Journal of Intelligent & Robotic Systems (2020) 98:19–37
https://doi.org/10.1007/s10846-019-01104-z

SMOOTH Robot: Design for a Novel Modular Welfare Robot

William K. Juel[1] · Frederik Haarslev[1] · Eduardo R. Ramírez[1] · Emanuela Marchetti[2] · Kerstin Fischer[3] · Danish Shaikh[4] · Poramate Manoonpong[4] · Christian Hauch[5] · Leon Bodenhagen[1] · Norbert Krüger[1]

Received: 28 January 2019 / Accepted: 20 September 2019 / Published online: 11 November 2019

Abstract

Demographic change is expected to challenge many societies in the next few decades if todays' standards of services in e.g. elder care shall be maintained. Robots are considered to at least partially mitigate this challenge, however, robots are rarely applied in the welfare domain. This paper describes the development of a concept for a novel welfare robot that is modular and affordable. The development is based on a participatory design process and by taking strengths and limitations of selected, commercially available robots into account. This work contributes a design methodology specific for welfare robots and a resulting robot concept that address three use cases in a care center. The concept includes multi-modal robot perception that facilitates a proactive robot behavior for achieving smooth interactions with end-users.

Keywords Robotics · Welfare · Healthcare · Design · HRI

1 Introduction

Many societies are facing a demographic shift. In 2015 8.5% of the global population was aged 65 or above and this number is projected to increase to 17% by 2050 [20]. As a result an increase in multi-morbidity is expected which causes prolonged, complex and transverse patient care. This creates an increased pressure on the healthcare system, both in terms of economic and staffing, which threatens the coherent patient pathways [3].

Healthcare systems are already under strain due to various challenges such as high workloads and difficulties with recruiting new staff, making it unlikely that existing structures can handle the challenges imposed by the demographic change without a substantial decrease in the level of service provided. It has been suggested that Robots have the potential to be a partial solution by supporting caregiving staff with selected tasks, however, robotic solutions are rarely found in this domain [36].

The general shortage of resources has motivated both company's and researchers from various disciplines to work towards applicable robotic solutions. Here we will focus on robots that arise from research projects while later we will discuss strengths and limitations of commercially available mobile robots applicable in the healthcare domain. An example of a robot arising from a research project is the HOBBIT robot [14]. This is a mobile service robot which is designed to be affordable and aimed at enabling aging in place, i.e. avoiding or delaying the need for an older adult to move to a care center. Core needs identified in expert workshops were the prevention of falls and the detection of emergencies. Subsequently workshops and questionnaires involving potential end users were used to identify user needs. These clustered in three groups: communication, household chores (e.g. fetch and carry tasks), and care activities e.g. detecting falls or giving reminders. Based on their identified use cases and their aim to develop an affordable robot, leading to a cost limit of EUR 15.000, the HOBBIT robot was designed with a low-cost manipulator

---

This research was part of the SMOOTH project (project number 6158-00009B) by Innovation Fund Denmark.

✉ William K. Juel
wkj@mmmi.sdu.dk

[1] SDU Robotics, Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark

[2] Department for the Study of Culture, University of Southern Denmark, Odense, Denmark

[3] Department of Design and Communication, University of Southern Denmark, Sønderborg, Denmark

[4] SDU Embodied Systems for Robotics and Learning, Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark

[5] Robotize Aps, Maglebjergvej 5B, 2800 Kgs. Lyngby, Denmark

and strategically arranged consumer grade sensors. The evaluation focused on the functionalities of the robot and potential costs, revealing that the majority of the users could imagine renting such a robot themselves.

The Toyota Home Service Robot (HSR) [18] is another mobile robot designed to support independent living and just as TIAGo [31] having similar overall properties regarding shape, sensor setup and size. Both robots seem though to have been developed without direct involvement of end users which at least for Toyota HSR has led to a mismatch between user's expectations and the actual capabilities and properties of the robot [18].

Other robots have been developed with a focus on tasks related to communication or monitoring, thus avoiding the need for manipulators, reducing the complexity of the setup and in some cases enabling the use of commercially available robots. ALIAS [35] is a mobile robot developed with entertainment of older adults being the main use case and the interaction between the robot and the end users being in focus. The project GiraffPlus [11] uses of a telepresence robot equipped with additional sensors for communication and monitoring to detect potential incidents or risks. The development was an iterative process involving both end users and professionals deriving the specifications of the system and for evaluation, resulting in the system being deployed in six homes in Spain, Sweden and Italy in common household areas [11] for evaluation.

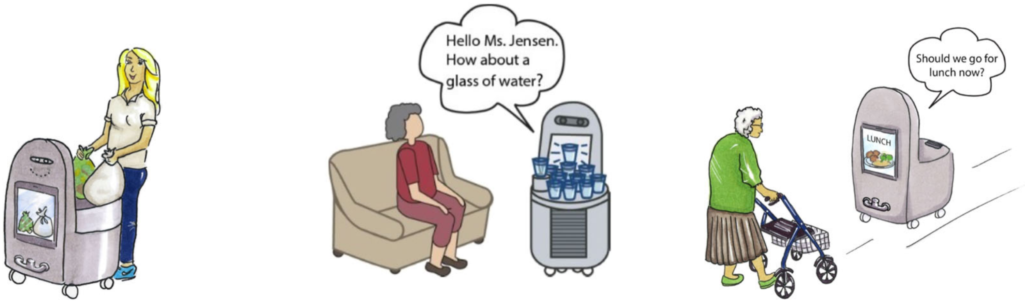
This paper describes the development process used the SMOOTH project [42] in which a design concept for a mobile welfare robot of limited complexity was developed by actively involving end users and professionals from Ølby elderly care center (ØECC). The overall aim was to develop a robot that contributes to mitigating the challenge imposed by the demographic change while aligning end-users' expectations with the capabilities of the robot. The process started by defining three use cases (see Fig. 1) that, based on dialogue with professionals from ØECC, exemplify user and professional carer needs.

The general design process can be split into four phases with some intermediate and overlapping steps as illustrated in Fig. 2. In Section 2 we describe phase 1 and 2: we present the participatory design process upon which the project is built, a detailed use case analysis based on an ethnographic study and a short explanation of our involvement with external experts in robot ethics and geronto-psychology. In Section 3 existing robotic solutions are discussed in the context of SMOOTH. The findings from this discussion combined with the knowledge extracted from phase 1 is used in Section 4 to identify and discuss four design requirements for welfare robots: *affordability, simplicity, modularity* and *acceptability* these are considered to be essential aspects to be taken into account during robot development.

Phase 3 is described in Section 5: we present three design concepts for the SMOOTH robot that via the design requirements can solve the three use cases. In this section we also initiate a selection process between these three designs. Phase 4 is described in Section 6: we present the final design concept of the SMOOTH robot and introduce our methods for realization of smooth robot control and navigation. In Section 7 we discuss a formative evaluation of the final design and in Section 8 we present the physical robot built from the process described in this paper.

2 Participatory Design Process

The design process for the robot concept started by following a participatory design (PD) approach [6, 45], to gain a rich understanding of how robot technologies could improve the lifes of elderly and improve the working environment for the care-givers. The goal of the project is, in accordance with the core principles of PD, to actively involve staff and residents from ØECC as much as possible as co-designers and to articulate relevant use cases and specifications so that we design a robot that best matches their articulated needs. Following PD principles, we conducted an iterative design



(a) Laundry and garbage handling (b) Water delivery (c) Guidance

Fig. 1 The three cases of the SMOOTH project (all illustrations made by the Danish Technological Institute)

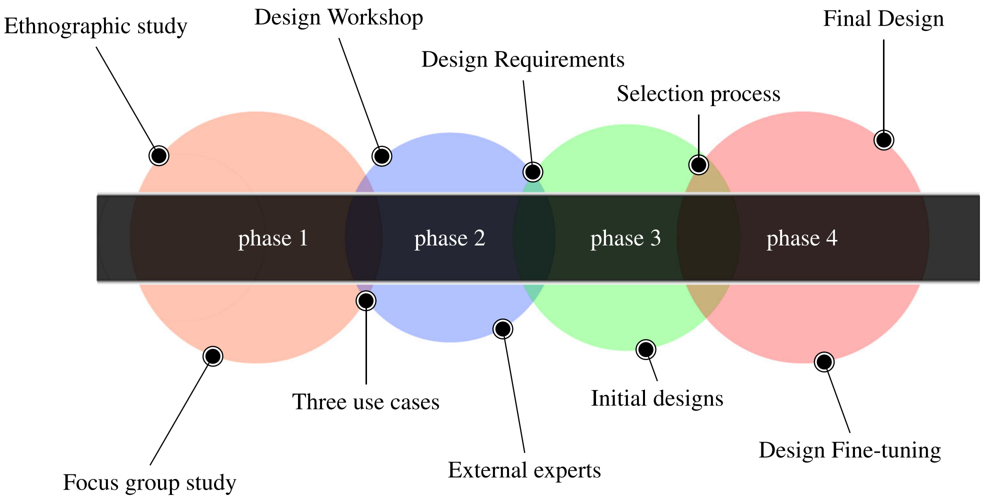


Fig. 2 The 4 phase design process of the SMOOTH project. Size of the blobs indicates the workload in each phase

process. The process started with a user study (phase 1), during which we conducted ethnographic observations and situated interviews [32] at ØECC together with caregivers and residents. From this study three use cases were identified. We then conducted a co-design workshop (phase 2) supported by situated interviews involving a focus group of 4 caregivers and 2 residents, during which the participants were asked to develop design concepts for a welfare robot. Further data were collected at the Robophilosophy 2018 Conference, where we organized a workshop[1] discussing the use cases within the areas of ethics, design and gerento-psychology. 2D and 3D visualizations was created to facilitate brainstorming within the research team on the shape and technical characteristics of the robot. Through this, which we call the selection process (phase 3, discussed in Section 5.2), a tangible, low-fidelity mock-up of the final design has been created in polystyrene and taken to ØECC for formative evaluation with caregivers and residents. A downscaled 3D tangible model of the robot design was also printed to show how the design concept look and elicit further comments regarding the aesthetics of the robot.

In the following sections we discuss the main work of phase 1 and 2: the ethnographic user study (Section 2.1), the use cases analysis (Section 2.2), the co-design workshop (Section 2.3), and finally the procedure and outcomes from the Robophilosophy Conference (Section 2.4).

2.1 Ethnographic Observation

To study the current workflows, a 24-hour ethnographic observation was done in two units at ØECC. Previous work, especially [30], had shown that the same robot can be perceived very differently, e.g. as helpful or as an obstacle, depending on the workflow the robot is embedded in and how the robot behaves. Furthermore, people may be so used to the current workflow that they do not notice the potential for improvement themselves [44]. A useful methodology to address these issues is ethnography, an exploratory methodology to document and describe cultural practices from the perspective of the stakeholders, i.e. from the inside [32]. We found that the residents at ØECC receive highly individualized care due to the small units and the high number of personnel per resident (two on the early shift and late shift, one on the night shift, for 5-6 residents). The caregivers often walked between residents' rooms and the laundry and garbage room, hence automating laundry and garbage collecting would relieve caregivers and free them for more caregiving and social interaction with the residents. Drinks was provided as necessary. Making the elderly drink was mainly stressed during mealtimes and different drinks are served in different and specific ways. There was little resident activity in the common areas, which could be countered by offering drinks between mealtimes in the common spaces. We observed that caregivers spend considerable amounts of time guiding residents on even short distances.

2.2 Use Case Analysis

Use case 1 (Fig. 1a) is a logistic task in which the robot transports laundry and garbage at distances between 10 and 50 meters. The laundry/garbage bin is usually located in the residents apartments, but when the staff members are doing their daily routine they will place the bin outside of the residents room, where the robot will pick it up and drive it to a drop off location close to the laundry/garbage room.

---

[1] Exploring Responsible Robotics Hands-On: A conference Lab on Three Use Cases (phase 2) http://conferences.au.dk/robo-philosophy-2018-at-the-university-of-vienna/workshops/

We expect little interaction between staff, residents and the robot during this use case. Use case 2 (Fig. 1b) addresses offering drinks to elderly people, who often lack a feeling of thirst. The important aspect is to construct a serving tray and motivate the elderly to drink, which can be encouraged by persuasive human-robot dialog. Use case 3 (Fig. 1c) addresses the problem of elderly people often requiring guidance to reach a certain place such as the dining area. Use cases 2 and 3 pose a significant challenge with respect to the smoothness and appropriateness of human-robot interaction (HRI). A core idea of the SMOOTH project is to design "pro-active" control which takes the expected actions of the humans into account (see Section 6.4).

The analysis of the use cases was done iteratively, using context evaluation and taking ethical, anthropological, design and geronto-psychological considerations into account, as well as aspects of computational and economic feasibility. This was done to ensure general user acceptance, a smooth integration of the robot into the existing workflows, and to align expectations between the different stakeholders.

2.3 Focus Group and Co-Design Workshop

Focus group interviews with the director and employees were carried out in combination with a prototyping workshop involving residents as well. Individual interviews with members of the staff were carried out to understand their specific ideas, hopes, needs and fears. Interviews provided us with a quantification of the processes observed during the Ethnographic Observation, i.e how many residents need guidance to and from their rooms, and how often laundry is collected.

The main goal of the design workshop was to include the users in the start-up phase and obtain user input for the robot development. By involving the staff at ØECC in the design process, we also try to increase their excitement about the robot. Participants were asked to build a robot prototype for the three use cases using various materials provided: cardboard, paper, scissors, tape, straws, rulers, egg trays, Lego etc. The Participants were asked to consider what the robot should say/hear and what behavior, appearance and, verbal feedback it should have. The Participants presented their prototypes which ended in a discussion. The output of the exercise revealed the importance of three main categories: Social skills, behavioral constraints and, appearance:

- Social skills: Polite voice, human-like voice, must not harm anybody or anything
- behavioral constraints: Global call system, be safe around humans, collect garbage and laundry one room at the time, lift outdoor garbage container lids
- Appearance: appealing design, no sharp edges, hygienic

2.4 Conference Workshop

A workshop was organized at the Robophilosophy 2018 Conference to further discuss the use cases. Four invited guests with different backgrounds – ethics, design, geronto-psychiatry and, HRI – provided us with many useful considerations regarding the concrete realization of the use cases and the robot prototypes. For instance ensuring that the anthropomorphic design of the robot could appeal more to people with dementia, and notes about choosing the colors of the robot, to fit the elderly care facility better.

Ethical discussions focused on the fact that we are designing not only robots, but future interactions, led to two questions:

> Which kinds of human-robot interactions should we promote?

and in contrast

> Which kinds of interactions are between a human and a 'social' robot?

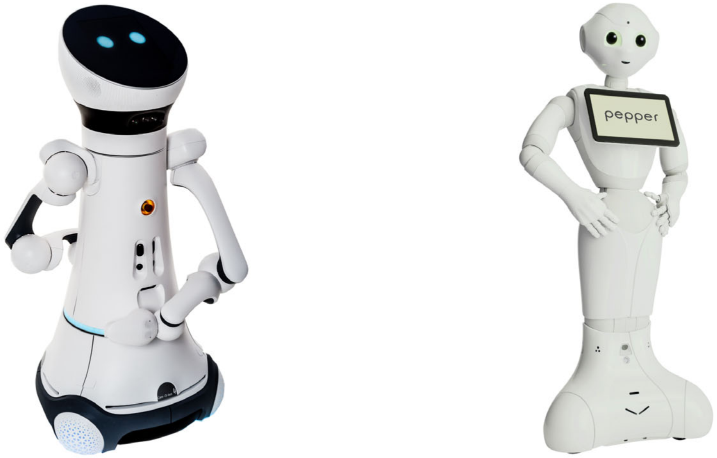
We must consider what the psychological, socio-cultural and economic dimensions of these interactions are, as well as who the people involved are. Juel et al. [23] describes the ethical considerations leading to the design conceptualization of the SMOOTH robot in more detail. It is out of scope for this paper to go in depth with this area but the considerations described in [23] were also used in the selection process and later also in the testing of the robot prototype.

3 Existing Technology in Context of SMOOTH

In this section, we present four robots (Fig. 3), which were chosen since they are commercially available and they fulfil at least one of these three requirements: service, personal assistance, and logistics needed for the execution of at least one of the three use cases. We will discuss and analyze their strengths and limitation in the context of the SMOOTH project and from that derive a set of design requirements for the SMOOTH robot.

The Care-O-bot 4 [24] (Fig. 3a) is a mobile service robot with grasping ability. Due to the design of the robots spherical joints the work space of the Care-O-bot 4 is extended compared to earlier versions while enabling a 360-degree rotation of head and torso. The robot is modular such that it can be equipped with up to two arms, trays, a 'head' or just be used as a mobile base.

Strengths: The novelty of the Care-O-bot 4 is its modularity, which makes it relevant to many scenarios, while the



(a) Care-O-Bot 4 (b) Pepper

(c) MiR100 (d) TIAGo

Fig. 3 Robots suitable for the healthcare system

height and design of the robot facilitates human-robot-interaction. Therefore, this robot could be applicable for the guiding use case and perhaps also for the water serving use case.

Limitation: The problem of grasping has been solved for controlled environments in industry but is still problematic in less constrained environments. This limits the robot to very specific serving tasks. Due to its lack of ability to carry a module or object (too large for the arm) the Care-O-bot 4 is not applicable for the logistic use case in SMOOTH. The price, depending on the modules chosen, is between 80.000 Euro and 130.000 Euro without the arms, and around 40.000 Euro per arm in addition.

Pepper [43] (Fig. 3b) is a personal assistant robot. Pepper is designed as a day-to-day companion with the emphasized ability to perceive emotions. Pepper is designed to communicate with humans through its body movements and by voice. The pepper robot has different sensors such that it can recognize faces and speech, and move autonomously.

Strengths: Pepper designed to be very appealing and express emotions while also being mobile. At a height of 1.2m it can talk to both standing and seated people and is therefore relevant to the guiding use case. Pepper could function as a companion that welcomes and guides elderly while being entertaining and emotionally aware. The price of Pepper is around 20.000 Euro.

Limitations: The Pepper is not modular, hence it cannot be used to solve any logistic or service tasks. Pepper can only lift up to 250g, hence it cannot be used to carry sufficient weight to fulfil the requirements of logistic (use case 1) or service (use case 2) tasks, which makes it not applicable for the SMOOTH project.

The MiR100 [28] (Fig. 3c) is a mobile logistic robot used for the automation of internal logistic tasks. The philosophy behind the robot is to optimize workflows and to release employee resources such that a company can increase productivity and reduce costs. MiR100 is used both in industry and the healthcare domain for logistic tasks, and also utilized as mobile base for other robots such as the UV Disinfection Robot described in [5]. Other mobile robots like TUG [1] are also used in logistic tasks in the healthcare system.

Strengths: The MiR100 can solve many types of logistic tasks with a payload of 100 kg. It is designed such that customized modules can be mounted on top, thus adapting the robot for various applications sharing the same base platform and navigation system. The price of this robot is between 20.000 Euro and 30.000 Euro.

Limitations: The bulky design is not optimal for interacting with humans, and has a lack of relevant sensors to detect what people are saying or doing due to its low height of 352 mm. In theory these could be solved by adding appropriate modules, but nothing appropriate exists yet to our knowledge. These limitations would generally limit its relevance for SMOOTH since seamless human-robot-interaction is a critical requirement.

The TIAGo [31] (Fig. 3d) is a mobile manipulator robot. The TIAGo base is a PMB-2 platform which is a mobile robot with a ø 54 cm footprint. The base of the TIAGo includes by default a laser range finder and an IMU and additional items such as ultrasound sensors and microphones can be included. It also has built-in speakers to inform its users about tasks. On top of the base, the "body" of the robot is installed which can be purchased in three different configurations. The TITANUAM configuration includes amongst other things a 7 DoF arm with a 5 finger gripper mounted, an extendable torso and a pan-tilt head. The arm has a 3 kg payload and an 87 cm reach while the torso can extend 35 cm.

Strengths: The TIAGo is generally modular with a lot of functionalities that makes it an interesting research platform. It facilitates HRI by having explicit anthropomorphic features, a voice system and, pan-tilt functionality. The robot also has the required sensors for state-of-the-art robot perception and HRI. For manipulation, the robot can utilize its long range combined with the lifting mechanism in the body. Hence, it can reach where other robots normally cannot.

Limitations: One of the primary goals of the SMOOTH project is to enable the robot to operate seamlessly with humans. From our experience working with complex robot design, we believe these robots have to be designed with simple mechanisms and functionalities. The TIAGo, with robot arm and gripper, lifting torso etc. has many complex mechanisms and functionalities which might give complications when implementing it in uncontrolled environments. The robot with its torso does not apply well to logistic tasks, but without the torso a hook mechanism or likewise could simply be added.

The robots outlined above represent available potential solutions to some of the identified use cases. However, none of them can be generalized to all three use cases without becoming economically infeasible and technically too complex. It follows from this that a modular design with simple mechanics which allows for the robot to solve multiple tasks is desirable. In the following we will describe the last steps in phase 2 of the design process, here we use the perceived strengths and weaknesses of the above robots to derive design criteria for the SMOOTH robot.

4 Requirements for Welfare Robots

In industry, robots have been accepted and adopted into the daily work environment for more than 20 years. This was enabled by having a controlled environment where humans and robots are separated by fences. Today, collaborative robots and humans share the environment rather than being separated, enabling collaboration on tasks. Such collaboration is essential in welfare robots that are autonomous and mobile, for handling less structured environments. Welfare robots are required to navigate and manipulate in changing environments, while communicating with the user. One of our primary goals when developing welfare robots is enabling the robot to operate alongside humans – in our case, the caretakers and residents at elderly care facilities.

In creating this collaboration, we face some specific technical constraints. In contrast to industrial robotics, grasping and manipulation is much more complex in general scenarios. In elderly care institutions and at hospitals the environment changes frequently and objects are different. Today's welfare robots that grasp physical objects are still only research projects, e.g. [14]. To create a technically feasible welfare robot that can be implemented and accepted at hospitals and care facilities we believe that we are for now required to avoid having manipulators in the form of robotic arms and instead use less dexterous devices to manipulate in the environment. Regarding the design, mobile welfare robots have to fit into the design context of the care center meaning a wild use of colors should be avoided (also to minimize irritation from people with dementia). Passive and lighter colors should be chosen and the shapes should be round without too much variation in the design. This will also help the robot being recognized as a machine without getting an industrial look. In comparison to industrial mobile robots, our type of robot should also be designed with some extent of anthropomorphic features to increase the surrounding humans understanding of the robot. For industrial mobile robots simple mechanism like light is used for this (see the MIR100 on Fig. 3c).

In order to facilitate the acceptance of a welfare robot by the staff, patients, and residents in hospitals and elderly care facilities, it is required to design the appearance and behavior of the robots in an appropriate way to ensure the dignity of the humans interacting with the robot [23]. Our concepts for realising navigation, perception and HRI will be described later, whats important here, is the robots ability to at some level perceive the expectations and capabilities of the residents and patients. Human interaction is successful because we are able to predict each other's actions and reactions. We believe it is essential that a welfare robot has the same ability, as much as possible. Therefore, it is important that it can read body language and understand a complex scene of interactions. Essentially the human perception of a welfare robot is shaped by its behaviors and physical design. In our view, it is crucial that welfare robots are able to anticipate human actions and proactively act on these, to arrive at smooth and hereby acceptable behaviors (see Sections 6.2, 6.3 and 6.4 for more details on this)

For the healthcare system to truly adopt and accept welfare robots we believe four design aspects are key in the development process of these:

Affordability: As all governmental institutes operate on limited budgets, it is important that welfare robots are not overly expensive. It is often hard to determine how much value a welfare robot will create, so the decision whether or not to buy it will often be based on price instead of created value. To ensure an affordable price, the robots need a simple design and mechanics of limited complexity.

Modularity: A strong business case can be facilitated through a modular design. This enables the robot to solve different tasks, allowing it to serve multiple purposes by using different modules for each task and thereby creating value. Such tasks could include logistics, aid in communication, guidance, service, and serving.

Simplicity: Installing a robot at a facility should not overly disrupt workflows, to avoid irritation and negative attitudes towards the robot before it is even put to use. Likewise, the daily use of the robot should be as uncomplicated as possible for the users – i.e users should not be interrupted in their daily work to service the robot. The interaction between the robot and humans using it should be simple and intuitive. This should be facilitated through a user-oriented design process when creating both the design of the robot and the graphical user interface.

Acceptability: While ease of use has a big role, the physical attributes of the robot also plays a role in whether users accept it into their workflow. It is important that the robots design allows it to convey its intentions as well as its internal state, and avoids facilitating misuse of the robot.

5 SMOOTH Welfare Robot

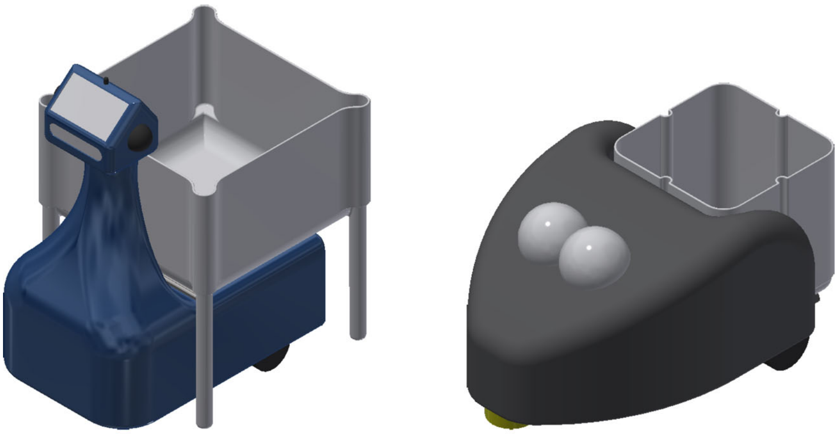
In phase 3, three initial robotic design concepts have been discussed for solving the three use cases in the SMOOTH project: The Swan, the Mouse and the Penguin (see Fig. 4). Each design was discussed with respect to the requirements defined in Section 4.

The requirement of affordability has been facilitated trough a economic distribution of sensors, in particular the safety approved laser scanner which is among the most expensive parts of the robot. Modularity is solved by all three designs by having interchangeable attachments, which can be chosen depending on the current use case. Simplicity of interaction has been partially solved in all designs through a richness of sensorial modalities. To increase acceptability we purposefully chose a design with minimal anthropomorphic features (as suggested in [12]).

5.1 Initial Designs

All three initial designs are built on the same three-wheeled mobile platform. The platform features two actuated wheels, and a single caster wheel, allowing the robot turn around the axis between its driven wheels. The platform uses a single safety laser scanner in the front – seen as the yellow circle in the bottom left of the designs in Fig. 4. This means that the robot is designed to drive forward, and is only in certain conditions allowed to drive backwards. For this cheaper 3D vision sensors and bumpers will be used to avoid collisions. The designs differ in the way they handle the modular attachments, as well as the sensor kit and user interface on top of the mobile platform.

The robots have one clear front facilitated by using screens and for the mouse sensors encapsulated in transparent plastic to create the anthropomorphic feature of a face, simply to make it easy for humans to know where to interact with the robot. The Swan and the Penguin have a touch interface, in tablet size, on the back of their



(a) Swan (b) Mouse

(c) Penguin

Fig. 4 Three different initial design suggestions all bases on the same wheel configuration

“head”. This is for input of navigation information and other steering functions. In all designs, sharp edges have purposely been avoided to create gentleness but also for safety. Different heights was explored to discuss safety towards humans and in terms of general stability. Some height for all three designs was required for the microphones being high enough to ensure good speech recognition.

The Swan: The design (Fig. 4a), uses a liftable platform to carry the attachments, which have legs to facilitate that the robot can drive underneath them. The design includes a elongated neck with a sensor head on top, which is the primary user interface (UI). It contains touch screens and vision sensors in the front and back, speakers on both sides, and a microphone on top.

The Mouse: The design (Fig. 4b) is a small logistic robot, that solves the problem of modularity by having attachments with wheels, which it can drag around. The design has speakers and a microphone, but does not have the same type of UI as the Swan.

The Penguin: This design (Fig. 4c) is a taller version of the mouse. It uses the same attachment system as the mouse. It has a tall body with a UI at the top, similar to the Swan. The height of the robot makes the design suitable for social interaction, while also allowing for control using a touch screen within standing reach.

5.2 Design Selection Process

in the last part of phase 3 one design concept of the three had to be chosen for further conceptualization and to serve as the basis for the creation of a low-fidelity mock-up to be tested at ØECC. We combined the input from the co-design workshop with the requirements and vision that we wanted to pursue in the project which included limitations related to technical feasibility, safety, and costs.

In regards to the category of acceptability we can primarily analyse the appearance of the robot. The Penguin and the mouse has only round edges which creates clean and simple design that can be more appealing for the users.

An advantage for the Penguin is the fact that it mainly consist of large surfaces which facilitate easy cleaning of the robot and promote better hygiene. The way the Penguin and the Mouse carry the attachments also makes them more hygienic than the Swan since the bottom of the laundry bin (part close to the ground) will not be in contact with the footprint of the robot like for the swan. The Swan also has a variational design that can course irritation and aggression for people with dementia.

In the context of modularity, the designs are indistinguishable due to the fact they can all carry the different attachments.

The category of affordability is quite equal for all robots, due to the use of a single safety laser and only simple mechanical elements. The Swan and Penguin are larger robots with added screens this makes them slightly more expensive but also significantly better for HRI, where the small footprint of the mouse makes the robot cheaper.

For simplicity, The legged design of the Swans attachments limits the safety lasers FoV. The way the Mouse and Penguin carry the attachments was made to avoid having the attachments block the safety laser FoV. The height of the swan and penguin makes them suitable for interaction, while also allowing for easy user control using a touch screen within standing reach. The protruding platform of the Swan might lead to potentially dangerous situations since it can be inviting for people to sit on. The Mouse has a low profile which makes it unsuitable for interacting with users in general and for the guidance use case in particular. The height also complicates the vision system of the robot as cameras would have to be tilted upwards, in order for the robot to see peoples faces. The low height improves stability and therefore an increase in safety compared to taller robots but people might easily stumble over the robot

Taking all of this into consideration the design concept of the Penguin was chosen for further conceptualization which is described in Section 6. Followed, will the extended conceptualization of the Penguin be presented with 3D prints, real-size mock-up and 3D drawings at ØECC (see Section 7).

6 The Penguin

Figure 5 shows the end result of phase 4, a refined version of the Penguin design. The refinement contains both changes to the mobile platform and the UI hub. The original three wheel design had the caster wheel placed behind the safety laser, but by moving the safety laser up the wheels can be placed beneath it providing better protection. To make sure that the full 270° FoV of the safety laser is unobstructed in the new placement, a groove has been added to either side of it (see Fig. 5f).

The changes to the UI hub includes placement of four vision sensors (explained in Section 6.1) as well as added detail to the screens. Figure 5c shows animated eyes on the front display, while Fig. 5a shows a map on the back display, which is visible during the guidance use case. Besides the changes to the UI hub, two vision sensors have been added on the body of the robot, one in the front and one in the back. The vision sensor in the front is used for general object detection for navigation where the one in the back is used when driving backwards to pick up add-on modules. Figure 5a and b shows possible attachments for the laundry and drink use cases. The bin can contain 75 liters, making it well suited for carrying laundry or garbage. It has a handle and wheels, which makes it easy to push it around when it is not attached to the robot. The rolling serving tray is one possible attachment for the drink use case.

In the following we will discuss the sensor head in more detail (Section 6.1), our sensor processing (Section 6.2), our navigation strategy (Section 6.3) and the proactive control scheme of the SMOOTH robot (Section 6.4).

6.1 Sensor Head

The head features two displays, four 3D vision sensors, two speakerphones for dialog, and a microphone array for sound localization (see Fig. 5g–i). The displays are 7" LCD screens each powered by a Raspberry Pi 3, which enables the rest of the system to communicate with the display via ROS. The front screen will show animated eyes, and not be used inputs by the user to avoid the discomfort of poking something in the "eyes". The back screen will facilitate the main way of instructing the robot non-verbally.

The vision setup consists of four cameras (Intel RealSense D435): one in the front, back, and on either side. The resulting field of view for both the cameras and the laser scanner (Hokuyo UAM-05LP-T301) is illustrated on Fig. 6. The laser scanner covers 270° which is shown as the red and dark blue areas. This means that the robot is not safety approved to reverse, and therefore cheaper sensors such as ultrasound or bumpers are to be explored. The head contains a single front facing camera for human robot interaction. The three backwards facing cameras are mainly used for the guidance use case, where the robot should drive in front of a person, while keeping the person in its field of view. Because of this, the cameras are placed in such a way that both the blindspot (shown as white area in Fig. 6) and the overlap (green area) are minimized.

6.2 Sensor Processing

The robot needs to use various sensor processing modules to solve the different use cases. Two of the main modules

Fig. 5 Visualization of the final design. **a** and **b** show different modules for the logistic and drink serving use cases and **c** illustrates the robot being applied to the guidance use case (background: stock image from Colourbox). **d**, **e** and **f** shows the robot without modules and the sensor head from different angles. **g**, **h** and **i** shows close-ups of the sensor head

are a people detector for the guidance and water delivery use cases, and a add-on module detector for the logistic and drink serving use cases, or future use cases using new add-on modules. The developed human detector uses the convolutional neural network (CNN) OpenPose [7] to detect 2D keypoints for all humans within the RGB field of view of the robot – see Fig. 7a. These keypoints are then converted to 3D by looking up the pixel coordinates in the point cloud. The 3D coordinates are used to crop the people pointclouds and determine the orientation of their torso. The 2D keypoints are also used to crop out the face of each person for facial recognition using FaceNet [40].

Besides detecting people, the robot also needs to detect the various add-on modules in order to pick them up. A simple solution to this problem is using AR markers, which enables unique identification and pose estimation of each module. This approach was used for the first demonstration of the logistic use case at ØECC – see Section 8. Figure 7b shows the bin module mockup used for the first demonstration, including the AR markers which are used to estimate the pose of it. The bin has AR markers on each of its sides, allowing it to be detected from multiple angles. However, the AR markers are not pleasant looking, especially in a nursing home setting. Therefore a

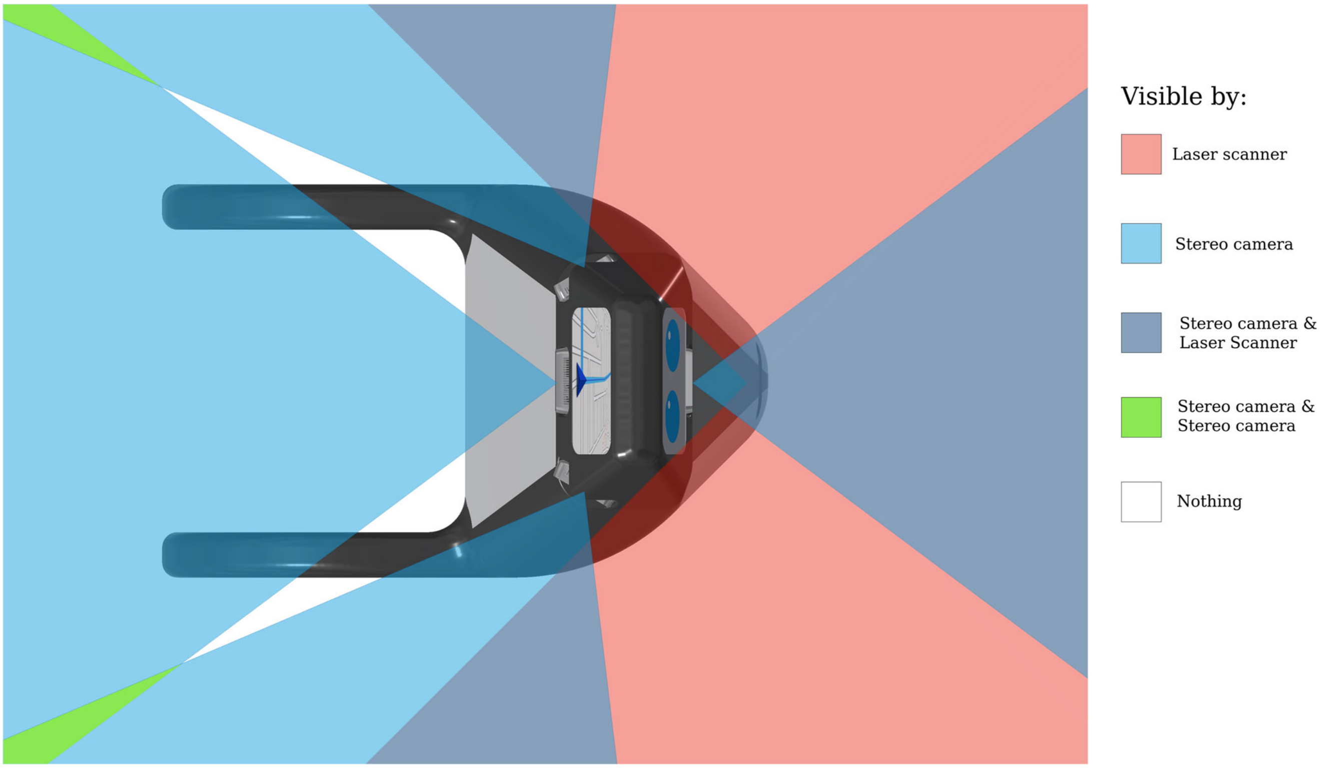


Fig. 6 Visualization of the field of view of the for cameras in the head, and the laser scanner

pose estimation solution which works on arbitrarily looking modules would be preferable.

Such a detection module has been developed using a CNN. The flow of the algorithm using a regular trash bin as example can be seen in Fig. 8. The module replaces a global pose estimation step with a pre-segmentation of the point cloud using a CNN for image segmentation. The network is trained on a data-set of auto-generated images, where cropped images of the bin is placed randomly on images from MS COCO [26]. This enables the method to work on new modules without requiring manual annotation of a data-set where the module is placed in relevant context. By making the background completely arbitrary the network learns to separate the bin from any background. After the add-on module is segmented the structured point cloud is cropped using the segmentation mask, leaving only points belonging to the bin. This allows for the pose to be estimated using only ICP [4].

Besides humans and our developed add-on modules, ØECC is also filled with various other types of objects. These can be detected using object detection networks such as Mask-RCNN [19]. Using Mask-RCNN has the added

(a) (b)

Fig. 7 Example of different sensor processing modules. **a** Human detector based on OpenPose [7]. It uses the detected keypoints to crop the person pointcloud, estimate torso direction, and recognize the face. **b** AR marker detection on bin module mockup using the *ar_track_alvar* ROS package [2]

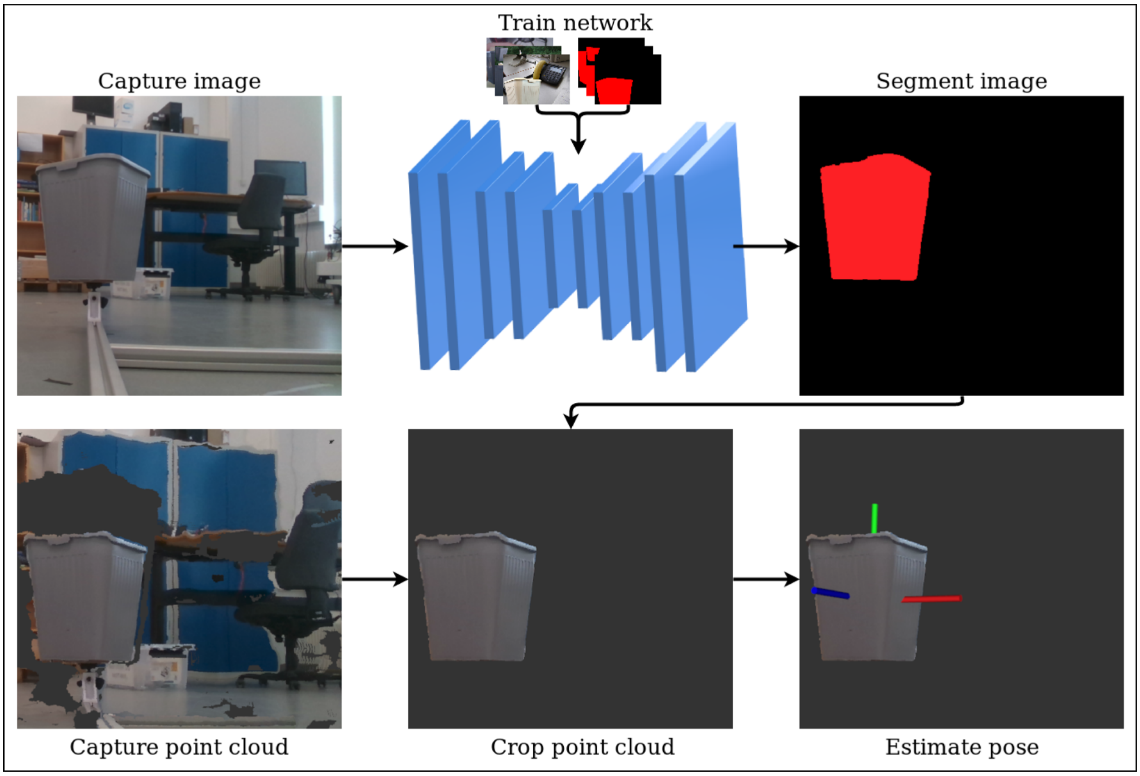


Fig. 8 Flow diagram of developed detection module for the bin add-on module used in use case 1

benefit of providing segmentation masks, which can be used to create cropped and semantically labelled pointclouds for each individual object. The process is as follows:

1. Use 2D image instance segmentation network to detect the objects in the image, and estimate their label and segmentation mask – see Fig. 9.
2. For each detected object, create a semantic pointcloud by extracting the points using the segmentation mask and label it with the object class.
3. For each object cloud remove points not associated to the actual object – e.g. using Euclidean cluster extraction or some simple median calculation.

To be able to run the network on the robot, we use an implementation [29] of Mask-RCNN based on the MobileNet [22] architecture.

6.3 Acceptable Navigation for Care Centers

For the navigation of the robot we use the ROS navigation stack [37]. The most basic configuration is to use the sensors to detect occupied areas and mark this area as an obstacle. Thus the robot will be aware of the placement of obstacles, but without any semantic information regarding the detected objects. This approach is sufficient when the objective is to maneuver and avoid collisions but does not enable the robot to smartly maneuver in the space of the encountered objects, based on object specific constraints. Including the semantic information from the sensor processing modules described in the previous section enables the robot to have such functionality. This is implemented in the ROS navigation stack using layered costmaps described by Lu et al. [27]. The different detection modules can thereby feed into their own costmap layer, which are then merged into a master costmap. Each layer has its own settings for generating the costs, which allows humans to have a larger cost gradient than chairs, etc. This also enables the robot to take into account various dynamics and social conventions regarding

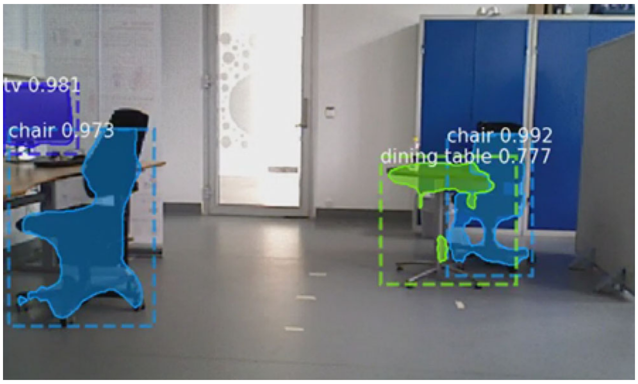
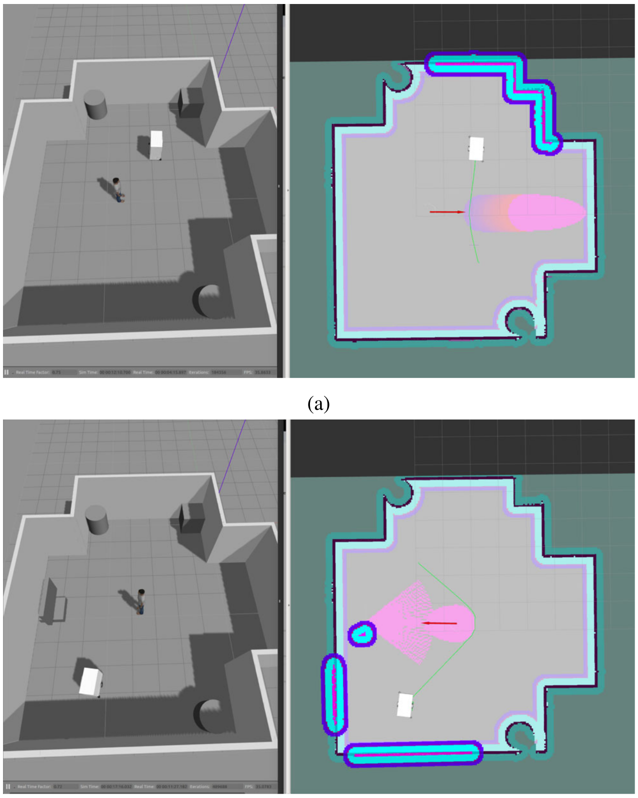


Fig. 9 Example of 2D output of Mask R-CNN in our laboratory


(a)

(b)

Fig. 10 Two different simulation scenarios: person moving + robot moving, person standing watching TV + robot moving

human-human and human-object interactions. This could for example be avoiding driving between two people having a conversation, or parking behind a chair which is sat on by a person.

On Fig. 10 two simple examples of our context aware planning and navigation is shown. The images on the left is from simulation and the image on the right is the costmap where also the robots trajectory is marked with a green line. This makes use of the human detector (red arrow) and object detection. On Fig. 10a the person is moving and by using the human detector we detect the human, it's motion and future trajectory. We then know where the human will be a time-step into the future and can therefore plan a trajectory for the robot that avoids collision with the human. On Fig. 10b a person is standing in front of a TV. Here we detect the human and the TV and set a cone-formed high cost in the space between, since we do not want the robot to intrude the space between the human and the TV, leading to the planed robot trajectory being behind the person.

6.4 Proactive Control

The control scheme of the SMOOTH robot combines two main components: 1) Multi-sensory integration for adaptively combing different sensor types (e.g., vision, sound, laser rangefinder) and 2) proactive control for autonomous learning to anticipate human behaviors and to perform proactive responses. This approach will result in predictable and comprehensive actions of the robot with natural human-robot interaction as shown in Haarslev et al. [17].

Human-human interaction is smooth and multi-modal, involving the processing of information from visual, auditory, and tactile senses. Smooth movements influence a robot's apparent animacy, unpleasantness and likability [8]. Multiple sensory modalities offer information redundancy and predictability, which can subsequently reduce overall movement errors and increase robustness as well as smoothness.

Conventional robot controllers, for human-robot interactions, typically require prior models of the environment (e.g., deliberative control with knowledge model [25]). Adaptive sensor-driven controllers on the other hand directly link perception to action. They can deal with unpredictable events better than controllers based on the sense-plan-act or deliberative control paradigm [9]. Multi-modal sensor-based control can be beneficial for smooth, naturalistic robot behavior.

The SMOOTH robot utilizes a multi-modal learning-based model for fusing sensor information [41] irrespective of modality and generating motor commands. An instance of this model is illustrated on Fig. 11. Here, we illustrate the use of the model to 1) adaptively fuse auditory and visual spatial information about target (e.g., sound source or human speaker) location relative to the robot and then 2) generate motor commands for smooth robot orientation behaviour. In this case, the auditory information acts as earlier (predictive) sensory feedback (i.e., the robot can hear the sound earlier in time from far away) while the visual information acts as later sensory feedback (i.e., the robot can see the target later in time in its proximity when approaching it). The adaptive sensory fusion is computed via a single multisensory neuron model with learning (blue block in Fig. 11). This neuron directly computes the angular motor velocity v of the robot as the weighted sum of auditory and visual directional information $x_a(t)$ and $x_v(t)$, respectively. This auditory-visual spatial information fusion is therefore simply modelled as $|v| = w_v x_v(t) + w_a x_a(t)$. Here, w_v and w_a are the weighting terms corresponding respectively to the visual and auditory spatial direction information. For learning, we employ two intramodal or unimodal learning rules (green blocks in Fig. 11), one for vision and the other for audition, along with two crossmodal learning rules (grey block in Fig. 11). The learning rules adapt the weights online on a moment-by-moment basis (see [41] for the learning equations). This implies that weighting of auditory spatial information is simultaneously and independently influenced by both the visual spatial information as well

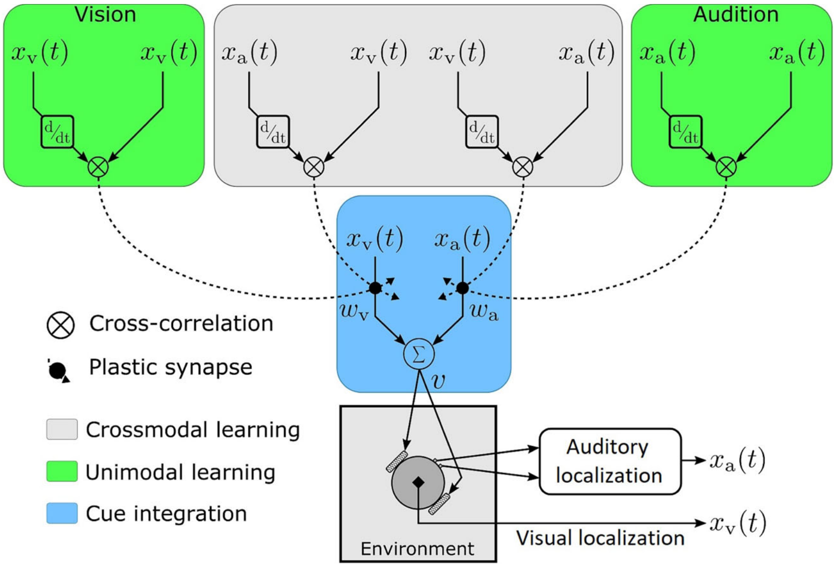


Fig. 11 Crossmodal learning-based controller architecture for sensor fusion that drives the orientation behaviour for the SMOOTH robot

as the previous sample of auditory spatial information and vice versa. The symmetry and independence of the crossmodal learning rules across modalities and intramodal learning rules within each modality ensures that the model compensates for errors in sensor data in either modality. The fast temporal correlation learning-based [33] algorithm used in the learning rules ensures that the weights quickly converge to their instantaneously optimal values when the auditory and visual sensor information are spatially congruent. This ensures that the robot executes smooth movements. This online learning-based sensor fusion model also allows for robot operation even if sensor information in a modality is unavailable, for example when a sensor suddenly fails. Exploiting temporal correlations between sensor modalities also allows this model to realize proactive control which generates proactive and smooth movements from predictive (earlier) sensory feedback (i.e., here, auditory feedback). This can help to determine intention in human-robot interaction; thereby leading to natural and smooth interaction.

7 End of Design Iteration and Formative Evaluation at ØECC

In conclusion of the current design iteration, we conducted a qualitative formative evaluation with a polystyrene low-fidelity mock-up of the robot concept visualized on Fig. 5 at ØECC [34]. During the evaluation, we engaged 2 residents, both males, and 4 caregivers, all females, in enacting the three use cases. During the test we applied the standard of ethics expected from scientific studies involving people, we have delivered a detailed informed consent form to the participants and anonymized the data recorded.

The test was framed in three stages: an introduction stage in which we described the goal of the visit, our mock-up and the program for the test. During the second stage, we observed and filmed the participants enacting the use cases. In the end, during the third stage, we conducted a shared discussion with the participants. The discussion was conducted as a semi-structured interview supported by design samples [34], hence, we engaged in a discussion starting from a few pre-written questions and we left the participants to talk freely on their experience during the test and only referred to our questions if needed. In order to trigger more comments from the participants, we showed them a series of animated videos showing the robot concept on Fig. 5 executing the use cases.

The filming was used to conduct a video ethnographic analysis, focusing on the dialogue taking place among the participants, and eventual remarks on the situation, body language, and facial expression [32]. Our analysis aimed at gathering requirements and impressions from the participants on four aspects:

- Aesthetics of the mock-up, including colors, shape and inspiring mood;
- Needed functionalities, feedback on the functionalities and affordances that we imagined and suggestions for new;
- Impressions on the overall experience, such as emotional responses, elicited mood and desires;
- Potential challenges and suggestions that we could not foresee, to re-conceptualize our mock-up and scenarios.

Starting from the aesthetic aspect, the shape of the robot was found pleasant for its round, chubby shape and the big eyes of the face. Imagining the robot as an assistant and tool was commented positively, however, one of the residents commented that "it had to look like a machine!" and not as a real person or pet. According to his perspective it was important for them that the robot communicated clearly that it is a machine, and the other participants agreed. It should look like an appliance that could serve specific purposes, while a too realistic anthropomorphic or animal-like look would be confusing, especially for residents affected by severe forms of dementia. As for the colors, the residents commented that a neutral color, like gray, was to be preferred to a more distinctive shade. One of the caregivers commented that the residents affected by severe dementia could react violently to colors like black or red. Moreover, the caregivers added that demented residents would prefer a complete face for the robot, saying: "It lacks a nose!" this could be perceived as strange by the residents. One of the caregivers suggested us to take inspiration from the dolls [39] that are used in therapy for people affected by dementia. These dolls are characterized by a round face with round eyes, a small round nose and stitched mouth arched into a smile. Finally, static facial expressions are preferable to dynamic changes of expressions, as these might be interpreted as a sign of initiating a conflict.

Positive feedback was provided on functionalities and overall experience with the mock-up. Interestingly, the participants controlling the prototype acted politely with the residents, addressing them with greetings and questions showing personal interest, such as: "Have you slept well?" or "How are you feeling?" and "Should you have something for lunch today?" The residents responded with the same politeness, answering the questions and adding specific remarks, for instance, one of the residences pointed at the courtyard outside saying: "See there? They throw cigarettes buds, isn't it bad?" At those remarks, the participant controlling the robot would answer nodding, as to show that their remark was heard and to show empathy. At the same time, when the caregivers were enacting the laundry and garbage scenarios, they would call the robot with a slightly loud and playful voice, as if calling a dog or a small child. Then the caregiver controlling the robot would come and answer cheerfully to the call as if confirming that the robot has heard the call and was ready to receive its task. We interpret this interaction as confirming our expectation that the robot has to be able to interact in a polite way as if engaging in a real conversation and more design explorations will be conducted in this front.

Critical questions were raised in relation to how the robot could concretely fit the scenarios and the need of ØECC. A requirement that emerged was that the robot had to keep a certain distance when approaching the resident, as some people, especially those affected by dementia, might feel uneasy if the robot would move too close. This requirement was already discussed within the consortium, however, testing the "robot" in context made us more aware of this issue.

One of the caregivers asked why the robot does not have any arms, as arms would enable the robot to grasp on things and be more independent in specific tasks. In fact, robotic arms were not included in the design purposely as described in Section 4 due to technical constraints. In this respect, additional critical comments were addressed by the caregivers to the drinking scenario. Already during the enactment of this scenario, it appeared evident from the affordance of the mock-up that a person has to place the glass on the robot. This means that the robot cannot independently tidy up and take the glasses back to the kitchen. Further, each individual resident has specific necessities and must be served individually, as some have diabetes and need unsweetened drinks, while others need other specific drinks containing milk or cream, depending on their blood pressure or other health conditions. This means that the caregivers have to be able to trust the robot to be able to serve the right drink to the right individual, demanding expensive technical features and critically increasing the price of the robot. For these reasons the drinking scenario has been under-prioritized and will be kept into account for the deployment of the robot in other contexts like conference centers or hotels where people can serve themselves from the robot. The other scenarios appeared instead to be promising and rich comments were given on the guiding scenario, which led to imagining other collateral scenarios in which the robot might be used to remind the residents of other tasks or events they have to participate in.

In conclusion, the participants expressed curiosity and interest for our mock-up and actively engaged in enacting the scenarios. In the final discussion, the participants provided relevant feedback and critical comments, which has been taken into account through later design iterations. Moreover, in this test we involved only male residents and female caregivers. This could undermine the validity of our insights for the totality of the residents, it would be ideal to also involve a small group of women residents and male caregivers in the next test, even if women are the most represented gender among the caregivers at ØECC, to eventually gain richer insights from both genders.

8 The SMOOTH Robot Prototype

The design process described throughout the article has led to the development of a robotic prototype, see Fig. 12. It

Fig. 12 First prototype of the SMOOTH robot. **a** shows the robot with all its electronics exposed. **b** and **c** shows front and back views of the robot complete with covers and everything installed. **d**–**f** shows close-ups of the sensor head

has the same basic shape as the penguin CAD model on Fig. 5, but alters it a bit due to practical and manufacturing constraints. Figure 12a–c shows the body of the robot from different angles, with and without covers. Internally it is comprised of three different parts. The bottom is a mobile base and includes the safety laser, batteries, motors, brakes and other robot electronics and mechanics. Also, the two object detection cameras reside here – one front facing and one back facing. Because of the three wheeled design, stability and balance are issues which needs to be thought of in the design. Because of this, the batteries are placed along the sides of the robot, moving the center of mass further back. This ensures that the robot does not fall, even when the caster wheel is rotated to the side and the robot operates on uneven surfaces.

The middle part with the dark gray cover contains all the processing units. This includes Nvidia Jetson TX2's for neural network processing, as well as Intel NUC's for navigation, speech recognition and speech synthesis. It also contains the microphone used for the speech recognition, see Fig. 12c. On Fig. 12a this is supported by metal plates. At a later stage of the building process when the covers were added, the metal support was removed. Instead the computers rests on shelves attached to the cover.

The top part is the sensor head which is showed in detail on Fig. 12d–f. It contains the two screens, as well as the four stereo cameras used for human-robot interaction. The head is constructed of various laser-cut plates and 3D printed parts. This allowed for rapid prototyping of sensor placements and screen positioning, leading to two of the backwards facing stereo cameras being positioned vertically instead of horizontally. This was done such that more of the guided person is in the field of view during the guiding use case. The two screens will be used to initialize use cases, display information, as well as showing intent and social cues, via animated eyes etc.

9 Discussion

The SMOOTH robot follows a technical framework based on ROS like the HOBBIT robot [14]. In comparison to HOBBIT, which uses 3D sensor data projected to 2D for SLAM, the SMOOTH robot uses a laser scanner placed in the mobile base. This provide the robot not only with a safety capability, but also with 2D perception. This is long range, has a wide FoV and high resolution. It is a safety certified device that brakes the motors when an obstacle is detected inside a certain range. This allows for the mapping and localization capabilities of the robot to consider bigger areas of the environment. For this the robot uses Cartographer [21], instead of Gmapping [16] and AMCL [15] as used in [14]. Cartographer reduces the computational requirements by matching the laser scanner readings with the submaps created, computing loop closure constraints and graph optimization. This allows the robot to map bigger areas, fitting the requirements of the environment.

Contrary to the HOBBIT robot, SMOOTH splits the path planning calculation into a global and local planner. While HOBBIT's SBPL global planner [10] considers the kinematics of the robot [14], SMOOTH's Dijkstra global planner [13] assumes the robot as a point, being able to calculate long paths faster in large maps, leaving the kinematics consideration to the local planner. The local planner is in charge of the definition of the robot command velocities based on the constraints around the robot in a small area (costmaps, robot kinematics, global path). The local planner used, Time-Elastic-Band Planner [38], optimizes different trajectories based on the topology of the environment, in this way the robot rarely falls into a local minima. It also performs calculations in order to estimate poses of dynamic obstacles appearing in the layered costmaps [27], see Section 6.3. By considering smaller areas for the local planner, the path planning computation time is reduced while keeping the reaction capability to moving objects.

Similar to HOBBIT, the SMOOTH robot is designed to operate in a human populated environment, and thus requires human detection capabilities. The developed human detector uses OpenPose [7] which is a bottom-up approach in comparison to the HOBBITS top-down approach. This entails that OpenPose's computation complexity does not scale linearly with the number of people in the frame. The point cloud extraction and orientation estimation build on top of OpenPose makes it possible to use the human detector directly in the navigation by creating semantic costmaps. This is further enabled by the other object recognition modules used on the robot like the add-on module detector or Mask-RCNN [19], see Section 6.3. Mask-RCNN provides instance segmented point clouds by segmenting the 3D data using using 2D masks. However, it is computationally expensive which is not specifically suited for embedded devices. It could easily be replaced with a newer more lightweight instance segmentation model using either 2D or 3D data.

For decision making, a waypoint server has been implemented to easily manage robot goals in the map, which could potentially be used within a graphical user interface. On top of the navigation and perception systems, a mission handler was implemented in order to handle the execution of the use cases. It acts as the coordinator between the different layers, assigning tasks to the robot based on triggers generated by the different actions. The implementation of the mission handler allows the system to have parallel decision making.

Summarizing, the systems used in SMOOTH aimed for the usage of the robot in a large, human populated and dynamic environment, optimizing already existing packages in order to adapt it to the robot needs.

10 Conclusion and Future Work

Based on a design process following four phases that include a participatory design scheme, an ethnographic study, a focus group study and an analysis of commercially available robots that are applied or at least applicable in the welfare domain, the design of a novel welfare robot has been derived.

The development of three initial design concepts were guided by and evaluated based on user input derived from workshops and interviews. Description of technical requirements and criteria of affordability, modularity, simplicity and acceptability were made in order to facilitate the development of an applicable welfare robot that can mitigate some of the challenges of the demographic change. Through a selection process involving the project partners and users, the design concept of the Penguin was chosen for further conceptualization.

During the participatory design process, we discovered that the appearance of a welfare robot should be designed with minimal anthropomorphic features and only provide limited features, e.g. resembling a face to facilitate interaction. The robot should rather be clearly recognized as a machine. The colors of a robot interacting with elderly and especially people with dementia should have natural colors where avoiding red is a necessity since it can have irritating effects.

To facilitate seamless interaction between the robot and its end-users a multi-modal sensor unit, including devices for providing feedback to the user, is integrated. We explain concepts for the sensor processing modules using the multi-modal sensor unit. We describe three software modules for realizing HRI, detection of add-on modules, perception and scene understanding. We also discuss how multi-modal sensor-based control is foreseen to generate proactive robot behaviors.

We compare and relate our technical concepts presented in this paper regarding used methods for realizing navigation, perception and HRI to that of the HOBBIT robot [14]. The robot navigation system uses some of the newest ROS packages in order to optimize the robot performance. It provides the SMOOTH robot with a safety grade, not only from a hardware perspective (safety rated sensors, and wide FoV covered by the laser scanner and cameras), but also from a perception and cognition side, being able to calculate optimal paths around the different facets of the environment.

We presented the first robot prototype arising from the design process described in this paper. The robot comprises of three different parts: The base (electronics and batteries), the middle part (processing units) and the top part (sensor head).

The next steps in the development process includes continued work on integrating new concepts for HRI and navigation, to create more stability and robustness when executing the use cases. At current state we have solved use case 1 by performing the garbage and laundry pick-up and transportation at ØECC. At the time of writing we are performing our first initial tests of the guiding use case at ØECC.

Acknowledgements The authors would like to thank the staff and residents at Ølby elderly care center for the fruitful discussions and valuable insights that have been shared. We would also like to thank Mobile Industrial Robots, Fraunhofer Institute, Softbanks Robotics and PAL Robotics for letting us use images of their robots.

References

1. Aethon: Tug. https://www.aethon.com/healthcare/, (accessed on 13/8/2018) (2014)
2. AR track Alvar: Ar track alvar. http://wiki.ros.org/ar_track_alvar, (accessed on 2019-07-06) (2013)
3. Bent Hansen, C.H.: Pres på sundhedsvæsenet derfor stiger sygehusudgifterne – sådan holder vi væksten nede (2015)
4. Besl, P.J., McKay, N.D.: A method for registration of 3-d shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992)
5. Blue Ocean Robotics: Uv disinfection robot. https://blue-ocean-robotics.com/uv-disinfection, (accessed on 13/8/2018) (2016)
6. Brandt, E., Binder, T., Sanders, E.: Tools and Techniques: Ways to Engage Telling, Making and Enacting. Routledge International Handbooks (2012)
7. Cao, Z., Simon, T., Wei, S., Sheikh, Y.: Realtime multi-person 2d pose estimation using part affinity fields. arXiv:1611.08050 (2016)
8. Castro-González, A., Admoni, H., Scassellati, B.: Effects of form and motion on judgments of social robots' animacy, likability, trustworthiness and unpleasantness. Int. J. Human-Comput. Stud. **90**, 27–38 (2016). https://doi.org/10.1016/j.ijhcs.2016.02.004
9. Cherubini, A., Passama, R., Fraisse, P., Crosnier, A.: A unified multimodal control framework for human–robot interaction. Robot. Auton. Syst. **70**, 106–115 (2015). https://doi.org/10.1016/j.robot.2015.03.002
10. Cohen, B.J., Chitta, S., Likhachev, M.: Search-based planning for manipulation with motion primitives. In: IEEE International Conference on Robotics and Automation. IEEE, Anchorage (2010)
11. Coradeschi, S., Cesta, A., Cortellessa, G., Coraci, L., Galindo, C., Gonzalez, J., Karlsson, L., Forsberg, A., Frennert, S., Furfari, F., Loutfi, A., Orlandini, A., Palumbo, F., Pecora, F., von Rump, S., Štimec, A., Ullberg, J., Ötslund, B.: GiraffPlus: A System for Monitoring Activities and Physiological Parameters and Promoting Social Interaction for Elderly, pp. 261–271. Springer International Publishing, Cham (2014)
12. Dereshev, D., Kirk, D.: Form, function and etiquette–potential users' perspectives on social domestic robots. Multimodal Technol. Interact. **1**, 2 (2017). https://doi.org/10.3390/mti1020012
13. Dijkstra, E.W.: A note on two problems in connection with graphs. Numerische Mathematik **1**, 269–271 (1959)
14. Fischinger, D., Einramhof, P., Papoutsakis, K., Wohlkinger, W., Mayer, P., Panek, P., Hofmann, S., Koertner, T., Weiss, A., Argyros, A., Vincze, M.: Hobbit, a care robot supporting independent living at home: First prototype and lessons learned. Robot. Auton. Syst. **75**, 60–78 (2016). https://doi.org/10.1016/j.robot.2014.09.029, assistance and Service Robotics in a Human Environment
15. Fox, D., Burgard, W., Dellaert, F., Thrun, S.: Monte Carlo localization: Efficient position estimation for mobile robots. In: Proceedings of the Sixteenth National Conference on Artificial Intelligence and Eleventh Conference on Innovative Applications of Artificial Intelligence, Orlando (1999)
16. Grisetti, G., Stachniss, C., Burgard, W.: Improved techniques for grid mapping with rao-blackwellized particle filters. IEEE Trans. Robot. **23**(1), 34–46 (2007)
17. Haarslev, F., Docherty, D., Suvei, S.D., Juel, W.K., Bodenhagen, L., Shaikh, D., Krüger, N., Manoonpong, P.: Towards crossmodal learning for smooth multimodal attention orientation. In: Social Robotics, pp. 318–328. Springer International Publishing, Cham (2018)
18. Hashimoto, K., Saito, F., Yamamoto, T., Ikeda, K.: A field study of the human support robot in the home environment (2013)
19. He, K., Gkioxari, G., Dollár, P., Girshick, R.B.: Mask r-cnn. In: 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2980–2988 (2017)
20. He, W., Goodkind, D., Kowal, P.: An aging world: 2015 international population reports. US Census Bureau (2016)
21. Hess, W., Kohler, D., Rapp, H., Andor, D.: Real-time loop closure in 2d lidar slam. In: 2016 IEEE International Conference on Robotics and Automation (ICRA), pp. 138–143. IEEE, Stockholm (2016)
22. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861 (2017)
23. Juel, W., Krüger, N., Bodenhagen, L.: Robots for elderly care institutions: How they may affect elderly care. Frontiers in Artificial Intelligence and Applications (2018)
24. Kittmann, R., Frölich, T., Schäfer, J., Reiser, U., Weisshardt, F., Haug, A.: Let me introduce myself: I am care-o-bot 4, a gentleman robot. Mensch und Computer 2015 Tagungsband (2015)
25. Lemaignan, S., Warnier, M., Sisbot, E., Alami, R.: Human-robot interaction: Tackling the ai challenges. Artificial Intelligence – Special Issue on Robotics (2015)
26. Lin, T., Maire, M., Belongie, S.J., Bourdev, L.D., Girshick, R.B., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft COCO: Common objects in context. arXiv:1405.0312 (2014)
27. Lu, D.V., Hershberger, D., Smart, W.D.: Layered costmaps for context-sensitive navigation. In: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 709–715 (2014)
28. Mobile Industrial Robots: Mir100. http://www.mobile-industrial-robots.com, (accessed on 13/8/2018) (2013)
29. Mobile Mask R-CNN: Mobile mask r-cnn. https://github.com/gustavz/Mobile_Mask_RCNN, (accessed on 2019-07-06) (2018)
30. Mutlu, B., Forlizzi, J.: Robots in organizations: The role of workflow, social, and environmental factors in human-robot

interaction. In: HRI 2008 - Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction: Living with Robots, pp. 287–294 (2008)
31. PAL Robotics: TIAGo. https://www.pal-robotics.com/en/home/, (accessed on 13/8/2018) (2015)
32. Pink, S.: Doing Visual Ethnography. Sage (2001)
33. Porr, B., Wörgötter, F.: Strongly improved stability and faster convergence of temporal sequence learning by utilising input correlations only. Neur Comput **18**(6), 1380–1412 (2006)
34. Preece, Y.J., Rogers, S.H.: Interaction Design: Beyond Human-Computer Interaction. Wiley (2015)
35. Rehrl, T., Geiger, J., Golcar, M., Gentsch, S., Knobloch, J., Rigoll, G., Scheibl, K., Schneider, W., Ihsen, S., Wallhoff, F.: The robot alias as a database for health monitoring for elderly people. In: Wichert, R., Klausing, H. (eds.) Ambient Assisted Living, pp. 225–245. Springer, Berlin (2014)
36. Riek, L.D.: Healthcare robotics. Commun. ACM **60**(11), 68–78 (2017). https://doi.org/10.1145/3127874
37. ROS Navigation Stack: Ros navigation stack. http://wiki.ros.org/navigation, (accessed on 2019-27-05) (2013)
38. Rösmann, C., Feiten, W., Wösch, T., Hoffmann, F., Bertram, T.: Efficient trajectory optimization using a sparse model. In: Proc. IEEE European Conference on Mobile Robots, pp. 138–143. IEEE, Barcelona (2013)
39. Barn, R.: Rubens barn. https://www.rubensbarn.com, (accessed on 13/8/2018) (2018)
40. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embedding for face recognition and clustering. arXiv:1503.03832 (2015)
41. Shaikh, D., Bodenhagen, L., Manoonpong, P.: Concurrent intramodal learning enhances multisensory responses of symmetric crossmodal learning in robotic audio-visual tracking. Cognitive Systems Research **54**, 138–153 (2019). https://doi.org/10.1016/j.cogsys.2018.10.026. http://www.sciencedirect.com/science/article/pii/S1389041718300172
42. SMOOTH: Seamless human-robot interaction for the support of elderly people. www.smooth-robot.dk/en/home/, (accessed on 13/8/2018) (2017)
43. Softbank Robotics: Nao, romeo and pepper. https://www.softbankrobotics.com/emea/en/robots, (accessed on 13/8/2018) (2014)
44. Spinuzzi, C.: The methodology of participatory design. Tech. Commun. **52**(2), 163–174 (2005)
45. Ventura, J., Bichard, J.A.: Design anthropology or anthropological design? towards 'social design'. Int. J. Des. Creativ. Innov. **5**(3–4), 222–234 (2017). https://doi.org/10.1080/21650349.2016.1246205

William K. Juel received a master's degree in Health- and Welfare technology from the University of Southern Denmark, in 2017. He is currently a Research Assistant of SDU Robotics at the University of Southern Denmark. His research interests include Robotics, Computer Vision and Machine Learning.

Frederik Haarslev is a PhD fellow at the University of Southern Denmark in the SDU Robotics group. His research interests include deep learning, convolutional neural networks, computer vision, robotics and human-robot interaction.

Eduardo R. Ramirez is research assistant in the SDU Robotics group. He holds a Master's degree in Automatic Control and Robotics, and a Bachelors in Mechatronics. For the last 2 years he has been working with the integration, deployment and cognition development of welfare mobile robots, both in the private sector as well as academia.

Emanuela Marchetti received the Ph.D. degree in 2016 at Aalborg University and is currently Assistant Professor at the Department for the Study of Culture at the University of Southern Denmark. Her research focuses on experience design and human-computer interaction. She is in particular experienced with user centered and participatory design methods.

Kerstin Fischer is professor (MSO) for Language and Technology Interaction at the University of Southern Denmark and leading the Human-Robot Interaction Lab in Sonderborg. After receiving her PhD in English Linguistics from Bielefeld University in 1998, she did postdoctoral work at the University of Hamburg on emotion in human-computer interaction. Her research focuses on the development of smooth, seamless interaction between humans and technologies, on usability and on participatory design techniques.

Danish Shaikh received his Ph.D. degree in BioRobotics from the University of Southern Denmark in 2012. Since 2018 he is employed as an Assistant Professor in Embodied AI and Neurorobotics at the University of Southern Denmark. His research interests include biologically-feasible learning, multi-sensory integration, computational neurorobotics, soft robotic perception and bio-inspired and biomimetic robotics.

Poramate Manoonpong received the Ph.D. degree in Electrical Engineering and Computer Science from the University of Siegen, Germany, in 2006. He is currently an Associate Professor of Embodied AI/Robotics at the University of Southern Denmark. His research interests include bio-inspired robotics, neural control, embodied cognitive robotics, exoskeleton control, and human-machine interaction.

Christian Hauch received a M.Sc. degree from the Danish Technical University in 1994 and a MBA from Melbourne Business School in 2005. He is currently employed as project manager at Robotize Aps and engaged in projects focusing on the development of mobile robots for the welfare domain.

Leon Bodenhagen holds a Ph.D. degree in robotics and is currently Assistant Professor at the University of Southern Denmark. His research interests cover topics related to robotics in the welfare domain such as human robot interaction, computer vision and multimodal perception.

Prof. Norbert Krügers research focuses on Cognitive Vision and Robotics, in particular vision based manipulation, learning and welfare robotics. He has published more than 60 papers in journals and more than 100 papers at conferences covering the topics computer vision, robotics, neuroscience as well as cognitive systems.