2012 IEEE/RSJ International Conference on Intelligent Robots and Systems
Vila Moura, Algarve, Portugal, October 7-12

IROS 2012 International Workshop on

Assistance and Service Robotics in a Human Environment

Full Day Workshop

October 12th, 2012, Vila Moura, Portugal

http ://emotion.inrialpes.fr/people/spalanzani/WorkshopIROS12/WorkshopIROS12.html

Organizers

Anne Spalanzani (INRIA Rhone-Alpes)
David Daney (INRIA Sophia Antipolis Mediterranée)
Olivier Simonin (INRIA Nancy Grand Est)
Jean-Pierre Merlet (INRIA Sophia Antipolis Mediterranée)

Proceedings of the IROS 2012 Workshop “Assistance and Service Robotics in a Human Environment”

2012 IEEE/RSJ International Conference on Intelligent Robots and Systems
Vila Moura, Algarve, Portugal, October 7-12

Foreword

This workshop aims to present **research activities in the area of robotic assistance of human in different contexts of human life**. The integral assistance systems are **robotic modules and technological aids in general for personal assistance, such as robots, mobile bases, electric wheelchairs, soft robot manipulator arm**. They can support **disabled and elderly people** with special needs in their living environment. Intelligent Service Robotics cover a broad spectrum of research axis, from intelligent robots acting as a servant, secretary, or companion to intelligent robotic functions such as autonomous wheelchair navigation, embedded robotics, ambient intelligence, or intelligent space. This workshop will focus on the **assistance of human in terms of its mobility, its social interaction, as well as its everyday chores that are especially pertinent to the elderly**. Intended audience concerns researchers and PhD students interested in assistive robotics, personal robotics, manipulation, Mobility aids, social interaction, teleoperated robots, Human-machine Interfaces for assistive robotics, Acceptance, Activity monitoring systems, Activity recognition, Elderly care assistive robots, Rehabilitation system, transfert machine, Sensor networks, Perception, Smart environments...

This workshop is composed with 4 invited talks and 11 selected papers. Four sessions have been organized :
Session I : Robot localization and people tracking
Session II : Navigation and manipulation in human populated environments
Session III : Behavioral modeling and Human/Robot Interaction
Session IV : Robotics for elderly and frail people

Anne Spalanzani, David Daney, Olivier Simonin, Jean-Pierre Merlet

Session I

Robot localization and people tracking

- Title : **Localization and Navigation of an Assistive Humanoid Robot in a Smart Environment .** Authors : E. Cervera, A. Abou Moughlbay, P. Martinet
- Title : **Tracking Mobile Objects with Several Kinect using HMMs and Component Labelling.** Authors : A. Dubois F. Charpillet
- Title : **Autonomous Shopping Cart Platform for People with Mobility Impairments.** Authors : L. Marchetti, D. Pucci, P. Morin

Localization and Navigation of an Assistive Humanoid Robot in a Smart Environment

Enric Cervera
Robotic Intelligence Lab
Jaume-I University
Castello-Spain
ecervera@icc.uji.es

Amine Abou Moughlbay
Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN)
Ecole Centrale de Nantes - France
Amine.abou-moughlbay@ec-nantes.fr

Philippe Martinet
IRCCyN - Ecole Centrale de Nantes
Institut Pascal - Clermont Ferrand
France
Philippe.Martinet@irccyn.ec-nantes.fr

***Abstract*—Assistive humanoids that manipulate objects in everyday environments are potentially useful to improve the lives of the elderly or disabled. To retrieve or deliver objects at home, precise localization is needed. But localization of humanoid robots is a challenging issue, due to rough odometry estimation, noisy onboard sensing, and the swaying motion caused by walking. To overcome these limitations, we advocate for the use of external sensors for localization and navigation in a smart home environment. As opposed to a stand-alone self contained robot, our humanoid benefits from the information coming from other sensing devices in the environment. In order to achieve robust localization while walking, and retrieve an object from the floor, we use RGBD camera information from an external Kinect sensor. Monte Carlo localization estimates the 6D torso pose estimation of the humanoid, which is then used for closed-loop navigation control. Experiments with a NAO humanoid point out that, by cooperating with the environmental sensors, the overall precision of robot navigation is dramatically improved.**

I. Introduction

Object retrieval is remarked as a high priority task for assistive robots by people with physical disabilities. Humanoid robots could potentially help people with motor impairments to retrieve dropped objects [1], [2]. But robust and precise robot localization is a prerequisite for such task, and humanoid localization remains a challenging issue due to inaccurate foot step odometry and noisy onboard sensor observations during walking [3].

Since our humanoid is targeted to human-friendly indoor environments, a smart home endowed with networked sensors would provide additional useful information to monitor both humans and robots. In smart environments, components are working together by exchanging information via the local network. The main idea behind the smart home concept is to use networked robots to integrate different services within the home as a means to control and monitor the entire living space [4]. Such services are not realized by a single full-equipped robot but by a combination of different elements such as environmental sensors, cameras and human communicating and cooperating through the network.

The Kinect is a powerful low-cost sensor which provides color and range images, suitable for human motion detection and tracking. Dingli et al. [5] have created an Ambient-Assisted Living application which monitors a person's position, labels objects around a room and raises alerts in case of falls. Stone and Skubic [6] have investigated this sensor for in-home fall risk assessment. Ni et al. [7] use color-depth fusion schemes for feature representation in human action recognition.

In this paper, we present a localization and navigation method for a humanoid robot in an indoor smart environment. The Kinect sensor is used to monitor and track the 6D pose of the robot. Localization and navigation towards a detected object can then be achieved precisely. The rest of the paper is organized as follows: related work on humanoid localization is discussed in Section II; Section III presents the architecture of the system; localization and control of the humanoid is described in Section IV; experimental results are presented in Section V; finally, Section VI draws some conclusions and outlines future work extensions.

II. Related work

Accurate localization, which is considered to be mainly solved for wheeled robots, is still a challenging problem for humanoid robots [3]. When dealing with biped robots, many problems arise such as foot slippage, stability problems during walking, and limited payload capabilities, preventing precise localization in their environment. In addition, humanoids usually cannot be assumed to move on a plane to which their sensors are parallel due to their walking motion.

In the last few years, Monte Carlo methods have been commonly used to perform localization on mobile robots [8], as well as other methods including grid-based Markov localization and Kalman filtering [9]. Furthermore, many studies have been made to solve the humanoid localization problem by tracking their pose in the two-dimensional space. Ido et al. [10] applied a vision-based approach and compare the current image to previously recorded reference images in order to estimate the location of the robot. Owald et al. [11] and Bennewitz et al. [12] compared visual features to a previously learned 2D feature map during pose tracking. Pretto et al. [13] tracked visual features over time for estimating the robot's odometry. Cupec et al. [14] detected objects with given shapes and colors in the local environment of the humanoid and determine its pose relative to these objects.

In addition to that, many techniques using laser range data have also been developed. Stachniss et al. [15] presented an approach to learn accurate 2D grid maps of large environments with a humanoid equipped with a Hokuyo laser scanner. Such a map was subsequently used by Faber et al. [16] for humanoid localization in 2D. Similarly, Tellez et al. [17] developed a navigation system for such a 2D environment representation using two laser scanners located in the feet of the robot.

Since a 2D map is often not sufficient for humanoid motion planning, several methods use 2.5D grid maps which additionally store a height value for each cell. Thompson et al. [18] track the 6D pose of a humanoid equipped with a laser scanner in such a representation. Hornung et al. [3] track a humanoid's 6D pose in a 3D world model, which may contain multiple levels connected by staircases.

In the cited methods, they used either embedded cameras and sensors on the robot or exteroceptive ones which are mounted on the robot's head or body. All these sensors are usually used to track the environment and subsequently localize the robot.

In our method, the sensors are fixed and used to localize the walking robot. The contribution of this paper is a robust localization system for humanoid robots navigating in indoor environments using Kinect cameras. The main goal is to develop a robust system which is able to track and localize the robot when walking, and control its motion precisely enough, to be able to retrieve an object from the floor.

III. System's Architecture

The system is composed of a small humanoid robot NAO, navigating in an indoor smart environment, where a 3D vision system, consisting of one or more low cost Kinect cameras, monitors and tracks both the user and robot activity, as depicted in Fig. 1.

In our current implementation, the system is able to detect small objects lying on the floor plane, as well as to localize the robot. In the future, we will incorporate the skeletal tracking of the Kinect sensor, to be able to interface directly with human gestures. This section briefly describes the hardware components along with the system software framework.

A. Kinect sensor

The used vision system is the Kinect camera, which consists of two optical sensors whose interaction allows a three-dimensional scene analysis. One of the sensors is an RGB camera which has a video resolution of 30 fps. The image resolution given by this camera is 640x480 pixels. The second sensor has the aim of obtaining depth information corresponding to the objects found at the scene. The working principle of this sensor is based on the emission of an infrared signal which is reflected by the objects and captured by a monochrome CMOS sensor. A matrix is then obtained which provides a depth image of the objects in the scene, called DEPTH. An investigation of the geometric quality of depth data obtained by the Kinect sensor was done by [19], revealing that the point cloud does not contain large systematic errors when compared with a laser scanning data.

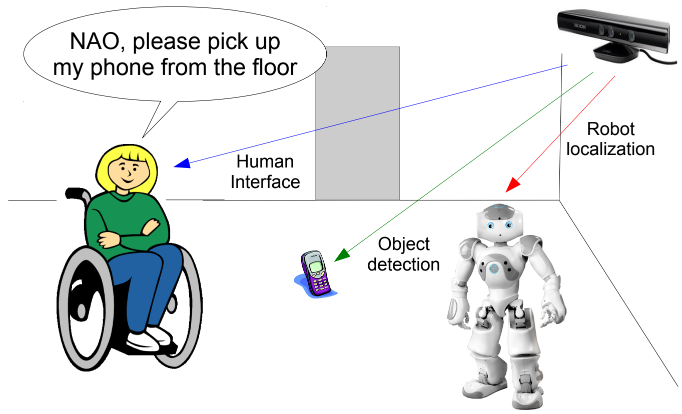


Fig. 1. Overview of the system: the Kinect sensor in the smart environment is able to monitor the user activity, detect objects on the floor, and localize precisely the robot.

B. NAO Robot

The NAO Robot [20], developed by Aldebaran robotics, is a biped robot with 25 Degrees of Freedom (DOF). It has 3-fingered robotic hands used for grasping and holding small objects (it can carry up to 300g using both hands). It is equipped with: 2 ultrasound devices situated in the chest that provide space information in a range of 1 meter, 2 cameras situated on the top and bottom of the head, 2 bumpers (contact sensors on the robot's feet), a gyrometer and an accelerometer (to determine whether the robot is in a stable or unstable position).

C. Software overview

The Point Cloud Library (PCL) is a standalone, large scale, open project for 3D point cloud processing. [21]. The PCL framework contains numerous state-of-the art algorithms including filtering, feature estimation, surface reconstruction, registration, model fitting and segmentation, as well as higher level tools for performing mapping and object recognition.

The tracking module of PCL [22] provides a comprehensive algorithmic base for the estimation of 3D object poses using Monte Carlo sampling techniques and for calculating the likelihood using combined weighted metrics for hyper-dimensional spaces including Cartesian data, colors, and surface normals.

ROS (Robot Operating System) is a software framework which provides libraries and tools to help developers create robot applications. It provides hardware abstraction, device drivers, libraries, visualizers, message-passing, package management, and more [23].

With its modular design, ROS makes it easy to develop network robot systems. Seamless interfacing with the Kinect sensor, NAO robot, and PCL is provided.

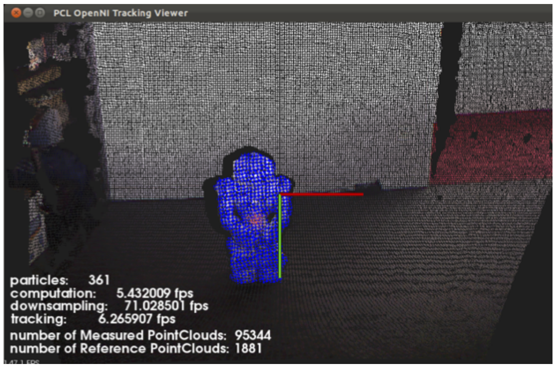


Fig. 2. 3D tracking of the NAO robot with Kinect.

IV. Localization and navigation

A. Localization

This technique consists of tracking 3D objects (position and orientation) in continuous point cloud data sequences. It was originally designed for robots to monitor the environment and make decisions and adapt their motions according to the changes in the world, but it is equally suitable for tracking the robot while it walks around the environment (Fig. 2). Localization has been optimized to perform computations in real-time, by employing multi CPU cores optimization, adaptive particle filtering (KLD sampling) and other modern techniques [22].

In our application, we use a rigid model of the torso and head parts of the robot, to track the system on real-time and find the 3D pose of the robot model even when walking. In future works, an articulated model could be used to track the whole body of the robot.

Using the PCL Cloud tracking technique we can find the actual pose of the robot and the desired one (the object) with respect to the Kinect. Thus the relative 3D pose can be calculated, and we can control the direction of walking in the plane: the position in X and Y directions and the orientation of the robot.

B. Navigation

The NAO robot is able to walk in velocity control mode. This enables the walk to be controlled reactively, as the most recent command overrides all previous commands. However, the walk uses a preview controller to guarantee stability. This uses a preview of time of 0.8s, so the walk will take this time to react to new commands. At maximum frequency this equates to about two steps [24].

The linear velocity of the robot $({}^cV_x, {}^cV_y)^T$ is calculated with respect to the pose error between actual and desired poses $(e_x, e_y, e_\theta)^T$ using a proportional gain λ. The aim is to move the robot linearly to the target.

$$\begin{pmatrix} {}^cV_x \\ {}^cV_y \end{pmatrix} = \lambda \begin{pmatrix} e_x \\ e_y \end{pmatrix} \tag{1}$$

The angular velocity of the robot ${}^c\omega$ may be calculated in two different manners: first, while the error is greater than a given threshold e_{min}, the robot will head towards the line joining its current location and final destination; second, when the error is lower than such threshold, the robot will head towards its final orientation.

$$ {}^c\omega = \begin{cases} \lambda \arctan(e_y/e_x) & \text{if } \sqrt{e_x^2 + e_y^2} > e_{min} \\ \lambda e_\theta & \text{otherwise} \end{cases} \tag{2}$$

C. Veering correction

Closed-loop control is able to converge in the presence of uncertainties in sensors and actuators. However, faster convergence is achieved if the system is properly calibrated. In a similar way to human beings [25], biped robots suffer from inability to maintain a straight path when walking without vision: slight differences between each leg stride caused by backslash, friction, or motor power will produce a veering behavior which can be estimated and corrected.

A simple veering correction procedure is now introduced: the robot is commanded in open-loop to walk straight away with a constant linear velocity. Its trajectory is recorded with the Kinect sensor, and two fitting steps are performed:

1) Circle fitting: the best fitting circle is computed, as shown in Fig. 3a.
2) Line fitting: the robot is now commanded to walk with a constant linear velocity, and a constant angular velocity, as computed from the previous fitting step. Fig. 3b depicts the recorded trajectory, and the LMS line fitting.

The error is significantly reduced: without correction, for a 1m walked distance, the lateral error grows to 35cm, and the orientation error is 30 (Fig. 3a). With angular correction (Fig. 3b), the orientation error is not noticeable, but a lateral error persists, 25cm for a 1.5m walked distance. This error is compensated an order of magnitude with a lateral velocity term, resulting in only 3cm for the same walked distance (Fig. 3c).

As a result, for a given command motion, the actual velocity sent to the robot $({}^rV_x, {}^rV_y, {}^r\omega)^T$ consists of the original commanded values $({}^cV_x, {}^cV_y, {}^c\omega)^T$ and compensation terms for the lateral and angular velocity. The radius of the fitting circle R and the slope of the fitting line m are used to compute the angular and lateral velocity compensation respectively:

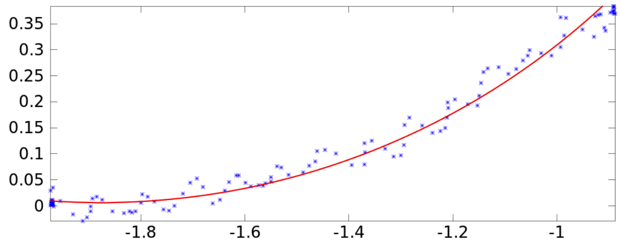
$$\begin{pmatrix} {}^rV_x \\ {}^rV_y \\ {}^r\omega \end{pmatrix} = \begin{pmatrix} {}^cV_x \\ {}^cV_y - m\,{}^cV_x \\ {}^c\omega - R\,{}^cV_x \end{pmatrix} \tag{3}$$

Finally, the velocity is translated to NAO's walk arguments (see [24] for details).

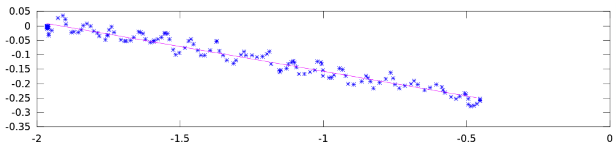
V. Experimental Results

In the experiment, the robot is commanded to approach an object lying on the floor as seen in Fig. 4. The initial distance to the object is 1.5m and the orientation of the initial pose with respect to the final pose is 25 degrees.

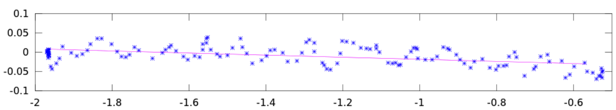
Figs. 5, 6 and 7 depict respectively the commanded velocity, the pose error, and the planar trajectory of the robot.



(a) Circle fitting.



(b) Line fitting.



(c) Final trajectory.

Fig. 3. Veering correction: the robot is commanded open-loop to walk straight with constant linear velocity. First, trajectory is recorded (a) and the best fitting circle is computed. Second, a new trajectory is recorded (b) and a linear fitting is computed. The final trajectory with angular and lateral compensation is shown in (c).

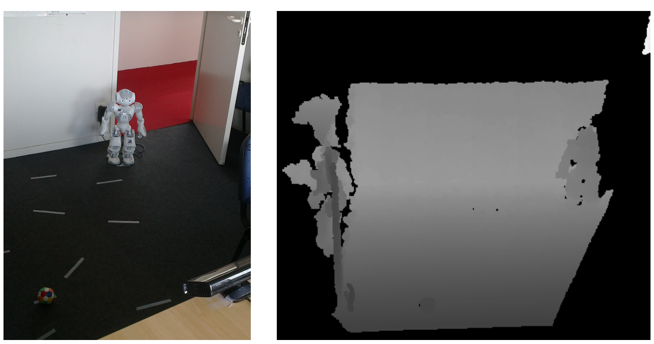


Fig. 4. Experimental setup (external and Kinect views): as NAO robot enters the room, it is commanded to move towards an object lying on the floor (the ball on the left-bottom corner of the image). Floor lines are not considered in the experiment.

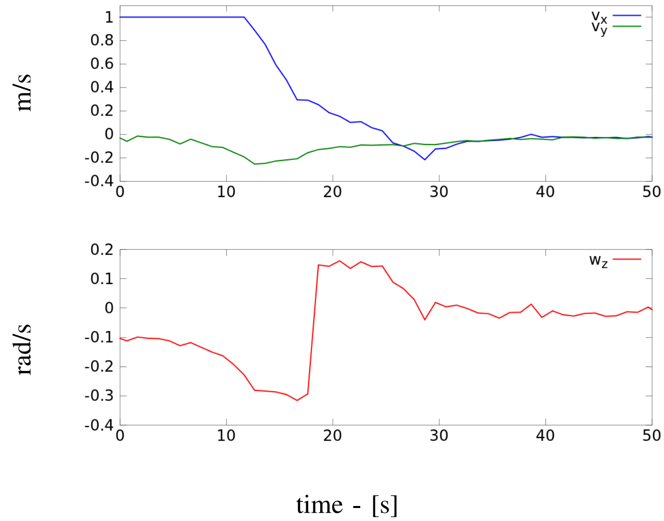


Fig. 5. Robot velocity in NAO's local frame: only planar and angular velocity is sent to the robot controller.

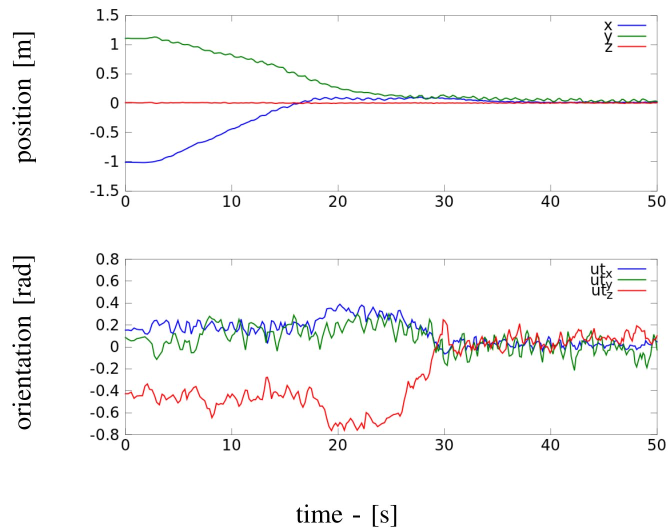


Fig. 6. Pose error of the robot torso (position and orientation), as measured by the Kinect sensor.

As can be seen in Fig. 5, the robot initially walks towards the destination at full speed, then it progressively decreases its velocity as the error is diminished. The profile of the angular velocity reflects the two stages defined in 2: from 0s to 19s, the robot turns towards the destination; after 19s the robot is nearer to the destination than the threshold (fixed to 30cm) and it turns to its final orientation.

This control strategy explains why the angular error ut_z in Fig. 6 does not decrease initially, since this error is measured with respect to the final orientation of the robot.

The trajectory of the robot in the room is depicted in Fig. 7. When using only odometry, the robot is not able to attain the destination goal, but the final pose is 40 cm away from the desired one. On the other hand, localization and closed-loop navigation allows the robot to attain the goal within a few centimeters precision. The final mean position error is $(x, y, \theta) = (0.012, 0.018, 0.07)$ and the standard deviation is $(\sigma_x, \sigma_y, \sigma_\theta) = (0.002, 0.004, 0.05)$. In preliminary experiments, it has been possible to fetch an object from the floor in such conditions.

VI. Conclusion

We have presented a localization and navigation method for a humanoid robot in a smart environment, where a Kinect sensor is used for monitoring and tracking the 3D pose of the robot.

This method is suitable for indoor environments, allowing not only to detect the robot but also to track people who

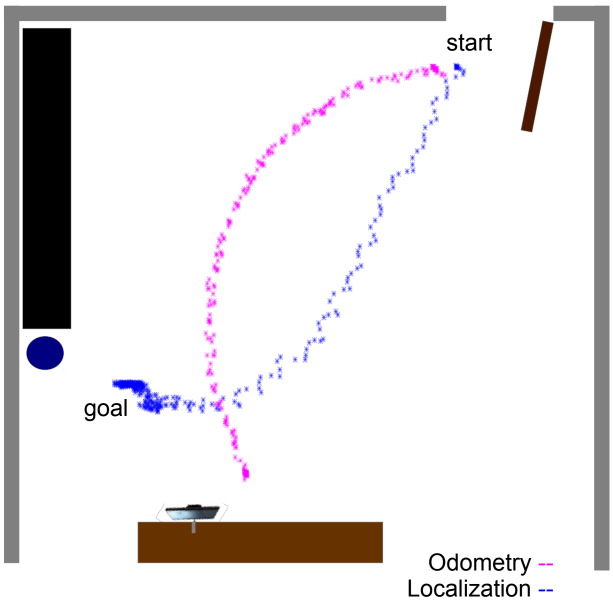


Fig. 7. Trajectories carried out in the experiments: while the odometry diverges, localization is able to attain the goal.

interact with the robot.

The accomplished precision of localization makes it possible to retrieve objects from the floor for assistance and service robotics. Further work will incorporate object manipulation to complete such task.

Further improvements are possible: obstacle detection with the Kinect sensor would allow the robot to navigate robustly in a cluttered environment. The workspace of the robot would be scalable by using more Kinect sensors in a networked system. Finally, multimodal user interaction can be achieved by gesture and voice recognition.

ACKNOWLEDGMENT

This research was partly funded by Ministerio de Ciencia e Innovación (DPI2011-27846), Ministerio de Educación (Programa "Salvador de Madariaga", referencia PR2011-015) Generalitat Valenciana (PROMETEO/2009/052) and Fundació Caixa Castelló-Bancaixa (P1-1B2011-54). We wish to thank W. Khalil for making it possible for one of the authors to stay at the IRCCyN, Ecole Centrale de Nantes.

REFERENCES

[1] A. Jain and C. Kemp, "El-e: an assistive mobile manipulator that autonomously fetches objects from flat surfaces," *Autonomous Robots*, vol. 28, pp. 45–64, 2010.

[2] C.-H. King, T. L. Chen, Z. Fan, J. D. Glass, and C. C. Kemp, "Dusty: an assistive mobile manipulator that retrieves dropped objects for people with motor impairments," *Disability and Rehabilitation: Assistive Technology*, vol. 7, no. 2, pp. 168–179, 2012.

[3] A. Hornung, K. M. Wurm, and M. Bennwitz, "Humanoid robot localization in complex indoor environments," in *2010 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems - IROS'10*, oct. 2010, pp. 1690 – 1695.

[4] A. Sanfeliu, N. Hagita, and A. Saffiotti, "Network robot systems," *Robotics and Autonomous Systems*, vol. 56, no. 10, pp. 793 – 797, 2008.

[5] A. Dingli, D. Attard, and R. Mamo, "Turning homes into low-cost ambient assisted living environments," *International Journal of Ambient Computing and Intelligence (IJACI)*, pp. 1–23, 2012.

[6] E. Stone and M. Skubic, "Evaluation of an inexpensive depth camera for passive in-home fall risk assessment," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on*, may 2011, pp. 71 –77.

[7] B. Ni, G. Wang, and P. Moulin, "Rgbd-hudaact: A color-depth video database for human daily activity recognition," in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, nov. 2011, pp. 1147 –1153.

[8] S. Thrun, D. Fox, W. Burgard, and F. Dellaert, "Robust monte carlo localization for mobile robots," 2001. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.18.8488

[9] J.-S. Gutmann, "Markov-kalman localization for mobile robots." in *ICPR (2)*, 2002, pp. 601–604. [Online]. Available: http://dblp.uni-trier.de/db/conf/icpr/icpr2002-2.html#Gutmann02

[10] J. Ido, Y. Shimizu, Y. Matsumoto, and T. Ogasawara, "Indoor navigation for a humanoid robot using a view sequence." *I. J. Robotic Res.*, vol. 28, no. 2, pp. 315–325, 2009. [Online]. Available: http://dblp.uni-trier.de/db/journals/ijrr/ijrr28.html#IdoSMO09

[11] S. Owald, A. Hornung, and M. Bennewitz, "Learning reliable and efficient navigation with a humanoid." in *ICRA*. IEEE, 2010, pp. 2375–2380. [Online]. Available: http://dblp.uni-trier.de/db/conf/icra/icra2010.html#OsswaldHB10

[12] M. Bennewitz, C. Stachniss, W. Burgard, and S. Behnke, "Metric localization with scale-invariant visual features using a single perspective camera." in *EUROS*, ser. Springer Tracts in Advanced Robotics, H. I. Christensen, Ed., vol. 22. Springer, 2006, pp. 195–209. [Online]. Available: http://dblp.uni-trier.de/db/conf/euros/euros2006.html#BennewitzSBB06

[13] A. Pretto, E. Menegatti, M. Bennewitz, W. Burgard, and E. Pagello, "A visual odometry framework robust to motion blur." in *ICRA*. IEEE, 2009, pp. 2250–2257. [Online]. Available: http://dblp.uni-trier.de/db/conf/icra/icra2009.html#PrettoMBBP09

[14] R. Cupec, G. Schmidt, and O. Lorch, "Experiments in vision-guided robot walking in a structured scenario," in *Industrial Electronics, 2005. ISIE 2005. Proceedings of the IEEE International Symposium on*, vol. 4, 20-23, 2005, pp. 1581 – 1586.

[15] C. Stachniss, M. Bennewitz, G. Grisetti, S. Behnke, and W. Burgard, "How to learn accurate grid maps with a humanoid." in *ICRA*. IEEE, 2008, pp. 3194–3199. [Online]. Available: http://dblp.uni-trier.de/db/conf/icra/icra2008.html#StachnissBGBB08

[16] F. Faber, M. Bennewitz, A. Grg, C. Gonsior, D. Joho, M. Schreiber, and S. Behnke, "The humanoid museum tour guide robotinho," in *in IEEE Int. Symp. on Robot and Human Interactive Communication*, 2009.

[17] R. A. Tllez, F. Ferro, D. Mora, D. Pinyol, and D. Faconti, "Autonomous humanoid navigation using laser and odometry data." in *Humanoids*. IEEE, 2008, pp. 500–506. [Online]. Available: http://dblp.uni-trier.de/db/conf/humanoids/humanoids2008.html#TellezFMPF08

[18] S. Thompson, S. Kagami, and K. Nishiwaki, "Localisation for autonomous humanoid navigation." in *Humanoids*. IEEE, 2006, pp. 13–19. [Online]. Available: http://dblp.uni-trier.de/db/conf/humanoids/humanoids2006.html#ThompsonKN06

[19] K. Khoshelham and S. O. Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensors*, vol. 12, no. 2, pp. 1437–1454, 2012.

[20] D. Gouaillier, V. Hugel, P. Blazevic, C. Kilner, J. Monceaux, P. Lafourcade, and B. Marnier et al., "Mechatronic design of nao humanoid," in *IEEE Int. Conf. on Robotics and Automation*, may 2009, pp. 769 –774.

[21] R. Rusu and S. Cousins, "3d is here: Point cloud library (pcl)," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, may 2011, pp. 1–4.

[22] R. Ueda, "Tracking 3D objects with Point Cloud Library," http://pointclouds.org/news/tracking-3d-objects-with-point-cloud-library.html, 2012, [Online; accessed 27-June-2012].

[23] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: an open-source robot operating system," in *ICRA Workshop on Open Source Software*, 2009.

[24] A. Robotics, "Walk control – NAO Software 1.12.5 documentation," http://www.aldebaran-robotics.com/documentation/naoqi/motion/control-walk.html, 2012, [Online; accessed 27-June-2012].

[25] C. S. Kallie, P. R. Schrater, and G. E. Legge, "Variability in stepping direction explains the veering behavior of blind walkers," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 33, no. 1, 2007.

Tracking Mobile Objects with Several Kinect using HMMs and Component Labelling

Amandine Dubois[1,2] and François Charpillet[2,1]

***Abstract*— This paper proposes a markerless system whose purpose is to detect falls of elderly people at home. To track human movements we use Microsoft Kinect camera which allows to acquire at the same time a RGB image and a depth image. One contribution of our work is to propose a method for fusioning the information provided by several Kinects based on an occupancy grid. The observed space is tesselated into cells forming a 3D occupancy grid. We calculate a probability of occupation for each cell of the grid independently of its nearby cells. From this probability we distinguish whether the cells are occupied by a static object (wall) or by a mobile object (chair, human being) or whether the cells are empty. This categorization is realized in real-time using a simple three states HMM. The use of HMMs allows to deal with an aliasing problem since mobile objects result in the same observation as static objects. This paper also deals with the issue regarding the presence of several people in the field of view of camera by using the component labelling method. The approach is evaluated in simulation and in a real environment showing an efficient real-time discrimination between cells occupied by different mobile objects and cells occupied by static objects.**

I. Introduction

During the last fifty years, the many progresses of medicine as well as the improvement of the quality of life have resulted in a longer life expectancy. The increase of the number of elderly people is a matter of public health because although elderly people can age in good health, old age also causes embrittlements in particular on the physical plan which can result in a loss of autonomy. If we look at elderly people (people of more than 65 years old), we realize that one of the main preoccupations at this age is fall. Indeed by looking at the figures of the INPES [1] we notice that one elderly people out of three falls in the year. Our work is related to fall prevention and detection.

Many systems exist to detect falls. One of the categories consists in systems with sensors, that the person wears on her. These sensors are either accelerometer, gyroscopes or goniometers. These various sensors can be integrated in devices detecting the fall automatically, as shown in article of Bourke *et al.* [2]. There exists also systems made of an alarm button, in this case it's the person who must press herself on a button to alert after the fall. Another category of fall detection devices are the systems using 2D or 3D cameras. For example Jansen *et al.* [3] used a 3D camera in order to classify the pose of a person among an *a priori* set of characteristic poses: "standing", "sitting" or "lying down". For working on detection falls, we interessed us in this last category of devices.

We chose a RGB depth camera, more precisely the Kinect camera. One of the disadvantages of the Kinect camera is its rather reduced field of view. The sensor can maintain tracking approximately 0.7–6 m. However we want to detect falls in parts of an apartment often larger than what the Kinect camera is able to see. Thus to cover a room entirely we should integrate several cameras in the same room.

This article is focused on the problem of human movements tracking with several Kinect. For this, three problems need to be solved. The first one is the fusion of the information provided by the different cameras. The second one is the discrimination between mobile objects and background. The last problem consists in gathering the mobile elements belonging to the same object. This stage makes it possible to identify several people as distinct in the same scene. In earlier work [4], we presented a method to detect pixels belonging to a mobile object adressing to the two first problems of human movements tracking. This paper entends previous work by adressing the third problem. We identify pixels belonging to the same object.

For making the fusion of the information provided by several 3D cameras, our method is to use a 3D occupancy grid. From the depth image, we create a spatial representation called an occupancy grid. In this representation, the space is divided into cells of a few centimeters with a probabilistic occupation state.

The second problem for tracking persons is to succeed at discriminating mobile objects from the background. Our approach is based on an extension of occupancy grids [5] using hidden Markov models such that each voxel of the grid is determined by a three state model (the voxel belongs to the background, the voxel belongs to a mobile object, the voxel is not occupied). Our work is related to others. Among them, let us quote Yapo *et al.* [6] who proposed a method to detect 3D objects using LIDAR sensors. Their approach is also based on the concept of occupancy grids. From a probabilistic representation they determine if the voxels are free, occupied or hidden.

The third problem of identification of the mobile objects is carried out thanks to the component labelling method which consist in gathering the voxels belonging to the same mobile object.

This paper is organized as follows. Section II is dedicated to the data fusion of several Kinect cameras. Then, section III describes the method for categorizing the voxels (cells) occupation state. Section IV explains the method to identify

[1]Université de Lorraine
[2]INRIA
LORIA – Campus scientifique, BP 239, 54506 Vandœuvre-lès-Nancy CEDEX

the different mobile objects. Finally in section V, we present experimental results obtained with our method.

II. Kinect fusion

We use an occupancy grid as a common coordinate system so as to fusion information from several Kinect cameras.

A. Occupancy grids

Occupancy grids such as defined in the article of Alberto Elfes [5] consist in dividing into cells a 2D or in our case a 3D space. The grid provides a representation of the environment. For each cell C_i (or voxels) we estimate its state, which can be either occupied or empty, from a probability of occupation $P(C_i|r)$. The probability of occupation is the probability of cell C_i to be in state occupied given sensor reading r. For the sake of simplicity, each cell is estimated independently of its nearby cells.

B. Calibration of Kinect

The fusion information of several Kinect consists in determining the position and orientation of the different Kinect compared to the common coordinate system (the grid). For that we calibrate the Kinect cameras. The aim of camera calibration is to determine the transformation matrix (represented in Figure 1) between the RGB camera of each Kinect (K_{RgbRgb}), between RGB camera and the grid ($K_{RgbGrid}$), between RGB and depth camera of the same Kinect ($K_{RgbDepth}$) and between grid and depth camera ($K_{GridDepth}$). So with all these transformation matrix, the cells of the grid can be projected to the different camera coordinate systems (Kinect) and we will be able to define an observation function for each Kinect presented as in section III-B.

The transformation matrix K_{RgbRgb} and $K_{RgbGrid}$ can be determined with the method of epipolar geometry [7] or chessboard calibration.

The transformation matrix $K_{RgbDepth}$ is known by manufacturer of Kinect (provided by openNI).

The transformation matrix $K_{GridDepth}$ is deduced from previous transformations with the following equation :

$$K_{GridDepth} = K^{-1}_{RgbGrid} \times K_{RgbDepth}$$

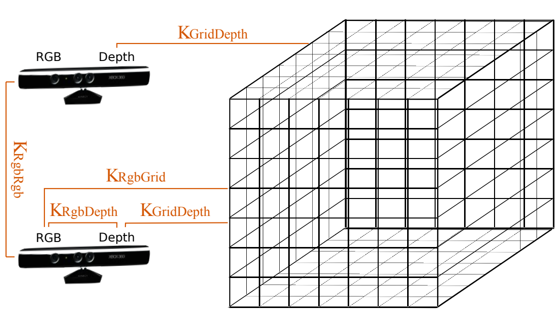


Fig. 1. Kinect grid calibration.

III. Classification of cells

A. HMM models

We want to distinguish mobile objects (chair, human beings) from static objects (walls). We define a cell containing a mobile object as being an occupied cell that has previously been empty. Whereas cells containing a static object are cells that are occupied and that have never been empty. In the classical occupancy grid method the state "empty" or "occupied" is calculated from the probability of occupation $P(r|C_i)$. We can consider the occupancy grid model as a two states HMM with no transitions between states as shown in Figure 2.

Fig. 2. Representation of the occupancy grid model (O: occupied; E: empty). Notice that $P(O) = 1 - P(E)$.

We modify the method of occupancy grid. For each cell, we use a three states HMM allowing to represent its dynamic and to determine its state. We remind that each cell is estimated independently of its nearby cells. The three states are:

- the state "O" meaning that the cell is occupied and has always been occupied;
- the state "M" meaning that the cell is occupied but has already been empty at least once;
- the state "E" meaning that the cell is not occupied.

In other words we can write:

- for mobile objects: $C_i = M$
- for statics objects: $C_i = O$
- for cells occupied by object: $C_i = M \vee O$

The representation of this HMM is shown in Figure 3.

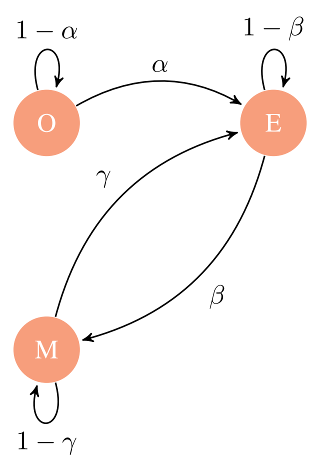


Fig. 3. The same HMM is used to model the evolution of each cell.

Probabilities of transition are $\alpha = 0.01$, $\beta = 0.1$ and $\gamma = 0.4$.

These probabilities respect the following assumptions :

- the detected cells in state O are regarded as being part of a wall. Consequently these cells have a weak probability of changing state. That's why the probability of transition of passing from O to another state is very weak. This probability is however not null because there is possibility that the model was mistaken since at the beginning it cannot make the difference between a wall and a mobile object which did not move yet (like a chair) ;
- when a cell appeared at least once free, there is no possibility for this cell to return to O because it is not a cell belonging to a wall. If not, this cell would have never passed in a free state ;
- the difference between a cell occupied in O and in M lies only in the passage or not in the state E that's why there is no transition between O and M.

B. Observation function

Each voxel C_i is represented by its center of mass, defined by coordinates (x,y,z). We can obtain at which distance is located the voxel from the camera by projecting the voxel to the camera coordinate system using the camera transformation matrix $K_{GridDepth}$. We denote as l this distance. The distance l of the voxel is compared to the depth, denoted as d, of the corresponding pixel provided by the Kinect camera (Figure 5). The observation r (see the section II-A) takes as value the error of distance (ε) between d and l calculated as $\varepsilon = d - l$. An observation function is built to evaluate the probability of occupation of the cell from the depth image $P(r|C_i) = f(\varepsilon)$. $f(\varepsilon)$ is represented in Figure 4.

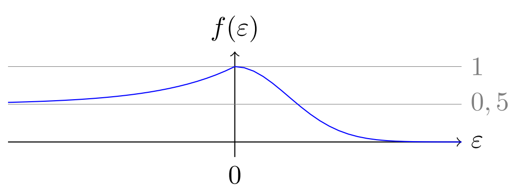


Fig. 4. Representation of occupation probability.

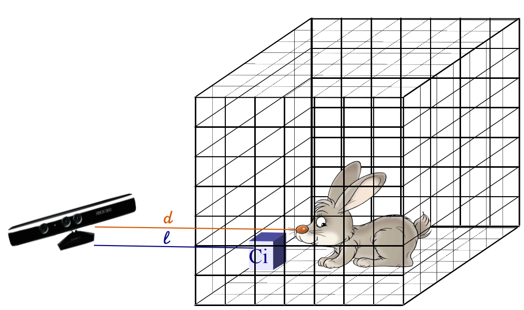


Fig. 5. d : distance between Kinect and object in occupancy grid, l : distance between Kinect and cells.

Assuming that the information provided by the different camera is conditionally independent, we can multiply the different observation functions:

$$P(r_1, ..., r_N|C_i) = \prod_{j=1}^{N} f(\varepsilon_j)$$

where N is the number of cameras and ε_j the error of distance calculated with camera j.

C. Inference

To calculate the probability to be in one of the three states of HMM, we use to the Forward procedure [8]. We don't use the procedure Backward because we want the result to be given online.

The three observation functions are given by :

$$\begin{aligned} P(r|C_i = O) &= f(\varepsilon) \\ P(r|C_i = M) &= f(\varepsilon) \\ P(r|C_i = E) &= 1 - f(\varepsilon). \end{aligned}$$

Cells are categorized by choosing the maximum *a posteriori* (MAP), that is to say the most likely state of the corresponding HMM.

We denote as po_t, pm_t, pe_t the probability of a cell being respectively occupied, mobile or empty at time t. We fix the initial probability with $po_0 = pe_0 = 0.5$ and $pm_0 = 0.0$.

IV. Identification of different mobile objects

In this part the aim is to gather the mobile cells belonging to the same object, so it will be possible to distinguish several persons in the same scene.

To gather the mobile cells we used the method «Component labelling» [9]. This method consists in assigning a label (a number) to each cell detected as mobile (state M of HMM). In Figure 6a), the colored cells are cells in state M. Thus in Figure 6b), the algorithm assign a different number to each colored cells. Then the technique is to look for each cell p if one of its neighbors has a smaller number. In this case, one assigns the number of the smallest neighbor to the cell p. This operation is shown in figure 6c). The cells at step t+2 take the smallest number among their neighbors at step t+1. One carries out this operation until there is no more change in the assignment of the cells labels. The Figure 6d) is the last step of the algorithm because the cells labels can't change. Thus all the cells having a same number will be gathered as being a same mobile object. In Figure 6d) the algorithm has detected two different objects, one object with cell carrying number 1 and another object with cell having number 7.

One key question in this algorithm is what is most judicious distance to consider for defining cell neighbors? Should we limit neighborhood at juxtaposed cells? Or is it necessary it to consider a wider range?

One of the problems if we consider only juxtaposed neighbors of the cell (neighborhood of size 1) is that it is possible that these neighbors aren't detected as mobile cell. As a consequence the leg or the arm can be cut of the body and the leg or the arm will be identified as an object separated from the rest of the body.

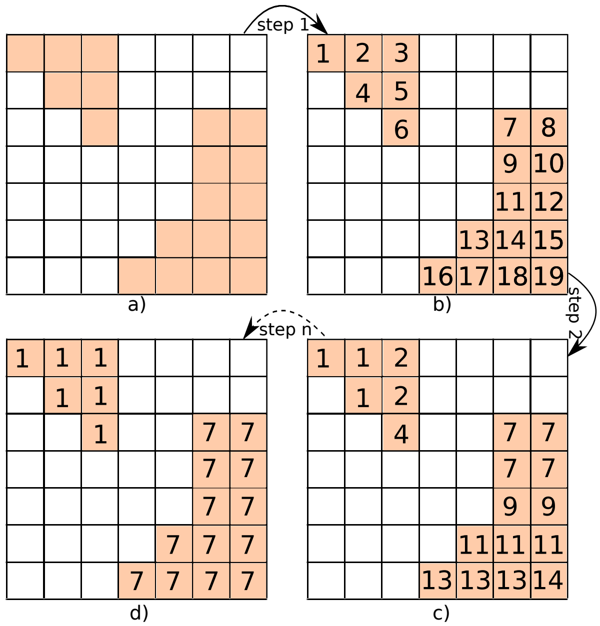


Fig. 6. a) mobile cells are colored, b) assignment of numbers to mobile cells, c) cells with number higher than it neighbor are modified, d) final result.

The second solution is to take farther cells in neighborhood, i.e. for cell C_i if we defined a neighborhood of size 2, the algorithm takes the neighbor cells positioned at C_{i-2}, C_{i-1}, C_{i+1} and C_{i+2}. The problem with this method is the risk of integrating cells which don't belong to the same object.

For illustrating these two methods we can give the following example. Figure 6 shows the case where we considered only juxtaposed neighbors cells. Two object in 6d) were detected. On the other hand if we had considered the neighbor of size 3 in Figure 6d), the two blocks of mobile cells would have gathered in only one object. In section V-B.2 we discuss this problem.

V. Results

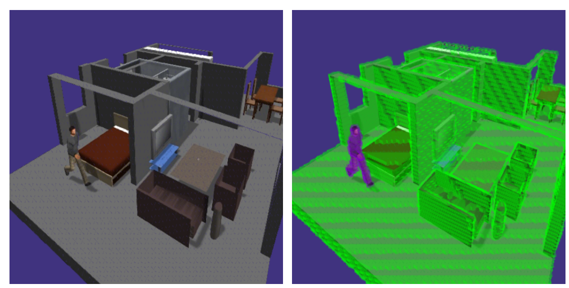
A. Simulation

This section describes our method for evaluating the sensitivity and the specificity of the system in a simulated environment. The sensitivity is the capacity to detect mobile objects when they are present and the specificity is the capacity of the system to detect the absence of mobile objects when there is no mobile object.

Figure 7 shows the output of the simulator and the result of the system classification. In order to perform the evaluation, the output of the system should be compared to a reference image pixel by pixel. Since it is impossible to evaluate the system in real conditions due to the fact that we need to index real images, we propose to limit the quantitative evaluation to a simulated environment. We have recorded a simulated human activity in a virtual scene and used these images as a reference. We simulate a Kinect by generating depth and RGB images from the virtual scene. In addition, a reference image that index each pixel in the scene as static or mobile object is also generated. RGB and depth images are supplied to our system to perform the classification. Finally we compare the output of the algorithm to the reference image. Results show a sensitivity of 87.14% and a specificity of 98.52% for a total of 430 frames (73.8M pixels). In spite of good specificity we can notice that there are as many false-positives (pixel detected as mobile whereas the pixel is static) as of true-positives (pixel detected as mobile and pixel is mobile). The problem comes from the fact that in the reference images only 2% of the pixels corresponded to the moving person whereas the other 98% were static objects or background. However visually the mobile points are always near to mobile object. The model has a some inertia and so if a pixel is mobile it takes a certain time before to return a static pixel. Table I shows the number of pixels for static and mobile objects obtained from the simulation and detected by our system.

		reference	
		pixels: mobile	pixels: static
detected	pixels: mobile	1 294 005	1 068 236
	pixels: static	190 958	71 278 387

TABLE I
Number of pixels in each category.



(a) Simulated apartment. (b) Distinction between static and mobile objects.

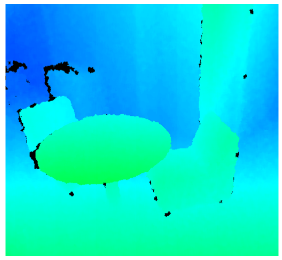
Fig. 7. Results in simulation.

B. Behavior in a smart room

We have tested our algorithm in an experimental apartment. Results presented here are qualitative. In Figure 8(a) and 8(b) we see the RGB and depth images. The image sent back by the Kinect is illustrated by Figure 8(c). Black spaces correspond to non reconstructed zones.

1) Results with one camera and one mobile object: We have tested the algorithm with one camera and one person walking in front of the camera. As illustrated by Figure 9(a), walls, furnitures and the ground are correctly detected as static objects represented by green color and the person as a mobile object represented by blue color. We can see that the feet of the person in figure are detected as a static object, it's due to the size of the voxels (6 cm) and the uncertainty of the observation. The feet of the person are integrated in voxels representing the ground. We can notice that there is very limited noise on the background where

(a) RGB image Kinect.

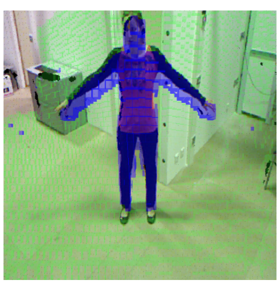


(b) Depth image Kinect.

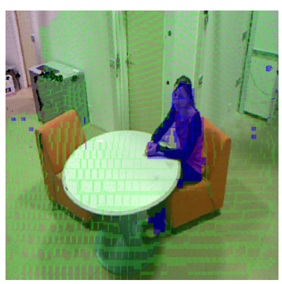
(c) 3D reconstruction of the scene using depth and RGB images.

Fig. 8. Image of Kinect camera.

a few badly positioned blue cubes remain. Moreover the tracking of mobile objects is fast enough to distinguish visually the members (leg, arm) of the person as it can be seen on Figure 9(a). A space without color is present above in the left of this figure. This is due to the size of the grid which is limited here to the perception range of the Kinect.



(a)



(b)

Fig. 9. Green color: static objects. Blue color: mobile object.

The obstacles in a room don't disturb the discrimination between mobile and static objects as shown on Figure 9(b).

When we move a furniture, this furniture, previously detected as a static object, is recognized as a mobile object. Figure 10 shows a chair becoming a mobile object. This result is allowed by the transition ($O \rightarrow E$) which models the fact that a furniture can be moved resulting in new empty space.

After a certain amount of time, it could be interesting to consider a furniture that has been moved as a new static object. This can be realized simply by adding a link ($M \rightarrow O$) with a small probability γ_2 to the transition matrix as illustrated by Figure 11.

(a) The chair is considered as a static object.

(b) The chair has been moved and is considered as a mobile object.

Fig. 10. Chair becoming a mobile object.

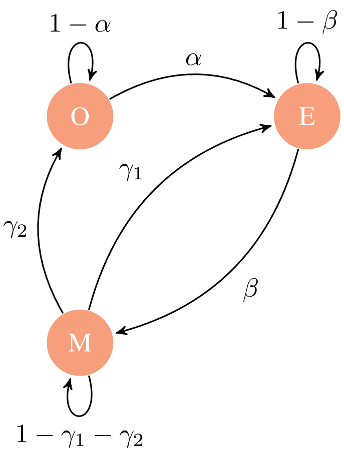


Fig. 11. Adding a link ($M \rightarrow O$) to the HMM so as to allow for a moved furniture to become a static object.

2) Results with one camera and several mobile objects: To test component labelling we realized a situation where several persons were in the field of view of the Kinect camera. The results of this test are showned in Figure 12(a). In this figure we haven't represented the occupancy grid, nor the scene but only the points representative of the center of mass of each cell of the object. Each group of points detected as one object is represented by a different color.

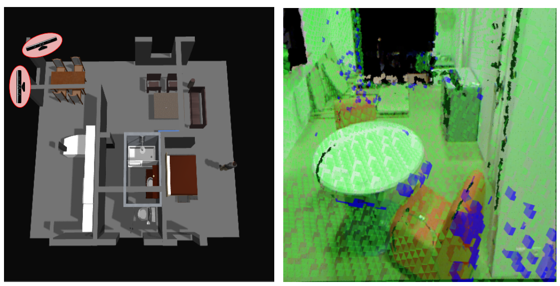
We said in section IV that to consider neighbor of size 1 is not a judicious choice because the parts of the body can be detected as separate objects. Thus to avoid this defect we increase the neighborhood to 3. But as before said the risk for taking a depth too large is to gather distinct object as shown in Figure 12(b). In this case two persons are in the room but they are too close and are considered as a single object (represented by a same color). When there is more space between the two persons as Figure 12(a), the algorithm correctly detected the two persons as separated mobile objects (as we can see because each object is in a different color).

3) Results with two cameras: To finish the experiment was then realized with two cameras placed as illustrated in Figure 13(a). We can see that the fusion of several cameras allows to discover more space while increasing the noise around static objects as illustrated in Figure 13. The noise is

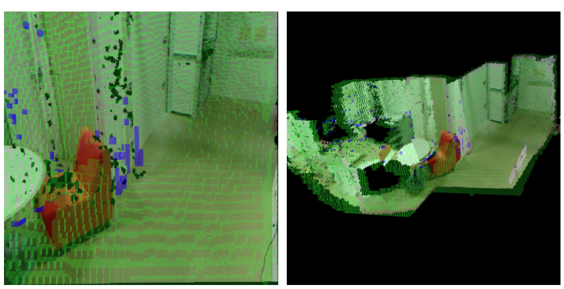
(a) Two persons are in the room and two objects are detected. (b) Two persons are in the room but only one object is detected.

Fig. 12. Results with component labelling.

due to interferences between the different infra-red images projected by the two camera Kinect.



(a) Position of the two cameras in the apartment. (b) View of one of the cameras.

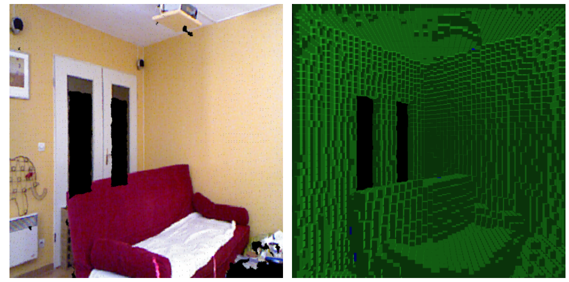


(c) View of the other cameras. (d) Fusion of the two cameras.

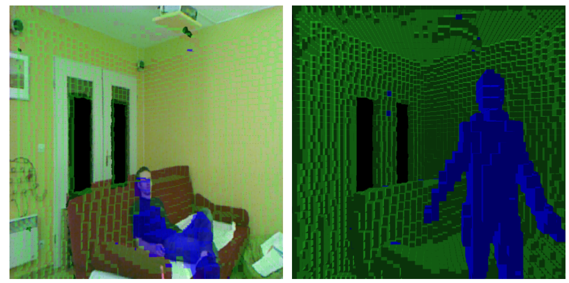
Fig. 13. Test with two cameras.

C. Behavior in realistic conditions

We have tested the algorithm, out of laboratory conditions, in a real apartment. One of the differences is on the level of lighting. The experimental apartment is located in a larger room with wall painted in black and not having windows. Thus lighting comes primarily from the artificial light. We wanted to test the algorithm in a more natural scene. We can see in Figure 14(c) that there are less noise compared to the experimental apartment. But we have noticed that when there is too much sun light on a white surface, the Kinect badly reconstructs the zone which is represented by black color in the lower right corner of Figure 14(a) which is the room where the test carryied out. We notice visually that the results are correct, the algorithm detects correctly the person.



(a) View of camera Kinect. (b) All the objects are detected as static (view without texture).



(c) A person is detected as a mobile object. (d) A person is detected as a mobile object (view without texture).

Fig. 14. The use of the algorithm in a real apartment.

VI. Conclusion

We presented a markerless system using Kinect cameras in the aim of tracking elderly people at home. First we proposed a system to merge several cameras by using a 3D occupancy grid. Secondly, compared to previous work on occupancy grid we proposed a method to allow the tracking of mobile objects. This method is based on a three states HMM: cell is empty, cell has always been occupied (static objects) and cell is occupied but has already been empty (mobile objects). This three state HMM is a simple yet elegant solution for solving a state aliasing problem (the observation for a static object is the same as the observation for a mobile object). Since each cell is updated independently one of the other, the process can be easily parallelized and implemented in a GPU allowing real-time (30 FPS) tracking with 2 cameras on a 1M cell grid. Thirdly, to solve the problem of distinguishing several persons in the field of view of the Kinect camera, we have implemented component labelling. This algorithm gather the cells belonging to the same object. Results in simulation allowed us to measure the quality of classification performed by the system in terms of sensitivity and specificity. Results on real images concerning the detection of cells occupied by mobile objects are visually satisfying. Concerning the detection of different persons the results are correct in most cases but the algorithm doesn't adress all the problems. For example, when two persons are too close the algorithm can't distinguish them. In the continuity of this work we will have to improve this part of the algorithm.

The aim of this project, as said in introduction, is to detect falls of elderly people. In this article, we presented the first part of this project. It would be necessary in continuation of this work to learn characteristics of a person (as her size...) so as to be able to recognize her, track her and detect her activity (sitting, standing...). The purpose is to learn the habits of a person for thus detecting when an unusual behavior occurs (for example lying on the ground).

VII. Acknowledgement

This work has been partly funded by Region Lorraine. The authors thanks Abdallah Dib for his contribution to this work and Cedric Rose for his valuable comments.

References

[1] H. Bourdessol and S. Pin, *Prévention des chutes chez les personnes âgées à domicile*. France: éditions Inpes, 2005.

[2] a. K. Bourke, J. V. O'Brien, and G. M. Lyons, "Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm." *Gait & posture*, vol. 26, no. 2, pp. 194–9, July 2007. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/17101272

[3] B. Jansen, F. Temmermans, and R. Deklerck, "3D human pose recognition for home monitoring of elderly," in *Proceedings of the 29th IEEE EMBS annual international conference*, August 2007.

[4] A. Dubois, A. Dib, and F. Charpillet, "Using hmms for discriminating mobile from static objects in a 3D occupancy grid," in *Proceedings of the International Conference on Tools with Artificial Intelligence (ICTAI'11)*, 2011.

[5] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, June 1989.

[6] T. Yapo, C. Steward, and R. Radke, "A probabilistic representation of LiDAR range data forefficient 3D object detection," in *Proceedings of the S3D (Search in 3D) Workshop, in conjunction with IEEE CVPR*, June 2008.

[7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. New York, NY, USA: Cambridge University Press, 2003.

[8] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," in *Proceedings of the IEEE*, 1989, pp. 257–286.

[9] K. Suzuki, I. Horiba, and N. Sugie, "Linear-time connected-component labeling based on sequential local operations," *Computer Vision and Image Understanding*, vol. 89, no. 1, pp. 1–23, Jan. 2003. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1077314202000309

Autonomous Shopping Cart Platform for People with Mobility Impairments

Luca Marchetti[1], Daniele Pucci[1], Pascal Morin[2]

***Abstract*— Providing a platform able to interact with a specific user is a challenging problem for assistance technologies. Among the many platforms accomplishing this task, we address the problem of designing an autonomous shopping cart. We assume that the shopping cart is set-up on a unicycle-like robot endowed with two sensors: an RGB-D camera and a planar laser range finder. To combine the information from these two sensors, a data fusion algorithm has been developed using a particle filter, augmented with a k-clustering step to extract person estimations. The problem of stabilizing the robot's position at a fixed distance from the user has been solved through classical control design. Results on a real mobile platform verify the effectiveness of the approach here proposed.**

I. Introduction

Assistive technologies focus their efforts on providing reliable solutions to help people in the everyday life. One of the key components of an assistive system is the ability to actively follow a user, a task well exemplified by a mobile robot that follows the user. An *autonomous shopping cart* is a simple application that provides a good test-bed for a whole class of problem: a robotic butler that helps on carrying heavy objects; a robotic lift that has to follow a companion to accomplish a coordinated task; an automatic walking aid that should support elderly people and so on.

Platform of such kind should be able to detect and recognize the user, among other people, and be capable of following continuously the same user. The environment should be modelled in such a way the robot can avoid obstacles, and pursue the user at the same time. We focused our attention on developing methodologies to accomplish a safe following of user's trajectory, while maintaining a certain degree of freedom on the reference position of the robot w.r.t. the user.

The involved scientific challenges can be summarised in two aspects. On one hand, a module must be developed in order to estimate the user position while identifying other people in the environment. The selection of the user should be effective in such a way the continuous following of the user will not be confused by the presence of other subjects. This estimation must achieved in a cluttered and noisy environment. On the other hand, one has to provide a reliable and feasible way to control the mobile platform, respecting the peculiarity of human motion and exploiting the robotic aid.

[1]INRIA Sophia Antipolis Meditérranée, 2004, Route des Lucioles - BP 93 Sophia Antipolis Cedex 06902 <name>.<surname>@inria.fr
[2]ISIR-UPMC, 4, place Jussieu 75005 Paris <surname>@isir.upmc.fr

Combining these two challenges represents the main contribution of our work. *People Position Estimation* is a well-known problem addressed in many scenarios, e.g. video-surveillance [1] or activity recognition [2]. In the field of object recognition, the human body represents probably the most challenging one. The complexity of shape, as well as the multiplicity of configuration it can achieve require complex sensors to be captured. Color cameras are amongst the most effective sensors, even if the information are limited to the image plane [3]. Recently the evolution of technology and availability of relatively cheap RGB-D sensors, capable of perceiving 3D structures, opened the possibility to extend the range image-based recognition [4]. It is then reasonable to choose such sensors to capture people positions[5].

These devices, however, usually have a limited field-of-view. While it is completely reasonable to use multiple camera to augment the virtual field-of-view [6], other aspects of the application guided us on choosing a different solution to cope with this problem. In the case of a robot following a person, it is important to take care of obstacles that can limit the motion of the robot. To address this problem, usually a *laser range finder* is employed to map the surrounding [7]. Its high precision on a 2D plane is an effective way to detect obstacles for the robotic motion. Moreover, the laser information can also be used to detect and track the legs of several people [8]. Thanks to the large field of view and range, and the sensor resolution, the user position can be estimated with high accuracy.

By exploiting the strengths of both sensors, an improved position estimation can be obtained. The result is an accurate person position estimation that can be given as input to a person following control module. While the human being is able to move along any direction, a wheeled robotic platform is usually subjected to kinematic constraints that limit the range of feasible trajectories [9]. Hence, to achieve human following by the robot, additional maneuvers may be necessary when the human trajectory does not satisfy the aforementioned constraints. This problem has been largely studied in the last decades and one can rely on existing techniques to address the control problem [9].

In this paper, we describe how a data fusion algorithm, that combines information from 2D and 3D sensors, can be effectively coupled with a trajectory stabilization method, able to drive the robot to solve the person following challenge.

II. Notation

We consider the class of unicycle-like robots sketched in Figure 1. The following notation is used. Let $\mathcal{I} = \{O; \vec{\imath}_0, \vec{\jmath}_0\}$

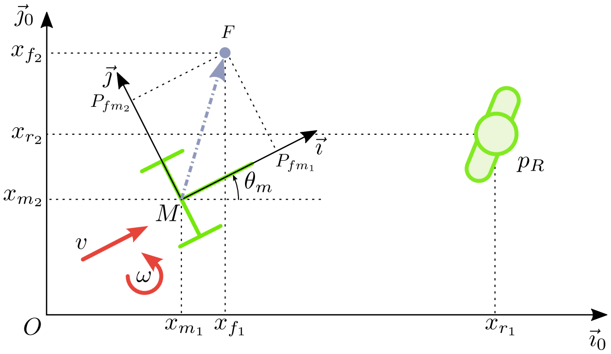


Fig. 1. Unicycle-like robotic platform for autonomous shopping cart.

be a fixed inertial frame with respect to (w.r.t.) which the robot's absolute pose is measured. The point M is the middle point of the wheel's axis, and $\mathcal{B} = \{M; \vec{\imath}, \vec{\jmath}\}$ is a frame attached to the robot. The vector $\vec{\imath}$ is perpendicular to the wheel's axis. The vector of coordinates of M in the basis of the fixed frame $\mathcal{I}$ is denoted as $x_m = (x_{m_1}, x_{m_2})^T$. Therefore, $\vec{OM} = x_{m_1}\vec{\imath}_0 + x_{m_2}\vec{\jmath}_0$. The robot's orientation is characterized by the angle θ_m between $\vec{\imath}_0$ and $\vec{\imath}$. The rotation matrix of an angle θ_m in the plane is $R(\theta_m)$. With $\{e_1, e_2\}$ we denote the canonical basis in $\mathbb{R}^2$. In view of this notation, the kinematic model of the robot writes [9]

$$\begin{aligned} \dot{x}_m &= vR(\theta_m)e_1 \\ \dot{\theta}_m &= \omega. \end{aligned} \tag{1}$$

with v the robot's rolling velocity and ω its rotational velocity considered as kinematic control inputs. The position of the user is represented by a *reference point* p_R. The vector of coordinates of p_R in the basis of the fixed frame $\mathcal{I}$ is denoted as $x_r = (x_{r_1}, x_{r_2})^T$. Therefore, $\vec{Op_R} = x_{r_1}\vec{\imath}_0 + x_{r_2}\vec{\jmath}_0$. The vector of coordinates associated with the linear velocity of p_R w.r.t. $\mathcal{I}$ is denoted as $\dot{x}_r$.

III. People Position Estimation

People detection is one of the main component needed to have a reliable autonomous shopping cart. In Section I, we highlighted the advantages of using two different sensors to achieve the position estimation. Next, we describe the data fusion architecture pointing out the main characteristics. The functional blocks are described in Figure 2.

The architecture is two-tiered. The lowest level has a person estimation method for each sensors. The highest level combines the two estimations to obtain a more reliable position estimation.

a) Laser-based position estimation: To detect people using the laser range finder, we use an implementation of the *Kalman Filter leg tracker*, described by Arras et al in [8]. Therefore, we follow the notation presented in that paper to briefly describe the functioning of this estimator.

The *KF-based multi-hypotheses tracker* describes a leg track as $p_L^i = (p_x, p_y, v_x^p, v_y^p)$, with p_x and p_y position on the laser plane, and v_x^p and v_y^p the components of the

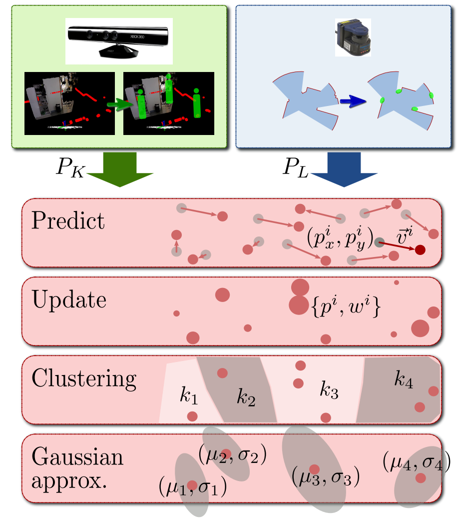


Fig. 2. The data fusion architecture for person position estimation.

velocity. The *state prediction* of the leg filter uses a constant-velocity model. The *observations* of legs are detected using an Adaboost algorithm [10] that classifies the segments found in the laser scan according to a set of features. Using these features, several weak classifiers are used to separate leg candidates. The combination of all the classifiers generates the leg observations. The training procedure and how to obtain a valid set of informative threshold to classify the segments is explained in [11]. The tracker labels assign each new measurement to existing leg tracks or creates new tracks. At any instant, a leg track can be *detected* (if measurements are assigned during the last observation phase) or *deleted* (if measurements are not assigned). New tracks are labeled as *new track* and a *false alarm* label is used when the measurements are mistakenly detected as track.

People tracks are extracted from leg tracks using the following heuristic:

- people have two legs;
- legs are close to each other;
- legs move in similar direction;
- legs have a higher probability of occluding each other, than being occluded by other people's legs or objects.

This model is implemented in such a way it takes into account the possible occlusions, thus avoiding deletion of track if legs are occluded for a short period of time. Other considerations about leg track labeling and probability association are described in details in [8] and will be omitted here.

b) Kinect-based person estimation: The availability of a ROS-integrated library for user skeleton detection simplifies the problem of detecting human shapes using the RGB-D data from the Kinect sensor. As for OpenNI library version 1.3, the user needed to perform a peculiar calibration procedure (the so-called ψ pose) to be detected. This limited the usability of the bundle software for multiple people detection. As for version 1.5, however, the users can be detected using a *User Generator* that does not require any calibration at all. The output of this module is a set of people position estimation P_K and can be extremely noisy, because the people's bodies are recognized applying statistical methods. This required to develop appropriate filtering procedure to establish correct person following.

c) Data-fusion for person position estimation: We decided then to use both information, from the laser and the RGB-D camera, to provide more reliable information to the control layer. The camera can be really accurate on estimating the complexity of human body, while the laser provides a larger field of view and a better precision on distance estimation.

To combine the advantages of both sensors, we designed a *particle filter* with *clustering*. Each particle represents a possible position estimation as $p^i = (p_x, p_y, v_x^p, v_y^p)$. The posterior density is approximated by:

$$\mathcal{P}(X|z) \approx \sum_{i=1}^{|P|} w^i \delta(X - x^i), \quad (2)$$

where the X is the current state of the probability density function and w^i is a weight associated to the sample $x^i \in X$. The interested reader can find more explanation about this representation in [12].

The algorithm is described below.

Algorithm 1: People Position Estimation

;
Data: $P_L := \{\text{position estimation from laser}\}$;
$P_K := \{\text{position estimation from Kinect}\}$;
$P := \{\text{particle set}\}$
for $p^i \in P$ **do**
  draw particles: $\widetilde{p}_t^i \sim \pi_t(p^i | P, \vec{v}_{t-1}^i)$;
  calculate weight: $\widetilde{w}_t^i \propto \mathcal{L}_L(\widetilde{p}_t^i) \times \mathcal{L}_K(\widetilde{p}_t^i)$;
resample: $\{p_t^i, w_t^i\}_{i=1}^{|P|} = resample(\{\widetilde{p}_t^i, \widetilde{w}_t^i\}_{i=1}^{|P|})$;
get position estimation clusters: $\mathcal{C} = KClusterise(P)$;

At time t, the particle filter algorithm requires a proposal distribution (π_t) from which it draws samples during the *prediction step*. We use the previous set of particles evolved using a constant velocity model. The *update step* uses information from the laser and the RGB-D sensors. A sensor model calculate the likelihood of each particle to belongs to the set of laser measurements P_L or camera measurements P_K. The likelihood is evaluated as:

$$\mathcal{L}_L = \frac{b(P_L|p_t^i) b(\widetilde{p}_t^i | p_{t-1}^i)}{\pi(\widetilde{p}_t^i | \widetilde{P}, \vec{v}_{t-1}^i)}, \quad (3)$$

$$\mathcal{L}_K = \frac{b(P_K|p_t^i) b(\widetilde{p}_t^i | p_{t-1}^i)}{\pi(\widetilde{p}_t^i | \widetilde{P}, \vec{v}_{t-1}^i)}, \quad (4)$$

where $b(\cdot)$ is the pdf of the system state. The estimated posterior represents the distribution of people over the sensor's space. This posterior is usually multimodal, given the noisy nature of the RGB-D camera and ambiguity on laser estimation. Therefore, a clustering phase is necessary to extract all the possible tracks. A track, or a person estimation, will be the Gaussian approximation (mean and variance) of a single cluster.

A selection procedure, not described here, selects the best candidate and assigns it to the control module.

A. K-Clustering

We implemented a *k-clustering* based technique[13], described in Algorithm 2. *KClusterise* tries to detect up to N_k clusters. Therefore, it is not a free-cluster algorithm, and this could potentially lead to a limitation. However, for the purpose of the presented applications, this is not a critical problem.

Algorithm 2: KClusterise

Data: P: particle set;
K: cluster set, of maximum size N_K;
O: outliers set
Initialise cluster set: $K \leftarrow \varnothing$;
Find cluster: $K = FindCluster(P)$;
Assign particle to cluster:
$O = ClassifyParticles(K, P)$;
Redistribute outliers: $SpreadOutliers(O, K)$;

Algorithm 3: FindCluster

```
// Find equally spaced out centroids
```
1 **for** $p \in P$ **do**
2   *isFar* = true;
3   **for** $k \in K$ **and** *isFar* **do**
4   *isFar* = $(\|p, k\| > \delta_{far})$;
5   **if** *isFar* = *true* **then**
6   Add particle p as centroid: $K \leftarrow p$
7 **return** K

First, the algorithm tries to find N_k points (with $N_k \geq 1$) that are equally spaced out (Algorithm 3). Adding particles in line 6 is done using a priority queue principle, considering the distance. At the end of procedure, K will contain the most distant N_K points, and they will be used as centroids.

Successively (Algorithm 4), for each point p, it calculates the Euclidean distance δ between p and clusters $k \in K$. Let δ_{min} be the minimum distance between p and a cluster k.

Two cases are possible: if δ_{min} is less than the *threshold distance* δ_{far} (discussed in Section III-B), p will be put in cluster k. Otherwise, p will be put into the *outliers set O*.

When the clusterisation phase is finished, the points in the outliers set O are evaluated. Each point in O will be put in the nearest cluster by ignoring the threshold distance.

B. Threshold Distance Function

One of the major problems in clustering techniques is to find a threshold distance to approximate the correct number of clusters. There are two possible ways: a fixed value

Algorithm 4: ClassifyParticles

// Classify particles
for $p \in P$ **do**
  added = false;
  for $k \in K$ **and not** *added* **do**
    $\delta_{min} = \|p, k\|$;
    if $\delta_{min} < \delta_{far}$ **then**
      Add particle p to cluster k: $k \leftarrow p$;
      added = true;
  if *added* = *false* **then**
    Add particle p to outliers set: $O \leftarrow p$
return O

Algorithm 5: SpreadOutliers

// Redistribute outlier particles
for $p \in O$ **do**
  $\delta_{min} = \infty$
  for $k \in K$ **do**
    $\delta = \|p, k\|$;
    if $\delta < \delta_{min}$ **then**
      $k_{candidate} = k$;
      $\delta_{min} = \delta$
  Classify particle p: $k_{candidate} \leftarrow p$

or a variable one. The first choice can be computationally efficient, but it is not flexible w.r.t. environment's changes.

We adopted the second strategy, by using a *dynamic threshold function*, that represents an approximation of the Mahalanobis distance. For each position estimation in P_L, P_K, we evaluate the average distance to other estimations as:

$$\delta_{far} = \frac{\sum_{i=1,j=2}^{|P_L|} \|p_L^i, p_L^j\| + \sum_{i=1,j=2}^{|P_K|} \|p_K^i, p_K^j\|}{|P_L| + |P_K|}, \quad i \neq j. \quad (5)$$

The idea behind this function is to keep the position estimation as far as possible to each other. Using the average of pairwise distance helps us to obtain well-balanced cluster while keeping them separated.

IV. Control Design

Achieving a reliable person following requires correct control laws to smoothly let the robot follow a trajectory constrained by the person's motion. Our objective is to describe the desired position of the person w.r.t. the robot, then minimize the distance between this desired position and the actual position. From Figure 1 the position of the desired (or *follower*) point F is given by $x_f = x_m + R(\theta_m)P_{fm}$, where P_{fm} is the vector of coordinates of $\vec{MF} = \vec{OF} - \vec{OM}$. Recall that the position of the person, is given by p_R. This is the results of person position estimation described in Section III.

Let $\tilde{x}$ be the position error in the fixed frame, defined by:

$$\tilde{x} \quad = \quad x_f - x_r, \quad (6)$$

and w.r.t. the mobile frame by:

$$\tilde{p} \quad = \quad R(\theta_m)^T \tilde{x}. \quad (7)$$

Note that here $x_r = (p_x, p_y)$. Therefore the error dynamics w.r.t. the mobile frame writes:

$$\dot{\tilde{p}} \quad = \quad -\omega S \tilde{p} + M u - R(\theta_m)^T \dot{x}_r(t), \quad (8)$$

with $M = \begin{bmatrix} 1 & -P_{fm_2} \\ 0 & P_{fm_1} \end{bmatrix}$, $S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $u = (v, \omega)^T$ and $\dot{x}_r = (v_x^p, v_y^p)^T$. Relying on the results in [9], one deduces the following lemma.

Lemma 1: Assume that $P_{fm_1} \neq 0$, so that $det(M) \neq 0$. Apply the control input

$$u \quad = \quad -M^{-1}\left[K\tilde{p} - R(\theta_m)^T \dot{x}_r(t)\right], \quad (9)$$

where $K = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$, $K > 0$. Then,

$$\dot{\tilde{p}} = -\omega S \tilde{p} - K \tilde{p}, \quad (10)$$

so that $\tilde{P} = 0$ is a globally asymptotically stable equilibrium point for the closed-loop system.

The stability analysis follows by verifying that $V(\tilde{p}) = \|\tilde{p}\|^2$ is a strict Lyapunov function for the closed-loop system. When $P_{fm_1} = 0$, the control law (9) is not well defined and other solutions to the tracking problem must be considered. This problem will be addressed in a forthcoming paper.

V. Results

The first implementations of the mobile platform have been conducted on an industrial wheeled robot. The robot (cfr. Figure 3) has a Kinect and the OpenNI[1] library is used to detect and estimate the raw positions of *all* people in the field-of-view of the camera. A laser is mounted on the front side of the robot and captures scans of the environment. This scans are used to estimate the odometry of the robot (using the *Canonical Scan Matcher*[2]). The onboard computer drives the platform and provides odometry from wheels. The attached laptop, instead, processes data from Kinect and laser, to model the pose of people and evaluate control inputs.

The software architecture has been developed using the ROS[3] framework. The actual architecture, illustrated in Figure 4 is composed of three layers: the *robot interface*, the *modeling* core, and the *behaviour* component.

The *robot interface* provides the abstraction layer between the actual platform and the software components. It is easy to adapt this interface for several platforms and keep unchanged the higher levels.

[1] http://www.openni.org
[2] http://ros.org/wiki/csm
[3] http://www.ros.org

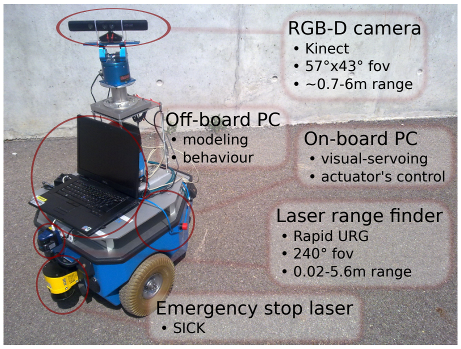


Fig. 3. The robot used to run the experiments.

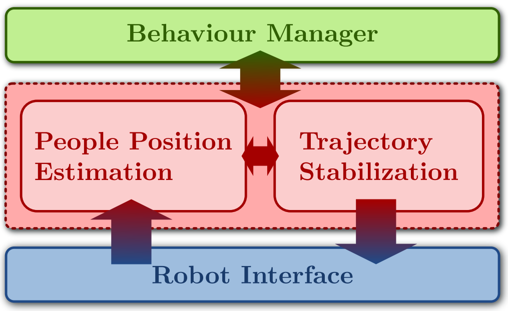


Fig. 4. The software architecture and modules developed for the autonomous shopping cart robot.

The *modeling* layer detects the user within the sights of the sensors and evaluates the control inputs. The user position estimation is evaluated as described in Section III, while the control inputs are the results of control laws developed in Section IV.

The last layer, the *behaviour* component, handles the safety measures to ensure the robots avoid close obstacles (detected by the laser), starts and stops services on request, enables initialization procedures and so on. In particular, it selects the first user to be followed, among possible candidates.

A. Experimental Results

In this section, we focus on the practical experiments obtained using the robot. We present two different configuration, considering the behaviour and estimation performed by the robot when the follower point x_f was placed in two different positions: $A = (1.5, 0)$ and $B = (1, 1)$. The positions are depicted in Figure V-A. Supplement material and high quality versions of the picture presented here can be found at `http://goo.gl/NFnjg`.

1) Test A: user in front of the robot: Figure V-A.2 presents the outcome of the data fusion procedure. The data-fusion trajectory represents the trajectory of the person passed to the control module. Despite the high number of hypotheses, the data fusion algorithm is able to consistently track the movement of the person. Since gathering ground truth of the person's position was not a straightforward procedure, the resulting trajectory from data fusion is given according

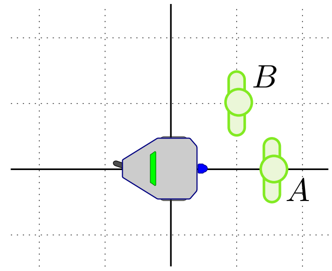


Fig. 5. The position of the follower point x_f used during tests.

the position of the robot in the environment. In Figure V-A.2 is presented the trajectory of the following point w.r.t. the person estimation: the control inputs are consistent with the estimation. The movement of the robot resulted much more smoothed than the estimation. This effect is due to the different frequency at which the control commands are sent to the motor (much lower than the estimation rate).

2) Test B: user at 45° in front of the robot: In Figure 6 are presented the same results as for test A, showing the effectiveness of the data fusion algorithm. In Figure 6 are presented the trajectory of the following point while at 45° w.r.t. the robot. The results are consistent with the test A.

VI. Conclusions

This article presented a reliable solution to the problem of person following. A data fusion algorithm has been presented to reliably detect and estimate the position of multiple people. Two kinds of sensors have been exploited to accomplish the position estimation: a laser range finder and a RGB-D camera. Based on this estimation, a feedback control law has been designed to track the user. To demonstrate the effectiveness of the proposed architecture, results have been presented using a real mobile platform. As future work, we are investigating how to address the singularity on the control design. A different control law needs to be developed to avoid the singularity, in order to have a completely arbitrary position for the following point x_f.

Acknowledgements

We would thank the Large Scale Initiative Action PAL (*Personally Assisted Living*), that funded this ongoing research, and the people working at INRIA Sophia Antipolis Meditérranée that help on pursuing this work.

References

[1] L. M. Fuentes and S. A. Velastin, "People tracking in surveillance applications," in *IEEE International workshop on PETS*, 2001.
[2] R. Bodor, B. Jackson, N. Papanikolopoulos, and H. Tracking, "Vision-based human tracking and activity recognition," in *Mediterranean Conference on Control and Automation*, 2003, pp. 18–20.
[3] S. J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler, "Tracking groups of people." *Computer Vision and Image Understanding*, vol. 80, no. 1, pp. 42–56, 2000.
[4] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2011.

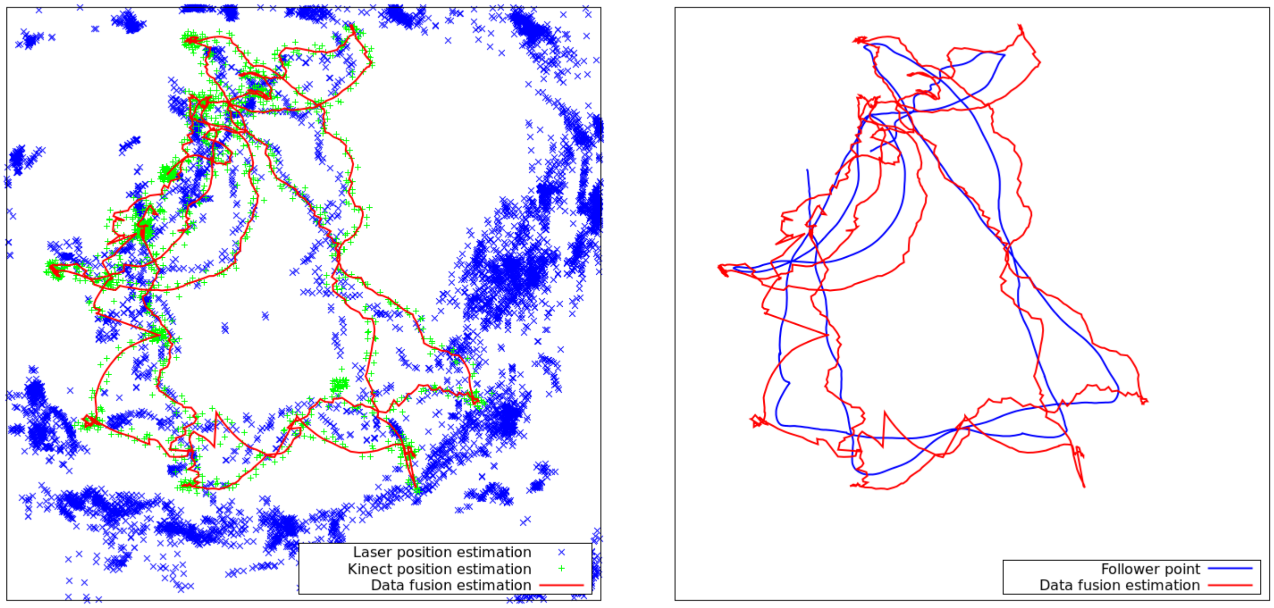


Fig. 6. Results for the data fusion algorithm during test A: $x_f = (1.5, 0)$.

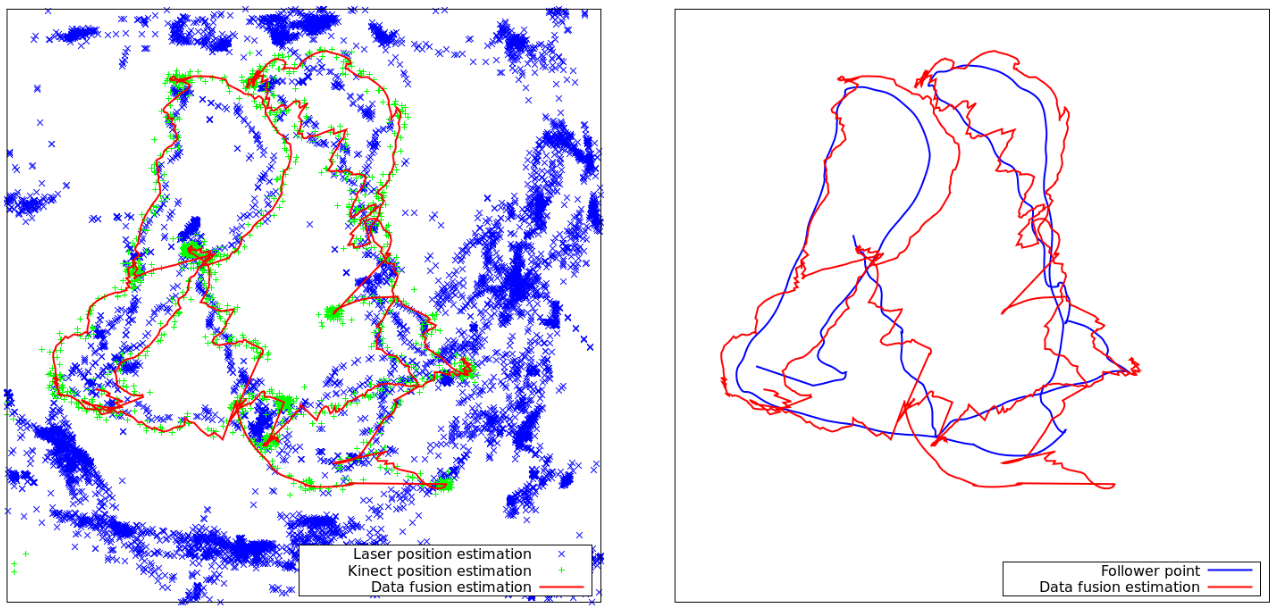


Fig. 7. Results for the data fusion algorithm during test B: $x_f = (1, 1)$.

[5] L. Spinello and K. O. Arras, "People detection in RGB-D data," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011.

[6] M. Luber, L. Spinello, and K. O. Arras, "People tracking in rgb-d data with on-line boosted target models." in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011.

[7] A. Censi, "An ICP variant using a point-to-line metric," in *IEEE International Conference on Robotics and Automation*, May 2008.

[8] K. O. Arras, S. Grzonka, M. Luber, and W. Burgard, "Efficient people tracking in laser range data using a multi-hypothesis leg-tracker with adaptive occlusion probabilities," in *IEEE International Conference on Robotics and Automation*, 2008.

[9] P. Morin and C. Samson, "Motion control of wheeled mobile robots," in *Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer, 2008, pp. 799–826.

[10] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *European Conference on Computational Learning Theory*. Springer-Verlag, 1995, pp. 23–37.

[11] K. O. Arras, scar Martnez, and M. W. Burgard, "Using boosted features for detection of people in 2d range scans," in *IEEE International Conference on Robotics and Automation*, 2007.

[12] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, 2001.

[13] A. K. Jain and R. C. Dubes, "Algorithms for Clustering Data," in *Prentice Hall, Englewood Cliffs New Jersey*, 1988.

Session II

Navigation and manipulation in human populated environments

- **Keynote speaker : Rachid Alami (Senior scientist at the Robotics and Artificial Intelligence Group from LAAS, CNRS, France)**
  Title : Equipping a robot with Human-aware decisional abilities.

- Title : **Intention Driven Human Aware Navigation for Assisted Mobility.**
  Authors : J. Rios-Martinez, A. Escobedo, A. Spalanzani, C. Laugier

- Title : **Robot Navigation Taking Advantage of Moving Agents.**
  Authors : P. Stein, A. Spalanzani, V. Santos and C. Laugier

- **Keynote speaker : Matei Ciocarlie (Research Scientist and Area Lead at Willow Garage)**
  Title : Mobile Manipulation Through An Assistive Home Robot

Keynote Speaker : **Rachid Alami**
(Senior scientist at the Robotics and Artificial Intelligence Group from LAAS, CNRS, France)

Equipping a robot with Human-aware decisional abilities

Abstract : This talk addresses some key decisional issues that are necessary for a robot which shares space and tasks with a human. We adopt a constructive approach based on the identification and the effective implementation of robot collaborative skills. These abilities include geometric reasoning and situation assessment based essentially on perspective-taking and affordances, management and exploitation of each agent (human and robot) knowledge in a separate cognitive model, human-aware task planning and human and robot interleaved plan achievement.

Keynote Speaker : **Matei Ciocarlie**
(Research Scientist and Area Lead at Willow Garage)

Mobile Manipulation Through An Assistive Home Robot

Abstract : We present a mobile manipulation platform operated by a motor-impaired person using input from a headtracker, single-button mouse. The platform is used to perform varied and unscripted manipulation tasks in a real home, combining navigation, perception and manipulation. The operator can make use of a wide range of interaction methods and tools, from direct tele-operation of the gripper or mobile base to autonomous sub-modules performing collision-free base navigation or arm motion planning. We describe the complete set of tools that enable the execution of complex tasks, and share the lessons learned from testing them in a real users home. In the context of grasping, we show how the use of autonomous sub-modules improves performance in complex, cluttered environments, and compare the results to those obtained by novice, able-bodied users operating the same system.

Mobile Manipulation Through An Assistive Home Robot

Matei Ciocarlie[1], Kaijen Hsiao[1], Adam Leeper[2] and David Gossow[3]

***Abstract*— We present a mobile manipulation platform operated by a motor-impaired person using input from a head-tracker, single-button mouse. The platform is used to perform varied and unscripted manipulation tasks in a real home, combining navigation, perception and manipulation. The operator can make use of a wide range of interaction methods and tools, from direct tele-operation of the gripper or mobile base to autonomous sub-modules performing collision-free base navigation or arm motion planning. We describe the complete set of tools that enable the execution of complex tasks, and share the lessons learned from testing them in a real user's home. In the context of grasping, we show how the use of autonomous sub-modules improves performance in complex, cluttered environments, and compare the results to those obtained by novice, able-bodied users operating the same system.**

I. Introduction

Independence, and a sense of control and freedom, are some of the key factors positively correlated with life satisfaction and health (both psychological and physical) for older adults and people with motor impairments. Confidence in the ability to undertake various tasks is core to one's psychological functioning [9], and a greater sense of control over life is positively correlated with better health [21] and a reduced mortality rate [22].

The vision for this study is that of a mobile manipulation platform sharing a living environment and operating side-by-side with its user, increasing independence and facilitating activities of daily living. In particular, we focus on the set of Instrumental Activities of Daily Living (IADLs) [23], which require manipulating the environment (e.g., performing housework) away from the user's body. A mobile robot could provide assistance with a variety of IADLs, operating in a large workspace without encumbering the user.

A key step for achieving this vision is enabling mobile robots to handle the complexity and variability inherent in real living environments. Despite impressive advances over the past decades, these factors have so far prevented versatile manipulators from achieving the level of reliability needed for long term deployment in real human settings.

We posit that the difficulties associated with the design of fully-autonomous systems could be mitigated by involving the care receiver in the loop, as a user and operator of the robot. The user's cognitive abilities can be tapped to deal with conditions that have proven difficult for autonomous systems to deal with. Autonomy would still play a vital role in this shared framework, but the human operator would provide the information that robots are incapable of deriving themselves. As additional autonomous components mature, they can be incorporated into the system to help reduce the load on the operator. Such a system could be reliable and robust enough for deployment in the near future.

Fig. 1. A mobile robot, operated by a motor-impaired user, performing a manipulation task in a real home environment.

With this directional goal in mind, we have developed a system that enables a motor-impaired user to command a robot in performing manipulation tasks in complex environments. The operator can make use of a wide range of interaction methods and tools, from direct teleoperation of the gripper or mobile base to autonomous sub-modules performing collision-free base navigation or arm motion planning. The operator receives feedback from the robot via a computer screen and provides input using only a mouse. In particular, this enables our system to be used by operators with severe motor impairments who can nonetheless still control a mouse cursor via a head tracking device.

This paper's contributions are as follows. We introduce what is, to the best of our knowledge, the first example of a mobile manipulation platform operated by a motor impaired person using only a head-tracker single-button mouse as an input device, and demonstrated for both varied and unscripted manipulation tasks in a real home and limited

*This work was partially supported by the NSF under Award IIP-1142743. A. Leeper is partially supported by an NSF GRFP Fellowship.
[1]Willow Garage Inc., Menlo Park, CA. Email: {matei,hsiao}@willowgarage.com
[2]Stanford University, Stanford, CA. aleeper@stanford.edu
[3]Technical Univ. of Munich, Germany. dgossow@cs.tum.edu

forms of social interaction. We describe the set of tools that enable the execution of complex tasks, their interplay, and their effect on the overall system. We share the lessons learned from deploying a real robot in a real user's home and attempting to manipulate the environment. In the context of grasping, we show how the use of autonomous sub-modules improves performance in complex, cluttered environments, and compare the results to those obtained by novice, able-bodied users operating the same system.

A variety of assistive robot systems have been tested in the past, including desktop workstations with robotic arms, wheelchair-mounted robotic arms, powered orthotic and prosthetic arms, and mobile manipulators[4], [7]. An example of an assistive robot tested for remote manipulation in real homes is shown in [15]. Recent assistive mobile manipulation systems with a number of important capabilities include Care-o-bot 2[6], SAM[16], and El-E[10]. In addition to showing and quantifying individual capabilities such as object grasping, in this paper we demonstrate operation in a real home for a complete mobile manipulation task, as well as limited social interaction through manipulation.

Despite previous advances in assistive mobile manipulators, none has been widely adopted to date. Part of the reason is their cost, as many of these robots, including the PR2 which we use here, are not suitable for widespread commercial adoption. Nonetheless, we believe that creating and demonstrating a system such as the one presented here, with a comprehensive suite of tools with varying levels of autonomous assistance for perception, navigation, and manipulation, can enable future flexible and competent platforms with widespread adoption

II. System Overview

Our system has two main components: the robot itself, a two-armed mobile manipulation platform, and the interface used by the operator to command and receive feedback from the robot. The interface can run on a commodity desktop or laptop computer. The system was designed for remote operation, with the goal of enabling the operator to perform tasks through the robot even from another room, or in other situations when there is no direct visual or audio contact.

A. Hardware Platform

The hardware we used was the PR2 personal robot [24]. The PR2 has two compliant, backdriveable 7-DOF arms with parallel-jaw grippers. We used two range sensors: a widely available Microsoft Kinect^{TM} mounted on the head of the robot (providing both range and color images), and a tilting laser rangefinder mounted on the chest (used for autonomous collision avoidance). The PR2 can communicate with the computer running the teleoperation interface via a commodity wireless network; we expect that a mobile robot in real households will have to be untethered to perform a large number of tasks. The PR2's form factor was designed for enabling operation in typical human environments. With arms folded in, it has a similar footprint (square, 668mm on each side) to a wheelchair, and can thus navigate in ADA-compliant spaces. It has a telescopic spine, allowing it to touch the floor with its arms, operate low cabinets, and manipulate objects on table- and counter-tops.

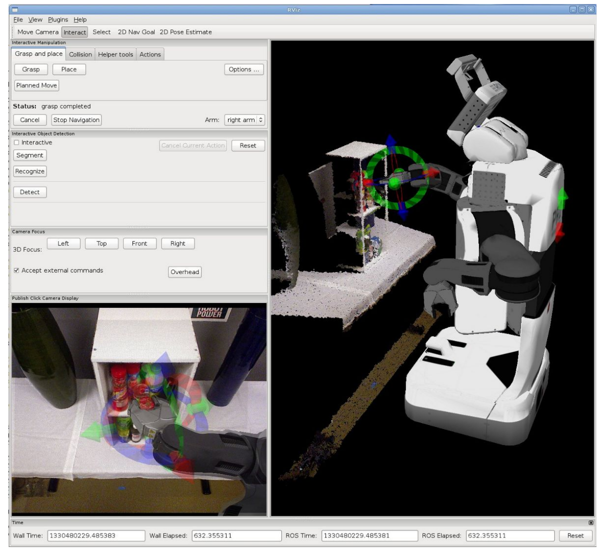


Fig. 2. Overview of user interface.

B. User Interface

We developed a "point-and-click" Graphical User Interface based on `rviz` [20], a 3D robot visualization and interaction environment that is part of the Robot Operating System (ROS) [14]. The choice of a point-and-click interface was motivated by accessibility: while higher-dimensionality input methods such as trackballs and haptic devices would provide additional benefits for teleoperation, a major advantage of a simple cursor-driven interface is the widespread accessibility of devices that provide cursor control, including, for example, head trackers for motion-impaired users.

Our user interface is shown in Fig. 2. It presents the user with two main displays: on the bottom left, a live image from the Kinect camera mounted atop the PR2's head; on the right, a rendered image of the PR2 in its current posture (using joint encoders for proprioception), along with a 3D point cloud showing the world as seen by the Kinect.

The user can point the robot's head by left-clicking anywhere in the camera view, which centers the robot's camera on the clicked point. Because the right image is only a rendering, its virtual camera can be dragged to any position by rotating, translating, and zooming the scene in and out, which is useful for seeing and understanding the 3D scene by viewing it in motion and from multiple angles.

We have found this dual visual feedback, one from a real 2D camera and one from an artificial rendering, to be highly useful for remote operation. The real image is easy to interpret for human operators, but the limited ability to change the viewpoint makes it difficult to rely on exclusively for tasks requiring depth perception. Conversely, the rendered image can be seen from any viewpoint, but must rely on 3D sensors for data acquisition, and requires the user to be able to manipulate a virtual camera around a rendered scene.

Our user interface allows the user to send commands to the robot in two main ways. The first one is through 3D widgets that are added to the rendered 3D world that the user can click on, drag, translate, rotate, etc. These are implemented using the `interactive_markers` ROS framework [5]. The second interaction type is through more conventional dialog windows. Fig. 2 shows examples of both interaction methods, with a 3D clickable control on the robot's gripper, and a set of dialogs in the upper left part of the window.

III. Shared Autonomy Tools for Mobile Manipulation

When an assistive robot is enabling a motor-impaired user to physically interact with the world, the human operator can use her cognitive skills to handle complex and unexpected situations, as long as she can receive relevant feedback from and send appropriate commands to the robot. However, a framework in which a teleoperator is always in charge of every aspect of the robot's behavior can be cumbersome.

We propose to address this problem by using a Human-in-the-Loop (HitL) framework that uses both low-level, direct teleoperation (when needed) and autonomous modules for completing sub-tasks (when possible). Low-level command tools are available for complex situations, requiring direct intervention by the operator. Autonomous capabilities can play the important roles of reducing the load on the operator and increasing efficiency for sub-tasks that can be performed reliably, or that require operator input in a form that is relatively effortless to provide. In this section, we introduce the components of our interface built along these directions, focusing on three main components of a mobile manipulation platform: perception, manipulation, and navigation.

A. Looking Around and Perceiving

Situational awareness is crucial for any teleoperation interface. Our framework has a number of tools to allow the user to both directly visualize the environment, and also help the robot in better understanding its world.

The user can click anywhere in the streaming camera image feed to point the head, and 3D point clouds from the Kinect allow the user to see the 3D world from any camera angle in the virtual camera view. If compressed, point cloud data can also be streamed in real time. However, it is often the case that crucial parts of the environment will be occluded by the robot. For instance, while manipulating, the robot's arms and grippers typically occlude the object being manipulated. Users can thus take a point cloud snapshot, which stays in the rendered image and is only refreshed when requested. Such a cloud is shown on the right side of Fig. 2.

When the robot is looking at objects of interest for manipulation, the user can ask it to autonomously segment flat surfaces and well-separated ($\geq$3 cm) objects on the surfaces, (Fig. 5, top left). In more complex scenes that autonomous segmentation cannot handle, the user can interactively aid the robot by drawing boxes around objects to segment. Currently there are two interactive segmentation tools available for our interface, one that uses a graph-cut algorithm as described in [13], and one that uses the algorithm from [2].

The user can also ask the robot to autonomously recognize specific object models stored in a database. Currently, we provide a 2-D, ICP-like algorithm that can be used on previously segmented objects [3], as well a textured object detection algorithm that operates on general scenes [19]. Three of the objects in Fig. 5, top left have been autonomously recognized, and their object model meshes are shown in the appropriate poses. If needed, the user can correct the robot's recognition results, by clicking through a set of possible object detections returned by the object recognition algorithms, or by rejecting all returned detections.

The aforementioned features all use the head-mounted Kinect camera. The robot also has a base laser, used for localization and obstacle avoidance while navigating, and a tilting laser rangefinder used for both navigation and collision-free arm motion planning. These sensors build what we refer to as collision maps, or occupancy grids showing the obstacles in the environment. As we describe below, autonomous motion planning for both the base and the arms can be extremely useful for moving the robot. However, moving obstacles or sensor noise can clutter collision maps and leave them unusable. Our interface therefore enables the user to visualize, clear or regenerate the robot's collision maps, or even ignore them altogether and move open-loop.

B. Manipulating the Environment

Manipulating objects presents a very diverse set of challenges, such as the high dimensionality of the movement space, the non-anthropomorphic characteristics of the arms, and the complex and cluttered scenes encountered in real homes. Our system combines direct, manual teleoperation controls with controls that offer autonomous functionality in order to give the user efficient and flexible ways to accomplish a variety of tasks.

1) Manual Teleoperation: Despite the large body of research in autonomous manipulation, there will always be tasks that are not anticipated or well-handled by any set of autonomous modules, especially in complex, unstructured settings like a home. In situations where human ingenuity is needed to complete a task, low-level control of the robot's arms can provide the necessary means.

End-Effector Pose: The 7-DOF arms of the PR2 present an interesting challenge for 2D cursor control. For most tasks it is the 6D pose of the end-effector that is of interest. Our gripper control consists of a set of rings and arrows that allow the user to instantly move the gripper, along one dimension at a time. Dragging on a ring will rotate the gripper, while dragging an arrow will translate it (Fig. 3).

Elbow Posture: The PR2 arm has 1 degree of redundancy with respect to the pose of the end-effector. Intuitively, this redundancy allows the elbow to float at different "heights" while keeping the gripper in the same pose. The user may not care where the elbow is during a task, so long as the gripper can reach the desired workspace. However, if the posture of the arm causes unwanted collisions, we provide a 1D ring

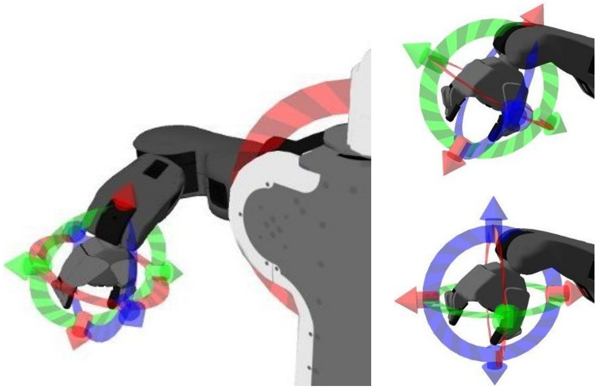


Fig. 3. Gripper control (6D) and shoulder ring (1D). Right images show a gripper-aligned control (top) and a world-aligned control (bottom).

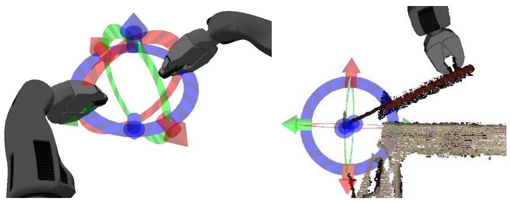


Fig. 4. Left: 2-armed Cartesian control. Right: Control frame moved to cabinet hinge for opening.

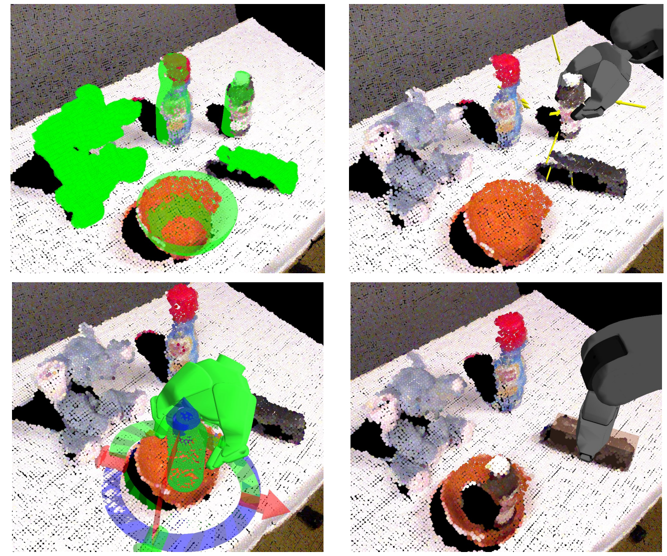


Fig. 5. Autonomous grasping for segmented and/or recognized objects. Top left: objects have been segmented, and three have been recognized (shown by superimposed meshes). Top right: autonomous grasping for a recognized object (peroxide bottle) using pre-computed database grasps. Bottom left: choosing a placing location for the peroxide bottle; the 3D mesh is available to the robot as the object had been previously recognized. Bottom right: Autonomous grasping for a segmented but unrecognized object (stapler).

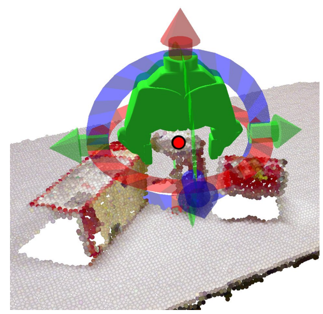


Fig. 6. The user indicates a desired grasp by placing a virtual gripper in the scene, and letting the robot plan a collision free execution path for it.

(Fig. 3) giving the user a measure of control of the elbow height, mapped loosely to the ring's rotary motion.

Changing the reference frame: Another useful feature of our low-level Cartsian control interface is the ability to move the reference frame for the Cartesian controller. Two common options, available through direct shortcuts, are to use either a gripper-aligned coordinate frame (e.g., for moving the fingertips directly into or out of grasps), or a world-aligned coordinate frame (e.g., for moving grasped objects directly up). Both types are shown in Fig. 3.

The user can also move the control reference frame to an arbitrary pose relative to the gripper; this is useful for tasks such as opening cabinets or using tools. Moving the control frame to the hinge of a cabinet (Fig. 4) allows a gripper grasping the cabinet handle to smoothly move in an arc around the hinge simply by rotating one control ring.

Our framework also allows both grippers to be moved at once by switching to a two-armed control mode, in which the reference frame is set to be halfway between the current poses for the two grippers; this mode is useful for moving around objects grasped by both grippers at once.

2) Autonomous Modules and Tools: For a certain class of well-defined and extensively-studied tasks, such as picking up an object, it can be more efficient to allow autonomous modules to handle most of the process.

Autonomous grasping: For objects that have been recognized, the user can right-click on the overlaid model and ask the robot to pick up the object; the robot then pulls a list of precomputed grasps for that object from its database and uses the first feasible one, as described in [3]. The user can also ask the robot to directly pick up unrecognized but segmented objects. In this case, the robot will compute grasps using the segmented object's point cloud, with the algorithm described in [8]. Both cases are illustrated in Fig. 5. In either case, upon successfully grasping an object, a collision model in the form of the object's bounding box will be attached to the robot's gripper, so that future planned motions avoid hitting the environment with the grasped object.

Collision free grasp execution: In situations where the robot cannot autonomously segment or recognize the object, or where a particular type of grasp is desired, a different interface allows the user to select just the final grasping pose, while still taking advantage of autonomous, collision-free motion planning. The user first clicks on a desired point in the 3D environment snapshot. A virtual gripper model is displayed at the clicked location; the operator can use a rings-and-arrows control to adjust its pose as desired (Fig. 6). As the user is adjusting the virtual gripper indicating the desired grasp, the robot continuously computes whether the grasp is feasible through collision-free motion planning. If it is, the virtual gripper control turns green indicating to the user that the grasp can be executed. Once the user is satisfied with a

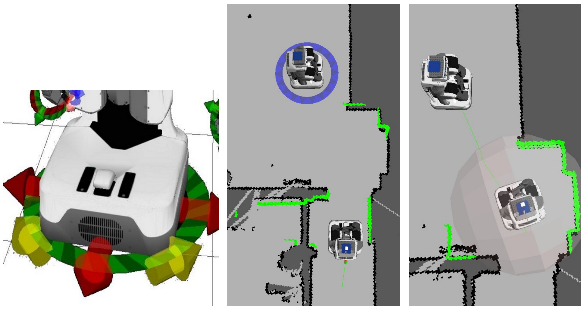


Fig. 7. Moving the base by clicking on rate-control arrows (left), or passing a goal to an autonomous navigation module (middle and right).

feasible grasp, the robot executes it autonomously.

Placing: The user can ask the robot to place a grasped object at a desired location in the environment, also selected through a virtual gripper rings-and-arrows control. If a model of the object is available, it can be shown along with the virtual gripper, so that the user can visualize the object at the desired place location (Fig. 5, bottom left). When the user is satisfied with the pose, the robot will autonomously plan a collision-free path to it and place the object.

Planned moves: The operator can use a virtual gripper control to simply move the gripper to a desired pose, letting the motion planner compute a collision free path that achieves the goal.

For grasping, placing, or planned moves, the motion planner can also be used with collision avoidance disabled; this type of open-loop movement can be useful in highly cluttered spaces, or when contact with the environment is desirable (e.g., pushing an obstacle). Advanced options available via an additional dialog box allow the user to further customize autonomous grasp and place execution with features such as reactive grasping [8], reactive placement, or slip detection/grasp force adjustment [18].

C. Moving Around: Base Movement and Navigation

For base movement and navigation, our framework provides tools for autonomous, collision-free navigation, as well as open-loop movement for small adjustments near or even into contact with obstacles. The user can ask the robot to navigate through free space by dragging a virtual robot model to a desired position and orientation in a scene. (Fig. 7, middle and right). If a map is available, it can be used as a reference when selecting the goal. 2-D map-making, global path planning, local/reactive path planning, and collision avoidance are provided by the PR2's navigation stack[12].

A similar control is used to perform small, precise, open-loop movements that 2-D autonomous navigation is incapable of or unwilling to execute (Fig. 8), such as moving the base under a table. Because the control takes on the current shape of the PR2 according to the robot's proprioception, the user can precisely position the goal relative to a 3D snapshot of the local environment. For both autonomous and open-loop goals, the robot can be stopped at any time by clicking on a translucent bubble that appears around the robot while it is driving (Fig. 7, right).

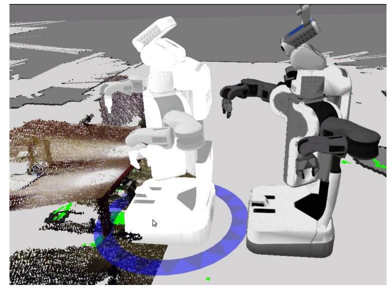


Fig. 8. Using an open-loop navigation goal allows the robot to get close to obstacles such as tables, or to push moveable obstacles with the base.

Fig. 9. Shelf environment used in the grasping study.

For less-precise but faster small adjustments, the user can also drive the base directly (both strafing and rotating) using rate-controlled arrows (Fig. 7, left).

IV. Demonstrations and Results

The cursor-based assistive teleoperation system described in this study was developed and tested in collaboration with a pilot user named Henry Evans. Henry is quadriplegic and mute due to a brainstem stroke, and can control a computer mouse via a headtracker and also issue single-button click commands with his thumb. Henry performed evaluations and proof-of-concept demonstrations of the system, and piloted its use in several real-life situations.

In this section, we present results from Henry operating the robot in three different contexts. The first attempts to quantify the performance of the manipulation tools for grasping in very cluttered environments. The second one is an example of object pick-and-place enabling social interaction, in the context of giving candy to Halloween trick-or-treaters. In the third demonstration, Henry uses the robot in his own home to retrieve a towel from the kitchen, combining navigation, door and drawer opening and closing, and object grasping.

A. Grasping in a Cluttered Environment

In this experiment, we quantified the ability of our system to execute grasping tasks, a key prerequisite for many manipulation tasks involving object acquisition or transport. Given the goal of operating in real users' homes, we focused

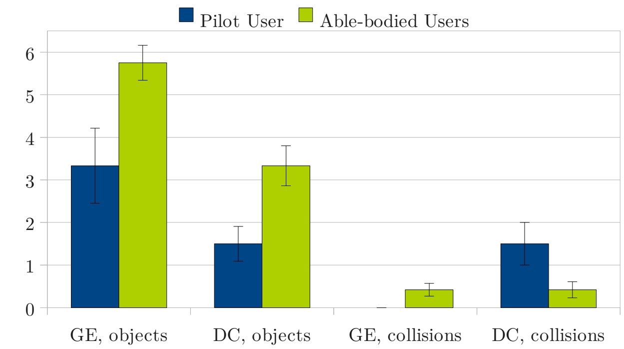


Fig. 10. Results for study on object grasping in cluttered environments for Grasp Execution (GE) and Direct Control (DC) strategies, showing the number of objects grasped and the number of collisions incurred.

on never-seen-before objects sitting in highly cluttered and constrained environments. As shown in Fig. 9, our test environment consisted of a small two-tiered shelf containing a large number of tightly packed objects.

The high degree of clutter and occlusion prevented the use of fully autonomous grasping, based on either object segmentation or recognition. We therefore quantified the performance of two grasping strategies:

- **Direct Control** (DC): this strategy involved the operator using the gripper Cartesian control tool in order to bring the gripper into a desired pose relative to a target object, before closing the gripper. When using this method for grasping, the user must essentially “drag” the gripper all the way from its starting position to the desired grasp location, while avoiding collisions along the way.
- **Grasp Execution**(GE): for this strategy, the operator used the final grasp point selection tool described in Sec. III-B. Once a desired grasp pose was confirmed, autonomous components were responsible for planning appropriate arm joint trajectories and avoiding unwanted collisions.

The operator’s goal was to securely grasp and lift as many objects from the shelf as possible in a limited amount of time (10 minutes per run), while avoiding collisions with the environment (the shelf or the table). Collisions were marked if unwanted contact was strong enough to displace the shelf or potentially cause damage to either the robot or its surroundings, as opposed to consequence-free occurences such as lightly brushing shelf walls.

With Henry Evans as our pilot user, we performed 3 trials using the Grasp Execution strategy, and 2 trials using the Direct Control strategy. Each trial was defined as a 10 minute run, over which we counted the number of objects grasped and collisions incurred. The results showed that our system can indeed enable an operator with motor impairments to execute grasps even in highly challenging environments. The robot successfully grasped 5, 3 and 2 objects respectively over the 3 trials using Grasp Execution. It also grasped 1 and 2 objects in the Direct Control trials. No collisions were encountered when using Grasp Execution; 1 and 2 collisions occured when using Direct Control.

Fig. 10 shows the complete results of our trials. For reference, we place them in the context of a previous user study [11] where we quantified the robot’s performance at the same task, but when operated by able-bodied novice users. However, the small sample size of our current Pilot User study limits the usefulness of quantitative comparisons performed between the two data sets.

Fig. 11. Henry (in the bottom right, using robot interface running on laptop) giving Halloween candy to children through the PR2 robot.

B. Trick-or-Treat: Social Interaction through a Robot

An important category of activities of daily living include social engagement and interaction with other people [1], [17]. In this study, we demonstrate a particular case where the ability to manipulate objects in the environment, and perform relatively simple pick-and-place operations, can serve as an enabler for social interaction. In the activity commonly referred to as trick-or-treating, occurring on the yearly Halloween holiday, children dress up in costumes and walk through the local community receiving candy from neighbors. In personal communication, Henry Evans described a desire to interact with trick-or-treaters and hand out candy, through the intermediary of the PR2 robot.

As a proof of concept implementation of this task, we set up a semi-structured environment designed for this type of interaction. The event took place on Halloween at a local mall, open to any children and their families present in the mall at that time. Henry and the robot sat behind a large table separating them from the public’s space. In addition, a small table with candy bars was placed to the robot’s side. The setup is shown in Fig. 11.

As a child would approach the robot and hold out his or her candy bag over the separating table, Henry would command the robot to pick up a candy from the side table and then place it inside the child’s bag. For picking up the candy, Henry used either the autonomous grasping tool based on object segmentation, or the Grasp Execution strategy described in the previous section, where he indicated the desired grasp point and the robot completed the grasp. For placing the candy in the bag, Henry moved the gripper to a pre-defined pose in front of the robot (either open-loop or using collision-free motion planning), then performed fine adjustments using direct Cartesian control in order to drop the candy inside the bag.

Over the course of one hour, Henry successfully handed out candy to more than a dozen local children. Occasional incidents included the robot dropping the candy bar on the way to a child's bag and having to re-grasp, or missing the bag on the first attempt to drop the candy inside and having to readjust. However, no major failures or disruptions occurred, and all interactions were completed successfully.

C. Mobile Manipulation in a Real Home

The vision driving our assistive robotics project is that of a robot operating in its user's home for indefinite periods of time. The variability and complexity encountered in real homes and the robustness needed for continuous operation over weeks and months, will be the ultimate reference criteria for such a system. As a step in this direction, Henry performed a number of tests with our system in his home. The environment was completely novel to the robot, with the single exception of a 2D floorplan of the house, acquired off-line, and used for localization.

In this setting, Henry demonstrated the ability of the system to perform a complete a task combining navigation, perception, and both prehensile and non-prehensile manipulation. Execution is illustrated in Fig. 12 and summarized below, along with the tools used for each component and the approximate time taken to execute:

- drive to from living room to kitchen: autonomous navigation combined with open-loop base movement (21 min).
- open kitchen cabinet door in order to inspect is contents: Grasp Execution tool for grasping handle, Cartesian end-effector control for opening door (6 min).
- close kitchen cabinet door: Cartesian end-effector control, used for pushing the door shut with the forearm (11 min).
- open kitchen drawer: Grasp Execution for grasping handle, Cartesian end-effector control for pulling drawer (6 min).
- grasping towel from drawer: Grasp Execution tool (3 min).
- bring towel to Henry's wheelchair in the living room: autonomous navigation combined with open-loop base movement (7 min).

The complete task was executed in a single continuous run (54 min), and succeeded on the first attempt.

V. Lessons and Limitations

While the ability to successfully perform a complete mobile manipulation task, on the first attempt and in a complex novel environment, is highly encouraging, our results also highlight numerous potential areas of improvement. Focusing on the time taken to perform the task, we note that:

- Due to a synchronization error between the robot's laser sensor and its computers, the autonomous navigation algorithm failed to process most goals received from the operator, who had to use manual base movement instead. As a result, navigation time accounted for more than half of the total task. This occurrence illustrates both the importance of having fallback mechanisms or multiple ways of achieving the same goal, and the potential efficiency gain obtained when more autonomous tools can be used.
- Cartesian end-effector control along with appropriate 3D perception can enable unforeseen tasks; for example, it enabled pushing a door shut with the forearm, a task that no module in our codebase was explicitly designed for. However, using it for more than small pose adjustments or short movements, especially relative to obstacles in the environment, can be laborious and time-consuming. Sub-tasks that could take advantage of autonomy (such as grasping the door knob, the cabinet handle, or the towel) were executed more efficiently.
- For grasping objects, the Grasp Execution strategy, where the operator only has to select the final gripper pose, proved more effective than Direct Control, as it leverages autonomous motion planning and collision avoidance. However, additional information from the robot can further increase efficiency. A significant part of the user's effort involves modifying a desired gripper pose so that the motion planner considers it reachable. We are currently augmenting our system to continuously sample the space of reachable, collision-free poses, and to provide suggestions in the area that the user is currently exploring.

In addition, we can draw a number of conclusions regarding mobile manipulators intended for in-home use, where few strong assumptions can be held. For example, a motion planning module for semi-structured settings could always expect the robot to be at a safe distance from any obstacle, and simply return an error if this assumption is violated. This, however, would greatly limit its usefulness in an unstructured setting, where unexpected contacts can and will occur.

The operator must also be equipped with tools to correct errors in the robot's view of the world. For example, in order to take advantage of the autonomous motion planner in as many cases as possible, the operator must be able to interact with the robot's representation of the world, adding obstacles that would otherwise be invisible to the robot's sensors (e.g. shiny or transparent objects), and removing non-existent obstacles that are the result of sensor noise.

The complexity of a real home can simply be beyond the capabilities of any autonomous algorithm. An illustrative example encountered in our pilot tests involved curtains billowing due to the air current from the robot's fan, and registering as obstacles in the robot's navigation map. Manual annotations, heuristic behaviors, or some level of altering the environment to suit the robot might be the only solutions for such extreme corner cases.

Finally, we have found that a simulated environment used for training was a key enabler to the successful execution of complex tasks. Even though a simulator can not accurately replicate all the complex interactions with real-life objects, it can still help the operator become familiar with the interface and robot, with no risk of injury or damage. We believe that appropriate training mechanisms will prove central to the effort of enabling non-roboticists to effectively use the widely-deployed assistive robots of the future.

Fig. 12. PR2 robot, operated by a pilot user, performing a mobile manipulation task in the user's home. From left to right: grasping a cabinet handle, pushing open a cabinet door, opening a drawer, grasping a towel inside the drawer, and navigating to desired drop-off location.

VI. Conclusions

In this paper we have shown how an assistive mobile robot, operated by a pilot user, can perform mobile manipulation tasks in rich, unstructured environments. The operator sends commands to, and receives feedback from the robot through an interface running on a commodity desktop or laptop computer, using only a head-tracker cursor as an input device. The available command tools range from low-level Cartesian movement commands for the base or gripper to autonomous modules for collision-free navigation or grasping.

Our pilot studies showed that a motor-impaired operator can command the robot to successfully grasp objects even amidst clutter and in constrained settings. We have also tested our system in a real home environment, where a pilot user completed a mobile manipulation task involving both prehensile and non-prehensile manipulation, as well as perception and navigation. The task was successfully completed on the first attempt, although we found that a number of components would greatly benefit from increased performance and faster execution time. Our experience performing these studies suggests that mobile manipulators have the potential to enable motor impaired users to perform a wide range of activities of daily living. Through a combination of human control and autonomous algorithms, assistive robots could one day gain the versatility and robustness needed for long term operation in real homes.

References

[1] J. M. Beer, A. Prakash, T. L. Mitzner, and W. A. Rogers. Understanding robot acceptance. Technical Report HFA-TR-1103, Georgia Institute of Technology: Human Factors and Aging Laboratory, 2011.

[2] M. Bjorkman and D. Kragic. Active 3d scene segmentation and detection of unknown objects. In *IEEE Intl. Conf. on Rob. and Automation*, 2010.

[3] M. Ciocarlie, K. Hsiao, E. G. Jones, S. Chitta, R. B. Rusu, and I. A. Sucan. Towards reliable grasping and manipulation in household environments. In *Intl. Symp. on Exp. Robotics*, 2010.

[4] J.L. Dallaway, R.D. Jackson, and P.H.A. Timmers. Rehabilitation robotics in europe. *IEEE Trans. on Rehabilitation Eng.*, 3(1), 1995.

[5] D. Gossow, A. Leeper, Hershberger D., and M. Ciocarlie. Interactive markers: 3-D user interfaces for ROS applications (ROS Topics column). *IEEE Robotics & Automation Magazine*, 18(4), 2011.

[6] B. Graf, M. Hans, and R.D. Schraft. Care-o-bot iidevelopment of a next generation robotic home assistant. *Autonomous robots*, 16(2):193–205, 2004.

[7] K.Z. Haigh and H.A. Yanco. Automation as caregiver: A survey of issues and technologies. *AAAI Workshop on Automation as Caregiver*, pages 39–53, 2002.

[8] K. Hsiao, S. Chitta, M. Ciocarlie, and E. G. Jones. Contact-reactive grasping of objects with partial shape information. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, 2010.

[9] S. E. Iso-Ahola. Social psychological foundations of leisure and resultant implications for leisure counselling. In E.T. Dowd, editor, *Leisure counselling: Concepts and applications*, pages 97–125. Springfield, IL: Charles C. Thomas, 1984.

[10] A. Jain and C.C. Kemp. El-e: An assistive mobile manipulator that autonomously fetches objects from flat surfaces. *Autonomous Robots*, 28(1):45–64, January 2010.

[11] A. Leeper, K. Hsiao, M. Ciocarlie, L. Takayama, and D. Gossow. Strategies for human-in-the-loop robotic grasping. In *Proc. of Human-Robot Interaction (HRI)*, 2012.

[12] E. Marder-Eppstein, E. Berger, T. Foote, B. Gerkey, and K. Konolige. The office marathon: Robust navigation in an indoor office environment. In *Intl. Conf. on Robotics and Automation*, 2010.

[13] B. Pitzer, M. Styer, C. Bersch, C. DuHadway, and J. Becker. J. becker, towards perceptual shared autonomy for robotic mobile manipulation. In *IEEE Intl. Conf. on Rob. and Automation*, 2011.

[14] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng. ROS : an open-source Robot Operating System. In *IEEE Intl. Conf. on Rob. and Automation*, 2009.

[15] M.A. Regalbuto, T.A. Krouskop, and J.B. Cheatham. Toward a practical mobile robotic aid system for people with severe physical disabilities. *J. of rehabilitation research and development*, 29(1), 1992.

[16] A. Remazeilles, C. Leroux, and G. Chalubert. Sam: A robotic butler for handicapped people. In *IEEE Intl. Symp. on Robot and Human Interactive Communication*, 2008.

[17] W. A. Rogers, B. Meyer, N. Walker, and A. D. Fisk. Functional limitations to daily living tasks in the aged: A focus group analysis. *Human Factors*, 1:111–125, 40.

[18] J. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K. Kuchenbecker. Human-inspired robotic grasp control with tactile sensing. *IEEE Trans. on Robotics*, 27, 2011.

[19] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. ORB: An efficient alternative to SIFT or SURF. In *Intl. Conf. on Computer Vision*, 2011.

[20] Rviz documentation page. http://www.ros.org/wiki/rviz.

[21] R. Schulz. Effects of control and predictability on the physical and psychological well-being of the institutionalized. *Journal of Personality and Social Psychology*, 33:563–573, 1976.

[22] M.S. Searle, M.J. Mahon, S.E. Iso-Ahola, H.A. Sdrolias, and J. Van Dyck. Enhancing a sense of independence and psychological well-being among the elderly: A field experiment. *Journal of Leisure Research*, 27:107–107, 1995.

[23] J.M. Wiener, R.J. Hanley, R. Clark, and J.F. Van Nostrand. Measuring the activities of daily living: Comparisons across national surveys. *Journal of Gerontology: SOCIAL SCIENCES*, 45(6):S229–237, 1990.

[24] Willow Garage, Inc. The PR2 Robot. http://www.willowgarage.com/pages/pr2/overview.

Intention Driven Human Aware Navigation for Assisted Mobility

Jorge Rios-Martinez[1], Arturo Escobedo[1], Anne Spalanzani[2], Christian Laugier[1] †
[1]INRIA Rhone-Alpes, Grenoble, France
[2]UPMF-Grenoble 2 - INRIA - LIG, France

***Abstract*— Ensuring proper living conditions for an ever growing number of elderly people is an important challenge for many countries. The difficulty and cost of hiring and training specialized personnel has fostered research in assistive robotics as a viable alternative. In particular, this paper studies the case of a robotic wheelchair, specifically its autonomous navigation and user adapted control. Integration of a technique to interpret user intention using head movements and a human aware motion planner is presented. Test results exhibit emerging behavior showing a robotic wheelchair interpreting gesture commands and taking the user to his desired goal, respecting social conventions during its navigation.**

***Index Terms*— Proxemics, Human aware navigation, User intention, Adapted control.**

I. Introduction

Ensuring proper living conditions for an ever growing number of elderly people is a significant challenge for many countries. The difficulty and cost of hiring and training specialized personnel has fostered research in assistive robotics as a viable alternative. In this context, an ideally suited and very relevant application is to transport people with reduced mobility as it can help them to preserve their independence. For such systems, it is crucial to take into account the actual needs and characteristics of both its users and the people around them. This paper studies the case of a robotic wheelchair, specifically its autonomous navigation and user adapted control, whose operation has been designed around the following requirements:

- Usability: People with motor disabilities or aging people often have problems using joysticks and other standard control devices. The system should account for this, for example by favoring the most reasonable actions when presented with an ambiguous command.
- Safety: The system should avoid collisions with both static and dynamic entities.
- Sociability: When moving, a robot may considerably disturb people around it, especially when its behavior is perceived as unsocial. Even worse, the wheelchairs passenger may be held responsible for that behavior. It is thus important to produce socially acceptable motion.

Social capability of planner chosen is based on the simple idea that, in a human populated environment, when people interact, they often adopt spatial formations implicitly forming "interaction zones". Thus, socially acceptable motion can be enforced not only by respecting personal space but also by detecting interaction zones and then computing the risk to invade them.

Usability can be improved by adding contextual information in order to ease interaction with the user, for example, the knowledge of interesting or frequently visited locations in a particular environment can be used by the system to infer the user's plan. Once the plan is identified the system can assist the user by executing the low level needed commands. The structure of this paper is as follows:

Section II offers an overview of related works. Section III presents our technique for achieving user adapted control. Section IV describes the RiskRRT method. In section V examples of execution on our real platform are exhibited. Section VI presents conclusions about the work and perspectives.

II. Related work

A. Semi-Autonomous Navigation

In latest years many efforts have been made to develop robotic wheelchairs that operate in a similar manner to an autonomous robot, where the user gives a final destination and supervises as the smart wheelchair moves (e.g.,NavChair [1], MIT Media Lab wheelchair [2]).

Other smart wheelchairs limit their assistance to collision avoidance and leave the majority of planning and navigation duties to the user. These systems do not normally require prior knowledge of an area or any specific alterations to the environment. They require instead more planning and continuous effort on the part of the user and are only appropriate for users who can effectively plan and execute a path to a destination.

Shared control is presented in situations in which the assisting device combines the control input coming from the robot and the user. This device may be a wheelchair, a tele-operated robot, a robotic travel aid for the visually impaired, or any other device where robot and human cooperate in a task [3].

The estimation of the user's plan is a key point in many shared control tasks because it allows to the automatic controller/robot to adequate its actions to the desire of its user. Inferring the user plan is necessary whenever the interface with the user doesn't allow him to explicitly dictate this to the robot as with many popular electric wheelchair interfaces.

Some methods aiming the implicit estimation of the users intention from simple joystick inputs have been proposed in [3], [4]. They model the users intent as possible trajectories

† {jorge.rios-martinez, jesus.escobedo-cabello, anne.spalanzani, christian.laugier}@inria.fr
Jorge Rios-Martinez has been partially funded by CONACYT 250140/308006.

to follow, then a probability distribution is maintained over the set of trajectories and finally the selection of the most probable one is done using the input from the user within a Bayesian framework.

In [5] a learned Partially Observable Markov Decision Process (POMDP) is used to estimate the intended destination into a predefined map of the environment in a high level topological manner. This means that the user focuses on driving the wheelchair from one spatial location to another without having to worry about all the low level control. Places of interest are selected as spatial locations in the environment where the user spends significantly most of his time.

The method presented in this article to infer the user's intended goal aims to build a model as simple as possible combining the Bayesian network approach of the first mentioned methods and the simple topological goal based representation of the environment used in the second one. A more natural human-machine interface based on a face tracking system is used to command the wheelchair (III-A) while the navigation is performed using a human aware planning algorithm (IV).

B. Human Aware Navigation

Human aware navigation is receiving an increasing attention in robotics community, this area of research appears once that robots navigate in human environments and safety solutions are not enough; now the main concern is related to produce solutions which also have to be understandable and acceptable by human beings. Next is a review of related works. A proposal of human aware navigation was presented in [6], where a motion planner takes explicitly into account its human partners. The authors introduced the criterion of visibility, which is simply based on the idea that the comfort increases when the robot is in the field of view of a person. Other work, [7], introduced an adaptive system which detects whether a person seeks to interact with the robot based on the person's pose and position, that system was presented as a basis for human aware navigation. Their results showed that the system was capable of navigate based in past interaction experiences and to adapt to different behaviors.
In [8] it was proposed a Spatial Behavior Cognition Model (SBCM) to describe the spatial effects existing between human-human and human-environment. SBCM was used to learn and predict behaviors of pedestrians in a particular environment and to help a service robot to take navigation decisions. Technique in [9] proposed an on-line method to learn generally occurring motion patterns in an office environment with a mobile robot. Navigation is realized by using these patterns, in form of sampled HMM, along with a Probabilistic Roadmap based path planning algorithm. Socially acceptable motion is achieved by minimizing social distractions, such as going through someone else's working space.

The work presented in [10] proposed rules that a single robot should obey in order to achieve not only a safe but also a least disturbance motion in a human-robot environment.

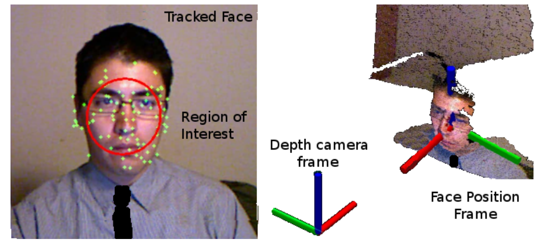


Fig. 1. In the presented approach the user drives the wheelchair by using his face. The face of the user is tracked by processing the RGB image received from a Kinect sensor (left image). The pose of the face is estimated from the depth data as shown in the right.

Rules define sensitive zones for both humans and robots, depending either on their security regions or on psychological feeling of humans.

Personal space, o-space and their relation to comfort were addressed in [11], where a risk based navigation was extended to include risk due to discomfort. Human's movement is supposed to be known by learning of typical trajectories in a particular environment. Optimization techniques that take into account social factors have been also proposed. In [12] a generalized framework for representing social conventions as components of a constrained optimization problem was presented and it was used for path planning, their results exhibited a more social navigation. In [13] an stochastic adaptive optimization method was used to minimize discomfort of humans in the environment, while robot navigate to the goal. Results show robot navigation respecting both information process space and personal space of people. Recently, legibility of robot navigation around humans was explored in [14]. A context depending cost model was developed to adjust robot behavior to human behavior in crossing scenarios.

III. Semi Autonomous Controller

A. Face Control Subsystem

The user controls the robotic wheelchair by using the movements of his head, this is accomplished by means of a face tracking system that estimates the direction of the sight of the user using data from a Microsoft's Kinect ©. The images taken by the 2D RGB camera are used to set up a region of interest over the depth data coming from the infrared sensor.

The face of the user is detected using a typical Haar detector then a set of SWIFT features are selected over the face and 2D tracking is performed using the Lucas-Kanade method as described in [15]. The identification of the face pose is done by a random forest classifier which takes as input the 3D data from the Kinect sensor and gives the estimated position of the face [16], the results of the face tracking are shown in figure 1.

B. User Intentions Estimation System

The user intentions are modeled as topological poses into a predefined map. Those locations are set by the user's

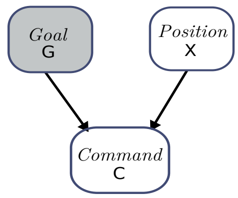


Fig. 2. The Bayesian network used to estimate the hidden user intention or goal G by knowing the current position X and the user command C.

habits (those places where the user spends most of his time are taken into account as probable goals) also interesting points taken from the map of the environment as doors, desks and other facilities are taken into account as probable destinations.

The reasoning method used is based on a Bayesian Network depicted in figure 2 that combines the information taken from the user interface input with the prior knowledge of the environment to infer a posterior probability distribution over the set of possible goals in the environment. This probabilistic model aims to take into account the uncertainty in the estimation of the desired goal and the inherent error over the user command read from the face tracking system.

To estimate the status of the *goal* variable, the command direction coming from the user-machine interface and the current user's position are applied as evidence. The posterior probability of the current goal given the position of the user and the direction of the command is expressed as $P(G_t^i|C_tX_t)$. The prior probability over each goal in the environment is denoted as $P(G_t^i|X_t)$.

Using Bayes' rule:

$$P(G_t^i|C_tX_t)) = \frac{P(C_t|X_t, G_t^i)P(G_t^i|X_t)}{P(C_t|X_t} \quad (1)$$

This can be simplified by using the normalizer η.

$$P(G_t^i|C_tX_t)) = \eta P(C_t|X_t, G_t^i)P(G_t^i|X_t) \quad (2)$$

We assume that the direction of the command given by the user C_t depends on the current position X_t and the intended goal G_t^i. Therefore this probability should be higher for goals that are in the direction of the current command input as shown in figure 3. $P(C_t|X_t, G_t^i)$ represents the probability of giving a command C_t when the user is located at position X_t and her goal is at position G_t^i at current time t. The notation G_t^i is used to express $G_t = g_i$ where g_i is one of the predefined goals in the environment as appear in Fig. 3.

$$P'(C_t|X_t, G_t^i) = \frac{1 - |a_i|}{\pi} \quad (3)$$

Where a_i term is the angle between both command and goal directions. The we normalize it as:

$$P(C_t|X_t, G_t^i) = \frac{P'(C_t|X_t, G_t^i)}{\sum_i P'(C_t|X_t, G_t^i)} \quad (4)$$

The prior probability table $P(G_t^i|X_t)$ was set manually for our current experimental set-up taking into account the user's

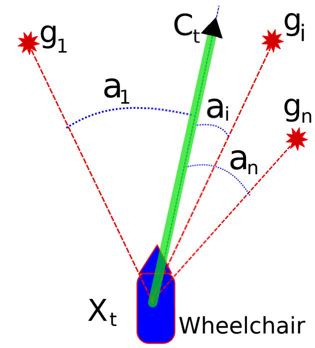


Fig. 3. The probability distribution for a given command C_t (big arrow) is proportional to the angle a_i formed with respect to each goal g_i in the environment.

habits. However, for a real-environment it is mandatory to learn its values autonomously using some machine learning method.

IV. RiskRRT approach

A. Planner main loop

Our planner is based on RiskRRT [17], a partial motion planner which integrates motion predictions to provide safe trajectories. At each loop the algorithm checks if a new goal is available. When a new goal is received, RiskRRT collects information about the static obstacles and humans' position in the environment. Then a prediction of pedestrians trajectories takes place. At this moment RiskRRT has enough information to proceed with the exploration of the environment by means of a Rapidly-Exploring Random Tree which is constantly updated with the new perceived data. New nodes are created by selecting possible compliant controls that conduct the robot towards randomly selected points. The process of exploration has a depth threshold for the nodes and limited time in order to achieve real time performance. A probability of collision is assigned to each node taking into account static obstacles and human predicted trajectories.
Finally, a best path is selected by choosing the branch with the lowest probability of collision and with the closest distance to goal. It is important to mention that RiskRRT generated paths include information about space, time and robot dynamics which is very important advantage to navigate in dynamic environments. In the fig. 4 planner execution at two distinct iterations can be observed. Tree shows the portion of the environment already explored. Nodes are represented by colored spheres, their color represents the time at which they would be reached by the robot, same interpretation is done for color in predicted trajectories of humans. Size of nodes represents the estimated risk, therefore a node very close to a predicted pedestrian position of the same color will have a big size.

This method has been extended by including a mechanism to obtain socially acceptable behavior which is explained in next section.

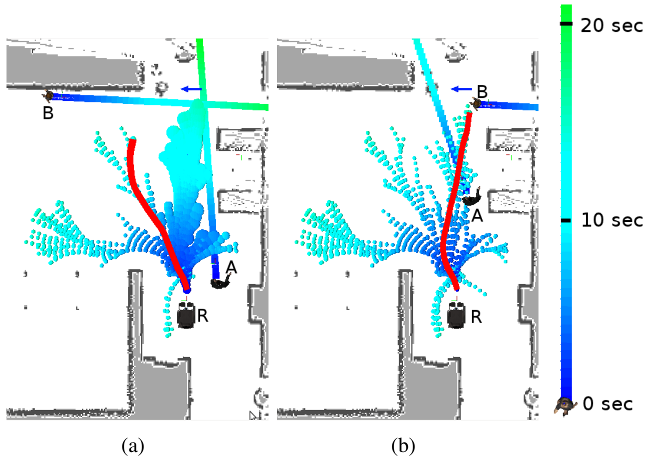


Fig. 4. Example of execution for the RiskRRT at two distinct iterations. One goal (blue arrow) has been passed to the wheelchair (R), while two people (A,B) are walking around. In a) partial path solution was found (red line) avoiding high risk zones identified in the image by the bigger size of nodes. Wheelchair is deliberately not moving then, after some instants, new observations and predictions produce different risk estimation and permits to select a better path.

B. Social conventions in human navigation

When the wheelchair is transporting a human, it will have to move in a populated environment where an "optimal" behavior may be perceived as unsocial. People will become uncomfortable if they are approached at a distance that is deemed to be too close, where the level of discomfort experienced by the person is related to the importance of his or her space. This simple idea was formalized introducing the concept of *personal space*, first proposed by Hall [18], which characterizes the space around a human being in terms of comfort to social activity.

Another interesting social situation arises when two or more of the persons in the environment are interacting. We model interactions using the concept of *o-space* which has been developed by sociologists [19]. This space can be observed in casual conversations among people where participants' position and orientation are used to establish boundaries of the space. This space is respected by other people and only participants are allowed to access to it, therefore the intrusion of a stranger causes discomfort. In our path planner, human friendly paths are generated by including a module called **"Social Filter"** which transforms those spaces into corresponding cost functions which lead the robot to avoid them. As a result, the choice of a best path done by RiskRRT is done by considering the probability of not encountering a collision along the path and not entering in a personal space or an o-space. Detailed explanation can be found in [11].

1) Modeling Personal Space: We have modeled personal space as a composition of two human centered Gaussians, one for the front and one for the back of the space, the front is larger as people is more sensitive to this space. Fig. 5 shows an example of personal space as provided by the Social Filter.

Fig. 5. Personal space calculated by Social Filter Module. Height of the Gaussian means Risk of disturbance then maximum disturbance is located at human position.

2) Modeling o-Space: When more than two people are in conversation, they tend to make a formation with circular shape. The o-space could be taken as a circle whose center coincides with that of the inner space. For the specific case of two people, some formations, called F-formations, have been identified as being particularly frequent [19]. The social filter identifies individual F-formations (Vis-a-vis, L-Shape, C-Shape or V-Shape) and builds the corresponding o-space. in Fig. 6, the calculated o-space for a Vis-a-Vis interaction is shown.

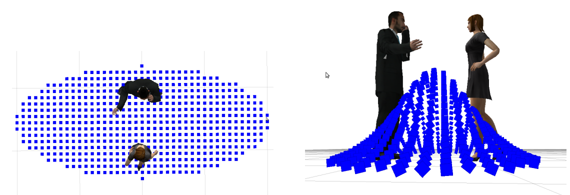


Fig. 6. O-space calculated by Social Filter Module for a Vis-a-Vis F-formation. Maximum risk of disturbance is located at o-space center, in the picture the disturbance is represented by height of Gaussian.

C. Planning under social constraints

This section explains how the social constraints are included in the RiskRRT framework, see [11], [17] for specific details. First we define PZ_i as the probability of disturbing by passing inside the o-space (sec. IV-B.2) of interaction i, and we calculate it as the maximum value of o-space model for that interaction evaluated in the intersection with the robot's path. PZ_i can be thought as a collision with a dynamic obstacle:

$$P_{cd} = 1 - \prod_{m=1}^{M} [1 - P_{cd}(o_m)] \prod_{i=1}^{r} [1 - PZ_i] \qquad (5)$$

where P_{cd} is the probability of dynamic collision considering the M humans in the environment and $P_{cd}(o_m)$ is the probability of collision with the human o_m taking into account the personal space. $P_{ps}(o_m)$ is the risk of disturbing by passing in such personal space and can be approximated as the probability that A, the area swept by the robot's path, intercepts the one represented by the personal space:

$$P_{ps}(o_m) = \int_A PS(o_m(t)) \qquad (6)$$

where $PS(o_m(t))$ is the model of personal space centered in $o_m(t)$ at time t as described in IV-B.1. To take into account this last constraint we use:

$$P_{cd}(o_m) = P_{dyn}(o_m) + P_{ps}(o_m)(1 - P_{dyn}(o_m)) \quad (7)$$

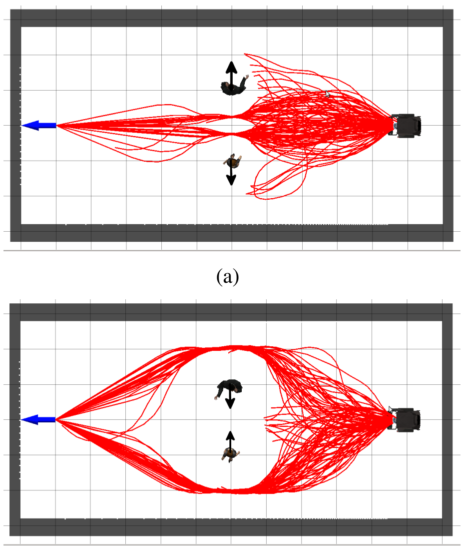
where $P_{dyn}(o_m)$ is the probability of dynamic collision between the robot and o_m considering only their trajectories. Last, total probability of collision, $collP$ is calculated for each node using:

$$collP = P_{cs} + (1. - P_{cs}) * P_{cd}; \quad (8)$$

Finally, also for each node we calculate a weight based on probability of collision and distance to goal. In order to compare paths worst node weight is chosen to represent a particular path.

D. Simulation results

In order to test socially acceptable behavior, we conducted several simulation tests.

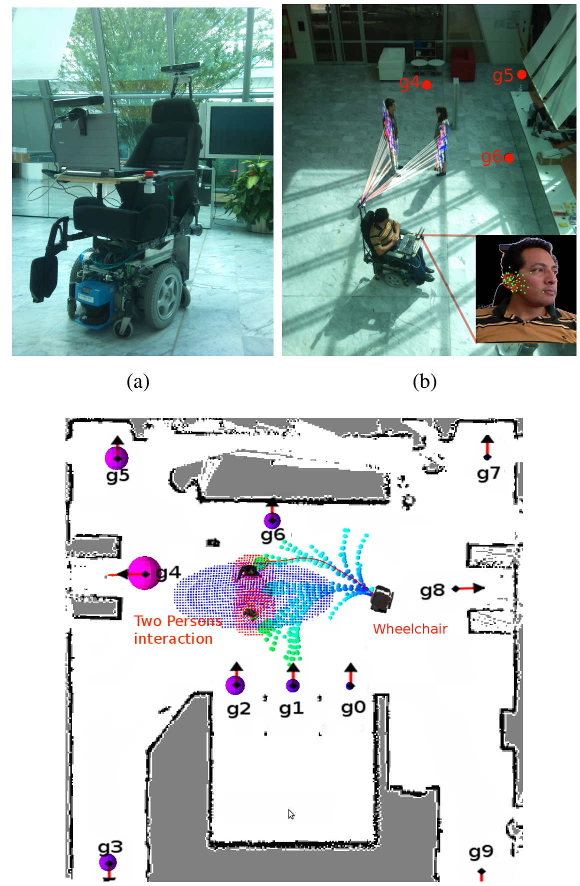


(a)

(b)

Fig. 7. Socially acceptable navigation in RiskRRT. Each figure shows a sample of generated plans (in red) for one hundred executions of RiskRRT, each execution run for twenty iterations of the algorithm. Goal is the blue arrow. In a) static persons are looking towards walls, therefore there is no social interacting zone, then navigation respects only their personal spaces. In b) both social spaces are respected.

Our first test scenario consisted of two people in the same corridor together with the wheelchair whose task is to navigate towards the goal. We realized one hundred executions of the planner in very similar conditions, as it can be seen in Fig. 7, when the social filter is turned on, all the plans managed to respect both personal space and interaction space without disturbing the involved people. We can observe that because of the random exploration of the environment some executions (almost ten percent) select as best path one that stops before to invade social space. Tests with the social filter off showed that thirty percent of plans passed in the middle of an interaction.



(a)

(b)

(c)

Fig. 8. Experimental scenario, in a) The wheelchair used for testing in b) the INRIA's hall. Possible goals are marked with a circle, the persons present in the scene are tracked by one of the Kinect sensors while the other Kinect is used to track the face of the user. c)Results of the system in a real scenario. The size of the spheres in the environment represent the value of the computed posterior probability for each destination. The computed trajectory and interaction regions are also shown.

V. Integration and Evaluation

Integration was done by using the semi-autonomous controller presented in sec. III as the source of the goals for the RiskRRT planner. Solution plans are executed in our experimental platform, details follow.

A. Experimental Platform

Our platform, is an automated wheelchair (Fig.8(a)) equipped with one Sick laser and two Microsoft Kinect, running ROS (Robotic Operating System) for achieving semi-autonomously mobility actions commanded by the wheelchair's user. Laser permits us to build a map of the environment, like shown on Fig. 8(c). Data coming from the upper Kinect allow us to have position and orientation of pedestrians in the scene while data from frontal Kinect collect face features to feed our intention recognizer algorithm.

B. Evaluation in a real scenario

The system was evaluated at scenario shown in Fig.8(b). User can start the movement at any location of the experimental scenario, he is asked to drive the wheelchair by seeing towards his desired destination. In the example, the user is seeing to the left so that it is more probable that he is aiming to go to the coaches located in that direction. The direction of his face is computed as previously explained in section III-A. Typical destinations were defined into the map, they are marked with small arrows in Fig.8(b).

Whenever a new command is read from the face control, the user estimation module computes the goal with the highest posterior probability, depicted in Fig.8(c) as the size of the sphere marking each goal, then it is sent to the navigation module to start the movement.

The navigation module receives the map of the environment, the currently computed goal and the list of people present in the field of view of the frontal viewing Kinect to compute the necessary trajectory to the goal as shown in Fig.8(c). In the example there are two persons in conversation, standing in the middle of the path between the wheelchair and the current estimated goal. Even if the user is pointing to the goal located in the other side of the two persons he does not have to worry about the necessary planning and commands to avoid interrupting the conversation because the autonomous navigation system is in charge of that.

VI. Conclusions and future work

The approach presented in this paper has integrated a human aware motion planner and a semi-autonomous controller adapted to the user with the aim to preserve independence of people with reduced mobility. The system was designed to improve both usability by reasoning about user intentions and sociability by including concepts like personal space and o-space in the planning algorithm. Experiments with a real wheelchair show that integration has been successfully achieved. The user intention algorithm has proved to be useful to translate simple input commands into high level orders or goals for our autonomous navigation system.

Using the pose of the face as input can be advantageous to assist the elderly because it provides a more natural way of interaction (we usually see where we are going) so they can be more confident when using the wheelchair. In this work we explored the use of head direction to infer the user desired destination in a navigation task. The system is capable of estimating the desired goal among a set of interesting points and to transport the user to it while avoiding humans in a social way.

Current work will be extended in two directions. First, by including as interesting points positions where the user should be located when he wants to join a group of people. Second, by adapting autonomously the user intention algorithm to user disability. Moreover other user-machine interfaces, like voice based, will be included to minimize ambiguities like that of the movements of the head that are intended to give a navigation command from those resulting from the natural observation of the environment.

References

[1] S. P. Levine, D. a Bell, L. a Jaros, R. C. Simpson, Y. Koren, and J. Borenstein, "The NavChair Assistive Wheelchair Navigation System." *IEEE transactions on rehabilitation engineering*, vol. 7, no. 4, pp. 443–51, Dec. 1999.

[2] T. Kollar, S. Tellex, D. Roy, and N. Roy, "Grounding Verbs of Motion in Natural Language Commands to Robots," in *International Symposium on Experimental Robotics*, 2010.

[3] E. Demeester, M. Nuttin, D. Vanhooydonck, and H. Van Brussel, "Assessing the user's intent using bayes' rule: application to wheelchair control," in *Proceedings of the 1st International Workshop on Advanced in Service Robotics*, 2003, pp. 117–124.

[4] E. Demeester, A. Huntemann, D. Vanhooydonck, G. Vanacker, A. Degeest, H. Brussel, and M. Nuttin, "Bayesian Estimation of Wheelchair Driver Intents: Modeling Intents as Geometric Paths Tracked by the Driver," *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5775–5780, 2006.

[5] T. Taha, J. V. Miro, and G. Dissanayake, "POMDP-based long-term user intention prediction for wheelchair navigation," *IEEE International Conference on Robotics and Automation*, pp. 3920–3925, 2008.

[6] E. A. Sisbot, L. F. Marin-Urias, R. Alami, and T. Simeon, "A human aware mobile robot motion planner," *IEEE Transactions on Robotics*, vol. 23, 2007.

[7] S. T. Hansen, M. Svenstrup, H. J. Andersen, and T. Bak, "Adaptive human aware navigation based on motion pattern analysis," *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, 2009.

[8] S.-Y. Chung and H.-P. Huang, "A mobile robot that understands pedestrian spatial behaviors," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, oct. 2010, pp. 5861 –5866.

[9] S. Sehestedt, S. Kodagoda, and G. Dissanayake, "Robot path planning in a social context," in *Robotics Automation and Mechatronics (RAM), 2010 IEEE Conference on*, june 2010, pp. 206 –211.

[10] C. P. Lam, C. T. Chou, K. H. Chiang, and L. C. Fu, "Human-centered robot navigation, towards a harmoniously human-robot coexisting environment," *IEEE Transactions on Robotics*, vol. 27, no. 1, 2011.

[11] J. Rios-Martinez, A. Spalanzani, and C. Laugier, "Understanding human interaction for probabilistic autonomous navigation using Risk-RRT approach," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sep. 2011, pp. 2014–2019.

[12] R. Kirby, R. Simmons, and J. Forlizzi, "Companion: A constraint-optimizing method for person acceptable navigation," *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, 2009.

[13] J. Rios-Martinez, A. Renzaglia, A. Spalanzani, A. Martinelli, and C. Laugier, "Navigating between people: A stochastic optimization approach," in *2012 IEEE International Conference on Robotics and Automation*, 2012.

[14] T. Kruse, P. Basili, S. Glasauer, and A. Kirsch, "Legible robot navigation in the proximity of moving humans," in *Advanced Robotics and its Social Impacts (ARSO), 2012 IEEE Workshop on*, may 2012, pp. 83 –88.

[15] "Pi face tracker webpage," 2012. [Online]. Available: www.pirobot.org

[16] G. Fanelli, J. Gall, and L. Van Gool, "Real time head pose estimation with random regression forests," *Computer Vision and Pattern Recognition*, 2011.

[17] C. Fulgenzi, A. Spalanzani, and C. Laugier, "Probabilistic motion planning among moving obstacles following typical motion patterns," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, oct. 2009, pp. 4027 –4033.

[18] E. T. Hall, *The hidden Dimension: Man's Use of Space in Public and Private*. The Bodley Head Ltd, London, UK, 1966.

[19] A. Kendon, "Spacing and orientation in co-present interaction," in *Development of Multimodal Interfaces: Active Listening and Synchrony*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010, vol. 5967, pp. 1–15.

Robot Navigation Taking Advantage of Moving Agents

Procópio Stein[1,2], Anne Spalanzani[1,3], Vítor Santos [2] and Christian Laugier[1]

***Abstract*— A crucial requirement for service robots is to be able to move in dynamic environments shared with humans as well as interact with them. Navigation in such environments is a challenging task, as the environment is constantly changing, future states have to be predicted and planning and execution must be carried on-line.**

However, even in very complex situations, humans can easily find a path that avoid both dynamic agents and static obstacles. This paper proposes a technique to take advantage of the human movement in such populated environments, selecting a leader to be followed in a probabilistic fashion, according to the robot's desired destination.

In this way, the robot can take advantage of the paths traveled by humans, effortlessly avoiding dynamic and static features as the human leader does, relieving the robot from the burden of having to generate its own path.

I. INTRODUCTION

With advances in mobile robotics and lowering costs of computers, it is becoming more and more common for us to find robots among groups of people. Service robots (home care, hospital, museum guides) are real example cases where robots have to be able to move and interact with humans in an ever changing environment. The success of interactions and human acceptance of service robots is directly related to the way they behave and approach others, as well as their capability to adapt to the environment.

However, navigation in dynamic environments is still an open and challenging issue for the robotic community. In such environments, sensor's measurements are prone to noise and the measurements have a short lifespan, being valid only for small time periods. Static features, that could guide a navigation algorithm, might not be detected. There are also limitations in the time spent by the navigation algorithm to provide solutions, and optimal approaches are unsuited for this task, as generated paths might be valid only for few time-steps.

Despite these difficulties, humans can easily navigate in dynamic environments. Moreover, their movement can provide indirect information of the environment, as they usually do not move at random. Instead of that, they move according to typical patterns, and their movements are related to features that they are interested in, such as doors, elevators, stairs or other people [1].

This research has been funded by the INRIA PAL Project and by Fundação para a Ciência e a Tecnologia SFRH/BD/46604/2008.

[1] Institut National de Recherche en Informatique et Automatique, INRIA Rhône Alpes, France anne.spalanzani@inria.fr, christian.laugier@inria.fr

[2] Department of Mechanical Engineering, Universidade de Aveiro, Portugal procopiostein@gmail.com, vitor@ua.pt

[3] UPMF, Grenoble, France

Taking advantage of that, several works propose to model those typical paths with probabilistic approaches as Gaussian Processes [2] [3] [4] or Hidden Markov Models [5], to address the problem of navigating in populated environments [6], [7]. Correctly detecting and learning a likely path permit the robot to avoid trajectories that have a risk of future collision with a pedestrian, as well as avoiding social disturbances [8], [9].

Those approaches, however, do not take into account changes people perform in their typical paths to avoid and adapt to other moving people [10]. These conditions allied to excessive future uncertainty may lead to situations where every generated path lead to colisions or frozen situations, as shown by [11].

The key insight of this paper is to present a navigation technique for dynamic environments that takes advantage of the typical movement of humans. This approach relies on the fact that people try to guarantee their safety and the safety of other people in the same environment, avoiding obstacles and avoiding hitting other persons.

In other words, people walking in populated scenarios can provide rich indirect information about their surroundings, as they are constantly dealing with large amount of high-level information and reacting to it, while following a goal.

A robot moving to a certain destination could identify behaviors of humans and detect a leader, someone moving along a typical pattern that would pass close to the destination point. After identification, the robot could follow people along that path, as "moving with the flow", relying more and more on people in front of it, as leaders or as part of a swarm.

A similar approach has been developed by [12]. With the difference that the main goal of that work is to implement a human-like motion behavior, and the choice of a leader is deterministic, based on the motion direction of the subject regarding the trajectory planned by the robot. Although such technique can be used for the same objectives that are proposed in this work, in some environments the algorithm may fail to find a leader as the extrapolation of the initial movement of a candidate may not match his/her actual goal, as illustrated in Fig. 1.

In this paper, however, the choice of a leader is performed probabilistically, using the Growing Hidden Markov Models (GHMM) technique, an extension to the HMM capable of learning the models parameters and structure in an incremental fashion. This technique not only provides a prediction of the future states of a moving agent, but also its goal, which allows the robot to plan further ahead, with a probabilistic knowledge of the goals pursued by the moving persons.

The global path planning is implemented with the

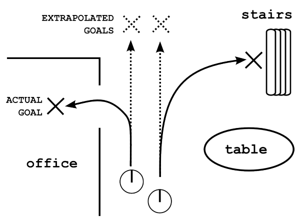


Fig. 1. Comparison of goal prediction using simple extrapolation and the actual goals of two moving agents

RiskRRT algorithm [7], which takes into account the risk of collision with dynamic obstacles while generating the tree. The use of this algorithm guarantees that the robot can find path solutions in dynamic environments when a leader is not found.

In section II the technique to choose a leader in dynamic environment is presented, followed by the explanation of a leader following technique, in section III. The experiments and their results are presented in section IV, and after that, the conclusions of this work are presented in section V.

II. CHOOSING THE LEADER

As the proposition of this work is to take advantage of the movements of a person, the method used to choose which person to follow plays a major role. Here, the choice of a leader is implemented based on the distance between the goal given to the robot and the predicted goal of the leader candidate.

The prediction of the motion and goal of a leader is not an easy task in dynamic environments. A simplistic approach may extrapolate the current orientation and speed of a moving person in an attempt to determine his/her likely goal. But due to environment or dynamic restrictions, the subjects can completely change their trajectory in the subsequent time steps, invalidating the extrapolation, as shown in Fig. 1. Here the dashed lines represent the goal and motion prediction using simple extrapolation or a Kalman Filter. Note however that due to the environment structure and interest points, the actual goal and paths, represented by solid lines, highly differ from the ones predicted using simplistic assumptions.

Knowing that the movement of humans is highly dependent on the environment structure and interest points, recent approaches take advantage of the typical paths in order to make predictions of motions and goals of humans. This approach can overcome the limitations posed by simple extrapolation techniques, as it allows to take into account the structure of the environment as well as the most common motion patterns.

In the current work, in order to predict the goal of moving agents, the Growing Hidden Markov Models (GHMM) algorithm is used [5]. It implements an approach where the *learning* and *prediction* phases are on-line concurrent processes, resulting in a *learn and predict* paradigm. The structure of the GHMMs are the same as the regular HMMs, with the difference that as new observations sequences are incorporated into the model, the transition structure and the

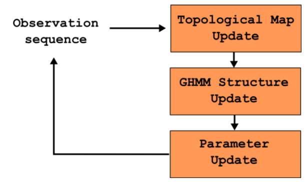


Fig. 2. GHMM learning algorithm overview (adapted from [5])

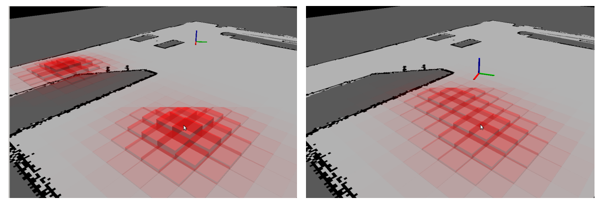


Fig. 3. Two instants in the goal prediction of a dynamic object. The height of the bars is proportional to the probability of a cell to be the final goal

number of states can change, updating the model, as seen in Fig. 2.

The GHMM algorithm consists of the use of the Growing Neural Gas (GNG) algorithm [13], used to estimate the model structure as well as the transition probabilities of a Hidden Markov Model (HMM). As the algorithm is adaptive, it is capable of creating or removing states to cope with new observations.

One very important aspect of the GHMM algorithm is that it is based on the hypothesis that moving agents always try to reach a goal in the environment. Therefore, each goal in the scenario has a different HMM associated to it. As the intention of the current work is to predict the most likely goal of a moving agent, the GHMM inherently provides a direct solution to that problem.

III. FOLLOWING THE LEADER

Once a moving agent has been detected as a leader, due to a similar goal with the robot's goal, there still remains the problem on how to follow the person. Simplistic approaches that try to use the current leader position as a subgoal may bring the robot to situations where undetected obstacles are present between the robot and the leader.

Following the leader path, generating subgoals along the tracked trajectory can overcome situations were obstacles that were not detected appear between the robot and the leader. However, if the robot is not close enough to the leader, it can lose track from the person or different agents that are not leaders may appear in the scene, blocking the tracked path.

Besides that, the motion algorithm has to be able to maintain a navigation solution even in situations where no leader is found, or after a leader is lost due to sensor occlusions or scene exit by the leader being followed. These constraints are addressed with the use of a variation of the RRT algorithm that efficiently explores the free space while at the same time takes into account the risk of collision with moving agents in a scenario.

A. Risk Rapid-exploring Random Tree

The Risk Rapid-exploring Random Tree (RiskRRT) is a variation of the classic RRT algorithm presented by [14] developed for navigation in dynamic environments. It takes into account the risk of traveling along generated paths according to predicted objects' motion. It combines a part dedicated to perception (of static and moving obstacles) with another for planning trajectories. Navigation and planning are done in parallel.

The configuration-time space is searched randomly, and a tree T is grown from the initial configuration all over the configuration space. The algorithm chooses a point P in the configuration space and tries to extend the current search tree toward that point.

The points P are randomly sampled on the map, but at the beginning, and then once every 100 times, the goal itself is chosen; this bias, which has been empirically set, speeds up the exploration toward the goal. The node chosen for extension is the most *promising* node: all the nodes in T are weighted taking into account the risk of collision and the estimated length of the total path:

$$\tilde{w}(q_N) = \frac{L_\pi(q_N)}{dist(q_0, q_N, P)} \tag{1}$$

$$w(q_N) = \frac{\tilde{w}_(q_N)}{\sum_q \tilde{w}_q} \tag{2}$$

At numerator, the likelihood of $\pi(q_N)$ is normalized with respect to the length of the path N; at denominator, $dist(.)$ is the sum between the length of the path from the root q_0 to the node q_N (which is known) and an estimation of the length of the path to P.

The weights are normalized over the set of nodes in the tree (2). The node to grow next is then chosen taking the maximum over the weights or drawing a random node proportionally to the weight. The new node q_+ is obtained applying an admissible control from the chosen node q toward P. The weight of q_+ is computed. If $w(q_+) \geq w(q)$ the tree is grown again from q_+ toward P otherwise another point is sampled from the space.

The likelihood of each partial path can also be expressed as the multiplication of the independent probability of collision with static (P_{cs}) and dynamic components (P_{cd}):

$$L_\pi(q_N) = \prod_{n=0}^{N}(1 - P_{cs}(q_N)) \cdot \prod_{n=0}^{N}(1 - P_{cd}(q_N)) \tag{3}$$

The prediction approach for forecasting the position of moving obstacles in the near future is done using the GHMM predictor. With the information of probable occupied positions in the future, the robot can anticipate the behavior of the agents. The selection of the best trajectories is done by taking into account the probability of collision for each path.

The probability of collision, or risk, can be seen in this case as a measure of the feasibility of a path, with the maximum accepted risk specified as a threshold. The RiskRRT algorithm also takes into account the interactions among humans so the robot can behave in a socially acceptable way. Therefore, the risk function must rely on safety but also in human friendly navigation.

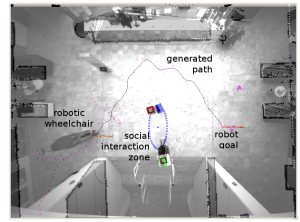


Fig. 4. RiskRRT algorithm avoiding a social interaction zone

The current version of the RiskRRT algorithm also takes into account the interactions among humans in order to behave in a socially acceptable way. The robot behavior should follow social conventions, respecting proximity constraints, avoiding people interacting or joining a group engaged in conversation without disturbing them. Therefore, the risk function must rely on safety but also in human friendly navigation.

After these extensions the "probability of success" calculated for every partial path is given by the probability of not encountering a collision along the path and not entering a social interaction zone. For more details about this method, refer to [15]. Fig. 4 depicts a situation where the RiskRRT algorithm generates a path that avoids the social interaction between two persons.

B. Algorithm

The developed program to follow a leader is shown in III-B. The program starts after receiving a desired goal for the robot, which is used to initiate the RiskRRT algorithm. At the same time, the tracking program starts, in order to detect moving agents in the scenario. The detected agents are fed into the GHMM predictor, which outputs the prediction of a goal for each person.

After that, the Euclidean distance between the robot's current goal and each subject's predicted goal is computed. If this distance falls inside a threshold, the corresponding subject is chosen as a leader.

In the case a leader is found, the robot starts to track its path, but does not immediately starts to follow the leader. This only takes place once the elected leader is closer to the goal than the robot. The reason for this criteria is to avoid situations where the robot would have to move away from the goal in order to follow the path of a leader.

Once the criteria to start following a leader is satisfied, the robot calls an $update$ routine in RiskRRT planning algorithm. This routine causes the RiskRRT algorithm to use the new subgoal instead of the original goal, in order to find a path.

As a result, the algorithm explores the open space and finds a path that poses the lesser risk to bring the robot to the chosen subgoal, which lies over the leader path. The $getSubGoal$ routine decides when to pass a new subgoal of the path to the RiskRRT algorithm, based on the robot

Leader Follower

```
procedure leader_follower
    goal ← readGoal()
    RiskRRT.init(goal)
    while goal not reached do
        agents ← Tracker()
        goalPred ← GHMM(agents)
        for i = 1 → agents.size() do
            d ← Distance(goal, goalPred[i])
            if d < thresh then
                foundLeader = true
                leader = i
                return
            else foundLeader = false
            end if
        end for
        if foundLeader = true then
            path ← trackPath(leader)
            if leader is closer then
                subgoal ← getSubGoal(path)
                RiskRRT.update(subgoal)
            end if
        else
            RiskRRT.update(goal)
        end if
    end while
end procedure
```

distance to the previously passed subgoal. This sequence of steps makes the robot follow the tracked leader's path.

This sequence of steps makes the robot follow the tracked leader path. The use of the Risk-RRT algorithm to reach and follow a leader's path has two main advantages. Firstly, it provides a reliable method to navigate until the leader's path start, since in a dynamic environment the space between the robot and its first subgoal may be occupied by moving agents or static obstacles.

In second place, once the robot reaches and starts to follow the leader path, the algorithm is capable of reusing nodes of its exploration tree to efficiently generate new paths for each new subgoal received by the *update* routine. The reuse of previously generated nodes, reduce the computational load of the algorithm, while still taking into account the risk of navigation.

Finally, in the case that a leader is not found, or the current leader is lost, the *update* routine sends once again to the RiskRRT algorithm its final goal, as chosen at the beginning of the program.

IV. EXPERIMENTS

The experiments were performed using several independent modules using the Robot Operating System (ROS) [16]. The modular architecture provided by ROS gives a series of advantages when implementing experiments, as modules can be added or removed, allowing the test of different techniques without the need of modification of the remaining modules.

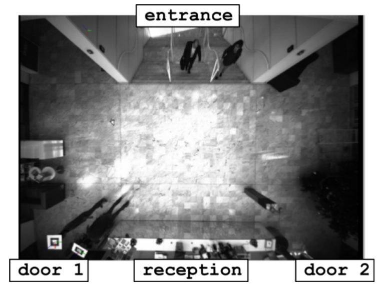


Fig. 5. INRIA Rhône Alpes entrance hall

Fig. 6. Subjects wearing hats with markers and camera with wide angle lens at INRIA's hall

To conduct the experiments, the main hall of INRIA Rhône Alpes has been chosen (Fig. 5). It is an interesting choice as it has a large flow of people during different times of the day, entering and leaving the building during lunch hours and at the beginning and the end of a working day. These conditions allow an easy understanding of the typical paths present in the scenario.

A. Real Data Acquisition

The trajectories used in the experiments to test the robot capability to chose and follow a leader are real human trajectories, that were previously tracked and recorded. An overhanging camera with wide angle lens provides an overview of the test area. The implemented tracker is based on the work of [17]. In the current work, fiducial markers were worn as hats by subjects in order to provide a robust and fast deployment tracker system, as shown in Fig. 6.

The GHMM is trained using a set of the real data acquired with the tracking system. Volunteers were asked to move naturally among interest points in the environment, as the entrance of the hall and the two doors. Fig. 7 shows a sample of the these trajectories.

B. Test Scenario

Two types of tests were conducted, one that evaluates the leader detection technique when several subjects move close to each other (Fig. 8), and another test that evaluates the advantage of the proposed technique to avoid moving agents (Fig. 9). The human perception and detection are performed by a second computer, using the techniques mentioned in the previous subsection.

The tests were conducted using a simulated robot, while the scenario agents represent real data recorded from the motion of humans. The scenario is simulated using

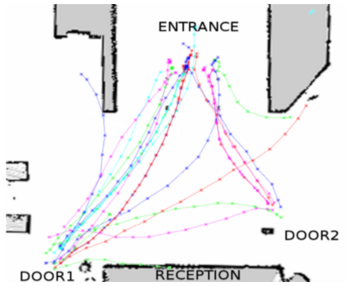


Fig. 7. Real data used in the GHMM initial training

PLAYER/STAGE and the robot has a copy of the environment map and localizes itself based on its odometry and a simulated laser range finder. The robot is represented as a light gray rectangle, and starts in the upper center of the scenario. The obstacles are colored dark gray and encompass walls, desks and sofas.

The circles are the persons detected by the overhanging camera, and the triangles are their respective predicted goals. They have a letter associated to identify their colors (Red, Green and Blue). The robot goal is marked as an X, located at the lower left of the test area. Finally, the dots represent the RiskRRT tree nodes and the solid line is the best path found by the planning algorithm.

C. Leader Detection Test

In the first test, shown in Fig. 8, three humans start to move just in front of the robot, and pursue one different goal each. After some iterations, as the subjects start to move in the scenario, the prediction algorithm gives an estimation for two of them (red and green). Based on that estimations, the leader following algorithm makes the choice of following the red subject, as its predicted goal lies within a distance threshold from the robot's goal.

Once that the leader is closer to the goal than the robot, by an empiric factor, the planned trajectory is computed from the robot to the subgoal. This corresponds to the first position of the chosen leader's path, and the robot starts to move along the trajectory taken by the human, in the direction of its desired goal.

D. Dynamic Agents Avoidance Test

The objective of this test is to evaluate the benefits of following a leader in order to avoid other dynamic agents and is shown in Fig. 9. The way the robot selects and follow a leader occurs in the same fashion as in the previous test. The robot goal is again in the left bottom corner of the image, but here there are now two humans that move from the door to the stairs, in the opposite direction of the robot's desired trajectory.

After the leader is chosen, using the same approach as in the previous test, the robot starts to follow him/her. As the leader approaches the two humans moving in the opposite direction, they naturally give room for him/her to pass. As the robot is closely following the path taken by the leader, it is able to continue to move without the need to take evasive measures to avoid the two incoming persons. As a result, the robot benefits from a straight trajectory toward its goal.

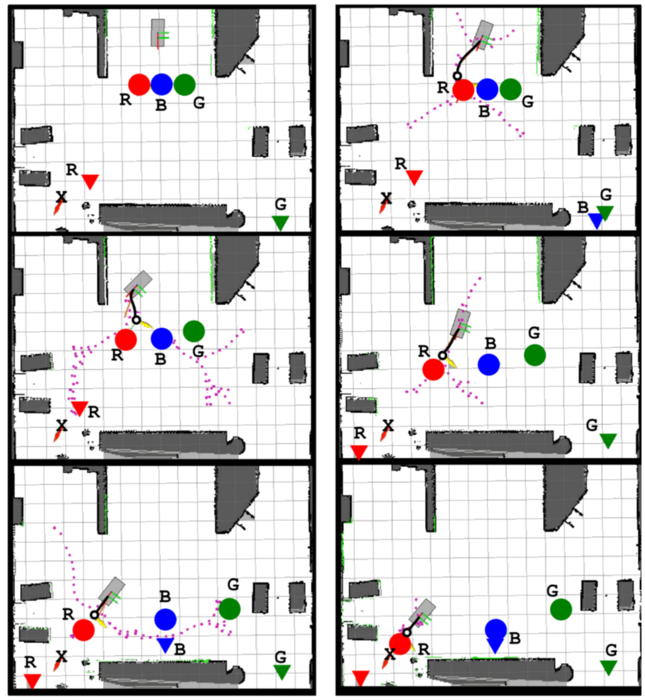


Fig. 8. Results of a typical test of leader detection and following

E. Discussion

The tests assessed the capability of the system to predict the goal of real moving agents, as well as the ability of the designed algorithm to properly follow a chosen leader, while avoiding other dynamic agents.

Results show that the leader following algorithm makes a proper choice of a leader, based on a probabilistic approach for goal prediction, even when the initial movement and is not directed toward his/her goal. This is an important advantage of a probabilistic approach for goal detection, based on previous knowledge of the most common trajectories in the environment.

The navigation technique employed to follow the leader's path continuously explore the surrounding space for alternative trajectories. In the case the leader is lost of the path being followed becomes blocked, the branches that were generated in different directions can be used to find a new path, without the need of replanning from scratch. This is a very important characteristic while navigating in dynamic environments, specially due to time constraints.

The resulting behavior is that the robot is able to follow the leader while at the same time exploring the open spaces to accommodate new and unpredicted situations.

In the second test scenario, the advantages of following a leader in a dynamic environment becomes evident. Classical approaches that would attempt to plan a trajectory taking into account the predicted motion of the incoming humans would fail to find an optimal solution, as a straight line to the robot's goal would be blocked.

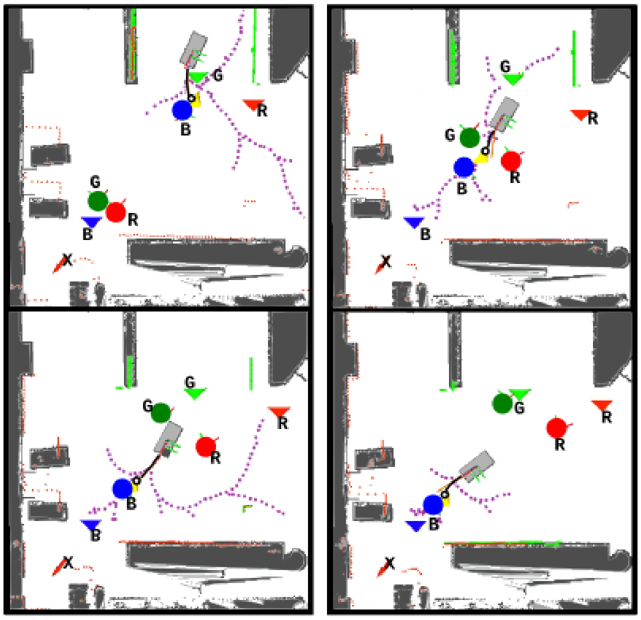


Fig. 9. Leader following allows the robot to avoid two incoming persons

However, as the robot follows a human that is able to correctly assume that the persons moving in the opposite direction will adapt their movement to avoid a collision, it is able to follow a straight trajectory to the goal. The result of this experiment clearly shows the benefit of the proposed technique, as the robot follows an optimal trajectory as a consequence of following a leader that has a better understanding on how to behave in such situations.

V. CONCLUSIONS

This work presents a method to take advantage of human motion in dynamic environments by selecting and following a leader. Its has two main contributions. The first is the methodology used to leader selection, which takes into account the typical paths in an environment and provides a probabilistic inference of subject's goal. The second is the modification of an algorithm designed for motion planning in dynamic environment, in order to adapt it to the task of leader following, while maintaining its original characteristics.

Tests used real data for the leader selection part, while the leader following algorithm was tested in a simulated and real environment. The results validated the proposed approach, with the robot being able to properly identify leaders among several subjects and follow him/her until its desired goal. It is our belief that the prediction algorithm can be further improved if it takes into account not only the position of the agents but also their speed and orientation, making it able to anticipate even more the future motion and goals of persons.

New ideas arose throughout this work, as the possibility to find leader that help the robot in portions of its path, which would required a more refined technique of leader selection, with the robot alternating between aided and unaided navigation.

Experiments will continue in different scenarios, with more tests in specific situations as leader obstruction/loss and also crowded environments. Currently a crowd simulator framework is in development, where extensive testing can take place and the impact of the robot behavior in the humans around can be taken into account.

ACKNOWLEDGMENT

We thank Dizan Vazques, Jorge Rios-Martinez and Arturo Escobedo for very useful insights and discussions.

REFERENCES

[1] K. Mombaur, A. Truong, and J. Laumond, "From human to humanoid locomotionan inverse optimal control approach," *Autonomous Robots*, vol. 28, pp. 369–383, Dec. 2009. [Online]. Available: http://www.springerlink.com/content/c571474575ux1t52/

[2] C. Tay and C. Laugier, "Modelling smooth paths using gaussian processes," in *Proc. of the Int. Conf. on Field and Service Robotics*, Chamonix, France, 2007.

[3] K. Kim, D. Lee, and I. Essa, "Gaussian process regression flow for analysis of motion trajectories," in *2011 IEEE International Conference on Computer Vision (ICCV)*, 2011.

[4] D. Ellis, E. Sommerlade, and I. Reid, "Modelling pedestrian trajectory patterns with gaussian processes," in *2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, Oct. 2009, pp. 1229–1234.

[5] D. A. Vasquez Govea, T. Fraichard, and C. Laugier, "Growing hidden markov models: An incremental tool for learning and predicting human and vehicle motion," *International Journal of Robotics Research*, vol. 28, no. 11-12, pp. 1486–1506, Nov. 2009.

[6] M. Bennewitz, W. Burgard, G. Cielniak, and S. Thrun, "Learning motion patterns of people for compliant robot motion," *The International Journal of Robotics Research*, vol. 24, no. 1, p. 31, 2005.

[7] C. Fulgenzi, A. Spalanzani, and C. Laugier, "Probabilistic motion planning among moving obstacles following typical motion patterns," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, 2009, pp. 4027–4033.

[8] S. T. O'Callaghan, S. P. N. Singh, A. Alempijevic, and F. T. Ramos, "Learning navigational maps by observing human motion patterns," in *2011 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2011, pp. 4333–4340.

[9] S. Sehestedt, S. Kodagoda, and G. Dissanayake, "Robot path planning in a social context," in *2010 IEEE Conference on Robotics Automation and Mechatronics (RAM)*. IEEE, June 2010, pp. 206–211.

[10] D. Helbing, I. Farkas, and T. Vicsek, "Simulating dynamical features of escape panic," *Nature, Vol. 407, pp. 487-490, 2000*, 2000.

[11] P. Trautman and A. Krause, "Unfreezing the robot: Navigation in dense, interacting crowds," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, Oct. 2010, pp. 797–803.

[12] J. Müller, C. Stachniss, K. Arras, and W. Burgard, "Socially inspired motion planning for mobile robots in populated environments," in *Proc. of International Conference on Cognitive Systems*, 2008.

[13] B. Fritzke *et al.*, "A growing neural gas network learns topologies," *Advances in neural information processing systems*, vol. 7, pp. 625–632, 1995.

[14] S. M. LaValle and J. J. Kuffner Jr, "Randomized kinodynamic planning," *The International Journal of Robotics Research*, vol. 20, no. 5, p. 378, 2001.

[15] J. Rios-Martinez, A. Spalanzani, and C. Laugier, "Understanding human interaction for probabilistic autonomous navigation using Risk-RRT approach," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, Sept. 2011, pp. 2014–2019.

[16] M. Quigley, B. Gerkeyy, K. Conleyy, J. Fausty, T. Footey, J. Leibsz, E. Bergery, R. Wheelery, and A. Ng, "ROS: an open-source robot operating system," in *ICRA Workshop on Open Source Software*, 2009.

[17] J. Snape, J. van den Berg, S. Guy, and D. Manocha, "Independent navigation of multiple mobile robots with hybrid reciprocal velocity obstacles," in *Proceedings of the International Conference on Intelligent Robots and Systems (IROS)*, 2009, pp. 5917–5922.

Session III

Behavioral modeling and Human/Robot Interaction

- **Keynote speaker : Hiroshi Ishiguro (Professor of Osaka University, Japan and Fellow of ATR, Ishiguro@sys.es.osaka-u.ac.jp)**
  Title : From teleoperated androids to cellphones as surrogates

- Title : **Multi-sensors engagement detection with a robot companion in a home environment.** Authors : W. Benkaouar and V. Dominique

- Title : **Social Inclusion of Senior Citizens by a Teleoperated Android : Toward Inter-generational TeleCommunity Creation.**
  Authors : R. Yamazaki, S. Nishio, H. Ishiguro, T. Minato, M. Norskov, N. Ishiguro, M. Nishikawa and T. Fujinami

- Title : **Can smart rollators be used for gait monitoring and fall prevention ?** Authors : C. Dune, P. Gorce, J.-P. Merlet

Keynote Speaker : **Hiroshi Ishiguro**
(Professor of Osaka University, Japan and Fellow of ATR, Ishiguro@sys.es.osaka-u.ac.jp)

From teleoperated androids to cellphones as surrogates

Abstract : In order to understand the meaning of human presence, we have developed Geminoid which is an teleoperated android of myself. With the android, we could learn how people can adapt the new media. Based on the knowledge, we have recently developed a simpler teleoperated android with the minimal humanlike appearance. The new android is called Telenoid. People can easily adapt to Telenoid and enjoy conversations by using it. Further, we are remaking it with a cell-phone size. It is called Elfoid. We believe the new type of cell-phone can transmit our presence to distant places and changes our life style again.

Multi-sensors engagement detection with a robot companion in a home environment

Wafa Benkaouar[1] and Vaufreydaz Dominique[1,2]
[1] PRIMA Team - Inria/LIG/CNRS, 655, avenue de l'Europe - 38334 Saint Ismier Cedex
[2] Université Pierre Mendès-France, BP 47 - 38040 Grenoble Cedex 9
Wafa.Benkaouar@inria.fr, Dominique.Vaufreydaz@inria.fr

***Abstract*— Recognition of intentions is an subconscious cognitive process vital to human communication. This skill enables anticipation and increases the quality of interactive exchanges between humans. Within the context of engagement, i.e. intention for interaction, non-verbal signals are used to communicate this intention to the partner. In this paper, we investigated methods to detect these signals in order to allow a robot to know when it is about to be addressed. Classically, the human position and speed, the human-robot distance are used to detect the engagement. Our hypothesis is that this method is not enough in the context of a home environment. The chosen approach integrates multimodal features gathered using a robot equipped with a Kinect. The evaluation of this new method of detection on our corpus collected in spontaneous conditions highlights its robustness and validates use of such a technique in real environment. Experimental validation shows that the use of multimodal sensors gives better precision and recall than the detector using only spatial and speed features. We also demonstrate that 7 multimodal features are sufficient to provide a good engagement detection score.**

I. INTRODUCTION

Social signal processing and affective computing have emerged as new areas of Computer Sciences over the last ten years (R. Picard [1]). These new areas explore the multimodal aspect of the human communication in order to develop more natural interaction between humans and computers or robots.

Speech is an important channel for communication and requires signal processing as well as semantics and linguistics domain. In addition to the semantics of speech, emotions, and the inner goals of humans are conveyed by other channels: body, gesture, etc. The research community is increasingly interested in this non-verbal (NV) communication. Recognition of intention is a basic skill acquired by infants early in their development. Vernon in [2] states that one of among other skills, the perception of the direction of the attention of others is crucial for the infant to master social interactions. The perception of intentions and emotions, present in newborn infants, helps to set their "preparedness" for social interaction. Human cognition has a high part of anticipation, allowing to read the intentions, and guessing the goal in order to react quickly to some stimulus.

Companion robots should also be able to detect the intentions of humans in order to adapt their behavior during interactions with humans. For natural human-robot interaction, the intention reading of the behavioral cues from an individual is fundamental.

Our goal for this research is to investigate techniques to detect and recognize signals for non-verbal communication reflecting intentions and in particular the engagement of a human with a robot. We define engagement as the phase during which one expresses, with NV cues, the intention of an interaction. Perception of engagement refers to the perception of the intention for interaction. Engagement is a real question especially when it comes to environments such as the work place or home; where people are not familiar to interacting with robots as shown in [3]. Engagement is fundamental for communication between human users and interactive robots.

Classically, the criterion for a user's engagement are spatial and speed information between the user and the communicant interface [4]. These studies made a simple assumption: if the user is close to the robot, he wants to interact. This detector of engagement based on distance and sometimes speed of the human gives good results for kiosk-like interfaces, but for an assistant living robot in real-life, close distance does not necessary signal a desire for engagement. Indeed, many times during the day one can pass in front of the refrigerator without the wish to open it. In the same vein, a robot in order to have more human acceptable behavior should be able to detect when it is about to be solicited, and to anticipate this interaction. In the context of a companion robot the proximity of the robot with a person should not be a continuous trigger for engagement. Other criterion can be taken into account such as the posture, the sound and other features described below.

We propose a multimodal approach for detecting engagement using the Kinect© sensors from Microsoft [5] to improve re-usability, and to enable us to build a detector deployable in real-life situations. From literature, in particular the cognitive sciences literature, we found some cues to measure the engagement of a person into an interaction. Hence, we propose to take into account the spatial information, body pose, frontal face detection, speech detection and sound localization in order to model the engagement detection system. An important contribution of this work is the multimodal dataset gathered from the robot point of view. Optimization of the acquisition process was needed to limit information loss and to facilitate the synchronization of the multimodal data. This corpus offers a realistic framework to test our hypothesis.

Evaluation using Multi-class Support Vector Machine and Artificial Neural Networks techniques to classify the features computed from the dataset have given significantly better results in the multimodal condition when compared to a unimodal spatial condition. We show that the spatial and speed features can be improved for engagement detection in a home environment. A subset of 7 multimodal features is proposed for the engagement detection task.

In the following sections, we first develop an overview of the approaches concerning engagement models in cognitive sciences and human-robot interaction. Then, we describe the recording of a robot centered corpus in a home environment and features we can extract from it. Finally, classification and space reduction evaluations are depicted to validate our hypothesis.

II. From Cognitive Sciences to Human-Robot Interaction

Humans are endowed by range of abilities called social intelligence [6]. They include the ability to express and recognize social signals produced during social interactions like agreement, politeness, empathy, friendliness, conflict, etc. They are coupled with the ability to manage these signals in order to get along with others while winning their cooperation.

An intelligent agent is commonly defined as an agent who perceives, learns, and adapts to the world. Social signals are manifested through a multiplicity of non-verbal behavioral cues including facial expressions, body postures and gestures, vocal outbursts like laughter, etc. , which are aimed to be analyzed by signal processing technologies, or automatically generated by synthesis technologies.

Social sensible computer systems and devices which are able to adapt their response to social signals in a polite, non-intrusive, or persuasive manner, in real-time, are likely to be perceived as more natural, efficient and trustworthy. In the context of assistance to personal living in a home environment, social adequacy seems to be crucial for the acceptance of a robot companion.

A. Intentionality in Human-Machine Interaction

Recognition of humans' intentions, goals and actions is important in the improvement of non verbal human-robot cooperation. Intention recognition is defined in [7] by the process of estimating the force driving humans actions based on noisy observations of humans' interaction with his environment. The DARPA/NSF in its final report on Human-Robot Interaction [8] recommends to improve the models of human-robot relationship and in particular to work on the intentionality issue.

In his study, Knight [9] points the importance for a robot to convey and to detect intentionality. It helps to clarify current activity and to anticipate the goals. Learning from the human engagement, the robot would be able to anticipate the interaction and also to learn adequate moments when the robot itself can engage an interaction. In [10] engagement is defined as the process by which two (or more) participants establish, maintain and end their perceived connection during interactions they jointly undertake.

Different modalities are used in the social signal analysis in computer science research field. The modality channels through which non-verbal communication can be measured are the audio, face, posture and gesture, the physiologic aspects, clothing, gender, age, etc. We focused on non-invasive aspect of social signal perception and present the modalities used in order to detect engagement.

B. Body Pose and Proxemic features

A way of detecting engagement would be to consider only proxemic metrics. Classical features included in proxemic features are the relative position of the individual to the robot and their relative speed. For a collaboration to be successful, the distance between the robot and the human should be optimum and the speed controlled. In [4], it is proposed to recognize intentional actions using relative movements of a human to a robot. Koo uses an Infrared sensor embedded on the robot to track and estimate the velocity of a person. He then infers intentional actions such as approach and depart using Hidden Markov Models (HMM) and position dependent model.

Spatial metrics can be useful measures to describe role, attention, and interaction. Psychologists have proposed many models to describe body pose metrics and their associated meaning. An overview of these metrics is presented in [11]. There is no consensus on the meaning and the emotional characteristics of a posture. Psychologists such as Hall, Mehrabian [12] and Schegloff [13] have proposed some metrics that have been used in computer assisted analysis of posture.

Posture is difficult to measure and evaluate using computer vision. Nevertheless, with the apparition of the Kinect sensor and other real-time 3D pose reconstruction techniques, we are able now to evaluate the pose of a person.

C. Audio Features

Pantic in [14] lists some features into the audio signal that can be used to spot basic emotions such as happiness, anger, fear and sadness. It can be agreed on, that some audio features such as pitch, intensity, speech rate, pitch contours, voice quality and silence are good parameters to classify the emotional state of an individual. Considering the recognition of the engagement in an interaction, only few papers in the literature use audio features in a multimodal frame. [15] proposes an engagement estimator using head pose associated to audio features in a face-to-face conversational agent interaction.

Even if we do not realize it, we are able to localize roughly a sound source. Sound spatialization is not often used for affect detection, but [16] invokes its interest in attention or focus estimation.

D. Facial Features

Concerning the engagement, the orientation of the head and the gaze seem to be crucial. As shown in [17] a speaker can be detected more easily with the combination of different features relative to the orientation of the face such as a mouth sensor. Face detection is already a first cue of interaction. The orientation of the face toward the interface seems to be a sign of attention.

III. A Corpus for Engagement with a Robot

A part of this study was to record a multimodal dataset including interaction with a robot companion enhanced with a Kinect device.

This section presents the method used to build the dataset needed to test our hypothesis.

A. The need for a dataset

In the context of a companion robot, we want to work with consumer devices in a natural environment. Even though the tendency is to use more and more physiological sensors (such as R. Picard's pulse bracelet Cardiocam, etc.), physiological devices are still invasive and expensive for the users to be released widely.

In order to evaluate our hypothesis, we confronted it to data. In the context of robot companion, the sensors considered are commonly microphones, video sensors, depth sensors, lasers... There exist datasets in the field of social signals processing dealing with non-verbal communication using multi sensors. Available datasets for affect recognition are unfortunately more often for face-to-face interaction with persons sitting and interaction with the speech only. The SSPNet association provided the SEMAINE-DB dataset [18] where several persons have been recorded in a face-to-face speech interaction. This database is suitable for a desktop environment for interaction with virtual communicant agent. Unfortunately, this dataset suits less human-robot interaction, especially if the non verbal cues of social signal that are involved in the engagement of interaction are more diverse than the facial expression and the speech characteristics. Other corpora exist that use the Kinect sensors and 3D information, such as [19] which presents a Cam3D dataset centered on facial and hand movement associated with audio recording. Yet, the proposition of a robot centered dataset for multimodal social signal processing has not been made.

1) Kompai robot: The Kompai robot has been lended by our partner Robosoft[1], allowed us to record our corpus. The Kompai robot, Figure 1, aims at helping elders and dependent persons. It is composed of a RobuLAB mobile platform containing the wheel actuators, obstacle detection system, manual remote control facilities, etc. The mobile platform is topped by a tablet serving as interface with the user, a pair of microphones, a motorized webcam and a speaker, to which we added a Kinect sensor.

In our recordings, we gathered every sensor available like the head-mounted webcam of the robot used to record videos during the experiment.

2) Kinect Sensor: The Kinect sensor is composed of several components which are represented in the Figure 2. Advantages of using such a sensor are its consumer price and its growing utilization in computer vision assisted system. During corpus gathering, we recorded several streams from the Kinect:

- *Depth Camera (using Infrared laser)*: the depth range is limited from 80 centimeters to 4 meters with a 2 millimeters accuracy.
- *Skeleton Tracking*: the Kinect supports up to two skeletons being tracked at the same time. Only the tracked skeletons with a high confidence score are stored in the dataset.
- *RGB Camera*: the resolution of the RGB image is 640x480 pixels by default. The RGB horizontal field of view is of 62.0 degrees.
- *Microphone Array*: the array is composed of four aligned microphones. It provides an angle of a detected sound with a confidence in the Kinect reference frame. It also outputs the more stimulated beam by the sound source.

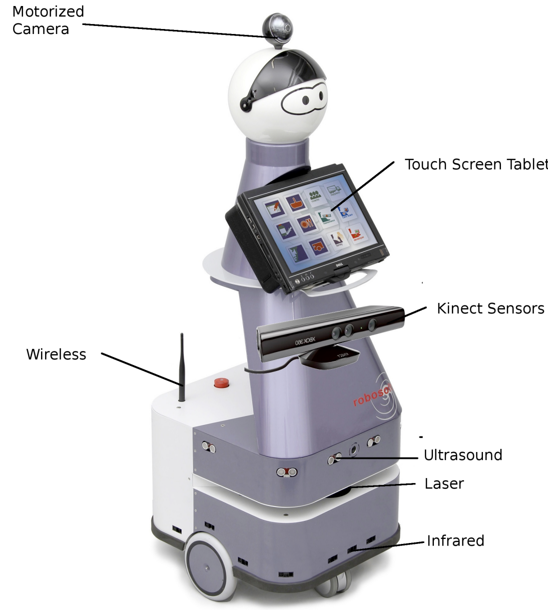


Fig. 1. The Kompai Robot from Robosoft.

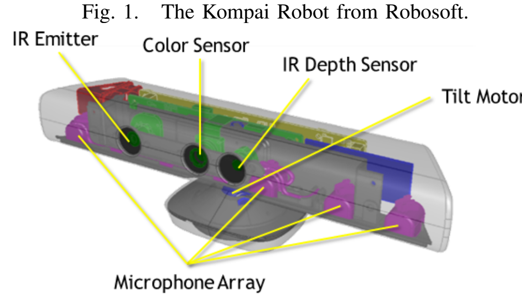


Fig. 2. Components of the Kinect Sensor [5]

B. Features extraction

The recorded data are presented in the table I. Some of them were analyzed to extract features for the engagement detection.

[1] http://www.robosoft.fr/

Data	Sensor	Maximal Frame Rate
Telemeters distances	Kompai	12.5Hz
Ultraound distances	Kompai	12.5Hz
Audio	Kinect	16kHz
Sound Source Beam and Position	Kinect	8Hz
Skeletons	Kinect	30Hz max
RGB Video	Kinect	30Hz max
Depth Video	Kinect	30Hz max
RGB Video 2	Webcam	15Hz
Button Press	Tablet	-

TABLE I

DATA RECORDED, ASSOCIATED SENSORS AND FRAME RATE

1) Features selection: Using all available sensors (see previous section), we must define which features to extract from the data. Looking at the literature, we decided to compute the following features:

- Using the laser telemeter, we can extract, in the frame of the robot, the *x* and *y* position, the *dx* and *dy* speed, and *dist*, the distance to the robot. These features are computed using a background subtraction on the telemeter input and a Kalman Filter to track moving people. This set of features will be, as expressed formerly in the article, our comparison point with the state-of-the-art technique for engagement detection. We named it the *telemeter* condition.
- Using the microphone array, we can add acoustic features: angle, activated beam and confidence of the acoustic source localization and speech activity detection using [20].
- Considering that facial information are important, we computed using OpenCV [21] in the RGB video stream from the Kinect *face_x*, *face_y* and *face_size* respectively the position and size of the biggest detected face in the image.
- Stance, hips, torso and shoulders positions and relative rotations depicted by [13] are computed from the tracked skeletons[2] and give 19 features.

Finally, our set consists in 32 features from different modalities captured from the Kompai enhanced with a Kinect.

2) Features fusion and synchronization: There are 3 main fusion techniques for multimodal corpora: data fusion, features fusion or decision fusion. Data fusion is more suitable when data are of the same kind (multiple video streams for instance). Second, fusion at the feature level aims to aggregate features extracted from the various sensors together before attempting to classify. Last, using late decision fusion has some advantages. The computational cost of the training is reduced and the strict synchrony of the inputs is not required since they bring complementary information. Its drawback is the expertise needed or the relative empiricism of the final decision fusion. According to [22], features fusion is considered more appropriate for closely temporally synchronized input modalities, such as speech and lip movements. As we considered that all our modalities synchronously express our engagement, we decided to use this method.

The common dimension of all the modalities is the time. As seen in table I, frame rate of inputs are different. We decided to synchronize all features on a fixed frame rate. Data from the Kinect present a variable frame rate when recording all streams and tracking people and skeletons at the same time. Only telemeters information is cadenced at fixed frame rate 12.5Hz using a micro-controller. We synchronized everything using the current value of features at the telemeter events timestamps.

C. Realistic Dataset

R. Picard in [1] gives five variables that may affect data collection. The first factor is the spontaneity of the expressed emotion. The emotion can be either elicited by a stimulus or asked to elicit (activated or acted). Another influence can come from the environment of the recording, and the question here is that are the emotions expressed and recorded similarly in a lab setting and in a real-life situation? Next question to be considered when recording affective data is: should the focus be on the expression of the emotions or on the internal feeling? The internal feeling would be measured by retrospective interviews of the participants. The awareness factor of the recording is another factor. Indeed, what is the influence of open-recording in comparison with hidden recording on the recorded data? Finally, should the emotion be presented to the subject as the purpose of the experiment or not?

Regarding our matter, the engagement is relatively spontaneous. It is asked to the participant to interact, yet its intention toward the interaction cannot be elicited artificially. The intention will show whenever the participant plan to interact. The participant is explained that the measurement is its reaction while playing the game. The goals of measuring intentions is still hidden, there is no awareness to the recorded factor by the participant. The recording is made in a smart environment, similar to a flat. For many of the participants, this room is new and this can create some fluctuations in the behaviors.

D. Scenarios

In order to test our hypothesis that the position and speed of the person is not enough to detect engagement, we propose to confront with scenarios where users pass close to the robot but with no intention of interaction. The robot is immobile in a waiting attitude.

We want to detect the pre-interaction phase where participant show social signals of their engagement. We made the assumption that these cues were detectable with the sensor that equipped our version of the Kompai robot.

The data were recorded with two different scenarios performed several times by different participants in a homelike environment with a Kompai. The room is similar to a small flat (Figure 3). It is randomly asked to the participant to enter the room by different doors, perform some realistic actions

[2] As we used the Windows version of the Kinect driver, we did not have specific initialization process for skeleton tracking while recording walking people.

and go out. One of the actions is to interact with the robot. The interaction consists in a small flash game on the tablet PC. The other actions were walking, sitting, or pouring water from the sink. Participants were not aware of our intention to measure their engagement with the robot.

1) Scenario "Passing By": In this first scenario, one participant is ask to go through the room twice by different doors (A), (B) or (C). Figure 3 shows the setting of this scenario.

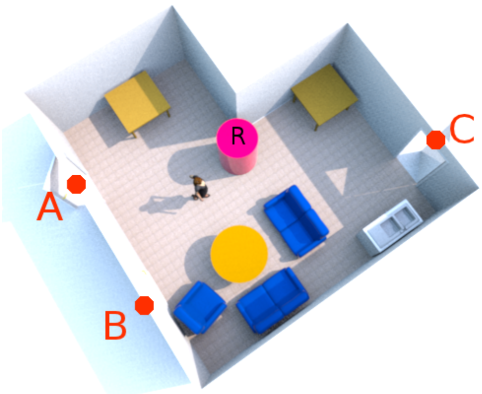


Fig. 3. Scenario 1 *"Passing by"*. A, B and C are access doors. R is the robot.

2) Scenario "Playing cards together": In this second scenario, 3 persons are asked to start a card game in the living-room part of the flat. A telephone placed in the room

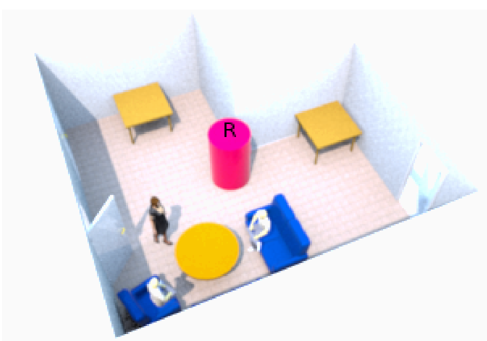


Fig. 4. Scenario 2 *"Playing cards together"*

is used to ask one of the participants to execute an action (gaming interaction with the robot, or using the sink for example). Figure 4 shows this scenario when one of the participant is entering in the room while the other two are already sitting.

E. The dataset in numbers

The recording of the corpus has been made during three sessions of one to two hours. The corpus includes 29 interactions with the robot, made by 15 different participants among more than 50 actions. In real life, all individuals do not express these signals the same way. Some variability has been introduced in the pool of participants. They are from 20 to 35 years old and are female and male. The voice, clothing, posture varies among the participants. The testing data are taken from different sessions of recording. To randomize the attributions of actions for the participants is also a way of controlling certain pattern in the parasite variables that can appear when experimenting with real data. The duration of the interaction also varies from 2 to 10 minutes according to the participant will.

The total size of the uncompressed data set is around 300 GB with more than 150.000 frames of 32 extracted features.

F. Corpus availability

The corpus is not, for now, available. An enlarged version of the corpus will be recorded with more participants and new scenarios. We plan to release it for the research community.

IV. Automatic Labeling

A. Steps of the interaction process

The process of interaction has been described by Sidner and Lee in [10]. They proposed a model in three steps: initiation of interaction (WILL_INTERACT in our labeling), maintenance of interaction (INTERACT) and disengagement (LEAVE_INTERACT). We added two more classes NO-ONE when nobody present and SOMEONE_AROUND when someone is around the robot and does not want to interact.

B. Labeling rules

Our scenarios were defined for helping us in the automatic labeling of the dataset. Before interacting, people are located in blind areas for the telemeters: outside the room or in the game area. Using laser telemeter information, we can detect when someone is moving towards the robot.

The interaction (INTERACT class) appears between the beginning and the end of user clicks on the tablet. The WILL_INTERACT phase preceding the beginning of interaction (first click) is labeled since appearance of a moving object just before the interaction on the robot tablet. In both scenarios, it can be done as people were coming from a blind area for the telemeter: outside for the first scenario, the playing cards area for the second one. LEAVE_INTERACT has been tagged during 5 seconds after the end of interaction. The idea behind this empirical choice is that leaving interaction with the robot is after a short leaving sequence, just like walking away from it. The SOMEONE_AROUND event is labeled when someone is in the room but with no wish of interacting with the robot. When nobody is in the room, it corresponds to the NO-ONE event.

Automatic labeling has been confronted and validated against manual pre-annotation of recorded sessions.

V. Evaluation

We focus on the engagement detection, i.e. on the WILL_INTERACT class. Other classification results are presented but will not be discussed in this paper.

A. *Classification Results*

In order to classify our features, we chose to use two kinds of classical classifications: Artificial Neural Networks (ANN) and Support Vector Machines (SVM) techniques. For these two techniques we built and tested two classifiers one for the multimodal dataset (including the whole 32 features) and one for telemeters condition (5 features).

1) Artificial Neural Networks: The Artificial Neural Network is a multi-layered model with perceptrons. We used the Weka [23] toolbox to perform this classification. The use of ANN is common to infer models from observation. In our case, we suppose that our features can characterize the engagement, the use of ANN technique can help us to test this hypothesis. ANN is a good classifier to build prospective detection especially with large feature vector. Results of the ANN classification are presented in the Table III for the telemeters and the Table II for the multimodal feature set.

Class	Precision	Recall	FPR	Accuracy
No-one	0,95	1,00	0,07	0,97
Will Interact	0,90	0,87	0,02	0,96
Interact	0,84	0,95	0,04	0,96
Leave Interact	0,21	0,01	0,00	0,99
Someone around	0,76	0,41	0,01	0,95
	0,91	0,91	0,02	0,96

TABLE II

RESULTS OF MULTIMODAL NEURAL-NETWORK 5-CLASS CLASSIFICATION.

Class	Precision	Recall	FP-Rate	Accuracy
No one	0,95	1,00	0,08	0,97
Will Interact	0,91	0,77	0,02	0,95
Interact	0,77	0,96	0,06	0,94
Leave Interact	0,00	0,00	0,00	0,99
Someone around	0,75	0,35	0,01	0,94
	0,90	0,90	0,03	0,96

TABLE III

RESULTS OF TELEMETER NEURAL-NETWORK 5-CLASS CLASSIFICATION.

First, these results show that the overall precision and recall of the classifier for our classes is slightly better in the multimodal approach. Concerning the engagement class, WILL_INTERACT, the system returns more relevant event as an engagement in the case of the multimodality and its accuracy is improved. For the engagement detection, in a practical point of view, the accent has to be put on the good performance in terms of recall and a low false-positive rate. The Neural Network classifier gave better recall rate in multimodal condition.

2) Multi-Class Support Vector Machine: Tests using Support Vector Machine were done using the Sklearn toolkit [24]. The results of the 5-classes classification using for the multimodal features are presented in Table IV. For the telemeters classification the results are presented by the Table V. We observe, comparing these tables, that the precision and recall scores for the WILL_INTERACT class are significantly improved by the multimodality. Also, for this same class, the False-Positive rate is higher in the case of the telemeters only. In particular, the aim of this detection was to decrease this rate of misclassifying an event as WILL_INTERACT, hence the system has less chance to predict an interaction when there will not be one and to disturb a user with no intention of interaction.

Class	Precision	Recall	FP-Rate	Accuracy
No one	0,92	0,88	0,11	0,89
Will interact	0,92	0,71	0,01	0,93
Interact	0,54	0,77	0,15	0,84
Leave interact	0,04	0,10	0,03	0,96
Someone around	0,52	0,29	0,02	0,93
	0,78	0,78	0,06	0,91

TABLE IV

RESULTS OF MULTIMODAL SVM 5-CLASS CLASSIFICATION.

Class	Precision	Recall	FP-Rate	Accuracy
No-one	0,68	1,00	0,65	0,72
Will interact	0,80	0,68	0,05	0,90
Interact	0,00	0,00	0,01	0,81
Leave interact	0,00	0,00	0,00	0,99
Someone around	0,76	0,01	0,00	0,93
	0,69	0,69	0,09	0,87

TABLE V

RESULTS OF TELEMETER SVM 5-CLASS CLASSIFICATION.

B. *Feature space reduction*

The Minimum Redundancy Maximum Relevance (MRMR) method [25] has been performed in order to highlight the best features for our detection system. Contrary to Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA), this dimensionality reduction technique has the advantage of selecting the most relevant features instead of building new features by combining the observed ones. MRMR uses mutual information to select features which jointly have the maximal statistical dependency while best characterize the statistical property of a target classification variable. Hence, it could allow discarding some less relevant features in order to optimize the detection of engagement process.

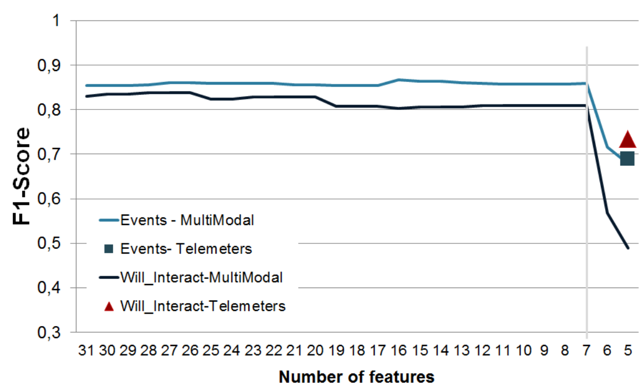


Fig. 5. F1-score evolution while decreasing the number of multimodal features in comparison with the telemeters for all the events and for the WILL_INTERACT event.

Figure 5 shows the impact on the f1-score[3] of the space reduction from 31 to 5 selected features with MRMR.

[3] F1-score is a combination of the *precision* and *recall* values (see `http://en.wikipedia.org/wiki/Precision_and_recall`).

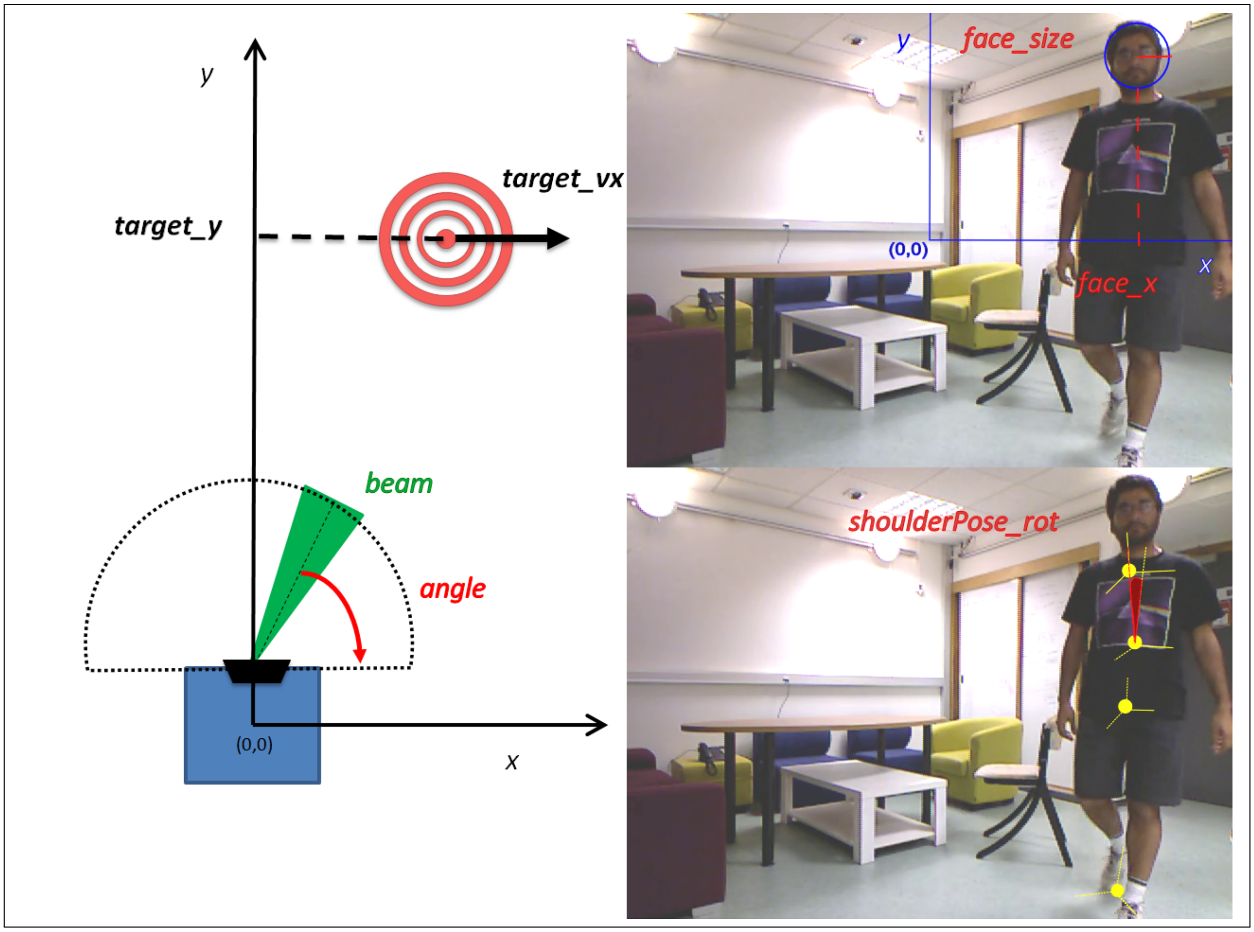


Fig. 6. Minimal multi-modal set with 7 features. The blue square represents the Kompai robot, the black trapezoid the Kinect. *Target_vx* and *target_y* are computed using telemeter information in the robot reference frame. Using the Kinect audio stream, video stream and skeleton tracking, we can respectively extract *angle* and *beam*, *face_size* and *face_x*, and the shoulder rotation (*shoulderPose_rot*).

The performance drops when six features are reached. Before, it remains pretty stable and even non significantly slightly increases along the feature reduction. These results confirm the fact that there are many correlations in the complete feature space. Some of these features seem to be fundamental for a better detection and to keep a precision higher than the telemeters' one.

The first remark on these results is that the 7 highest rated features are coming from heterogeneous modalities. The *shoulderPose_rot* corresponds to the relative orientation of the shoulder in the body, and is extracted from the skeleton information. MRMR classes it as the principal feature. Next, some telemeters information are considered as relevant: *target_vx* and position *target_y*. The *face_size* and *face_x* are respectively the relative size and position of the face in the video of the Kinect. The *beam* and the *angle* are the sound localization features from the Kinect's microphone array. These features are illustrated in Figure 6.

From these results, our intuitions based on cognitive sciences studies of the engagement recognition are comforted. Indeed, the importance of the body pose, such as the orientation of the shoulder is exposed. Position and size of the face in the image show that the person is facing the robot which a priori confirms its engagement. Some moving criteria complete this features list but not all of them. Distance to the robot, y and dy in the reference frame of the robot are not selected whereas x position and speed are significant in our experiment.

VI. Conclusion and Future Work

In this article, we presented our multimodal approach for engagement detection in a homelike environment with a robot companion enhanced with a Kinect. We recorded a multimodal robot-centered corpus for engagement detection following mono-user and multi-users scenarios. In comparison with the usual spatial features set and using this corpus, we increased precision of multimodal engagement detection respectively from 71% up to 87% with recall staying at 90%.

With feature space reduction technique, we highlight the 7 most relevant multimodal features for engagement detection from our features set. Shoulder rotation, face position and size, user distance and lateral speed, sound localization information were found to be coherent with the results on engagement described in cognitive sciences researches.

With a more powerful embedded system and a computation limited to 7 features, we are currently working on a real-time detection on the robot. Prediction of engagement is a first step toward a smoother human-robot interaction.

REFERENCES

[1] R. W. Picard, *Affective Computing*. International Series in Experimental Social Psychology, The MIT Press, 2005.

[2] D. Vernon, C. Hofsten, and L. Fadiga, *A Roadmap for Cognitive Development in Humanoid Robots*, vol. 11 of *Cognitive Systems Monographs*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.

[3] L. Wang, P.-L. P. Rau, V. Evers, B. K. Robinson, and P. Hinds, "When in Rome: the role of culture & context in adherence to robot recommendations," *ACM*, pp. 359–366, Mar. 2010.

[4] S. Koo and D.-s. Kwon, "Recognizing Human Intentional Actions from the Relative Movements between Human and Robot," *Nonlinear Dynamics*, pp. 939–944, 2009.

[5] Microsoft Reasearch, "Kinect for windows programming guide,", `http://msdn.microsoft.com/en-us/library/hh855348.aspx`, 2012.

[6] SSPNet, "Social signal processing network,", `http://sspnet.eu/about/`, 2012.

[7] P. Krauthausen and U. D. Hanebeck, "Situation-Specific Intention Recognition for Human-Robot Cooperation," in *LNAI*, vol. 6359, pp. 418–425, 2010.

[8] J. L. Burke, R. R. Murphy, E. Rogers, V. J. Lumelsky, and J. Scholtz, "Final Report for the DARPA / NSF Interdisciplinary Study on Human Robot Interaction," *IEEE, Transactions on Systems, Man, and Cybernetics - PART C: Applications and Reviews*, vol. 34, no. 2, pp. 103–112, 2004.

[9] H. Knight, "Eight Lessons Learned about Non-verbal Interactions through Robot Theater Motivation : Use Theater to Improve Robot Sociability Background : Non-verbal Interaction," pp. 42–51, 2011.

[10] C. L. Sidner, C. Lee, and N. Lesh, "Engagement Rules for Human-Robot Collaborative Interactions," 2003.

[11] R. Mead, A. Atrash, and M. J. Pantic, "Proxemic Feature Recognition for Interactive Robots : Automating Metrics from the Social Sciences," pp. 52–61, 2011.

[12] A. Mehrabian, "Pleasure-Arousal . Dominance : A General Framework for Describing and Measuring Individual Differences in Temperament," *Learning*, vol. 14, no. 4, pp. 261–292, 1996.

[13] E. A. Schegloff, "Body Torque," *Social Research*, vol. 65, no. 3, pp. 535–596, 1998.

[14] M. Pantic and L. J. M. Rothkrantz, "Toward an Affect-Sensitive Multimodal Human - Computer Interaction," *Organization*, vol. 91, no. 9, 2003.

[15] R. Ooko, R. Ishii, and Y. I. Nakano, "Estimating a User's Conversational Engagement Based on Head Pose Information," pp. 262–268.

[16] J. Maisonnasse, *Estimation des Relations Attentionelles dans un Environnement Intelligent*. PhD thesis, Universite Joseph Fourier de Grenoble, 2007.

[17] J. M. Rehg, K. P. Murphy, and P. W. Fieguth, "Vision-Based Speaker Detection Using Bayesian Networks," *Pattern Recognition*, no. Cvpr 99, pp. 110–116, 1999.

[18] SSPNet, "Social signal porcessing network,", `http://sspnet.eu/2010/04/semaine-corpus/`, 2010.

[19] M. Mahmoud, T. Baltru, P. Robinson, and L. Riek, "3D corpus of spontaneous complex mental states," *Corpus*, 2011.

[20] D. Vaufreydaz, R. Emonet, and P. Reignier, "A Lightweight Speech Detection System for Perceptive Environments," *3rd Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms, Washington : United States*, 2006.

[21] WillowGarage, "Facedetection.", `http://www.opencv.willowgarage.com/wiki.FaceDetection`.

[22] A. Jaimes and N. Sebe, "Multimodal human-computer interaction: A survey," *Computer Vision and Image Understanding*, vol. 108, pp. 116–134, Oct. 2007.

[23] Weka, "Weka 3: Data mining software and toolkit in java.", `http://www.cs.waikato.ac.nz/ml/weka/`.

[24] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[25] P. Hanchuan, L. Fuhui, and D. Chris, "Feature Selection Based on Mutual Information : Criteria of Max-Dependency, Max-Relevance and Min-Redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.

Social Inclusion of Senior Citizens by a Teleoperated Android: Toward Inter-generational TeleCommunity Creation*

Ryuji Yamazaki[1, 2], Shuichi Nishio[1], *Member, IEEE*, Hiroshi Ishiguro[3, 1], *Member, IEEE*, Takashi Minato[1], Marco Nørskov[4], Nobu Ishiguro[5], Masaru Nishikawa[6], and Tsutomu Fujinami[2]

***Abstract*—As populations continue to age, there is a growing need for assistive technologies that help senior citizens maintain their autonomy and enjoy their lives. We explore the potential of teleoperated androids, which are embodied telecommunication media with humanlike appearances. Our exploratory study focused on the social aspects of Telenoid, a teleoperated android designed as a minimalistic human, which might facilitate communication between senior citizens and its operators. We conducted cross-cultural field trials in Japan and Denmark by introducing Telenoid into care facilities and the private homes of seniors to observe how they responded to it. In Japan, we set up a teleoperation system in an elementary school and investigated how it shaped communication through the internet between the elderly in a care facility and the children who acted as teleoperators. In both countries, the elderly commonly assumed positive attitudes toward Telenoid and imaginatively developed various dialogue strategies. Telenoid lowered the barriers for the children as operators for communicating with demented seniors so that they became more relaxed to participate in and positively continue conversations using Telenoid. Our results suggest that its minimalistic human design is inclusive for seniors with or without dementia and facilitates inter-generational communication, which may be expanded to a social network of trans-national supportive relationships among all generations.**

I. INTRODUCTION

This paper overviews our work and explores the next generation of communities in which senior citizens will remotely participate. Due to increasing longevity and declining fertility, worldwide societies have never been as old as they currently are, and the aging of populations is predicted to continue as the following UN data [1] show. The global proportion of senior citizens over 60 reached 11% in 2011, and

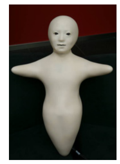


Fig. 1 Telenoid R2

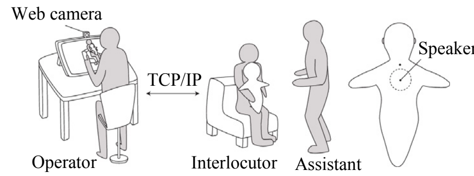


Fig. 2 Conceptual diagram of Telenoid system

the figure will double by 2050. The amount already exceeds 23% in Japan. Less developed regions also face rapidly aging populations. In 1950, the number of children under 15 in the more developed regions was more than twice the number of elder persons. By 2011, the proportion of senior citizens in the more developed regions had surpassed that of children, and by 2050, the proportion of senior citizens is expected to be about double that of children. In the acceleration of demographic changes, the growing number of senior citizens is also rapidly increasing the amount of senile dementia. The care of such elderly people and inter-generational communication require immediate attention. On the other hand, traditional social ties have been frayed or torn by such shifts as the concentration of populations in urban areas, the separation of working and living places, and the trend toward nuclear families.

The social isolation of seniors includes such problems as solitary deaths, crimes, and the loss of valuable components of their lives. Discussion has focused on such themes as changes in the family structure that reflect increases of paid labor, changes in interpersonal relationships, changes in medical and care environments by policy developments, and poverty [2, 3]. There have been two main approaches to resolving the social isolation of the elderly: reconstruction of local communities and improvements in healthcare policies or systems. Since the decline of local communities often socially isolates the elderly, reconstruction is required; however, there are limits to locally resolve such problems by families and neighbors due to the shifts and the "hollowing out" of local communities. A new approach is required that creates a supportive and inclusive telecommunity for senior citizens by telecommunication technology, which may also be promoted by a technology policy that seeks sustainable development for aging societies.

How can robotics contribute to the development of social networks of senior citizens? Institutionalization is one solution to social isolation among elderly people; nevertheless, staff shortages in care services complicate maintaining healthcare systems, and there is also increased demand from seniors who want to enter aging-in-place, which advocates allowing residents to remain in their living environments, especially in their own homes, despite the physical or mental decline that commonly occurs with aging. Therefore, there is a growing need for new technologies that can assist the elderly in their daily living environments: care facilities and their homes. We seek a sustainable solution that addresses the social isolation

*This research was partially supported by JST, CREST and SPP, and MEXT KAKENHI 24300200.

1: Advanced Telecommunications Research Institute International, Keihanna Science City, Kyoto 619-0288, Japan.(e-mail: ryuji-y@atr.jp)

2: School of Knowledge Science, Japan Advanced Institute of Science and Technology, Nomi, Ishikawa 923-1211, Japan.

3: Department of Systems Innovation, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-0043, Japan.

4: Institute for Culture and Society – Philosophy, Aarhus University, Jens Chr. Skousvej 7, 8000 Aarhus C, Denmark.

5: Studies in Language and Society, Graduate School of Language and Culture, Osaka University, Minoo, Osaka 562-8558, Japan.

6: Center for the Study of Communication-Design, Osaka University, Toyonaka, Osaka 560-0043, Japan.

of the elderly and improves their quality of life. We propose a teleoperated android robot as an embodied interface for the elderly and their remote communication with others. Our exploratory study identifies the key issues by verifying the influences on elderly people of a teleoperated android robot named Telenoid. We verify whether its design technology is universally acceptable for supporting seniors' quality of life and for creating a remotely inclusive community for them. Our research question focuses on how the elderly, their caretakers, the operators, and their societies accept Telenoid. As pilot cases, we first conducted field research at a residential care facility in Japan and next at care facilities and homes in Denmark to see the cross-cultural reactions of seniors to it. In addtion, we explored how the android affects communication between the elderly and schoolchildren in Japan.

In the study of assistive technology for senior citizens, designing communication aids is crucial, and computer-based assistance has been investigated for improving communication among people with dementia, care personnel, family members, and volunteers. The importance of conversation has also been identified. Interviews with caretakers revealed that such behavioral disturbances as wandering and verbal abuse were reduced after conversations with volunteers who listened by videophones [4]. Poor/limited social interaction and networks have also been recognized as a main factor that increases the incidence of dementia [5]. To improve the quality and quantity of communication between the elderly and others, we conducted experiments with a field work approach in their real environments by introducing android robots. In the past, many studies on social robots in elderly care have focused on the pet-like companionship a robot might provide. The following relationships have been investigated with zoomorphic robots that resemble animals: AIBO, a robot resembling a dog; iCat, a cat-like robot; Phyno, a penguin-like robot; and Paro, a seal robot designed for therapy with demented seniors. After Paro was introduced into a nursing home, participants treated it like a new companion and felt less lonely [6]. The influence of humanoid robots on the elderly remains uninvestigated, especially the uncanny valley problem, which is the subtle differences from humans that evoke anxiety in observers [7].

Exploring the influences of teleoperated robots on human relationships is challenging. Most recent studies on types of robots have concentrated on autonomous schemes, but teleoperation systems have been used as supplements or substitutes. The use of substitutes reflects the difficulty of implementing sufficient intelligence. Remote disadvantages have been resolved by video conferencing and telepresence robots, but one issue raised by our research is whether it is possible to determine the primacy of remote communication over face-to-face [8]. We must try to develop tools that people might even prefer to use when they have the option of interacting with physical proxies [9]. Social robot studies have advanced the application of robots to education, although most works have concentrated on autonomous robots. Sick students remotely controlled PEBBLES, which are mobile video conferencing platforms [10] that are housed in egg-shaped shells with huggable contours. Even though theories on robot design are expected, they remain undeveloped. Recently, robots have been produced whose appearances closely resemble humans, and research has begun on such minimal designs [11]. Designing teleoperated androids allows us to probe the effects of minimizing the shape or function of a human body during human interactions.

II. TELENOIDS R1 & R2

Telenoids R1 and R2 are a new type of teleoperated android robots with a minimal human likeliness design that can represent anybody (Figs. 1, 2). Both have nine degrees of freedom for their eyes, jaw, neck, and hand motions. Telenoid R1 is 80 cm long and weighs about 5 kg. Its silicon skin feels pleasantly similar to human skin. Telenoid R2, which is 70 cm long and weighs about 4 kg, is covered with soft vinyl chloride. The teleoperation system in both versions is common and only requires a single laptop; with an internet connection, both Telenoids can be operated from anywhere in the world. The operator's face direction and lip movements are captured by a face recognition system and remotely sent to the robot. GUI buttons control the specific movements of the arms and head to remotely embody the operator's behaviors and emotions, such as waving good-bye or hugging. Such unconscious motions as breathing and blinking are generated automatically to give a sense that the android is alive.

Telenoid is modeled to resemble a human and can be perceived as either male or female, young or old. This minimal design feature allows people to feel as if a distant acquaintance is actually close. Our objective is to create a minimal human representation/embodiment that allows any person to transfer her own presence to a distant location by mediation. Key features include: 1) an omni-human likeness that enables users to feel any person's presence (e.g., a feeling of being there), 2) holdability that facilitates physical interaction, and 3) mobility that encourages people to use Telenoid in a variety of situations. There have been developed other robots that are comparable to the Telenoid, with respect to certain aspects of their functionality, such as for instance the "IP RobotPHONE", which also addresses the telepresence communication [12]. However, for a minimal design, the robot's appearance should avoid preconceived ideas about robots. As a minimal human, Telenoid was created by removing as many unnecessary features as possible by using the following procedure: 1) choosing features for communication with humans, e.g., voice, and eliminating the non-neutral, gender-specific ones, e.g., beard, 2) re-evaluating the chosen features to fit the design requirements by eliminating unnecessary ones, and finally, 3) acquiring the essential features.

III. FIELD TRIAL 1: ELDERLY CARE FACILITY IN JAPAN

First, in Japan, we introduced Telenoid R1 into a residential care facility to observe the various reactions of seniors to it. We wanted to see the natural reactions and responses of demented elderly citizens.

A. Method

The experiment was conducted at a long-term care facility in Ishikawa in Japan. Prior to its implementation, we received approval for the experiment from the facility's administrator, its primary doctor, the family members of all the participants, and the ethical committee at Advanced Telecommunications Research Institute International (approval code: 10-507-4).

The participants were ten elderly women with dementia (mean age, 86.6 years), as evaluated by a Japanese government standard [13]. The severity of dementia can here be categorized within the following ranks: Ranks I-IV and M. The participants' dementia was diagnosed as Rank II (mild), III (moderate) and M (severe) respectively. Rank II elderly are capable of live independently with supervision, but Rank III seniors require constant nursing supervision. Rank "M" indicates the emergence of severe psychological symptoms, problematic behaviors, and severe physical disorders. In a one-day field trial, the participants stayed in their living space as usual and were invited by their caretakers to sit in front of Telenoid in turns (Fig. 3). The operator's role was carried out by involved researchers and a chief caretaker at the facility who also observed the reactions of the participants. The time for each participant to talk with Telenoid was limited to about 20 minutes since there was only one robot. Conversation topics included health, hobbies, and family.

B. Results

1) Positive and interactive reactions

We observed the conversations between Telenoid and the seniors with various levels of dementia and described the notable cases that showed reaction patterns. Telenoid made a fine first impression on the participants, especially on the elderly with mild dementia. At first glance, almost all the elderly persons reacted positively to the interaction with Telenoid, often saying something like "*You are really cute.*" Participant E's (Rank II) reaction was typical: She maintained a positive attitude like the other participants. When the operator asked her to hold it, she said, "*Do you want me to hold you? You said, 'Yes' Haha! OK, I can hold you.* [An attendant handed Telenoid to E]. *Oh, you are so cold! But nice! I like you. Oh, great! You are so cute! Ha, but you are quite heavy, aren't you.*" [She warmly rubbed its back]. Participant attachment to Telenoid increased after holding it.

Although it was hard for participants I and J with severe dementia (Ranks III and M) to maintain verbal communication, they intermittently caressed its back and arms and interacted slowly with it. All other participants held interactive conversations with it. For example, participant E said, "*Haha, you asked me to hug you tighter.*" [She patted and rocked Telenoid in her arms.] She expressed a desire for her own Telenoid after being asked to hold it tighter by the operator. She added, "*Oh, yes. I want a doll like this! Oh dear, so cute. I really want a doll like this!*" Even though the senior imagined Telenoid as a well designed doll, she became reluctant to leave it. As the operator prepared to leave, she said, "*I will miss having you around. I don't want you to leave.*" The operator replied, *"It is already my bedtime,"* and she said, *"You mean good-bye? Why are you getting sleepy? Oh dear, how cute you are!"* The caretaker said, "*you really seemed pleased with him. You told me you were a bit hesitant to come*," and E replied "*Haha, it was fun. I had a great time! I'm going to live a long time.*" Participants increased their attachment to Telenoid during their conversations and physical contact with it.

2) Imaginative conversations

Participants projected their images onto Telenoid. Participant G (Rank II; she had a moderately severe or severe

Fig. 3 Elderly interacting with Telenoid

cognitive impairment based on another test) showed a typical case of imaginative conversations. At first glance, she commented, "*Hello, boy. You seem about to smile. You are so cute!*" She was asked to hold Telenoid and said, "*Can I hold you?. . . Oh, you are so heavy! When you were a little child, you were probably lighter.*" The operator asked the senior to guess its age, and she replied, "*I think you are about five years old,*" although she volunteered that one of her daughters was eight and confessed to easily forgetting their names. She hugged Telenoid, "*You are getting quite heavy, aren't you. But that's OK. That's natural because you are growing up*."

A female caretaker took a turn operating Telenoid, which had been operated by a male researcher. After a while she asked G whether it was male or female, and even though the caretaker said that it was female, G claimed that it was a boy. On the other hand, G said, "*You do not have to worry about it. It'll be about a year before you look like a girl. . . . When your hair grows longer, you'll look more feminine.*" After they sang together, the operator said that Telenoid was female and asked who seemed to be female among those in front of G. She insisted, "*(Telenoid) is female.*" The operator admitted that she was a nurse working in the facility and asked G to guess her age. G replied, "*Maybe seven or eight . . . . You look young, but you're not a child.*" G recognized Telenoid-system as human, a person who would age, and grow up. She changed her image of Telenoid during the interactive conversation, which was imaginatively conveyed. She positively engaged in it despite severe cognitive impairment. When she finished talking with Telenoid, she seemed very interested in it, asked her caretaker what Telenoid was going to do, and expressed concern about who would care for it until its mother came.

IV. Field Trial 2: In Senior Homes in Denmark

In Denmark, we tested Telenoid R2 in care facilities and the homes of seniors. By setting it up in private homes, we expected residents to act freely and react strongly toward it. Here we focus on two cases of experiment results in the homes of elderly Danish participants in the Svendborg municipality.

A. Method

We conducted a two-day trial with two male participants who were living alone in houses attached to care facilities: a healthy 92-year-old (participant K) and a 75-year-old with mild Alzheimer's (participant L). In both cases, we set Telenoid up in the living room where they could relax and receive visitors (Fig. 4). For convenience in setting up the operating system, the bedrooms next to the living rooms were utilized for teleoperation by the researchers, nursing students, and a friend. Each participant interacted with Telenoid for about two hours. Conversation topics included health, hobbies, and family, and a cooperative map game/task. To identify the degree of the participant engagement in the conversation, we extracted three minutes of conversation from each of four situations (two situations each for participants). We

Fig. 4 Participants K (left) and L (right): interaction with Telenoid at home

accumulated the participant times, the operator utterance intervals, and the non-speech intervals and calculated the ratios of the total times to the entire three minutes. We chose those three minutes because they were representative of the whole conversation and to illustrate the possibility of meaningful interactions.

B. Result

1) Positive reactions across cultures

In the individual homes, Telenoid encouraged positive responses and behaviors of both participants K and L, who were willing to talk using it and maintained the conversations. They naturally started conversing with Telenoid from the beginning. We observed that their attitudes toward it were consistently positive and they actively reacted to it.

In the case of K, he positively continued to talk to it to the extent that the rate of his utterance intervals (the accumulation of such times) to the entire three minutes exceeded 70%, whereas the rate of the operator/nursing student was about 20%. He loved to talk and had many things to share, because Telenoid-system became a good listener. On the first day, two nursing students alternately played the operator, but K was not aware of the changes. This suggests that Telenoid's identity reflected the interlocutor's imagination more than the real operators. On the second day, his senior female friend became the operator and talked with him. He continued to talk to Telenoid; his friend talked with him at a 40% utterance rate, and their conversation proceeded more interactively than the previous day while they nearly filled every second.

Participant L with mild Alzheimer's preferred to act calmly and to contemplate more than K. It was possible to have conversations about literature, politics, and health with L. We tested whether the conversation content could go beyond small talk. When a researcher/operator discussed poetry, L became a listener; after another researcher assumed the discussion, L continued to talk about the poems and his utterance rate doubled from 20 to 44%, but the rate of their non-speech intervals was halved. L became a speaker, and his conversation with Telenoid interactively progressed.

2) Dialogue strategy

Both participants continued to enjoy interaction by changing their dialogue strategies, including reading poems, talking about TV programs, hugging Telenoid, showing favorite items to it, playing the piano, singing, introducing friends, and sharing time with it. Participant K was highly motivated to talk to Telenoid and encouraged others to join their conversations. He also changed the situation and began to amuse Telenoid by playing the piano and singing. On the second day, to enjoy the conversation in a different way, he invited a senior friend. For participant L, he established not only verbal but also nonverbal contact with Telenoid by giving it a big, silent hug while standing up. This big hug was one of his characteristic behaviors. He seemed to feel this reward to be a sense of reassurance and greater comfort from his experience with the robot. Even when talking to it, he used many gestures and communicated with Telenoid in various ways. Such behaviors as touching, holding, hugging, and imitating were observed to build a relationship for both K and L. They were attracted and encouraged to engage in conversation by the embodied communication medium.

V. Field Trial 3: Inter-Generational Communication

In Japan, we set Telenoid R2 up in a care facility and our teleoperation system in an elementary school to determine the influence of the teleoperated android on both the elderly and the children. In this section we focus on the operator's side.

A. Method

We conducted a two-day trial. The schoolchildren telecommunicated with the elderly on the first day, and on the second day they received training for developing embodied communication skills inside the school. First, they operated two Telenoids set up at a facility in Kyoto that was far from their school in Ishikawa and communicated with demented seniors (left in Fig. 4). The participants were 16 children (9-10 years, six boys and ten girls) and ten senior citizens with mild to moderate Alzheimer's and dementia (mean age, 92 years, two males and eight females). Both sides of the participants were divided into two groups, and they had one-on-one communication in each group, although the time was limited to ten minutes for each person.

On the second day, three Telenoids were brought into the class for developing the children's operation skills. They reflected on and explored their relationships with the Telenoids in order to improve their communication with the seniors. Under the guidance of a dancer who taught the children nonverbal ways of communication. They explored how people can humanize a robot from the interlocutor side and more fundamentally identify it as part of their own body on the operator side despite mechanical heterogeneity. To explore their relationships with the robots, they were told to imagine the voices. After naming each Telenoid, talking to them, and listening to the robot voices in their own minds, the children in pairs imitated the bodily expressions that they shared with the Telenoids and carefully observed their own movements in a mirror and also those of Telenoid in front of the operator (right in Fig. 4.)

A. Results

1) Easy communication

When conversing with the elderly participants, the children felt nervous at first, but began to relax as they continued to talk with the senior citizens. Here are examples of comments made by the children: "*I did not feel embarrassed as much as face-to-face*," "*It was easier to talk through Telenoid because I could see the other but she couldn't see me*," "*Telenoid is convenient because it let me ask what I couldn't ask face-to-face*," "*While I was talking through Telenoid, it became fun to talk, and I want to talk with a senior citizen again*," "*It was fun to talk through Telenoid, and I was glad that I could see so many smiles on the elderly*," and "*The grandma was always laughing, so it was fun.*" Most of the children felt more relaxed talking through Telenoid than face-to-face and had positive impressions about the elderly with dementia. The children also developed clues for communicating with the elderly by carefully observing the other operators and were positively involved in the interaction with the senior citizens.

2) Teleoperation skill

On the first day, the children began to explore their new relationships with elderly while adapting to the conversations. "*The grandma hugged and kissed Telenoid, it was so funny*," and "*I talked with my friend about ideas for using Telenoid, and we came up with ideas like 'please move your hand over my head.' and 'please shake my hand.' When I asked her to shake my hand, she responded, so I was pleased.*" They also asked the elderly to pass her hand over the face of Telenoid and they tried to sing together. They spontaneously changed their dialogue strategy to explore nonverbal communication ways to utilize their varied embodiments. Training the teleoperators is required so that the communication can more fully exploit the embodied medium. The dancer, who taught embodied communication on the second day, focused on the children's conceptions of Telenoid. He asked them about becoming friends with it and encouraged them to gradually integrate their conceptions into their own body images so that they could conceive of it as part of themselves.

During the session, the children started to establish relationships with Telenoid by approaching closer and closer, talking to and listening to it, even during their breaks and after the session. After representing the imagined voices inside it by saying, e.g., "*Telebow (a given name to it) asked me to play*," although they were told to hear the voices, the children told that the Telenoids seemed to have their minds. They operated the robots in turns and noticed the differences and the relation between their bodies and Telenoid's by comparing each movement. They realized that its left-right motion is the opposite of their own self-image reflected in a mirror. They adjusted their movements to fit its limited functionality, e.g., the range and speed of their head movements. Their conceptions of Telenoid were articulated in relation to themselves, and they slowly overcame the gap. The children began to accept Telenoid not only as a member of the class but also as a part of their own bodies. With the embodied medium, the children explored ways to communicate nonverbally with each other. Although the system could only track the operator's head movement, the operators moved and made sounds rhythmically; the interlocutors responded to the rhythm by clapping and swaying. They experienced the emergence of a rule or protocol among themselves and developed repertoires of communication strategies.

VI. DISCUSSION AND FUTURE DIRECTIONS

In our field trial at a care facility in Japan, we focused on demented senior citizens and found strong attachment to Telenoid's minimalistic human design and willingness to converse with it. A notable feature was shown in participant G who had an imaginative and interactive conversation with Telenoid. She treated it as a child growing up, for example, with respect to its weight and hair length. She talked to Telenoid as if it were alive and even changed her image of it during the interactive conversation. Telenoid's personality or identity reflected her own state of mind, due to its minimized design that allows people to put their own images onto the robot. In fact, the elderly with dementia projected images of the operators, onto Telenoid and interactively created its identity. Hence, the concept of minimalistic human design seems suitable for demented senior citizens. In Denmark, we also found that Telenoid elicited enthusiastic and empathetic attitudes from the seniors with or without dementia and encouraged interactive communication with others using it. Their generally positive attitudes and attachment to its minimal design were cross-culturally shared in both countries.

Fig. 5 Left: child conversing with senior citizen;
Right: child comparing his movements with Telenoid's

On the other hand, we received skeptical and even negative reactions from the press in Denmark; we never saw this reaction in Japan. For example, a university ethics professor commented: "*Relational technology can lead to the curtailment of human emotions*" [14]. Negative effects, however, must be discussed and analyzed in light of the positive results, and vice versa. Even if emotional involvement is reduced, the result has ambiguous aspects and can also be seen positively. In our third trial aimed at children, many of whom unfortunately have few opportunities to meet senior citizens, Telenoid helped them feel relaxed and lowered communication barriers with demented seniors; the children showed greater interest in communication with the seniors. In interviews with the facility personnel after the inter-generational communication, one staff member said, "*The resident seniors made more informal conversations than face-to-face with visitors. Even if the visitors were children, the elderly would become really tired. There must be situations where they can talk to a doll when it's hard to contact people*." At another facility in the first trial, a caretaker said, "*The elderly interactively talked much more to Telenoid than they would to a doll*."

Our results suggest that teleoperated androids can be a universally acceptable medium that fosters an inclusive relationship for senior citizens, even those with severe cognitive impairment, by promoting conversation. Telenoid has potential to promote better understanding of seniors by helping operators elicit their imagination and enter their world. The operator's responses also affected the reactions of the senior citizens, and in dementia care, Telenoid might become an educational training tool for understanding the needs and desires of such seniors. Furthermore, it functioned as a buffer for the children as well as the elderly and encouraged the children to change their strategies to communicate with the elderly. To more fully exploit the embodied medium, we need to develop a teleoperation training program. Embodied communication might develop a wide repertoire of communication strategies for operators. As an assistive technology, we propose to employ Embodied Communication Technology (ECT), such as Telenoid, that promotes the social inclusion of senior citizens and inter-generational communication. We envision creating a telecommunity where all generations are sustained through mutual assistance to take care of others by anyone, at anytime, and anywhere.

Based on the results from the above trials, we identified the following issues to be solved in future work:

1) *Efficacy of telecommunication*: We observed that senior citizens spontaneously and positively talked to and interactively developed conversation with Telenoid both in Japan and Denmark. The importance of conversation has been identified in dementia care and prevention, and it is reported that patient's behaviors as symptoms of dementia like wandering and verbal abuse (such behavior disorders are called BPSD) were reduced after conversations [4, 5]. We need to verify if the conversation promoted by Telenoid reduces BPSD in seniors. A longitudinal study is required to firmly establish the efficacy of Telenoid's use. In past studies, people tended to lose interest in robots after several days. Since teleoperated androids are remotely driven by the operator, this is unlikely to happen, but we need verification. Moreover, the manifold effects of telecommunication must be investigated: the primary effect during conversation, the secondary effect on ordinary life after the Telenoid sessions, and the tertiary effect on those connected to the seniors like staff members and families.

2) *Embodied communication training*: Seniors projected their images onto Telenoid, and due to its minimized design, its personality or identity reflected their states of mind. It might become a new medium that enables operators to explore the inner worlds of seniors by performing dialogues. We need to verify whether it can be used as a training tool for knowing their needs and desires. Using Telenoid provides another advantage; operators can easily ask others for advice about communication ways. They can develop nonverbal communication skills by identifying their relation to a new body. We need to develop an effective training program for operators to react to and elicit responses from seniors. From the interlocutor side, the question remains how caretakers can facilitate relationships between seniors and Telenoid.

3) *Cross-national and trans-disciplinary study*: We received negative reactions from non-users in Danish media reports, which were not seen in Japan. The issues emphasized in the reports were related to the theme of losing relationships with humans by replacing humans with machines, and included the ethical question of what is considered good or evil in various human cultures and by the international community. Our research intensions are not to replace humans with robots but to develop new potentials for human users (operators as well as interlocutors). On the other hand, there are those who say that Telenoid cannot contribute to saving manpower as long as an operator is required and that we need to consider how to deal with this issue. One solution is developing a semi- or completely autonomous system, but an ethical issue arises in leaving elderly to a robot: losing their relationships with other people and resulting in their neglect. Also, a reporter asked if using a robot such as Telenoid with seniors by humanizing is deceitful. We may also ask if perspective operators are in danger of losing, and missing an opportunity to grow in terms of their compassion and empathy with others. Further investigation is required on both sides regarding what people can gain and lose by using teleoperated androids, and on how a consensus can be reached on the usage of such robots.

To create a worldwide telecommunity, a cross-national study is inevitable, where observations are made both on societies and on such units as individuals within societies. By comparison of reactions in each country, we will be able to further explore the essential features of a minimal design for human beings and a universally inclusive medium for e.g. seniors. Furthermore, we need to investigate differences concerning the social acceptability of teleoperated androids based on different cultural backgrounds and collect information from the various stakeholders involved. Put simply, technological knowledge can be used for good or ill alike: e.g., a knife can be made for benign or malign purposes. Technology, technique, etc., were derived from the ancient Greek term *technē* that refers to the knowledge of an expert like a physician. As for the nature of such knowledge, philosophers discussed *technē* as productive knowledge that is instrumental, serving other ends. Its effectiveness offers no guarantee that the ends are good, but the type of knowledge is intrinsic in humans: tools and technology may be fundamentally constitutive of human consciousness [15, 16]. More trans-disciplinary research is nescessary, considering the cultural dimensions/complexity of social acceptability, as well as ethical and philosophical issues such as the nature of technological knowledge and the new possibilities of body/mind relations mediated by a teleoperated android.

REFERENCES

[1] Population Division, United Nations, *World Population Prospects: The 2010 Revision*, New York, 2011.

[2] P. Townsend, *The Family Life of Old People*, London, 1957.

[3] Cabinet Office, Government of Japan, *Fiscal Year 2010 (April 2010 - March 2011) Annual Report on the Aging Society*.

[4] N. Kuwahara, et al., "Networked reminiscence therapy for individuals with dementia by using photo and video sharing," *Proc. of ASSETS '06*, pp. 125-132, 2006.

[5] L. Fratiglioni, et al., "Influence of social network on occurrence of dementia: a community-based longitudinal study," *Lancet*, Vol. 355, No. 9212, pp. 1315–9, 2000.

[6] K. Wada, et al., "Robot Therapy in a Care House: Results of Case Studies," *Proc. of the 15th International Symposium on Robot and Human Interactive Communication*, pp. 581-6, 2006.

[7] M. Mori, "Bukimi no tani [The uncanny valley]," *Energy*, Vol. 7, No. 4, pp. 33–5, 1970 (in Japanese).

[8] J. Hollan, et al., "Beyond being there," *Proc. of CHI'92*, pp. 119-25, 1992.

[9] R. Yamazaki, et al., "How Does Telenoid Affect the Communication between Children in Classroom Setting?" *CHI'12 Ext. Abst.*, pp. 351-66, 2012.

[10] D. I. Fels, et al., "Video-mediated communication in the classroom to support sick children: a case study," *International Jour. of Industrial Ergonomics*," Vol. 28, No. 5, pp. 251-63, 2001.

[11] K. Ogawa, et al., "Exploring the Natural Reaction of Young and Aged Person with Telenoid in a Real World," *Jour. of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 15, No. 5, pp. 592-7, 2011.

[12] D. Sekiguchi, et al., "Robot-PHONE: RUI for Interpersonal Communication," *CHI'01 Ext. Abst.*, pp. 277-8, 2001.

[13] Meeting for study of the Long-Term Care Insurance system: *Handbook for investigator of certification of the long-term care insurance system*, Tokyo, Chuo-hoki, 2003 (in Japanese).

[14] "Den kan også give en krammer [It can also give a hug]" http://www.information.dk/263649 [Accessed on June 25, 2012] (in Danish)

[15] Aristotle, *Nicomachean Ethics*, Trans. M. Ostwald, New York, Macmillan, 1962.

[16] B.Stiegler, *Technics and Time, 1: The Fault of Epimetheus*, Trans. R. Beardsworth and G. Collins, Stanford, Stanford University Press, 1998.

Can smart rollators be used for gait monitoring and fall prevention ?

C. Dune, P. Gorce, J.-P. Merlet.

***Abstract*— Clinical evaluation of frailty in the elderly is the first step to decide the degree of assistance they require. This evaluation is usually performed once and for all by filling standard forms with macro-information about standing and walking abilities. Advances in robotics make it possible to turn a standard assistance device into an augmented device. The existing tests could then be enriched by a new set of daily measured criteria derived from the daily use of standard assistance devices. This paper surveys existing Smart Walker to figure out whether they can be used for gait monitoring and frailty evaluation, focusing on the user-system interaction. Biomechanical gait analysis methods are presented and compared to robotics system designs, to highlight their convergences and differences. On the one hand, monitoring devices try to estimate accurately biomechanical features, whereas, on the other hand, walking assistance and fall prevention do not systematically rely on an accurate human model and prefer heuristics on the user-robot state.**

I. Introduction

Ageing in society is a worldwide issue that especially impacts northern countries. In France, due to the high care cost and to the limited number of rooms in care institution, the solution that has been chosen by care-givers, frail people and their family is to maintain elderly at home the longest and in the best conditions by giving them an *adapted assistance*.

Clinical evaluation of frailty in the elderly is the first step to decide the degree of assistance they require. This evaluation is usually performed once and for all by filling standard forms with macro-information about standing and walking abilities, e.g. by measuring the time taken to walk $10m$. Advances in robotics make it possible to enhance a standard assistance device by adding sensors and actuators. The existing tests could then be enriched by adding a new set of daily measured criteria derived from the daily use of standard assistance devices. This monitoring will allow to evaluate gait in ambulatory conditions, to measure the evolution of some pathologies, to refine diagnostics and to distinguish autonomy levels. The assistance device is not meant to be an alternative for clinical frailty observation but rather as a complementary tool that gives field information. The data acquired on-line could also be used to control a robotics walker in order to prevent a fall. These new characteristics can extend the use of walkers to more diverse population.

C. Dune and P. Gorce are with Handibio, EA4322 Université du Sud-Toulon Var, France.

J.-P. Merlet is with Coprin team Inria Sophia Antipolis, France

In this survey, we will focus on the use of a standard *rollator* equipped with sensors and actuators for monitoring. We will try to determine whether relevant information can be obtained from an equipped *rollator* without adding any sensor on the human body. This study then tries to answer the following questions :

1) What are the relevant features for walk analysis ?
2) How the use of a rollator distort the walking gait ?
3) Some smart walker already exists. What kind of data are they acquiring about human state and what kind of input are they using to maintain the system balance ?

In the first section, we will start with a biomechanical point of view by surveying studies led on elderly and rollator walking, and we will partially see to which extend studying the walk with a rollator may be relevant. In a second part, we will survey the existing robotics frame and their user interfaces. The first section is dedicated to walk *monitoring* with a smart rollator, the second one deals with walking *assistance* method and the last presents *fall detection* systems.

II. Walking Gait Analysis with our without rollator

Depending on the degree of assistance they need, people are prescribed canes, crutches or walkers [1]. Le latter can be legged walker or wheeled walkers (rollators). A rollator can be defined as a frame with three, or four wheels. It has handles with brakes, and in some case a seat, a basket and a tray (see fig. 1).

Fig. 1. 3 wheeled and 4 wheeled rollators

Static equilibrium is maintained when the body's center of pressure is positioned over the base of support. Loss of balance can result when the center of mass is displaced in relation to the base of support because of voluntary movements or external perturbations.The use of a walker increases the base of support, thereby allowing a greater tolerated range for center of mass positions. They can also prevent instability by allowing stabilizing reaction forces such as holding on or pushing against the ground.

Basically, standard four legs are dedicated to people that need assistance to maintain their balance or for those that require partial weight support [1]. Their use changes gait pattern and posture during a gait and requires good coordination for lifting up and placing forward the device during gait [2]. On the opposite rollators induce a more natural gait pattern but lack in stability. They are designed for people that need less weight bearing [2].

A. *Standard measurement of walking gait*

In order to design a useful system for clinicians we have first to understand what are the common features they use for walk gait analysis and what are the standard tools. The medical walk analysis can be divided in three steps : i) patient qualitative observation, ii) a description phase and iii) a biomechanical analysis. Description phase and biomechanical analysis depends on the equipment available in the medical center. The observed features range from spatio-temporal gait analysis to fine body motion analysis.

a) Elderly specific gait pattern: Ageing decreases the muscular force and changes the postural control and gait. It is not obvious to determine what is a ”normal walk” for elderly since each individual develops his own adaptation strategy to maintain balance. Some believe a slowed gait is a disordered gait, and others believe that any aesthetic abnormality, e.g., deviation in smoothness, symmetry, and synchrony of movement pattern, constitutes a gait disorder. However, a slowed or aesthetically abnormal gait may in fact provide the older adult with a safe gait pattern that helps maintain independence [3]. Patients are separated into three classes : *autonomous*, if they can walk freely without assistance, *frail* and *dependant* if they can not walk without assistance device.

To assess the frailty degree, standard tests are performed : *10 meters walk test*, *Tinetti balance test* and *Timed Up and Go* test [4]. In the latter, the patient sits on a chair with his back against the chair back. On the command ”go”, the patient rises from the chair, walks 3 meters at a comfortable pace, turns, walks back to the chair and sits down [4]. Tinetti test studies user postural abilities. Other tests can be added to evaluate transfer abilities and muscular force are also evaluated. Balance is tested by evaluating intrinsic balance by changing arm height, extrinsic balance by pushing the patient and simple support capabilities opened eyes or blind.

Physicians observe patient overall posture and arm swing during the walk. They also examines gait parameter such as gait width, step length, cadence and high of heels during a walking cycle, metatarsus-tibia angle, position of the foot when walking.

Indeed, some specific patterns can be detected by studying walking gait :

- After a fall or a stroke, people are subject to retropulsion syndrome which make them walk on the heels, enlarge their support base, and increase the knee flex.
- A stepping may be related to antero lateral leg muscles paralysis along with a loss in foot’s dorsi-flexion, making the patient lift his feet higher that necessary.
- Parkinsonian festination corresponds to a speed up of the pace. The patient bends with an increase flexion of the knees.
- Hemiplegic pyramidal spastic gait induces a rigid leg and foot sliding on the ground
- Multiple infarcts syndromes are related to small steps where the heel of one foot does not reach the toes of the other foot
- Heeled walking can be related to sensory diseases
- Charcot’s gait increase the support base, i.e. the gait width
- Waddling gait can be related to muscular force loss or D vitamin deficiency for elderly
- Zigzag gait is linked with vestibular syndrome

Notice that these pattern rely on standard spatio-temporal analysis of gait. 3D feet positions during the walk seem to be the key component in gait analysis. But it is may be due to the fact that the current gait analysis tools can accurately measures this feature. If other characteristics could be as easily extracted, the pattern might be enriched with other dimensions.

a)

b)

Fig. 2. From locometer to lokomat a) Bessou’s locometer (Walkmeter ©), consists of two wired that are tied to the feet. Wired expansion is measured in time, witch allow to compute spatio-temporal analysis.b) Lokomat ©is a complex treadmill for walk analysis and rehabilitation

b) Standard gait analysis tools: A standard inexpensive gait analysis device is the Bessou’s Locometer. It consists in measuring the length of two cables tied to the feet (see fig. (2)). It allows to accurately measure simple/double support phases, step length, cadence and velocity. Walkway or treadmills equipped with force sensors may give additional information about foot pressure distribution, centre of pressure trajectory and center of gravity trajectory. Further biomechanical investigation can be obtained using body segment tracking thanks to motion capture equipment (eg. Vicon or Qualisys motion capture system).

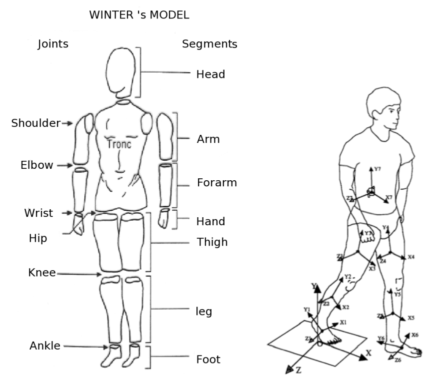


Fig. 3. Winter's human 12 link model and body frames.

The human model kinematics can be represented as a system made of articulated rigid bodies (cf fig. 3). The number of links depend on the analysis to perform. For walking analysis, the model is sometimes simplified, assuming the walk to be symmetrical. It is then limited to the study of one lower limb in the sagital plane. Head and trunk are then fused in one only body and the arms motions are neglected [5]. Joints are marked with reflecting spheres and their trajectories is tracked thanks to a motion capture system [6].

In addition to the kinematics analysis, extrinsic forces can be analysis thanks to force-plates or walkways, giving information about system mass dynamics, which could give useful information to infer balance capabilities. And finally, if pressure sole are available, the foot pressure distribution can be studied, e.g. to check if the elderly is walking on the heel or have a proper use of the toes for propulsion.

A grasping force test is sometimes added to estimate the muscular force, assuming that the loss is uniform on the whole body.

The first objective is to provide physicians with the features they are used to process to evaluate elderly frailty, while maintaining a low cost and ensuring a good ease of use and by embedding all the sensors on the walker without equipping the patient. And at best, the intelligent walker could deliver others relevant features that will enriched the existing feature set. But, can we measure relevant features with a rollator for patients that need it as well as patients that are still autonomous ? Does the gait pattern change significantly using a rollator ?

B. Standard biomechanical analysis of rollator Walking

Very few work have studied rollator walking specificities, although such information are relevant to make a decision on its use or whether the user needs some additional muscular or balance training. This information is also mandatory for a smart walker to assist the walk or compensate for loss of balance.

In a early study, it has been shown that the walking performance in elderly subjects measured in terms of distance, cadence and velocity is improved when they walk with a rollator [7]. And the rollator users are generally satisfied with their rollator and consider it a prerequisite for living a socially active and independent life [8].

Studies on walking with canes or poles have shown that these walking aids reduce the load on the lower extremities [9], [10]. The rollator might also reduce the load on the leg muscles and the joint to some extend as well however, the walking change in gait parameters when walking with a rollator been quantified in very few studies.[11] observes a reduction in the vertical ground reaction force during rollator walking. This study has been refined in [12] that studied the biomedical effects of walking with a rollator on the walking pattern of healthy subjects.

The set up consists in two force platforms and a marker based video tracking system. Methodology follows [13] for the setting up of 15 spherical markers and the use of a three-dimensional inverse dynamics method. Results, for tested healthy people, show that rollator walking did not result in an overall unloading of the muscles and joints of the lower extremities but on a selective one. The unloading of the ankle and knee seemed to be partly compensated by an increase in the hip extensor moment, which probably was needed to push the rollator in a forward direction and keep its horizontal velocity. It is due to the increased forward flexion of the trunk during rollator walking. It concurs with other studies that have observed increased hip flexion along with an increase in the hip extensor moment [14]

Dealing with the pattern gait characteristics, they prove the following specificities :

- increased hip flexion;
- decreased ankle dorsi-flexion and knee flexion;
- significant decreased ankle and especially knee joint moments;
- increased hip extensor.

Theses characteristics have to be taken into account when examining rollator walking.

C. Partial conclusion

Physician usually study spatio-temporal gait parameters. Standard biomechanical tools provide either some of these features or all of them. It can be shown that some walking patterns give relevant indications on specific diseases. Yet, using a rollator may alter the walking gait, changing at the same time the measure pattern. For example, decrease ankle dorsi-flexion is induced by rollator walking and could be interpreted as antero lateral leg muscles paralysis in a free walk. Before drawing diagnosis on patient by using a rollator, we have to make sure that the criterion are valid for a rollator walking. Yet, some features, such as pace and asymmetrical walk can still be observed directly during rollator walking.

III. Intelligent walkers

In this section, we will survey the existing smart walkers, and we will limit our scope to standard medical frame equipped with sensors. Recently, *hand-free* walkers have been introduced, e.g. the KineAssist [15] or the Walkaround [16]. These systems are promising in terms of balance recovering and they have the advantage of leaving the hands free for daily tasks. Yet, in this paper we will focus on common wheeled walker that connects to a person at the hands.

We will investigate the type of data acquired about user's state for three applications : i) gait monitoring, ii) user intent estimation for navigation and iii) fall prevention. Smart walkers may be used to analyse either the environment or the user's behaviour. Environmental data is dedicated to navigation purpose, such as obstacle avoidance [17], wall following [18], slope compensation [19] or localisation [20] [21]. Even though these functionalities are relevant for people autonomy, especially for the visually impaired, these functionalities are out of the scope of this survey. A thorough survey on assistance mobility device, focusing on smart walkers can be found in [22], [23].

A. Monitoring the user state

Some of the existing Smarts walkers aim at tracking the trajectories of gait features in order to monitor health. The great advantage of such systems is that the user stands at a roughly known position with regards to the walker. Body segment localisation is then made easier.

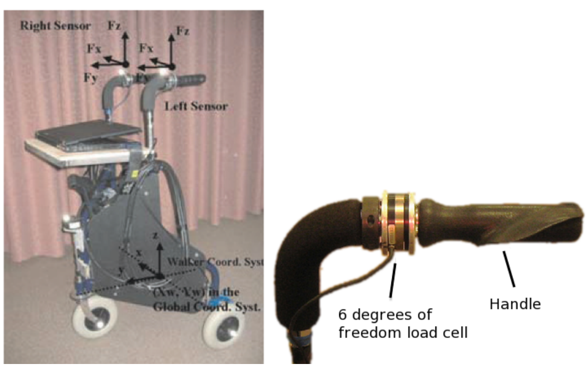


Fig. 4. Medical Automation Reasearch Centrer (MARC) Smart walker [24]

c) Using force sensor to evaluate the walk gait parameters: Walker can be equipped with force-moment sensors mounted on the walker handles [24], [25], or under the forearm [22], [26] to passively derive some gait characteristics. In both cases it is assumed that the force and moment recorded have cyclic changes reflecting the gait cycle and that these changes depend on basic gait features (cadence, stride time, gait phases).

The iWalker [25] quantifies loads exerted through the handles an frame and standards spatio-temporal parameters (such as speed and distance). In [24], a direct comparison between motion capture and force-moment data was studied to detect significant pattern in the force signal (cf. fig. 4). The lateral sway motion of the upper body reflects in peaks in vertical direction and in the corresponding forward moment signal. These peaks coincided with the heel initial contacts and. The forward propulsion force applied by the user is related to the toe-off event from the right and left toe. Finally, the stride (ie. duration of a gait cycle) can be computed from two heel contacts.

In [26], a method based on Weighted Frequency Fourier Linear Combiner, is introduced for the same standards gait parameters extraction from force data.

d) From odometers and accelerometers: Walker wheel motion measurement can also be used to estimated the user state [17], [27]. The Personal aid for mobility and monitoring project (PAMM) [17] developed health monitoring tools. The PAMM smart walker is an omnidirectional walker design for walking assistance with navigation and monitoring functionalities (cf fig. 5). Its sensors record user speed and compute the stride-to-stride variability, which have been shown to be an effective predictor of falls. A power spectrum analysis on PAMM's velocity allows to estimate user's stride length and frequency. Besides, the shape of the power spectrum is related to the gait symmetry. Indeed, for a symmetric gait, the energy is located at twice the stride frequency. However, the system can detect asymmetric gait as spectrum with energy located at the stride frequency and at higher frequency. An asymmetrical gait could be an indicator of a physical injury or a minor stroke.

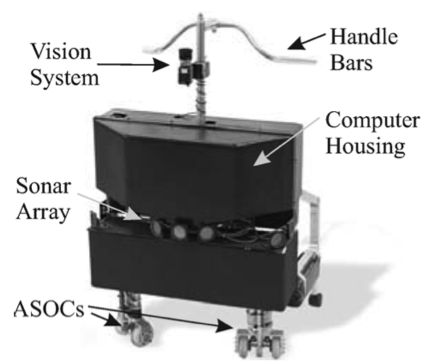


Fig. 5. The PAMM Smart Walker for walk assistance and health monitoring [17]

e) Feet/User position using ultrasonic sensors or cameras: Direct measurement of body segments may be obtained by using ultrasonics sensors or cameras [22], [28]. A vector of ultrasonic sensors can be mounted on the walker to scan the space between the user and the walker and determine coordinate of each leg without adding any marker on the patient [22]. In [28], a camera is mounted on the frame and observes markers on the toes. This marker based toe tracking algorithm allows to calculate step width and provide an accurate assessment of foot placement during rollator use.

The main issue when using that kind of device is that the accuracy of the leg localisation depends strongly on the clothes that the user wears. It is a drawback with regards to method based on odometers or force sensors. The method proposes in [28] by-passes this point by adding markers on the toes. Yet it also by-passes our constraint not to equip the user in order to ensure acceptance and ease of use.

B. Assisting the walk by estimating user intent

Some intuitive command schemes based in user intent analysis have been proposed. The point is to determine how to give the control to a user with taking into account his possible cognitive degeneration. Furthermore, data could be distorted by a pathological gait.

f) From user force interaction: User intent can be inferred from upper-body force interaction in rolator assisted gait [29], [17], [30]. In [30], two 3D force sensors are installed under the forearm supporting platforms (see). The elbow lateral motion is then restricted and the weight of the user on the sensor is increased, leading to a more stable configuration. The force data are processed to isolate three components : high frequencies that are due to the vibration induced by the wheeled friction on a non smooth floor, users's trunk oscillations that are directly related to user gait and user navigation command. The two first components can be monitored and analysed for gait analysis [26], whereas the third component can be used to estimate user intention [30].

PAMM systems [17] use a six axis force/torque sensor attached to the handle as the main user control interface. The force/moment signals are interpreted for motion control by using an admittance controller. They contain the user's intention as well as support and stability information about the user. The admittance model can be tuned for each individual user. It can be made manoeuvrable and light for agile users and slow and stable for someone who needs more support. The admittance is the transfer function from the user's force and torques to the PAMM's velocities. The response of the PAMM is obtained by solving the dynamics equations and then solving the inverse kinematics of the physical system to get the actual actuator velocity. The challenge is then to design the appropriate dynamic model to give the user a comfortable feeling. This is done by choosing a metric to evaluate the performance of the model so that the operator effort is minimized.

The MARC Smart Walker [24] is a three wheeled walker also equipped with force/torque sensors on handles (in addition with sonar and infra-red sensors). The control of the MARC Smart Walker is performed by fusing user intent and walker intent to control the orientation of the front wheel. Walker intent is computed from the information on the surroundings that is acquired by the sonar and infra-red sensors.

g) Using inertial data and inverse pendulum model: The Murata walker has been unveiled during fall, 2011 (see fig.6). It is a two wheeled walker that uses inertial information to maintain human-walker balance, considering human body as an inverse pendulum. If a fall is detected, the system brakes to help the user recovering his balance.

C. Preventing the fall

h) User/Walker distance: The walker-uses relative distance can be used to classify the states between a *walking state*, a *stopped state* and an *emergency state* [31]. A laser range finder mounted on the walker acquired the position of the knee with regards to the walker frame. It is assumed that the position of the feet is at the vertical of the knee position, and fixing the human frame in the middle of the feet. User velocity is estimated from the walker velocity obtained by odometers. The *stopped state* occurs when both the walker and the human velocities are null. To distinguish the *walking state* from the *emergency state*, user-walker distance is used. A normal distance distribution is computed to determine the *walking state* based on user data. The robot control tries to bring the user distance right to the mean of the walking state distance distribution.

Fig. 6. Murata walker has been unveiled at CEATEC Japan exhibition during fall, 2011

i) Observing human posture: In [32] the RT-Walker (cf fig 7) is equipped with laser range finder and perform an estimation of the kinematics of a 7-link human model. The model is used to estimate the position of the user center of gravity (CoG) in 3D. A stable region is determined by analysing the distribution of the C.o.G. position for three subjects with different physiques who walked for 100 seconds with a walker. If the C.o.G is out of the region, the user may fall. The system then brakes enough to compensates for its lightweight and prevent the fall. Notice that the fall detection is restricted to the sagital plane.

In [33], The RT Walker is equipped with vision sensor to classify the user state among four classes : sitting, standing, falling, and walking. The classifier is based on heuristic on the distance between user head, hands and shoulder. Basically, the vision algorithms are based on head tracking, and skin detection. Shoulder detection is performed by finding the higher points of a uniform color region under the head, which seems to lack robustness with regards to environment properties and user clothes.

j) Observing walker odometers: The Assisted Navigation Guide (ANG walker) is based on a 4 wheeled Rollator [27] equipped with accelerometers and wheel encoders (cf. 8). The rear wheels are motorized, a bistable clutching mechanism allowing to clutch and unclutch the actuators at will. Two modes are then available : *free mode* were the motors are unlashed but the wheel rotation are measured and the *motor mode* where the motor are clutched and produce motion help or servo brakes. The acceleration of the walker and the velocity of the wheels are directly linked to the user

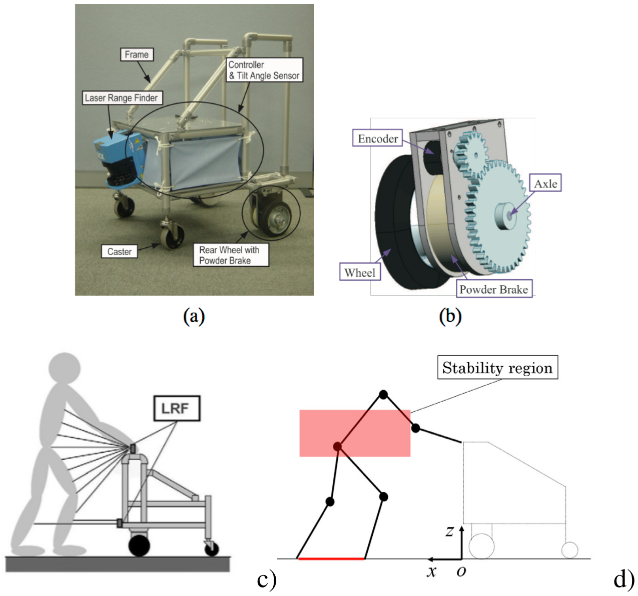


Fig. 7. a) RT Walker prototype [31] is a passive device using servo brakes (b) rather than actuators for obstacle avoidance and fall prevention. c) is the Laser range finder set up that is used to estimate a 7-link sagital human body model [32]. d) depicts the C.o.G. stability area.

Fig. 8. The Assisted Navigation Guide (ANG walker) [34] is equipped with various sensors and a bistable clutching mechanism

state. The fall detection system is based on a processing of these values.

D. Partial conclusion

Monitoring systems allow to estimate biomechanical features in ambulatory conditions, i.e. in uncontrolled conditions. Information about the walker itself (odometry, inertial parameters) or user-walker interaction seems more robust in these condition than laser data, video or distance sensors because they do not rely on environmental parameters such as user clothes and lightning condition. Yet, these sensors can provide useful and complementary information about user posture.

Walking assistance based on user intend allows to control the direction of the walker. Fall prevention algorithm tend to draw the boundaries between "normal situations" and "risk of fall". Most of the time, the control strategy consists in braking to stop the system.

IV. Conclusion and future works

Standard biomechanical features such as walking speed, cadence, step length can be estimated from observing rollator walking. Yet, rollator walking constraints the walk and modify posture and gait. For example, arm swing is a relevant parameter for balance estimation that can not be observed. Some other information seems hard to obtained without equipping the user (3D feet positions, force pressure distribution on the ground). On the opposite, it provides additional information with regards to a standard locometer, such as gait width.

This survey on existing smart walker raises several questions. In previous works, the equilibrium model is restrained to sagital plane and most of the works assume the walk to be symmetrical. Would it be possible to use a rollator to evaluate out of the plane motions and falls ? Regarding modification of the posture and gait, further study should focus on human-walker model to asses if we can evaluate autonomous walker with an assistance device they do not need. If the user does not use the device properly, could we still measure some relevant gait parameters ? And finally could we obtain an estimation of a free gait with studying rollator walking ?

Use of 6 D.o.F. sensor allow to have an idea of the force someone put on a handle. It might be linked with its global muscular force.

References

[1] B. Joyce and K. R.L., "Canes, crutches and walkers," *American Family Physician*, vol. 2, no. 43, pp. 535–542, 1991.
[2] D. W. B. Van Hook F.W., Demonbreun, "Ambulatory devices for chronic gait disorders in the elderly," *Americain Familly Physician*, vol. 8, no. 67, pp. 1717–1724, 2003.
[3] A. G. Alexander, "Gait disorders: search for multiple causes," *Cleve Clinінical Journal Medecine*, vol. 72, pp. 586–600, 2005.
[4] S. Mathias and U. Nayak, "Balance in elderly patients: the" get-up and go" test." *Archives of physical medicine and rehabilitation*, vol. 67, no. 6, p. 387, 1986.
[5] D. A. Winter, *Biomechanics and Motor Control of Human Movement*. New York: Wiley, 1990.
[6] G. Wua, F. van der Helmb, H. Veegerc, M. Makhsouse, P. Van Royf, C. Angling, J. Nagelsh, M. K. Kardunai, A.R., X. Wangk, W. Wernerl, F., and B. Buchholzm, "Isb recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—part ii: shoulder, elbow, wrist and hand," *Journal of Biomechanics*, vol. 38, p. 981–992, 2005.
[7] J. Mahoney, R. Euhardy, and M. Carnes, "A comparison of a two wheeled walker and a three wheeled walker in geriatric population," *Journal of American Geriatric Society*, vol. 40, pp. 208–212, 1992.
[8] S. S. A. Brandt, 1. Iwarsson, "Satisfaction with rollator among community-living users: a follow-up study," *Disability and Rehabilitation*, 2003.
[9] P. Levangie, M. Guihan, P. Meyer, and K. Stuhr, "Effect of altering handle position of a rolling walker on gait in children with cerebral palsy," *Phys Ther*, vol. 69, pp. 130–134, 1989.
[10] D. M. K. T. S. J. Willson, J. Torry M.R., "Effects of walking poles on lower extremity gait mechanics," *Medecine and Science in Sports and Exercise*, vol. 33, pp. 142–147, 2001.
[11] J. Youdas, "Partial weight-bearing gait using conventional assist devices," *Archive of Physical Medecine and Rehabilitation*, vol. 86, pp. 394–398, 2005.
[12] T. Alkjaer, P. Larsen, G. Pedersen, N. L.H., and S. E.B., "Biomechnical analysis of rollator walking," in *BioMedical Engineering OnLine*, 2006, pp. 1–7.
[13] C. Vaughan, B. Davis, and J. O'Connor, *Dynamics of human gait*. Champaign, Illinois: Human Kinetics Publishers, 1992.

[14] P. Devita and T. Hortobaygyi, "Gait biomechnics are not normal after anterior cruciate ligament reconstruction and accelerated rehabilitation," *Medecine and Science in Sports and Exercise*, vol. 30, pp. 1481–1488, 1998.

[15] M. Peshkin, D. Brown, J. J. Santos-Munee, A. Maklin, E. Lewis, J. E. Colgate, J. Patton, and D. Schwandt, "Kineassist: A robotic overground gait and balance training device," in *IEEE International Conference on Rehabilitation Robotics*, Chicago, IL, USA, June 28 - July 1 2005, pp. 241–246.

[16] A. Veg and D. Popovic, "Walkaround: Mobile balance support for therapy of walking," *IEEE Trans Neural Syst Rehabil Eng*, vol. 16, no. 3, pp. 264–269, June 2008.

[17] M. Spenko, H. Yu, and S. Dubowsky, "Robotic personal aids for mobility and monitoring for the elderly," *IEEE Transaction on Neural Systems and REhabilitation Engineering*, vol. 14, no. 3, pp. 344–351, 2006.

[18] H. Yu, M. Spenko, and S. Dubowsky, "An adaptive shared control system for an intelligent mobility aid for the elderly," *Journal of Autonomous Robots*, vol. 15, no. 1, pp. 53–66, 2003.

[19] Y. Hirata, A. Hara, and K. Kosuge, "Motion control of passive intelligent walker using servo brakes," *Robotics, IEEE Transactions on*, vol. 23, no. 5, pp. 981 –990, oct. 2007.

[20] S. MacNamara and G. Lacey, "A smart walker for the frail visually impaired," in *IEEE International Conference on Robotics and Automation*, vol. 2, 2000, pp. 1354–1359.

[21] S. Kotani, H. Mori, and N. Kiyohiro, "Development of the robotic travel aid hitomi," *Journal of Autonomous Robots*, vol. 17, pp. 119–128, 1996.

[22] A. Frizera, R. Ceres, J. Pons, A. Abellanas, and R. Raya, "The smart walkers as geriatric assistive device. the simbiosis purpose." in *International Conference of the International Society for Gerontechnology*, 2008, pp. 1–6.

[23] M. Martins, C. Santos, A. Frizera, and R. Ceres, "Assistive mobility devices focusing on smart walkers: classification and review," *Robotics and Autonomous Systems*, December 2011.

[24] M. Alwan, A. Ledoux, G. Wasson, P. Sheth, and H. C., "Basic walker-assisted gait characteristics derived from forces and moments exerted on the walker's handles: Results on normal subjects," *Medical Engineering and physics*, vol. 29, p. 380–389, 2007.

[25] J. Tung, "Development and evaluation of the iwalker : An instrumented rolling walker to assess balance and mobility in everyday activitiess," Ph.D. dissertation, University of Tonroto, 2010.

[26] A. Frizera, J. Gallego, E. Racon de Lima, A. Abellanas, J. Pons, and R. Ceres, "On line cadence estimation through force interaction in walker assisted gait," in *ISSNIP Biosignals and Biorobotics Conference*, Vitoria, Brazil, 2010, pp. 1–5.

[27] J. Merlet, *New trends in Mechanism Science: Analysis and Design*. Springer, 2012, ch. Preliminary design of ANG, a low cost automated walker for elderly, pp. 529–536.

[28] *Feasibility of objectively assessing foot placement during rollator use*. Toronto Rehabilitation Ressearch Day, Toronto, Canada, 2009.

[29] Y. Hirata, T. Baba, and K. Kosuge, "Motion control of omni-directional type walking support system walking helper," in *IEEE Workshop Robot Human Interactive Commun*, 2003, pp. 85–90.

[30] A. Frizera, "Extraction of user's navigation commands from upper body force interaction in walker assisted gait," in *BioMedical Engineering Online*.

[31] Y. Hirata, A. Muraki, and K. Kosuge, "Motion control of intelligent passive-type walker for fall-prevention function based on estimation of user state," in *Robotics and Automation, ICRA 2006, IEEE International Conference on*, May 2006, pp. 3498–3503.

[32] Y. Hirata, S. Komatsuda, and K. Kosuge, "Fall prevention control of passive intelligent walker based on human model," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, sept. 2008, pp. 1222 –1228.

[33] S. Taghvaei, Y. Hirata, and K. Kosuge, "Vision-based human state estimation to control an intelligent passive walker," in *IEEE/SICE International Symposium on System Integration (SII)*, 2010, pp. 146–151.

[34] J. Merlet, "Preliminary design of ang, a low cost automated walker for elderly," in *3rd European Conference on Mechanism Science*, Cluj-Napoca, Romania, 14-18 Septembre 2010.

Session IV

Robotics for elderly and frail people

- **Keynote speaker : Charles Fattal (MD, PhD, President of the Scientific Committee for the National Expert Center on Assistive Robotics Laboratoire m2h, France)**
  Title : Assistive robotics for gestual facilitation and disability compensation in persons with quadriplegia

- Title : **User studies of a mobile assistance robot for supporting elderly : methodology and results.** Authors : A. Garzo, L. Martinez, M. Isken, D. Lowet and A. Remazeilles

- Title : **ANG, a family of multi-mode, low-cost walking aid.**
  Authors : J-P. Merlet

- Title : **Velocity Control for Walk Assistance by Endeffector Force in the Leg Coordinate based on the Biarticularly-actuated System.**
  Authors : S. Sonokawa, Y. Kim, S. Oh and Y. Hori

Keynote Speaker : **Charles Fattal**
(MD, PhD, President of the Scientific Committee for the National Expert Center on Assistive Robotics Laboratoire m2h, Centre Mutualiste Neurologique Propara, Montpellier, France)

Assistive robotics for gestual facilitation and disability compensation in persons with quadriplegia

Abstract : For the past 20 years, assistive robotics for manipulation offers practical, useful solutions that are available on the market for persons with quadriplegia. However, there is still the need to validate the effectiveness and reliability of the impact of these technologies in the quest for a more independent life and better quality of life in this population. In order to achieve these objectives, it is essential to proceed with fine and thorough analysis of the users needs and expectations and ensure that the evolution potential, adaptation and intuitivism of the command interfaces are optimal in order for the user to appropriate these new technologies. Finally it is also important that maintenance and after-sale services be reactive and efficient to avoid adding a detrimental technological dependence on top of physical impairments.

Assistive robotics for gestual facilitation and disability compensation in persons with quadriplegia

Fattal C[1,2,3,4]. Leynaert V[1,2,3], Petit C[1,3], Gilbert C[1,3], Clement B[1,3], Emica C[1,3].

I. Introduction

Quadriplegia is most often caused by trauma to the spinal cord with an injury at and above C8-T1. Medical advances have increased life expectancy of persons with quadriplegia who are ventilator-dependent (C1 à C4) or with a high level of injury but without the need for ventilation (C5-C6). However, even when the cervical injury is lower (C7-C8-T1) and the upper limbs are less impaired, medical and interacting life events (accident or secondary disease, aging) can lead to an « added disability » and increased dependence. The need to «save energy » on certain movements and/or limit efforts, can, in persons with quadriplegia with a low level of cervical injury justify the use of technologies for disability compensation.

For the past 20 years, assistive robotics for manipulation offers practical, useful solutions that are available on the market for persons with quadriplegia. However, there is still the need to validate the effectiveness and reliability of the impact of these technologies in the quest for a more independent life and better quality of life in this population. **In order to achieve these objectives, it is essential to proceed with fine and thorough analysis of the users' needs and expectations and ensure that the evolution potential, adaptation and intuitivism of the command interfaces are optimal in order for the user to appropriate these new technologies**. Finally it is also important that maintenance and after-sale services be reactive and efficient to avoid adding a detrimental technological dependence on top of physical impairments.

II. Assistive robotics for manipulation tasks: applications

The main applications appear to be: drinking and eating (grabbing and bringing up to the mouth, cutting up foods, opening containers), self-care and facial care, getting dressed/undressed and finally reaching, grasping and releasing ordinary daily life objects located up or down. The patient's hopes for functional improvement should not distract from the fact that, regardless of the person's level of dependence, it is essential for the user to remain in charge of the robotic device interface. However, robotics is a dense and complex technology, put in the hands of a non-specialized user, in a poorly structured and highly diverse environment. This underlines the need for technological development of intuitive, versatile, flexible and diversified command interfaces in order to meet the various levels of gestual impairments and/or upper limb movement limitation and improve the user's appropriation of the device. It is also important for all levels of command to be represented: *manual* when the user controls the device, *automatic* when the user chooses a task and lets the robot perform it, *shared or semi-automatic* (the user and the robot cooperate to perform the task). Safety and reliability are essential components, and several safety devices should be mounted with the robot to respond in real time and an emergency switch-off system is mandatory.

III. A wide array of robotic devices for assisted manipulation

Assistive robotics for manipulation is one of the facets of a large panel of robotic devices aimed at persons with loss of autonomy either to compensate the functional impairment or to assist functional rehabilitation (table 1).

Robotic arms mounted on wheelchairs are the devices that benefited from the most advanced technological knowledge and are more widely available on the market than any other device. The underlying concept is to propose a user-controlled device that would replace what several technical or human aids could offer. The cost-related argument is also quite important because it would save on acquiring several technical resources and human caregivers could focus on more social-related tasks for the patient (outings, taking time to play games or watch a movie, etc.)

We can differentiate several devices:

- Robotic arms that are already on the market, ready to be launched or part of a « technological display » allowing researchers and engineers to benefit from the advanced know-how of other teams.
- Uni or bilateral robotic arms mounted on a wheelchair or mobile base.

The reference product today is the ***JACO®* robotic arm**, on the one hand it is the most promising arm on the market in terms of functional contribution and ergonomics (fast moving speed, easy learning, lightweight, and easily

1. CEN Robotique, 34090 Montpellier
2. Association Approche, 34090 Montpellier
3. CMN Propara, 263 rue du Caducée, 34090 Montpellier
4. Laboratoire M2H (Staps, UM1) 34090 Montpellier

transportable) and on the other hand, it is, along with the ***I-ARM®*** a true reflection of the research works that for the past 30 years have tried to mount on wheelchairs an assistive device for manipulation in order for users to grasp, release and manipulate objects above and under shoulder level.

In the history of marketed robotic arms, the ***MANUS®*** was for a long time considered the leader of them all. Manufactured and marketed in the nineties by Exact Dynamics in the Netherlands, this robotic arm succeeded in proving that inserting robotics in the environment of a disabled user and surrounding caregivers was indeed possible. This robotic arm was actually an order from the Ministry of Health in the Netherlands in the context of a government program. It is the perfect illustration that the government's desire to promote new technologies for assisting persons with disabilities could result in home-based high-technological devices while providing a common maintenance service.

This 6-degree of freedom arm with a 2-finger extremity pinch, like the Manus initially and with the I-Arm nowadays, is mounted on the seat or arm rest. With an initial weight of 20 kilos it has been replaced by a product that is now twice as light – the I-Arm is globally similar in terms of technical parameters especially with its "step-by-step" command interface, which is not very intuitive and could be improved. The Manus has now been scaled down to be a part of a « technological display ». We now find this arm associated with mobile bases or other types of prototype devices (**TAURO**, **ANSO, ARPH**, projects)

The most recent Jaco arm, marketed by a Canadian company, took another technological approach in terms of:

- material – carbon fiber cuts the weight of the arm in half-
- the intuitive command interface – the system is open and accessible to configure adapted interfaces. It can be interfaced with an environment control panel mounted on the wheelchair.
- amplitude (90 cm) and functional capacity (7 degrees of freedom). It can reach objects on the floor as well as objects located right in front of the patient or above the user's head.

In the cases of the Jaco or I-arm, the input of the user in decoding the grabbing strategy remains important and demanding.

Regarding the American arm, the ***RAPTOR®*****,** designed in 2000, it was granted FDA approval in the USA. It is not as advanced as its competitors but is much cheaper.

Finally the ***BRIDGIT®***, like the other two arms -**MATS** and **ASIMOV-**confined to an experimental stage, offer a bilateral manipulation either with a railing mounted at the back of the wheelchair's seat, allowing the arm to move from the left to the right of the user and vice versa, or with the option to mount the arm on a control station on both sides of the wheelchair.

IV. Conclusion

Fiction has met reality and robotic arms for manipulation are available on the market. However, everything still remains to be done to implant these devices in the daily lives of persons with disabilities because the encounter between a high-technology device and a person is not limited to a sole user. Family and professional caregivers play an important role in the appropriation and acceptation of the device.

The role of the National Expert Center in Robotics (CENRob) is to conduct a thorough inventory of the existing marketed products and to evaluate other products awaiting industrial transfer. Among other objectives, the CENRob works on the transfer of such devices in the users' daily life environment and, analyze the role of high-technology aids in a caregiver environment. The CENRob, in partnership with users, patients' associations, researchers and engineers, clinical networks and other partners has the challenge to convey the credibility of these devices in daily uses but also to « impose » that potential users play a part in the development of prototypes before and after manufacturing in order to make sure that the device meets the functional needs and expectations of end users. Assistive robotics should not be developed for persons with quadriplegia but in partnership with them.

Table 1 : Classification of Assistive Robotics
(CENRob, Montpellier)

1) *1. ASSISTIVE ROBOTS FOR MANIPULATION*

1.1 Work stations

1.2 Fixed single-task robots

1.3 Robotic manipulator arms mounted on wheelchairs

1.4 Robotic manipulator arms mounted on mobile supports

1.5 Neuroprosthesis, prostheses and exoskeleton of the upper limb

1.6 Robots dedicated to handling and transporting wheelchairs

2) *2. ASSISTIVE ROBOTS FOR DOMESTIC TASKS*

2.1 Single-task robot

2.1.1 Robots for vacuuming

2.1.2 Robots for cleaning

2.1.3 Robots for mowing the lawn

2.1.4 Other robots

2.2 Multi-task humanoid and android robots

3) *3. ASSISTIVE ROBOTS FOR SOCIAL INTERACTIONs*

3.1 Friendly robots / videomonitoring robots

3.2 Robots for sensory awakening (social or cognitive robots)

4) *4. MOBILITY ROBOTS*

4.1 Exoskeleton of the lower limbs

4.2 Neuroprostheses and prostheses of the lower limbs

4.3 Smart wheelchairs and walkers

4.4 Robots to help caregivers with locomotion

5) *5. ASSISTIVE ROBOTS FOR REHABILITATION*

5.1 Rehabilitation robots for the lower limbs

5.2 Rehabilitation robots for the upper limbs

User studies of a mobile assistance robot for supporting elderly: methodology and results

A. Garzo, L. Martinez, M. Isken, D. Lowet and A. Remazeilles,

Abstract— **The FP7 European project Florence aims at investigating how current state of the art robotic technology can be used to support elderly to live longer independently at home. During the whole development for our robotic platform users and caregivers were strongly involved through focus groups, Wizard of Oz tests and functional validation within controlled environments. In this project we focus on type of services: on the one hand we consider lifestyle activities like tele-presence and coaching services and on the other hand we also propose safety services like fall handling. This article describes the user-centered mechanisms we put in place during this project and compiles the information we could gather during interaction with potential end-users. The collected data naturally strongly influences our development. They are provided here as well as general guidelines for any further assistive mobile robot development.**

I. Introduction

The European project Florence is interested in improving the well-being of elderly (and their beloved ones) as well as the efficiency in care through Ambient Assisted Living (AAL) services supported by a mobile robot platform. The Florence project investigates the use of such robots to deliver new kinds of AAL services to elderly persons and their care providers. Florence places the robot as *the* connecting element between several traditional stand alone AAL services deployed in a living environment as well as between these AAL services and the elderly. It is needed supposed that through these safety-, care-, coaching- and connectedness-services, as proposed by Florence, the elderly could remain much longer independent.

The project aims to create a low-cost solution based on current state of the art technology. The Florence robot is a mobile platform[1] (Fig. 1), 1,5 meter tall, with a touch screen interface to perform traditional interaction (no physical manipulation capabilities are so far embedded).

*Research supported by the 7th Framework Programme by the EC, grant ICT-2009-248730.

A. Garzo (corresponding author), L. Martinez and A. Remazeilles are with the Assistive Technologies Dep. from TECNALIA, San Sebastian, 20009 Spain (phone: (+34) 946-430-850; fax: (+34) 944-041-445 e-mail: ainara.garzo@tecnalia.com, leire.martinez@tecnalia.com, anthony.remazeilles@tecnalia.com).

M. Isken is with OFFIS Institute of Information Technology, Escherweg 2, Oldenburg, 26121 Germany (melvin.isken@offis.de).

D. Lowet is with Philips Innovation Group Research, Eindhoven, 5656 AE Netherlands (dietwig.lowet@philips.com).

[1] In its first version the Florence robot was based on the PekeeII platform of Wany Robotics (http://www.wanyrobotics.com/)

Figure 1. Florence robot platform

The robot is equipped with several sensors such as a 2D laser scanner, a RGB-D camera (Kinect[2]) and a regular color camera. The robot software is developed with the Robotic Operating System, ROS [1] – the emerging de facto standard in robotic software.

The project follows a **user-centered** and **iterative development process** [2] [3]. According to this methodology users are involved in all the steps of the project from the definition of requirements to the design and development of the system, which is permanently revised and reconsidered from the requirements and guidelines of the end-users. This paper highlights the methodology we followed. In the section II of this article some background about similar projects can be found. Sections III, IV, V and VI include information about the different testing that were carried out with final users: brainstorm sessions, focus groups, Wizard of Oz and testing in controlled environments. The conclusions of this study are explained in section VII.

II. Background

Service robotics is at the intersection of different areas of robotic research. Already over 20 years ago, Engelberger classified service robotics into different application fields [4]. Amongst them are: medical robotics, health care and rehabilitation, assistance to the handicapped and the elderly, commercial cleaning, household tasks, military services and construction or surveillance.

As technology advances, robotic solutions are getting more autonomous and flexible. In the case of health care, robotics research started with fixed workstations [5], going over wheelchair-mounted systems and intelligent wheelchairs [6] [7], to autonomous mobile robot systems [8]. Today

[2] http://www.xbox.com/es-es/kinect

several mobile robots for health assistance are available (commercially, or for research purposes). Systems like Nursebot [9], Robocare [10], Wakamaru[3], or Care-O-bot [8] propose to help, guide, and assist people at home. To do so they generally provide services like tele-presence (video conferencing-based), automatic reminders (food, drug intake), automatic emergency management (detect harmful events and notify doctors or care providers) and companionship (conversation, playing games). Most of these platforms are still in (even though advanced) research states, improving aspects like the autonomy in home-like environments [11], the learning of environmental factors and user behaviors [12] as well as the robot design itself [13].

For the presented kind of assistive robotics the involvement of the stakeholders in the design and development is crucial because these robots are developed to work and cooperate in environments with humans [14] [15]. In Florence project, as illustrated on Fig. 2, several interaction tools taken from user-centric methodologies were used [16], depending on the current stage of the project.

In all the sessions (Focus Groups, Wizard of Oz and Controlled Testing) information and consent form were handed out. In all cases, these forms were read out by the moderator of each session explaining the project and the legal-ethical considerations to be signed by the participants.

In the following sections, the different sessions, the procedure, the collected data and the results are described.

III. Brainstorm Sessions

In order to get a first list of potential lifestyle and AAL services for the elderly a brainstorming session were done by the project team. The researchers who participated in these brainstorming have wide experience in elderly needs and assistive robotics. The key areas identified during this initial work were the following:

- *Coaching*, by giving feedback on specific activities like physical exercises, and advices on activities of daily living, with the objective to encourage the person to be active.
- *Social inclusion*, by supporting access to the social networks, including web-2.0, aiming at keeping a connection to family and friends.
- *Safety*, by using Florence as additional ears and eyes in comfort or safety situations, controlled by service providers or the elderly themselves (crisis or emergency detection, smoke detection, personal alarm, water-damage …), with the clear objective to get a quicker management of critical situations.
- *Care monitoring,* by monitoring the activities of the person, and reminding health-related tasks (like taking pills).

After the brainstorm sessions some use cases were defined to prepare the focus groups. The definition of these use cases can be found in the TABLE I.

[3] http://www.mhi.co.jp/en/products/detail/wakamaru.html

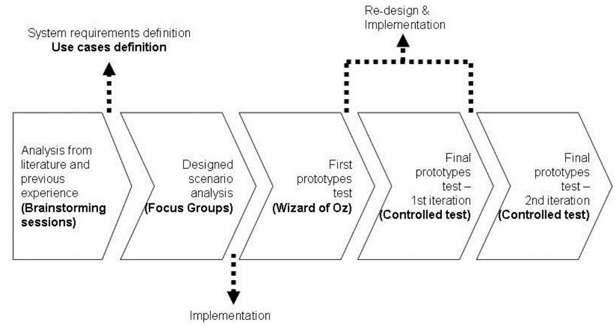


Figure 2. Phases (arrow) and tools (in bold) used along the Florence project

IV. Focus Groups

The Focus Group (FG) technique is one of the most famous procedures for opinion and experience capturing. It can be used to bring together a cross-section of different views in a discussion group format. This tool is useful in requirements specification and provides a multi-faceted perspective on them [16]. Even though there is a large bibliography on FG and its variants, there is still controversy about the implementation of FGs. That is the reason why the project consortium decided to follow an adaptive approach focused on the flow of the discussion instead of a desktop pre-plan.

A. Procedure

Once the Use Cases were defined some FGs were conducted for defining context and early design with stakeholders[4]. Also the FGs were used to validate the defined scenarios and use cases with the final users and to identify user requirements and needs.

TABLE I. Use Cases

Acronym	Definition	Area
KEETOU	Innovative tool for communication with their familiy and friends.	Social inclusion
HOMINT	Advanced home interface for remote control of the home (lights, doors, windows, etc.).	Safety
FALHAN	Fall situation handling.	Safety
AGEREM	Agenda reminder.	Care monitoring
LIFIMP	Lifestyle improvement encourages the user to do exercise.	Coaching
DEVCOA	Device coach. Sharing information with other users for doing difficult tasks as cook, sewing, …	Coaching
COLGAM	Collaborative gaming using videoconference to share the same physical game site.	Social inclusion
LOGSYS	A service providing an overview of constantly logged data like daily activities, physiological parameters etc.	Care monitoring

[4] Stakeholder: individual or organization having a right, share, claim or interest in a system or in its possession of characteristics that meet their needs and expectations [3]. In this case final users, caregivers and relatives where involved.

FGs were held in Netherlands, Germany and Spain to compare the culture differences. For these sessions the homogeneity of the participants' profiles was considered as an important aspect because people tend to be more comfortable in a group with similar features. In the case of Spain the FGs were made not only with final users, but also with professionals (9 people). All the participants representing final users in the FGs were +65 and with some sort of experience with technology. In total around 30 participants took part in this FG sessions. A facilitator conducted the session to focus on the use cases defined.

B. *Results*

The following conclusions were taken for each use case:

- **KEETOU.** Users suggested keeping track of the amount of visits the parent had in the near past and how many visits are expected in the near future. Another suggestion of the users was to have the possibility of measure some mood-related aspects (stress level, low-activity level, and sleep patterns) and make them accessible to the caregivers such that they could decide whether this could be good moment to virtually visit the elderly person.
- **HOMINT.** Participants transmitted an interest not only in the automatic home control (like *close windows*) but also in the addition of sensors monitoring the place and eventually detecting emergencies situations, like infrared cameras to detect heat and fire or smoke detectors.
- **FALHAN.** In a first step we wanted to use a wearable emergency button but the participants asked the robot should be able to detect the fall without the need of external hardware. The service should be designed that the user is able to cancel the alarm.
- **AGEREM.** The service should be designed to allow for the addition of events such as relative's birthdays. Ease of use should be obviously a primary focus in the design of the service.
- **LIFIMP.** The comments of the participants during FGs do not directly affect this use case. In this case the feedback from different FGs was somehow contradictory so the collected information was not taken into account.
- **DEVCOA.** Participants decided that this scenario was less relevant than others because they would rather consult other sources of information such asking kids, neighbors … So, after the FGs this use case has been discarded.
- **COLGAM.** Including gesture pointing recognition was well accepted by the users.
- **LOGSYS.** The data access control is a critical aspect of this scenario. In this use case the privacy of the user personal data was considered very important by the users. The user has to be able to exactly choose the persons that should have access.

The comments from the users were discussed technically and most of them were included in the re-design of the system, which was tested by a Wizard of Oz (WoZ) as it is explained in the next section.

V. Wizard of Oz

The origin of this test dates back to 1984 [17]. This tool takes its name from the story *The Wizard of Oz*, and more exactly it takes the name from the figure of the character under the curtain that controls everything without revealing his presence. It is a technique used to test a product or a service in a detailed way by observing the interaction of a potential user with the object without revealing the evaluator's presence [18]. WoZ is very useful when new services or systems want to be tested, considering that these first trials are generally performed with a kind of remotely control prototype, reducing the need of complete autonomy. WoZ permits thus to use a system at an early stage and to test already the user acceptance, enabling this way to better align the next developments.

A. *Procedure*

With the first implementation of the robot platform WoZ testing with end users were held in Novay, Philips and OFFIS facilities to have the chance of comparing results from different countries. For WoZ the supervision of the regional Ethical Committee was asked in each country.

In the WoZ conducted the time for each participant was about 40-45 minutes organized with the following structure:

- Welcome the participant
- Explain her the goal and the structure of the session
- Ask participant about first impressions about the robot (e.g. what do you like / dislike about the robot's looks / moves; what do you think it can do)
- The participant uses the robot
- And at the end, ask again participants about second impressions and reflections; e.g. Now that you've played with the robot: i) what are your overall impressions; ii) what is it good in or bad at?; iii) what aspect should we improve, from its appearance to the way it moves.

The robot platform was active for maximum 30 minutes per participant due to the robot battery autonomy. For the WoZ this time was considered sufficient to get some relevant feedbacks without exhausting the user. At a longer term, this element could be critical and should be addressed by improving the battery lifetime and/or providing the robot with some automatic docking for battery reload capacities.

In each WoZ session a participant (final user, elderly), a researcher (observer) and a wizard took part. In total 17 final users and caregivers participated in these sessions.

B. *Results*

After WoZ some conclusions were taken:

- **KEETOU.** The participants specially focused on privacy, safe communication and transfer of health data. They also proposed having an appointment manager and connection to pharmacy resources which should include remote payment capability or payment when delivered at home.
- **HOMINT.** The participants suggested adapting interfaces in other environments: functionalities provided by the different spaces such as hospitals, care centers, supermarkets or offices. This suggestion was discarded because Florence project is focused in home environments.
- **FALHAN.** We also saw that the robot should not always follow the user because it could be annoying.
- **AGEREM.** The users liked the voice interaction to add some items in the calendar (an generally speaking for any type of action) and commands were easy to learn for them. In addition to the voice control, a touch screen is also desired by the users. Messages and notifications should stay on the screen until they have been noticed and therefore they need to be acknowledged by the user. Depending on the situation, for routinely interruptions the participants prefer a low profile melody whereas for something urgent they prefer an alarm sound.
- **LIFIMP.** The robot should not be too "pushy" regarding giving advices. The user would appreciate to view on the screen health status and progress towards health-related goals, such that the user is informed and motivated to perform an activity.
- **COLGAM.** According to the users, the robot should not move by itself during collaborative activities and should only move on user's demand. User should always be in control. The users considered this service less important than others, so that we decided to discard it from the following testing.
- **LOGSYS.** As in FALHAN the researchers concluded the robot should not approach nearer than 50cm from the user unless interaction is required.

VI. CONTROLLED TESTING

Using the WoZ conclusions, the development of the robotic platform was refined again before the next testing done in controlled environments (or Home Lab) specifically designed for this kind of evaluation.

Controlled experiments permit to test the usability of a prototype in a place similar to the envisioned environment of use, and giving the impression to the user to interact with a finalized system. Representative final users are asked to perform tasks with the prototype, enabling this way to analyze, and eventually improve and extend the specifications of the user requirements [16].

In the Florence project, tests in controlled environment were conducted in two sites, the Philips and the OFFIS labs that are now described.

A. Controlled Home Environments

Philips Experience Lab

The Philips Home Lab (Fig. 3) is a permanent fully functional home laboratory enabling the Philips researchers to better understand the user needs and motivations to use technology, and bring better products to market in the quickest possible timeframe.

The Philips Home Lab is built as a two-storey house with a living room, a kitchen, two bedrooms and a bathroom. The observation room adjacent to the flat has a direct view into the Home Lab. Signals captured by the cameras can be monitored on any of the four observation stations, each observation station being equipped with two monitors and one desktop computer.

The user tests at the Philips Home Lab took place mainly in the living room and the kitchen of the lab. The layout of both rooms in terms of furniture placement and equipment can be considered typical and no specific alterations have been made, except for the installation of the home sensors and actuators for the user test (a doorbell, a window-state sensor, a temperature sensor, an emergency button and a wireless weighing scale).

OFFIS IDEAAL Living Lab

The IDEAAL Living Lab (Integrated Development Environment for Ambient Assisted Living) consists of a senior apartment in the OFFIS institute building (Fig. 4). It is a fully functional two room flat consisting of a living room, a sleeping/working room, a kitchen, a bathroom and a corridor. The apartment reproduces all areas of normal life. Beside a home automation infrastructure to present today's technologies, the Living Lab demonstrates research and development results of several AAL projects at the OFFIS institute. The realization of the IDEAAL apartment was geared to the taste of the focus group by including the opinion of an example couple.

The user tests at the OFFIS IDEAAL Living Lab were performed mainly within the living room. The room was used "as it was", i.e. neither furniture nor equipment was altered for the user tests. Only additional hardware was installed, such as a weight scale for the LIFIMP service. For the use cases involving home automation, the already available infrastructure has been used.

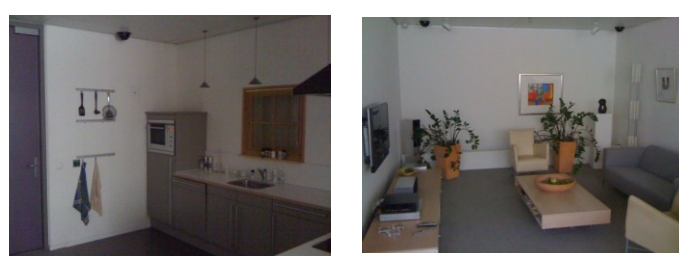


Figure 3. Philips Home Lab pictures

Figure 4. OFFIS Home Lab pictures

B. Procedure

The controlled home environment tests, presented in this section, evaluated the first version of the Florence software implementation. To ensure the safety of the test persons, these tests were conducted under constant supervision. The results of these tests were used to identify limitations and improve the Florence system in preparation to the Living Lab tests. For this testing the supervision of the Ethical Committee of each region was asked.

17 persons participated at the Philips Experience Lab: 8 persons aged between 66 and 82 together with a close acquaintance (mostly a son or a daughter), plus an elderly person. In addition, two professionals of a care institute took part in the tests as well. The elderly participants were selected based on the following inclusion criteria: i) living independently at home; ii) living alone; iii) being healthy.

18 elderly persons (aged between 60 and 80) participated at the user tests in the OFFIS IDEAAL Living Lab. Participants were selected using the same inclusion criteria as in Philips Lab and, if possible (but not necessary) having some kind of history with home accidents like falls.

The difference on the last criterion in the sites enables to get an even broader range of feedback. Furthermore Philips concentrated their effort more on the social aspect of the system whereas OFFIS highlighted the safety aspects.

During the user tests at the Philips Home Lab the elderly tested the Florence platform under the guidance of a user experience researcher, while a second researcher took notes (and assisted if necessary). A third researcher interviewed the close acquaintance while they observed the session from the supervision room.

During the user tests at OFFIS two researchers were present. One researcher led the interview and took notes, and the other was handling any technological aspect, like controlling the robot and home automation.

Due to the limited time frame for each test, not all services could be tested. A subset of services was selected, trying to cover the main functionalities of the intended system. Services not used were presented through presentations to get some feedback. In general the tests started with a short introduction to the user, followed by a testing of the services and ended with a feedback / evaluation phase.

C. Results

Some general points can be extracted from both experiences:

- Feedback depended strongly on:
    - Their experience with technology: surprise or comparison with other products.
    - Their experience with elderly people care (e.g. relatives suffering from MCI).
- Users' evaluations were very mixed, from no interest at all to very much interested. There was a strong correlation between the interest of the elderly in the robot services and the amount of care/support they currently received or had received in the past.
- The solution's evaluation heavily depends on what people have already experience with. People that were less familiar with what can be done by technology already available (e.g. Skype), were more positive as the entire demo surprised them more. The others often mentioned that it already exists.
- Robotic tele-presence was significantly appreciated by the close acquaintances population that appreciated to get extended videoconference solutions to interact with their parent and eventually get more quickly information when a potential risky situation has been detected.
- The interaction level in between the two agents seem to be another important aspect: if most of the users appreciate to get some advices from the robot, none would appreciate if theses indication turns to be considered as orders.
- Several users mentioned that even though the added value of the proposed services was interesting, this would not be enough to justify the need of buying such equipment. Several persons indeed mentioned that they would prefer the robot to be able to perform some additional household tasks like: cleaning the floor, bringing tea, taking product out of the cupboard, opening the gas burner, make up the bed.
- Many different people would like to use different services and also like different ways of communication. The configuration aspect (that would of course need to be totally transparent at the user's level) is thus another key aspect for user acceptance.
- From the technical point of view the services KEETOU, HOMINT, FALHAN and LOGSYS were well accepted, so no strong changes were defined for the next test iterations.
- For timing issues, AGEREM service was not tested at this stage, but we decided to include it in the next trials, using the using the Google Calendar[5].

[5] www.google.com/calendar?hl=en

- In the LIFIMP service, we were proposing to monitor the person weight, and, depending on the status, the service was suggesting some activities. From the feedbacks we received, we decided to change the parameter observed for the next trials, shifting to the blood pressure that was perceived as more health related and less intrusive.

VII. Conclusion

This paper has described the overall user involvement methodology the Florence project has put in place to get the users involved in the development and evaluation of a mobile robotic platform aiming at providing assistance to the elderly through well-designed services. We have been describing each evaluation iteration that has already taken place: brainstorming, Focus Group, Wizard of Oz and tests in Controlled environments.

During discussion, the participants in the sessions distinguished two different scenarios from the implemented services: communication (gathering services like KEETOU, FALHAN and COLGAM) and support including health care (through HOMINT, AGEREM and LIFIMP). From their point of view, the differences in between the services from a same group were not sufficient to justify a separate scenario. Starting from this point of view, we decided to reorganize the presentation and implementation of our services to present hem to the user using these two main concepts.

Technically speaking, the obstacle avoidance was also considered as an important issue by the technicians during the testing. At this moment, the obstacle detection was done with a single laser range scanning an horizontal plane close to the floor. This does not permit to correctly detect and handle objects that are located on other higher planes such as a table or a desktop. We are thus investigating the use of 3D obstacle avoidance, using the tridimensional occupancy grid implicitly provided by the 3D vision sensor we are now using.

Those new changes will be tested by elderly and their relatives and caregivers at their own homes at the end of this year, with the aim of verifying again the functionalities and acceptance, but also to analyze the usability.

For the future, we recommend to use a similar methodology to involve users in the design, the development and the validation of such assistive device, like for any other product. We saw that the definition of use cases must be done with final users and technical people, for checking the viability and the usefulness of the proposed solution in the same time. Some of the methods used in this research can be replaced by other relevant tools. In any case, we would recommend to perform at least three iterations of controlled tests before taking the robot to the houses, instead of only two as we did in this research. The first iteration is needed to check if the user needs and requirements are well covered. A second iteration can be used to validate the functionalities according to the results of the first evaluation. One last iteration is needed to test the final version of the system.

Acknowledgment

We would like to thank all partners from the Florence project and all the participants in the test sessions. Furthermore, we are very grateful to Ingema[6] for their contribution in the definition of elderly needs.

References

[1] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng, "ROS: an open-source Robot Operating System" in *Proc. IEEE ICRA, Workshop on Open Source Software*, Kobe, 2009.

[2] B. W. Boehm, "A Spiral Model of Software Development and Enhancement," *IEEE Computer*, vol. 21, May. 1988, pp. 61-72.

[3] ISO 9241-210: 2010 Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems.

[4] J. F. Engelberg, *Robotics in Service.* Cambridge, MA. MIT Press, 1989.

[5] J. Hammel, H. V. der Loos, and J. Leifer, "DeVAR transfer from R&D to vocational and educational settings," in *Proc. 4th Int. Conf. On Rehabilitation Robotics*, Wilmington, 1994, pp. 14-16.

[6] Y. Matsumoto, T. Ino, and T. Ogasawara, "Development of intelligent wheelchair system with face and gaze based interface" in *Conf. 2001 IEEE Ro-man*, pp. 262–267.

[7] M. Hillman, K. Hagan, S. Hagan, J. Jepson, and R. Orpwood, "A wheelchair mounted assistive robot" in *Conf. ICORR 1999*, pp. 86-91.

[8] U. Reiser, C. Connette, J. Fischer, J. Kubacki, A. Bubeck, F. Weisshardt, T. Jacobs, C. Parlitz, M. Hägele, and A. Verl, "Care-O-bot 3 - Creating a product vision for service robot applications by integrating design and technology," in *Conf. 2009 IEEE/RSJ IROS*, pp. 1992–1997.

[9] M. E. Pollack, S. Engberg, J. T. Matthews, S. Thrun, L. Brown, D. Colbry, C. Orosz, B. Peintner, S. Ramakrishnan, J. Dunbar-Jacob, C. McCarthy, M. Montemerlo, J. Pineau, and N. Roy, "Pearl: A Mobile Robotic Assistant for the Elderly," in *AAAI 2002 Workshop on Automation as Eldercare*.

[10] A. Cesta, G. Cortellessa, M. Giuliani, F. Pecora, M. Scopelliti, and L. Tiberio, "Psychological implications of domestic assistive technology for the elderly," in *PsychNology Journal*, vol. 5, 2007, pp. 229-252.

[11] A. Petrovskaya, and A. Ng "Probabilistic mobile manipulation in dynamic environments, with application to opening doors," in *Conf. 2007 IJCAI*, pp. 2178-2184.

[12] C. L. Breazeal, "Sociable Machines: Expressive Social Exchange between Humans and Robots," Ph.D dissertation, Dept. of Electric. Eng. & Comp. Sc., MIT, Cambridge, MA, 2000.

[13] C. Ray, F. Mondada, and R. Siegwart, "What do people expect from robots?" in *Conf. 2008 IEEE/RSJ IROS*, pp. 3816-3821.

[14] A. Weiss, R. Bernhaupt, M. Tscheligi, D. Wollherr, K Kühnlenz, and M Buss, "A methodological variation for acceptance evaluation of Human-Robot Interaction in public places," in *Conf. 2008 IEEE Ro-man*, pp. 713-718.

[15] M. Heerink, B. Kröse, V. Evers, and B. Wielinga, "The influence of a robot's social abilities on acceptance by elderly users," in *Conf. 2006 IEEE Ro-man*.

[16] M. C. Maguire, "User-Centred Requirements Handbook, " RESPECT Consortium.

[17] J.F. Kelley, "An Iterative Design Methodology for User-Friendly Natural Language Office Information Applications," *ACM TOIS*, vol. 2, Jan.1984, pp. 26-41.

[18] R. Tassi, "Design della comunicazione e design dei servizi. Il progetto della comunicazione per la fase di implementazione," Ph.D. dissertation, Politecnico di Milano, Milan, Italy, 2008.

[6] Ingema is a non-profit organization dedicated to generating knowledge in the field of the elderly and the disabled people. http://www.ingema.es/en

ANG, a family of multi-mode, low-cost walking aid

J-P. Merlet

***Abstract*— ANG is a family of low cost modular walking aids based on commercially available Rollator. We present two models of this family: the simplest `ANG-light` which is intended to be used as a diagnosis tool for walking and as a fall detection tool and `ANG-II`, a motorized walker with over 20 on-board sensors. Trajectography obtained through `ANG-light` with 24 healthy subjects has been used to obtain gold standards of walking pattern and the walker is currently being used at Nice hospital with 30 elderly end-users in order to determine if trajectory records obtained during the daily use of a walker may allow doctors to objectively characterize abnormal walking patterns, to follow the progress of a rehabilitation process and to detect emerging pathologies. The more sophisticated `ANG-II` is intended to be used as a test platform for original functionalities, some of which will be presented.**

I. INTRODUCTION

In aging societies the number of people suffering from locomotion problem because of the decline of muscular strength is increasing. A related problem is fall: in France fall account for 80% of accidents, 2/3 of them at home, and 9300 deaths per year are a direct consequence of a fall. This paper focuses on the design of walker-type support systems with the following motivations, which have been established after a 2 years discussion period with end-users:

1) *social acceptance*: the design is based on commercially available walker that are already accepted. Added elements may be hidden in the walker frame. The walker should look familiar, of reasonable size and not obviously intrusive.
2) *low cost*: the cost of systems such as humanoid robots is by far too high to be afforded by a vast majority of elderly. Our purpose is to design low-cost systems,
3) *modularity*: elderly have their own trajectory of life which requires a constant adaptation of the walker functionalities
4) *information provider for doctors*: a walker which is used daily may provide a wealth of information for following the health evolution of elderly users.
5) *safety*: as mentioned above fall is a major problem for elderly people. Hence a walker should be an element of a fall prevention/detection scheme and should be able to help if such an event occurs
6) *energy autonomy*: running out of battery power may lead to dramatic situation with frail people. Hence a strict power management and possibly on-board power source is a critical part of a walker system

Various passive and active walker have been proposed in the past: the passive walker of MARC at Virginia University [1],the active PAM-AID and its extension GUIDO [2], HLRP [3], PAMAID [4], NURSEBOT [5], the sit to stand devices MONIMAD of LRP [6] and of Chugo [7], IWalker [8], RT-Walker [9], the sophisticated CARE-o-BOT [10] and the omnidirectional walker of Chuy [11]. Although most of these walkers provide interesting functionalities, none of them fully satisfy the above requirements especially in terms of costs, intrusivity, safety and autonomy.

J-P. Merlet is with INRIA Sophia-Antipolis, email: `Jean-Pierre.Merlet@inria.fr`

II. THE `ANG` FAMILY

A. Motivations

Current walking aids provide basically a single function, motricity help, are largely accepted and low cost. Our motivation in this project is to divert daily life objects that are already accepted, so that they can provide additional functionalities. These added functionalities will depend upon the trajectory of life of the elderly and therefore must be modular. First we want to address a major request of the medical community: *provide objective information on the walking behavior of the end-user not only during visit but also at home*. This *monitoring* aspect is a major request for us. The second major request that we want to address is a safety issue: fall detection/prevention. Fall is a major problem for elderly people: in France over 9300 deaths of elderly people are a direct consequence of a fall. Apart of these two major objectives several functionalities may be added, that will be presented in the next sections.

However we have other guidelines for developing assistance walker:

- *low cost*: this is an essential requirement as the social security system cannot afford expensive devices
- *low energy comsumption*: this is an often neglected subject which is however an important issue. Battery charge must not be an issue for elderly people and hence we have to provide devices with long lasting life and possibly with an alternative energy source that will allow the device to ensure minimal funtionalities,
- *smart devices*: assistance devices cannot be considered as isolated systems. They must be able to communicate with the outside world and with other devices (e.g. to deal with an emergency)
- *social acceptance*: the devices must be accepted by the end-users but also by the medical community and the caregivers

To illustrate the modularity and the various possible functionalities we have decided to develop a whole family of walking aids, the Assisted Navigation Guide (`ANG`). All the members of this family will be based on existing walker that are already accepted. Currently we have two members of this family, that will be presented in the next sections.

B. ANG-light

ANG-light is the simplest walker in the family with functionalities that are centered on monitoring and fall detection. It is based on based on a commercially available 3-wheels Rollator walker, the two rear wheels being fixed while the front is a caster wheel(figure 1). We have simply added to this walker incremental encoders in the rear wheels, a Phidgets 3D accelerometer/gyrometer and an SBC computing unit, together with a small battery. Using the encoders we are able to determine accurately the trajectory of the walker, while the gyrometer allows us to determine the rotation speed of the walker around its z axis (figure 1). Our initial assumption for developing this walker was that the trajectory of the walker on a typical path will provide useful information on the walking pattern of the end-user. We have initially validated this assumption by asking healthy subjects and one lightly handicapped subject to perform a 10m straight line trajectory with the walker. When comparing the trajectories we were able to determine that the handicapped subject has a problem with the left leg although this was completely invisible with a naked eye. Furthermore the doctors (that have not seen the test) to which we have showed the trajectory records were able to diagnosis the pathology of the handicapped subject. Furthermore the trajectory records enable us to calculate various objective numerical indices that characterize the walking pattern. Such indices may be precious to follow a rehabilitation process but we assume also that they will enable an early detection of emerging pathologies. To confirm this assumption we have performed a full test on 24 healthy subjects. Each subject was instrumented with a 3D accelerometer on the knees and on the wrists and were wearing special shoes for measuring the feet pressure. Each subject has to perform two typical trajectories two times, with and without the walker. Our purpose was to determine a golden standard for normal walking behavior and to determine the influence of the walker on the walking behavior. Figure 1 shows the record for a straight line trajectory with the walker together with the z gyrometer data (in which the step of the end-user can be easily detected). We are now in the process of performing the same test with 30 elderly people of the Nice Memory Center (NMC), of the Nice Hospital. Our purpose is to establish walking indices that will allow doctors to characterize walking patterns and to establish interval values for these indices that are typical of healthy subjects or pathological subjects.

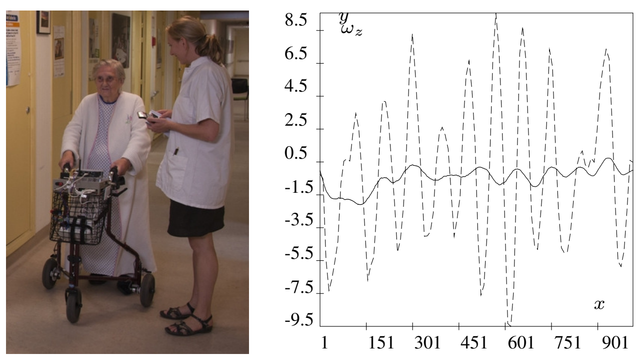


Fig. 1. The walking aid ANG-light and trajectory record and angular rate record (dotted) of the walker for a straight line trajectory (unit:cm, rd/s)

A second motivation for the development of ANG-light is fall detection. Forward or rearward fall may result in abnormal walker speed/acceleration that can be detected by the on-board sensors. When detecting such an event the SBC computer may send an alert through its wifi/bluetooth/optical network connection.

The intrusivity of the device is minimal as the on-board sensors and computer units may be fully integrated in the frame of the walker. The energy consumption is also minimal: the current battery allows for one week of typical use without recharge. We are now working to reduce the consumption by substituting the SBC board (which is much more powerful than needed) by an Arduino board and using energy harvester modules. In spite of the interesting functionalities provided by ANG-light the added cost is very low (a few dozens of euros).

C. ANG-II

ANG-II is the second walker we have developed and the most sophisticated one. It is based on a commercially available 4-wheels Rollator walker, the two rear wheels being fixed while the two front are caster wheels (figure 2) and hence is similar in principle to the NURSEBOT and iWalker. This Rollator offers the possibility of sitting on it.

The two rear wheels are equipped with 155W permanent magnet motors with a maximal torque of 15.5 NM leading to a maximal walking velocity of about 3.8 km/h. Electromagnetic friction clutches (HUCO SO26) allows to connect on demand the motors to the wheels with a controlled coupling friction. The total weight of ANG is about 20kg and the motors (9.2kg) and batteries (45Ah, 4.5kg) are located on a plate that is 4cm over the ground, leading to a center of gravity (CoG) of ANG-II that is about 30 cm below the CoG of the original walker, thus ensuring a better stability. Both the motors and clutches are controlled through a DC controller (Phidgets 1064). ANG-II has the same basic equipment than ANG-light (encoders in the rear wheels, 3D accelerometer/gyrometer) to which we have added:

- a force sensor (FSR01) on each handle
- 4 proximity on-off switches (at the front and at the rear)
- 4 infra-red range sensors (SHARP GP2Y3A003KOF, range: 40-300 cm at the front, GP2Y0A02 on left, right,rear, range: 20-150cm). Each of these sensors have 5 leds allowing to perform distance measurement in an arc of $\pm$ 25 degrees
- on-board GPS
- a set of on-off switches (called *push-buttons*) in the immediate proximity of the user.

Fig. 2. The ANG-II walker

- two webcams, one front-looking and one rear-looking
- 3D joystick, remote control through IR (Phidgets 1055), RC receiver

Interface to the motors and sensors are managed by two Phidgets 1019 interface kits (http://www.phidgets.com): these kits have 8 analog input ports, 8 input logical inputs and 8 output logical ports. A hub of 6 USB ports allows to connect USB sensors or other interface kits, while a main USB port allows one to connect the kit to the main computer.

A single computer manages the walker: currently we use a fit-pc2 (http://www.fit-pc.com, 101 × 115 × 27 mm), running a 1.6 GHz Intel Atom Z530 CPU. This is a low-consumption computer (maximal consumption: 8W), running Linux, offering infrared, bluetooth and wifi connection and several USB ports.

Programming is done in C with a configuration file indicating what resources are available. Then the Phidgets library is used to get the sensor information, pass it to our control algorithms which output the necessary motors/clutches instructions.

An on-board 15W solar panel provides additional energy if needed. Our test have shown that on a sunny day the panel provides about twice the energy that is used by the computer (between 4 and 8W), interface kits and sensors (about 3W). We have also performed tests with a Phidgets SBC board running a customized Linux on a 266 Mhz ARM processor, whose power consumption is 1.2W (without the wifi), without noticing any degradation in term of performances. If the motors are used at 50% power (which correspond to mounting a steep slope with a 60kg load) with the clutches on, the autonomy is about 1 hour. In summary ANG offers a sufficiently large energy autonomy.

The next section will describe the operating mode of ANG-II.

III. OPERATING MODES AND FUNCTIONALITIES

Several basic operating modes are available for ANG-II:

- *free mode*: the motors are unclutched and the system basically behave like the initial Rollator, although all sensors are running
- *motor mode* in which the motors are clutched and provide either braking power or motion help while the user is using the walker
- *teleoperated mode*: the walker moves under the direct remote control of the user
- *autonomous mode*: the walker moves autonomously while the user is away
- *reeducation mode*: this mode may be used to regulate the walking stance of the patient [12] (possibly using a sound system as proposed by Kanai [13]) or to oppose a selective resistance to the walking for training or navigation purposes

For normal walking we believe that the free-mode should be preferred as much as possible for:

- maintaining the cognitive/physical activities of the elderly
- active walkers usually present a larger task completion time than passive walkers [4]. It is indeed difficult to determine the intents of the elderly and control delay are very negatively resented
- the motor mode should be used only in case of emergency or after a direct request from the user

Beside these modes we have implemented several other functionalities, that will be explained in the next sections.

A. Navigation, street mapping and monitoring

Indoor and outdoor motion planning for walkers is a subject that has been often addressed in the literature [10], [8], [14] with the purpose of moving safely an elderly from a location to another one. But clinical evaluation shows that users are not very comfortable with this functionality [15]. ANG-II is able to provide a similar help: it can follow walls, avoid obstacle and plan an outdoor trajectory with its GPS and also "suggest" a motion direction by using differential friction in the clutches.

But we plan another use of the walker which emanates from a request of an handicapped people association which was proposing to map lowered kerbs on sideways so that if will be possible to provide an itinerary for people using wheelchairs or walkers. The initial proposition was to use a camera on a vehicle, computer vision being then used to detect the lowered kerbs. But this approach has the drawbacks to provide only static maps while the detection of the kerbs in a cluttered environment may not always be possible. We come up with another approach which is based on a collaborative map such as OpenStreetMap (http://www.openstreetmap.org). The idea is to annotate each city map with the data collected from the walkers of its inhabitants. Indeed lowered kerbs may be detected by on-board sensors of ANG-II: we may either use the webcam or more simply the accelerometer data. As seen in figure 3 an appropriate filtering of the accelerometer allows one to detect easily the kerb. Then using GPS data or the odometry of the walker we may locate the kerb on the map and annotate the location. Adding a time stamp for the annotation allows to build a dynamic map. Figure 3 shows a map of INRIA site whose OSM file has been automatically modified to take into account detected kerbs.

Furthermore we are considering building dynamically a ranking of the sideways using the same data. For example a sudden drop of the frequenting of a high-ranked sideways may indicate a temporary obstruction.

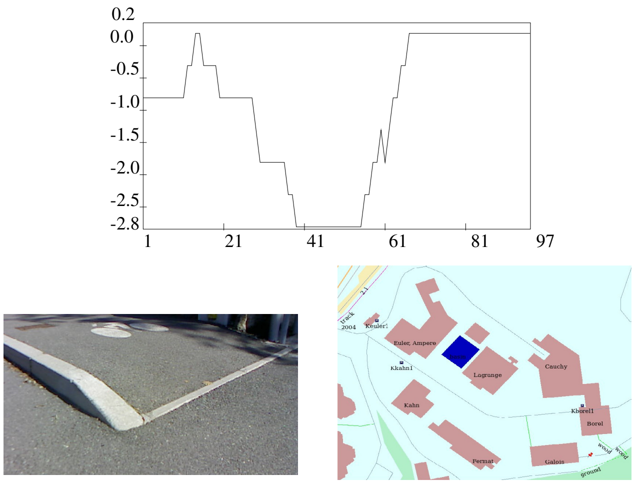


Fig. 3. Filtered record of the angle between the walker vertical and the gravity for the shown lowered kerbs as measured by the accelerometer: the kerb is easily detectable. On the right a map of the INRIA site which has been automatically annotated with lowered kerbs (denoted by K) that have been detected during a walk.

B. Fall detection/prevention

Fall prevention is obtained by monitoring the IR rear sensor and the velocity and acceleration sensors of ANG-II. The IR sensor is used to detect the presence of an end-user behind the walker. Then if an abnormal velocity or acceleration is detected the two motors will be clutched. In that configuration the walker provides a strong support that may help to prevent an axial fall. However a lateral fall cannot be prevented with this method. In that case if the IR sensor indicates that the end-user distance to the walker is larger than a fixed threshold or that he/she cannot be detected, then the two motors will be clutched and the walker will start a slow backward motion retracing its trajectory toward the last walker position $\mathbf{X_1}$ at which the end-user was detected by the IR sensor. If at some point the IR sensor detect the end-user, then the distance measurements along the different sensing directions allows one to roughly determine the direction of the end-user on the ground. The walker will then maneuver until its axis is perpendicular to the end-user axis and close to him/her. In that configuration the walker provides a strong support that can be used by the elderly to stand up. Figure 4 shows a typical recovery experiment. If the end-user is not detected by the IR sensor before the walker reaches the position $\mathbf{X_1}$, then the walker rotates alternatively on the left and on the right while keeping fixed the walker center location in order to increase the area covered by the IR distance measurements.

Note that the above strategy allows also the end-user to move away from the walker (for example to fetch an object) with the walker backtracking automatically towards him/her.

As an alternative to using the IR sensor we may also use the on-board rear looking webcam to detect the fallen end-user. Figure 5 shows one image of the elderly as seen from

Fig. 4. A typical fall experiment on a slope: fall occurs (top left), the walker start accelerating on the slope (top right), brakes and start its backward motion (bottom left), until it comes close to the user (bottom right).

the on board webcam, the contour that is detected and its approximation by a set of rectangles.

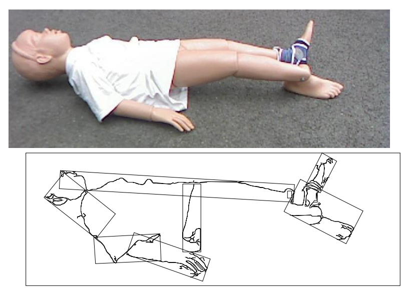


Fig. 5. The image of an elderly after a chute as seen by the on board webcam, the contours that are detected and their approximation by a set of rectangles.

As a second alternative we are considering equipping clothes of the end-user with a Lilypad Arduino (`http://www.arduino.cc`), connected to a verticality sensor and an accelerometer. The Arduino and the walker will be in constant communication through a bluetooth connection so that the Arduino is aware of the walker position. In case of a fall the Arduino will send an alert to the walker through its bluetooth connection, that includes the walker position at the moment of the fall. We have also imagined that the Arduino may light up IR diodes whose location may be determined by the walker sensors to facilitate the motion of the walker toward the fallen end-user.

An important aspect of the fall system is that ANG-II is able to use its wifi, bluetooth, radio and infrared ports to emit a fall warning, allowing other assistance devices to converge toward the end-user and provide additional help.

C. Shopping cart mode

Shopping is one of the major use of walkers, most of which have a shopping basket in front. However fetching goods while still maneuvering the walker is a difficult task for elderly people. Having the walker following automatically the end-user will be appreciated. Classical approaches for providing this functionality will be to use either the IR front sensor or the forward looking webcam. After some preliminary test it has appeared that the IR distance sensor was too sensitive to be used as a person tracker in a shop. As for the webcam, although robust people trackers are available, it appears that they have difficulty to manage full occultation of the end-user, an event that may frequently occur in a shop. Furthermore a non contact link does not allow to define a social bubble around the walker and the end-user. We have hence adopted a much simpler method: a passive 2 dof serial arm has been installed in front of the walker and a retractable dog leash is attached at the end-effector of this arm (figure 2). When the end-user pull the leash the elevation angle of the arm increases and the motor start running, moving the walker in the direction provided by the measurement of the rotation of the arm. If the walker come close to the end-user or he/she stops pulling the leash, then the elevation of the arm decreases and the motor will unclutch and stop turning. We plan to have a similar system at the rear so that the walker will also follow the end-user, ready to provide a support if necessary.

This simple system has proved to be very efficient: a social bubble is created with clear limits and such a system, which is already familiar to end-user, is easily accepted.

D. Walking aid and rehabilitation

Providing a walking aid is a major objective of `ANG-II`. As mentioned before we favor the free mode but the walker is able to provide motion help. For example we monitor the force sensors at the handles and the direction of gravity with the accelerometer to determine the slope of the walker ground. If both force sensors are used and the slope is 0 or positive the motor will be clutched and will rotate at a speed that is proportional to the exerted forces: this allows to have a temporary motion help for example when walking along a steep slope. If the slope is negative (i.e. the walker is going downward) we use the electromagnetic clutches to regulate the walker velocity. In another mode a simple push on the force sensor will make the walker move in a straight line at a constant velocity, that will be stopped as soon as the IR backward looking sensor does not detect the end-user. It may also be possible that at some point the end-user is too weak to walk. To deal with this problem we can transform `ANG-II` as a temporary wheel chair. For that purpose we have installed at the rear two plates with rolling balls to accomodate the feet of the end-user while sitting on the walker. Then he may control the walker by using the on-board joystick or the sensors in the handles (figure 6).

`ANG-II` may also be used as a walking rehabilitation device. We have already mentioned the monitoring mode in which the walker trajectory is stored for providing walking evaluation. We have also a *semi-active mode* in which the walker uses the motors in specific part of the gait in order to decrease the muscular effort of the patient. The intensity of this help may decrease in time for an optimal training.

Fig. 6. `ANG-II` used as a temporary wheelchair.

In the *active mode* the walker may provide a resistance (by adjusting the friction in the clutches) or regulate the gait (with the motors) to train the elderly. In the active modes we may use a sound system to help the end-user to regulate his stance [13]).

E. Sun mode

In order to improve the walker autonomy we have designed a simple *solar mode* to optimize the energy production of the solar panel. This mode may be used when the walker is outdoor but not in use. The on-board electronic compass is used to determine North and the computer calculates the sun azimuth (possibly using the GPS data to determine the latitude and longitude of the walker).The walker, after checking if there is no obstacle around it, simply rotates around the z axis so that the solar panel is oriented toward the sun (figure 7). In that location the walker goes into a sleep mode, switching off most equipments. The process is repeated every 15 minutes.

Fig. 7. The *sun mode*: the walker rotates until the solar panel is perpendicular to the sun and then goes into a sleep mode. After 15 minutes it rotates clockwise until the light measurement is maximum.

F. Remote control

In some occasions the walker may have to be remotely controlled either by the end-user, caregivers or in emergency situation by rescuers. For short distance `ANG-II` may be controlled through a wired 3D joystick, infrared tv remote set or by the incline of a smart phone through the bluetooth link. For middle distance we use a Graupner RC X412 set, 35 Mhz. The RC set may control the walker motion at a distance from up to 100 meter and is also able to control the pick-up reacher (see section III-G). We are currently investigating the use of Enocean 868 MHz RCM-TCM module that have the advantage of having a very low energy consumption.

For long distance we have developed several interfaces to control remotely the walker through the web relying on a wifi connection or using an on-board smart phone as a modem.

G. Domestic tasks

Although ANG-II is basically a mobility aid it can also be used for other tasks. For example it has been equipped with a powerful vacuum cleaner (figure 8) and an actuated pick-up reacher: it can be remotely operated to pick up fallen objects such as keys, that are then made available to the elderly at hand level (figure 9). As it is quite powerful ANG-II has

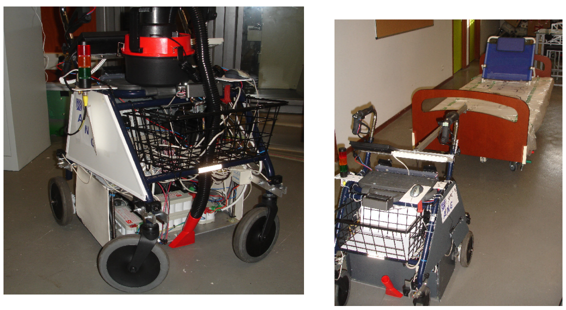


Fig. 8. ANG with its vacuum cleaner and towing a medical bed.

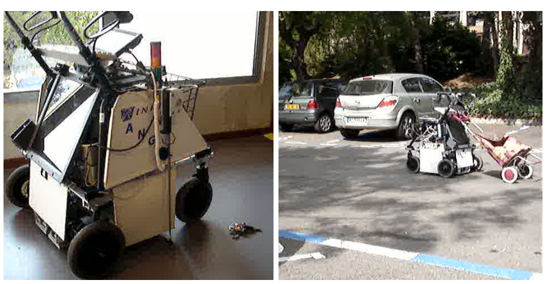


Fig. 9. ANG picking up fallen keys using the actuated pick-up reacher and towing a shopping cart.

also been used as a tug to move pieces of furniture and to tow a shopping cart (figure 8,9).

IV. CONCLUSIONS AND FUTURE WORKS

This project has presented two members of a family of devices which are a result of long term discussion with elderly people, caregivers and doctors working with elderly, that has allowed us to identify the needs and requirements of these communities. Maybe the most important aspects that have emerged from these discussions are the need for adaptability (in terms of platform and interface) and social acceptance concern (with a direct impact on the intrusivity and external appearance).An important added value of these devices is its monitoring function that allows doctor to objectively characterize walking patterns, not only during patient visit but also at home.

With this project we also want to show that by using modern IT tools it is possible to design a low cost assistant device with a large number of sensors (currently ANG-II has 21 sensors but may accommodate up to 30 analog sensors and 16 logical input/output sensors). This hardware has allowed us to design a platform with a large number of functionalities. Although being primarily designed as a walking aid (with built-in safety mechanisms to prevent/detect fall), including possible use as a rehabilitation device. But we have shown that the walker may perform efficiently many other tasks.

Only experimental work on a large panel of users will allow to validate the utility of the functionalities. We have already used 24 healthy subjects for testing the monitoring function, which may allow doctor to analyze walking patterns and possibly detect emerging pathologies.After a lengthy legal process tests are going on currently at Nice hospital with 30 elderly people.

Acknowledgment: this work has been partly funded by F.E.D.E.R. through the CPER TELIUS and by INRIA Large Scale Initiative PAL (Personally Assisted Living).

REFERENCES

[1] M. Alwan *et al.*, "Stability margin monitoring in steering-controlled intelligent walkers for the elderly," in *AAAI Fall Symposium*, Arlington, 4-6 November 2005, pp. 1509–1514.

[2] G. Lacey and S. MacNamara, "User involvement in the design and evaluation of a smart mobility aid," *J. of Rehabilitation Research & Development*, vol. 37, no. 6, pp. 709–723, 2000.

[3] R. Bostelman and J. Albus, "Robotic patient transfer and rehabilitation device for patient care facilities or the home," *Advanced Robotics*, vol. 22, no. 12, pp. 1287–1307, 2008.

[4] A. Rentschler *et al.*, "Intelligent walkers for the elderly: performance and safety testing of VA-PAMAID robotic walker," *J. of Rehabilitation Research & Development*, vol. 40, no. 5, pp. 423–432, September-October 2003.

[5] J. Glover *et al.*, "A robotically-augmented walker for older adults," CMU, Pittsburgh, Tech. Rep. CMU-CS-03-170, August 1 2003.

[6] P. Médéric, V. Pasqui, F. Plumet, P. Rumeau, and P. Bidaud, "Design of an active walking-aid for elderly people," in *3rd International Advanced Robotics Program : International Workshop on Service, Assistive and Personal Robots*, Madrid, 2003. [Online]. Available: http://www.isir.upmc.fr//files/2003ACTI76.pdf

[7] D. Chugo *et al.*, "A moving control of a robotic walker for standing, walking and seating assistance," in *Int. Conf. on Robotics and Biomimetics*, Bangkok, 21-26 February 2008, pp. 692–697.

[8] V. Kulyukin *et al.*, "iWalker: toward a rollator mounted wayfinding system for the elderly," in *IEEE Int. Conf. on RFID*, Las Vegas, 16-17 April 2008, pp. 303–311.

[9] Y. Hirata, S. Komatsuda, and K. Kosuge, "Fall prevention of passive intelligent walker based on human model," in *IEEE Int. Conf. on Intelligent Robots and Systems (IROS)*, Nice, 22-26 September 2008, pp. 1222–1228.

[10] B. Graf, "An adaptive guidance system for robotic walking aids," *J. of Computing and Information Technology*, vol. 17, no. 1, pp. 109–120, 2009.

[11] O. Chuy *et al.*, "Motion control algorithms for a new intelligent robotic walker in emulating ambulatory device function," in *IEEE Int. Conf. on Mechatronics and Automation*, Niagara Falls, July 2005, pp. 1509–1514.

[12] M. Takanokura, "Theoretical optimization of usage of four-wheeled walker and body posture of elderly users for comfortable steady walking," in *IEEE Engineering in Medicine and Biology Conference*, Vancouver, 20-24 August 2008, pp. 4519–4522.

[13] T. Kanai and T. Muto, "The dynamics of the gait training process with rhytmic cueing for para-hemiplegic patients," in *IEEE Int. Conf. on Intelligent Robots and Systems (IROS)*, San Diego, 22-26 September 2007, pp. 2275–2280.

[14] A. Morris *et al.*, "A robotic walker that provide guidance," in *IEEE Int. Conf. on Mechatronics and Automation*, Taipeh, 14-19 September 2003, pp. 25–30.

[15] A. Rentschler *et al.*, "Clinical evaluation of Guido robotic walker," *J. of Rehabilitation Research & Development*, vol. 45, no. 9, pp. 1281–1294, 2008.

Velocity Control for Walk Assistance by Endeffector Force in the Leg Coordinate based on the Biarticularly-actuated System

Shinta Sonokawa[1], Yunha Kim[1], Sehoon Oh[1] and Yoichi Hori[2]

Abstract— **This paper proposes a novel velocity control of the center of mass (COM) of a human body with attached ankle foot orthosis (AFO) during the stance phase. We propose a novel coordinate system for COM that achieves model simplificatio with the biarticularly-actuated system. This allows for simple control design of the velocity and position of COM. In addition to simplifie control, the proposed mathematical model for AFO has a simple structure that reproduces the biarticularly-actuated system using passive elements such as springs. Simulation results and comparison with conventional methods verify the effectiveness of the proposed control design.**

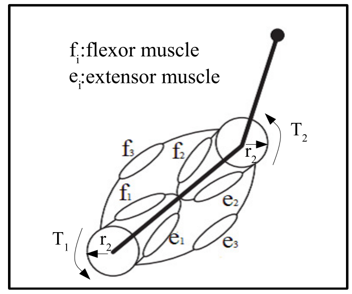


Fig. 1. 3 pairs 6 muscles model

I. INTRODUCTION

Hemiplegia is known as a sequela related to cerebrovascular diseases, such as apoplexy. People who are suffering from hemiplegia fin walking difficult Especially in case of elderly people hemiplegia imposes a danger of fatal injuries which include joint dislocation and fracture, for them being easy to fall to the paralyzed side during walking.

In this paper, we perform coordinate conversion using biarticular actuation, and propose a simple velocity control method based on the end-effector force control for a two-link manipulator. Based on this result, we are fabricating the new AFO which can ease the patients' burden by substituting the patients' gastrocnemius muscle with a passive element.

Various kinds of AFO have been developed for hemiparetic gait assistance. Blaya et al. developed AFO with variable impedance of an ankle, which prevents foot drop during stance phase, aggravation of symptoms, and accidents [1]. Yamamoto et al. developed a hydraulic AFO with variable rigidity of the ankle, which assists dorsifl xion and plantarfl xion during walking [2]. However these orthoses only help the patients who can walk by themselves. Thus, there are needs for AFOs which generate propulsion force for forward walking for more severely disabled people.

Some peculiar muscles called biarticular muscles which characterize animal limbs are receiving attention in recent years. Biarticular muscles are attached ranging over two joints, and drive the two joints simultaneously. Conventional robotics, however, seldom takes biarticular muscles into consideration despite the fact that they are the essential actuators in animal limbs. Kumamoto and Hogan verifie the effectiveness of biarticular muscles in end-effector force characteristics using the animal limb model [4],[5] using biarticular actuators.

Animals with limbs perform various movements such as walking, running, and jumping by harnessing relevant muscles. These animal movements can be described by a spring model called SLIP (Spring Loaded Inverted Pendulum) model [3]. The SLIP model implies that it is not one joint or one muscle, but mutually working multiple joints and muscles that achieves animal locomotion. This property inspired many research works on biarticular muscles. For example, Lewis et al. tried to reproduce the dynamic movement of animals driving two joints of the legs by using only one motor [6], Iida et al. used a spring between the joints of the leg [7]. And Klein et al. tried to reproduce the biarticular muscle movement during human walking using belt [8]. As seen in the aforementioned research works, the consistency in animal movements are explained by applying biarticular muscles.

In Section II, biarticularly actuated drive mechanism is shown. In Section III, we defin a supporting leg model during walking, and a two-link manipulator in consideration. Then we defin the leg coordinate system which is fi ed to Link 1. Based on this coordinate system the kinematics and statics are described. In Section IV, we propose a velocity control method of the center of mass using the end-effector force control in the leg coordinate system. In Section V, a simplifie method is examined for the limb model using passive biarticular elements. Finally Section VI concludes the work.

II. BIARTICULARY-ACTUATED MODEL AND MATHEMATICAL EXPRESSION OF JOINT TORQUES

Animal bodies are driven by various muscles. However if we restrict limb movements into planar ones, they can

[1]Department of Electrical Engineering, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656 Japan
sonokawa@hflab.k.u-tokyo.ac.jp
yunha.kim@hori.k.u-tokyo.ac.jp
sehoon@hori.k.u-tokyo.ac.jp
[2]Department of Advanced Energy, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba, 277-8561 Japan
hori@k.u-tokyo.ac.jp

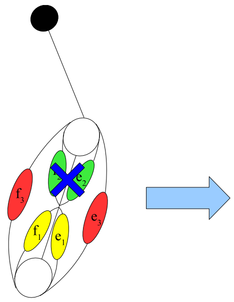


Fig. 2. Mono-Bi system without τ_2

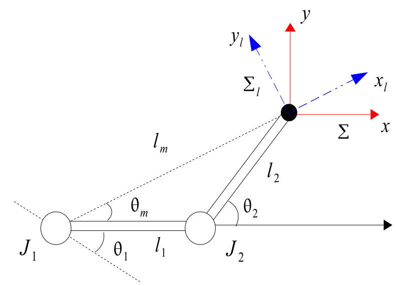


Fig. 3. Definitio of novel leg coordinate

be expressed with the 3 pairs 6 muscles models as shown in Fig.1 where $f_i, \{i = 1, 2, 3\}$ represents extensor muscles and $e_i, \{i = 1, 2, 3\}$ represents fl xor muscles. f_1, e_1 are monoarticular muscles of the firs joint, f_2, e_2 are monoarticular muscles of the second joint, and f_3, e_3 are biarticular muscles over two joints. These are driven as a pair, and if torques that each pair exerts are define as $\tau_i\{i = 1, 2, 3\}$, then the joint torques T_1, T_2 are denoted as the following formula (1) (2).

$$T_1 = \tau_1 + \tau_3 \tag{1}$$

$$T_2 = \tau_2 + \tau_3 \tag{2}$$

Since there are three inputs and two outputs as shown in the formulae above, biarticular-muscle-driven system has redundancy. However, Oh et al. clarifie that this redundancy contributes greatly to direction control of end-effector force [9]. It was shown that with biarticular muscles it is easy to realize straight forward movements since the straight line connecting the firs joint and end-effector ease the control.

III. Coordinate System and Kinematics Using Biarticular Muscles

In human walking behavior, the velocity of COM is known to be obtained by the propulsive force exerted by the muscles of the leg [4]. In this paper, we focus on generating propulsive force which pushes the COM using the supporting leg with an AFO during stance phase of a patient with hemiplegia. In definin the supporting leg model during stance phase, we propose the Mono-Bi system shown in Fig.2 which is equipped only with a monoarticular muscle pair τ_1 of the firs joint and a biarticular muscle pair τ_{12} extracted from 3 pairs 6 muscles model. That is, the torque of the monoarticular muscle pair of the second joint is zero, $\tau_2 = 0$. Moreover, with such actuator configuration we also propose the leg coordinate system which is suitable for describing movements at the end-effector of a manipulator.

A. Proposal of the Leg coordinate suitable for biarticular muscle inputs

J_1 and J_2 represent the firs and the second joint, and l_1, l_2 are the lengths of each link. Joint angles are denoted as θ_1 and θ_2. m_1, m_2, and I_1, I_2 are the mass and the inertia of each link member. Joint torque is denoted as T_1 and T_2. And finall the mass of the point P, the COM of the upper body is written as m_0.

B. Derivation of kinematics and statics in the proposal Leg coordinate

In the case of two link manipulator, it is common to use the absolute coordinate system. In this paper, however, we propose a coordinate system which is fi ed to the tip of the end-effector. When considering two muscles present in this work, it becomes easier to discuss the movements of end-effector P using this kind of coordinate system. We call this Σ_l coordinate leg coordinate system.

Link lengths of the manipulator shown in Fig. III-B are as mentioned above, the length of the straight line which connects firs joint J_1 and end-effector P is l_m, θ_1 is an angle between the absolute coordinate and the Σ coordinate system, where $\theta_{12} = \theta_1 + \theta_2$. $\Delta x_l, \Delta y_l$ are the displacements of each axis, and $\Delta\theta_1, \Delta\theta_{12}$ are the displacement of angles θ_1 and θ_{12}. For simplification each length is set in a way that $l_1 = l_2 = l$. The kinematics and inverse kinematics in the leg coordinate are written as follows (3) (4).

$$\begin{aligned}\begin{pmatrix}\Delta x_l\\ \Delta y_l\end{pmatrix} &= \frac{1}{l_m}\begin{pmatrix} l^2\sin\theta_2 & -l^2\sin\theta_2\\ l^2(1+\cos\theta_2) & l^2(1+\cos\theta_2)\end{pmatrix}\begin{pmatrix}\Delta\theta_1\\ \Delta\theta_{12}\end{pmatrix}\\ &= \frac{l^2}{l_m}\begin{pmatrix}\sin\theta_2(\Delta\theta_1-\Delta\theta_{12})\\ (1+\cos\theta_2)(\Delta\theta_1+\Delta\theta_{12})\end{pmatrix}\end{aligned} \tag{3}$$

$$\begin{aligned}\begin{pmatrix}\Delta\theta_1\\ \Delta\theta_{12}\end{pmatrix} &= \frac{1}{l_m\sin\theta_2}\begin{pmatrix}\cos\theta_2+1 & \sin\theta_2\\ -\cos\theta_2-1 & \sin\theta_2\end{pmatrix}\begin{pmatrix}\Delta x\\ \Delta y\end{pmatrix}\\ &= \frac{\cos\theta_2+1}{l_m\sin\theta_2}\begin{pmatrix}1\\ -1\end{pmatrix}\Delta x + \frac{1}{l_m}\begin{pmatrix}1\\ 1\end{pmatrix}\Delta y\end{aligned} \tag{4}$$

Torque exerted by the monoarticular muscle of the firs joint is τ_1, and one exerted by the biarticular muscle is denoted as τ_3. The force outputs of the P in the leg coordinate are f_l^x and f_l^y.

$$\begin{aligned}\begin{pmatrix}\tau_1\\ \tau_{12}\end{pmatrix} &= \frac{1}{l_m}\begin{pmatrix} l^2\sin\theta_2 & l^2(1+\cos\theta_2)\\ -l^2\sin\theta_2 & l^2(1+\cos\theta_2)\end{pmatrix}\begin{pmatrix}f_l^x\\ f_l^y\end{pmatrix}\\ &= \frac{l^2\sin\theta_2}{l_m}\begin{pmatrix}1\\ -1\end{pmatrix}f_l^x + \frac{l^2(\cos\theta_2+1)}{l_m}\begin{pmatrix}1\\ 1\end{pmatrix}f_l^y\end{aligned} \tag{5}$$

$$\begin{aligned}\begin{pmatrix}f_l^x\\ f_l^y\end{pmatrix} &= \frac{1}{l_m\sin\theta_2}\begin{pmatrix}\cos\theta_2+1 & -\cos\theta_2-1\\ \sin\theta_2 & \sin\theta_2\end{pmatrix}\begin{pmatrix}\tau_1\\ \tau_{12}\end{pmatrix}\\ &= \frac{1}{l_m\sin\theta_2}\begin{pmatrix}(1+\cos\theta_2)(\tau_1-\tau_{12})\\ \sin\theta_2(\tau_1+\tau_{12})\end{pmatrix}\end{aligned} \tag{6}$$

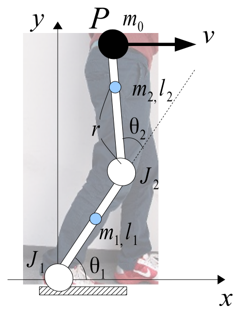


Fig. 4. Lower extremity modeled as 2 link manipulator

As mentioned above, f_l^x, f_l^y, the outputs in the leg coordinate system are obtained by the difference and the sum mode of each joint torque, respectively.

IV. Proposal of the velocity control method using outputs of biarticular muscles

The two link supporting leg model including the upper body used in this work is shown in Fig. 4. For simplification movements of the supporting leg model are limited to a two-dimensional plane. Inclination of the upper body rarely changes when non-disabled people walk, and even in case of patients with hemiplegia, it is possible to control so that the upper body does not fall according to muscles of the trunk. This kind of control, however, is not a part of the control for the supporting leg. Thus in this paper, we consider the mass of the upper body equivalent to a point mass at P. v_{lx} and v_{ly} are the linear velocities of the COM in terms of the leg coordinate system. J_1 is regarded as ankle, and we set each link length $l_1 = l_2 = l$ for simplification J_1 is fi ed at this time since we are considering supporting phase, and movements of the COM shows a circumferential movement with a radius of l_m.

A. Velocity control using biarticulary actuation in the leg coordinate

We propose velocity control of the COM, making a supporting leg follow reference value l_m^{ref} by using outputs seen at the point P in the leg coordinate.

Using the relation of kinematics and leg length $l_m = 2l\cos\theta_m$ the actual velocity in each direction, v_{lx} and v_{ly} in the leg coordinate can be expressed as follows.

$$v_{lx} = l\sin\theta_m(\dot{\theta}_1 - \dot{\theta}_{12}) \quad (7)$$

$$v_{ly} = l\cos\theta_m(\dot{\theta}_1 + \dot{\theta}_{12}) \quad (8)$$

If we put the gain for the leg length error as K_p, and the gain for the speed error in each direction asK_{dx} and K_{dy}, respectively, then the control inputs, f_l^x and f_l^y for each direction seen at point P can be written as follows.

$$f_l^x = K_p\left(l_m^{ref} - l_m\right) + K_{dx}\left(v_{lx}^{ref} - v_{lx}\right) \quad (9)$$

$$f_l^y = K_{dy}\left(v_{ly}^{ref} - v_{ly}\right) \quad (10)$$

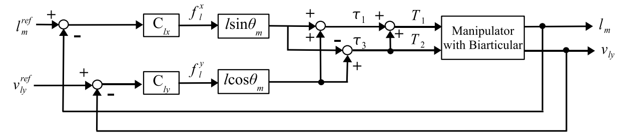


Fig. 5. Control design using leg coordinate

Note that we set $v_{lx}^{ref} = 0$ in this paper. In this way, the control input f_l^x becomes equivalent to a PD controller for the leg length l_m.

B. When biarticular muscles are not passive elements

From (9) (10), and considering $v_{lx}^{ref} = 0$, the torques of each pair of muscles τ_1, τ_{12} can be written as follows, by using the relationship (5) of the statics and leg length l_m in the leg coordinate system, then these become the input torques.

$$\tau_1 = K_p l\sin\theta_m\left(l_m^{ref} - l_m\right) + K_d l\cos\theta_m v_{ly}^{ref} - K_d l^2\left(\dot{\theta}_1 + \cos\theta_2\dot{\theta}_{12}\right) \quad (11)$$

$$\tau_{12} = -K_p l\sin\theta_m\left(l_m^{ref} - l_m\right) + K_d l\cos\theta_m v_{ly}^{ref} - K_d l^2\left(\cos\theta_2\dot{\theta}_1 + \dot{\theta}_{12}\right) \quad (12)$$

As the counterpart for comparison, the velocity reference of COM vref and the torque references (T_1, T_2) are given. As shown also in the block diagram shown in Fig. 5, because complex calculation of Jacobian matrix is unnecessary in the proposed method, a simple controller can be realized.

C. When biarticular muscles are passive elements

In addition to the input torques in (11) and (12), we assume there is a passive element which is a biarticular torque τ_{12} generating elastic torque. That is, (12) can be rewritten as follows.

$$\tau_{12} = K_{fix}\Delta\theta_{12} \quad (13)$$

Where, $\Delta\theta_{12}$ is the angle displacement from the equilibrium position θ_{12}^0, and it is define as follows.

$$\Delta\theta_{12} = \theta_{12}^0 - \theta_{12} \quad (14)$$

Moreover, gain K_d for the speed error becomes (15) below, by introducing the spring coefficien K_{fix} and considering (12) and (13).

$$K_d = \frac{K_{fix}\Delta\theta_{12} + K_p l\sin\theta_m\left(l_m^{ref} - l_m\right)}{l\cos\theta_m v_{ly}^{ref} - l^2\left(\cos\theta_m\dot{\theta}_1 - \dot{\theta}_{12}\right)} \quad (15)$$

From these observations, when Biarticular torque τ_{12} is included as a passive mechanism to movements, it becomes the input torques in (11) and (13).

TABLE I

PARAMETERS OF SIMULATION MODEL

g = 9.8[m/s^2]	m_0 = 50[kg]	m_1= 1.8[kg]
m_2= 3.6[kg]	l_1 = 0.4[m]	l_2 = 0.4[m]
I_1 = 0.217[kgm^2]	I_2 = 0.434[kgm^2]	l_m^{ref} = 0.78[m]

D. Forward walking simulation by the proposal method

Since we consider an AFO which is attached under the knee, it is assumed for the simulations that the firs joint J_1 of two link manipulator described in Section Ⅲ corresponds to the ankle of the supporting leg model. The parameters used in the simulations are indicated in table I. The initial posture is given at (θ_1, θ_2) = (1.8 [rad], 0.4 [rad]). The simulation is done from treading in until kicking out in the supporting leg model. The velocity reference of each axial direction is given in the forward direction at (v_l^{ref}, v_l^{ref}) = (0 [m/s], 1.0 [m/s]).

1) When biarticular muscles are not passive elements: In case input torques don't contain passive elements, the governing equations are written as (11) and (12). K_p, K_d can be considered as the P and D gain for the position references in the direction of x_l, respectively. These two gains are computed by pole assignment. In this case, the plant can be written as (16) below.

$$P(s) = \frac{1}{Ms^2} \tag{16}$$

Here, M is equal to the COM point mass m_0. Using the plant and the PD controller above, pole assignment is done. Simulation is done with a given pole at $\omega = 50$[rad/s].

2) When biarticular muscles are passive elements: In case input torques contain passive elements, the governing equations are written as (11) and (13). We use the same pole at $\omega = 50$[rad/s], and the spring coefficien at $K_{fix} = 550$[Nm/rad]. The variable gain K_d is determined by substituting gains K_{fix}, K_p into (15).

E. Simulation result

The simulation results of both cases are shown. These results are fairly consistent to that of human walking, even when the biarticular muscle τ_{12} is a passive element.

When not included a passive element, values follows well in reference values. On the other hand, we can see the characteristics of a passive element when included.

V. APPLICATION OF THE SUPPORTING LEG CHANGE ALGORITHM

In the previous section, the velocity control only for one step of the supporting leg model was considered. However, since changing legs is essential in order to extend the algorithm to walking, leg-change should be discussed. Thus here, the inverted pendulum model is introduced [10].

A. The changing algorithm of the leg in walk operation

Fig. 12 is a walking model including upper body and the leg using the inverted pendulum model. For simplification the model is limited to the sagittal plane and the mass of legs are ignored. m, J, g, and l are the mass, the inertia of the upper body, the gravitational acceleration, and the length from the waist to the center of mass of the upper body, respectively. r is the length from a tip of a foot to the waist, and ϕ_1, ϕ_2 are the inclination angles of a leg and the upper body from the vertical. r, ϕ_1, and ϕ_2 are state variables changing with time. τ is the torque acting on the waist, and F is the generalized force corresponding to the length of the leg r.

The state equation of the inverted pendulum model is decoupled and linearized, and the poles where the system is asymptotically stable are assigned. The body can be stabilized by the pole assignment, and when ϕ_2 approaches 0, ϕ_1 has an unstable pole.

If we put the required time to take one step as T, the state variables $(\phi_1, \dot{\phi}_1)$, the inclining angle of the leg can be written as follows.

$$\begin{pmatrix} \dot{\phi}_1 \\ \ddot{\phi}_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{g}{r+l} & 0 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \dot{\phi}_1 \end{pmatrix} \tag{17}$$

Equation (17) can be discretized by walking cycle T,

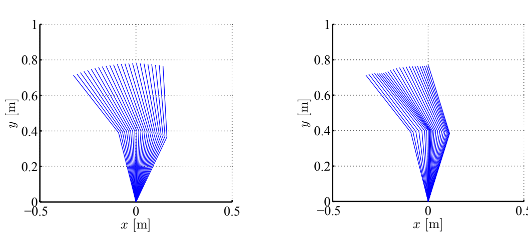
$$\phi_D[k+1] = \begin{pmatrix} \cosh bT & \frac{1}{b}\sinh bT \\ b\sinh bT & \cosh bT \end{pmatrix} \phi_D[k] - \begin{pmatrix} \cosh bT \\ b\sinh bT \end{pmatrix} u_D[k] \tag{18}$$

where, $b^2 = \frac{g}{r+l}$. This discrete system can be stabilized by determining h_1 and h_2 by assigning poles when the input is u_D from (19). Input u_D is the angle between both the legs for one step required at the instant of landing. In practice, u_D is define as (19) below.

$$u_D = h_1(\phi_{end}[k] - \frac{\phi_r}{2}) + h_2(\dot{\phi}_{end} - \frac{v - v^{ref}}{r}) + \phi_r \tag{19}$$

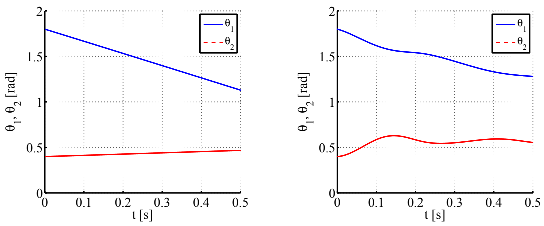
h_1 and h_2 are coefficient which assign the arbitrary poles in a discrete system, and can be written as below.

$$\begin{cases} h_1 = 1 - \lambda_1\lambda_2 \\ h_2 = \frac{(1+\lambda_1\lambda_2)\cosh bT - \lambda_1 - \lambda_2}{b\sinh bT} \end{cases} \tag{20}$$



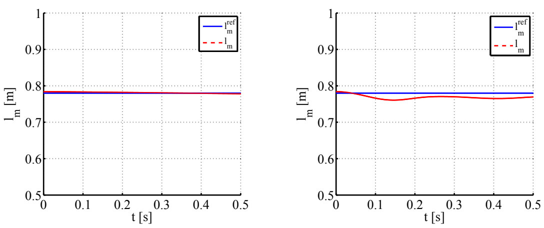
(a) W/O passive element (b) W passive element

Fig. 6. Stick diagram of supporting leg



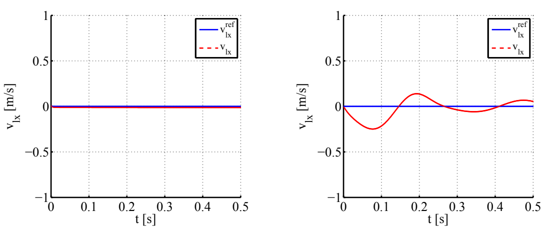
(a) W/O passive element (b) W passive element

Fig. 7. Joint angles θ_1, θ_{12}



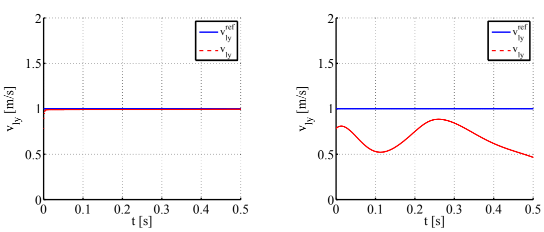
(a) W/O passive element (b) W passive element

Fig. 8. Change of leg length l_m



(a) W/O passive element (b) W passive element

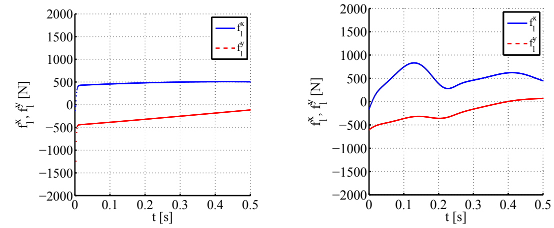
Fig. 9. Velocity of center of mass v_{lx}



(a) W/O passive element (b) W passive element

Fig. 10. Velocity of center of mass v_{ly}

ϕ_r and v^{ref} are reference step width and the velocity reference, respectively. Fig. 13 shows the leg-change model of the supporting leg.



(a) W/O passive element (b) W passive element

Fig. 11. Input forces f_l^x, f_l^y in the leg coordinate

B. State variables in changing legs

If we set ϕ_1 and r of the inverted pendulum model, as $\phi_{1,end}$ and r_{end} right before the leg-change, and as $\phi_{1,st} and r_{st}$ right after the leg-change, the parameters shown in Fig. 13 can be estimated by the equations below.

$$\phi_{1,st} = \phi_{1,end} - u_D \tag{21}$$

$$r_{st} = \frac{r_{end}\cos\phi_{1,end}}{\cos\phi_{1,st}} \tag{22}$$

$$\dot{\phi}_{1,st} = \frac{\dot{r}_{end}}{r_{st}}\sin u_D + \frac{r_{end}}{r_{st}}\dot{\phi}_{1,end}\cos u_D \tag{23}$$

$$\dot{r}_{st} = \dot{r}_{end}\cos u_D - r_{end}\dot{\phi}_{1,st}\sin u_D \tag{24}$$

And these equations can be rewritten, by using the state variables $(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2)$ of the Mono-Bi system, as follows. Note that $\theta_m = \frac{\theta_2}{2}$.

$$\phi_{1,st} = \frac{\pi}{2} - \theta_1 - \theta_m \tag{25}$$

$$r_{st} = l_m \tag{26}$$

$$\dot{\phi}_{1,st} = -\dot{\theta}_1 - \dot{\theta}_m \tag{27}$$

$$\dot{r}_{st} = v_{lx} \tag{28}$$

By using the state variables $(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2)$ obtained from the equations above, extension to walking of a supporting leg model can be realized.

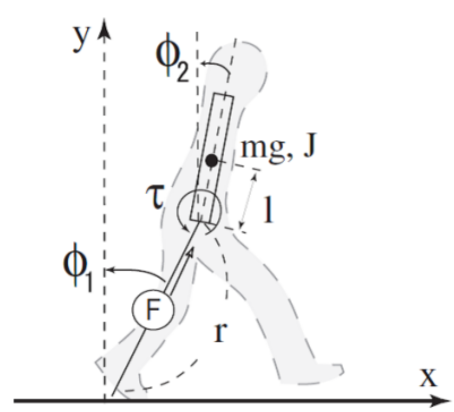


Fig. 12. The figur of the walk model

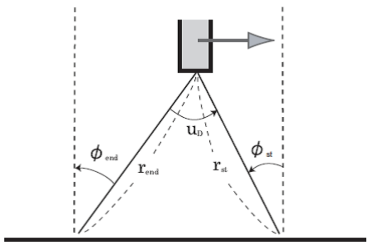


Fig. 13. The walk model at the time of change a supporting leg

C. Simulation including changing legs of walking operation

Simulation is done by placing the pole at $\omega = 50$[rad/s] for the plant in (16), and the spring coefficien of the biarticular torque τ_{12} is $K_{fix} = 550$[Nm/rad].

VI. CONCLUSION

We proposed the mono-bi system eliminating the monoarticular muscle pair of the second joint compared to the conventional 3 pairs 6 muscles model. Then we simplifie the kinematics of the system using a novel leg coordinates. Velocity control of COM by using this Mono-Bi system as the supporting leg model was applied to show very good tracking characteristics. Also even when the biarticular muscle τ_{12} was of a passive element, the result was fairly consistent to that of human walking. Moreover, simulations on leg-change during walking using the inverted pendulum model were done to show the feasibility of the new AFO which can reduce patients' burden.

As the next step, a relation between the reference generation and parameters of walking needs to be clarified Future work will include findin solutions to such problems, and fabrication of the prototype of the novel AFO.

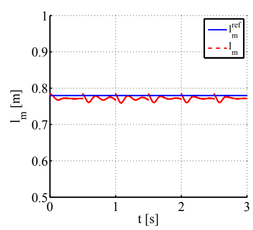
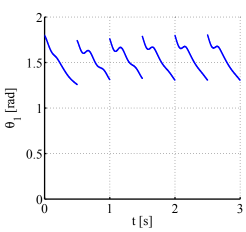
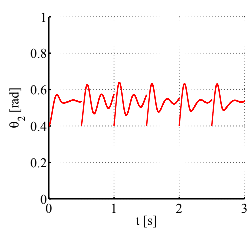


Fig. 14. Change of leg length l_m

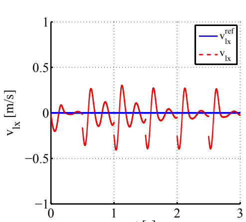


(a) First joint θ_1

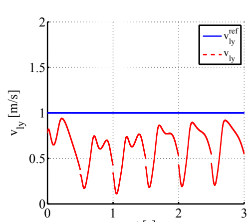


(b) Second joint θ_2

Fig. 15. Joint angles θ_1, θ_{12}

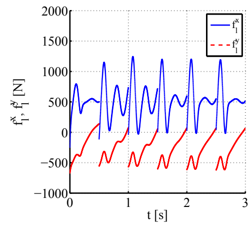


(a) Velocity v_{lx}

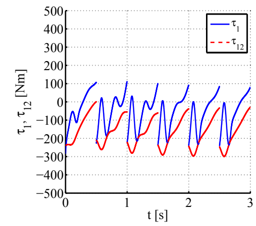


(b) Velocity v_{ly}

Fig. 16. Velocity of center of mass v_{lx}, v_{ly}



(a) Input forces f_l^x, f_l^y



(b) Input torques τ_1, τ_{12}

Fig. 17. Input forces and torques in the leg coordinate

REFERENCES

[1] Joaquin A. Blaya, Hugh Herr: "Adaptive Control of a Variable-Impedance Ankle-Foot Orthosis to Assist Drop-Foot Gait", *IEEE Transactions on Neural System and Rehabilitation Engineering*, vol. 12, no. 1, pp. 24-31, March 2004

[2] S. Yamamoto: "Development of Ankle Foot Orthosis for Hemiplegic Patients Based on Gait Analysis", *Rigakuryoho Kagaku*, vol. 18, no. 3, pp. 115-121, 2003(in Japanese)

[3] R. Blickhan: "The spring-mass model for running and hopping", *Journal of Biomechanics*, vol. 22, no. 11-12, pp. 1217-1227, 1989

[4] T. Oshima, T. Fujikawa and M. Kumamoto: "Functional Evaluation of Effective Muscle Strength Based on a muscle Coordinate System Consisted of Bi-articular and Mono-articular Muscles - Contractile Forces and Output Forces of Human Limbs - (in Japanese)", *Journal of the Japan Society for Precision Engineering*, vol. 65, No. 12, pp. 1772-1777, 1999.

[5] N. Hogan: "Adaptive control of mechanical impedance by coactivation of antagonist muscles", *IEEE Transactions on Automatic Control*, vol. 29, no. 8, pp. 681-690, Aug. 1984

[6] M. A. Lewis, M. R. Bunting, B. Salemi, H. Hoffmann: "Toward Ultra High Speed Locomotions: Design and Test of a Cheetah Robot Hind Limb", *2011 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, May, 2011, pp. 1990-1996

[7] F. Iida, J. Rummel, A. Seyfarth: "Bipedal walking and running with spring-like biarticular muscles", *Journal of Biomechanics*, vol. 41, no. 3, pp. 656-667, 2008

[8] T. J. Klein, M. A. Lewis: "A robot leg based on mammalian muscle architecture", in *2009 IEEE International Conference on Robotics and Biomechanics (ROBIO)*, IEEE, pp. 2521-2526, December, 2009

[9] Sehoon Oh, Yasuto Kimura, Yoichi Hori: "Reaction Force Control of Robot Manipulator Based on Biarticular Muscle Viscoelasticity Control", *IEEE/ASME International Conference*, July 2010

[10] Ferdinand Gubina, Hooshang Hemami, Robert B. McGhee: "On the Dynamic Stability of Biped Locomotion", IEEE Transactions on Biomedical Engineering, pp. 102-108, vol. BME21, no. 2, March (1974)