

Human-centered manipulation and navigation with Robot DE NIRO

Fabian Falck, Sagar Doshi, Nico Smuts, John Lingi, Kim Rants, Petar Kormushev

Abstract—Social assistance robots in health and elderly care have the potential to support and ease human lives. Given the macrosocial trends of aging, long-lived populations, robotics-based care research mainly focused on helping the elderly live independently. In this paper, we introduce Robot DE NIRO, a research platform that aims to support the supporter (the caregiver) and also offers direct human-robot interaction for the care recipient. Augmented by several sensors, DE NIRO is capable of complex manipulation tasks. It reliably interacts with humans and can autonomously and swiftly navigate through dynamically changing environments. We describe preliminary experiments in a demonstrative scenario and discuss DE NIRO’s design and capabilities. We put particular emphases on safe, human-centered interaction procedures implemented in both hardware and software, including collision avoidance in manipulation and navigation as well as an intuitive perception stack through speech and face recognition.

I. INTRODUCTION

Social assistance robots for elderly care or general nursing have been subject to extensive research in recent years. They may serve to counterbalance the global nursing shortage caused by both demand factors, such as demographic trends [1], and supply factors, such as unfavorable working environments or egregious wage disparities [2] [3]. Most systems are focused on directly assisting the care recipient – often an independently living elderly person – with social companionship or simple household services [4] [5]. However, elderly care today is still predominantly administered by human caregivers, who may themselves benefit from a robot assistant. Instead of seeking to replace caregivers, we propose Robot DE NIRO (Design Engineering’s Natural Interaction RObot)¹ as a tool for caregivers. DE NIRO is a collaborative research platform that can aid geriatric nurses by performing well-defined, repeated auxiliary tasks, such as retrieving a bottle of medicine and taking it to the care recipient. In designing DE NIRO, we have put an emphasis on natural and safe human-robot interaction procedures across multiple components, including speech and face recognition and collision avoidance.

This paper explains more of the context and functionalities of DE NIRO. In section II, we briefly discuss related work on care and social assistance robots. Section III gives an overview of DE NIRO’s initial hardware design and the

The authors are affiliated with Imperial College London, Robot Intelligence Lab, Department of Computing and Dyson School of Design Engineering. fabian.falck17, sagar.doshi17, nicoalaas.smuts17, john.lingi09, kim.rants17, p.kormushev@imperial.ac.uk

¹ Further information on DE NIRO can be found under http://www.imperial.ac.uk/robot-intelligence/robots/robot_de_niro/

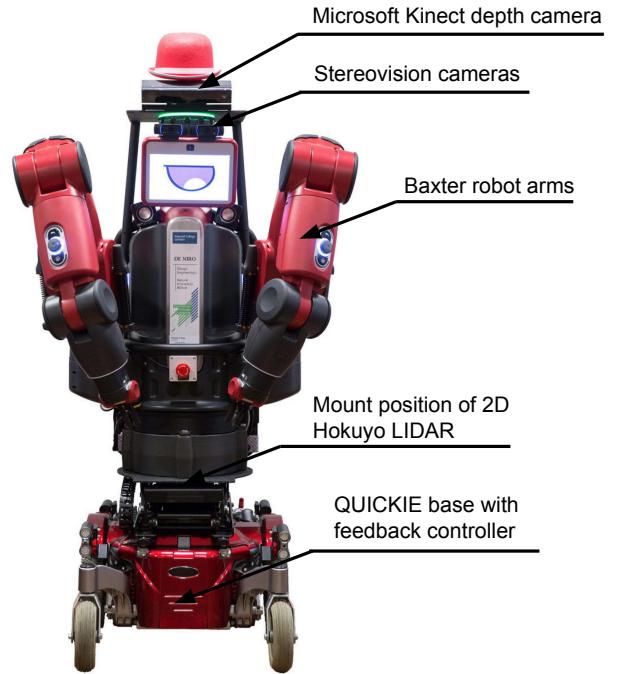


Fig. 1. Robot DE NIRO - a collaborative research platform for mobile manipulation. The figure shows its design, main components and sensors.

applied software frameworks. Section IV relates our preliminary experiments with DE NIRO. We give an overview of its current perception capabilities and possible actions, with focus on manipulation and LIDAR-based navigation. Finally, section V concludes the paper and gives an outlook for future work on this research platform.

II. RELATED WORK

In current literature, robot systems for elderly care are typically built as autonomous systems that directly support personal independence through social companionship, routine household services, and telepresence [5]. Examples of state-of-the-art platforms in service robotics – some of them combining all three categories – are Care-O-Bot 3 [6], ASIMO [7], HRP-3 [8], various solutions for ambient assisted living such as the DOMEQ RobuMate [9] and a social assistant robot for people with mild cognitive impairments [4] [10]. Willow Garage built the more general platform PR2, which has been used by various universities to build human-centered applications that include supporting the elderly [11]. Advances for better social interaction and improved navigation have been made, such as gesture recognition for

service ordering [12] or efficient navigation in unstructured household environments in an emergency situation [13]. Target care recipients are those who suffer from psychological diseases or, more commonly, those with limited mobility, including people who are elderly, disabled, temporarily or chronically sick, pregnant, or otherwise constrained [14].

With these recent advances in social assistant robots, public and academic debate has not yet settled on whether and for which purposes such systems are an ethical and desirable outcome for our society [15] [16]. For instance, Sharkey and Sharkey point out six main ethical concerns with robot assistance for elderly care, including human isolation, loss of control and personal liberty, and deception and infantilization [17]. Furthermore, [18] note varying attitudes and preferences regarding social assistance robots. While this discussion is not settled, we believe it is appropriate in the interim to focus robot assistance on the caregiver. This allows the caregiver to delegate simple, repetitive tasks to a social robot assistant and gain time for more complex, empathetic tasks.

Much less work has sought to support caregivers as they shoulder typical nursing tasks to care for elderly persons. Among those, [19] propose a transfer assistant robot to lift a patient from a bed to a wheelchair through a model-based holding posture estimation and a model-free generation of the lifting motion. [20] demonstrate complex manipulation skills for bottles and cans based on sparse 3D models for object recognition and pose estimation. They integrate these skills with navigation and mapping capabilities, both in a static and a dynamic environment.

III. DESIGN

DE NIRO's core design idea is to combine the industrial Baxter dual robot arms with autonomous navigation into a mobile manipulation research platform. The Baxter arms are a common standard in human-robot interaction and allow complex manipulation of objects [21]. A particular safety feature of the arms is their passive compliance through series elastic actuators. This allows the robot to interact with humans in close proximity to the robot safely, since in the case of a contact, most of the physical impact is absorbed. The Baxter arms are mounted on a QUICKIE movable electric wheelchair base. Its differential drive is operated with a custom PID angular position and velocity controller, allowing primitive motion commands for navigation. The controller itself is implemented through an integrated Mbed microcontroller [22]. On the hardware side, multiple layers of safety are operating for the event of an emergency. An automated interrupt procedure stops the movement of QUICKIE if a time out appears. Furthermore, both on-board and wireless e-stop buttons allow the user to brake the robot immediately.

These core elements are augmented with the following sensors: a Microsoft Kinect RGB-D camera, various stereovision cameras with built-in microphones, ultrasonic and infrared proximity sensors, speakers for audio output, and a Hokuyo 2D LIDAR scanner. Equipped with this extensive list

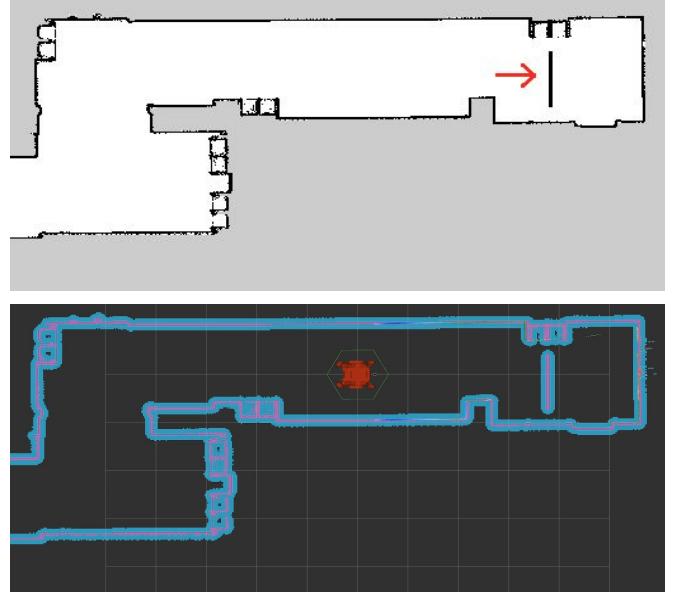


Fig. 2. A static map of a corridor (top). The red arrow points at a synthetic barrier manually added to the map. A costmap of 10cm surrounds all static barriers. The hexagonal shape is a synthetic barrier around the QUICKIE base used for collision avoidance (bottom).

of sensors and actuators, DE NIRO is capable of performing a wide variety of the typical, repetitive tasks of a caregiver, such as serving drinks and food, grasping, fetching and carrying of objects, and helping others come to a standing position [14]. Figure 1 visualizes the hardware design of DE NIRO as a whole.

To handle concurrent execution and both synchronous and asynchronous communication between components, we use Robot Operating System (ROS) as middleware. We define distinct functionalities of the robot with a finite-state machine, such as `listening` (for command input) or `grasping` (to physically pick up an object). The state machine handles the control flow among these states. Furthermore, we set up a wireless LAN network for concurrent communication between the Baxter core and the two controlling laptops mounted on the back of DE NIRO. For testing and debugging purposes and in order to integrate all sensor outputs and log messages, we built an rqt-based GUI illustrated in figure 4 and useful mainly to the technical user [23].

IV. IMPLEMENTATION AND EXPERIMENTS

Our primary work focused on the development and integration of state-of-the-art algorithms to perform a particular demonstration scenario. This scenario involved interacting with a user to receive a command indicating which object to grasp; navigating to and from an object warehouse; and grasping, manipulating, and passing back the requested object to the requester.

Perception and user interaction. One challenge was to recognize the user's face and distinguish that individual from others. To solve this, we used a pre-trained machine

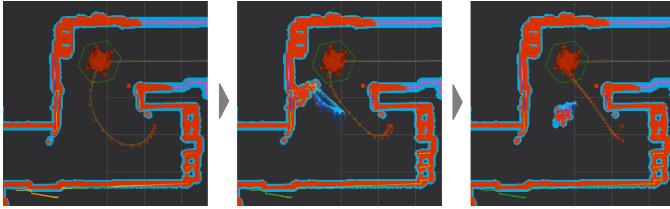


Fig. 3. The original optimal path (left), encountering a dynamic obstacle (middle), and adjusting in response (right) during trajectory planning.

learning model based on the Residual Learning for Image Recognition (ResNet) approach that we applied to video frames retrieved by the Kinect camera [24] [25]. The model has reached a 99.38% accuracy on a standard benchmark [26]. It compares the output vector encodings of known faces (via saved images) with others extracted from the processed frames by computing a distance metric between the saved and incoming vector encodings. If that distance is below a threshold, it predicts a positive match. We tuned the model to predict with a very low false-positive rate at the cost of a slightly increased false-negative rate, in order to be less vulnerable to unintended interactions.

To naturally interact with a user, we implemented a speech recognition system using the offline library *CMU Sphinx* [27]. We defined a JSpeech Grammar to allow voice commands in a specific, yet flexible format tested on a variety of accents. Furthermore, the system regularly calibrates to background noise levels. Compared to several online APIs, this implementation achieved the most robust results in varying environments. For audio output, we elected to use *eSpeak*, a simple speech package, over more sophisticated candidates, due to its high reliability, rapid response time, low required processing power, and – especially – offline implementation [28].

Navigation, Mapping, and Planning. We implemented a reliable navigation stack consisting of mapping, localization, and trajectory planning. To generate an initial static bitmap, we apply a SLAM-based approach to the LIDAR sensor in order to detect any spatial boundaries and 2D artifacts in a predefined space [29]. Then, we localize the robot by overlaying a dynamic map onto the static map. This dynamic mapping feature is also useful for collision avoidance, which is particularly important when DE NIRO is in the vicinity

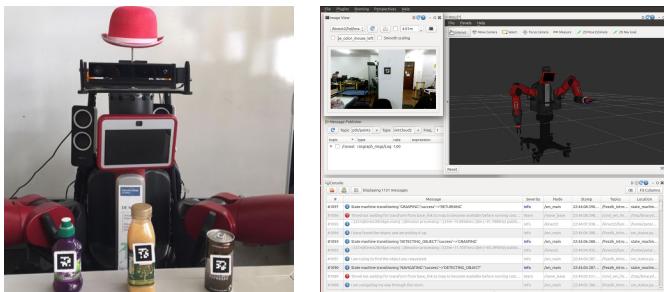


Fig. 4. The 2D fiducial markers attached to test objects (left) and the qt-based technical GUI for testing purposes (right).



Fig. 5. The three grasping stages of moving to an intermediate position (left), grasping the object (middle), and executing handover-mode (right).

of untrained humans. For additional safety, we impose a costmap to create a virtual cushion around all static and dynamic obstacles and around the QUICKIE base itself. The static map is illustrated in figure 2.

For efficient trajectory planning, we use a “timed elastic band” approach, conceiving of trajectory planning as a multi-objective optimization problem [30] [31]. This approach minimizes the costs that are assigned to variables like total travel time and obstacle proximity simultaneously. This helps DE NIRO to maintain a safe distance from users. The planned linear and angular velocities of the optimal path are then scaled and smoothed by the custom PID controller discussed above. Finally, an electric signal to the motor produces actuating rotational movement [22]. Figure 3 illustrates a path replanning scenario for when DE NIRO detects a dynamic obstacle.

Object Recognition and Manipulation. To recognize the target object and localize it in 3D space, we relied on 2D fiducial markers attached to the object [32]. We also experimented with various other more generic and scalable solutions, including recognizing objects on a planar surface. This solution and other attempts we made did not work robustly enough during grasping. The 2D fiducial markers depicted in figure 4 were much more consistent.

To control the Baxter arms, we employed an inverse kinematics solver to compute each of the seven joint angle trajectories needed to reach an object [33]. We designed a dynamic awareness procedure so that DE NIRO selects the most appropriate arm to make a grasp attempt; reacts to changes in the object’s location during grasping; and actively avoids collisions, say, with the unused arm. We experimented with various constraints on possible joint angles and settled on a grasping procedure with one intermediate point that achieves its goal most frequently. After grasping the object, the user can retrieve it easily during handover mode by imposing a small force along the z-axis. The experimental grasping process is illustrated in figure 5.

V. CONCLUSION

In this research ², we presented the design and implementation of Robot DE NIRO to support geriatric nurses in

² We open-sourced our object-oriented *code* base in Python together with an extensive *documentation* for it that will be continuously updated. Furthermore, we have published a *video* illustrating some of the current core skills of DE NIRO. Publications of *sensor data* from DE NIRO will follow soon. All resources are linked at <http://www.imperial.ac.uk/robot-intelligence/software/>.

interaction tasks with care recipients. DE NIRO's (current) design is limited in various ways: First, the robot design is nonholonomic, being limited to only forward and backward translational and rotational (but no side-ways) movement. Second, with a maximum payload of 2.2 kg per arm, DE NIRO is limited to relatively light weight tasks, e.g. incapable of lifting a human body. Third, due to limited sensor capabilities in the current design, we constrain DE NIRO to trajectories using forward motion which can result in the robot getting stuck in corners.

DE NIRO can, however, go further. Future work may explore increased awareness, such as through safety improvements with a 360-degree camera rig, the application of a 3D LIDAR (already operational), more robust localization that does not require predefined mapping, human pose estimation, visuospatial skill learning by demonstration [34], a more persistent autonomy during navigation without deadlock situations [35], and further improvements to point cloud based object detection [36] [37]. The work we have accomplished here, nevertheless, shows that DE NIRO's current capabilities can be used to provide reliable, efficient support to tasks requiring frequent, natural interaction with humans.

REFERENCES

- [1] A. Tapus, M. J. Mataric, and B. Scassellati, "Socially assistive robotics [grand challenges of robotics]," *IEEE Robotics & Automation Magazine*, vol. 14, no. 1, pp. 35–42, 2007.
- [2] N. Super, "Who will be there to care? the growing gap between caregiver supply and demand," 2002.
- [3] J. A. Oulton, "The global nursing shortage: an overview of issues and actions," *Policy, Politics, & Nursing Practice*, vol. 7, no. 3_suppl, pp. 34S–39S, 2006.
- [4] C. Schroeter, S. Mueller, M. Volkhardt, E. Einhorn, C. Huijnen, H. van den Heuvel, A. van Berlo, A. Bley, and H.-M. Gross, "Realization and user evaluation of a companion robot for people with mild cognitive impairments," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1153–1159.
- [5] D. Fischinger, P. Einramhof, K. Papoutsakis, W. Wohlkinger, P. Mayer, P. Panek, S. Hofmann, T. Koertner, A. Weiss, A. Argyros, *et al.*, "Hobbit, a care robot supporting independent living at home: First prototype and lessons learned," *Robotics and Autonomous Systems*, vol. 75, pp. 60–78, 2016.
- [6] B. Graf, C. Parlitz, and M. Hägele, "Robotic home assistant care-o-bot® 3 - product vision and innovation platform," in *International Conference on Human-Computer Interaction*. Springer, 2009, pp. 312–320.
- [7] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, and K. Fujimura, "The intelligent ASIMO: System overview and integration," in *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, vol. 3. IEEE, 2002, pp. 2478–2483.
- [8] K. Kaneko, K. Harada, F. Kaneko, G. Miyamori, and K. Akachi, "Humanoid robot HRP-3," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. IEEE, 2008, pp. 2471–2478.
- [9] "Domeo RobuMate," http://www.aat.tuwien.ac.at/domeo/index_en.html.
- [10] P. Rashidi and A. Mihailidis, "A survey on ambient-assisted living tools for older adults," *IEEE journal of biomedical and health informatics*, vol. 17, no. 3, pp. 579–590, 2013.
- [11] S. Cousins, "Ros on the pr2 [ros topics]," *IEEE Robotics & Automation Magazine*, vol. 17, no. 3, pp. 23–25, 2010.
- [12] X. Zhao, A. M. Naguib, and S. Lee, "Kinect based calling gesture recognition for taking order service of elderly care robot," in *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*. IEEE, 2014, pp. 525–530.
- [13] K. Berns and S. A. Mehdi, "Use of an autonomous mobile robot for elderly care," in *Advanced Technologies for Enhancing Quality of Life (AT-EQUAL)*, 2010. IEEE, 2010, pp. 121–126.
- [14] B. Graf, M. Hans, and R. D. Schraft, "Care-o-bot II Development of a next generation robotic home assistant," *Autonomous robots*, vol. 16, no. 2, pp. 193–205, 2004.
- [15] R. Sparrow and L. Sparrow, "In the hands of machines? the future of aged care," *Minds and Machines*, vol. 16, no. 2, pp. 141–161, 2006.
- [16] W. Wallach and C. Allen, *Moral machines: Teaching robots right from wrong*. Oxford University Press, 2008.
- [17] A. Sharkey and N. Sharkey, "Granny and the robots: ethical issues in robot care for the elderly," *Ethics and Information Technology*, vol. 14, no. 1, pp. 27–40, Mar 2012.
- [18] K. Zsiga, G. Edelmayer, P. Rumeau, O. Péter, A. Tóth, and G. Fazekas, "Home care robot for socially supporting the elderly: focus group studies in three european countries to screen user attitudes and requirements," *International Journal of Rehabilitation Research*, vol. 36, no. 4, pp. 375–378, 2013.
- [19] M. Ding, T. Matsubara, Y. Funaki, R. Ikeura, T. Mukai, and T. Ogawara, "Generation of comfortable lifting motion for a human transfer assistant robot," *International Journal of Intelligent Robotics and Applications*, vol. 1, no. 1, pp. 74–85, Feb 2017.
- [20] S. S. Srinivasa, D. Ferguson, C. J. Helfrich, D. Berenson, A. Collet, R. Diankov, G. Gallagher, G. Hollinger, J. Kuffner, and M. V. Weghe, "Herb: a home exploring robotic butler," *Autonomous Robots*, vol. 28, no. 1, p. 5, 2010.
- [21] R. Robotics, "Baxter industrial robot," <https://web.archive.org/web/20180514010257/https://www.rethinkrobotics.com/baxter/>.
- [22] E. S. Aveiga, "State estimation and feedback controller design for autonomous navigation of a high-performance mobile robot," Imperial College London, 2017.
- [23] D. Thomas, D. Scholz, and A. Blasdel, "rqt documentation," <http://wiki.ros.org/rqt>.
- [24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [25] A. Geitgey, "Face recognition dlib Python interface," https://github.com/ageitgey/face_recognition/blob/master/README.md, 2018.
- [26] "dlib face recognition documentation," https://web.archive.org/web/20170302162536/http://dlib.net/dnn_face_recognition_ex.cpp.html, 2018.
- [27] C. M. University, "CMU Sphinx documentation," <https://cmusphinx.github.io/wiki/>, 2018.
- [28] eSpeak, "eSpeak text to speech," <http://espeak.sourceforge.net/index.html>, 2018.
- [29] S. Kohlbrecher, "Hector mapping ROS package," http://wiki.ros.org/hector_mapping, 2018.
- [30] C. Rosmann, W. Feiten, T. Wosch, F. Hoffmann, and T. Bertram, "Efficient trajectory optimization using a sparse model," European Conference on Mobile Robots Barcelona pp 138-143, 2013.
- [31] C. Rosmann, "Timed elastic band algorithm implementation," http://wiki.ros.org/teb_local_planner, 2018.
- [32] S. Lemaignan, "ROS Markers Chili tags," https://github.com/chili-epfl/ros_markers.
- [33] R. Robotics, "Inverse kinematics solver service," http://sdk.rethinkrobotics.com/wiki/IK_Service_-Code_Walkthrough.
- [34] S. R. Ahmadzadeh, P. Kormushev, and D. G. Caldwell, "Visuospatial skill learning for object reconfiguration tasks," in *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems (IROS 2013)*, 2013.
- [35] P. Kormushev and S. R. Ahmadzadeh, "Robot learning for persistent autonomy," in *Handling Uncertainty and Networked Structure in Robot Control*, L. Busoni and L. Tam'as, Eds. Springer International Publishing, February 2016, ch. 1, pp. 3–28.
- [36] P. Gajewski, P. Ferreira, G. Bartels, C. Wang, F. Guerin, B. Indurkha, M. Beetz, and B. Sniezynski, "Adapting everyday manipulation skills to varied scenarios," *arXiv preprint arXiv:1803.02743*, 2018.
- [37] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz, "Towards 3d point cloud based object maps for household environments," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 927–941, 2008.