Autonomous Multi-Sensory Robotic Assistant for a Drinking Task

Felix Ferdinand Goldau
Institute of Automation
University of Bremen
Bremen, Germany
fgoldau@uni-bremen.de

Tejas Kumar Shastha
Institute of Automation
University of Bremen
Bremen, Germany
tejas@uni-bremen.de

Maria Kyrarini
Institute of Automation
University of Bremen
Bremen, Germany
mkyrar@iat.uni-bremen.de

Axel Gräser
Institute of Automation
University of Bremen
Bremen, Germany
ag@iat.uni-bremen.de

Abstract—**Assistive robots have the potential to support people with disabilities in their Activities of Daily Life. The drinking task has a high priority and requires constant assistance by caregivers to be executed regularly. Due to incapacitating disabilities such as tetraplegia, which is the paralysis of all limbs, affected people cannot use classic control interfaces such as joysticks. This paper presents a robotic solution to enable independent, straw-less drinking using a smart cup and no physically attached elements on the user. The system's hardware and software components are presented and the overarching control scheme described. The cup approaches the mouth utilising a user-friendly and vision-based robot control based on head pose estimation. Once contact has been established, the user can drink by tilting the cup with a force sensor-based control setup. Two experimental studies have been conducted, where the participants (mostly able-bodied and one tetraplegic), could separately experience the cup's contactless approach and the contact-based sequence. First results show a high user acceptance rate and consistent positive feedback. The evaluation of internal data showed a high reliability of the safety-critical components with the test groups perceiving the system as intuitive and easy to use.**

Keywords—assistive robots, human-robot interaction, force control, head pose estimation

I. INTRODUCTION

People suffering from a severe disability, like tetraplegia, have difficulties performing Activities of Daily Living (ADLs). Tetraplegia is the paralysis of four limbs, thus limiting voluntary motor function of everything below the neck and, even though the treatment of paralysis has undergone great progress, people require the assistance of a caregiver to perform ADLs [1].

Assistive robotic manipulators have the potential to support individuals with tetraplegia to regain some of their independency in performing ADLs. One example is the wheelchair-mounted robotic manipulator FRIEND, which was used as a personal assistant for a tetraplegic end-user in performing ADLs and tasks in a working environment [2, 3]. A survey with potential end-users of robotic manipulators shows that drinking and eating are highly prioritized tasks [4].

This paper presents a robotic solution as an attempt to enable a person with tetraplegia to independently perform a drinking task using a cup, without external human aide and without any physically user-attached elements. The drinking task is executed without a straw or similar device and must therefore be accomplished with contact between the user and the robot-handled cup. The goal to be achieved, is to give the user a feeling of sovereignty over their own drinking and the perception of performing the drinking themselves, as opposed to being served a drink. First results and user feedback based on two small studies are presented.

The main contribution of the presented work is a consistent, user-friendly, and fully flexible concept of control with the human in the loop as opposed to a system using predefined positions. This allows a safe and intuitive human-robot interaction. The drinking is fully personalised with the user gaining control and comfort using natural head movements for the delivery of the cup to the mouth, and achieving adaptive control of the drinking process using cup-mounted force sensors.

The paper is organized as follows: In section II, related work is reviewed. Section III describes the proposed framework and Section IV presents the experimental results and the evaluation. Finally, Section V discusses the conclusion with directions for future work.

II. RELATED WORK

An important topic in assistive robotics is Human-Machine Interfaces (HMIs), as users are often restricted in their movements and standard computer interfaces might not be usable. Based on a study using a vision-guided robot arm, [5] shows that many systems are too complicated for their respective end-users who have to follow tedious HMI-sequences. It also shows that systems should adopt higher velocities when not close to the user to minimise waiting times and user frustration.

Various systems have been developed within the field of assistive robotics, most of them based on a Wheelchair Mounted Robot Arm (WMRA). For example, the FRIEND systems I-IV [2] where FRIEND IV was capable of enabling a tetraplegic person to work as a librarian. The used HMI consists of a set of buttons operated by the user with head motions, a chin joystick, and a monitor for system feedback. This setup can be seen as a head operated computer mouse which allows the user to move and click with a cursor on the monitor to control the wheelchair and WMRA. Previous versions also addressed the drinking task, but were for safety reasons restricted to drinking with a straw [3].

Another approach to assistive drinking shown by [6] uses a Brain-Machine Interface (BMI) to control a robotic manipulator holding a cup for drinking without a straw. An externally mounted RGB-D camera is used to estimate the user's mouth by applying a golden-ratio approach on the detected face, thus bypassing issues of face occlusion by the cup and the robot. The user controls the scenario by giving GO-signals via the BMI and can thereby accelerate or stop the procedure in seven discrete predefined steps.

Moreover, hands-free human-robot interfaces could be used to assist individuals suffering from tetraplegia. Two interfaces using eye gestures [7] and Brain-Computer Interface (BCI) [8] have been developed to allow users with tetraplegia to control a 7 Degrees-of-Freedom (DoF) robotic arm and its gripper. A manipulation task was selected to evaluate both interfaces for the robot control. For research purposes the robotic degree of automation was set to zero, thus the user had to control the entire process step by step. It was shown that this kind of robot control burdens the user with high cognitive load due to the fast response needed during the drinking process.

III. Proposed Method

This section proposes a solution to the drinking task with contact between the user and the cup, maintaining user safety as an active element inside the loop and without using predefined steps. The drinking procedure is divided in two consecutive sequences which are individually described below:

- The 'Vision-based Robot Control to Serve a Drink' handles the delivery of a grasped cup to the user's mouth based on camera input, whereas
- the 'Robot Force Control for the Drinking Process' deals with the process of tilting the cup based on the force applied onto the cup by the user, thus enabling drinking.

A. *Hardware Setup*

The Kinova Jaco 2 [9] 7-DoF spherical ultra-lightweight robotic arm with a three-finger gripper attached as the end-effector (Fig. 1a) is used as the main assistive robotic manipulator. The Jaco 2 has been specifically designed for use as an assistive robot and has been thoroughly tested in a scientific context [10].

The robot arm grasps a smart cup (Fig. 1b), developed in previous work [11], consisting of a feeding cup with a beak, two force sensors, and a Bluetooth module which wirelessly transmits the force values to the operating computer. The force sensors are attached to the beak of the cup, just above and below the mouth piece. In the context of this work, the beak of the mouth piece is considered as the cup's origin.

a b

Fig. 1. The hardware components a) assistive robotic manipulator with vision sensor, b) smart cup with force sensors [11]

An Intel Realsense Depth Camera D435 [12] is selected as the vision sensor and is mounted on the robot's end-effector between the gripper and the last joint using a smooth 3D-printed attachment clamp. The vision sensor is USB powered and provides RGB-D (Red Green Blue - Depth) information of the scene ahead, with an RGB resolution of up to 1920x1080 pixels and a depth sensor range between 0.105 m and 10 m. The short distance between the camera and the robot's end-effector minimises the occlusion of the scene by the manipulator, as only the smart cup and the gripper's fingertips can be seen obstructing the scene. For this application, the camera is configured to allow minimal depth measurements, thus being able to detect the distances of objects right up to the tip of the cup.

A UNIX computer is used to combine all systems, interpret the sensor data and control the robot manipulator accordingly. The proposed system is developed on Robot Operatic System (ROS) [13].

B. *Vision-based Robot Control to Serve a Drink*

The concept of vision-based robot control, which is used to serve a drink, is shown in Fig. 2. The main idea is that the robot delivers the smart cup to the user's mouth using data gathered by the vision sensor. The information from the vision sensor is processed by the following modules: Vision-based User Face Detection and Tracking, Head Pose Estimation, and Mouth Pose Estimation. The task control module calculates the necessary robot action based on the mouth pose estimation and the robot pose, and controls the robot accordingly. Furthermore, the task control module uses the force data from the smart cup to ensure safe human-robot interaction. The modules in Fig. 2 are explained in detail as follows.

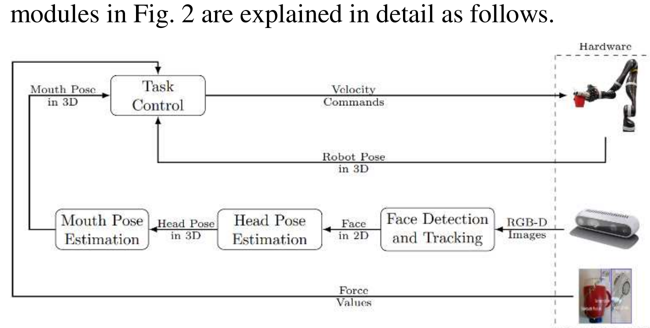

Fig. 2. Vision-based Robot Control to Serve a Drink

1. *Face Detection and Tracking*

The frontal-face detector of the dlib open source library [14] is used to search the 2D RGB-data of the vision sensor (camera) for a human face in order to detect the user. This detector is based on linear Support Vector Machine (SVM) using features of Histograms of Oriented Gradients (HOG) and has been trained on images of frontal faces from the Labeled Faces in the Wild database [15]. The output of this module is the 2D bounding box of the user's face.

In case the face is not detected, but information about a recent valid face detection from a previous iteration is available, a Discriminative Correlation Filter with Channel and Spatial Reliability (CSR-DCF) tracker [16] is used to estimate the bounding box of the face. This tracker compares the frequency domain of the relevant area and compares it to its surroundings to determine an object's movement in a sequence of frames.

2. Head Pose Estimation

Based on the detected user's face, a predefined 3D model of a human face is aligned with the data by applying an active shape model [17] based approach. The model is fitted by applying a cascade of regression trees, which have been previously trained on the 300-W dataset with 68 landmarks per face [18].

Using the landmark knowledge of the predefined 3D model, a PnP solution is calculated based on the RANSAC paradigm [19]. The result is a 3D position and orientation for the camera reference system (${}^{H}_{C}T$) in relation to the landmark coordinate system with the mouth at its origin. The head pose in relation to the world reference frame ${}^{W}_{H}T$ can be calculated by applying (1), with ${}^{W}_{E}T$ as the transformation of the robot's end-effector reference to the world reference system as result of the direct kinematics, and ${}^{E}_{C}T$ as the transformation of the camera reference system to the robot's end-effector system as defined by the rigid connection of the attachment clamp.

$$ {}^{W}_{H}T = {}^{W}_{E}T \cdot {}^{E}_{C}T \cdot ({}^{H}_{C}T)^{-1} \qquad (1) $$

The final head pose is validated twice to ensure user safety. This is done by partially reversing the previous procedure, thus reprojecting the face landmarks of the final pose onto the image plane of the camera. This projection outputs new 2D landmarks which are compared to the equivalent landmark projection of the fitted active model. If the error is too large, a misdetection is assumed and the pose is omitted. The second validation is performed in case information about the user's pose is available from previous iterations, by comparing this prior information to the most recently calculated user pose. If either the orientation or the position shows major changes, the chances of an error are raised and the head pose is again omitted for a higher user safety.

Even though not intended by the user, it is possible that they move their head too much during two iterations, thus causing the second validation step to fail. This behaviour is intended for an increased user safety, as the situation is considered dangerous if too much user movement occurs and the possibility of false detections rises.

3. Mouth Pose Estimation

After the head pose has been estimated successfully, the pose ${}^{W}_{M}T$ of the mouth with respect to the world reference frame can be directly derived from the result as shown in (2), because both poses share the same reference system and the user's mouth as their origin.

$$ {}^{W}_{H}T = {}^{W}_{M}T \qquad (2) $$

When the robot is advanced close to the user (distance between the beak and the mouth is approx. 5cm), the head pose cannot be estimated anymore. Therefore, another method to estimate the mouth pose is developed in this work. The proposed mouth pose estimation calculates a result based on information gathered in previous iterations and predefined knowledge of the underlying path control logic, by comparing the previously tracked position of the mouth with an updated prediction. Two assumptions are considered:

1. If they want to drink, the user does not move, and
2. the robot moves the smart cup in a direct path towards the mouth.

The first assumption is only a minor constraint on the user's behalf as they would also refrain from moving if another person served them instead of a robot; and the second assumption is system-defined. If the user were allowed to move and this assumption were not made, the user's mouth position would be undefined until the transformation is available again. This would result in a safety-critical situation.

Using those two assumptions, an estimation of the mouth's position with respect to the camera can be calculated. Initially, the position obtained by previous head pose estimations is used and adjusted afterwards, while the position change of the camera is known by applying the robot's direct kinematics. The vector between this new camera pose and the previous head pose is then scaled, such that its length corresponds to a selection of distance values measured by the vision sensor.

The calculated position of the mouth is reprojected into the image plane. From the projected landmark points, a bounding box of the mouth-and-nose-region is calculated. In the meantime, a CSR-DCF tracker [16] trained during previous head pose estimations tracks the same region based entirely on RGB information and also introduces a bounding box. The resulting rectangles of both bounding boxes are compared with respect to their relative overlapping area. If the area exceeds a predefined threshold, it is assumed that the correct position of the mouth is known, and thus the position of the mouth is accepted.

The result of the head and mouth pose estimation is shown in Fig. 3 with the 3D-landmarks as red dots and the resulting pose as a color-coded coordinate system (red = x-axis, green = y-axis, blue = z-axis) at the user's mouth. A white rectangle represents the area used to calculate the safety-critical user distance.

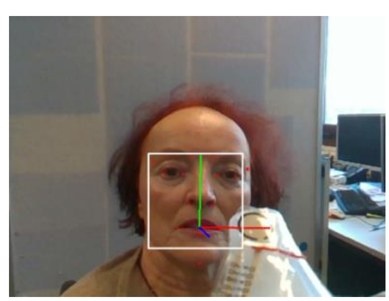

Fig. 3. Result of the Head and Mouth Pose Estimation with 3D-landmarks (red dots) and coordinate system (red = x-axis, green = y-axis, blue = z-axis)

In the event that the threshold is not met, it is implied that either the user has moved or another error has occurred. In either case, reliable information cannot be assumed and the calculated mouth pose is discarded. This also cancels future mouth pose estimations until a new head pose has been identified and the tracker is redefined.

4. Task Control

As defined by the control concept, the robot control is based on head pose information. The task control module is designed based on an analogy of a human assistant who serves a cup to the user and reacts to the user's head orientation. This results in the cup being brought to the user's mouth only if the head is oriented towards the cup. If this orientation is not given, the robot motion is stopped and, after a small delay, reversed towards a predefined home position of the robot.

When the head of the user is oriented towards the cup and the cup is within a predefined field of view with angle θ (in this work $\theta = 15°$) in either direction of the user's central z-axis, the robot moves the cup on a smooth and parabolic trajectory towards the mouth. The goal of this process can be described by (3) with ${}^{W}_{B}T$ as the pose of the cup's beak in relation to the world reference frame.

$$ {}^{W}_{M}T = {}^{W}_{B}T \qquad (3) $$

The pose of the beak ${}^{W}_{B}T$ can be calculated using the direct kinematics of the manipulator and structural information about the grasped cup as shown in (4).

$$ {}^{W}_{B}T = {}^{W}_{E}T \cdot {}^{E}_{B}T \qquad (4) $$

Fig. 4 shows an example path (purple) for the end-effector during the cup's approach towards a user with the respective coordinate systems of the user and the end-effector (red = x-axis, green = y-axis, blue = z-axis), and the θ-based field of view. Following a parabolic path, the cup initially converges towards the z-axis of the mouth before closing the distance. Once the cup has reached a distance of less than 5 cm to the mouth, the path is no longer defined by parabolic curves but moves the cup on a straight path directly to the user's mouth. The task control module sends velocity commands to the robot, in order to control it along the path.

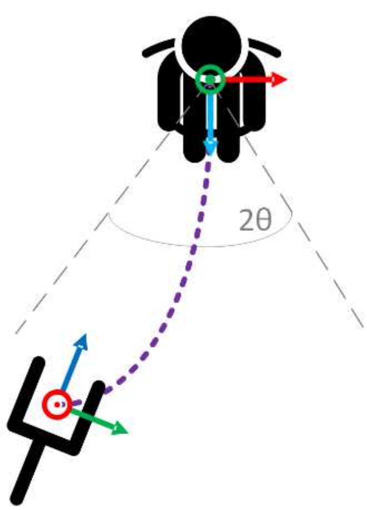

Fig. 4. Example path of the cup's approach

Moreover, in case the force sensors detect a contact at any point before the mouth is reached, the approach is stopped. If the contact has been initiated intentionally, the second sequence of the drinking process commences. (section III.C).

If, at any point, no head pose information is available, or if the user's orientation does not meet the requirements, the advance of the cup towards the user is stopped. If not updated within a short time period, the cup is retracted towards the home position, while the camera remains oriented towards the last known head position, thus enabling the user to regain control at any given point. The home pose is defined in a way that the user's head is seen by the camera in a standard scenario.

C. Robot Force Control for the Drinking Process

Once the smart cup, controlled by the vision-based robot control (Section III.B), gets close to the user's mouth, the force controller is initialised to enable the drinking process. The objective of this controller is to support the active drinking task wherein the user interacts with the smart cup by applying force to the force sensors (shown in Fig. 1b). Fig. 5 presents an overview of the force control for the drinking process. The inputs of this controller are the values read by the two force sensors on the cup and the current pose of the robot. The output is the Cartesian velocity command supplied to the robot. The controller drives the cup along a vertical plane which runs along the nose of the user, perpendicular to the face, to emulate a natural drinking motion.

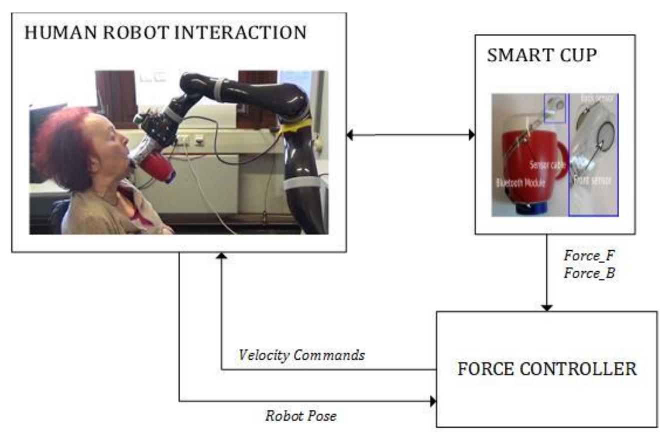

Fig. 5. Overview of force control for the drinking process

When the user applies force on the back sensor (Force_B), the cup's beak rotates down by a discrete angle (i.e.: feed motion). When the user applies force on the front sensor (Force_F), the cup's beak rotates up by a discrete angle (i.e.: non-feed motion). When forces are applied on both sensors simultaneously, the cup initiates a fall-back motion, moving laterally away from the user and stopping at the home position. The drinking task is completed.

To achieve this behaviour, three threshold values are defined for each of the force sensors:

- Trigger threshold - 0.5 N - Minimum amount of force to be applied to either sensors to trigger the respective action. Allows for slight pressure exerted when user swallows water.

Authors´ accepted version. © IEEE https://doi.org/10.1109/ICORR.2019.8779521

- Fall-back threshold – 1.0 N – Simultaneous force required on both sensors to initiate fall-back.
- Safety threshold - 2.5 N - Maximum amount of force allowed to be exerted before an emergency halt is issued to the robot, stopping all motion.

This control schema is initially build based on intuition and available hardware. The next section explains how it was tested and proven to be effective.

IV. Experimental Results and Evaluation

Two small studies were conducted: the first to evaluate the safety and usability of vision-based robot control and the second to evaluate the general usability of the force-controlled feeding system.

At the end, each user was presented with a general feedback questionnaire based on a 5-point Likert scale. The questions are agreement based and the Likert scale ranges from 1 (“strongly disagree”) to 5 (“strongly agree”). All participants gave their informed signed consent to participate in this study.

A. Serving a Drink

The first study is designed to evaluate the safety and usability of the cup serving scenario with the proposed solution. 25 users, one of whom is tetraplegic, participated in the experiments. The average age was 31.08 ± 14.55 years with a gender distribution of 13 males and 12 females (11 able-bodied and one tetraplegic). The experiments were conducted individually and independently for each subject.

Seven tasks were performed by each user. In each task, the smart cup is already grasped by the robot. The user is seated on a chair (or a wheelchair for the tetraplegic user) at a pre-defined pose relative to the robot. The concept of control described in section III.B.1 is applied in each task. The tasks were as follows:

0. The user freely tested the system for one minute to get familiar with it.

1. The user constantly oriented their head towards the robot’s gripper, thus effectively commanding the robot to bring the cup to their mouth on a direct path.

2. The user performed so-called abort actions: The robot advanced the cup towards the user, but the user aborted the task by turning their head away. When the robot detected such an abort action, it stopped and withdrew towards its home position. The user was encouraged to compare the abort actions as reactions to motions to the top, bottom, left, and right.

3-6. The last four scenarios were performed with different starting positions for the robot and with the user imagining a standard drinking application, but also allowing themselves to become distracted once in a while to include an abort action.

Fig. 6 shows the robot’s end-effector paths of five different users during scenario 1 with the home position in the top right and the users’ mouths on the left. Though every user sat on the same chair with a defined position relative to the robot, the pose of the mouth, and with it, the final position of the end-effector, differs vastly. This is due to a variety of reasons including different heights and head orientations and it shows the necessity of an adaptive system as proposed in this paper.

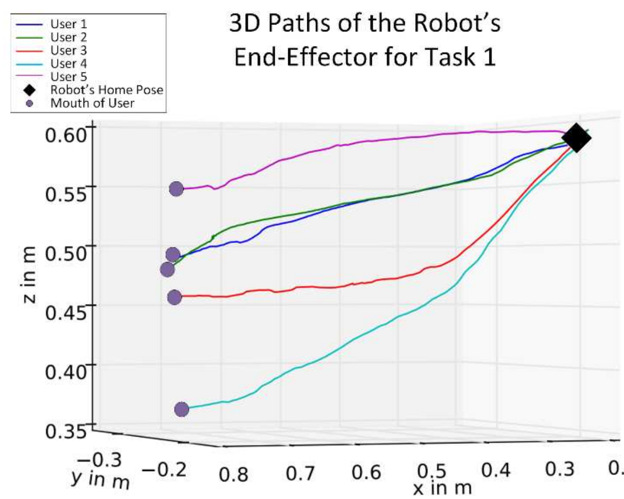

Fig. 6. 3D paths of the robot’s end-effector for 5 selected users in task 1

As the safety-critical elements during the cup serving sequence are mostly defined by the abort actions as well as the distance information obtained from the vision sensor and head pose estimation, these components are discussed in detail.

The outcome of abort actions during scenario 2 was manually sorted into the three categories: successful, delayed, and failed. Actions were considered delayed if the system did not react immediately and failed if an emergency stop was pressed or the user was forced to change the current abort motion in order to cause a reaction of the system. The users performed between 3 and 24 abort actions each, with 214 actions in total. Approximately 7 % of all actions were resolved with a delay and 4% of all actions failed, most of which occurred during upwards abort motions. The failing of upwards abort motions is caused by the fact that the cup already approached the user from above, which already requires an upwards tilting of the head. For the abort motion the users had to rotate their head very far upwards to use this specific abort action.

For redundancy purposes, the distance between the cup and the user is obtained using two separate methods. While one method uses the result of the previously mentioned head pose estimation, the other one is based on averaged values of the depth image in the region of the face. The robot control logic uses both methods and compares their values to lie within a defined offset for a single redundancy check. To evaluate the two methods during the experiments, the values of sequence 1 are compared to a reference value, which is calculated as the remaining distance of the robot’s end-effector to its final position. The results show an average error of 0.194 ± 0.038 m for the first method, and an average error of 0.142 ± 0.011 m for the second method respectively. The values show comparably high and constant systematic errors, which are mostly due to the offset of the final robot’s end-effector position to the mouth. This offset is a known factor and taken into account by the control setup. The remaining random errors are very small and considered as such in terms of safety critical actions.

The participants were able to test the system thoroughly in a realistic scenario during the final four sequences and reported their feedback on a questionnaire afterwards. The results of the Likert scale are shown in Fig. 7 as the mean values across subjects. The feedback is positive throughout with a rapid user familiarisation and high acceptability. The previously mentioned issues with the abort actions to the top are only slightly reflected by users' feedback. This is probably due to different home positions for the different tasks, thus not always approaching the user from above, but also from lower directions. The user feedback also shows a requirement to increase the velocity of the system.

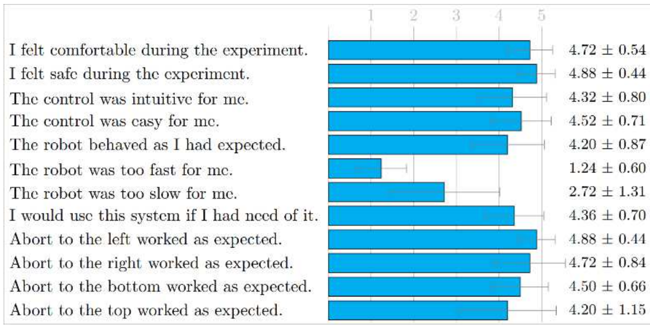

Fig. 7. Subjective user feedback for the first study

The feedback of the tetraplegic user aligns with the feedback of the other participants and is very positive in total, even verbally referring to the system as her "favourite one". As a very experienced user of assistive robotic systems, she did not express any concerns regarding safety or comfort, but instead listed the system as being rather too slow. Due to the user's restricted motion capabilities, abort motions to one side were not possible, but all others worked without any problems.

B. Drinking Process

The experiments for the drinking process were conducted with 16 participants (15 able-bodied and one tetraplegic). The average age was 26 ± 10 years. Out of 16 participants, seven (six able-bodied and one tetraplegic) were females. The users were allowed a brief acclimatization period followed by an attempt at one full successful run for the drinking process, terminating in a fall-back.

Fig. 8 shows the orientation of the end-effector in Euler angles (Fig. 8a) and the concurrent force sensor readings (Fig. 8b) during one of the trials of the drinking process. These values were collected as output values of the manipulator for the experiments of each user. Force_B (red curve) crossing the trigger line (yellow line) causes an increment in the roll angle (feed motion) and Force_F (blue curve) crossing the trigger line causes decrement in the roll angle (non-feed motion). Both sensors crossing the trigger line causes the smooth decrement in the roll angle, which is the fall-back. The initial increment in the roll angle seen before any force input is a pre-defined rotation for user convenience.

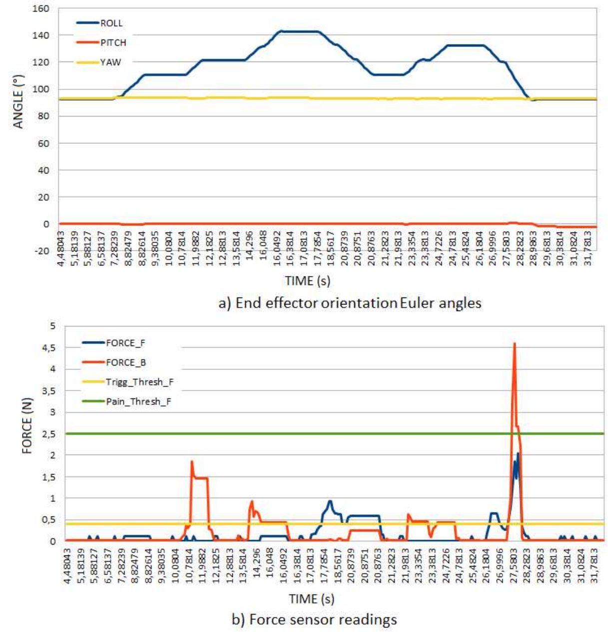

Fig. 8. Plots of feeding experiment of one of the trials

Since this is the second trial for the same user, the user is already familiar with the system, as evidenced by the fact that there are no overshoots above the pain threshold and that the user is able to manipulate the cup with only the minimum amount of force required.

Some of the feedback provided by the subjects is summarized in Fig. 9 as the mean feedback across the subjects. It can be safely concluded from the analyses of Fig. 8 and Fig. 9 that the force controller designed for the drinking process performs satisfactorily well. With regards to the drinking process, in alignment with the rest of the subjects, the tetraplegic user felt that the system was indeed comfortable and intuitive to use. She did have some critical feedback to provide, which forms the basis for some opportunities of further development:

- She preferred to use a normal cup without a beak
- She felt the need for a user-controlled emergency stop signal, such as a voice or eyes activated trigger.

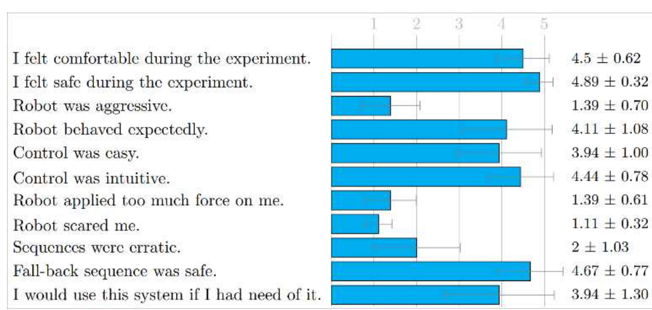

Fig. 9. User feedback

V. Conclusion and Future Work

The proposed solution functions as a user-friendly and safe concept of control for an assistive drinking task with the user in the loop. The mechanical setup requires no physical attachment on the user or in close proximity to them, thus allowing a flexible out-of-the-box usage with no preparations on the user's side.

The integration of all subsystems has been successful and first results could be obtained with potential users. The experiments show a high reliability of the safety-critical systems and quick responses on fall-back commands. The system has a high user acceptance and consistent positive feedback with an easy and intuitively perceived control scheme.

Future work will focus on replacing the current smart cup with a standard cup without the beak, as requested by the tetraplegic user, and increasing the overall velocity of the system in situations that are not safety-critical. To implement a system capable of executing the whole drinking operation, the autonomous filling and grasping of the cup will be necessary, as well as defining a comfortable and safe home position for the robot. The complete system will be tested within a larger user study with a higher ratio of potential end-users.

Acknowledgment

The research was supported by the German Federal Ministry of Education and Research (BMBF) as part of the project MobiLe (Physical Human-Robot-Interaction for Independent Living). The authors would like to thank all the persons involved, in particular the tetraplegic user, Mrs. Lena Kredel, who took part in both studies.

References

[1] W. Strubreither, M. Neikes, D. Stirnimann, J. Eisenhuth, B. Schulz und P. Lude, Klinische Psychologie bei Querschnittlähmung, Vienna: Springer, 2015.

[2] A. Graser, T. Heyer, L. Fotoohi, U. Lange, H. Kampe, B. Enjarini, S. Heyer, C. Fragkopoulos und D. Ristic-Durrant, „A supportive friend at work: Robotic workplace assistance for the disabled," *IEEE Robotics & Automation Magazine 20.4,* pp. 148-159, 2013.

[3] I. Volosyak, O. Ivlev and A. Graser, "Rehabilitation robot FRIEND II-the general concept and current implementation," *9th International Conference on Rehabilitation Robotics (ICORR),* pp. 540-544, 2005.

[4] C.-S. Chung, H. Wang und R. A. Cooper, „Functional assessment and performance evaluation for assistive robotic manipulators: Literature review," *The journal of spinal cord medicine. Vol 36. No. 4,* pp. 273-289, 2013.

[5] D.-J. Kim, R. Lovelett und A. Behal, „An empirical study with simulated adl tasks using a vision-guided assistive robot arm," in *IEEE International Conference on Rehabilitation Robotics (ICORR)*, 2009.

[6] S. Schröer, I. Killmann, B. Frank, M. Völker, L. Fiederer, T. Ball und W. Burgard, „An autonomous robotic assistant for drinking," *IEEE International Conference on Robotics and Automation (ICRA),* pp. 6482-6487, 2015.

[7] S. Alsharif, O. Kuzmicheva und A. Gräser, „Gaze Gesture-Based Human Robot Interface," *Technische Untersützungssysteme, die die Menschen wirklich wollen,* p. 339, 2016.

[8] B. Mindermann, *Untersuchung eines hybriden Brain-Computer Interfaces (BCIs) zur optimalen Auslegung als Mensch-Maschine-Schnittstelle,* University of Bremen, 2018.

[9] "KINOVA Gen2 Ultra lightweight robot," [Online]. Available: https://www.kinovarobotics.com/en/products/robotic-arms/kinova-gen2-ultra-lightweight-robot. [Accessed 1 Feb 2019].

[10] V. Maheu, P. S. Archambault, J. Frappier und F. Routhier, „Evaluation of the JACO robotic arm: Clinico-economic study for powered wheelchair users with upper-extremity disabilities," *IEEE International Conference on Rehabilitation Robotics (ICORR),* pp. 1-5, 2011.

[11] T. L. Koller, M. Kyrarini und A. Gräser, „Towards robotic drinking assistance: Low cost multi-sensor system to limit forces in Human-Robot-Interaction [accepted]," in *12th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA)*, Rhodos, Greece, 2019.

[12] Intel Corporation, "Intel RealSense Depth Camera D435," Intel, [Online]. Available: https://click.intel.com/intelr-realsensetm-depth-camera-d435.html. [Accessed 1 February 2019].

[13] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler und A. Y. Ng, „ROS: an open-source Robot Operating System," in *ICRA workshop on open source software. Vol. 3, No. 3.2*, Kobe, Japan, 2009, p. 5.

[14] D. E. King, „Dlib-ml: A Machine Learning Toolkit," *Journal of Machine Learning Research,* Bd. 10, pp. 1755-1758, 2009.

[15] G. B. Huang, M. Ramesh, T. Berg und E. Learned-Miller, *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments,* Amherst: University of Massachusetts, 2007.

[16] A. Lukezic, T. Vojir, L. C. Zajc, J. Matas und M. Kristan, „Discriminative Correlation Filter With Channel and Spatial Reliability," *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 6309-6318 , 2017.

[17] T. F. Cootes, C. J. Taylor, D. H. Cooper und J. Graham, „Active shape models-their training and application," *Computer vision and image understanding,* Bd. 61, Nr. 1, pp. 38-59, 1995.

[18] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou und M. Pantic, „300 faces in-the-wild challenge: Database and results," *Image and vision computing. Volume 47,* pp. 3-18, 2016.

[19] M. A. Fischler und R. C. Bolles, „Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Commun. ACM Vol. 24 No. 6,* pp. 381-395, Juni 1981.