A Unified Human-Likeness Criterion for Evaluating Human-Like Motion Retargeting on Bimanual Manipulation Tasks

Andre Meixner[a], Mischa Carl[a], Franziska Krebs[a], Steffen Steudle[a], Noémie Jaquier[b], and Tamim Asfour[a]

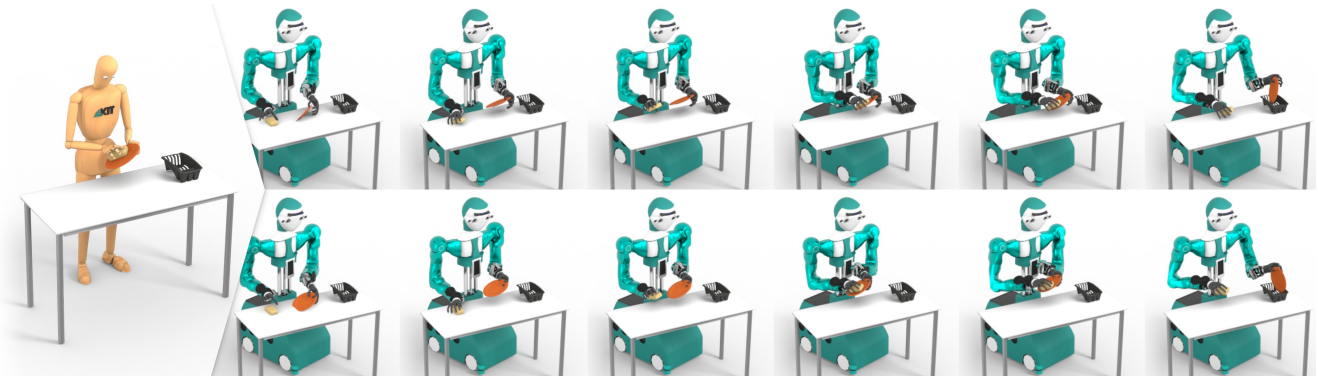

Fig. 1: Comparison of two motion retargeting approaches for wiping a plate on ARMAR-6 in a kinematic simulation. An importance-based approach (*top*, $m_{hl} \approx 0.63$) leads to an increase in the proposed unified human-likeness criterion $m_{hl} \in [0, 1]$ compared to end-effector pose tracking (*bottom*, $m_{hl} \approx 0.56$). Both non-linear optimization-based approaches are augmented with the same human-likeness criterion as additional objective and qualitatively provide similar human-like robot motions.

***Abstract*— Understanding bimanual and human-like motion is pivotal to equip humanoid robots with human-like capabilities and manipulation skills and to enable intuitive human-robot interaction. In this extended abstract, we present a multi-modal dataset of accurate whole-body human bimanual manipulation actions. Moreover, we conceptualize and derive a novel unified human-likeness criterion to assess human-like robot motions, which we evaluated across applications related to motion retargeting on bimanual manipulation tasks. Building on these results, we propose an importance-based motion retargeting approach improving human likeness.**

I. Introduction

To assist humans with a variety of household tasks, humanoid robots must be able to learn new skills and adapt them to different scenarios. In order to increase efficiency and tackle more complex tasks, humans tend to perform such daily activities using both hands simultaneously. Robot manipulation similarly benefits from the integration of both arms and the use of bimanual coordination strategies [1]. A promising method to achieve such human-like manipulation capabilities involves learning from human demonstration [2]. This poses several challenges, including learning suitable bimanual task representations on symbolic and subsymbolic levels, and reproduction across various humanoid robot kinematics. Another important factor is the quality of the executed robot motion. Human-like motions are key to enable intuitive human-robot interaction and increase the predictability and acceptability of robot movements in collaborative and assistive robot scenarios [3], [4].

The research leading to these results has received funding from the German Federal Ministry of Research, Technology and Space (BMFTR) under the Robotics Institute Germany (RIG), the European Union's Horizon 2021 Research and Innovation programme under grant agreement No 101070292 (HARIA), and the German Federal Ministry of Education and Research (BMBF) under the project CATCH-HEMI (01KU2012).

[a]Institute for Anthropomatics and Robotics, High Performance Humanoid Technologies Lab (H2T), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany. {andre.meixner, asfour}@kit.edu

[b]Division of Robotics, Perception and Learning (RPL), KTH Royal Institute of Technology, Stockholm, Sweden.

II. The KIT Bimanual Manipulation Dataset

To better understand bimanual and human-like motion, we provide a comprehensive multi-modal dataset (about 176 minutes) of whole-body human bimanual manipulation tasks with variety of short actions [5] and longer tasks [6]. The recorded sensor data is converted to the Master Motor Map (MMM) data format [7][1], a unified format for whole-body motion data that decouples the motion capture process from further post-processing tasks. The dataset contains accurate whole-body human motion and hand kinematics mapped to the MMM model — a reference model of the human body including statistical, kinematic, and dynamic properties based on biomechanical studies — and the 6D poses of objects involved in the task. The motions are manually segmented and annotated with individual actions for each hand (e. g., approach, hold, pour, stir) and, to some extent, with bimanual categories based on our bimanual manipulation taxonomy (e. g., loosely coupled, tightly coupled symmetrical) [1]. The dataset is publicly available as part of the KIT Whole-Body Human Motion Database[2]. It offers valuable insights into

[1]https://mmm.humanoids.kit.edu

[2]https://motion-database.humanoids.kit.edu/details/datasets/3521

Fig. 2: Derived unified human-likeness criterion as a hierarchical combination of 11 criteria quantifying human likeness [9].

bimanual human motions and enables the evaluation of approaches for automated action segmentation of bimanual human motions [6], classification into bimanual categories [1], learning movement primitives [8], or human-like motion retargeting on bimanual manipulation tasks [9].

III. A Unified Human-Likeness Criterion

Various criteria for assessing human likeness have been proposed in the literature, spanning various fields such as neuroscience, biomechanics, and robotics. Although numerous criteria have been carefully designed, guidelines to select the appropriate criteria for a given context remain largely overlooked. To better understand these human-likeness criteria and their interrelationships in the context of robotics, we analyze a dataset of paired human and corresponding human-like robot motions. In our context, this pairing is essential as some criteria require the comparison of resulting robot motions with the human reference, and human-like robot motions ensure that the selected criteria remain applicable. We generate this dataset using existing motion retargeting approaches across various robot kinematics, while incorporating human-likeness criteria as additional optimization objectives. We analyze different human-likeness criteria using correlation metrics on the resulting dataset, first for each kinematic individually, and then combined as an average correlation matrix $\mathbf{C}$. To derive a *unified human-likeness criterion*, we perform a hierarchical cluster analysis of the matrix $\mathbf{D} = \mathbf{1} - \mathbf{C}$ and simplify the resulting hierarchy by merging close clusters and discarding dissimilar ones. The unified human-likeness criterion can then be expressed as a weighted sum of normalized individual criteria [9].

We analyzed 11 criteria quantifying human likeness of upper-body robot motions (e. g., elbow elevation angle, joint-space jerk, similarity to the reference motion on joint level or to learned synergies extracted via principal component analysis) using bimanual human motions from our KIT Bimanual Manipulation Dataset mapped to 3 different robot kinematics. The hierarchical human-likeness criterion derived from this dataset is shown in Fig. 2. Note that the analysis and derived unified criterion depend on the generated dataset and individual criteria. However, the results align with expected relationships between criteria. We then applied the proposed unified human-likeness criterion to compare the performance of bimanual motion retargeting approaches in two applications, namely absolute and relative end-effector pose tracking. We observed that a non-linear optimization-based inverse kinematics solver (similar to [10]) augmented with human-likeness criteria achieved the highest score with respect to the proposed unified criterion and functional criteria.

IV. Importance-based Motion Retargeting on Bimanual Manipulation Tasks

In the context of bimanual manipulation tasks, human-like motion retargeting approaches should both fulfill required task-space objectives and preserve the human-like characteristics of the human reference motion. In our previous work [9], we tracked either the absolute or relative end-effector poses from the reference motion. We argued that relative end-effector tracking provides additional redundancy which enables to further optimize the human likeness. However, this is only valid in certain situations such as for asymmetric bimanual actions. Overall, the priority or importance of such objectives can either be learned implicitly from multiple demonstrations [11] or inferred from the context of a single human demonstration [12], [13].

Our goal is to design an approach based on a single reference motion that can dynamically change the relative importance of objectives on bimanual manipulation tasks, thereby resulting in more freedom for optimizing human likeness. We propose to reformulate our previous optimization-based solver in order to adjust the individual weights of objectives and parameters of their loss functions based on the context, e. g., using a *Swamp Groove* loss [14]. We specify the context by leveraging constraints arising from the human interaction with the environment [12], [15] and the coordination of the hands [1], [13]. In particular, we use calculated hand-object and object-object distance relationships in the reference motion, as well as changes in the bimanual category. In a bimanual pouring task, for instance, the non-dominant end-effector holding a cup may deviate from its absolute position depending on its distance to the table, i. e., the allowed deviation is greater the further the cup was lifted. Meanwhile, the dominant hand pouring the liquid moves relative to the non-dominant hand, following a leader-follower approach. Fig. 1 shows a preliminary qualitative comparison between an importance-based motion retargeting approach and the approach that directly follows both absolute end-effector poses. As anticipated, an importance-based approach, which allows for more flexibility in achieving task objectives, leads to an increase of our unified human-likeness criterion.

V. Conclusion

This extended abstract highlighted the potential of leveraging bimanual human motion data to better understand bimanual and human-like motions. We derived a novel unified human-likeness criterion that enables the evaluation of bimanual human-like robot motions, as well as importance-based human-like motion retargeting. Future work will address additional challenges posed by retargeting bimanual manipulation tasks which involve complex motion patterns and dexterous manipulation of objects, including the adaptation to changes in the environment. To this end, we plan to evaluate suitable motion representations that capture these capabilities at the symbolic and/or subsymbolic level using our proposed unified human-likeness criterion.

REFERENCES

[1] F. Krebs and T. Asfour, "A bimanual manipulation taxonomy," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11031–11038, 2022.

[2] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Handbook of robotics*, pp. 1371–1394, Springer Berlin Heidelberg, 2008.

[3] A. Kupferberg, S. Glasauer, M. Huber, M. Rickert, A. Knoll, and T. Brandt, "Biological movement increases acceptance of humanoid robots as human partners in motor interaction," *AI & society*, vol. 26, no. 4, pp. 339–345, 2011.

[4] M. V. Liarokapis, P. Artemiadis, C. Bechlioulis, and K. Kyriakopoulos, "Directions, methods and metrics for mapping human to robot motion with functional anthropomorphism: A review," *School of Mechanical Engineering, National Technical University of Athens*, vol. 10, 2013.

[5] F. Krebs, A. Meixner, I. Patzer, and T. Asfour, "The KIT Bimanual Manipulation Dataset," in *IEEE/RAS Intl. Conf. on Humanoid Robots (Humanoids)*, pp. 499–506, 2021.

[6] A. Meixner, F. Krebs, N. Jaquier, and T. Asfour, "An evaluation of action segmentation algorithms on bimanual manipulation datasets," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pp. 4912–4919, 2023.

[7] C. Mandery, Ö. Terlemez, M. Do, N. Vahrenkamp, and T. Asfour, "Unifying representations and large-scale whole-body motion databases for studying human motion," *IEEE Trans. on Robotics*, vol. 32, no. 4, pp. 796–809, 2016.

[8] T. Daab, N. Jaquier, C. Dreher, A. Meixner, F. Krebs, and T. Asfour, "Incremental learning of full-pose via-point movement primitives on riemannian manifolds," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 2317–2323, 2024.

[9] A. Meixner, M. Carl, F. Krebs, N. Jaquier, and T. Asfour, "Towards unifying human likeness: Evaluating metrics for human-like motion retargeting on bimanual manipulation tasks," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 13015–13022, 2024.

[10] D. Rakita, B. Mutlu, and M. Gleicher, "RelaxedIK: Real-time synthesis of accurate and feasible robot arm motion," in *Robotics: Science and Systems (R:SS)*, 2018.

[11] J. Silvério, S. Calinon, L. Rozo, and D. G. Caldwell, "Learning task priorities from demonstrations," *IEEE Trans. on Robotics*, vol. 35, no. 1, pp. 78–94, 2019.

[12] H. J. Shin, J. Lee, S. Y. Shin, and M. Gleicher, "Computer puppetry: An importance-based approach," *ACM Trans. Graph.*, vol. 20, p. 67–94, Apr. 2001.

[13] D. Rakita, B. Mutlu, M. Gleicher, and L. M. Hiatt, "Shared control–based bimanual robot manipulation," *Science Robotics*, vol. 4, no. 30, 2019.

[14] Y. Wang, P. Praveena, D. Rakita, and M. Gleicher, "RangedIK: An optimization-based robot motion generation method for ranged-goal tasks," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pp. 9700–9706, 2023.

[15] K. Otani and K. Bouyarmane, "Adaptive whole-body manipulation in human-to-humanoid multi-contact motion retargeting," in *IEEE/RAS Intl. Conf. on Humanoid Robots (Humanoids)*, pp. 446–453, 2017.