This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Performance Analysis of Learning From Demonstration Approaches During a Fine Movement Generation

Aljaž Baumkircher, Marko Munih, *Member, IEEE*, and Matjaž Mihelj, *Member, IEEE*

***Abstract*—Learning from demonstration (LfD) is a well-established method of movement demonstration; however, the performance of different LfD approaches during a fine movement generation is still unknown. In this study, we compare kinesthetic teaching, teleoperation, and cooperative robot tool approaches on two different tasks, where a submillimeter accuracy is required. Additionally, we analyze the influence of a visual enhancement feature on each of the approaches and the influence of a spatial scaling feature on the teleoperation approach. The participants are a well-balanced group (regarding age, gender, and expertise), with 65% having no previous experience using robots. In our study, we found that all approaches achieved a submillimeter median positioning error. However, when no additional features are used, the cooperative robot tool (CRT) approach outperforms other approaches since it consistently achieves the lowest positioning error. Besides the positioning error, the generated velocity and the participants' feedback (via a questionnaire) also indicates that it is the most suitable approach for an accurate submillimeter movement generation. We also concludes that the visual enhancement feature and the spatial scaling feature has a significant influence on the performance of all approaches. When the two features are used, the generated positioning error drops considerably. When the visual enhancement feature is used, kinesthetic teaching performs in some cases as good as the CRT approach, while the teleoperation with the spatial scaling feature approach in some cases even outperforms the CRT approach. However, we still consider the CRT to be the best approach for fine movement generation since these features cannot be used in every possible scenario.**

***Index Terms*—Cooperative robot tool (CRT), kinesthetic teaching, learning from demonstration (LfD), submillimeter accuracy, teleoperation, visual enhancement.**

I. Introduction

Learning from demonstration (LfD) is a concept of robot task learning, where the task is learned through a human demonstration [1], [2]. The often-used methods to perform a demonstration are kinesthetic teaching and teleoperation. The two approaches allow users to teach robots in ways that are usually faster and more intuitive than the conventional robot programming [3].

Teleoperation is a method that requires an input (usually haptic) device, through which the robot movement is controlled [4]. Kinesthetic teaching, on the other hand, does not require an input device, since the robot movement is controlled directly by the user grabbing and moving individual segments [5], [6]. In comparison with teleoperation, kinesthetic teaching turns out to be faster and easier to use [3], [7]. The quality of demonstrations is similar between the two approaches, although the type of the input device does influence the performance of teleoperation [3].

Depending on the task, the demonstration can contain discrete keyframes (through which the robot motion is planned), a trajectory (that robot will follow), or a combination of the two [8]. While trajectory demonstrations are more intuitive for naïve users and allow complicated skills to be taught, it might be hard for users to manipulate a high degree-of-freedom (DoF) robot, while sustaining a smooth trajectory. Keyframe demonstrations, on the other hand, allow for a more generalized description of a task, which makes it easier to adapt the task to new situations [9].

However, the comparison of different LfD approaches has been made on tasks, composed of coarse movements. With these movements, the required accuracy of the demonstration is still small enough for users to successfully transfer the desired movement on a robot, using kinesthetic teaching. Some examples of coarse movements are stacking blocks on top of one another [8], scooping [7], placement of a circuit board [10], hitting a table tennis ball with a racket [11], and writing tasks [12],[13].

None of the studies so far focused on the comparison of LfD approaches for fine movements. Compared to coarse movements, the fine movements require submillimeter accuracy. These types of movements are often required in the field of surgical robotics [14],[15]. Teleoperation [4],[16]–[18] and cooperative robot tools [19],[20] are the two established approaches for the demonstration of fine movements. In the case of cooperative robot tools (CRTs), the user manipulates the mechanism's end effector. The mechanism itself is a non-backdrivable system, which is controlled with an admittance control scheme. The movement of the mechanism is defined by the forces with which the user manipulates the end effector. This approach allows for a fine movement generation since it reduces the influence of the

Manuscript received December 9, 2020; revised July 13, 2021; accepted August 10, 2021. This work was supported by Slovenian Research Agency (research core funding No. P2-0228). This article was recommended by Associate Editor J. Yang. *(Corresponding author: Aljaž Baumkircher.)*

The authors are with the Laboratory of Robotics, Faculty of Electrical Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia (e-mail: aljaz.baumkircher@gmail.com; marko.munih@robo.fe.uni-lj.si; matjaz.mihelj@robo.fe.uni-lj.si).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/THMS.2021.3107523.

Digital Object Identifier 10.1109/THMS.2021.3107523

user's tremor and scales the input forces to accordingly small movements [21], [22].

Despite many benefits of kinesthetic teaching, it is still not clear how well can a user demonstrate a fine movement. In this study, we compared kinesthetic teaching, with two well-established LfD approaches that are often used during the fine movement demonstration (e.g., teleoperation and CRT). The demonstrations were done with Franka Emika Panda 7-DoF robot, for which we purposely designed an end-effector tip that enabled the user to assume robot's position with submillimeter accuracy. We compared the LfD approaches on two different tasks. The keyframe demonstration task required the user to move the robot's end effector from the start toward the target position and place the end effector within a 0.5-mm radius of the target position. The trajectory demonstration task, on the other hand, required the user to move the robot's end effector along the target, 0.5-mm wide, trajectory from the beginning until the end. With the first task, we were able to test the ability to precisely position the robot's end effector, while with the second, we were able to test the ability to precisely generate the desired movement. For both tasks, the visual references (e.g., target position and target trajectory) were displayed on a horizontally placed LCD monitor. As part of our research, we also tested the influence of visual enhancement and spatial scaling on the user's accuracy. We designed the visual enhancement feature, in order to enable the user to magnify the area under the robot's end effector, and in turn, increase the user's visual sensitivity. The spatial scaling feature was used as part of the teleoperation approach and allowed for a variable spatial scaling factor between the user's and the robot's movement. This, in turn, influenced the robot's sensitivity to the user's generated movement. Both the experimental setup and the tasks are described in detail in Sections II-A and II-B, respectively. In order to improve the accuracy of the robot manipulator, we performed a kinematic model calibration. The kinematic model calibration is described in detail in chapter I of the supplementary file. Besides the objective measures (e.g., positioning error, movement velocity, movement smoothness, etc.), we took into consideration also some subjective measures (e.g., ease of use, sense of accuracy, and room for improvement) by letting users fill out a questionnaire after the end of the experiment. This enabled us a broader discussion comparing different LfD approaches. The questionnaire is described in detail in Section II-F.

II. Materials and Methods

A. Experimental Setup

The experimental setup consisted of robot manipulator Panda (Franka Emika, GmbH), haptic interface Phantom Premium 1.5 (3D Systems, Inc.), force sensor Nano17 (ATI Industrial Automation, Inc.), a 1-DoF joystick, and a horizontally placed LCD monitor (see Fig. 1, top left). The Panda robot was used as a manipulator, while also serving as an input device during the kinesthetic teaching. For the teleoperation and the CRT approach, the haptic interface Phantom and the force sensor Nano17, respectively, served as input devices. During the teleoperation approach, users also used the 1-DoF joystick as an input device, with which they controlled the spatial scaling between the input motion (Phantom), and the generated motion (Panda). The horizontally placed LCD monitor served as a user interface, as it rendered the target positions and trajectory (depending on the task), indicated the current spatial scaling value, while also enabling the visual enhancement feature.

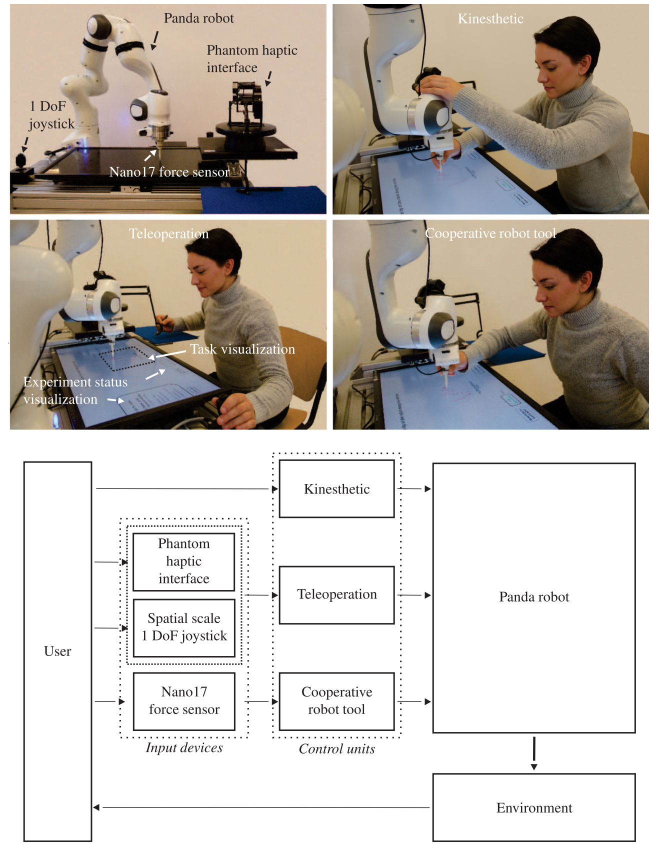

Fig. 1. (Top) Experimental setup overview. (Bottom) Corresponding schematic overview.

For a better understanding, we grouped the aforementioned devices into different functional units such as Panda robot control unit, teleoperation control unit, CRT control unit, and user interface unit (see Fig. 1, right). For these units to work as a whole, a high-level control was created using MATLAB Simulink 2019b (The MathWorks, Inc.). Besides connecting the individual units, it was also used for high-level decision making (e.g., switching the control modes) and saving the measured data. Recorded data contained all relevant sensor information from the Panda robot (end-effector position and orientation, joint positions, joint velocities, etc.), Phantom haptic interface (end-effector position and velocity), Nano17 force sensor (3-D force and torque vectors), and 1-DoF joystick (joystick angle). The data from each sensor were acquired with 1-kHz sampling frequency and sent to MATLAB in real time using UDP communication. Each functional unit will be presented in the following paragraphs.

1) Robot Manipulator Panda: The experiments were done by manipulating 7-DoF collaborative robot Panda. Its controller was connected via Ethernet cable to a PC (Ubuntu 16.04 with a real-time kernel), which was used to execute a low-level control program taking care of the robot's motion planning. Motion planning depended on the demonstration type. During kinesthetic teaching, the generated motion was based on the robot's compensated dynamical model. In contrast, during teleoperation and CRT, the generated motion was based on the reference positions calculated by the teleoperation or CRT control units.

We designed a custom end-effector with a narrow tip that enabled the user to assume end-effector's position with submillimeter accuracy. The Nano17 force sensor was also integrated as part of the end effector since that allowed an optimal CRT experience (see Fig. 1, top left). Throughout the experiment, the robot's end-effector orientation was constant, and the movement was limited to a horizontal plane. The constraints were done in order to simplify the comparison between different demonstration types. The frequency of the robot's control loop was 1 kHz.

A crucial issue that also had to be overcome was the robot's kinematic model accuracy. Due to tolerances in mechanical construction and assembly, small errors in end-effector pose can occur. Errors distributed along the arm are amplified due to the robot's open chain kinematic structure. During testing, it turned out that end-effector position errors as large as 5 mm occurred when moving the robot to the same desired end-effector position under different joint configurations. Therefore, a calibration of the kinematic model was required. In our case, after the calibration, the average error of the robot's end-effector position was reduced to 0.1 mm. The calibration procedure is described in detail in chapter I of the supplementary file.

2) Teleoperation Control Unit: For the teleoperation control, we used Phantom Premium 1.5 haptic interface, a backdrivable 3-DoF robot manipulator, as an input device. It was controlled using an industrial PC and MATLAB Simulink Real Time. The control sampling frequency was 5 kHz.

We controlled the Phantom's position on a horizontal plane in order to limit the user's movement in the vertical direction. We also introduced a spatial scaling factor $K_S \in (0,1)$ in order to set the spatial scaling between the user's input and Panda robot's generated motion. It was set by the user, using a 1-DoF joystick. The K_S value equaled 0 when the joystick was in the neutral pose and increased toward 1 with an increased offset from the joystick's neutral pose.

The desired Panda robot's position $\mathbf{x_R}(t)$ was, therefore, defined as

$$\mathbf{x_R}(t) = K_S(t) \int \dot{\mathbf{x}}_\mathbf{P}(t)dt \tag{1}$$

where $\dot{\mathbf{x}}_\mathbf{P}(t)$ represented Phantom robot's movement velocity, and $K_S(t)$ the spatial scaling factor value.

3) CRT Control Unit: For the Panda robot to function as a CRT, precise force measurements were required. We used Nano17 high sensitivity force and torque sensor. Its values were sampled using a DAQ interface, connected to the aforementioned industrial PC.

We defined the Panda robot's desired position as

$$\mathbf{x_R}(t) = \int \left(\int \frac{\mathbf{F_S} - \mathbf{b}\dot{\mathbf{x}}_\mathbf{R}}{\mathbf{m}} dt \right) dt \tag{2}$$

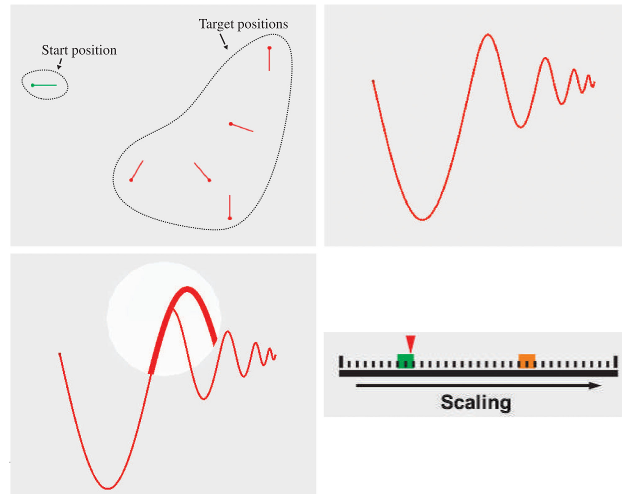

Fig. 2. Depiction of the user interface—The keyframe (top left) and the trajectory (top right) task, the visual enhancement (bottom left), and the teleoperation spatial scaling indicator (bottom right).

with $\mathbf{F_S}$ representing measured forces, while $\mathbf{m}$ and $\mathbf{b}$ parameters represent inertia and damping of the admittance control [12], [13]. We defined $\mathbf{m}$ and $\mathbf{b}$ as a 3-D vector, since we set the inertia and damping for each direction in the cartesian space. By modifying the parameter values, we influenced the dynamics between the measured forces and the reference trajectory in order to optimize the user's control of the Panda robot when using a force sensor as an input device. Since we wanted Panda robot's movement to be constrained to a horizontal plane, we set $m_Z = \infty$.

4) User Interface Unit: The user interface unit was used to project the required task to the user, and to show relevant information regarding the ongoing experiment.

Besides displaying reference keyframes and trajectories (described in Section II-B), a visual enhancement feature was also implemented (see Fig. 2, right). Its purpose was to visually enlarge the area under Panda robot's end effector (3x magnification) in order to increase user's positioning error detection, and therefore, influence the accuracy of a demonstration. It was created using the Unity game engine (Unity Software Inc.). Another essential feature of the user interface unit was teleoperation spatial scale indicator (see Fig. 2, bottom). It was used to inform the user about the current K_S. The indicator had a green and an orange zone. The green zone indicated the desired K_S value range ($K_S = 0.25 \pm 0.05$), when greater accuracy was required, while the orange zone represented the appropriate value range ($K_S = 0.70 \pm 0.05$) when larger movements were desired. The environment was projected on a horizontally placed 28-in LCD monitor.

B. Tasks

As part of our experiment, we defined two different tasks on which to compare the LfD approaches. Following Akgun *et al.* [8], we designed a keyframe and a trajectory demonstration task. Both tasks will be explained in detail in the following paragraphs.

1) Keyframe Demonstration: Keyframe demonstrations are useful when it is necessary to set start, end, and intermediate positions to define a task, with the generated path between the set positions being irrelevant. In some cases, the task requires a rather precise definition of a keyframe position. An example from a microbiological environment is a bacterial colony picking task, where the colonies with an approximately 1 mm in diameter have to be accurately picked.

Therefore, we designed a task to determine whether it is possible to position the robot's end effector with sufficient accuracy using different LfD approaches. The target position was defined as a circle with a 1-mm diameter (see Fig. 2, left). The user was required to move the robot's end effector from the start position and position it over the target position. Besides the circle, each target position also contained a line that indicated the desired direction from which the user should approach the area. Since the users tried to follow the line, we could not only analyze whether the user reached the target position, but also analyze how they approached the target. During the experiment, the target position was projected on five different positions, with each having a different orientation. Multiple goal destinations, in turn, allowed us a more generalized analysis.

2) Trajectory Demonstration: Trajectory demonstrations are useful when a specific movement has to be demonstrated. In this task, we tested the influence of the target trajectory's spatial dimensions on the accuracy of a demonstration. We designed a reference trajectory with gradually smaller spatial dimensions (see Fig. 2, middle). It was based on a sine wave, whose amplitude decreases with each period

$$y_i = A_i \sin(\omega_i x_i), \quad \mathbf{A} = \begin{pmatrix} 0.06 \\ 0.02 \\ 0.01 \\ 0.005 \\ 0.0025 \end{pmatrix} m. \tag{3}$$

In order to ensure a smooth trajectory, the ω_i had to be scaled accordingly

$$\omega_i = \frac{A_{i-1}}{A_i}\omega_{i-1}, \quad \omega_0 = 0.01256. \tag{4}$$

During this task, the user had to move the robot's end effector along the target trajectory, while trying to follow the line as accurately as possible. A gradual decrease in amplitude allowed for a smooth transition from coarse to fine movements. The target trajectory represented a fine movement during the last two periods (A_4, A_5). We will refer to parts of the trajectory with the same A_i as trajectory segments (i.e., trajectory segment A_3 represents the part of trajectory where $A_i = A_3$).

C. Experiment Protocol

The tasks were performed by 31 participants (20 male, 11 female; 11 robotic experts, 20 robotic non-experts; average age 32 ± 9.8 years). Before the experiment started, each had an approximately 20-min training session, testing all LfD approaches, visual enhancement feature, and spatial scaling factor influence. Preliminary studies showed that the amount of training needed to minimize the learning effect during the experiment varied from participant to participant based on their motoric skills and previous experience. Therefore, enough time was given to each participant in order for them to feel comfortable with controls. A set of small tests helped us indicate whether participants are still improving or started to stagnate, and are therefore, ready to start with the experiment. A training session lasted approximately 20 min.

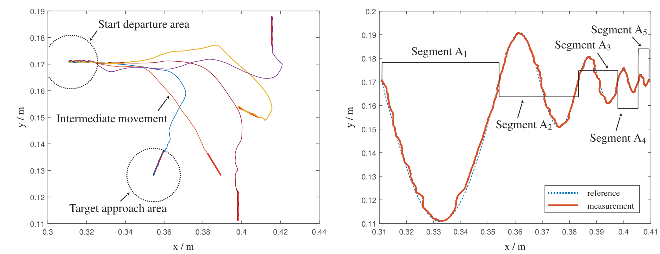

Fig. 3. Examples of a demonstration for each (left) keyframe task and (right) trajectory task.

After the training, the experiment started with the keyframe demonstration task. For each of the LfD approaches (kinesthetic, teleoperation, and CRT), the user moved from the start position to each of the five target positions. This was repeated four times with the visual enhancement turned OFF and four times with the visual enhancement turned ON. For the trajectory demonstration task, the user had to follow the reference trajectory for each of the LfD approaches. This was repeated four times with the visual enhancement turned OFF and four times with the visual enhancement turned ON. To test the influence of spatial scaling during teleoperation, we modified the aforementioned protocol. Users still performed four repetitions under each condition. However, three of these repetitions were done using a variable spatial scale, and one was done with a fixed spatial scale (K_S(t) = 1). In the end, the user had to fill out a questionnaire, which consisted of a seven-point Likert-scale evaluation for each of the LfD approaches, and two open-ended questions. The questionnaire is described in detail in Section II-F.

Each participant required 2–3 hours for the whole experiment.

D. Data Analysis

1) Keyframe Demonstration Task: The goal of the keyframe demonstration was to analyze the positioning error during the departure from the start keyframe, and approach to the target keyframe. In our study, we acquired 3720 keyframe demonstrations.

When the robot was positioned within a 1-cm radius of the start or target keyframe (see Fig. 3, left), the positioning error was determined as the smallest distance from the robot's end effector to the keyframe approach/departure line.

Movement velocity and smoothness were also analyzed. The velocity depended on the movement's trajectory path and the time required to perform the movement. Movement smoothness was determined using a well-established spectrum arc length (SAL) method [23]. The smoothness is based on the movement's velocity Fourier magnitude spectrum and its spectral arc-length since the arc-length decreases with the increased motion smoothness. The author defined SAL as a negative value of the arc-length. Therefore, a movement is smoother if it has a less negative SAL value. Movement velocity and smoothness were analyzed for the intermediate movement (coarse movement between the start and target keyframe) and the target keyframe's fine approach.

2) Trajectory Demonstration Task: The goal of the trajectory demonstration was to determine the accuracy with which the user can follow the reference trajectory. In our study, we acquired 744 trajectory demonstrations. In order to compare measurements with the reference, a time-alignment preprocessing was necessary. For this purpose, we used a well-established method dynamic time warping [24]. The method determines the optimal match between two given sequences by nonlinearly transforming the time dimension of each sequence.

The time alignment allowed us to compare each demonstrated trajectory with the reference and calculate the positioning error that occurred along the demonstration. The positioning error was determined as the Euclidean distance between the robot's position and the reference position. The errors were calculated along the whole trajectory and grouped based on the trajectory segment they were a part of (see Fig. 3, right).

Movement velocity and smoothness analysis were also carried out. They were calculated for each of the trajectory's segments. The movement velocity depended on the movement trajectory's path, and the time required to get from one segment to the other. The movement smoothness was again determined using the SAL method [23], additionally explained also in Section II-D.1.

E. Statistical Analysis

A correct statistical method had to be chosen, to determine if there is a statistically significant difference between demonstrations under different conditions (LfD method, use of visual enhancement, and use of spatial scaling). Since the same group of participants executed the experiments under different conditions, a paired statistical test is appropriate. Due to repetitive violation of sphericity [25] and occurrence of outliers in the dataset, we decided nonparametric tests should be used [26],[27]. A two-tailed Wilcoxon signed-rank test [28] was used to compare two conditions, whereas Friedman's test [29] was conducted when more than two conditions were compared simultaneously. If the null hypothesis of Friedman's test was rejected, a two-tailed Wilcoxon signed-rank test was used for multiple comparisons. The significance threshold was set to 0.05. In order to prevent Type-I error inflation due to multiple comparisons, we used the Bonferroni correction [30]. The results of the statistical significance are presented in chapter II of the supplementary file.

F. Questionnaire

The questionnaire was designed with a purpose to get additional insight into each of the LfD approaches, visual enhancement feature, and variable spatial scaling. The users were asked to grade each task on a seven-point Likert scale based on the ease of use, perceived accuracy, and possible room for improvement. Ease of use represented the difficulty with which the user accurately positioned the robot's end effector, with 1 representing "highest difficulty" and 7 representing "lowest difficulty." Perceived accuracy represented the user's opinion about the achieved accuracy using each of the LfD approaches, with 1 representing "lowest accuracy" and 7 representing "highest accuracy." Room for improvement represented the user's opinion about the possible improvement with a certain LfD approach if they would have more time to train, with 1 representing "no improvement" and 7 representing "a lot of improvement."

Besides the seven-point Likert scale, we also asked two open-ended questions. The first question stated, "Which of the LfD approaches would you choose if you would have to accurately demonstrate a task and why?" The second question stated, "Did you find visual enhancement useful? Please elaborate the answer." The answers to these two questions gave us additional insight into each of the LfD approaches and the visual enhancement feature.

III. Results

In the following chapter, we present the results, which are relevant for the discussion. The results are presented with box and whiskers plots, and in order to describe them we used median (Med), 25th Percentile (Q1), 75th Percentile (Q3), minimum (Min), and maximum (Max) values. In order to improve the readability of the following chapter, authors omitted detailed description for some experiment condition results. A comprehensive overview of results is provided in chapter III of the supplementary file. All comparisons were statistically evaluated. The results of the statistical analysis are presented in chapter II of the supplementary file.

During the data analysis, we determined that the results of each LfD approach did not significantly differ depending on the target keyframe (keyframe demonstration task). Therefore, we combined the results of all target keyframe demonstrations performed under the same experimental conditions (e.g., LfD approach and use of features). We focused on differences in performance between different LfD approaches. For both tasks, we also found no significant difference in demonstrations depending on the gender (male/female) or expertise (robotic expert/non-expert). This allowed us to combine all participants in one single statistical population. Sufficient and flexible training procedure and the fact that the experiment required from participants good motoric skills instead of extensive experience with robots could be two reasons behind finding no significant difference in participants based on their expertise.

Moreover, for the trajectory task, we determined that the results from segments A_1 and A_2 did not deviate from what we observed in segments A_3, A_4, and A_5. Therefore, we omitted the results from the first two segments in order to focus on segments that required motion generation of smaller spatial dimensions.

A. Keyframe Demonstration Task

In Fig. 4, we use box and whiskers plots to present the positioning error users made during the departure from the start keyframe (white background) and during the target keyframe approach (gray background). The figure depicts the performance of each LfD approach with and without the influence of the visual enhancement feature and the spatial scaling feature (teleoperation only).

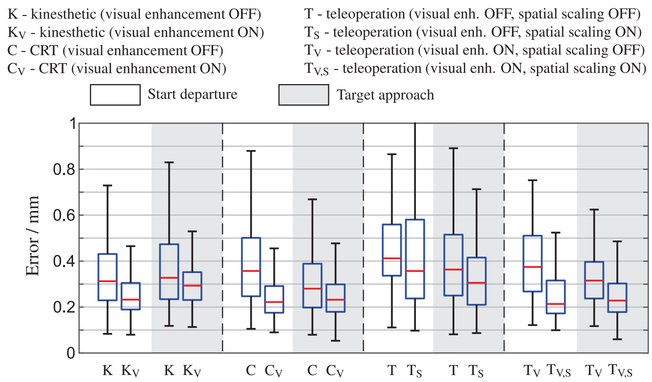

Fig. 4. Positioning error during the start keyframe departure and the target keyframe approach.

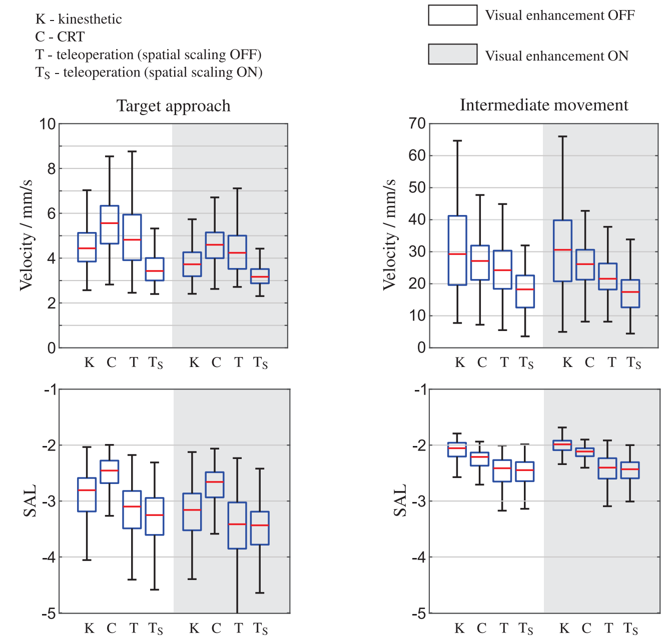

Fig. 5. Movement velocity and smoothness during the intermediate movement and the target approach movement.

When the visual enhancement feature was not used during the departure from the start position, kinesthetic teaching (Med = 0.33 mm; Q1 = 0.23 mm; Q3 = 0.47 mm; Min = 0.12 mm; and Max = 1.43 mm) performed similarly as CRT and teleoperation with the spatial scaling feature, providing no statistically significant difference. Teleoperation without the spatial scaling feature performed worse than the kinesthetic teaching and CRT. However, no statistically significant difference was found between the two teleoperation approaches. During the target approach, the CRT generated the lowest positioning error (Med = 0.28 mm; Q1 = 0.19 mm; Q3 = 0.39 mm; Min = 0.07 mm; and Max = 0.87 mm), followed by kinesthetic teaching, teleoperation with the spatial scaling feature, and teleoperation without the spatial scaling feature.

When the visual enhancement feature was used during the departure from the start position, CRT (Med = 0.22 mm; Q1 = 0.17 mm; Q3 = 0.29 mm; Min = 0.09 mm; and Max = 0.79 mm) and teleoperation with the spatial scaling feature performed similarly, providing no statistically significant difference. These were followed by kinesthetic teaching and teleoperation without the spatial scaling feature. During the target approach, kinesthetic teaching (Med = 0.23 mm; Q1 = 0.19 mm; Q3 = 0.31 mm; Min = 0.08 mm; and Max = 0.84 mm), CRT, and teleoperation with the spatial scaling feature performed similarly, providing no statistically significant difference. Teleoperation without the spatial scaling feature generated a higher positioning error compared to other approaches.

In all cases, the visual enhancement feature provided statistically significant improvement in the positioning error.

In Fig. 5, we use box and whiskers plots to present the movement velocity, and smoothness for each of the LfD approaches. The results were plotted separately for the intermediate movement and the target keyframe approach movement. Furthermore, the figure also depicts the influence of the visual enhancement feature and the spatial scaling feature (teleoperation only).

When the visual enhancement feature was not used during the target keyframe approach movement, the CRT (Med = 5.6 mm/s; Q1 = 4.7 mm/s; Q3 = 6.3 mm/s; Min = 2.8 mm/s; and Max = 9.7 mm/s) proved to be the fastest, followed by teleoperation without the spatial scaling feature, kinesthetic teaching, and teleoperation with the spatial scaling feature. Regarding the movement smoothness, the CRT (Med = −2.46; Q1 = −2.69; Q3 = −2.28; Min = −4.49; and Max = −2.00) generated the smoothest movement, followed by kinesthetic teaching, teleoperation without the spatial scaling feature, and teleoperation with the spatial scaling feature. During the intermediate movement, kinesthetic teaching (Med = 29.3 mm/s; Q1 = 19.6 mm/s; Q3 = 41.2 mm/s; Min = 7.8 mm/s; Max = 64.7 mm/s) turned out to be the fastest, followed by CRT and teleoperation without the spatial scaling feature, with no statistically significant difference between the two. Teleoperation with the spatial scaling feature generated the slowest movement. Regarding the movement smoothness, kinesthetic teaching (Med = −2.05; Q1 = −2.20; Q3 = −1.96; Min = −3.43; Max = −1.79) generated the smoothest movement, followed by the CRT, teleoperation without the spatial scaling feature, and teleoperation with the spatial scaling feature. There was no statistically significant difference between the two teleoperation approaches.

When the visual enhancement feature was used during the target keyframe approach movement, the CRT (Med = 4.6 mm/s; Q1 = 4.0 mm/s; Q3 = 5.1 mm/s; Min = 2.6 mm/s; and Max = 7.7 mm/s) proved to be the fastest, followed by teleoperation without the spatial scaling feature, kinesthetic teaching, and teleoperation with the spatial scaling feature. Regarding the movement smoothness, the CRT (Med = −2.66; Q1 = −2.93; Q3 = −2.48; Min = −4.07; Max = −2.06) generated the smoothes movement, followed by kinesthetic teaching, teleoperation without the spatial scaling feature, and teleoperation with the spatial scaling feature, where the latter two provided no statistically significant difference. During the intermediate movement, kinesthetic teaching (Med = 30.6 mm/s; Q1 = 20.7 mm/s; Q3 = 39.8 mm/s; Min = 4.9 mm/s; and Max = 70.2 mm/s) proved to be the fastest, followed by the CRT, teleoperation without the spatial scaling feature, and teleoperation with the

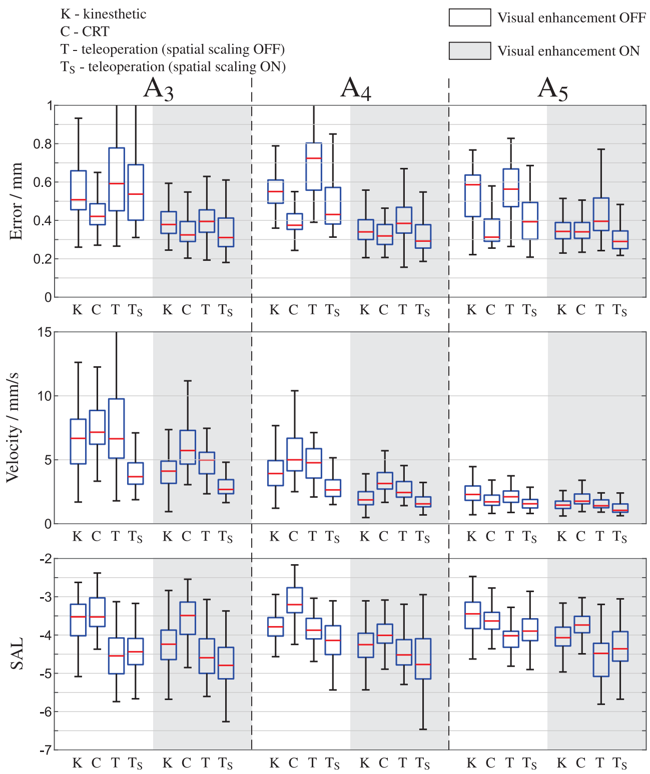

Fig. 6. Overview of positioning error, movement velocity, and smoothness along different trajectory segments.

spatial scaling feature. Regarding the movement smoothness, kinesthetic teaching (Med = −1.98; Q1 = −2.09; Q3 = −1.92; Min = −2.78; and Max = −1.68) produced the smoothest trajectory, followed by CRT, teleoperation without the spatial scaling feature, and teleoperation with the spatial scaling feature, where the latter two provided no statistically significant difference.

The visual enhancement feature significantly reduced the velocity of the target approach movement and reduced the movement smoothness. However, it had a minor effect during the intermediate movement, with no statistically significant change in velocity for kinesthetic teaching and teleoperation without the spatial scaling feature.

B. Trajectory Demonstration Task

In Fig. 6, we use box and whiskers plots to present the demonstrated movements positioning error, velocity, and smoothness for each of the LfD approaches. The results are plotted separately for each of the last three trajectory segments (A_3, A_4, and A_5). Furthermore, the figure also depicts the influence of the visual enhancement feature and the spatial scaling feature (teleoperation only).

When the visual enhancement feature was not used, CRT (Med = 0.42 mm; Q1 = 0.38 mm; Q3 = 0.49 mm; Min = 0.27 mm; and Max = 0.75 mm) approach generated the lowest positioning error in the segment A_3. It was followed by kinesthetic teaching, teleoperation with the spatial scaling feature, and teleoperation without the spatial scaling feature. However, there was no statistically significant difference between kinesthetic teaching and teleoperation with the spatial scaling feature. There was also no statistically significant difference between the two teleoperation approaches. The generated movement velocity was the highest with CRT (Med = 7.2 mm/s; Q1 = 6.2 mm/s; Q3 = 8.9 mm/s; Min = 3.3 mm/s; Max = 17.8 mm/s), followed by teleoperation without the spatial scaling feature, kinesthetic teaching, and teleoperation with the spatial scaling feature. There were no statistically significant differences between teleoperation without the spatial scaling feature and kinesthetic teaching or CRT. Regarding the movement smoothness, CRT (Med = −3.53; Q1 = −3.78; Q3 = −3.03; Min = −4.37; Max = −2.38) generated the smoothest trajectory, followed by kinesthetic teaching, teleoperation without the spatial scaling feature, and teleoperation with the spatial scaling feature, with no statistically significant difference between the latter two. In segment A_4, the CRT (Med = 0.37 mm; Q1 = 0.35 mm; Q3 = 0.44 mm; Min = 0.24 mm; Max = 0.75 mm) generated the lowest positioning error, followed by teleoperation with the spatial scaling feature, kinesthetic teaching, and teleoperation without the spatial scaling feature. The generated movement velocity was the highest with CRT (Med = 5.0 mm/s; Q1 = 4.1 mm/s; Q3 = 6.7 mm/s; Min = 2.5 mm/s; and Max = 12.7 mm/s), followed by kinesthetic teaching and teleoperation without spatial scaling feature, with no statistically significant difference between the two. The movement was the slowest using teleoperation with the spatial scaling feature. The generated movement was the smoothest with the CRT (Med = −3.21; Q1 = −3.42; Q3 = −2.76; Min = −4.24; and Max = −2.17), followed by kinesthetic teaching, and teleoperation without the spatial scaling feature, with no statistically significant difference between the two. Teleoperation with the spatial scaling feature generated the roughest movement. In segment A_5, CRT (Med = 0.31 mm; Q1 = 0.29 mm; Q3 = 0.41 mm; Min = 0.26 mm; and Max = 0.79 mm) achieved the lowest positioning error, followed by teleoperation with the spatial scaling feature, kinesthetic teaching, and teleoperation without the spatial scaling feature. The generated movement velocity was the highest with kinesthetic teaching (Med = 2.3 mm/s; Q1 = 1.8 mm/s; Q3 = 2.9 mm/s; Min = 0.7 mm/s; and Max = 6.2 mm/s), followed by teleoperation without the spatial scaling feature, CRT, and teleoperation with the spatial scaling feature. However, there was no statistically significant difference between CRT and teleoperation with the spatial scaling feature. Kinesthetic teaching (Med = −3.45; Q1 = −3.83; Q3 = −3.14; Min = −5.48; and Max = −2.47) also generated the smoothest movement, followed by the CRT, teleoperation with the spatial scaling feature, and teleoperation without the spatial scaling feature. There was no statistically significant difference between the two teleoperation approaches.

When the visual enhancement feature was used in the segment A_3, the CRT (Med = 0.32 mm; Q1 = 0.29 mm; Q3 = 0.39 mm; Min = 0.20 mm; and Max = 0.59 mm) and teleoperation with the spatial scaling features (Med = 0.31 mm; Q1 = 0.26 mm; Q3 = 0.41 mm; Min = 0.18 mm; and Max = 0.67 mm) achieved a better positioning error than kinesthetic teaching and teleoperation without spatial features. No statistically significant difference was observed between the CRT and teleoperation with the spatial scaling features. The generated movement velocity was the highest with the CRT (Med = 5.7 mm/s; Q1 = 4.7 mm/s; Q3 = 7.3 mm/s; Min = 3.1 mm/s; and Max = 15.3 mm/s) and teleoperation without the spatial scaling feature, followed by kinesthetic teaching and teleoperation with the spatial scaling feature. There was no statistically significant difference between CRT and teleoperation without the spatial scaling feature. Regarding the movement smoothness, CRT generated the smoothest movement (Med = −3.49; Q1 = −3.98; Q3 = −3.14; Min = −5.42; and Max = −2.54), followed by kinesthetic teaching, teleoperation without spatial scaling feature, and teleoperation with spatial scaling feature. However, there was no statistically significant difference between kinesthetic teaching and teleoperation without spatial scaling feature, and also no difference between both teleoperation approaches. In segment A_4, teleoperation with the spatial scaling feature (Med = 0.29 mm; Q1 = 0.26 mm; Q3 = 0.38 mm; Min = 0.19 mm; and Max = 0.58 mm), and the CRT achieved the lowest positioning error, followed by kinesthetic teaching and teleoperation without the spatial scaling feature. No statistically significant difference was observed between the CRT and teleoperation with the spatial scaling features. The highest movement velocity was generated by the CRT (Med = 3.1 mm/s; Q1 = 2.7 mm/s; Q3 = 4.0 mm/s; Min = 1.7 mm/s; and Max = 7.8 mm/s), followed by teleoperation without the spatial scaling feature, kinesthetic teaching, and teleoperation with the spatial scaling feature. Regarding the movement smoothness, the CRT (Med = −4.01; Q1 = −4.21; Q3 = −3.72; Min = −4.89; and Max = −3.09) generated the smoothest movement, followed by kinesthetic teaching, teleoperation without the spatial scaling feature, and teleoperation with the spatial scaling feature. However, there was no statistically significant difference between kinesthetic teaching and teleoperation without the spatial scaling feature. There was also no statistically significant difference between both teleoperation approaches. In the segment A_5, teleoperation with the spatial scaling feature (Med = 0.29 mm; Q1 = 0.25 mm; Q3 = 0.35 mm; Min = 0.22 mm; and Max = 0.62 mm) achieved the lowest positioning error, followed by the CRT, kinesthetic teaching, and teleoperation without the spatial scaling feature. There was no statistically significant difference between the CRT and kinesthetic teaching. The highest movement velocity was achieved by the CRT (Med = 1.8 mm/s; Q1 = 1.6 mm/s; Q3 = 2.3 mm/s; Min = 0.9 mm/s; and Max = 4.0 mm/s), followed by kinesthetic teaching, teleoperation without the spatial scaling feature, and teleoperation with the spatial scaling feature. However, there was no statistically significant difference between kinesthetic teaching and teleoperation without the spatial scaling feature. There was also no statistically significant difference between both teleoperation approaches. Regarding the movement smoothness, the CRT (Med = −3.74; Q1 = −3.94; Q3 = −3.51; Min = −4.74; and Max = −2.72) achieved the smoothest movement, followed by kinesthetic teaching, teleoperation with the spatial scaling feature, and teleoperation without the spatial scaling feature.

C. *Questionnaire*

In Fig. 7, we use box and whiskers plots to present the results of a questionnaire. Participants scored the LfD approaches (with and without the visual enhancement, and spatial scaling features) based on the ease of use, the achieved accuracy, and the possibility for improvement on a seven-point Likert scale.

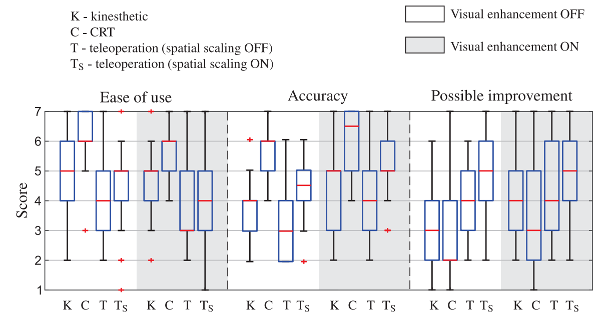

Fig. 7. Overview of subjective measures about the LfD approaches, and the influence of the visual enhancement feature and the spatial scaling feature. Red + signs represent data outliers (i.e., values that are more than 1.5 times the interquartile range away from the bottom or top of the box).

When the visual enhancement feature was not used, users found the CRT easiest to use, followed by kinesthetic teaching, teleoperation with the spatial scaling feature, and teleoperation without the spatial scaling feature. They presumed that they were the most accurate with the CRT, followed by teleoperation with the spatial scaling feature, kinesthetic teaching, and teleoperation without the spatial scaling feature. However, they had the most room for improvement with teleoperation with the spatial scaling feature approach, followed by teleoperation without the spatial scaling feature, kinesthetic teaching, and CRT.

When the visual enhancement feature was used, users found the CRT easiest to use, followed by kinesthetic teaching, teleoperation with the spatial scaling feature, and teleoperation without the spatial scaling feature. They presumed that they were the most accurate with the CRT, followed by teleoperation with the spatial scaling feature, kinesthetic teaching, and teleoperation without the spatial scaling feature. However, they had the most room for improvement with teleoperation with the spatial scaling feature approach, followed by teleoperation without the spatial scaling feature, kinesthetic teaching, and CRT.

The majority of users stated in answer to the first open-ended question that they would use the CRT approach if they would have to demonstrate a fine movement. They would use it because it allowed them to generate a smooth movement using insignificant input forces. Furthermore, the majority of users also stated in answer to the second open-ended question that they found the visual enhancement feature useful since it allowed them a greater positioning accuracy. However, a significant number of answers also stated that the visual enhancement feature slightly increased the complexity and psychological stress due to the increased positioning error detection.

IV. Discussion

In our study, we compared kinesthetic teaching, teleoperation, and CRT on two different tasks in order to determine their performance during fine movement generation. We implemented a visual enhancement feature, which was used with each approach, and a spatial scaling feature, which was used with the teleoperation approach only. Different approaches were compared based on the positioning error, movement velocity, movement smoothness (SAL), and the questionnaire results. Although the positioning error was the most important measure on which we compared the approaches, other measures also provided us with important and interesting insight. Movement velocity and smoothness, for example, provided us with information about the level of control participants had over the Panda robot, using different control approaches. Less negative values of SAL (i.e., a smoother movement), and higher levels of velocity, would indicate that it was easier for participants to achieve their desired positioning precision using the selected control approach.

During the keyframe demonstration task, it is evident from the positioning error distribution (see Fig. 4) that all LfD approaches generated a submillimeter median positioning error. When no features were used during the target approach, the CRT approach generated the lowest positioning error (Med = 0.28 mm; Q1 = 0.19 mm; Q3 = 0.39 mm; Min = 0.07 mm; and Max = 0.87 mm). If we compare the two teleoperation approaches, the generated positioning error was lower when the spatial scaling feature was used. This is due to a better motoric sensitivity of the robot's movement. Furthermore, it is evident that the visual enhancement feature significantly reduced the generated error for all LfD approaches. This was due to an enhancement, which improved the user's positioning error detection. However, it could also be that by increasing the error detection, the users became more cautious about the achieved accuracy, and consequently, increased their effort. The generated errors during the start of the demonstration were generally higher than the ones during the target approach. The reason could be that users perceived the start movement with lesser caution since it was followed by an intermediate movement, which required no positioning accuracy. With the use of the visual enhancement feature, however, the error dropped considerably, indicating that users took greater care at being accurate already from the start of the demonstration.

An advantage of kinesthetic teaching can be observed during the intermediate movement (see Fig. 5) since the generated movements had higher velocity compared to other approaches. This is because it allows a relatively easy generation of larger movements. It also generated the smoothest movement. Higher movement velocity generated larger momentum, which in turn decreased the inference of higher frequency movements, and consequently, improved the movement smoothness. However, during the target approach, the CRT achieved higher velocity and smoother movement than kinesthetic teaching. Therefore, when higher positioning accuracy was required, the kinesthetic teaching approach no longer excelled in the generated velocity. Teleoperation with the spatial scaling feature generated the slowest movement, which is due to the increased motoric sensitivity. Using the visual enhancement feature during the target approach reduced the generated movement velocity, due to users being more cautious about the achieved accuracy. However, visual enhancement had a minor effect on the velocity and smoothness of the intermediate movement, which was due to the lack of a movement reference. Therefore, nothing could be visually enhanced.

The difference in the generated positioning error is more evident during the trajectory demonstration task (see Fig. 6). Throughout the last three segments (A_3, A_4, and A_5), the CRT achieved the lowest positioning error. Kinesthetic teaching generated a submillimeter positioning error but performed significantly worse than the CRT. In comparison to the keyframe demonstration task, the difference is more evident because of the task's increased complexity. Each segment required a movement of smaller spatial dimensions, which was more challenging for users to generate. Consequently, the movement velocity dropped significantly from one segment to the next in order to achieve similar accuracy. Greater positioning accuracy was achieved by using the visual enhancement feature and the spatial scaling feature. Without the spatial scaling feature, teleoperation performed worse than kinesthetic teaching. However, when the spatial scaling feature was used, the positioning error dropped significantly. The increased motoric sensitivity allowed a more accurate demonstration throughout the whole trajectory. The improvement in teleoperation is even more evident when the visual enhancement feature was also used since it slightly outperforms the CRT with the visual enhancement feature. Kinesthetic teaching with the visual enhancement feature also achieved similar improvement, as it performed as good as CRT with the visual enhancement feature. It is interesting how the visual enhancement feature provided little benefit to the CRT. It could be that the users reached a certain plateau of accuracy, and therefore, the influence of the visual enhancement feature on the CRT is not as significant as in other approaches.

From a practical standpoint, the results achieved in this study are also on the limit of the possible achieved accuracy, due to external factors such as nonnegligible reference trajectory width, end-effector tip width, robot manipulator size, and the kinematic model errors.

Overall, the CRT achieved the highest levels of accuracy while allowing relatively fast and smooth movements. Its performance is also reflected from the questionnaire results (see Fig. 7) since it achieved the best score for ease of use and perceived accuracy. It was also the preferred LfD approach based on the first open-ended question since users found this approach to generate a smooth movement without a struggle, using insignificant input forces. Kinesthetic teaching, due to the imperfect dynamic model compensation, required more effort for accurate robot positioning, which is the reason why it received a lower score for ease of use and achieved a lower perceived accuracy. None of the teleoperation approaches scored high regarding the ease of use, which is due to the higher complexity of the approach. The influence of the spatial scaling feature was also noticeable since the teleoperation with the spatial scaling feature achieved a higher perceived accuracy score as the one without. Both teleoperation approaches also scored high regarding the possible room for improvement. The users thought that they could, through time, still improve the positioning accuracy for both teleoperation approaches. This was not the case for kinesthetic teaching and CRT. This is again due to the higher complexity of both teleoperation approaches. The majority of users also confirmed the usefulness of the visual enhancement feature. Based on the second open-ended question, the benefit of the increased positioning error detection outweighed the increased complexity to most users.

Among the participants, there was a similar male/female and expert/nonexpert representation. The overall performance was not gender dependant, but moreover, it was not qualification dependant. This means that these results can be achieved by an operator that has no previous experience in robotics.

V. Conclusion

In our study, we compared three different LfD approaches (kinesthetic teaching, teleoperation, and cooperative robot tool) on two different tasks (keyframe demonstration and trajectory demonstration) to determine whether these approaches are suitable for a movement generation where submillimeter accuracy is required. All approaches achieved a submillimeter median accuracy when no additional features were used. However, we conclude that the CRT is the most suitable approach to achieve submillimeter accuracy since it consistently achieved a median positioning error lower than 0.5 mm even when the visual enhancement feature was not used. Teleoperation and kinesthetic teaching achieve similar accuracy when additional features are used (visual enhancement and spatial scaling). The suitability of the CRT for fine movement demonstration is further confirmed by users, who graded the CRT the highest in terms of ease of use and achieved accuracy.

References

[1] H. Ravichandar, A. S. Polydoros, S. Chernova, and A. Billard, "Recent advances in robot learning from demonstration," *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 3, pp. 297–330, 2020.

[2] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robot. Auton. Syst.*, vol. 57, no. 5, pp. 469–483, 2009.

[3] K. Fischer *et al.*, "A comparison of types of robot control for programming by demonstration," in *Proc. 11th ACM/IEEE Int. Conf. Hum.-Robot Interaction*, 2016, pp. 213–220.

[4] C. Preusche, T. Ortmaier, and G. Hirzinger, "Teleoperation concepts in minimal invasive surgery," *Control Eng. Pract.*, vol. 10, no. 11, pp. 1245–1250, 2002.

[5] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, *Survey: Robot Programming by Demonstration*. Berlin, Germany: Springer, 2008.

[6] K. Kronander and A. Billard, "Learning compliant manipulation through kinesthetic and tactile human-robot interaction," *IEEE Trans. Haptics*, vol. 7, no. 3, pp. 367–380, Jul.–Sep. 2014.

[7] B. Akgun, K. Subramanian, and A. L. Thomaz, "Novel interaction strategies for learning from teleoperation," in *Proc. AAAI Fall Symp. Ser.*, 2012.

[8] B. Akgun, M. Cakmak, K. Jiang, and A. L. Thomaz, "Keyframe-based learning from demonstration," *Int. J. Social Robot.*, vol. 4, no. 4, pp. 343–355, 2012.

[9] C. Schou, J. S. Damgaard, S. Bøgh, and O. Madsen, "Human-robot interface for instructing industrial tasks using kinesthetic teaching," in *Proc. IEEE Int. Conf. Int. Saf. Robot.*, 2013, pp. 1–6.

[10] Y. Huang, L. Rozo, J. Silvério, and D. G. Caldwell, "Kernelized movement primitives," *Int. J. Robot. Res.*, vol. 38, no. 7, pp. 833–852, 2019.

[11] A. Paraschos, C. Daniel, J. Peters, and G. Neumann, "Using probabilistic movement primitives in robotics," *Auton. Robots*, vol. 42, no. 3, pp. 529–551, 2018.

[12] F. Steinmetz, A. Montebelli, and V. Kyrki, "Simultaneous kinesthetic teaching of positional and force requirements for sequential in-contact tasks," in *Proc. IEEE-RAS 15th Int. Conf. Humanoid Robots*, 2015, pp. 202–209.

[13] D. Berio, S. Calinon, and F. F. Leymarie, "Learning dynamic graffiti strokes with a compliant robot," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 3981–3986.

[14] A. E. Abdelaal, P. Mathur, and S. E. Salcudean, "Robotics in vivo: A perspective on human-robot interaction in surgical robotics," *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 3, pp. 221–242, 2020.

[15] A.N. Sridhar, T. P. Briggs, J. D. Kelly, and S. Nathan, "Training in robotic surgery-an overview," *Current Urol. Rep.*, vol. 18, no. 8, pp. 1–8, 2017.

[16] S. Calinon, D. Bruno, M. S. Malekzadeh, T. Nanayakkara, and D. G. Caldwell, "Human-robot skills transfer interfaces for a flexible surgical robot," *Comput. Methods Programs Biomed.*, vol. 116, no. 2, pp. 81–96, 2014.

[17] B. Huang, M. Ye, Y. Hu, A. Vandini, S.-L. Lee, and G.-Z. Yang, "A multirobot cooperation framework for sewing personalized stent grafts," *IEEE Trans. Ind. Inform.*, vol. 14, no. 4, pp. 1776–1785, Apr. 2018.

[18] J. Liang, J. Mahler, M. Laskey, P. Li, and K. Goldberg, "Using dVRK teleoperation to facilitate deep learning of automation tasks for an industrial robot," in *Proc. 13th IEEE Conf. Automat. Sci. Eng.*, 2017, pp. 1–8.

[19] J. J. Abbott, G. D. Hager, and A. M. Okamura, "Steady-hand teleoperation with virtual fixtures," in *Proc. 12th IEEE Int. Workshop Robot Human Interactive Commun.*, 2003, pp. 145–151.

[20] A. Bettini, P. Marayong, S. Lang, A. M. Okamura, and G. D. Hager, "Vision-assisted control for manipulation using virtual fixtures," *IEEE Trans. Robot.*, vol. 20, no. 6, pp. 953–966, Dec. 2004.

[21] R. Kumar, G. D. Hager, A. Barnes, P. Jensen, and R. H. Taylor, "An augmentation system for fine manipulation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2000, pp. 956–965.

[22] B. Mitchell *et al.*, "Development and application of a new steady-hand manipulator for retinal surgery," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2007, pp. 623–629.

[23] S. Balasubramanian, A. Melendez-Calderon, A. Roby-Brami, and E. Burdet, "On the analysis of movement smoothness," *J. Neuroeng. Rehabil.*, vol. 12, no. 1, pp. 1–11, 2015.

[24] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *Proc. KDD Workshop*, Seattle, WA, USA, 1994, vol. 10, no. 16, pp. 359–370.

[25] H. Huynh and L. S. Feldt, "Conditions under which mean square ratios in repeated measurements designs have exact f-distributions," *J. Amer. Stat. Assoc.*, vol. 65, no. 332, pp. 1582–1589, 1970.

[26] E. Lehmann and H. D'Abrera, *Nonparametrics: Statistical Methods Based on Ranks*. San Francisco, CA, USA: Holden-Day, Inc., 1975.

[27] L. A. Marascuilo and M. McSweeney, Nonparametric and Distribution-Free Methods for the Social Sciences. Monterey, CA, USA: Cole Publishing Company, 1977.

[28] R. C. Blair and J. J. Higgins, "Comparison of the power of the paired samples t test to that of Wilcoxon's signed-ranks test under various population shapes," *Psychol. Bull.*, vol. 97, no. 1, pp. 119–128, 1985.

[29] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *J. Amer. Statist. Assoc.*, vol. 32, no. 200, pp. 675–701, 1937.

[30] S. Lee and D. K. Lee, "What is the proper way to apply the multiple comparison test?," *Korean J. Anesthesiol.*, vol. 71, no. 5, pp. 353–360, 2018.