Intelligent Service Robotics manuscript No.
(will be inserted by the editor)

Navigation Assistance and Guidance of Older Adults across Complex Public Spaces: the DALi Approach

Luigi Palopoli · Antonis Argyros · Josef Birchbauer · Alessio Colombo · Daniele Fontanelli · Axel Legay · Andrea Garulli · Antonello Giannitrapani · David Macii · Federico Moro · Payam Nazemzadeh · Pashalis Padeleris · Roberto Passerone · Georg Poier · Domenico Prattichizzo · Tizar Rizano · Luca Rizzon · Stefano Scheggi · Sean Sedwards

Received: date / Accepted: date

Abstract The Devices for Assisted Living (DALi) project is a research initiative sponsored by the European Commission under the FP7 programme aiming for the development of a robotic device to assist people with cognitive impairments in navigating complex environments. The project revisits the popular paradigm of the walker enriching it with sensing abilities (to perceive the environment), with cognitive abilities (to decide the best path across the space) and with mechanical, visual, acoustic and haptic guidance devices (to guide the person along the path). In this paper we offer an overview of the developed system and describe in detail some of its most important technological aspects.

Keywords Assistive Robotics · Navigation · Guidance · Haptics

1 Introduction

A growing body of research [70] suggests that a reduced out-of-home mobility can have widespread, detrimental effects for older adults and ultimately accelerate the process of ageing. Adults for whom mobility is a problem certainly experience a reduction in the quality of their social life. They have fewer choices in terms of where and when they can shop, and they have been found to have problems in maintaining a balanced diet. Finally, they have a limited access to leisure and to other social activities.

Several factors have and adverse effect on mobility, the most obvious being physical impairments, loss or reduction of visual and auditory ability and of the key function of balance. Less recognised but as important is the decline of cognitive abilities such as timely reaction to external stimuli, sense of direction, peripheral vision and navigation skills [65]. Cognitive problems like these are difficult to recognise and with a very few counter-strategies of proven effect. The afflicted gradually perceives such places as shopping malls, airports or train stations as unfamiliar and intimidating and starts to withdraw [42].

The goal of the DALi project is to pursue autonomous mobility by using robotic technologies. The concrete embodiment of DALi's research is the so called cognitive walker (c-Walker). The c-Walker provides physical, cognitive, and emotional support to older adults in public environments such as shopping centres and airports. Its key features are: 1. the ability to identify the path across the space that best supports the user preferences and goals; 2. the detection and recognition of anomalies along the way; 3. the observation of humans in the area of interest of the device and the prediction of their

---

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n. ICT-2011-288917 "DALi - Devices for Assisted Living"

L. Palopoli, D. Fontanelli, A. Colombo, R. Passerone, T. Rizano, F. Moro, D. Macii, P. Nazemzadeh and L. Rizzon
Università di Trento
E-mail: palopoli@disi.unitn.it

A. Argyros and P. Padeleris
Institute of Computer Science (ICS), Foundation for Research and Technology-Hellas (FORTH)

J. Birchbauer and G. Poier
SIEMENS

A. Legay and S. Sedwards
INRIA

A. Garulli, A. Giannitrapani, D. Prattichizzo and S. Scheggi
DIISM, Università di Siena

future position; 4. the possibility to locally re-shape the path to avoid potential risks and collisions with other humans; 5. a rich set of interfaces that the system can use to recommend a path to the user, which include passive interfaces (visual, acoustic, and haptic) and active interfaces (electromechanical brakes, and motorised turning wheels). The device has a merely assistive role: the user is in charge of the final decision about the route to follow and can override the device suggestions.

The complex functionalities of the c-Walker are implemented relying in part on its embedded sensing and intelligence, in part on the ambient intelligence, and in part on the information shared with other c-Walkers present in the field.

DALi's emphasis on cognitive deficiencies makes the c-Walker unique in the landscape of the robotic walking assistants proposed in the recent literature (see Section 2.3). The c-Walker is original in its reliance on low cost mechatronic solutions that reduce its cost within acceptable boundsm, in its use of sophisticated models and algorithmic solutions to sense and perceive the surrounding environment and in its possibility to extend its sensing range by querying remote sensors and other devices of similar type.

A careful consideration of user's actual needs has been the distinctive mark of DALi's approach. Senior users have been engaged to extract and consolidate a set of requirements that has triggered the development of the different components. The design of the components has been fine–tuned using their feedback. Finally, the system as a whole has been validated with field studies and experimental trials. The outcome of these studies is out of the scope of this paper, which presents the technical highlights of the project. The interested reader can find the user evaluation results in the project public reports [11], while specific scientific papers on the matter are currently in preparation.

The paper is organised as follows. In Section 2 we offer an overview of the system functionalities. In Section 3 we describe the sensing abilities of the system. In Section 4 we discuss how to generate path plans and adapt them to the contingencies. In Section 5 we show the different solutions to guide the user along a path. Finally, in Section 6 we offer our conclusion.

2 System Overview

2.1 An example Use-Case

A general picture of DALi's functionalities is best offered using the following use-case.

An old lady loves to go shopping in a nearby mall. She suffers from the consequences of a recent fall, and the emotional stress makes her feel increasingly unsafe. Sometimes she gets confused and has difficulty to find her way across a large space.

When she reaches the mall, she uses the touch screen on her c-Walker to select the places she wants to visit. The c-Walker queries the server in search of the optimal route to visit the different shops. The route is planned accounting for the preferences stored in her profile and for the contingent situation on the field. For instance, the recent fall could suggest to avoid overcrowded spaces as well as areas where alert signals (e.g. wet floor) could cause her anxiety.

When she starts walking, her next turn is shown on the tablet display in connection with clear landmarks (e.g. "turn left when you meet the shoe shop"). As long as she walks within a safety band around the planned path, the c-Walker remains passive, no different from an ordinary walker or shopping trolley. If she deviates significantly from the planned route, the system pushes her back through a gentle vibration of the wireless haptic bracelets she wears. The display also shows a clear big arrow to adjust the route. If she ignores the direction signal the system generates a pop-up window on the display to check if the user changed her mind on her destination and re-plans a new route. When she meets a group of bystanders potentially obstructing her motion, the system predicts the evolution of their position. In case of evident risks of collision, the system plans a deviation that ensures better safety margins. The user is warned through the bracelets and through the display. For users with visual impairments, it is possible to use a headphone that modulates a 3D sound to show the correct path accounting for the pose of her head. Since the user seems unaware of the risk and keeps moving along her previous route, the guidance system operates forcing a gentle turn with a differentiated action of the two motorised brakes operating on the back wheels.

Later on, the system detects another anomaly: some items fallen from a shelf on the ground. A warning sign is in place to detour the visitors around while the attendants resolve the problem. In this case, not only does the system spot the sign and re-plan a safe route for the user, but it also broadcasts the information and share it with other walkers operating on the field to avoid problems for their users (e.g. through a timely re-plan of their route).

2.2 Functionalities and Modules

The DALi system relies on a number of functionalities implemented in different modules, as shown in Figure 1. The functionalities can be roughly classified in three dif-

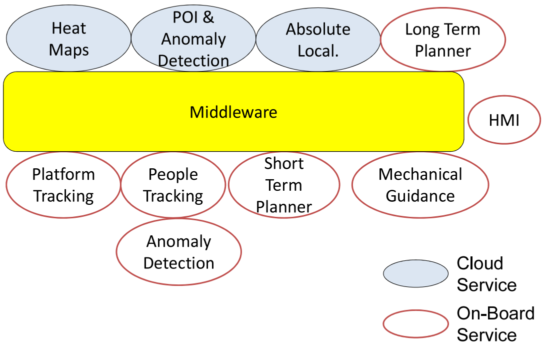


Fig. 1 The main modules of DALi.

ferent groups: 1. sensing, storing and processing information on the user and on her/his environment; 2. taking decisions and making plans to suggest and present them to the user; 3. guiding the user through the implementation of the plan.

The first group contains: representation of the environment, user localisation, tracking of people and detection of anomalies in the proximity of the c-Walker, detection of crowded areas, people of interest and anomalies in distant areas. Crowded areas are detected by computing heat maps on the presence of humans in the different areas viewed by surveillance cameras [37]. The same sensors are also used to detect people of interest (who can be recognised by the colour patterns of their uniforms [25]). This information is updated on a periodic basis (with a frequency in the order of a few times in a minute), published by software services deployed in the cloud and used for long term decisions. On the contrary, information on anomalies and presence of people in the proximity of the c-Walker is used for short term decisions and has to be updated several times in a second. Therefore, it is implemented by embedded software modules that are hosted on-board of the c-Walker. Localisation relies both on embedded modules (for relative positioning) and on occasional calls to a computation intensive cloud service (for absolute positioning). Some of the embedded modules have a bi-directional connection to the cloud since they can publish information of potential interest to other c-Walkers.

The second group of functionalities contains: long term planning and short term planning. Long term planning amounts to selecting a route to connect the starting point with the user first destination in a way that satisfies the requirements coded in her/his profile. In the current version, the user profile contains simple information expressing the user necessities, i.e. select the route that stays closer to toilettes or benches. The algorithm utilises the topological and metric information of the map and the real–time information on the presence of crowded areas and of anomalies in the place. The need for global information suggests to implement long term planning as a cloud service, although a degraded version (that relies solely on the knowledge of the map) can be executed locally in case of lack of connectivity. Short term planning is a "local" modification of the planned route, required to react to potential risks detected on the ground and to maximise the comfort of the user. This module utilises local anomaly detection and people tracking and is implemented by an embedded software module.

The third group contains: visual interface, haptic and acoustic interfaces and mechanical guidance. Visual interfaces are used to take the user input (by a touch screen), to show her/him information on the navigation and to notify relevant messages. Haptic and acoustic interfaces can be used to raise the user attention (e.g. on turns to take or on potential dangers) and to guide her/him along safe paths. Mechanical guidance differs in the way the user is guided. While with haptic and acoustic interfaces the system remains passive shifting the final decision on where to go entirely to the user, with mechanical guidance it plays a more active goal "forcing" changes in the direction of motion of the walker by its motorised brakes. In any case, the user can override the system decision forcing her/his way through the path. All interfaces and guidance functionalities require response time in the order of few tens of milliseconds and are operated by embedded modules.

2.3 Related work

Robotic walkers have gained an undisputed popularity in the research community on ambient assisted living [43,63,49]. Closely related to DALi is the Assistants for Safe Mobility (ASSAM) project [2], where a system consisting of a set of modular navigation assistants deployed on several devices is used to encourage physical exercise. ASSAM has a major focus on the seamless transition from indoors to outdoors, while DALi specifically considers large indoors environment. Also, the behaviour of people in the surroundings is not considered in ASSAM.

The iWalkActive project [38] is more directly focused on the development of an active walker that facilitates and supports physical exercise, both indoors and outdoors. E-NO-FALLS [16] is another project focused on the development of a smart walker. In this case the main emphasis is on preventing falls. Although of interest, these aspects are complementary with those of DALi, whose main focus is on fighting cognitive decline.

The different facets of a vast undertaking like DALi have strong and widespread relations with the existing literature in several scientific disciplines. In the following, we will discuss with some detail a selection of the functionalities, the ones in which lies most of the project innovation, in each of the three groups.

3 Sensing and processing information on user and environment

A robust and efficient system for sensing and processing information on user and environment is crucial for the c-Walker operation. Part of the sensing functionalities from visual data (i.e. heat maps and person of interest detection) could be delivered adapting and engineering solutions taken from the portfolio of technologies of an industrial partner[1]. For other problems, such as mapping, localisation and short range detection of the environment, a more substantial research effort was required.

3.1 Representing the Environment

The information produced by the sensing system is complex and heterogeneous and is used in a variety of ways by the different modules compounding the c-Walker. Hence, a flexible and easy-to-integrate representation of the map of the environment is a crucial element of the project.

3.1.1 Related work

Long term navigation requires the knowledge of the environment in which the c-Walker operates. Every motion planning methodology, like roadmap methods, cell decomposition or artificial potential fields [45], relies on a map of the environment for its success. The problem of developing robotic systems able to autonomously build a map of an unknown environment has been extensively addressed in recent years, and a number of possible solutions have been proposed [46,53]. The map construction process, however, is much easier when a digital planimetry of the building is available. The approach adopted in the DALi project exploits the information extracted from a CAD file in order to build a mixed metric/topological representation of the environment, consisting of a cell decomposition of the free space combined with a connectivity graph linking adjacent cells together [8].

The advantages of this choice are manifold. First, the reconstruction of topological and metric information on the place is entirely automated. Second, additional layers of information can be inserted to define the presence of points of interest and of dynamically changing anomalies. Finally, the production of motion plans can be made using efficient (Dijkstra–like) algorithms.

3.1.2 The DALi approach

A first requirement for the mapping system is the ability to represent the topology of the place (i.e. the number of free areas and their interconnection) along with metric information on the shape and the extension of the free areas and on the position of obstacles. Two additional information layers are required for planning and decision making. The first one is a semantic layer associating certain areas with points of interest such as shops, offices and bathroom. The second one contains dynamic information on the different areas such as temporary anomalies and obstructions or presence of visitors. A possible solution, flexible enough to contain all the relevant data, is one based on *a cell decomposition* of the environment. This technique builds a structured representation of the free space which turns out to be especially useful for path planning purposes.

The data structure resulting from the map building procedure can be used to populate a spatial and geographic database (e.g. PostGIS) which can be queried for efficient on-line computation of optimal routes. The use of a geographic database is particularly convenient to easily aggregate different layers of information on the map, as well as to retrieve and manipulate information on the environment through the use of an easy and standard API.

Cell decomposition The cell decomposition approach is a well known methodology, widely used in robotics for planning safe paths between two different robot configurations [45]. In the DALi project, it has been tailored to the map building problem to provide an input to the long term planner (see Section 4.3). The overall map building procedure can be summarised as follows. Starting from a CAD model of the environment, a preprocessing is performed in order to get rid of unrelated data and to retain only relevant information (walls, stairs and other static structures). Then, the actual cell decomposition is carried out. The main idea is to decompose the free areas of the environment into rectangular cells with variable size. This is done by recursively splitting a cell into four equal sub-cells until one of the following conditions is satisfied: i) the cell is completely free, ii) it is completely filled by an obstacle or iii) it has reached a predefined minimum size. The minimum cell size is a tuning knob which can be used to trade-off complexity and spatial resolution of the resulting map. This process naturally leads to organising all the cells into a *quadtree*. Labels are at-

---

[1] `http://www.visual-tools.com/en`

tached to selected groups of cells, in order to enrich the map with semantic information like the presence of stairs, exit doors or other points of interest. Finally, a graph is built, in order to capture the connectivity of the free space. The nodes of the graph are the free cells and the edges connect only nodes corresponding to adjacent cells, i.e. cells that share a part of their border. Figure 2 illustrates the results of this process on an actual building.

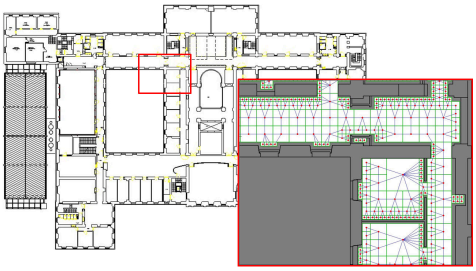


Fig. 2 Preprocessed CAD model of an actual building (about 3000 m^2, left) and resulting map of a part of it (about 100 m^2, right): free cells (green) and connectivity graph.

Geo-referenced database A geo-reference database is a database that is optimised for storing and retrieving spatial information. It is used as a persistent storage of the accumulated information of the environment and as a hub where information from different modules or sensors is collected. SpatiaLite[2] is the spatial database used in this project due to its simple architecture.

Using SpatiaLite it is possible to store multiple tables that represent: the map as a graph, point of interests, heat maps and people tracking result.

One of the advantages of geo-referenced databases is the ability to run queries that are related to the geometric objects, e.g. finding distance between two points, finding the area of a polygon, checking if two polygons overlap. The query differs from a normal SQL query in that it allows the use of simple geometric objects such as points, lines and polygons.

3.2 Localisation

Localising and tracking the c-Walker in the map is a key enabler for many of the functionalities developed in DALi, including planning (Section 4) and guidance (Section 5).

3.2.1 Related work

Pure dead-reckoning localisation approaches relying on incremental sensors are subject to measurement bias and drift [7]. Therefore, absolute position and orientation data are necessary not only to initialise the system, but also to adjust target location periodically. This problem is addressed in our project by using absolute measurements.

The use of the front camera as a bearing-only sensor is widespread in robotics for localisation and map construction due to its relatively low computational burden and good accuracy [50]. Well known is also the use of RFID tags [32]. Our unique combination of dead-reckoning sensors, RFID tags markers is described in detail in our previous papers [54, 55].

3.2.2 The DALi approach

The planning module requires that the absolute positioning accuracy of the c-Walker is at most 50 cm, while the guidance module requires position updates at least 10 times per second. The system is expected to operate in large indoor spaces where the GPS signal is unavailable or unreliable. Finally, the localisation module can benefit from the presence of a local infrastructure (e.g. tags in the environment, cloud services), but it is required to remain functional even when the connection with the external infrastructure is lost and no landmark is in sight.

This challenging group of requirements mandates a flexible solution based on the complementary use of different technologies. A relative based on the complementary use of different technologies. The idea is illustrated in Figure 3. A relative localisation module is activated on a periodic basis and estimates the relative motion of the c-Walker from a known position. This is done by integrating the information collected from egomotion sensors: incremental encoders and gyroscopes. The use of such sensors enables high-rate position estimates. However, the initial position is not observable. Also, the uncertainty on the reconstructed position inevitably grows in time because it comes from the integration of noisy measurements. For these reasons, relative localisation has to operate hand-in-hand with additional systems for absolute localisation. In DALi we have two different technologies for absolute localisation: one is based on the detection of RFID tags and visual markers spread out in the place. The second one is based on the use of 3D models of the environment. It is offered as a cloud service and does not require any kind of instrumentation on the environment. In both cases, absolute localisation is event–triggered. The module based on the

---

[2] `http://www.gaia-gis.it/gaia-sins/`

detection of tags and markers is activated when the system comes across one of these landmarks. Localisation based on 3D modules has a relevant footprint on communication bandwidth, so it is activated only when the uncertainty on the position is above a preset thresholds. The fusion between absolute and relative localisation takes place in a position estimation block that updates the estimated position with the frequency required by the applications.

Relative Localisation Relative localisation uses a multi-sensor data fusion technique based on data collected from two incremental encoders installed in both rear wheels and a triaxial gyroscope located in the central top part of the walker. The encoders are used to estimate the relative position and orientation of the c-Walker by using a simple kinematic model. Relative orientation estimation can be improved also by integrating the angular velocity values collected from a high-performance gyroscope. The fusion between these sensor data takes place within an Extended Kalman Filter (EKF) [71].

Absolute Localisation based on RFID tags and markers

Short-range passive RFIDs are an attractive option to adjust position, since they are cheap and can be easily stuck on the floor. Also, position measurement uncertainty can be kept smaller than the RFID reading range (i.e., in the order of a few tens of cm). One potential limitation of RFIDs is that they do not convey any information on the bearing of the platform. To this purpose it is useful to use visual markers (e.g. stickers with arrows, or anything that could indicate a specific orientation) that can be attached on the floor as well and can be detected by a front camera. QR codes provide this feature since it is possible to distinguish the four corners and therefore determine the orientation of the marker. The resulting configuration is shown in Figure 4. The computational workload for visual marker detection on an embedded computer is not irrelevant. However, the detection algorithm is activated at a low rate (i.e. 10 Hz).

Absolute Localisation based on 3D models

When the c-Walker travels across areas that are not covered by RFID and visual markers a different technology is available for absolute localisation. In this work, we used a slightly modified version of the incremental Structure from Motion (SfM) system described in [36]. At run-time, the position of the c-Walker is discovered solving the Perspective Pose Estimation from an image taken from the Kinect device mounted in the front of the device. From the operational point of view, this is done by invoking a remote service in the cloud.

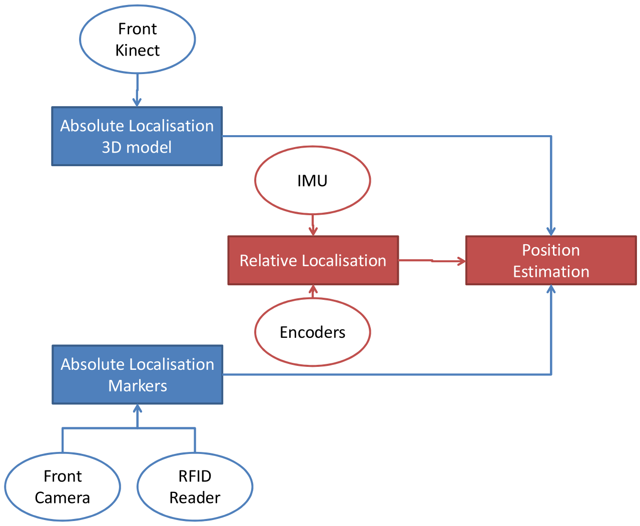


Fig. 3 Block diagram of the proposed multi-sensor data fusion technique for rollator position tracking.

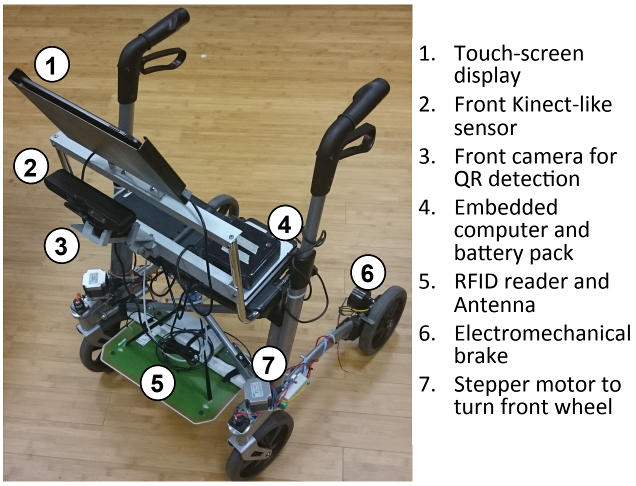


Fig. 4 The c-Walker.

The Structure from Motion (SfM) method computes camera positions and orientations from point correspondences between images. 3D points are obtained from these correspondences implicitly. Given a set of unordered input images, SIFT features [51] are extracted from each image. We use an approach based on a vocabulary tree for coarse identification of similar images [56]. Hence, we can greatly reduce the computational effort of pairwise image matching, by only matching the most relevant images as reported by the vocabulary scoring. Figure 5(a) shows the 3D reconstruction we obtain for a factory hall.

The 3D model is densified for the purpose of better visualisation.

Localising the platform by solving the Perspective Pose Estimation Problem with a platform sensor image makes it necessary to find a correspondence between the

(a) (b)

Fig. 5 3D point cloud data derived from our “Structure from Motion”-Workflow.

2D points from the sensor image and the 3D SfM point cloud. Each 3D point is associated with a set of feature descriptors from the SfM input images. From these correspondences the pose of one perspective image can be computed using the well-known RANSAC and Three Point algorithms [33]. Figure 5(b) illustrates the estimated camera positions along a path of the platform with respect to the factory hall environment. According to the experiments conducted, our approach seems viable for reliable anchor point localisation as long as it is not forced to work at any place and in all environments.

Fusing different information on localisation The data on relative localisation from dead-reckoning sensors are fused with the ones on absolute localisation using an Extended Kalman Filter (position estimation block in Figure 3). Every time new information on the relative motion is produced by the relative localisation block, the EKF updates its estimated absolute position. Likewise, when a new absolute localisation measurement is available, it is used to correct the current estimate accounting for the covariance of the noise. The covariance matrix can be obtained using the inverse Hessian of the reprojection error function [67]. An example of use is shown in Figure 6.

This experiment was made by replicating a small market place. We used the position measured by a laser ranger as a ground truth (dashed line). In this specific example we show the combined operation of relative and absolute localisation based on visual markers and RFID tags (solid line). Position estimates are updated every 4 ms on average. The observed deviation is below 50 cm with 95% probability.

3.3 Short range perception of the environment

A correct perception of the environment in the surroundings of the c-Walker is essential to make decisions. The information of interest can be classified as follows:

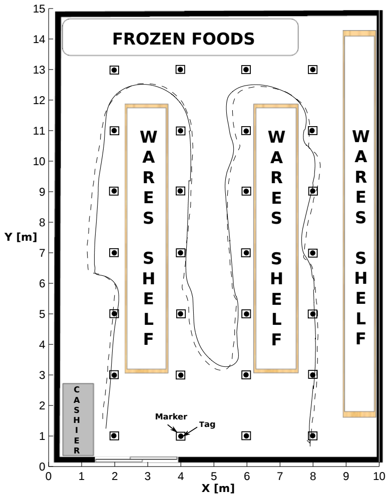
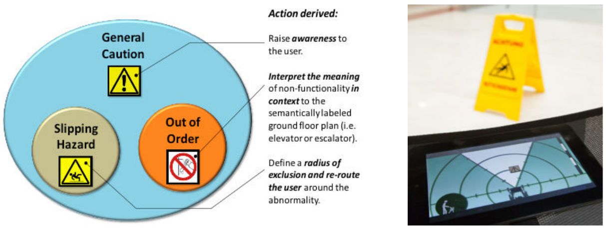


Fig. 6 An experiment to test the performance of the localisation algorithm.

- Vision-based 3D perception of the environment (obstacle detection and map construction);
- Observation of humans (tracking the 3D position and velocity of humans in the immediate surroundings of the platform);
- Smart OCR for real world text recognition (reading signs to identify anomalies);
- Observation and monitoring the user behavior (understanding the emotional state of the c-Walker user and/or her/his intentions).

3.3.1 Related work on anomaly detection

Textual information extraction from images is known to be a challenging problem. The focus of many applications for Optical Character Recognition (OCR) has been primarily on reading digitized printed documents. Thus, only little work has been undertaken traditionally on reading text in natural images, denoted as Real-World OCR. During the last ten years this topic has received an increased attention [72,69]. The approach advocated in the DALi project to Real-World OCR relies on a clear separation of concerns between text detection or localisation, on one hand, and OCR on the other. For the first part, state-of-the-art approaches reach disappointing values of F-measures in the range 0.65 to 0.7 on a commonly used database [64]. Also, in the conclusion from the latest ICDAR Robust Reading Competition, the topic is expected to remain a hot topic of research for the coming years [62] . Our main contribution is a runtime efficient processing chain - 3-5 fps on barebone-PC hardware - that covers robust localisation by exploiting 3D plane constraints as well as automatic rectification of non-camera parallel text.



(a) (b) (c)

Fig. 7 Anomaly detection in DALi

3.3.2 The DALi approach to anomaly detection

Detecting anomalies in the surrounding of the c-Walker is critical for safe navigation. Some examples of anomalies are a wet floor (recognisable by a dedicated floor sign), a physical object at a previously uncovered location, e.g. goods dropped down from a shelf, or an unusual aggregation of people. A subset of these anomalies can be detected by means of real world OCR. Warning signs show a wide variety in pictogram representations with the same meaning, i.e. indicating the same desired action. Hence, it is difficult to train a robust pictorial classifier to cover all variations. Reading and analysing the text makes the problem more feasible, albeit it requires semantic interpretation of text blocks.

In contrast to document OCR, real world images bear a lot of challenges including image quality, lighting conditions, perspective distortion, highly cluttered scene content as well as occlusions. Our main contribution is a run-time efficient processing chain - 3-5 fps on barebone-PC hardware - that covers robust localisation via exploiting 3D plane constraints as well as automatic rectification of non-camera parallel text.

In order to yield viable information in the context of DALi, appropriate action must be derived from the interpreted text. This ranges from an input to the navigation system for avoiding a potentially dangerous area to raising awareness to the elderly person that there is something that requires his/her special attention. The classes of anomalies we currently focus on in our work are shown in Figure 7(a). Figure 7(b) illustrates the action we are able to derive from each class. Finally Figure 7(c) shows what our running system looks like as we provide a radar paradigm of visual attention in real-time.

3.3.3 Related work on detection and tracking of moving people

Our work is related to multi-target tracking, camera egomotion estimation/localization and 3D scene structure estimation/map construction. A complete review of the related work constitutes a huge task even for any of the individual subproblems and is beyond the scope of this paper.

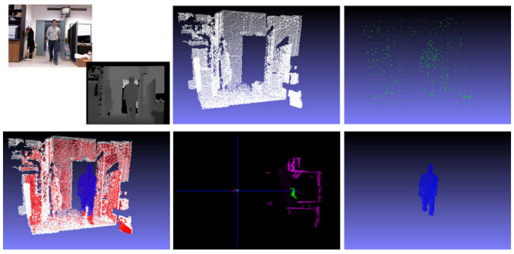


Fig. 8 Detection and tracking of moving people. Top left: an RGBD frame. Top middle: The current map of the environment. Top right: A sparse set of 3D points used for egomotion estimation. Bottom left: registration of the whole point cloud with the current map of the environment. Bottom middle: a top view showing the camera position, the local environment map and the independently moving objects. Bottom right: The point cloud of the foreground moving object.

A number of methods have been proposed for the problem of tracking objects in RGB images [73] and several algorithms are tailored specific to humans [66, 21, 20]. To address the challenges of tracking objects from a moving platform several approaches have been recently proposed that use multiple detectors and various sensor configurations [19, 9]. In our work we are limited to a single RGB-D sensor and thus we need to conform with the limitations of the hardware.

For the problem of simultaneous localization and mapping (SLAM), a recent review is provided in [24]. The consideration of SLAM in dynamic environments leads to the so-called SLAMMOT problem [68] which involves SLAM together with detection and tracking of dynamic objects. Their approach combines SLAM and moving object tracking that are performed based on a 2D laser scanner.

The introduction of RGB-D sensors has provided a cheap way of acquiring relatively accurate 3D structure information. This has enabled the development of SLAM and object tracking methods with impressive results [39, 17]. The goal of our work is to address the combined SLAMMOT problem by investigating the coupling of the individual subproblems under the limited computational resources of the smart walker platform.

3.3.4 The DALi approach for detecting and tracking moving people

Our approach to detect and track humans in the surrounding of the c-Walker is based on segmenting and tracking objects that move independently in the field of view of a moving RGB-D camera (see Figure 8). The camera is assumed to move with 6 degrees of freedom

Fig. 9 Snapshot of the execution of people tracking algorithm. Bottom left: the depth and RGB maps corresponding to independently moving people. Right: A top view of the local environment map showing the motion hypotheses for two people. Top left: the bounding boxes of the identified moving persons.

(DOFs), while moving objects in the environment are assumed to move on a planar floor. This last assumption is the only a-priori information about the environment. Motion is estimated with respect to a coordinate system related to the static environment. In order to segment the static background from the moving foreground, we first select a small number of points of interest whose 3D positions are estimated directly from the sensory information. The camera motion is computed by fitting those points to a progressively built model of the environment. A 3D point may not match the current version of the map either because it is a noise-contaminated observation, or because it belongs to a moving object, or because it belongs to a structure attached to the static environment that is observed for the first time. A classification mechanism is used to perform this disambiguation. Based on its output, noise is filtered, points on independently moving objects are grouped to form moving object hypotheses and static points are integrated to the evolving map of the environment. More details regarding the adopted approach are reported in [57]. Sample results obtained from the execution of the algorithm are shown in Figure 9.

Several experimental results demonstrate that the proposed method is able to track moving objects correctly. Interestingly, the performance of egomotion estimation and map construction practically remains unaffected by the presence of independently moving objects. From a computational point of view, the method works at a frame rate of 50 fps on a laptop with an "Intel i7" CPU without the use of GPU acceleration, and can perform at near real-time speeds on ARM-based embedded platforms.

4 Planning

The c-Walker motion planner divides its task into long term planning and short term planning. The long term plan can be considered the a-priori best route to achieve the user's specified goals (e.g. places to visit in order, within time constraints), considering the known physical layout of the environment and the preferences of the user. The long term planner also accounts for the level of crowdedness recorded by the environmental cameras and for anomalies possibly detected by other walkers. On the contrary, it cannot take into account the instantaneous positions of other pedestrians and/or the presence of occasional anomalies. Hence, the short term planner is required to react to possible contingencies on the field and to produce a new plan that remains as close as possible to the original plan, while avoiding risks and stress for the user. If the short term planner finds that the long term plan is no longer applicable, it requests the long term planner to find an alternative route. The long term planner may also be triggered if the user's objectives change.

4.1 Related work

The work proposed in this paper is related to motion planning in crowded environments [47,48,44]. We tackle the problem using a two-levels approach, in which a global and a local planning problem are solved. This strategy is common in the literature [52,31], with different proposals on how to independently solve the two problems.

Our global planner is based on the well-known Dijkstra shortest-path algorithm [14]. However, two problems have not been addressed to a satisfactory level as yet. The first one is how to insert the user's requirements in the planning problem. The second one is how to account for the presence of time-dependent anomalies (e.g., crowds) along the way. Adaptations of the Djikstra algorithm to cope with time varying graph topologies is know to be a hard problem. Different groups have proposed algorithmic solution to this tough problem [15,12,13,23], but with no clear winner at the time of this writing. The definition of user preferences as constraints or cost functions and their appropriate encoding into the Shortest Path problem is largely an unexplored area. DALi offers its solutions to both problems, and the results in terms of efficacy and efficiency are promising.

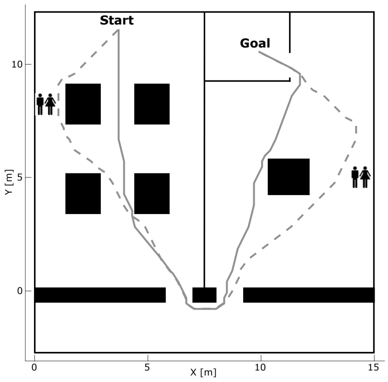


Fig. 10 The long term planner. The continuous line is the shortest path between the two endpoints, while the dashed line represents the optimal path that keeps the restrooms within reach.

4.2 Long term Planner

As discussed above, DALi relies on a geo-referenced data base (SpatiaLite) that represents the environment in its different facets (see Section 3.1.2). The database is constructed through a cell decomposition of the environment and has several layers: the topological/metric layer, the semantic layer (to represent the points of interest) the heat maps layer (to represent the level of crowdedness of different cells) and the anomaly layer.

All this knowledge is used by the long term planner. In its simplest instance, the planner solves a Dijkstra optimisation problem on the graph made by the cells and by their interconnection. In this case, only the topological/metric and the semantic layers are used and the weight of each edge is given by the time to cover (at a given reference speed) the distance between the centres of the cells at the ends of the edge itself. The optimisation problem is solved by a simple query to SpatiaLite.

The presence of such anomalies as obstructions or wet floor is coded by temporarily removing some edges from the graph. Likewise, heat-maps revealing the presence of humans can be accounted for by changing the weight of the edges. Manipulations of edge weights is also our means to enforce user preferences. For instance, if a user needs frequent access to the bathroom, the edges toward the bathroom cells will have a reduced weight. This is shown in Figure 10. The continuous line is the shortest path between the starting and the destination points and it is suggested to a user without particular preferences. The dashed line shows instead how the path is modified for an older adult willing to always have a restroom within reach.

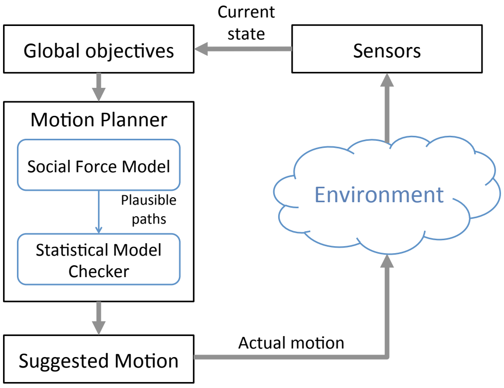


Fig. 11 The functional diagram of the short term planner.

4.3 Short term Planner

The short term planner [10] takes as input the long term route as well as the position and the velocity of nearby objects (provided by sensors – see Sections 3.2 and 3.3). It makes prediction on the evolution of the situation in a short time horizon (from 3 to 10 seconds), and re-shapes the path to avoid potential problems. The architecture of the short term planner is illustrated in Figure 11.

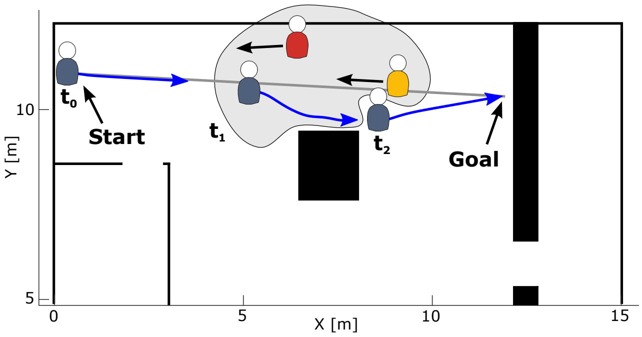


Fig. 12 An example run of the short term planner. The c-Walker (blue-coloured icon) is moving from "start" (time t_0) to "goal" following the gray-coloured long term plan. At time t_1 two agents (red and yellow-coloured icons) appear within the sensor range and the short term planner suggests the c-Walker an optimal trajectory (blue-coloured path) that avoids the agents with high probability. The planner then gently takes the user back to long term plan (time t_2).

The mathematical model used to foresee the future positions of humans over a given area is derived from the *social force model* (SFM) [34]. The SFM is based on a stochastic differential equation that includes *attractive* terms to represent pedestrians along their desired trajectories, and *repulsive* terms to model their reluctance to stay too close to each other. The DALi's variation of the SFM includes stochastic terms to account for imponderable aspects of human intent that could determine unexpected changes in the direction of motion.

The optimal trajectory sought by the short term planner has to: 1. be smooth; 2. comply with the social rules encoded in the SFM; 3. remain close to the long term plan and 4. avoid collisions and unpleasant situations. The algorithm hypothesises different candidate trajectories. Each of them splits away from the current path with a different orientation and evolves according to the SFM. The current path exerts an attractive force, while human obstacles are repulsive. As a result, the computed path joins back into the original path as soon as the obstacles are overcome.

Each candidate trajectory is tested to verify whether the safety properties are fulfilled. Safety is modelled by a temporal logic formula and can be different for different user profiles. The logic can express sequences of events in time and space and is a convenient way to describe the long term plan and the desired personal space of the user. For each candidate direction a number of simulations are carried out to model the stochastic fluctuations in human motion (both for the user and for the other actors in the scene). Each simulation (solution of the equations) produces a unique realisation on which the property is tested. This can be done efficiently using *statistical model checking* (SMC) [74], which accumulates statistics on the violation of the property. Eventually, the system planner selects the trajectory that is the closest to the original plan and that assures safety with a target probability. An example is shown in Figure 12 where the c-Walker is following the long term plan (gray-coloured), while two agents appear along the route. The short term planner suggests the optimal trajectory (blue-coloured) that is able to avoid the two unforeseen obstacles with a high probability.

5 Guidance

Once a motion plan is decided, the c-Walker has to assist the user in its execution. DALi's philosophy is to play a minimally invasive role. So, as long as the user moves in a virtual corridor including the planned path, the only role of the c-Walker is to provide him/her directions on the next turn to take. This is done mainly through visual indications on the screen, or via acoustic or haptic messages.

The role of the c-Walker changes when the user approaches the boundary of the virtual corridor. In this case, the guidance system kicks in to drive the user back on the safe course. The c-Walker can perform this task passively generating haptic and acoustic signals, or it can play a more active role by braking the back wheels and/or by turning the front caster wheels.

5.1 Mechanical Guidance

5.1.1 Related work

When developing robotic aids for supporting assisted people with limited motor and/or cognitive abilities the primary concern is safety. The paradigm of *passive robotics* for designing intrinsically safe devices was first introduced by Goswami et al. [30]. This approach has been also adopted in [29] for realising the Cobot walking assistant. A Cobot is a robotic device composed of a cane with a caster wheel equipped with a servo motor for actuating the steering angle. The users supplies the motive power. This is the rationale of the mechanical guidance for the DALi walker. In particular we envisage two different passive guidance approaches: steering by brakes [35, 22] and a more straightforward steering wheels control.

5.1.2 The DALi approach

For mechanical guidance the c-Walker can utilise two different actuators: the electrical brakes mounted on the back wheels or a stepper motor mounted on the front wheels. These actuators are used to implement a path following control algorithm. Paths are generated by the long or short term planners described in the previous section, thus having a known geometry (see Figure 10).

Given the dynamic model of the walker, if we had a direct control on the torques applied to the rear wheels, the path following task could be achieved by simply applying desired torque values coming from a feedback control law. Since we do not have any motor to directly generate such torques, the path following problem is solved by roughly estimate the user imposed thrust and then applying a dissipative braking action. Additionally, to maximise the user comfort, the braking action is computed by solving an optimal problem minimising the dissipative action [22].

Another important aspect is our idea of a virtual corridor, such that when the user is in the middle the control is loose, while it becomes increasingly authoritative when he/she approaches the boundaries. This way, the user perceives that he/she still remains in charge of navigation decisions.

A similar control framework has been set up using the steering wheels. In this case, we use a technological characteristic of the stepper motor controlling the rotation of the wheel, which allows us to control the torque applied to hold the wheel in a desired position by setting a different value for the armature current. When the current is maximum, the wheel is stiffly held in the control position. This is useful when the user is close to the boundary of the corridor and we seek an authoritative control action. Notice that, from a control theoretic viewpoint, the steering wheel control is more straightforward and is carried out directly on the kinematic model of the moving platform.

5.2 Haptic Guidance

5.2.1 Related work

A large body of literature exists on the theme of haptic guidance. In [28] the authors investigate the design of a stiff rein which enhances human trust and confidence in cooperative human-robot navigation. Haptic feedback for human navigation has been used in [1], where the authors used a grounded haptic manipulator to apply force and position signals to the user's hand in order to assist the operator in reaching a desired position in large remote environments. Although haptic devices which provide kinesthetic feedback generate strong forces and effectively guide human motion, they are bulky, grounded and not portable in large environments. To address this issue, in [18] a haptic belt is used for waypoint navigation. However, the system relies not only on vibrotactile stimuli, but also on GPS information which is not available indoors. In [61,60], the authors proposed a vibrotactile feedback policy for the coordination of human-robot teams and cooperative human-robot navigation.

Differently from such previous works, we present the design of vibrotactile devices to improve the navigation of older adults in large and crowded environments. The vibrotactile bracelets were designed to be efficient in terms of causal chain of stimulus-perception-decision and *aftereffect* problem. They can be easily used in conjunction with the assistive walker in order to softly "suggest" to the older adult the optimal path and to guide the user with a minimal impact on his/her freedom of motion.

5.2.2 The DALi approach

Guidance Algorithm The haptic guidance is based on the generation of haptic signals that have to be followed by the user. Additionally, control signals are in this case quantised as we are only to suggest left turn and right turn. One could potentially adapt the algorithm described for mechanical guidance to quantised actuators, e.g. by using bang-bang or sliding mode strategies. However, in our practical experience with users, we found out that a person cannot be treated as a motor: signals require a deliberation phase before being accepted, and we cannot generate signals in the opposite direction within a short interval. This type of considerations lead us to opt for a heuristic algorithm, using thresholds on the difference between θ and θ_d and on l to decide the signals to generate. As usual, when the user in the middle of the virtual corridor, no signal is emitted. When he approaches the boundary and the difference between θ and θ_r exceeds the threshold, a vibration is generated. Such thresholds are adapted on the single users "learning" from her/his reaction to past signals.

Maximising the efficacy of haptic signals Studies have demonstrated that vibration is best on hairy skin due to skin thickness and nerve depth, and that vibrotactile stimuli are best detected in bony areas [26]. In particular, wrists and spine are generally preferred for detecting vibrations, with arms next in line [41]. Movement can decrease detection rate and increases response time of particular body areas [41]. In the design of the vibrotactile device, we have to keep in mind the reduced perception of vibrotactile feedback in older adults. Given that a decline in the main sensory modalities (i.e. vision, hearing, taste, and smell) is well reported to occur with advancing age, one would expect similar change to occur with touch sensation and perception. Studies on the effects of aging in the sense of touch have been reported in [27] where experimental results revealed that vibration threshold is the most rapidly affected by age and is maximal after the age of 65 years.

Following these considerations, we designed a wearable haptic bracelet in which two cylindrical vibro-motors can be independently controlled via an external PC using the Bluetooth communication protocol, and generate vibratory signals to warn the human (see Figure 13). The user wears one vibrotactile bracelet on each arm in order to maximize the stimuli separation while keeping the discrimination process as intuitive as possible. Vibration of the left wristband signals the participant to turn left and vice versa. The vibrotactile device is fitted to the arm, just below the elbow. This configuration is optimal to distinguish the haptic stimuli from the vibrations of the c-Walker induced when the cart moves along bumpy areas. On each bracelet the distance between the two motors is about 80 mm; the minimal distance between two stimuli to be differentiated is about 35 mm on the forearms. In two point discrimination perception, there is no evidence for differences among the left and right sides of the body [26].

From a technical point of view, two Precision Microdrives 303-100 Pico Vibe 3.2mm vibration motors were placed into two fabric pockets on the external surface of the bracelet (the width of the wristband is about 60 mm), with shafts aligned with the elbow bone. The

Fig. 13 Final design of the vibrotactile bracelets.

motors have a vibration frequency range of 100-280 Hz (the maximal sensitivity is achieved around 200-300 Hz while the human perceptibility range is between 20 Hz and 400 Hz), a typical normalized amplitude of 0.6 g, a lag time of 21 ms, rise time of 32 ms and a stop time of 35 ms. The motors are controlled by applying a voltage which determines both frequency and amplitude, to provide a multi-dimensional stimulation. In order to maximize the vibrotactile perception and minimize the *aftereffect* problem, we use a periodic vibrational pattern with period 0.4 s instead of a continuous signal, with a frequency of 280 Hz.

5.3 Acoustic guidance

5.3.1 Related work

Several methods have been reported in the literature to re-create sounds that appear coming from a precise location (3D sound). The most frequently used method to render 3D sound is based on the Head Related Transfer Function (HRTF) which represents the response of the human ears and body for a given direction of the incoming sound and individual anthropometry [4]. HRTFs change and are measured for each individual. This method requires large memory to store the filter coefficients and the resulting filtering process is computationally demanding. It also lacks the sensation of distance and movement, which have to be separately implemented by means of proper algorithms (e.g. Doppler effect, reverberation).

Another approach is based on sound propagation modeling. One of the major cues for sound localisation is the Interaural Time Difference (ITD), that is the transient time difference between the two ears due to the differential distance between ears and the sound source [6]. Another cue for the localization of sounds lying on the horizontal plane is the Interaural Level Difference (ILD) or Intensity Difference. The effect of the path is easily inserted in the model considering that the attenuation in amplitude is equivalent to the inverse of the path length. The shadowing effect introduced by the head is frequency dependent. However, this effect can be modeled as a single pole/single zero filter.

5.3.2 The DALi approach

Guidance algorithm Traditional acoustic interfaces provide the user with explicit directions in the form of instructions, such as “turn left”, or “go straight”. The limitation of this approach, especially for those who are visually impaired, is that it does not provide an enhanced representation of the environment, and the information cannot easily be used to make independent decisions. We have therefore followed a different approach, in which the auditory stimulus is synthesized to paint a *soundscape* of the actual environment. This is achieved by processing the acoustic signals with 3D sound techniques, which make sounds appear as if they are generated from precise points in space. The stimulus, instead, is provided to the user through a pair of headphones. This way, selected sounds can be used to direct the attention towards the path to be followed, and potentially to pin-point the position of obstacles. This allows us to guide the user by simply generating a sound from a virtual source located on a point moving along the desired path with the desired speed.

Generation of 3D sounds The ability of humans to judge the position in space where a sound originates from is called *sound localisation*. Sound localisation relies on the body shape and the displacement of human ears, resulting in differences in amplitude and phase (binaural cues), and on some spectral cues of the sounds generated by the pinna folds (monaural cues), to determine the direction of arrival of sound waves. Several techniques can be used to create 3D sounds. One is to measure the impulse response between points in space and the ear, and determine an overall transfer function, known as HRTF. Our approach is based instead on modeling the sound propagation phenomena and compute the response based on anthropometric information. This gives us more flexibility, since we are not tied to discretized measurement points, and the possibility to tune the parameters of the model to fit the individual user. This solution is also convenient for an embedded implementation, since it does not require access to large databases of transfer functions (one for each point), and is generally considerably faster to compute [58,59]. This last aspect is important, given that the sound position must be updated in real-time as the user moves in the environment.

The acoustic feedback module takes as input the current position of the user, the position of the obstacles to be avoided, and the desired path, which are provided by the planner. The module can synthesize a sound that moves a certain distance along the path to be followed, with a configurable speed. Alternatively, the sound may be activated only when the user is requested to change direction, contextually with the haptic feedback described in the previous section.

Reverberation has the effect of de-correlating the signal at the two ears, increasing the sensation of immersion and spaciousness. We introduce reverberation using the Image Source Method, which consists of replicating the sound as virtual images resulting from the reflections of the sound waves on the walls [5]. More complex reverberation algorithms [40] are not necessary, since it has been shown that it is only the first reflections that are useful to help with the sound localisation process [3]. Likewise, sound localisation can be considerably improved by tracking the position of the head, and adjusting the sound sources accordingly. This is accomplished by mounting an inertial platform on top of the headphone, which provides real-time information on the head orientation. The improvement in sound localisation is due to the ability of the user to integrate information in time as the head moves. This is particularly useful, especially to reduce the front to back confusion that may arise with 3D sound engines.

5.4 A guidance example

As an example of the guidance system choices, we present here guidance algorithm tests carried out with several aged users, replicating a shopping centre scenario. In this section we offer some experimental results. In Figure 14, we show the same planned path followed by different users. The figure shows three different guidance mechanisms in use. The results should not be compared against each other because each mode is used by a different user. As a general observation, the different guidance mechanisms work acceptably for a large part of the users (we tested 20 aged users and the same number of care-givers). Different users express a different preference on the guidance mechanisms. Some of them feel comfortable with the direct action of the mechanical guidance, other feel uneasy and prefer passive signalling. Acoustic guidance meets the appreciation of users only if they have visual impairments. This diversity in the evaluation underscores the importance of having a portfolio of solutions to pick from and fine-tune to meet the specific requirements of every single users.

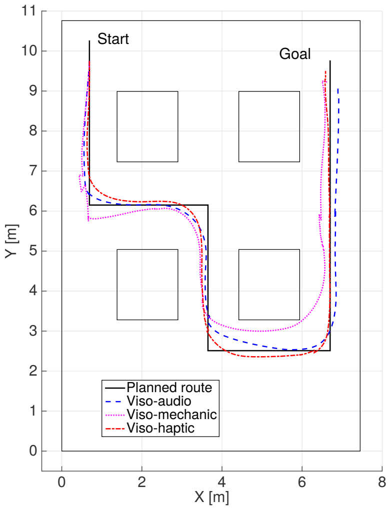


Fig. 14 Experimental results of several guidance mechanism with different users.

6 Conclusions

In this paper we have offered an overview on the DALi project and briefly described some of its most important technological highlights.

A particular focus was on sensing, planning and guidance techniques but we also tried to put them in the context of the overall design of the system. Additional information on the system are available on the project website (`http://www.ict-dali.eu`), on the personal webpages of the authors and on the cited publications.

6.1 Lesson Learned

There are a few lessons we have learned from DALi's activities, which will underpin new research activities and help us set up solid foundations for a possible commercial developments.

From the technical point of view, the most important finding of the project is that a careful applications of modern computer technologies can help us reduce the production costs offering a service of sensible quality within a moderate target price. This is an alternative approach to building devices with sophisticated and costly mechanical engineering. For instance, the user of appropriate processing algorithms can help us extract a significant amount of information using low cost sensors. This aspect can further be emphasised by networking the devices in order for then to share information and by giving them access to the information retrieved by environment sensors. Of particular relevance is the ability to offload part of the computation into a cloud of external services when needed (e.g., 3D localisation).

Modern approaches to planning, which leverage the most recent developments of techniques at the confluence between operation research and formal methods, allow us to capture the user requirements and effectively react to several types of contingencies. An open issue here is to "learn" from the user past behaviours in order to produce a profile that actually matches her/his needs.

A final technical consideration is on guidance. A careful combination of passive (haptic and headphone) and active methods enables the implementation of soft strategies for human guidance, leaving the human in charge of the last decision while preserving her/his safety. The use of electromechanical brakes is apparently a good solution when the floor is rough and offers a good grip. The use of motorised caster wheels seems to have a more general applicability.

From the user perspective, the "stigma" of using unusual and strange looking devices (such as haptic bracelets) seems to be the most pressing concern. Much we will have to do to minimise the visual impact of the design. Users are particularly reluctant to use headphones, whose application is suggested for people with visual impairments and conditioned to an adequate propagation of external sounds.

To summarise, this paper has proposed a soft approach to the assistance of older adults, where assistive devices are used to offer physical and emotional support to older adults augmenting their cognitive abilities. The interfaces remain by and large hidden underneath the design, thus reducing the perceived distance from a standard walker. Users appear to be open to this paradigm. Newer developments are required to narrow down the distance between the device that instantiates this paradigm and the users' needs.

References

1. Arias, A., Hanebeck, U.: Wide-area haptic guidance: Taking the user by the hand. In: Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst., pp. 5824–5829 (2010)
2. Assistants for safe mobility (assam). `http://assam.nmshost.de/`
3. Begault, D., Wenzel, E., Lee, A., Anderson, M.: Direct comparison of the impact of head tracking, reverberation, and individualized HRTFs on the spatial perception of a virtual speech source. J. Audio Eng. Soc **49**(10), 904–916 (2001)
4. Blauert, J.: Spatial Hearing-Revised Edition: The Psychophysics of Human Sound Localization. MIT press (1996)
5. Borish, J.: Extension of the image model to arbitrary polyhedra. In: Journal of Acoustic Society of America, vol. 75 (1984)
6. Brown, C., Duda, R.: A structural model for binaural sound synthesis. Speech and Audio Processing, IEEE Transactions on **6**(5), 476–488 (September 1998)
7. Byoung-Suk Choi, Joon-Woo Lee, Ju-Jang Lee: Localization and map-building of mobile robot based on RFID sensor fusion system. In: Proc. International Conference on Industrial Informatics (INDIN), pp. 412–417. Daejeon, Korea (2008)
8. Cai, C., Ferrari, S.: Information-driven sensor path planning by approximate cell decomposition. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics **39**(3), 672–689 (2009)
9. Choi, W., Pantofaru, C., Savarese, S.: A general framework for tracking multiple people from a moving camera. IEEE Trans. Pattern Anal. Mach. Intell. **35**(7), 1577–1591 (2013)
10. Colombo, A., Legay, A., Palopoli, L., Sedwards, S., Fontanelli, D.: Motion Planning in Crowds Using Statistical Model Checking to Enhance the Social Force Model. In: 52nd IEEE Conference on Decision and Control, pp. – . IEEE, Florence, Italy (2013)
11. Coventry, L., Targher, S., et al.: Dali deliverable 6.3: System evaluation. Tech. rep., European Commission (2015)
12. Delling, D., Wagner, D.: Time-dependent route planning. In: R. Ahuja, R. Möhring, C. Zaroliagis (eds.) Robust and Online Large-Scale Optimization, *Lecture Notes in Computer Science*, vol. 5868, pp. 207–230. Springer Berlin Heidelberg (2009)
13. Demiryurek, U., Banaei-Kashani, F., Shahabi, C.: A case for time-dependent shortest path computation in spatial networks. In: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '10, pp. 474–477. ACM, New York, NY, USA (2010)
14. Dijkstra, E.: A note on two problems in connexion with graphs. Numerische Mathematik **1**(1), 269–271 (1959)
15. Ding, B., Yu, J.X., Qin, L.: Finding time-dependent shortest paths over large graphs. In: Proceedings of the 11th International Conference on Extending Database Technology: Advances in Database Technology, EDBT '08, pp. 205–216. ACM, New York, NY, USA (2008)
16. The e-no-falls project. `http://www.e-nofalls.eu`
17. Endres, F., Hess, J., Sturm, J., Cremers, D., Burgard, W.: 3d mapping with an RGB-D camera. IEEE Transactions on Robotics (T-RO) (2013)
18. Erp, J.B.F.V., Veen, H.A.H.C.V., C. Jansen, C., Dobbins, T.: Waypoint navigation with a vibrotactile waist belt. ACM Trans. Appl. Percept. **2**(2), 106–117 (2005)
19. Ess, A., Leibe, B., Schindler, K., Gool, L.J.V.: Robust multiperson tracking from a mobile platform. IEEE Trans. Pattern Anal. Mach. Intell. **31**(10), 1831–1846 (2009)
20. Felzenszwalb, P.F., Girshick, R.B., McAllester, D.A., Ramanan, D.: Object detection with discriminatively trained part-based models. IEEE Trans. Pattern Anal. Mach. Intell. **32**(9), 1627–1645 (2010)
21. Ferrari, V., Marin-Jimenez, M., Zisserman, A.: Progressive search space reduction for human pose estimation. In: Proceedings of the IEEE Computer Vision and Pattern Recognition. Alaska (2008)

22. Fontanelli, D., Giannitrapani, A., Palopoli, L., Prattichizzo, D.: Unicycle Steering by Brakes: a Passive Guidance Support for an Assistive Cart. In: Proc. IEEE Int. Conf. on Decision and Control, pp. 2275–2280. IEEE, Florence, Italy (2013). DOI 10.1109/CDC.2013.6760220
23. Foschini, L., Hershberger, J., Suri, S.: On the complexity of time-dependent shortest paths. Algorithmica **68**(4), 1075–1097 (2014)
24. Fuentes-Pacheco, J., Ruiz-Ascencio, J., Rendn-Mancha, J.: Visual simultaneous localization and mapping: a survey. Artificial Intelligence Review pp. 1–27 (2012)
25. Gallagher, A., Chen, T.: Clothing cosegmentation for recognizing people. In: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, pp. 1–8 (2008). DOI 10.1109/CVPR.2008.4587481
26. Gemperle, F., Hirsch, T., Goode, A., Pearce, J., Siewiorek, D., Smailigic, A.: Wearable vibro-tactile display (2003). Carnegie Mellon University
27. Gescheider, G.A., Bolanowski, S.J., Hall, K.L., Hoffman, K.E., Verrillo, R.T.: The effects of aging on information-processing channels in the sense of touch: I. absolute sensitivity. Somatosens Mot Res. **11**(4), 345–357 (1994)
28. Ghosh, A., Alboul, L., Penders, J., Jones, P., Reed, H.: Following a robot using a haptic interface without visual feedback. In: Proc. Int. Conf. on Advances in Computer-Human Interactions, pp. 147–153 (2014)
29. Gillespie, R., Colgate, J., Peshkin, M.: A general framework for Cobot control. IEEE Transactions on Robotics and Automation **17**(4), 391–401 (2001)
30. Goswami, A., Peshkin, M., Colgate, J.: Passive robotics: An exploration of mechanical computation. In: Proceedings of the 1990 American Control Conference, pp. 2791–2796 (1990)
31. Guo, D., Wang, C., Wanf, X.: A hierarchical pedestrians motion planning model for heterogeneous crowds simulation. In: International Conferenceon Information and Automation (ICIA '09), pp. 1363–1367 (2009)
32. Hahnel, D., Burgard, W., Fox, D., Fishkin, K., Philipose, M.: Mapping and localization with rfid technology. In: Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on, vol. 1, pp. 1015–1020. IEEE (2004)
33. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2003)
34. Helbing, D., Molnár, P.: Social force model for pedestrian dynamics. Physical review E **51**(5), 4282–4286 (1995)
35. Hirata, Y., Hara, A., Kosuge, K.: Motion control of passive intelligent walker using servo brakes. IEEE Transactions on Robotics **23**(5), 981–990 (2007)
36. Hoppe, C., Klopschitz, M., Rumpler, M., Wendel, A., Kluckner, S., Bischof, H., Reitmayr, G.: Online feedback for structure-from-motion image acquisition. In: BMVC, pp. 1–12 (2012)
37. Ihaddadene, N., Djeraba, C.: Real-time crowd motion analysis. In: Pattern Recognition, 2008. ICPR 2008. 19th International Conference on, pp. 1–4. IEEE (2008)
38. The iwalkactive project. `http://www.iwalkactive.eu/`
39. Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., Fitzgibbon, A.: Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In: Proc. UIST, pp. 559–568 (2011)
40. Jot, J.M.: Efficient models for reverberation and distance rendering in computer music and virtual audio reality. In: Proc. International Computer Music Conference (1997)
41. Karuei, I., MacLean, K.E., Foley-Fisher, Z., MacKenzie, R., Koch, S., El-Zohairy, M.: Detecting vibrations across the body in mobile contexts. In: Proc. Int. Conf. on Human Factors in Computing Systems, pp. 3267–3276 (2011)
42. Kollmuss, A., Agyeman, J.: Mind the gap: why do people act environmentally and what are the barriers to pro-environmental behavior? Environmental education research **8**(3), 239–260 (2002)
43. Krishnan, R., Pugazhenthi, S.: Mobility assistive devices and self-transfer robotic systems for elderly, a review. Intelligent Service Robotics **7**(1), 37–49 (2014)
44. Lasovsky, Y., Joskowicz, L.: Motion planning in crowded planar environments. Robotica **null**, 365–371 (1999)
45. Latombe, J.: Robot motion planning. Kluwer Academic Publishers (1991)
46. Laugier, C., Vasquez, D., Yguel, M., Fraichard, T., Aycard, O.: Geometric and bayesian models for safe navigation in dynamic environments. Intelligent Service Robotics **1**(1), 51–72 (2008). DOI 10.1007/s11370-007-0004-1. URL `http://dx.doi.org/10.1007/s11370-007-0004-1`
47. LaValle, S., Sharma, R.: Robot motion planning in a changing, partially predictable environment. In: Proceedings of the 1994 IEEE International Symposium on Intelligent Control, pp. 261–266 (1994)
48. LaValle, S.M.: Planning Algorithms. Cambridge University Press (2006)
49. Lee, G., Ohnuma, T., Chong, N.: Design and control of JAIST active robotic walker. Intelligent Service Robotics **3**(3), 125–135 (2010)
50. Lemaire, T., Lacroix, S., Sola, J.: A practical 3d bearing-only SLAM algorithm. In: Proc. Int. Conf. on Intelligent Robots and Systems (IROS), pp. 2449–2454. Edmonton, Alberta, Canada (2005)
51. Lowe, D.: Distinctive image features from scale-invariant keypoints. In: Int. Jour. of Computer Vision, vol. 20, pp. 91–110 (2003)
52. Miura, J., Shirai, Y.: Hierarchical vision-motion planning with uncertainty: Local path planning and global route selection. In: Proceedings of the 1992 lEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 3, pp. 1847–1854 (1992)
53. Morioka, K., Yamanaka, S., Hoshino, F.: Simplified map representation and map learning system for autonomous navigation of mobile robots. Intelligent Service Robotics **7**(1), 25–35 (2014). DOI 10.1007/s11370-013-0143-5. URL `http://dx.doi.org/10.1007/s11370-013-0143-5`
54. Nazemzadeh, P., Fontanelli, D., Macii, D.: An Indoor Position Tracking Technique based on Data Fusion for Ambient Assisted Living. In: 2013 IEEE Intl. Conf. on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, pp. 7–12. IEEE, Milan, Italy (2013)
55. Nazemzadeh, P., Fontanelli, D., Macii, D., Palopoli, L.: Indoor Positioning of Wheeled Devices for Ambient Assisted Living: a Case Study. In: Proc. IEEE Int. Instrumentation and Measurement Technology Conference (I2MTC), pp. 1421–1426. IEEE, Montevideo, Uruguay (2014)
56. Nister, D., Stewenius, H.: Scalable recognition with a vocabulary tree. In: Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, vol. 2, pp. 2161–2168. IEEE (2006)
57. Panteleris, P., Argyros, A.A.: Vision-based SLAM and moving objects tracking for the perceptual support of a smart walker platform. In: Workshop on Assistive Computer Vision and Robotics (ACVR 2014), in conjunction with ECCV 2014 (2014)

58. Rizzon, L., Passerone, R.: Embedded soundscape rendering for the visually impaired. In: Proceedings of the 8^{th} IEEE International Symposium on Industrial Embedded Systems, SIES13, pp. 101–104. Porto, Portugal (2013). DOI 10.1109/SIES.2013.6601480
59. Rizzon, L., Passerone, R.: Spatial sound rendering for assisted living on an embedded platform. In: A. De Gloria (ed.) Applications in Electronics Pervading Industry, Environment and Society, *Lecture Notes in Electrical Engineering*, vol. 289, chap. 6, pp. 61–73. Springer International Publishing (2014). DOI 10.1007/978-3-319-04370-8_6
60. Scheggi, S., Aggravi, M., Morbidi, F., Prattichizzo, D.: Cooperative human-robot haptic navigation. In: Proc. IEEE Int. Conf. on Robotics and Automation, pp. 2693–2698. Hong Kong, China (2014)
61. Scheggi, S., Morbidi, F., Prattichizzo, D.: Human-robot formation control via visual and vibrotactile haptic feedback. IEEE Trans. on Haptics (2014)
62. Shahab, A., Shafait, F., Dengel, A.: Icdar 2011 robust reading competition challenge 2: Reading text in scene images. In: Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR-2011), 11th, September 18-21, Beijing, China. IEEE (2011)
63. Sinyukov, D., Desmond, R., Dickerman, M., Fleming, J., Schaufeld, J., Padir, T.: Multi-modal control framework for a semi-autonomous wheelchair using modular sensor designs. Intelligent Service Robotics **7**(3), 145–155 (2014)
64. Sosa, L.P., Lucas, S.M., Panaretos, A., Sosa, L., Tang, A., Wong, S., Young, R.: Icdar 2003 robust reading competitions. In: Proceedings of the Seventh International Conference on Document Analysis and Recognition, pp. 682–687. IEEE Press (2003)
65. Van Cauwenberg, J., Van Holle, V., Simons, D., Deridder, R., Clarys, P., Goubert, L., Nasar, J., Salmon, J., De Bourdeaudhuij, I., Deforche, B., et al.: Environmental factors influencing older adults' walking for transportation: a study using walk-along interviews. Int J Behav Nutr Phys Act **9**(1), 85 (2012)
66. Viola, P., Jones, M., Snow, D.: Detecting pedestrians using patterns of motion and appearance. In: Proceedings of the 9th IEEE International Conference on Computer Vision. Nice, France (2003)
67. Voigt, R., Nikolic, J., Hurzeler, C., Weiss, S., Kneip, L., Siegwart, R.: Robust embedded egomotion estimation. In: Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on, pp. 2694–2699. IEEE (2011)
68. Wang, C., Thorpe, C., Hebert, M., Thrun, S., Durrant-whyte, H.: Simultaneous localization, mapping and moving object tracking. International Journal of Robotics Research (2004)
69. Wang, K., Babenko, B., Belongie, S.: End-to-end scene text recognition. In: Proceedings of the 2011 International Conference on Computer Vision, ICCV '11, pp. 1457–1464. Washington, DC, USA (2011)
70. Warburton, D.E., Nicol, C.W., Bredin, S.S.: Health benefits of physical activity: the evidence. Canadian medical association journal **174**(6), 801–809 (2006)
71. Y. Bar-Shalom X. Rong Li, T.K.: Estimation with Application to Tracking and Navigation – Theory, Algorithm and Software. John. Wiley and Sons (2001)
72. Yao, C., Bai, X., Liu, W., Ma, Y., Tu, Z.: Detecting texts of arbitrary orientations in natural images. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1083–1090. Washington, DC, USA (2012)
73. Yilmaz, A., Javed, O., Shah, M.: Object tracking: A survey. ACM Comput. Surv. **38**(4) (2006)
74. Younes, H., Simmons, R.: Probabilistic verification of discrete event systems using acceptance sampling. In: CAV, vol. 2404, pp. 23–39. Springer (2002)