Kinestethic vs Imitation: Analysis of Usability and Workload of Programming by Demonstration Methods

Bruno Maric*, Filip Zoric*, Frano Petric and Matko Orsag

***Abstract*— Programming by demonstration (PbD) enables unskilled operators to easily demonstrate tasks and guide robots. In this paper we present comparison of demonstration methods with comprehensive user study. Each participant had to demonstrate drawing simple pattern using virtual marker and kinesthetic teaching with a cobot. We conducted user study with 24 participants which filled out NASA raw task load index (rTLX) and system usability scale (SUS). We also evaluated the quality of demonstration. We concluded study with finding that human demonstration using a virtual marker is on average 8 times faster, superior in terms of quality and imposes 2 times less overall workload than kinesthetic teaching.**

I. Introduction

The lack of intuitive and fast programming methods has been the main deterrent to application of manipulators in agile production lines, where products and services change daily, since programming robots requires the use of trained robotics engineers, who often lack the practical experience for the task at hand. Since the advent of cobots [1], Programming by Demonstration (PbD) [2] has gained widespread popularity in both academia and industry. PbD, or programming without coding [3], aims to simplify robot deployment and eliminate explicit task programming. Demonstration can be divided into kinesthetic teaching, teleoperation, and passive observation [4].

This paper focuses on PbD using a specially developed virtual marker and compares passive demonstration using the marker with kinesthetic teaching using robot manipulator via a comprehensive user study. We estimated imposed workload and system usability, whilst measuring demonstration duration and quality via a simple drawing task that involves reaching of several waypoints, path following between the waypoints and requires maintaining the contact with the surface at all times. Theses aspects of the task encompass a wide range of currently manual operations in the industry, in tasks such as welding, sanding, cutting, engraving etc. The task is designed after a popular puzzle given to children, ensuring that all participants are familiar with the task, as is expected in the envisioned industrial setting, and that the only novelty for a participant in the study is the method for demonstration of the task. *As the main contribution of the paper we present a comprehensive analysis of the conducted comparative study that explores workload and usability differences between kinesthetic teaching and human demonstration.*

II. Related work

Different methods, such as kinesthetic teaching [5] or teleoperation [6], can be used for the learning phase of PbD, followed by the representation phase with robot movement mapping and task execution. Various representation approaches are available, such as probabilistic models [7], and the popular Dynamic Movement Primitives (DMP) [8].

Wang et all. [9] proposed novel approach where the robot is able to recreate precise insertion task by passively observing human, with visual servoing enabling human hand tracking. Instead of visual servoing, in our work the motion capture system is used to track the virtual marker, which enables recording and further processing of recorded motion, introducing flexibility and allowing for demonstrations from different operators. Authors in [10] propose PbD system that besides kinesthetic teaching incorporates different modalities that humans use when naturally communicating some physical task or a mission, such as gaze and speech. As authors report, using multimodal PbD can lead to overtrust and automation bias in the long term, which is why it makes sense to explore HRI through PbD with different modalities of kinesthetic teaching and human demonstration, but also to include other modalities that are synchronized with virtual marker motion, such as force or human pose measurements. Building on the work presented in [11], our setup aims to capture both the position and force profiles simultaneously, for which we deploy a 6DOF F/T sensor alongside the virtual marker.

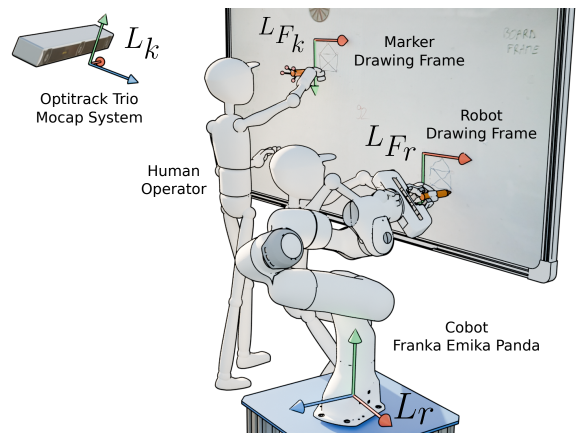

Fig. 1: The experimental setup for the user-experience study comprised a motion capture system, specially developed virtual markers, and the collaborative robot *Franka Emika Panda*. Participants were tasked to draw simple pattern using the virtual marker (*human on the left*) and kinesthetic teaching with cobot (*human on the right*).

III. Virtual Marker

This study focuses on the application of using virtual marker in PbD scenarios, since we strongly believe there are numerous different applications where kinestethic teaching

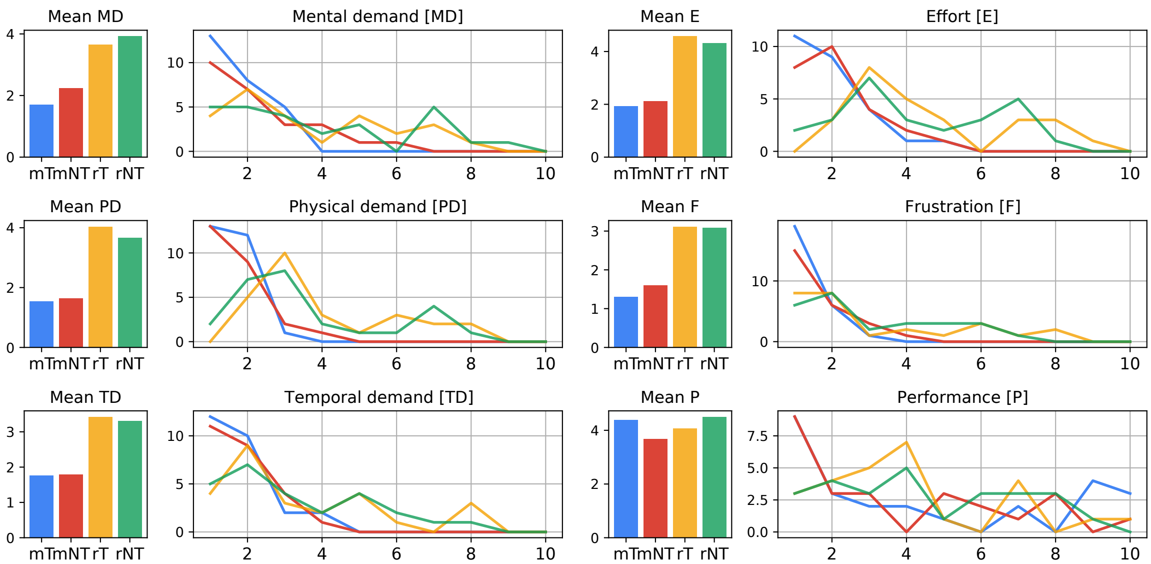

Fig. 2: NASA raw task load index (rTLX) questionnaire results. Bar charts show mean ratings of all participants. Line plots show frequency of certain rating for each task in the data acquired. For the workload measurement, *lower rating is better*. Graphs show that virtual marker introduces significantly less workload compared to the robot guidance.

fails to capture true essence of the demonstrated skill.

Optitrack mocap system tracks the global pose $\mathbf{T}_W^I$ of the fiducials placed on the virtual marker in world frame. The specifically developed calibration process is used to derive the relative transformation between the virtual marker tip and the Optitrack fiducials, represented as $\mathbf{T}_I^P$, resulting in the global pose of the virtual marker tip $\mathbf{T}_W^P = \mathbf{T}_W^I \cdot \mathbf{T}_I^P$.

The virtual marker measures and marks points in the global reference frame with submillimeter accuracy, which enables the marker to be used in various applications, especially those that require precise end-effector positioning over extended period of time, such as welding, sanding and drilling. The recordings obtained with marker can be synchronized and augmented with other important modalities for a given task, such as force/torque measurements.

IV. Experiment Methodology

Experimental setup (Fig. 1) consists of the virtual marker tool, the *Optitrack* mocap system, a collaborative manipulator *Franka Emika Panda*, force/torque sensor and a whiteboard. This setup allows us to concurrently track the contact force along with tracing the virtual marker.

To assess the ergonomics and user-friendliness of the virtual marker in PbD tasks, we conducted a user study. Participants were tasked with demonstrating drawings using both the virtual marker and by guiding the flange of the collaborative manipulator *Franka Emika Panda*. Our primary objective was to validate the potential of this Human-Robot-Interface (HRI) to enhance the adoption of robot manipulators, even among unskilled robot operators. To simulate demonstrations of tasks where the effect of the tool can be observed during demonstration and those where such demonstration is not feasible (safety, delicacy of the part etc), we used two markers of same dimensions versions: one that leaves a trace on the whiteboard, and the other one that does not.

1) Drawing frames: To standardize user drawings created with both the virtual marker and the robot, and facilitate uniform post-analysis, the drawing frames were marked on the whiteboard, and a 3D printed template was aligned within the frame to mark five waypoints for the drawing. This ensured consistency and comparability across all participant drawings. The virtual marker drawing frame is labeled as L_{F_k}, while the robot drawing frame is denoted as L_{F_r}. The drawing frames were indexed using both the virtual marker tip $\mathbf{T}_W^P$ and the robot tip $\mathbf{T}_B^P$, enabling transformation of points between frames.

2) Task and Population samples: To show difference between demonstration methods for the same task, we divided participants in 8 different groups to mitigate influence of knowing tasks beforehand. Each participant received instructions on how to use marker and the robot, and then had to complete the drawing in four different ways: A) virtual marker that leaves trace, B) virtual marker that leaves no trace, C) trace-leaving marker on a cobot, D) traceless marker on a cobot. There were 24 study participants, divided into 8 groups, and each group executed demonstration in different ordering (ABCD, ABDC, BACD, BADC, CDAB, DCAB, CDBA, DCBA).

V. Experimental results

After each of 4 tasks performed, participant was prompted to fill out NASA raw Task Load Index (rTLX) [12] and the system usability scale (SUS) test [13].Participants had to estimate each of the following workload categories: physical demand (PD), mental demand (MD), frustration (F), performance (P), temporal demand (TD) and effort (E). Compared to the original NASA TLX we have reduced rating scale from 0-21 to 0-10 because it is simpler for untrained users to populate such questionnaire. In the Fig. 2 it can be seen that almost all of the workload

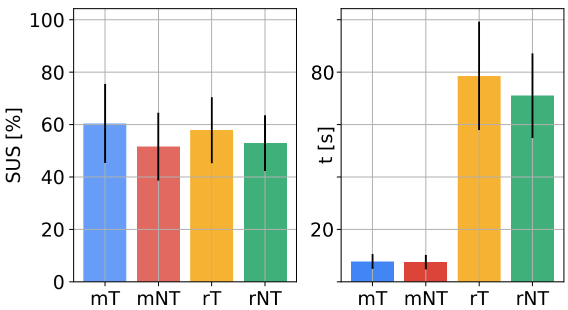

Fig. 3: Left: mean System usability score (SUS) (*higher is better*). Right: average time for different programming modality (*lower is better*). Users determined that virtual marker that leaves trace is best in terms of usability. Using virtual marker speeds up demonstration process eight times.

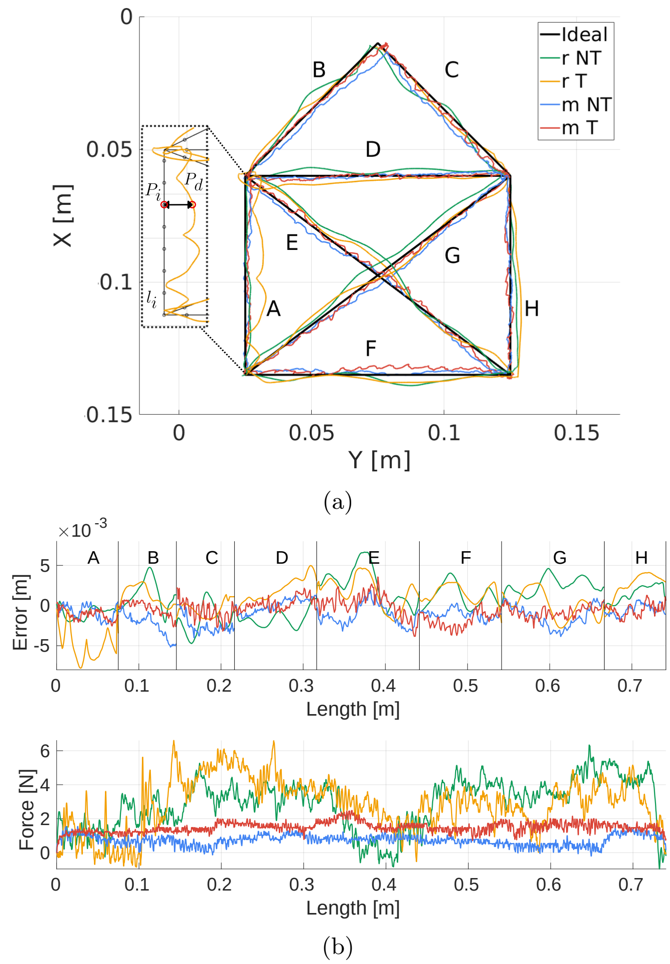

Fig. 4: Top: participant demonstration with the cobot (green: no-trace, yellow: with trace) and the virtual marker (blue: no-trace, red: with trace). The detail demonstrates the sampling of demonstration, with marked point P_i on ideal line l_i and point P_d on demonstrated trajectory. Middle: drawing errors for each segment (A-H). Bottom: contact force for each demonstrated trajectory.

categories (except performance which has inverted scale which probably confused at least some participants), when averaged across all study participants were significantly lower for the virtual marker than for the robot manipulator. Such results show that virtual marker induces less operator workload compared to the kinesthetic teaching. Averaged SUS results across participants and average demonstration time for programming modality can be found in Fig. 3.

We also performed a drawing task quantitative assessment. We evaluated the deviation of demonstrated trajectories from the ideal lines connecting these marked waypoints by transformed all recordings into a common drawing frame. Then we segmented trajectories into eight straight line sections, and resampled them to ensure an equal number of points. Each resampled point P_i on segment l_i was paired with a corresponding point P_d on the demonstrated trajectory as shown in Fig. 4. The orthogonal projection of P_d on l_i represents the drawing error. This process was conducted for each participant and each drawing case, paired with contact forces exerted between the marker and the whiteboard. Fig. 4 illustrates the task evaluation process for a single participant, showing ideal and demonstrated trajectories (robot and marker), drawing errors, and contact forces on the whiteboard.

The depiction in Fig. 5, obtained by overlapping all demonstrated trajectories, highlights that the kinesthetic demonstration spans a broader area compared to the ideal waypoints. We also computed the mean trajectory, obtained from the drawings errors. Alongside the mean trajectory, we presented the standard deviation, representing the dispersion around the mean trajectory.

The histogram depicted in Fig. 6 illustrates the distances of the trajectory points from the ideal line, with the corresponding count of such points. We highlight the *Epsilon* area in blue, representing a 3 mm wide zone around the ideal lines, where we anticipate the majority of trajectory points to fall. It is evident that for virtual marker demonstrations, a greater number of trajectory points lie within the *Epsilon* area.

The force amplitude (Fig. 4) is higher and the signal is significantly more variable in the robot demonstration. In contrast, the contact force signal in virtual marker demonstrations has a lower amplitude and is more consistent. To investigate further, we conducted Fast Fourier Transformation (FFT) on the force signal F_z, with results shown in Fig. 7. It is evident that force signals from robot demonstrations, along higher amplitude, encompass a broader range of frequencies. This can be attributed to the ergonomic challenges associated with guiding the robot flange, often resulting in the loss and re-establishment of contact between the marker and the whiteboard.

VI. Conclusion

Survey confirms ($p \ll 0.05$) that the marker induces significantly less operator workload compared to guiding the cobot (see Fig. 2), and suggests ($p \approx 0.13$) that participants found the virtual marker system more useful (see Fig 3). The task evaluation showed that trajectories demonstrated with the virtual marker were closer to the ideal task, exhibiting lower error rates and less variation than those demonstrated with the cobot (see Fig. 5 and Fig. 6). Furthermore, demonstrations with the cobot exerted

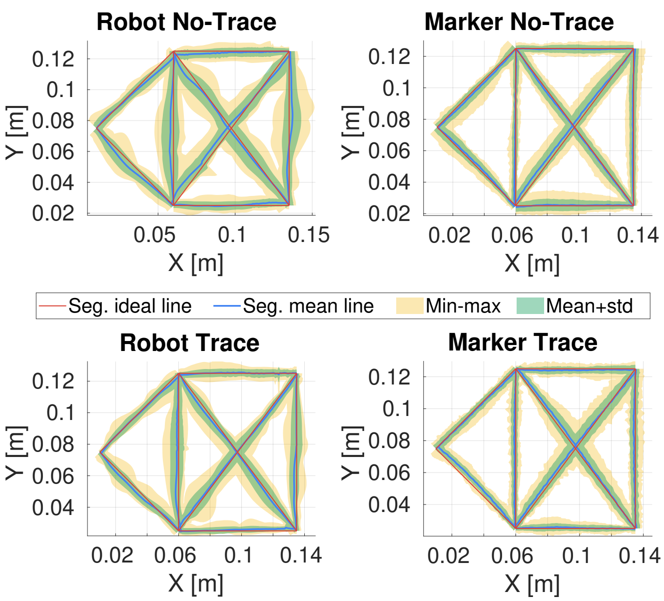

Fig. 5: The means of all demonstrations (blue) compared to the ideal trajectory (red). Yellow indicates the envelope of trajectory values, while the green area portrays the standard deviation from the mean trajectory.

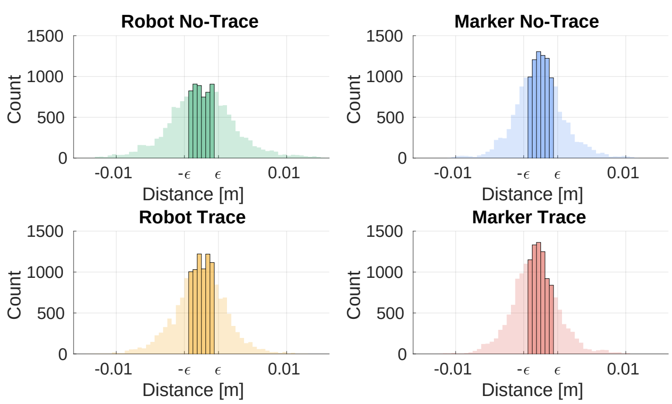

Fig. 6: Histogram of distances of trajectory points from the ideal segment's lines. The narrow ϵ area, indicating where the majority of well-demonstrated trajectory points are expected, is highlighted.

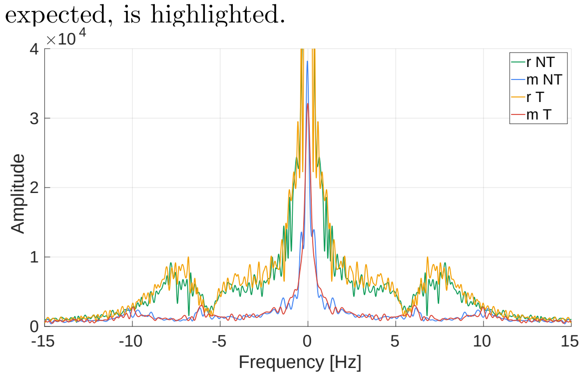

Fig. 7: Fast-Fourier-Transformation of force signals. Demonstrations where marker does not leave trace are shown in green (robot) and blue (marker), while yellow (robot) and red (marker) show demonstrations with the trace.

higher contact force with greater variations compared to virtual marker demonstrations (see Fig. 7). These findings suggest that directly guiding the cobot to demonstrate a specific task can be difficult, and the quality of the demonstration may suffer from the operator being outside their comfort zone. This may lead to demonstrations that fail to capture the true essence of the demonstrated motion, as indicated when comparing the two approaches in the case of robotic deep-micro-hole drilling of moulds in the glass manufacturing industry. The first approach [14] involves kinesthetic teaching. Work in [15] focuses on capturing the operator skill with the virtual marker coupled with exerted forces and demonstrates significant performance improvement compared to [14], underscoring the importance of comfort zone for the precise capture and successful robotic reproduction of expert skills.

References

[1] A. Grau, M. Indri, L. Lo Bello, and T. Sauter, "Robots in industry: The past, present, and future of a growing collaboration with humans," *IEEE Industrial Electronics Magazine*, vol. 15, no. 1, pp. 50–61, 2021.

[2] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, *Robot Programming by Demonstration*, pp. 1371–1394. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.

[3] G. Lentini, G. Grioli, M. G. Catalano, and A. Bicchi, "Robot programming without coding," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7576–7582, 2020.

[4] H. Ravichandar, A. S. Polydoros, S. Chernova, and A. Billard, "Recent advances in robot learning from demonstration," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, no. 1, pp. 297–330, 2020.

[5] M. Tykal, A. Montebelli, and V. Kyrki, "Incrementally assisted kinesthetic teaching for programming by demonstration," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 205–212, 2016.

[6] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Goldberg, and P. Abbeel, "Deep imitation learning for complex manipulation tasks from virtual reality teleoperation," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5628–5635, 2018.

[7] S. Calinon, *Robot Programming by Demonstration*. Boca Raton, FL, USA: CRC Press, Inc., 1st ed., 2009.

[8] P. Pastor, M. Kalakrishnan, S. Chitta, E. Theodorou, and S. Schaal, "Skill learning and task outcome prediction for manipulation," in *2011 IEEE International Conference on Robotics and Automation*, pp. 3828–3834, 2011.

[9] K. Wang, Y. Fan, and I. Sakuma, "Robot programming from a single demonstration for high precision industrial insertion," *Sensors*, vol. 23, no. 5, 2023.

[10] G. Ajaykumar and C.-M. Huang, "Multimodal robot programming by demonstration: A preliminary exploration," 2023.

[11] P. Kormushev, S. Calinon, and D. G. Caldwell, "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input," *Advanced Robotics*, vol. 25, p. 581–603, Jan. 2011.

[12] E. A. Bustamante and R. D. Spain, "Measurement Invariance of the Nasa TLX," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 52, no. 19, pp. 1522–1526, 2008.

[13] J. Brooke, "Sus: A quick and dirty usability scale," *Usability Eval. Ind.*, vol. 189, 11 1995.

[14] H. Ochoa and R. Cortesão, "Impedance control architecture for robotic-assisted micro-drilling tasks," *Journal of Manufacturing Processes*, vol. 67, pp. 356–363, 2021.

[15] B. Maric, F. Petric, D. Stuhne, V. Ranogajec, and M. Orsag, "Replicating human skill for robotic deep-micro-hole drilling," in *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)*, pp. 2238–2244, 2022.