the remaining data delivers point clouds of objects. A

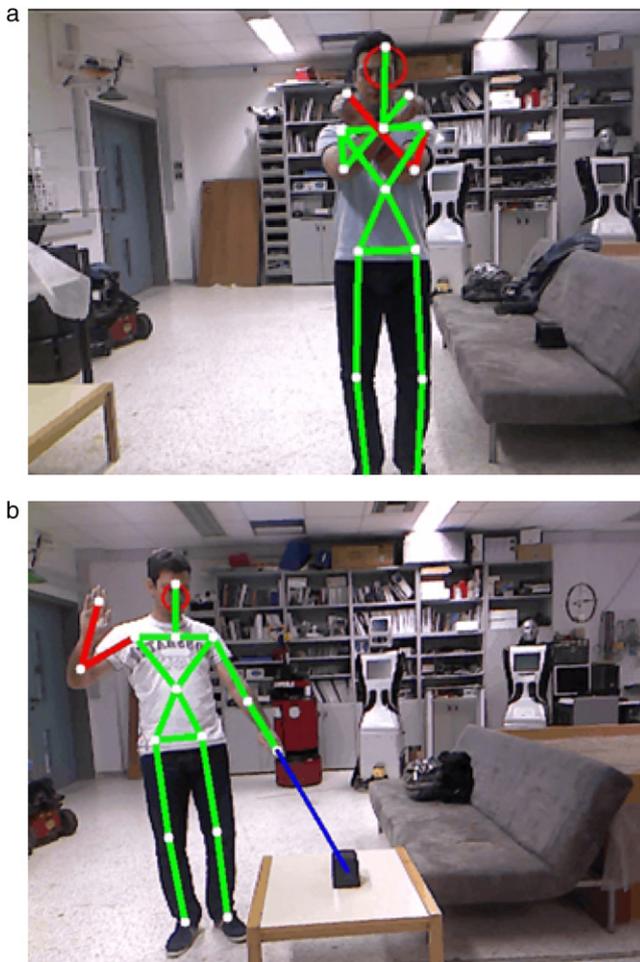


Fig. 13. In (a) the “Help” gesture is demonstrated, crossing both wrists at the height of the chest. In (b) the “Pointing” gesture is performed. The user points to an unknown object in 3D space. The blue line indicates the calculated 3D direction specified by the extended arm towards an object of interest on a table. In both images the skeletal model of the standing subject is also rendered in green–red lines for the main body limbs and white dots for the joints. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

procedure tests if a point cloud is suitable for grasping, taking into account object size and object position. If the number of points is below a threshold value ($n = 200$ for the first user trials) or above a maximum number of points (to rule out objects which are too big for grasping respectively transporting), point clouds are not used as grasping targets. Similarly, in the case an object is detected at a position where grasping will probably fail it will also not be grasped; for example when the robot detects an object below a table (maybe the table leg) in the clean floor task (see Section 5.4 for details). To eliminate the latter case, we compare each point cloud position with the map recorded for navigation. In this map we define graspable areas. In the case of grasping known objects, grasp point detection is limited to the point cloud identified as the desired object. Grasp points are calculated with Height Accumulated Features (HAF). For a thorough description of this method we refer to [37] where it was used to unload a box of unknown items and to [38] where single standing objects as well as items in a pile of objects were grasped. This method calculates feature values based on height differences on the object and uses these feature values to detect good grasp points by applying a grasp classifier that was trained using Support Vector Machines (SVMs). A heuristic selects the best rated grasp point from all detected grasp points. Path planning for the robot arm including

obstacle avoidance is performed with the simulation environment OpenRAVE [39].

Lessons learned

The limited arm kinematics related to the 5 DOF arm mentioned in 3.3 often makes it impossible to approach an object on a defined straight path keeping the desired hand orientation. The limited arm kinematics was compensated by an accurate fine positioning of the platform, using the additional 2 DOF of the robot. The iterative fine positioning step was time consuming and will be replaced for PT2 by a more flexible calculation of the arm position when grasping. In the PT1 user studies a grasp was accepted as successful if the object was moved from its original place after the arm moved out of the view of the head camera to a defined position at the side of the robot. For faster operation/grasping (as required by the users) a method is implemented for PT2 that checks a successful grasp after the gripper has closed, taking into account the deformed gripper fingers when an objects was actually grasped.

4.5. Object learning and recognition

For learning and recognizing objects, 3D shape descriptors [40] are calculated from views of the object, coming in the form of RGB-D data from the Kinect camera in the “head” of the robot.

In the learning stage, objects are placed on a turntable (Fig. 15(b)) and while rotating the arm each new view of the object is stored in a database [41] and later matched against in the recognition phase using random forests [42]. The re-training of the forest is done immediately after new views of an object are added to the database. This system design allows great flexibility, e.g. a standard set of object classes can already be present before the user teaches the robot specific objects.

In the recognition stage when the robot is sweeping for objects by panning the camera, objects on flat surfaces (e.g. tables and on the floor) are recognized on the fly and reported back to the search algorithm.

Lessons learned

3D object classification and recognition on a robot has to deal with greatly varying working and sensing distances. Learning objects on a turntable at a distance of 80 cm and recognizing these objects on tables (1.0–1.5 m) and on the floor (1.5–3 m) is challenging given the different resolution and noise level of objects at these distances. Clutter in the environment is a major performance factor and has to be considered in the training phase by including a special clutter-class in the classification algorithm.

Reporting false objects as well as not finding objects will not increase confidence of the user in the robot. Hence, single-shot classification should be replaced by a more sophisticated approach where the camera is centering on object candidates for validation and thus eliminating false classifications at image borders through cut-of-objects. In a second step, the robot should move closer to the object for repeated recognition under different approach directions for increasing the detection rate. This needs to be done in cooperation with grasp planning to position the robot ready for grasping.

From the user side, recognition of small objects (ear-ring, glasses) was requested but this is currently out of scope as the camera is too far from the floor/table and offers too low resolution for this task. High resolution 2D image recognition algorithms, novel 3D sensors, or bringing the camera closer to the floor/table could address this user request.

5. Robot tasks

In the development of the Hobbit robot we followed a user-centered design approach, meaning that the system design should be adapted to the needs and expectations of the potential user

group. Contrary to participatory design, where design decisions are made together with representative target users, we followed an interdisciplinary approach where final system design decisions were made together with experts from robotics, gerontology, and industrial design. As primary target user group of Hobbit we considered older people of age 70+ who may need assistance for living at home in the near future. In the following we will give a short impression about the user requirement phase and how we derived the first set of functional requirements. Subsequently, we will describe the main tasks that were implemented for the first laboratory studies with the Hobbit PT1. The following sections are not meant to be a complete list of all Hobbit tasks, but an overview on tasks that were implemented and subsequently studied with real target users in a controlled experiment.

5.1. User requirements

The starting point for the development of Hobbit was prior expert knowledge on two aspects namely that (1) fall prevention and (2) emergency detection and handling are crucial for aging in place. From here, we conducted expert workshops with gerontologist, followed by workshops with potential end users (in Austria [43] and Sweden [44]). These workshops then served as a basis for questionnaire studies and subsequent interviews in Austria, Sweden, and Greece. All in all, in terms of robot tasks that users were interested in according to the requirement studies, three main categories were expressed repeatedly in different data sets. (1) Aspects of communication and social contact (e.g. speech interface, assist social relationships, provide entertainment), (2) Help with household activities (e.g. search for objects, fetch-and-carry objects from the floor and shelves, water flowers) (3) Care activities (e.g. detect emergencies, reminder functions, standing-up and walking support). Among these, communication/talking to/with the robot, bringing and searching objects, and emergency functionalities were of highest interest for the users who participated in the assessments. Additionally, design and appearance issues have been surveyed and analyzed, which give further input to the development of a robot system that fulfills usability expectations and is also accepted by its target group.

Our requirement studies as well as the research of others [45] indicate that older adults mainly expect assistance in various household maintaining tasks but also to improve social connectedness. However, state-of-the-art robots can often not meet users' expectations. Therefore, the idea of Hobbit is that the robot performs meaningful tasks for the user and cooperatively performs them together with the user where it needs help (e.g. learning a new object). In this way, older adults can remain active and the robot only compensates their limitations by assisting the task, such as picking up something from the floor. We called this interaction paradigm Mutual Care, more details on it can be found elsewhere [46,47].

5.2. Call Hobbit

To facilitate easy calling of the robot to a specific place when user and robot are not in the same room, self-powered (by piezoelectricity) and wireless (EnOcean standard) call buttons are used as part of the Ambient Assisted Living (AAL) environment. Such stationary buttons can be placed e.g. near the bedside, in the kitchen or in the living room wherever the user is expected to be frequently. When the user presses the call button, the robot will directly navigate to the known place so that it brings itself into a closer interaction distance and pose relative to the user which is suitable for touchscreen, ASR, and GRI operation.

For the call buttons and the sensors of the AAL environment tests were performed in an AAL lab [48] (see Fig. 14) with different

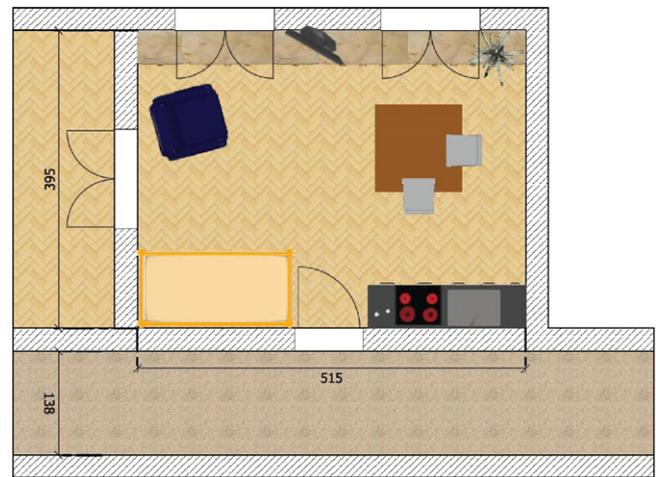


Fig. 14. Map of AAL laboratory used for robot trials in an intelligent environment.

zones modeled similar to a realistic home environment. Up to our knowledge the interaction can be initiated easily with any other care robot for older adults. We consider this input modality as substantial aid to ease the interaction for older adults.

5.3. Introduction phase—user specific setup

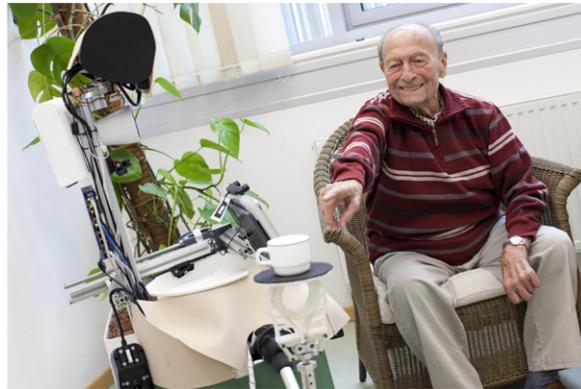
The default settings of Hobbit are a good starting point for most users. To allow for individual adaptation a so-called Initialization Script, which runs upon first introduction of the robot to the user and later on user request, guides the user through a set of questions. The user is asked for preferences on sound volume and robot speed as well as gender of the speech output voice; the user is invited to try out speech, gesture, and screen input and can give the robot an individual name it will answer to. The final prototype will also allow to configure the individual behavior settings, such as different robot personalities (more companion-like or more machine-like) and proxemics parameters. The selected values are directly demonstrated during the process to give the user immediate feedback. We know from HRI literature how important the adaptation of robots to user preference is (see eg. [12]), but are not aware of other care robots that allow this type of personalization.

5.4. Clear floor

Triggered by voice or touch screen, Hobbit is capable of cleaning the floor from objects laying around. The robot first detects the floor as the main horizontal plane and eliminates all points corresponding to the floor and clusters the remaining data to objects. The use of lower and upper limits for the size of point cloud clusters enables the elimination of objects that are too big (too heavy to be lifted by Hobbit) or too small (sometimes the floor is slightly rippled which leads to an insufficient ground floor elimination). The robot uses structural information about the domestic environment gathered during mapping phase to eliminate objects that are unlikely or impossible to grasp. As an example, if an object cluster is placed at the position of a wall, Hobbit does not try to grasp it since it is probably a segmented part of the wall. If Hobbit finds an object on the floor, it moves towards the object, grasps it and brings it to the user. If no graspable object was found, Hobbit changes its position and searches again on the floor until the floor is emptied or a stopping criterion is fulfilled (e.g. time spent on the task or the number of tries exceed predefined thresholds).



(a) Clear floor.



(b) Learn object.



(c) Fall detection.

Fig. 15. User trials: Hobbit brings an object from the floor to a user (left, for completed user trials Hobbit did not use human pointing directions, this feature will be included for coming user trials in real apartments); Hobbit learns a mug; Hobbit detects a user fall and calls for help (right).

5.5. Learn new objects

To learn a new object, the robot has to see the object from multiple views and – for objects like a pack of aspirin which can be found in any pose – from upside-down. To achieve this, the robot uses a small turn-table (see Fig. 15(b)). The turntable is designed in a way that the gripper can hold it in a defined pose. The user is asked to put the new object onto the turntable. The robot then slowly rotates its arm and captures views of the object while its turning. After a full rotation, the user is asked to put the object upside-down to now learn the previously unseen sides of the object. The turntable rotates again and views are captured and stored. Now the user has the choice of teaching the robot another object or remove the current one. After finishing learning, the newly learnt object can be used in other tasks such as “Bring Object”.

We are aware of the fact that this way of learning limits the size of potential objects to learn. However, it also offers the advantage of close cooperation between the user and the robot, which is desirable for elderly care robots to make the user still feel in charge and needed [45].

5.6. Bring object

Users can command Hobbit to search and bring a previously learnt object. For objects often needed by the user, Hobbit saves the typical object location, (e.g. the kitchen table). Hobbit first searches at this place, grasps the object, puts it on its tray and brings it to the user. To simplify scenarios during user trials, we used predefined arm positions for grasping. After the searched object was found, Hobbit places itself in a predefined position with respect to the object and executed a fixed arm movement to grasp the object.

5.7. Fall detection and help function

Fall detection of older adults is a major health risk and several systems have been proposed for the automatic early detection and prevention of such emergency cases [49–51]. To this end, fall prevention and detection is a crucial functionality that Hobbit is designed to support in order to help elderly users to feel safe in their home, by identifying body fall/instability or the user lying on the floor and handling emergency events appropriately.

A fall detection function is continuously run by the system as a background process. In the first place, it is able to recognize abrupt motion of a detected and tracked human body that indicates instability or an ongoing fall. Additional events can be captured as emergency alerts by the help function based on the GRI and ASR modules of the system (see 3.5), such as a predefined emergency



Fig. 16. The GUI page that provides access to entertainment options.

gesture or voice command, with which the older adult can ask the robot for help.

On a technical level, body fall detection is based on information related to 3D body skeletal tracking that relies on visual data acquired by the “head” camera of the robot and the 3D human observation functions (see 4.2). A 3D bounding box of the detected human body is calculated for each frame and emergency detection is performed by analyzing the length, velocity, and acceleration of each dimension of the calculated 3D bounding box in time. Fig. 15(c) illustrates a relevant case during lab trials. Our methodology bears some resemblance to the method in [52].

In case of a detected emergency, a subsequent part of the help function is triggered, namely the emergency handler, that enables the robot to safely approach the user’s position, initiate an emergency dialog to calm him and perform a phone call for help, if necessary.

5.8. User entertainment and social connectedness

Hobbit offers entertainment by allowing the user to listen to favorite music, watch videos, and play games. For the first prototype (only) some examples were integrated in the menu of the GUI (see Fig. 16). For the final prototype these will be extended adding also access to social media. Hobbit offers services for social communication including an Internet phone used for the emergency scenario during the first empirical trials, but which can also be used to stay in touch with friends and relatives. With respect to these two functionalities Hobbit offers similar functionalities as comparable state-of-the-art care robots e.g. DOME0 [5] (see Fig. 17).



Fig. 17. The GUI page that provides access to a variety of useful information.

6. First user studies

First empirical user studies in a controlled laboratory setting with the Hobbit PT1 were carried out in Austria, Greece, and Sweden with a total of 49 primary participants. The studies were based on six representative interaction scenarios that should demonstrate the core tasks. The aim of the study was threefold: (1) exploring the Mutual Care interaction paradigm with two experimental conditions (see [47]), (2) exploring the impact of different impairment grades on the usability of Hobbit (see [53]), (3) deriving implications for improvement of the PT1 with respect to usability, acceptance, and affordability. The first two perspectives are published elsewhere with detailed descriptions of the experimental design, the instruments used, and results. In this article we present the study design (setting and procedure), instruments and measures, and results for the third aspect. In order to derive implications how PT1 can be improved we posed the following research questions:

- How do older adults (with representative age impairments) perceive the multimodal interaction possibilities of Hobbit in terms of usability?
- Do older adults accept Hobbit as assistive household robot after interacting with it in the laboratory?
- How do older adults perceive the value of Hobbit as support to enable independent living at home with respect to affordability and willingness to pay for it?

6.1. Sample

As mentioned before, the ultimate goal of the Hobbit robot is to enable older adults to live independently at home as long as possible. In Austria the age of older adults moving to a care facility is around 81 (according to the in-house statistics of the care facility in Austria we cooperated with), with men on average being slightly younger (76 years). Therefore, we decided to conduct our studies with participants aged 70 plus, as these will be the users who will have a Hobbit at home. Additionally, we tried to have a representative sample in relation to the typical age impairments [54]. In order to identify impairments we used self-reporting via telephone in the recruitment phase to assess the grade of impairments in the field of vision, hearing, and mobility. Many of our participants experienced impairments in more than one of the three categories. In total, 44 (89.8%) had some form of multiple impairment (e.g. moderate vision and minor mobility problems) and 78% of the sample fulfilled the impairment requirement of having at least one impairment graded as “moderate”.

In 35 cases the PUs were accompanied by secondary users (SU)—relatives or friends, whose presence was assumed to help primary users feel more comfortable during the experiment. In Austria 12 PUs and 9 SUs took part in the study; in Sweden 21 PUs and 11 SUs and in Greece 16 PUs and 15 SUs.

6.2. Representative tasks

The user studies were based on six tasks, which were representative for the core functionalities of Hobbit PT1, in order to allow participants to reasonably assess the acceptability and affordability of the robot.

6.2.1. Introduction

This task served as an ice-breaker to familiarize the participant with the robot. Hereby the robot introduced itself and explained its functionalities; Hobbit guided the user through a configuration dialog to define setup attributes like robot voice, sound volume and user name. Additionally, the user could try out speech, gesture, and screen input.

6.2.2. Clear floor

This task demonstrated the clear floor functionality. The user had to command Hobbit to pick up an object from the floor, put it on its tray and bring it to the user.

6.2.3. Learn object

This task demanded that the participants help the robot (one aspect of Mutual Care) to learn a new object. In order to learn an object the participant was asked to put the object on a specific “learning turntable”, which had to be put into the gripper of Hobbit. When the task was finished, half of the participants (i.e. Mutual Care condition) were thanked by the robot for teaching it a new object and were offered that Hobbit could return that favor. If participants wanted the favor returned, Hobbit offered a surprise (a randomly chosen joke, video or music file). The other half of the participants (i.e. the control group) were just told by the robot that it successfully finished learning at the end of the task. In other words, although participants of both groups had to help the robot, only the Mutual Care group received the stimulus that the robot wants to return the favor of helping it to learn an object.

6.2.4. Bring object with failure

This task was set-up intentionally in a way that Hobbit first failed to bring the object after the user commanded it to do so. In the Mutual Care group Hobbit then returned and asked the user if he/she might help it finding the object. In case the participant agreed he/she could specify the whereabouts of the object via touchscreen. After another search using this information the robot returned with the object. It thanked the participants for the received help and offered to return the favor by letting them choose from its entertainment menu. On the contrary, in the control group the robot returned to the participants and only reported that it could not fulfill the task. In other words, no help was demanded or given at all.

6.2.5. Bring object

This task was exactly the same again for both groups. Hobbit searched for another object and successfully brought it to the participants. This was intended to demonstrate participants of both groups that Hobbit in general is capable of bringing a specified object on its own.

6.2.6. Emergency

This last task was again the same for both groups and should demonstrate to the participants how an emergency call scenario with Hobbit would look like. Therefore an actor played a senior falling on the floor in front of Hobbit. Hobbit detected the accident, started a calming dialog and then established an emergency call, which was then handed over to the participant.



Fig. 18. Briefing Area (left) and Main Testing Area (right), both in Austria.

6.3. Setting and procedure

We began the user studies at the Austrian test site in March 2013, and then continued in Greece in April, and finally conducted the trials in Sweden in early May. The trials consisted of three parts: (A) the introduction phase, including a pre-questionnaire and briefing on how to use Hobbit and what it can do (B) the actual user study with the robot (six representative tasks) and (C) the debriefing phase. The setting for the user studies was very similar at the three test sites: It always consisted of two adjacent areas with separation screens and a doorway in between. We had a Briefing Area at all sites (see Fig. 18, left) and a Main Testing Area (see Fig. 18, right). This area was decorated as a living room including a cozy chair for the PU and a space in the background for the SU and the study facilitator.

The following people were present during the trials:

- The primary user,
- the secondary user,
- the facilitator: a researcher who introduced the robot and guided the user through the trial tasks,
- a scientific observer: a researcher who remained in the background and observed the users' behavior and reactions or incidences during the studies, such as unexpected reactions from the participants and technical problems,
- a technician: a researcher who also remained in the background to navigate the robot with remote control and assure that the robot functioned correctly, especially during learning, object recognition and grasping, which were autonomously done by the robot.

This semi-autonomous setting ensured the same study conditions for every participant. In total, one trial lasted on average 2.5 h (including introduction and debriefing questionnaire). However, if wanted, users could take breaks in between phases or tasks.

6.4. Instruments and measures

The user studies were based on a multi-informant approach taking into account data generated by the PUs, SUs, and the scientific observer. We used observational protocols filled in by the SU and the scientific observer, moreover, questionnaires were filled in by the PU together with the study facilitator in an interview-like manner. All trials were also video-recorded to fill gaps in the observation protocols after the study. In the following we will describe our measures for the three research aims respectively.

6.4.1. Usability measurements

In order to measure how participants perceive the usability of interacting with Hobbit they had to answer the following three usability-related questions after every task (*post-task questionnaire*) on a 4-point scale, with “1” always being the negative pole and “4” the positive one.

- How easy was the task to accomplish?
- How intuitively could you operate Hobbit in this task?
- How was the pace of the task?

Moreover, we developed a *debriefing questionnaire*, which had to be filled in by all participants at the end of the trial (all items had to be rated on a 4-point scale, with 1 always being the negative pole and 4 the positive one). This questionnaire contained eight selected items from the System Usability Scale questionnaire [55]. Additionally, they were asked to rank the three input modalities (speech, gesture, and touch screen) according to the usage preference and subsequently three usability detail questions regarding the touch screen were posed.

6.4.2. Acceptance measurements

In order to measure if participants accept Hobbit as an assistive household robot we posed the following questions in the *debriefing questionnaire*

- Which pick-up functionality is the most important/helpful for you?
- How important would it be for you, if the robot transports objects?
- Could you imagine having the robot for a longer period in your home?
- Could you imagine having a robot taking care of you?
- How helpful do you think the robot would be in your home?
- How did you like being entertained by the robot?

6.4.3. Affordability measurements

Similarly, the perceived value of Hobbit and if participants consider it affordable was measured using the following items in the *debriefing questionnaire*.

- Would you buy such a robot for 14.000 Euro?¹
- Could you imagine buying such a robot for yourself in general?
- Could you imagine your relatives buying such a robot for you?
- Could you imagine renting such a robot if you needed it?²
- Could you imagine buying such a robot, if it could postpone your moving into a care institution by one year?

6.5. Results

In general, PUs were rather skeptical in the beginning if the robot could assist them. However, after working with the robot

¹ This cost assessment was based on the overall component costs of Hobbit PT1.

² For this question intentionally no renting costs were provided. We were just interested in gaining insights if primary and secondary users could identify more with the concept of renting or buying the robot.

Table 1
Results of the post-task questionnaire on usability.

Question	Task	Min	Max	Mode	Median	N
How easy?	T1	2	4	4	4	48
	T2	2	4	3	3	48
	T3	1	4	3	3	46
	T4	3	4	3	3	46
	T5	2	4	3	3	46
	T6	2	4	3	3	48
How intuitive?	T1	2	4	4	3.5	48
	T2	1	4	3	3	49
	T3	1	4	3	3	46
	T4	2	4	3	3	46
	T5	3	4	4	4	46
	T6	2	4	3	3	47
Right speed?	T1	1	3	3	2	33 (15)
	T2	1	3	1	2	47 (1)
	T3	1	4	3	3	44 (2)
	T4	1	3	1	2	44 (1)
	T5	1	3	3	2	45 (1)
	T6	2	4	3	3	43 (4)

for the few tasks, PUs mostly enjoyed the trial situation and found the tasks easy to accomplish and the interaction with Hobbit understandable and traceable. In the following we will present the results on our three research aims in more detail.

6.5.1. Usability

Table 1 presents the results on the post-task questionnaire and gives a first impression on how participants perceived the interaction with Hobbit in terms of the usability aspects ease of use, intuitiveness, and efficiency (in terms of task speed). The answering scale was a 4-point Likert scale (1 negative pole–4 positive pole); the numbers in brackets for question 3 indicate the number of participants who thought it was just the right speed, neither too fast nor too slow.

In other words, on a reflective level participants perceived all tasks as rather easy and intuitively to perform together with Hobbit (at this point, it needs to be considered that participants had the free choice to decide which input modality: speech, gesture or touch to use). When PUs were asked after the all six tasks to rank which mode of operation they preferred ($n = 49$), results showed the following order: voice commands (49%), touch screen (42.9%), gestures (6.1%). SUs ($n = 35$) were asked to rank the operation options as well. Again voice was chosen most often as the most preferred option (49%), touch screen was in second place (16.3%) and then gestures (2%). The results on the speed of the interaction shows that it was overall perceived as rather slow and that only for the very first task 15 participants considered that Hobbit had the right speed. So it seems the navigation speed was perceived well, however more complex interaction scenarios need improvement in terms of speed, which however is a trade-off with safety aspects.

Additionally, the observational data revealed that most participants were rather skeptical or insecure in the beginning, but then became more and more confident in the interaction with Hobbit. Moreover, it became apparent in the observation protocols that participants often began interacting with Hobbit with speech as input modality and then switched to the touchscreen. For Task 3 (Learn Object Task), it could be observed that this task was most challenging for the participants (putting the turntable in the gripper, following the instructions of the robot and being dependent on understanding the instructions to successfully complete the task). Thus, this robot task needs improvement in order to be successfully performed by older adults together with the robot.

The debriefing questionnaire which was based on items of the System Usability scale [55] revealed the following results (see Table 2). The answering scale was again a 4-point Likert scale (1 negative pole–4 positive pole). Overall as question 8 indicates

participants felt confident using the robot and assessed a positive perceived usability rating.

Finally, participants additionally assessed the graphical user interface of the touch screen, indicating a satisfying design (median: 4, mode: 4), font size (median: 4, mode: 4), and symbol recognition (median: 4, mode: 4).

6.5.2. Acceptance

Regarding the core functionality of Hobbit to pick-up objects PUs ranked fetching objects from the floor as the most important/helpful functionality (49%) among all pick-up alternatives, followed by fetching objects from high shelves (32.7%), whereas fetching objects from tables was only considered as most important by 10.2%.

Regarding the other questions on acceptance, results are presented in Table 3.

Interestingly, 77.6% of the PUs answered the question on transporting small objects for them as rather or very much important, but only 53.1% of the SUs considered that.

In total 57.2% of the PUs could imagine to have the robot at home for a longer period of time and even 65.3% could imagine that Hobbit takes care of them. Interestingly, 49% of the PUs considered the robot as rather or very helpful at home, but almost an equal number of PUs (44.9%) were skeptical about its helpfulness. Moreover, when asked if they could imagine having a robot taking care of them, frequent comments from PUs were that they would prefer a human being. Similarly, some also voiced the opinion that the robot could indeed be helpful, but that they themselves were still too healthy or active to need such a device now. We consider this partly as an answer effect [56], as it would be stigmatizing for an older adult to admit that they need a robot to independently live at home.

Finally, the entertainment functionality was considered as very enjoyable by the PUs during the user studies (N is only 25 as this question was only posed to participants who used this functionality during the study). In total 92% of the PUs stated that they rather or very much liked it to be entertained by Hobbit. Hereby, participants mentioned memory training, music, audio books, and fitness instructions as most interesting, while cooking recipes and computer games were rather unpopular among our participants.

6.5.3. Affordability

The question if PUs would be willing to spend 14.000 Euro for the robot (a production estimation made by the project consortium at that time with the used components) was not surprisingly rated rather low (only 4.1% answered this question with “rather”, nobody with “very much”). However, the question if one could independently from the price imagine to buy such a robot was rated better. In total 34.7% of the PUs could imagine to buy such a robot, however, they were skeptical if their SUs would be willing to buy such a robot for them.

The willingness of having such a robot however increased when we asked for renting options: 77.6% of the PUs could imagine to rent the robot and 81.6% could imagine to have such a robot in case it could postpone the movement into the retirement home. Even though the last question can be considered a leading question, the answer behavior nevertheless demonstrates that the willingness of independent living at home out-rules potential fears and rejection tendencies of robotic technology.

6.6. Summary

To summarize, we now want to answer our three research questions. Regarding RQ1 (Perceived Usability), it can be said that the questionnaire results showed that the interaction with Hobbit

Table 2
Results of the debriefing questionnaire usability.

	Min	Max	Mode	Median	N
I thought the robot was easy to use.	2	4	4	3.5	48
I think I would like to use the robot frequently.	1	4	3	3	48
I found the robot complex.	1	4	1	1	48
I think that I would need assistance to be able to use robot.	1	4	1	2	47
I imagine that most people learn to use the robot very quickly.	1	4	3	3	48
I found the robot very cumbersome/awkward to use.	1	2	1	1	48
I felt very confident using the robot.	1	4	3	3	48

Table 3
Results of the debriefing questionnaire on acceptance.

	Min	Max	Mode	Median	N
Could you imagine having the robot for a longer period in your home?	1	4	3	3	47
Could you imagine having a robot taking care of you?	1	4	3	3	48
How helpful do you think the robot would be in your home?	1	4	3	3	46
How important would it be for you, if the robot transports objects for you?	1	4	4	4	48
How did you like being entertained by the robot?	1	4	4	4	25

was perceived as easy and intuitive. However, observations during the study showed that improvements are still necessary for the initialization dialog and wording of robot instructions in general. The robot was furthermore mostly perceived as being rather slow in the tasks. On the whole, the multimodal approach of Hobbit with interaction possibilities via voice, touch screen, and gestures was confirmed as beneficial by the users. Voice and touchscreen were the possibilities, which were used most often. The Learn Object task, however, will need to be adjusted and made more intuitive for older adults, including instructions from the robot and easier handling of the turntable for objects.

Regarding RQ2 (Acceptance) it could be demonstrated that the most relevant and helpful household functionalities for our participants was picking up objects from the floor (from all pick-up options) and transporting small objects. Entertainment functionalities were highly appreciated by the participants, whereby according to the participants memory training, music, audio books, and fitness instructions are most preferable. More than half of our participants could imagine to have the robot at home for a longer period of time and that it could take care of them, even if the majority clearly preferred a human to do that, but overall the robot was positively perceived as care-giver.

Finally, with regards to RQ3 (Affordability), answers in the debriefing questionnaire clearly indicated that participants were skeptical of buying such a robot, but could imagine renting it for some time if needed. From the results, it can be assumed that SUs are more likely to be a buying target group.

Lessons learned

From our first user trials we could derive several relevant methodological lessons learned for fellow researchers. During the recruitment procedure we noticed that telephone reports are a resource-saving option for the categorization of impairments, but that they do not in all cases depict reality (as participants do not want to stigmatize themselves or they are not aware of the severity of an impairment). Therefore for the next trials we will use self-reports only as a first selection criteria and follow up with simple exercises that give insights on the impairment grade.

During the trials we noticed, that the effect, that older adults are insecure or afraid of using the robot vanished after the ice-breaker task. Therefore we recommend to use an initialization phase in which the participant can get used to the robot for every laboratory or field trial study with care robots that involve older adults as target group, as it reduces the novelty effect bias in the data. Additionally, having SUs present during the trials was of high added

value for our studies, as the PUs were more relaxed during the trials similarly to what has been shown in studies for child–robot interaction [57]. A lot of additional qualitative reflection data could be gathered this way from both PUs and SUs. Moreover, involving SUs as observers not only increased the interpretability of the observation results, but also ensured that they do not get too much involved in the interaction with the robot (it was still the PU we explored and not the SU).

Answering the questionnaire items in an interview-like manner together with the facilitator also proved its value to ease the overall study procedure for older adults and enabled us to ensure that the questions were correctly understood by the PU. However, we are aware of the fact that it might also have increased the amount of socially desirable answers, a phenomenon which can be even more observed in user studies with older adults [56]. Finally, the semi-autonomous Wizard-of-Oz design enabled us on the one hand to provide comparable situations for all participants due to the remote-controlled parts, but on the other hand also allowed to test key behaviors autonomously.

7. Conclusions and outlook

In this article we presented results from the development of the first Hobbit robot prototype (PT1) and the first set of user trials in a controlled laboratory setting focusing on the development of a socially assistive care robot for older adults, which has the potential to promote aging in place and to postpone the need to move to a care facility. Hobbit is designed especially for fall detection and prevention (e.g. by picking up objects from the floor, patrolling through the apartment and by employing reminder functionalities) and supports multimodal interaction for different impairment levels. The results from the user studies with the first prototype demonstrate that the robotic system can perform its core tasks in a satisfying manner for the target group. All participants were capable of performing all tasks together with the robot and assessed it as usable and acceptable. This was in particular astounding as users first approached the robot with great skepticism and doubted it could help or assist them.

The desirable long-term goal is that Hobbit enters real homes of older adults and that it provides a feeling of being safe and supported to its owner. Therefore, in the next period of the project we will test if our methods for autonomous navigation in the domestic environments, the strategies for human detection and tracking and object recognition and grasping, as well as the multimodal interface for interaction constitute a suitable framework for the overall



Fig. 19. Design of the second Hobbit Prototype, PT2.

scenario of a socially assistive robot for fall prevention and detection. These studies will be performed with the next prototype version PT2 which is currently under development (see Fig. 19). This time the hardware platform will be built by the German company *Metralabs* in order to assure compliance with all safety standards for a commercial product. At its current stage it seems promising that we can keep our price limit of 15 000 Euro in the development. After extensive testing we will conduct one of the (up to our knowledge) first long-term household trials with Hobbit in 20 private households (again in Austria, Greece, and Sweden) in order to explore how the user reception of robot and the self-efficacy of the user changes over time in a three weeks (per user) period.

We believe that methods, results, and lessons learned presented in this article constitute valuable knowledge for fellow researchers in the field of service robotics and serve as a stepping stone towards developing affordable care robots for the aging population.

Acknowledgments

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 288146, Hobbit and from the Austrian Science Fund (FWF) under grant agreement T623-N23, V4HRC.

References

- [1] L.N. Gitlin, Conducting research on home environments: lessons learned and new directions, *Gerontologist* 43 (5) (2003) 628–637.
- [2] C.B. Fausset, A.J. Kelly, W.A. Rogers, A.D. Fisk, Challenges to ageing in place: understanding home maintenance difficulties, *J. Hous. Elderly* 25 (2) (2011) 125–141.
- [3] P. Parette, M. Scherer, Assistive technology use and stigma, *Educ. Train. Dev. Disabil.* 39 (3) (2004) 217–226.
- [4] “ksera”. [Online]. Available: <http://ksera.ieis.tue.nl/>.
- [5] “domeo”. [Online]. Available: http://www.aat.tuwien.ac.at/domeo/index_en.html.
- [6] “cogniron”. [Online]. Available: <http://www.cogniron.org/>.
- [7] “companionable”. [Online]. Available: <http://www.companionable.net/>.
- [8] “srs”. [Online]. Available: <http://srs-project.eu/>.
- [9] “care-o-bot”. [Online]. Available: <http://www.care-o-bot.de/>.
- [10] “accompanyproject”. [Online]. Available: <http://www.accompanyproject.eu/>.
- [11] “herb”. [Online]. Available: <http://www.cmu.edu/herb-robot/>.
- [12] J. Broekens, M. Heerink, H. Rosendal, Assistive social robots in elderly care: a review, *Gerontechnology* 8 (2) (2009) 94–103.
- [13] K. Wada, T. Shibata, T. Saito, K. Sakamoto, K. Tanie, Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged, in: *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on, IEEE, 2005*, pp. 2785–2790.
- [14] B. Graf, U. Reiser, M. Hägele, K. Mauz, P. Klein, Robotic home assistant care-o-bot® 3-product vision and innovation platform, in: *Advanced Robotics and its Social Impacts (ARSO), 2009 IEEE Workshop on, IEEE, 2009*, pp. 139–144.
- [15] C. Schroeter, S. Mueller, M. Volkhardt, E. Einhorn, C. Huijnen, H. van den Heuvel, A. van Berlo, A. Bley, H.-M. Gross, Realization and user evaluation of a companion robot for people with mild cognitive impairments, in: *Robotics and Automation (ICRA), 2013 IEEE International Conference on, IEEE, 2013*, pp. 1153–1159.
- [16] “IGUS Robolink”. [Online]. Available: <http://www.igus.eu/wpck/6076/roboLink>.
- [17] “Festo Fin Ray”. [Online]. Available: http://www.festo.com/net/SupportPortal/Files/53886/AdaptiveGripper_DHDG_en.pdf.
- [18] J. Oberzaucher, K. Werner, H.P. Mairböck, C. Beck, P. Panek, W. Hlauschek, W.L. Zagler, A videophone prototype system evaluated by Elderly users in the living lab Schwechat, in: *HCI and Usability for inclusion*, vol. 5889, 2009, pp. 345–352.
- [19] M. Wölfel, J.W. McDonough, *Distant Speech Recognition*, John Wiley & Sons, 2009.
- [20] P. Panek, P. Mayer, Challenges in adopting speech control for assistive robots, in: *Advanced Technologies and Societal Change to be Printed in Ambient Assisted Living, 7th German AAL Congress, 21–22 January 2014, Berlin, Springer, 2014*.
- [21] M.L. Walters, K. Dautenhahn, S.N. Woods, K.L. Koay, Robotic etiquette: results from user studies involving a fetch and carry task, in: *Proceedings of the ACM/IEEE International Conference on Human–robot Interaction, HRI’07, ACM, New York, NY, USA, 2007*, pp. 317–324. [Online]. Available: <http://doi.acm.org/10.1145/1228716.1228759>.
- [22] G. Grisetti, C. Stachniss, W. Burgard, Improved techniques for grid mapping with Rao–Blackwellized particle filters, *IEEE Trans. Robot.* 23 (1) (2007) 34–46.
- [23] P. Besl, N. McKay, A method for registration of 3-D shapes, *IEEE Trans. PAMI* 14 (2) (1992) 239–256.
- [24] S. Thrun, W. Burgard, D. Fox, *Probabilistic Robotics*, MIT Press, Cambridge, MA, 2005. <http://www.probabilistic-robotics.org>.
- [25] R. Labayrade, D. Aubert, J.P. Tarel, Real time obstacle detection in stereovision on non flat road geometry through “v-disparity” representation, in: *Intelligent Vehicle Symposium 2002 IEEE, vol. 2, IEEE, 2002*, pp. 646–651.
- [26] M. Phillips, M. Likhachev, Planning in domains with cost function dependent actions, in: *Twenty-Fifth Conference on Artificial Intelligence (AAAI), San Francisco, USA, 2011*.
- [27] M. Likhachev, D. Ferguson, G. Gordon, A. Stentz, S. Thrun, Anytime dynamic A*: an anytime, replanning algorithm, *Science* (2005) 262–271.
- [28] T.B. Moeslund, *Visual Analysis of Humans: Looking at People*, Springer, 2011.
- [29] K. Papoutsakis, P. Panteleris, A. Ntelidakis, S. Stefanou, X. Zabulis, D. Kosmopoulos, A. Argyros, Developing visual competencies for socially assistive robots: the HOBBIT approach, in: *6th International Conference on Pervasive Technologies for Assistive Environments, 2013*.
- [30] “No Title”. [Online]. Available: <http://www.microsoft.com/en-us/kinectforwindows/>.
- [31] C. Plagemann, V. Ganapathi, D. Koller, S. Thrun, Real-time identification and localization of body parts from depth images, in: *Robotics and Automation (ICRA), 2010 IEEE International Conference on, 2010*, pp. 3108–3113.
- [32] J. Shotton, R. Girshick, A. Fitzgibbon, T. Sharp, M. Cook, M. Finocchio, R. Moore, P. Kohli, A. Criminisi, A. Kipman, A. Blake, Efficient human pose estimation from single depth images, in: *IEEE PAMI*, vol. 99, 2012.
- [33] “No Title”. [Online]. Available: <http://www.openni.org/wp-content/uploads/2013/02/NITE-Algorithms.pdf>.
- [34] “openni”. [Online]. Available: <http://www.openni.org/>.
- [35] P. Paderis, X. Zabulis, A.A. Argyros, Head pose estimation on depth data based on Particle Swarm Optimization, in: *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on, 2012*, pp. 42–49.
- [36] “nite”. [Online]. Available: <http://www.primesense.com/solutions/nite-middleware/>.
- [37] D. Fischinger, M. Vincze, Empty the basket—a shape based learning approach for grasping piles of unknown objects, in: *IROS, 2012*.

- [38] D. Fischinger, Y. Jiang, M. Vincze, Learning grasps for unknown objects in cluttered scenes, in: International Conference on Robotics and Automation (ICRA), 2013.
- [39] R. Diankov, J. Kuffner, OpenRAVE: a planning architecture for autonomous robotics, Tech. Rep. CMU-RI-TR-08-34, Robotics Institute, 2008, no. July.
- [40] W. Wohlkinger, M. Vincze, Ensemble of shape functions for 3D object classification, in: ROBIO, 2011, pp. 2987–2992.
- [41] W. Wohlkinger, A. Aldoma, R.B. Rusu, M. Vincze, 3DNet: large-scale object class recognition from CAD models, in: ICRA, 2012, pp. 5384–5391.
- [42] A. Criminisi, J. Shotton, E. Konukoglu, Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning, *Found. Trends Comput. Graph. Vis.* 7 (2–3) (2012) 81–227.
- [43] T. Körtner, A. Schmid, D. Batko-Klein, C. Gisinger, A. Huber, L. Lammer, M. Vincze, How social robots make older users really feel well—a method to assess users' concepts of a social robotic assistant, in: *Social Robotics*, Springer, 2012, pp. 138–147.
- [44] S. Frennert, B. Östlund, H. Efring, Capturing seniors' requirements for assistive robots by the use of attention cards, in: Proceedings of the 7th Nordic Conference on Human–Computer Interaction: Making Sense Through Design, NordiCHI'12, ACM, New York, NY, USA, 2012, pp. 783–784. [Online]. Available: <http://doi.acm.org/10.1145/2399016.2399146>.
- [45] J.M. Beer, C.-A. Smarr, T.L. Chen, A. Prakash, T.L. Mitzner, C.C. Kemp, W.A. Rogers, The domesticated robot: design guidelines for assisting older adults to age in place, in: *Human–Robot Interaction (HRI)*, 2012 7th ACM/IEEE International Conference, 2012, pp. 335–342.
- [46] L. Lammer, A. Huber, W. Zagler, M. Vincze, Mutual-care: users will love their imperfect social assistive robots, in: Proceedings of the International Conference on Social Robotics, 2011, pp. 24–25.
- [47] L. Lammer, A. Huber, A. Weiss, M. Vincze, Mutual care: how older adults react when they should help their care robot, in: AISB2014: Proceedings of the 3rd International Symposium on New Frontiers in Human–Robot Interaction, 2014.
- [48] P. Mayer, P. Panek, A social assistive robot in an intelligent environment, *Biomed. Eng.* (2013) *Biomedizinische Technik*.
- [49] N. Noury, A. Fleury, P. Rumeau, A.K. Bourke, G.O. Laughin, V. Rialle, J.E. Lundy, Fall detection—principles and methods, in: *EMBS*, 2007, pp. 1663–1666.
- [50] C. Rougier, E. Auvinet, J. Rousseau, M. Mignotte, J. Meunier, Fall detection from depth map video sequences, in: Proceedings of the 9th International Conference on Toward Useful Services for Elderly and People with Disabilities: Smart Homes and Health Telematics, ICOST'11, Springer-Verlag, Berlin, Heidelberg, 2011, pp. 121–128.
- [51] Z.-P. Bian, L.-P. Chau, N. Magnenat-Thalmann, Fall detection based on skeleton extraction, in: Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry, VRCAI'12, ACM, 2012, pp. 91–94.
- [52] G. Mastorakis, D. Makris, Fall detection system using Kinect's infrared sensor, *J. Real-Time Image Process.* (2012) 1–12.
- [53] T. Körtner, A. Schmid, D. Batko-Klein, C. Gisinger, Meeting requirements of older users?—robot prototype trials in a home-like environment, in: Proceedings of HCI International 2014 Conference, Crete, 2014.
- [54] U. Lindenberger, *Die Berliner Altersstudie*, Akademie Verlag, 2010.
- [55] J. Brooke, SUS—a quick and dirty usability scale, in: *Usability Evaluation in Industry*, vol. 189, 1996, p. 194.
- [56] R. Eisma, A. Dickinson, J. Goodman, A. Syme, L. Tiwari, A.F. Newell, Early user involvement in the development of information technology-related products for older people, 2004, pp. 131–140.
- [57] H. Knight, A. Chang, Robot design rubrics for social gesture categorization and user studies with children, in: *HRI Workshop on ...*, 2009, pp. 2–5.



David Fischinger is a researcher and Ph.D. candidate at the Vienna University of Technology at the "Automation and Control Institute" since 2010. His main research interests are robotic grasping, machine learning and vision for robotics. He graduated in Technical Mathematics at the Vienna University of Technology in 2005. In 2007, he achieved Master's degrees in "Computer Science Management" and in "Computational Intelligence" (Informatics)—both with honors. From 2007 to 2010 he worked as Senior Programmer at the management and consulting company Accenture in Munich, Vienna and Hyderabad (India).



Peter Einramhof studied Electrical Engineering with a focus on computer technology at Vienna University of Technology, where he graduated with distinction. In 2003 he joined the vision for robotics group (V4R) of the university's Automation and Control Institute. For ten years he was working as research assistant in robotics-related EC-funded projects. The scope of his research includes real-time processing of data from stereo and depth cameras for self-localization and safe navigation of service robots, and visual attention algorithms. In July 2013 he joined the Institute for Applied Systems Technology in Bremen.



Konstantinos Papoutsakis is a Ph.D. candidate at University of Crete, Greece and a Research Assistant at the Institute of Computer Science, FORTH. He graduated with a Bachelor's Degree in Computer Engineering and Informatics from University of Patras and received a Master's degree in Computer Science from University of Crete. His main research interests are computer vision, robotics and machine learning with emphasis on visual object tracking, human motion analysis, activity recognition and human–robot interaction.



Walter Wohlkinger is the Co-Founder and CEO of Blue Danube Robotics. During his masters in Computer Graphics at the Vienna University of Technology, and also after graduating with a Ph.D. in Electrical Engineering, he worked on robot vision and grasping, especially in the context of personal assistive robots. In 2013, he founded Blue Danube Robotics together with Michael Zillich, to realize a new class of affordable robots truly designed around personal use at home, helping people with disabilities to maintain their independence.



Peter Mayer graduated in 1985 from the Vienna University of Technology with a diploma degree in Electrical Engineering. During his studies he gained practical experience in Quality Assurance at Schrack and in the programming of control and tomographic measurement of electron beams for welding at the Institute of Industrial Electronics of the Vienna University of Technology. Since then he worked in the area of rehabilitation engineering as a research assistant at the Vienna University of Technology in many national and EU-funded R&D projects. He specialized on assistive devices for the disabled and old. During the last years the focus of his work shifted to assistive robotics and smart environments (AAL). His special interest is on speech input and output technology, embedded systems, modern communication services, smart sensors, mainstreaming education.



Paul Panek was born in 1966 and studied Communication Engineering at the Vienna University of Technology. Since 1993 he is a member of the Fortec group working in the field of Rehabilitation Technology. His main areas of interest are man machine interfaces for multiple impaired persons, alternative and augmentative communication (AAC), environmental control systems (ECS) and Ambient Assisted Living (AAL). In 1997/1998 he did an industrial R&D project at the Carinthian Tech Research. Since 2006 he also works at Ceit Raltec institute in AAL Living Lab Schwechat.



Stefan Hofmann made his Master in "System Design" (with focus on "Control System") at the advanced technical college in Villach/Austria with honors. He is now working for Hella Automation in the field of control theory for robotic arms.



Tobias Koertner was born on 25.09.1980 in Bielefeld, Germany. He graduated in Psychology and Anglistics at the University of Vienna. He is a trained clinical and health psychologist and is working in science projects with the "Academy for Aging Research" since 2011.



Astrid Weiss is a postdoctoral research fellow in HRI at the Vision4Robotics group at the Institute of Automation and Control (ACIN) at Vienna University of Technology (Austria). She holds a Master's degree in Sociology and a Ph.D. in Social Sciences from the University of Salzburg, Austria. During her studies she specialized on methodologies of empirical social research and applied statistics. Her current research focuses on user-centered design and evaluation studies for Human–Computer Interaction and Human–Robot Interaction. She is especially interested in the impact technology has on our everyday life and what

makes people accept or reject technology. Before her position in Vienna she was a postdoc researcher at the HCI&Usability Unit, of the ICT&S Center, University of Salzburg, Austria and at the Christian Doppler Laboratory on “Contextual Interfaces” at University of Salzburg.



Antonis Argyros is a Professor of Computer Science at the Computer Science Department, University of Crete (CSD-UoC) and a Researcher at the Institute of Computer Science—FORTH, in Heraklion, Crete, Greece. He received B.Sc. (1989) and M.Sc. degrees (1992) in Computer Science, both from the CSD-UoC. On July 1996, he completed his Ph.D. in Visual Motion Analysis at the same Department. He has been a postdoctoral fellow at the Computational Vision and Active Perception Laboratory, KTH, Sweden. He is an area editor for the Computer Vision and Image Understanding (CVIU) Journal, member of the Editorial Board

of the IET Image Processing Journal and a General Chair of ECCV'2010. He is also a member of the Executive Committee of the European Consortium for Informatics and Mathematics (ERCIM). His research interests fall in the areas of computer vision

with emphasis on tracking, human gesture and posture recognition, 3D reconstruction and omnidirectional vision. He is also interested in applications of computational vision in the fields of robotics and smart environments. In these areas he has (co-)authored more than 100 papers in scientific journals and conference proceedings.



Markus Vincze received his diploma in Mechanical Engineering from Technical University Wien (TUW) in 1988 and a M.Sc. from Rensselaer Polytechnic Institute, USA, 1990. He finished his Ph.D. at TUW in 1993. With a grant from the Austrian Academy of Sciences he worked at HelpMate Robotics Inc. and at the Vision Laboratory of Gregory Hager at Yale University. In 2004, he obtained his habilitation in robotics. Presently he leads a group of researchers in the “Vision for Robotics” laboratory at TUW. With Gregory Hager he edited a book on Robust Vision for IEEE and is (co-)author of over 250 papers. His special

interests are computer vision techniques for robotics solutions situated in real-world environments and especially homes.