Autonomous Approach to Human Physical Assistance by a Humanoid

Anastasia Bolotnikova, Sébastien Courtois, Abderrahmane Kheddar

Autonomous Approach to Human Physical Assistance by a Humanoid

Anastasia Bolotnikova[1,2], Sébastien Courtois[1], Abderrahmane Kheddar[2]

***Abstract*— We study the use of humanoid robot technology for physical assistance in motion for a frail person. A careful design of a whole-body controller for humanoid robot needs to be developed in order to ensure efficient, intuitive and secure interaction between humanoid-assistant and human-patient. In this work, we present a design and implementation of a whole-body controller that enables a humanoid robot with mobile base to autonomously reach a person, perform audiovisual communication of intent, and establish several physical contacts for initiating an assistance in a given human motion. Our controller uses (i) visual human perception as a feedback for navigation and (ii) joint residual signal based contact detection for closed-loop physical contact creation. We assessed the developed controller on a healthy subject and report on the experiments achieved and the results.**

I. Introduction

According to the most recent predictions, by the year of 2050, 1 out of every 6 people in the world will be over the age of 65, compared to 1 out of every 11 for the same age group in 2019 [1], [2]. As a result, demand for the care workers, who can help frail elderly with everyday routine tasks, is expected to increase dramatically in the years to come. However, with a shortage of people of working age who could supply this help rethinking of the caregiving operating industry is required. Moreover, the recent Covid19 outbreak brought to light the crucial role such robots can play in future pandemics.

Robots are expected to play a pivotal role in the future of healthcare for elderly and disabled, improving both quality of life of patients and quality of work of human caregivers [3]. Particularly, user-friendly humanoid robots [4] may be a good fit for the job thanks to their familiar shape [5], potential multi-functionality and natural communication capabilities of a *social robot* specifically designed to interact with people in a natural, intuitive and useful way [6].

In this work, we consider a general interaction scenario which may occur in everyday care. A humanoid robot is required to autonomously reach a person and establish several physical contacts with a human participant with intention to initiate physical assistance in motion. A whole-body quadratic programming (QP) controller framework[1] is used to develop a Finite State Machine (FSM) controller that enables a humanoid robot to autonomously participate in the described interaction scenario.

A visual human perception is used for closed loop navigation towards a human. The joint residual signal based contact detection is used to determine established contacts with the human in closed-loop [7], [8] after a postural planning according to the task to be achieved [9]. Additionally, verbal, visual and body language communication are included in the controller design to enable the robot to autonomously communicate its intentions to a human in an audiovisual seamless way.

More specifically, our contributions in this work are the following:

- the design of the whole-body controller architecture for the interaction scenario (Sec. III-A);
- the implementation of a visual feedback based navigation towards a human (Sec. III-B);
- the integration of multi-modal communication of intent in the controller design (Sec. III-C);
- finally, we present the results of using the controller for interaction trials with few healthy subjects (Sec. IV).

II. Background

In the field of research on physical human-robot interaction for assistance in motion or power augmentation, a large majority of the works consider either the scenarios where human is creating a contact on the robot body surface [10], [11] or the application of exoskeletons [12]. In our work, however, the roles are reversed, it is the robot who is responsible for autonomously establishing a physical contact. An interaction scenario where robot itself is actively and autonomously creating contacts on a surface of a human participant body is rarely considered. In [13] a control of a robot establishing a contact on human body was studied with consideration of safety limits in terms of the human pain tolerance limits. In [14] a humanoid robot was used in a sit-to-stand assistance context to contact a human and perform a motion while maintaining a contact.

Human-aware navigation is a basic skill that robot must have when operating in the same environment as human [15]. A sensor fusion approach was developed and tested in [16]. More often a human avoidance strategies in navigation are studied [17], [18]. Visual servoing was used to make robot navigate towards human, maintain constant distance between itself and human, and follow a walking human. This visual servoing based control was presented in [19], [20], however without including the details on the navigation method implementation in the publication itself, however, a demo of the navigation is available online[2].

Yet, none of the previous works have considered a full autonomous interaction scenario with integration of all com-

[1]SoftBank Robotics Europe, Paris, France

[2]University of Montpellier–CNRS LIRMM, Interactive Digital Humans, Montpellier, France

[1]https://jrl-umi3218.github.io/mc_rtc/

[2]https://youtu.be/QDmDY5koKIE

ponents of visual feedback based navigation towards a person, multi-modal communication of intent and closed-loop physical contact creation. In the next section, we describe the implementation of the whole-body QP controller that enables the robot to autonomously participate in the studied interaction scenario.

III. Controller Architecture

The whole-body controller is developed using `mc_rtc` QP task-space control framework. In this work we consider a complex interaction scenario which is comprised of several distinct stages, each of which consists in achieving multiple objectives. Therefore, the controller is implemented in a form of an FSM, where every state corresponds to a specific stage of the interaction process. Figure 1 presents the general structure of the implemented FSM controller. In Sec. III-A we describe this controller implementation in detail.

A. FSM QP controller implementation

The robot actions are controlled by an acceleration-resolved QP controller. This means that an optimization problem is formulated with a quadratic objective function that consists of a weighted sum of *tasks* formulated as errors between actual and desired setpoints in task space, as well as first and second derivatives of these task errors. Robot joints acceleration are the decision variables of the problem. A set of linear *constraints* in the optimization problem formulation ensure that solution is physically feasible and safe.

As indicted in Fig. 1, the interaction starts with controller initialization (*init*). At this stage, a robot description module is used to build the base control problem as QP with default tasks and constraints (Eq. 1) that remain in the problem formulation till the end of the experiment. In this study, an upgraded prototype of Pepper humanoid robot platform is used in the experiments and corresponding robot description module is created for building the control problem.

$$\min_{\ddot{\boldsymbol{q}},\boldsymbol{f}} \quad \mathcal{P}+\mathcal{B}+\mathcal{C} \tag{1a}$$

$$\text{s. t.} \begin{cases} \text{joint position, velocity and torque limits} & \text{(1b)} \\ \text{self–collision avoidance} & \text{(1c)} \\ \text{sliding ground contact constraints} & \text{(1d)} \\ \text{bound mobile base velocity and acceleration} & \text{(1e)} \end{cases}$$

The objective function of the QP contains a default posture task $\mathcal{P}$, default mobile base position task $\mathcal{B}$ and center of mass task $\mathcal{C}$. Interested reader can refer to [21] for detailed definition of these common QP tasks and constraints.

After the controller initialization, transition to *InitialPosture* state is triggered. At this point we assume that the robot is positioned such that it is far from a human participant (at least 1.5m) and facing the participant. At the start of this state, the default posture task $\mathcal{P}$ is given a new setpoint q_d, which is an upright standing default posture of the robot. We define a threshold $\delta_{\mathcal{P}}$ and consider this state to be completed when the posture task error $\epsilon_{\mathcal{P}}$ is less than of equal to this threshold, which triggers transition to the *NavigateToHuman*.

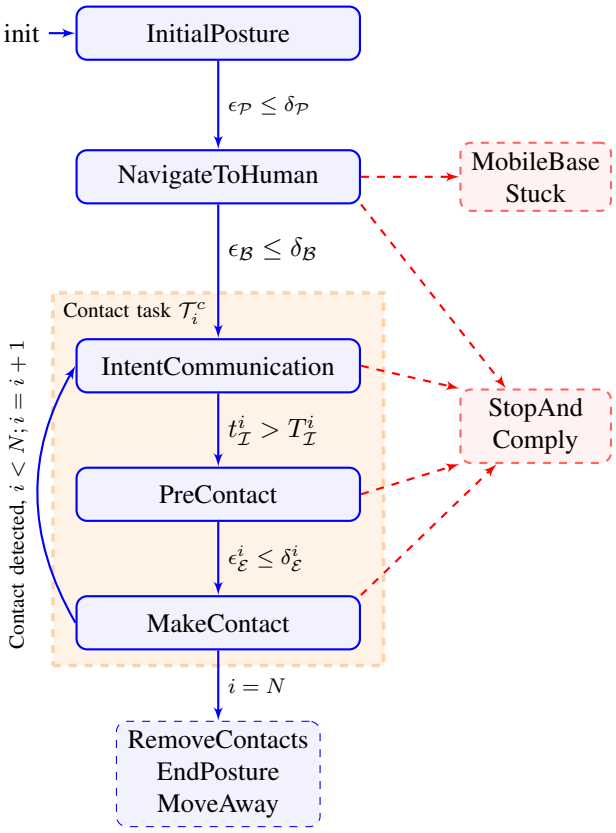


Fig. 1: FSM based controller design for interaction scenario.

In the *NavigateToHuman* state a mobile base position task is given a new set-point ${}^{w}X_b^*$ which is defined as a relative position and orientation w.r.t human. We detail how this set-point is computed and reached in closed-loop in Sec. III-B. Termination of this state is triggered when mobile base position task error $\epsilon_{\mathcal{B}}$ reaches a predefined threshold $\delta_{\mathcal{B}}$.

In case of normal flow of the experiment, the robot starts executing a sequence of predefined contact tasks $\mathcal{T}_i^c, i = 0, \ldots, N$. For every contact task from the sequence, the first state (of the sub-FSM) is *IntentCommunication* where the robot uses verbal, visual and body language communication modalities to explain and illustrate its further intentions specific to $\mathcal{T}_i^c$ to the human participant (Fig. 3). This state is implemented to ensure smooth and comprehensive transition to the states where the robot establishes physical contacts with the human participant. We detail the purpose and implementation of this state in Sec. III-C. The time for intent communication for every contact task is predefined $T_{\mathcal{I}}$. Once state time $t_{\mathcal{I}}$ exceeds $T_{\mathcal{I}}$ intent is considered to be successfully communicated and transition to *PreContact* state is triggered. Establishing physical contact on human body is done using methods presented in [7], [8]. When all contact tasks are finished the experiment ends with *RemoveContacts* state, final posture and moving away from a human participant.

B. Nearby surrounding navigation towards human

We define a fixed mobile base position and orientation target w.r.t human reference frame ${}^{h}X_{b}^{*}$. Human reference frame (*Pelvis* link) pose in the robot camera frame are obtained from human body detection by the Azure Kinect camera installed on the robot, we denote this transformation as ${}^{c}X_{h}$. From the joint encoder readings and known robot kinematics we compute transformation between camera frame and mobile base frame ${}^{b}X_{c}$, which at the start of the experiment is equal to the world frame ${}^{b}X_{c} = {}^{w}X_{c}$ at $t = 0$, where t is the time elapsed from the start of the experiment. The desired setpoint pose ${}^{h}X_{b}^{*}$ in the camera frame is computed as follows

$$ {}^{c}X_{b}^{*} = {}^{c}X_{h}\,{}^{h}X_{b}^{*} \tag{2} $$

This update is performed on every iteration of the controller. Due to the fact that a camera and the body tracking frame rate (30Hz and 20Hz respectively) is much lower than the control frame rate (83Hz in case of Pepper), new detection data is not available for every controller iteration. This may result in sudden task error jumps and discontinuities. To ensure a smooth PBVS task error evolution and convergence, in this work, we set lower limits for the mobile base speed and acceleration.

The error between current and desired frames is used as a feedback error for the mobile base task $\mathcal{B}$ to navigate to the desired position:

$$ \epsilon_{\mathcal{B}} = {}^{c}X_{b}({}^{c}X_{b^*})^{-1} \tag{3} $$

All frames involved in these computations are illustrated on simulation setup in Fig. 2.

Detected human head frame is used for a camera orientation target task implemented in the controller as an Image Based Visual Servoing (IBVS) task $\mathcal{V}$. With ${}^{\text{head}}X_{c}$ being a human head farem pose in the camera frame detected by Azure Kinect, the $\mathcal{V}$ task error which ensures that human head is kept as close as possible to the center of the field of view, (FoV) is following

$$ \epsilon_{\mathcal{V}} = {}^{\text{head}}X_{c}.\text{translation} \tag{4} $$

Based on our experience, keeping human head in the center of FoV results in a better overall human body tracking results from the Azure Kinect, especially once some human body parts are no longer in the FoV.

In the future we consider to combine this method with V-SLAM[3] technology for highly robust performance. This is necessary for better safety considerations and for handling low frame rate and hight frame latency of Azure Kinect depth sensor.

[3]We choose SLAM as it will certainly be a mature technology to navigate in between rooms indoor senior citizens' home, personal houses, hospitals, etc.

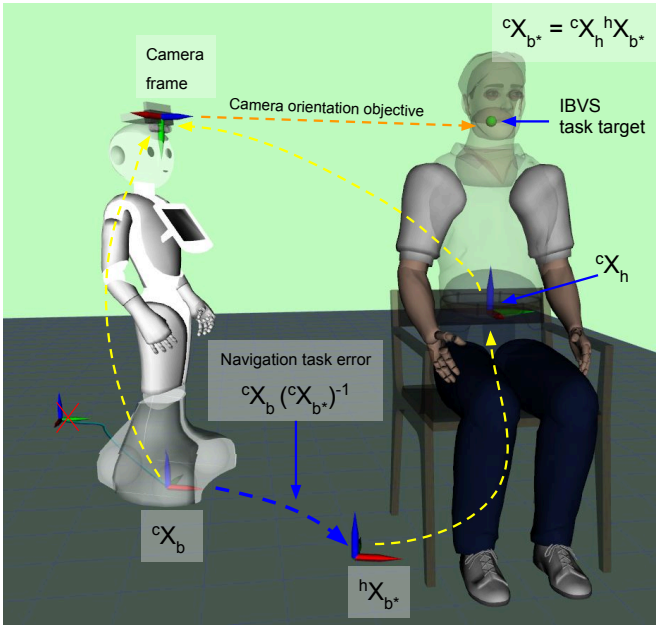


Fig. 2: *NavigateToHuman* FSM state visual tasks

C. Intent communication for user comfort and safety

Humans express interest to touch a humanoid robot when they see one. However, when the roles are reversed and it is the robot that will actively establish physical contacts on a human body, the loss of control from a human participant side, the lack of understanding of the robot intentions, and the low communicative cognitive capabilities of the robot can cause discomfort and even fear during the interaction. The closer a robot is to a person, the more considerations need to be taken into account for human comfort and safety. And in the case of direct physical contact, human participant safety and comfort (both physical and mental) are of extreme importance for successful human-robot interaction.

Indeed, human participant needs to trust the robot to feel comfortable to allow it to establish a physical contact. And this trust can be established only if the human can predict or know what the robot intentions are []. Therefore, we carefully integrate a good user experience design considerations as parts of the FSM controller. In order for human participant to feel comfortable while robot is in close proximity and is establishing contact on a surface of participant's body. For that robot is programmed to clearly communicate its intentions to the user, using different communication modalities (verbal, visual, body language).

In order to ensure that the robot interacts with human in a socially acceptable fashion and conveying its intention in a human-perceptible way, when the robot reaches the person closely enough and prior to transitioning directly to establishing multi-contacts, the robot reminds the participant that he can be stopped at any point.

Then prior to making any further motions robot is communicating verbally what it intends to do in the next phase of the experiment. Using body language the robot draws attention of a participant to its tablet screen where an illustration of

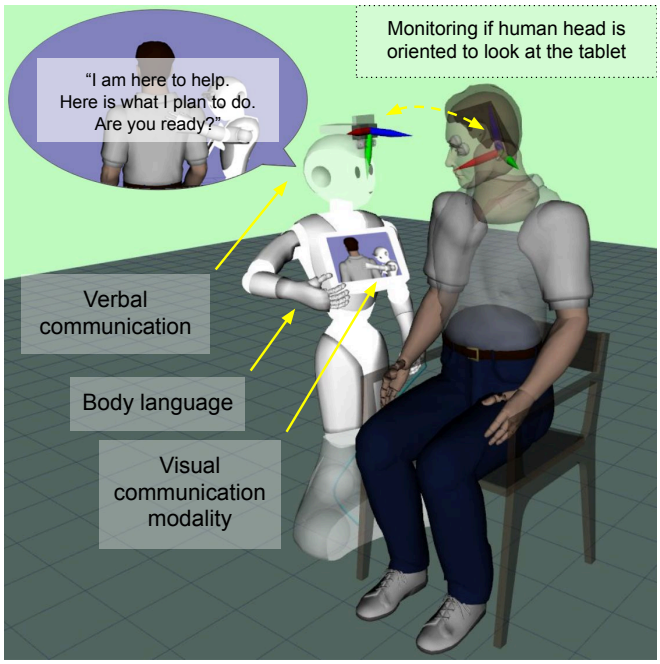


Fig. 3: Multi-modal *IntentCommunication* FSM state

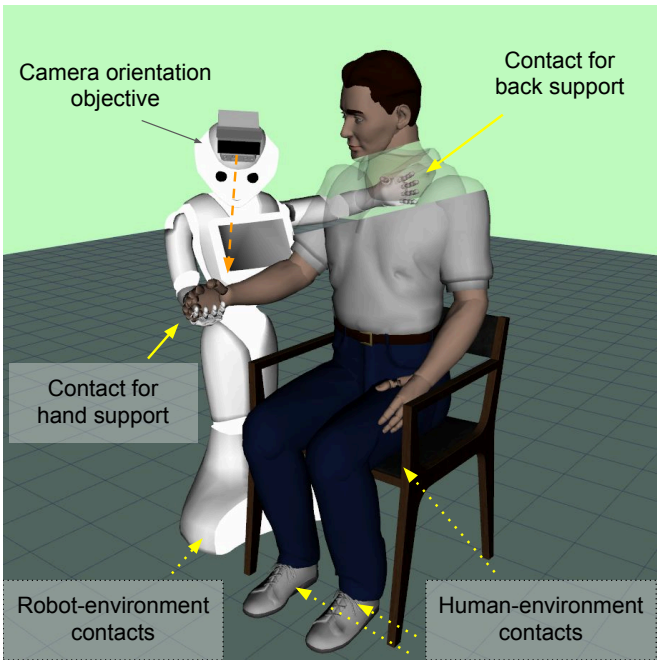


Fig. 4: *MakeContact* FSM state, all contacts established

the intended physical interaction motion figure is displayed as shown in Fig. 3.

With these familiar tools for information transfer, we enrich the interaction with different communication modalities and enable robot to make its intentions very clear prior to establishing physical contact.

D. Physical contact establishment

The physical contacts are established using end-effector QP tasks for robot's right and left hand links. Contacts are established in closed-loop via monitoring of residual signal between predicted (learned) and measured joint position tracking error [7], [8]. The exact feasible placement of the contacts can be planned by analyzing the human point cloud using the method presented recently [9].

E. Axillary states

Additionally to the main states, there are also several axillary states meant for safety and recovery from mobile base being stuck (*MobileBaseStuck*) or cases when contact is not detected (*StopAndComply*).

IV. Experimental Results

We have achieved preliminary experiments that will be enhanced further.[4]

[4]For the time being, these experiments couldn't be achieved with more persons and in good conditions because of the restrictions of the Covid19 outbreak. They were performed at the first author private house. This section will be enhanced with more trials and better conditions for the final version, shall the paper be accepted.

A. Platform description

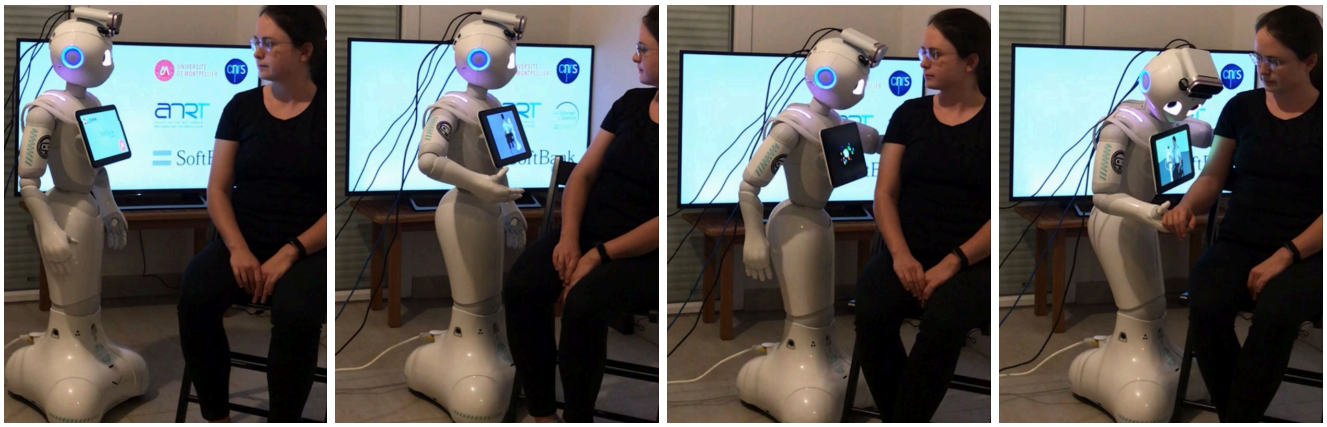
We are using the Pepper wheeled humanoid robot that have been customized for later in-situ experiments with real patients. With respect to the commercial version, this robot has additional Azur Kinect camera mounted on the top of the head, mainly for human body tracking. The robot is also equipped with the RealSense D434 camera for SLAM and additional source of IMU measurement that we are not using in this paper. Under the tablet, this Pepper embed also with a Jetson TX2 for image processing and other heavy computation. For the time being, we are running the controller on an external PC for easy debugging and programming purposes.

B. Results

At the start of the preliminary experiment, the robot is placed about 1.2 meters away from a chair where human participant is sitting. First, robot navigates towards the person, then communicates its intent, and then proceeds to initiate several physical contacts. First contact robot establishes one on the right shoulder of the human participant. Then, robot invites participant to place a hand in its right end-effector (as being demonstrated on the tablet screen of the robot). The experiment ends with all contacts being autonomously removed and robot thanking the participant.

Fig. 5 illustrates 4 main stages of the experimental controller run with real human participant. Interested reader is invited to see the video accompanying this paper for the full experiment demonstration.

Plot on Fig. 6 shows evolution of the PBVS task errors in translational axes during the *NavigateToHuman* state (Fig. 5a). As can be seen from the plot, the task error evolution and convergence is mostly smooth. The only



(a) Navigation (b) Intent communication (c) First contact (d) Second contact

Fig. 5: Screenshots illustrating main parts of the experiment with human participant

sudden jump occurs at the end of the state for the Y axis, when the robot is very close to the human participant. Typically, at this stage the depth perception deteriorates due to human being too close to the camera and, as a result, discontinuities in human body tracking are also likely to occur. Nevertheless, the *NavigateToHuman* state completion criteria is successfully reached and transition to the next FSM state is triggered.

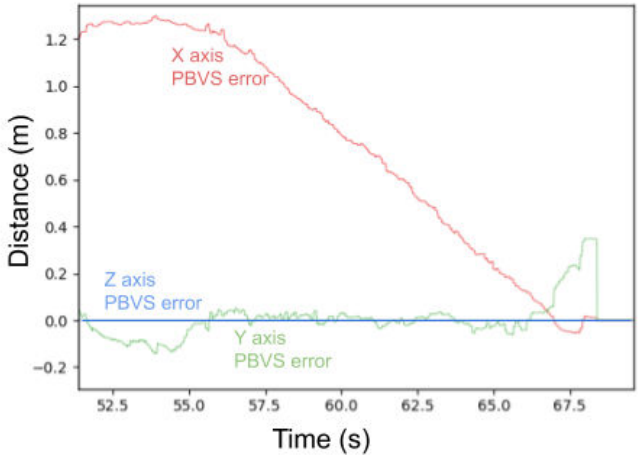


Fig. 6: Evolution of navigation task errors

Next state in the FSM (Fig. 1) is *IntentCommunication*. This part of the experiment can be seen on Fig. 5b. While communicating the intent using multiple communication modalities, as described in Sec. III-C, IBVS task is responsible for keeping human head in the field of view. The orientation of the human head frame is monitored and robot ensures that human face was oriented towards robot's tablet at least once for paying attention to the figure displayed on the screen. This is an additional criteria for exiting *IntentCommunication* state.

Once intent is communicated, the transition to *PreContact* and then *MakeContact* states for the robot's left end-effector are triggered. The robot tries to establish a physical contact on the right shoulder of the participant. The monitoring of the residual between measured position tracking error and predicted expected position tracking error allows to detect a contact event (as mentioned in Sec. III-D). Fig. 7 shows the residual signal for the left robot shoulder roll joint. As can be seen from the plot, the contact is detected (when residual reached a threshold) and maintained for several seconds after the detection occurs as is requested according to the *MakeContact* state design.

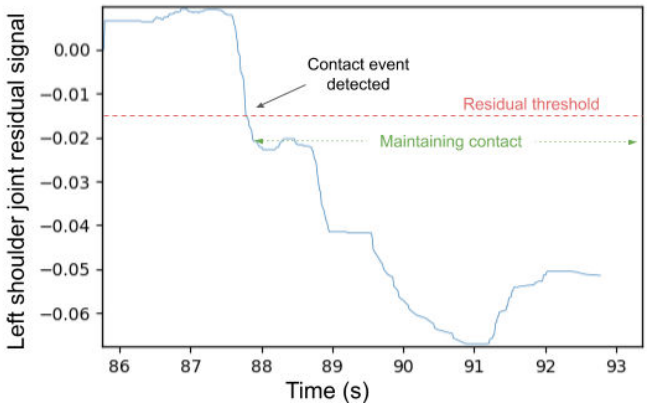


Fig. 7: Left end-effector contact detection

Similar process is repeated for the robot's right end-effector. However, in this case robot invites user participant to place her hand into robot's right end effector which is brought up in front of the user as is shown in (Fig. 5d). The plot on the Fig. 8 shows the residual signal for the robot right elbow roll joint and indicates successful detection of the contact event. This contact is also maintained for few seconds before being removed.

The experiment ends with all contacts being autonomously and carefully removed by the robot. The robot thanks human user for participation at the end of this experiment.

V. Conclusion and Future Work

In this paper we integrated and enhanced different bricks to build an architecture that enables the humanoid robot

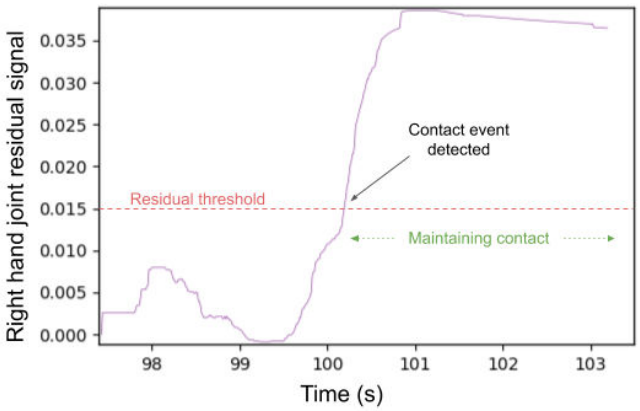


Fig. 8: Right end-effector contact detection

Pepper to reach a person in need for physical assistance to achieve some daily motions. We developed a task-space optimization controller instance derived from [21] and enhanced with visual servoing [22], which allows reaching safely a person in various postures dictated from a multi-contact planning [9]. In close contact interaction the robot initiate contacts [7] to achieve compliant contact motion [8]. The controller was enhance with communication plug-ins. We also discuss vision shortcomings when the robot is in close physical interaction with a human.

We believe that all the bricks that have been developed are a viable ground technologies to effectively achieve frail pHRI purpose in real experiments that are planed to be conducted *in-situ* a retirement house with few real patients by the end of this year. Engineering efforts are now undertaken to clean the existing software to meet safety requirements and recovery procedures. SLAM technology will certainly mature to a real-time reliable and commercially available software and is integrated (but not used in this paper) to navigate in indoor spaces. We forecast the same maturation for human body tracking. Task-space optimization have been proved extremely flexible and reliable in many industrial set-up. In near future work, the human model will be integrated as part of the multi-robot QP control model [21]. Thereby, we are synthesizing an adaptive force controller in the task space to allow the robot to effectively assist the site-to-stand motion, human limbs manipulation, etc. The latter problem is challenging because a human is not equipped with encoders or torque sensing which would allow closing the loop and computing the right amount of assistive forces. We will also redesign the balance criteria for the humanoid to increase safety of the dyad during pHRI.

REFERENCES

[1] "World population prospects 2019. highlights," report ST/ESA/SER.A/423, United Nations, Department of Economic and Social Affairs, Population Division, New York, US, 2019.

[2] "World population ageing 2019. highlights," report ST/ESA/SER.A/430, United Nations, Department of Economic and Social Affairs, Population Division, New York, US, 2019.

[3] M. Niemelä and H. Melkas, "Robots as social and physical assistants in elderly care," in *Human-Centered Digitalization and Services*, vol. 19, pp. 177–197, 2019.

[4] A. Pandey and R. Gelin, "A mass-produced sociable humanoid robot: pepper: the first machine of its kind," *IEEE Robotics & Automation Magazine*, vol. 25, no. 3, pp. 40–48, 2018.

[5] M. Staffa and S. Rossi, "Recommender interfaces: the more human-like, the more humans like," in *International Conference on Social Robotics*, (Kansas City, US), pp. 200–210, 1–3 November 2016.

[6] S. B. Daily, M. T. James, D. Cherry, J. J. P. III, S. S. Darnell, J. Isaac, and T. Roy, "Affective computing: historical foundations, current applications, and future trends," in *Emotions and Affect in Human Factors and Human-Computer Interaction*, pp. 213–231, Elsevier, 2017.

[7] A. Bolotnikova, S. Courtois, and A. Kheddar, "Contact observer for humanoid robot pepper based on tracking joint position discrepancies," in *IEEE International Conference on Robot and Human Interactive Communication*, (Nanjing, China), pp. 29–34, 27–31 August 2018.

[8] A. Bolotnikova, S. Courtois, and A. Kheddar, "Compliant robot motion regulated via proprioceptive sensor based contact observer," in *IEEE-RAS International Conference on Humanoid Robots*, (Beijing, China), pp. 854–859, 6–9 November 2018.

[9] A. Bolotnikova, S. Courtois, and A. Kheddar, "Multi-contact planning on humans for physical assistance by humanoid," *IEEE Robotics and Automation Letters*, vol. 5, no. 1, pp. 135–142, 2020.

[10] Y. Tirupachuri, G. Nava, L. Rapetti, C. Latella, and D. Pucci, "Trajectory advancement during human-robot collaboration," in *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, (New Delhi, India), pp. 1–8, 14–18 October 2019.

[11] F. Romano, G. Nava, M. Azad, J. Čamernik, S. Dafarra, O. Dermy, C. Latella, M. Lazzaroni, R. Lober, M. Lorenzini, *et al.*, "The codyco project achievements and beyond: Toward human aware whole-body controllers for physical human robot interaction," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 516–523, 2017.

[12] L. M. Vaca Benitez, M. Tabie, N. Will, S. Schmidt, M. Jordan, and E. A. Kirchner, "Exoskeleton technology in rehabilitation: Towards an emg-based orthosis system for upper limb neuromotor rehabilitation," *Journal of Robotics*, 2013.

[13] Y. Yamada, Y. Hirasawa, S. Huang, Y. Umetani, and K. Suita, "Human-robot contact in the safeguarding space," *IEEE/ASME transactions on mechatronics*, vol. 2, no. 4, pp. 230–236, 1997.

[14] A. M. López, J. Vaillant, F. Keith, P. Fraisse, and A. Kheddar, "Compliant control of a humanoid robot helping a person stand up from a seated position," in *IEEE-RAS International Conference on Humanoid Robots*, (Madrid, Spain), pp. 817–822, 18–20 November 2014.

[15] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1726–1743, 2013.

[16] M. Tee Kit Tsun, B. T. Lau, and H. Siswoyo Jo, "An improved indoor robot human-following navigation model using depth camera, active ir marker and proximity sensors fusion," *Robotics*, vol. 7, no. 1, p. 4, 2018.

[17] 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, pp. 99–112, 2010.

[18] A. R. Araujo, D. D. Caminhas, and G. A. Pereira, "An architecture for navigation of service robots in human-populated office-like environments," *IFAC-PapersOnLine*, vol. 48, no. 19, pp. 189–194, 2015.

[19] G. Claudio, F. Spindler, and F. Chaumette, "Vision-based manipulation with the humanoid robot romeo," in *IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pp. 286–293, 2016.

[20] D. J. Agravante, G. Claudio, F. Spindler, and F. Chaumette, "Visual servoing in an optimization framework for the whole-body control of humanoid robots," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 608–615, 2016.

[21] K. Bouyarmane, K. Chappellet, J. Vaillant, and A. Kheddar, "Quadratic programming for multirobot and task-space force control," *IEEE Transactions on Robotics*, vol. 35, no. 1, pp. 64–77, 2019.

[22] A. Paolillo, K. Chappellet, A. Bolotnikova, and A. Kheddar, "Interlinked visual tracking and robotic manipulation of articulated objects," *IEEE Robotics and Automation Letters*, vol. 3, pp. 2746–2753, October 2018.