2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

Towards Unifying Human Likeness: Evaluating Metrics for Human-Like Motion Retargeting on Bimanual Manipulation Tasks

Andre Meixner, Mischa Carl, Franziska Krebs, Noémie Jaquier, and Tamim Asfour

***Abstract*— Generating human-like robot motions is pivotal for achieving smooth human-robot interactions. Such motions contribute to better predictions of robot motions by humans, thus leading to more intuitive interaction and increased acceptability. Human likeness in robot motions has been conventionally measured and realized via the optimization of human-likeness metrics. However, the abundance of such metrics and the absence of standardized criteria impede their usage in novel contexts. In this work, we introduce a unified human-likeness metric built from a hierarchically weighted sum of individual metrics. The proposed metric is derived from a thorough analysis of eleven existing human-likeness criteria and is applicable across various tasks and robot models. We evaluate its performance in the context of motion retargeting of bimanual tasks with three different humanoid robots.**

I. Introduction

Human-like motions of robots that assist, interact, and collaborate with humans significantly contribute to intuitive and effective human-robot interactions. Human-like robot motions are pivotal as they enable humans to better anticipate robot actions [1], thus leading to increased acceptability [2]. As such, the generation of human-like robot motions was investigated within various areas of robotics, including social interactions [3], [4], learning from demonstrations [5], [6], and motion planning [7], [8], among others. In this context, roboticists explored numerous criteria to evaluate the human likeness of robot motions [1], [9], [10]. These include metrics based on biologically-inspired motion patterns [3], [11], predicted joint angles [12], [13], or joint synergies extracted from human motion data [7], [14], which are then utilized for motion generation. Human-likeness can also be targeted by optimizing the robot motion directly based on kinematic similarity to a reference motion [15]–[17].

Despite the careful design of various metrics, guidelines to choose the appropriate one for a given setting remain overlooked. Specifically, several metrics were introduced without comprehensive assessment [8], [18], [19] or their evaluations are mostly achieved via high-effort user studies [3], [4], [20], [21] or by comparing the robot motion to a reference human demonstration [13]. Moreover, optimizing a trajectory for a specific human-likeness metric ensures that this metric achieves the best score compared to other criteria, thus leading to flawed evaluations.

The research leading to these results has received funding from the European Union's Horizon Europe 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). The authors are with the Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, Germany {andre.meixner, asfour}@kit.edu

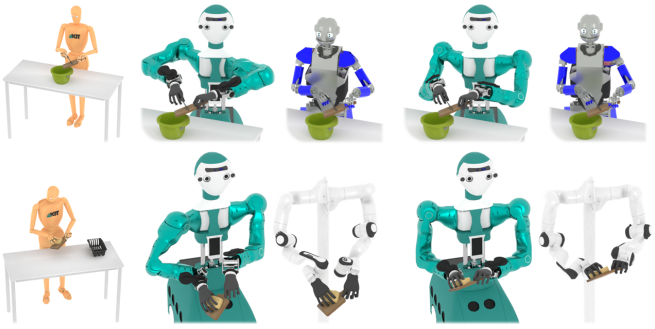


Fig. 1: Motion retargeting of two bimanual motions from the MMM reference model (left) to two robots each (middle, right) via a non-linear optimization-based approach (OPT). *Top:* End-effector pose mapping for a transfer task. *Bottom:* Relative pose mapping for a wiping task. The right motions are mapped via OPT augmented with two human-likeness criteria (*top:* SYN, HJA^r, *bot:* SYN, SYN^r).

In this paper, we aim to bring various human-likeness metrics together into a unified human-likeness metric applicable across various tasks and robots. To do so, we create a large dataset of retargeted robot motions (Section IV-A) by leveraging joint mappings between the MMM reference model [15] and three different humanoid robot kinematics (Section III-A), as well as existing human-likeness criteria (Sections III-B and III-C) to assess single and dual-arm robot configurations. Based on this large dataset, we then analyze correlations between the different human-likeness criteria (Section IV-B) and derive a unified hierarchical metric for human likeness (Section IV-C). We showcase the practical application of our metric in two human-like bimanual motion retargeting scenarios based on tasks from the KIT Bimanual Manipulation Dataset [22] (see Fig. 1 and Section V).

The contribution of this work is three-fold: (i) We analyze the correlations among 11 metrics quantifying human likeness of robot motions; (ii) we derive a novel hierarchical human-likeness metric for unified assessment of human likeness of robot motions; and (iii) we compare the performance of various human-likeness criteria to generate human-like motions on three simulated humanoid robots. A video showcasing qualitative results of several motion retargeting approaches accompanies this paper and is available at https://youtu.be/JmSRgW2cbic.

II. Related Work

The foundational research for human-like motion planning and generation draws from late 20th-century neuroscience and psychology studies [11], [23], [24] and primarily focused on unimanual arm movements. Soechting and Flanders [12], [25] showed that human arm movements are planned using

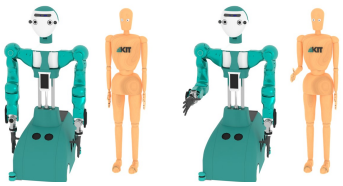


Fig. 2: Example of sampled joint configuration for ARMAR-6. *Left*: MMM default pose. *Right*: moved elbow e_x joint.

a shoulder-centric spherical coordinate system. Building on this insight, they introduced a linear model to predict the natural human arm posture from a target wrist position. A systematic dependency between the target of a pointing movement and the elevation of the elbow was found in [26]. This dependency was then leveraged by Kim et al. [27] to measure human likeness. Zanchettin et al. [13] further extended the approach to consider the complete 6D hand pose. The elbow elevation angle itself was utilized as a measure of human-likeness in [5], [8]. In contrast, human-like motions were generated in [28] based on the predicted angles of [12]. Gäbert et al. [8] proposed a distance metric based on the human range of motion. Other metrics aiming at evaluating human-likeness of unimanual arm configurations include Cartesian joint distance [29], minimum jerk [3], minimum energy [30], spatio-temporal correspondences [31], [32], and legibility and predictability [33].

In contrast, only few approaches were investigated in the context of bimanual manipulation [10], a deficiency in the field also noted in [9]. Lamperti et al. [34] quantified muscular effort in bimanual tasks, and used this metric for redundancy resolution in bimanual human-like motion planning. Suárez et al. [14] built on the concept of human joint synergies [35] to plan bimanual human-like motions using the first four principal components (PCs) which account for over 70% of the variance. The PCs were extracted from bimanual motion capture recordings using principal component analysis (PCA). PCA was also applied in [5] and a similar dimensionality reduction was achieved using neural networks in [36]. García et al. [7], [37] further expanded the concept of bimanual synergies [14] by introducing first-order synergies based on PCA in the joint velocity space.

Most of the aforementioned works address human-like motion generation in the context of motion planning. However, human-likeness metrics are important for motion retargeting and imitation learning. In recent years, machine learning methods led to significant advancements in bridging the kinematic gap between human demonstrators and target kinematics [38], [39]. However, difficulties remain in explicitly considering task constraints. The two most common approaches to incorporate task constraints treat motion retargeting either as non-linear optimization problem [15]–[17] by minimizing a weighted set of cost functions, or via inverse kinematics [6], which can include additional criteria as a weighted sum of nullspace gradients. For instance, Rakita et al. introduced RelaxedIK [40], a non-linear optimization approach, which they applied for bimanual telemanipulation switching between various functional constraints depending on the recognized bimanual actions [41].

Fig. 3: Joint mapping of bimanual human motion on various robots.

Similar to [10], we contend that methods for evaluating human likeness were developed independently and overlook transferability and compatibility. We address this concern and undertake a comprehensive analysis by evaluating 11 individual human-likeness metrics on bimanual motions.

III. Metrics for Motion Human Likeness

In this section, we first develop a closed-form conversion between similar joint spaces with high perceptual similarity. Then, we introduce the different human-likeness metrics that will be analyzed in Section IV-C. We apply unimanual metrics to our bimanual use case by averaging the result of the unimanual metric for the left and right arm.

A. Joint Mapping from MMM Reference Model

A mapping from the human joints to the robot model on a kinematic similarity level is required to apply some human-likeness criteria. We aim for a closed-form solution mostly based on one-to-one joint mappings, which requires similar kinematic structures. This approach increases performance, is not ambiguous, and also allows direct application of the human joint limits criteria [8].

We map the 7-DoF joints of the left and right arms of the MMM model [15] — a reference model of the human body including kinematics and dynamics based on biomechanical studies, and used for unified representation of human motion — to a target robot kinematic. To do so, we define linear mappings $q_i^B = a_i q_i^A + b_i$ that identify the joint configuration q^A of kinematic A with the joint configuration q^B of kinematic B. Note that special care must be applied to account for the different Euler conventions used in various robot models. To determine a and b, we first require a common default pose which, in our case, is based on the MMM reference model (see Fig. 2). The joint values of the target kinematics are then perceptually adjusted to match the MMM pose and used to calculate the static offset b. Next, we move each individual joint i of the MMM reference model by a small angle and adjust the robot kinematic accordingly to compute the linear factor a_i. We try to avoid any dynamic offset, i.e., dependence on more than one reference joint, by choosing a suitable default pose for the robot and moving individual joints only. Although this does not perfectly align the centers of the rotation axes, it leads to a one-to-one joint mapping and preserves potential joint synergies.

Table I provides the calculated parameters of the joint mapping for the humanoid robots ARMAR-III [42] and ARMAR-6 [43] and for a bimanual arm setup based on two Franka Emika Panda, similar to [44], hereinafter called 2-Panda. Although we tried to avoid dynamic offsets, the second shoulder joint ($i = 2$) of ARMAR-6 exhibits an

TABLE I: Determined joint mapping parameters from MMM to target robot. s_x, s_y and s_z are based on the Euler xyz-convention.

		ARMAR-III				ARMAR-6				2-Panda			
i	joint	Left		Right		Left		Right		Left		Right	
		a	b	a	b	a	b	a	b	a	b	a	b
1	s_x	-1	0	-1	0	-1	0	1	0	-1	0	1	0
2	s_y	1	β_{A_3}	-1	β_{A_3}	1	$\beta_{A_{6_l}}$	-1	$\beta_{A_{6_r}}$	1	$-\frac{\pi}{2}$	1	$\frac{\pi}{2}$
3	s_z	-1	0	1	0	-1	0	-1	0	-1	$\frac{\pi}{2}$	-1	$\frac{\pi}{2}$
4	e_x	1	$-\frac{\pi}{2}$	1	$-\frac{\pi}{2}$	1	0	1	0	-1	0	-1	0
5	e_z	-1	$\frac{\pi}{2}$	1	$\frac{\pi}{2}$	-1	$\frac{\pi}{2}$	-1	$-\frac{\pi}{2}$	-1	$-\frac{\pi}{2}$	-1	$\frac{\pi}{2}$
6	w_x	1	0	1	0	1	0	-1	0	-1	π	1	π
7	w_z	1	0	-1	0	-1	0	1	0	0	$\frac{3\pi}{4}$	0	$\frac{3\pi}{4}$

$\beta_{A_3} = -0.2417$, $\beta_{A_{6_l}} = \begin{cases} (2\pi - s_x)\frac{0.55}{\pi}, & \text{if } s_x \geq \pi \\ s_x \frac{0.55}{\pi}, & \text{otherwise} \end{cases}$, $\beta_{A_{6_r}} = s_x \frac{0.55}{\pi}$

angular offset relative to the ground plane and is therefore also dependent on s_x to achieve perceptual similarity to the MMM configuration. We would like to emphasize that our solution is only one of many, i.e., alternative solutions may be more suited to other applications. Results of this joint mapping for a bimanual human motion are shown in Fig. 3.

B. Demonstration-Independent Metrics

Next, we introduce human-likeness metrics that are independent of a human reference demonstration and, thus, are generally applicable to uni- and bimanual motion generation.

Elbow Elevation Angle (EEA) [8], [13]. EEA measures the normalized distance h_{EEA} (original c'_Ψ) between the robot's actual elbow elevation angle Ψ_{robot} and the predicted human-like elbow elevation angle Ψ_{human} [8, Eq. 4].

Soechting Angles (SOA) [12], [25], [28]. We use formulas derived from *pointing in the dark* [28, Eq. 26] to compute the natural arm angles [12], [25]. In addition, the input spherical coordinate r of the wrist position is scaled based on the target arm length as $r = \frac{0.6}{l_u + l_f} r$. The dividend is the arm length of the MMM model scaled to $1.8\,\text{m}$, with l_u, l_f denoting the robot's upper and lower arm lengths in meter. We define two Soechting metrics: (i) The task space metric ***SOA$_x$*** measures the distance in cm between the actual and predicted elbow position $p_{\text{elbow}} = (l_u \sin q_e^u \sin q_y^u, l_u \sin q_e^u \cos q_y^u, -l_u \cos q_e^u)^\top$ [12, Eq. 1] based on the upper arm elevation q_e^u and yaw angle q_y^u. The final distance is normalized by l_u; (ii) The joint space metric ***SOA$_q$*** measures the mean squared distance[1] between the robot configuration and predicted human joint angle targets converted to the robot's joint space. The predicted human shoulder angles are computed as $s_x = -\arctan2(\vec{l_{uy}}, \vec{l_{uz}}) + \pi$ and $s_y = -\arcsin(\vec{l_{ux}})$ where $\vec{l_u}$ denotes the normalized upper arm vector from shoulder to elbow. The elbow angle e_x is the angle between the predicted upper and lower arm vector [28, Eq. 29].

Human Joint Limits (HJL) [8]. HJL measures the distance to human joint limits in the robot's joint space [8, Eq. 2], mapped via the previously-defined joint mapping. We use the human joint limits described in [45]. In order to have the required fixed joint limits, the dynamic offset for ARMAR-6 is ignored by setting $\beta_{A_{6_r}} = \beta_{A_{6_l}} = \overline{h}_{s_x} \frac{0.55}{\pi} \approx 0.2$. This shifts the human joint limit center $\overline{h}_{s_y}$ by β_{A_6} but enables a more precise metric for s_x.

Synergy-based Reconstruction (SYN). For each robot, a large dataset of retargeted motions is created using the previously defined joint mapping. We use all motions from the KIT Extended Bimanual Manipulation Dataset [46], excluding motions used for evaluation in Section IV-B and V. We discard the first and last seconds of each recording. In addition, all motions are mirrored to account for missing left-handed subjects. Similar to [14], the principal components (PCs) are computed for each robot individually. In our case, the first five PCs account for about 70% of variance, which is explained by the two additional DoF of the MMM reference model arms. The metric is computed as the standard reconstruction error using the first five PCs.

First-Order Synergy Misalignment (MIS) [7], [37]. MIS measures the misalignment to the first-order bases obtained from human movements. The same motion data is used as for SYN. The required joint velocities are derived using a Savgol-Golitzky filter with order 3 and window length 21.

Minimum Jerk (MJE) [40]. MJE provides an approximation of the jerk using backward finite differences with the last four joint configurations.

C. Metrics with Reference Motion

In contrast to the metrics of Section III-B, the human-likeness metrics presented next require either a raw or a learned representation of a human reference motion.

Human Joint Angle Similarity (HJA^r). HJA^r measures the mean squared joint distance[1] between the mapped joint configuration from the MMM model using the predefined joint mapping and the actual robot joint configuration.

Synergy-based Pose Similarity (SYN^r). SYN^r measures the distance between the human reference and robot configuration in a lower-dimensional subspace. As such, it abstracts the joint configuration space and emphasizes bimanual joint coordination and synergy aspects. The same bimanual human motion data is used as for SYN. In this metric, the MMM model configuration is mapped to the robot joint space using the defined joint mapping. Both the mapped and actual joint configurations are further mapped to the low-dimensional space spanned by the first five PCs, and the mean squared error is computed as metric.

Elbow Elevation Angle Similarity (EEA^r). EEA^r extends EEA [8, Eq. 4] by taking the normalized angular distance to the elbow elevation angle Ψ_{MMM} of the MMM model in the reference demonstration into account. EEA^r with $h_{\text{EEA}^r} = \min(h_{\text{EEA}}, \pi^{-1} d_\theta(\Psi_{\text{robot}}, \Psi_{\text{MMM}}))$ should allow a better adaptation to situations where the predicted angle cannot be reached due to task or environmental constraints.

Manipulability Tracking (MAN^r) [47]. Most previous metrics are based on the kinematic similarity between the human and the robot. In contrast, MAN^r encourages the robot to match the velocity manipulability ellipsoid of the normalized MMM model. As the ellipsoid size depends on the kinematic chain, we compute a manipulability scaling factor λ. This is achieved by computing the 3×3 manipulability ellipsoid

[1]Computed as angular distance for the limitless joints, e. g., of ARMAR-6.

TABLE II: Empirically determined cost function weights γ_i.

EEA	$\text{SOA}_{\boldsymbol{x}}$	$\text{SOA}_{\boldsymbol{q}}$	HJL	SYN	MIS	HJA^r	SYN^r	EEA^r	MAN^r
4	10	2	1	6	1	2	2	6	0.5

$\boldsymbol{M}$ at many randomly sampled configurations and setting the scaling factor as $\lambda_R = \sqrt[3]{\max\det(\boldsymbol{M}_R)/\max\det(\boldsymbol{M}_H)}$ with $\boldsymbol{M}_R$ and $\boldsymbol{M}_H$ as the manipulability ellipsoids of the robot R and normalized MMM model, respectively. This led to $\lambda_{\text{ARMAR-III}} \approx 7.25$, $\lambda_{\text{ARMAR-6}} \approx 8.76$ and $\lambda_{\text{2-Panda}} \approx 4.7$. The metric is finally given as the geodesic distance [47, Eq. 4] between the current robot manipulability scaled by λ_R^{-1} and the target manipulability of the normalized MMM model.

IV. A Unified Metric for Human Likeness

To derive a unified human-likeness metric based on existing individual metrics, we must first understand how these relate to each other. To do so, we generate a dataset of retargeted motions, which we then leverage to analyze and correlate individual metrics for various robots. Building on our analysis, we propose a unified human-likeness metric as a hierarchical combination of individual ones.

A. Motion Retargeting Approaches

Here, we generate a dataset of retargeted motions by leveraging existing motion retargeting approaches based on human-likeness metrics. To do so, we use two different motion retargeting methods based on inverse kinematics (IK) to transfer human hand task space trajectories: (i) A Jacobian-based velocity controller *(JV)* with joint limit avoidance [48], and (ii) a non-linear optimization-based IK solver *(OPT)* using sequential least-squares programming (SLSQP) in NLopt [49] implemented similarly as RelaxedIK [40]. Both IK solvers compute the joint configurations for each arm of ARMAR-6 (8 DoF) and of 2-Panda (7 DoF), and for the arms (7 DoF) and hip (2 DoF) of ARMAR-III. They are initialized at the mean joint configuration at the first timestep. To map task space trajectories, we added a tool center point (TCP) virtual node for each arm of each robot, which we placed roughly in the middle of the fingers during a spherical grasp.

The human-likeness criteria are added as additional cost functions to both IK solvers using scaling weights γ_i shown in Table II to better align the values of the different criteria. For OPT, each cost h_i is used in combination with a groove loss $l_i = (-1)^n \exp(\frac{-(\gamma_i h_i - s)^d}{2c^2}) + r(\gamma_i h_i - s)^g$ with $n=1$, $s=0$, $c=0.1$, $r=10$, $d=g=2$, similar to RelaxedIK [40]. This parametrization of the groove loss function is also used for other criteria such as end-effector position where we additionally set $\gamma_p = 10$. The weight of the loss function l_i is set to $w_i=10$ for each human-likeness criteria, $w_p=10$, $w_o = 9$ for position and orientation, and $w_v = 7$, $w_a = 2$, $w_j = 1$ for velocity, acceleration, and jerk minimization, respectively. The respective gradients are computed via finite differences. The elbow elevation gradients are capped due to the non-continuous nature of the computation of the actual elevation angle. The weight for MJE is not listed in Table II as it is part of the implementation of OPT. In addition, MJE is excluded when using JV as it mostly worsened its performance. The gradient of the cost functions ΔH [48, Eq. 10] for JV is computed using sixth-order finite differences with a nullspace step size $k = -0.5$, weights $w_i = 0.1\gamma_i$ and normalized to $\max_i |\Delta H| \leq 0.2$. This excludes the metrics $\text{SOA}_{\boldsymbol{x}}$, HJA^r, and MAN^r which are computed using the analytic gradient. The analytic gradient for MAN^r is computed based on the manipulability Jacobian [47] with $\gamma_{h_{\text{MAN}^r}} = 0.1$ where we added the joint limit avoidance matrix [48, Eq. 13] as weight matrix. We also add a cost function for joint limit avoidance [48, Eq. 11] with $\gamma_{h_{\text{HJLA}}} = 4$. The IK is solved iteratively with JV until reaching target poses, joint value convergence, or iteration limit $i = 100$.

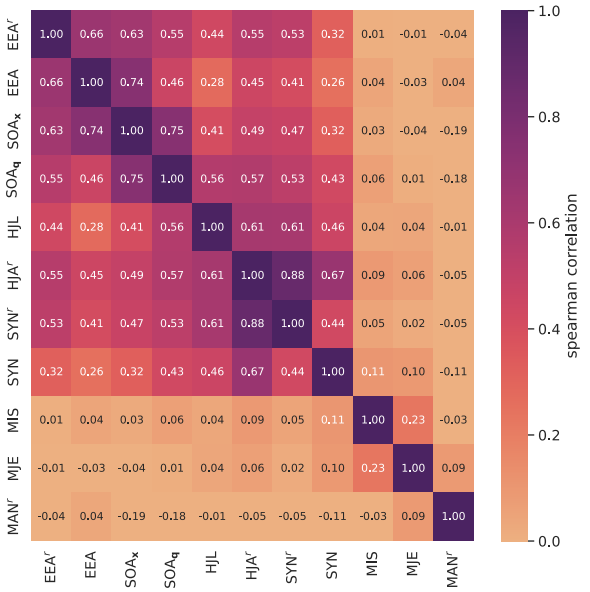


Fig. 4: Spearman correlation matrix of evaluated human-likeness metrics sorted based on hierarchical clustering.

In the following, we evaluate the performance of JV and OPT augmented with zero, one, and two human-likeness criteria for retargeting motions on various robotics platforms. Their weights w_i are divided by the number of added human-likeness cost functions. We consider the naive joint mapping approach *(JM)* introduced in Section III-A as an additional baseline. Note that JM maps the joint trajectory directly from the MMM reference model to the robot without considering any task constraints. Before retargeting motions, we find a suitable initial placement for each robot. To do so, we compute the oriented reachability maps (ORMs) [50] of the executed human left and right hand trajectories. We extend the method in [50, Sec. IV-C] to bimanual trajectory execution by selecting the minimum entry of all ORMs for left and right hand. Moreover, we added a simple mean filter to the final result before selecting the best robot placement.

B. Correlation Analysis

To analyze the relationship between existing human-likeness metrics, we select three repetitions of ten bimanual tasks (e. g., *cut*, *pour*) split equally between the two subjects of the KIT Bimanual Manipulation Dataset [22]. We use the motion retargeting approaches to create a dataset of retargeted robot motions (on ARMAR-III, ARMAR-6,

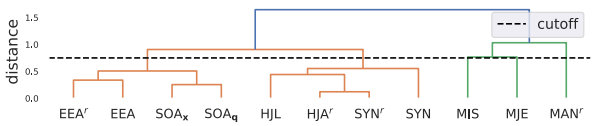


Fig. 5: Hierarchical clustering of human-likeness metrics.

TABLE III: Parameters for hierarchical human-likeness metric

	EEA	SOA$_x$	SOA$_q$	HJL	SYN	MIS	MJE	HJA^r	SYN^r	EEA^r	MAN^r
x_0	0.3	0.2	0.34	0.48	0.12	0.49	$2e^{-4}$	0.35	0.47	0.1	1.9
w_{hl}^{-1}	20	20	20	15	15	5	5	30	30	20	5

and 2-Panda) using zero or one human-likeness criteria for optimization. We restrict the dataset for OPT and JV to the robot configurations where both robot hand positions are at least within a position threshold of $10\,\mathrm{cm}$ to the target position. Then, we compute the Spearman correlation coefficients of the respective human-likeness metrics. Such coefficients handle non-linear correlations and outliers [51]. The resulting correlation matrix $\boldsymbol{C} \in \mathbb{R}^{11\times11}$, averaged over the three robots, is displayed in Fig. 4. To better analyze similarities within the human-likeness metrics, we further perform a hierarchical clustering using Ward's method [52] on the distance matrix $\boldsymbol{D} = 1 - \boldsymbol{C}$. The resulting hierarchy is shown in Fig. 5. We observe that metrics such as MAN^r, MJE, and MIS are rather uncorrelated to others. Interestingly, these metrics focus on different aspects of the motion, i.e., manipulability, joint velocity, and jerk. Instead, metrics based on kinematic similarities belong to the same tree branch and are closely correlated when based on the same underlying concept, see, e. g., SYN^r and HJA^r which are based on the reference motion and defined joint mapping, or SOA and EEA which are derived from unimanual human motion studies. We also observe high correlations between demonstration-independent metrics (SYN, HJL, SOA$_q$) and the joint mapping-based metric HJA^r, making these particularly interesting in the absence of human demonstrations. We hypothesize that the correlation values increase when the joints or synergies of both arms are considered. This would also illustrate why metrics derived from bimanual human movements (SYN, SYN^r) do not correlate closely with metrics derived from unimanual studies (EEA, SOA). Moreover, the mean absolute deviation of the correlation coefficients across evaluated robots ($\mathrm{MAD}_{\mathrm{robots}} = 0.126$) and retargeted motions ($\mathrm{MAD}_{\mathrm{motions}} = 0.1$) is non-zero, thus indicating that the correlation coefficients differ to a certain extent across motions and slightly more across robot models.

C. Hierarchical Human-Likeness Metric

Building on our analysis, we propose a unified human-likeness metric as a hierarchical combination of individual metrics. To do so, we first normalize the cost function corresponding to each metric to an interval $[0, 1]$. Namely, the metric m_i is computed over a time window T as

$$m_i = \frac{1}{T}\sum_{t=1}^{T} f_{x_{0_i}}(h_i(t)), \tag{1}$$

with the logistic function $f_{x_{0_i}}(x) = L(1 + e^{-k(x-x_{0_i})})^{-1}$, $k = -5{x_{0_i}}^{-1}$ and $L = 1$. The value of the function's middle point x_0 with $f_{x_0}(x_0) = 0.5$ is chosen for each criterion based on its median on the retargeted dataset (see Table III). Based on the correlation results, we derive a hierarchical human-likeness metric by merging close clusters ($\Delta_{\mathrm{distance}} < 0.15$ in Fig. 5) and averaging all metrics m_i in each cluster below a *cutoff* distance of 0.75 (see Fig. 5). Note that opting for a higher threshold merges less-correlated metrics, whereas a lower value reduces emphasis on those metrics. Our hierarchical human-likeness metric is obtained as

$$m_{hl} = \sum_{i}^{N} w_{hl_i} m_i, \quad \text{with} \sum_{i}^{N} w_{hl_i} = 1, \tag{2}$$

with w_{hl_i} as in Table III. To apply this hierarchical human-likeness metric along with certain functional constraints, we combine the hierarchical metric (2) with a criterion m_f that scores task-specific functionalities and constraints. The final metric is computed as $m_{hl,f} = \alpha m_{hl} + (1-\alpha)m_f$ with $\alpha \in [0, 1]$. In this paper, we use $\alpha = 0.5$ to equally weight these criteria and focus on evaluating human likeness.

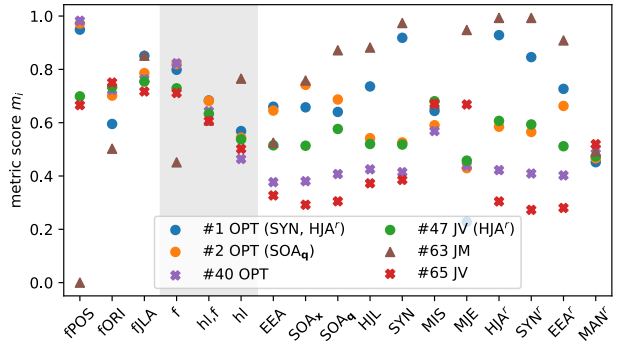
V. Evaluation

We evaluate the proposed hierarchical human-likeness criteria on two motion retargeting applications, namely (i) end-effector pose tracking and (ii) relative pose tracking. Fig. 1 and the accompanying video provide qualitative results.

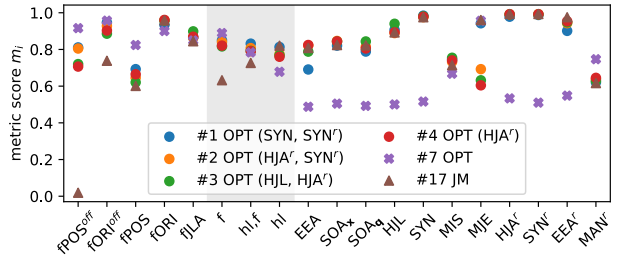
A. End-Effector Pose Retargeting

We evaluate our unified human-likeness criteria on a test dataset obtained similarly as in Section IV-A with different motions. Namely, we use the same tasks but executed by the other subject, another variation of the same task, or both. We use the same weights, approaches, and cost functions as in Section IV-A with up to two human-likeness criteria. We set m_f as mean of the normalized joint limit avoidance *(fJLA)* [48, Eq. 11], mean Euclidean distance in meter *(fPOS)* and angle in radian *(fORI)* between the actual robot and target hand poses, as these are the basic constraints in robot motion generation. The normalization parameters are empirically set to $x_{0,\mathrm{fJLA}} = 0.5$, $x_{0,\mathrm{fPOS}} = 0.03$, and $x_{0,\mathrm{fORI}} = 0.35$.

Fig. 6a shows the hierarchical and individual metric scores achieved by selected retargeting approaches. Overall, OPT largely outperforms JV for both human-likeness and functionality criteria. We hypothesize that this is due to a large amount of difficult and unreachable end-effector poses, especially for ARMAR-III, which has a large influence on Jacobian-based approaches. However, we observe that optimizing human likeness increases scores in both approaches with low effect on the functionality. Overall, SYN and HJA^r (see also Fig. 1) achieve the best scores, while SOA$_q$ as an individual optimization criterion also performs exceptionally well due to high functional scores. Moreover, OPT with additional criteria can outperform JM for functional constraints while remaining comparable in several human-likeness metrics. Note that the human-likeness scores are not improving significantly in case of ARMAR-III due to its already human-like joint angle space (see video).



(a) End-effector pose retargeting (two best general approaches, best JV approach, OPT and JV without cost functions, and JM).



(b) Relative pose retargeting (four best general approaches, OPT without cost functions, and JM).

Fig. 6: Individual metric scores m_i for selected motion retargeting approaches averaged across robots and motions. The number in the legend represents the placement of the approach based on $m_{hl,f}$. Combined functional m_f and hierarchical human-likeness m_{hl} metric scores are highlighted ( ).

Multiple optimization criteria can also have a negative impact, particularly in terms of functional constraints. Reducing α which puts a stronger focus on functional constraints would therefore generally improve the scores of optimization approaches based on single human-likeness criteria.

B. Relative Pose Retargeting

Similar to [34], we solve the IK for the relative pose between both hands extracted from the human demonstration. Relative pose retargeting is conducted for a subset of the test dataset, namely four asymmetric bimanual actions (*peeling*, *wiping*, *pouring* and *closing*) for a total of 12 motions. The success of these actions is mostly based on the relative motion of both hands. In this section, we focus the evaluation on OPT and JM, as OPT outperformed JV in Section V-A.

The initial pose at the first timestep is determined based on the end-effectors as in Section V-A. The following cost functions are then added to the optimization problem,

$$h_{\text{fPOS}^{\text{off}}} = \left\| \mathbf{t}_{\text{left}} + \mathbf{R}_{\text{left}} \cdot \mathbf{t}_{\text{left}}^{\text{right}} - \mathbf{t}_{\text{right}} \right\|_2, \tag{3}$$

$$h_{\text{fORI}^{\text{off}}} = d_\theta(\mathbf{R}_{\text{left}} \cdot \mathbf{R}_{\text{left}}^{\text{right}}, \mathbf{R}_{\text{right}}), \tag{4}$$

where $\mathbf{t}_i$, $\mathbf{R}_i$ denote the global position and orientation of the robot's TCP, respectively, and d_θ the absolute angle between rotations. These costs aim at minimizing the distance to relative position $\mathbf{t}_{\text{left}}^{\text{right}}$ and orientation $\mathbf{R}_{\text{left}}^{\text{right}}$ offsets between the left and right TCPs extracted from the human demonstration.

Compared to Section V-A, the relative retargeting increases the dimension of the nullspace from 2 to 8 DoF for 2-Panda and from 4 to 10 DoF for ARMAR-III and ARMAR-6. To cope with this additional redundancy, we scale the weight of the human-likeness cost functions by 10. Additionally, the relative pose weights are empirically set as $w_{\text{fPOS}^{\text{off}}} = 3$, $w_{\text{fORI}^{\text{off}}} = 15$, and the weights for minimum velocity, acceleration, and jerk are set as $w_v = 40$, $w_a = 10$, and $w_j = 5$. The end-effector position w_p and orientation w_o distance weights are set to 0. Moreover, we add the position offset (*fPOS^{off}*) and orientation offset (*fORI^{off}*) to m_f. The normalization parameters are set to $x_{0,\text{fJLA}} = 0.5$, $x_{0,\text{fPOS}^{\text{off}}} = 0.03$, $x_{0,\text{fORI}^{\text{off}}} = 0.35$, $x_{0,\text{fPOS}} = 0.3$, and $x_{0,\text{fORI}} = 1$.

Fig. 6b shows selected retargeting approaches and their individual and combined metric scores. Interestingly, the four best approaches all use a metric with reference motion as optimization criteria, such as SYN^r or HJA^r. Moreover, their human-likeness score significantly increased compared to the previous evaluation in Fig. 6a. This can be explained by the higher nullspace dimension. The best approaches also outperform OPT without additional human-likeness criteria and perform comparably to JM in terms of human-likeness scores. However, the functionality scores for the position offset of OPT are reasonably better.

VI. Conclusion and Future Work

This paper provided a thorough analysis of correlations between existing human-likeness criteria for robot motions. By building on our analysis, we introduced a novel hierarchical human-likeness metric that unifies existing criteria into a single measure. In particular, our unified criteria, made of a weighted combination of individual metrics, consistently accounts for similarities among existing metrics. We compared the performance of various bimanual motion retargeting approaches evaluated with respect to our criteria. We observe that SLSQP-based IK solver augmented with several human-likeness criteria consistently achieved the highest scores.

It is important to notice that the analysis and proposed unified metric is subject to the generated motion retargeting dataset, which itself depends on the design of the individual human-likeness metrics and joint mapping. This is due to the fact that a different dataset may result in slightly different correlation coefficients, which then affect the proposed unified metric. However, our unified metric reflects the expected relationships between metrics. Therefore, we expect that it will undertake only limited modifications when considering different datasets. This will be verified in future work by extending our dataset with different humanoid robots and additional metrics with the aim of obtaining a more accurate human-likeness measure and correlation scores with lower inter-robot variance. In this work, several parameters of the unified human-likeness metric and the utilized optimization approaches were set empirically based on our analysis. In future work, we will investigate methods to automatically select and fine-tune these parameters, while accounting for our hierarchical analysis. Finally, we plan to evaluate the unified human-likeness metric and the different human-like motion retargeting approaches on real robots.

REFERENCES

[1] 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.

[2] 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.

[3] E. De Momi, L. Kranendonk, M. Valenti, N. Enayati, and G. Ferrigno, "A neural network-based approach for trajectory planning in robot–human handover tasks," *Frontiers in Robotics and AI*, vol. 3, p. 34, 2016.

[4] R. Rasch, S. Wachsmuth, and M. König, "A joint motion model for human-like robot-human handover," in *IEEE/RAS Intl. Conf. on Humanoid Robots (Humanoids)*. IEEE, 2018, pp. 180–187.

[5] S. Y. Shin and C. Kim, "Human-like motion generation and control for humanoid's dual arm object manipulation," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2265–2276, 2015.

[6] N. Jaquier, L. Rozo, and S. Calinon, "Analysis and transfer of human movement manipulability in industry-like activities," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2020, pp. 11 131–11 138.

[7] N. García, J. Rosell, and R. Suárez, "Motion planning using first-order synergies," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2015, pp. 2058–2063.

[8] C. Gäbert, S. Kaden, and U. Thomas, "Generation of human-like arm motions using sampling-based motion planning," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2021, pp. 2534–2541.

[9] G. Gulletta, W. Erlhagen, and E. Bicho, "Human-like arm motion generation: A review," *Robotics*, vol. 9, no. 4, p. 102, 2020.

[10] U. Trivedi, D. Menychtas, R. Alqasemi, and R. Dubey, "Biomimetic Approaches for Human Arm Motion Generation: Literature Review and Future Directions," *Sensors*, vol. 23, no. 8, p. 3912, Jan. 2023.

[11] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," *Journal of Neuroscience*, vol. 5, no. 7, pp. 1688–1703, 1985.

[12] J. F. Soechting and M. Flanders, "Errors in pointing are due to approximations in sensorimotor transformations," *Journal of neurophysiology*, vol. 62, no. 2, pp. 595–608, 1989.

[13] A. M. Zanchettin, P. Rocco, L. Bascetta, I. Symeonidis, and S. Peldschus, "Kinematic analysis and synthesis of the human arm motion during a manipulation task," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*. IEEE, 2011, pp. 2692–2697.

[14] R. Suárez, J. Rosell, and N. García, "Using synergies in dual-arm manipulation tasks," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2015, pp. 5655–5661.

[15] 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.

[16] S. Choi and J. Kim, "Towards a natural motion generator: a pipeline to control a humanoid based on motion data," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2019, pp. 4373–4380.

[17] S. Masuda, K. Ayusawa, and E. Yoshida, "Optimization framework of humanoid walking pattern for human motion retargeting," in *IEEE/RAS Intl. Conf. on Humanoid Robots (Humanoids)*, 2018, pp. 270–276.

[18] G. Averta, D. Caporale, C. D. Santina, A. Bicchi, and M. Bianchi, "A technical framework for human-like motion generation with autonomous anthropomorphic redundant manipulators," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2020, pp. 3853–3859.

[19] G. Gulletta, E. Costa e Silva, W. Erlhagen, R. Meulenbroek, M. F. P. Costa, and E. Bicho, "A human-like upper-limb motion planner: Generating naturalistic movements for humanoid robots," *International Journal of Advanced Robotic Systems*, vol. 18, no. 2, 2021.

[20] L. Ren, A. Patrick, A. A. Efros, J. K. Hodgins, and J. M. Rehg, "A data-driven approach to quantifying natural human motion," in *SIGGRAPH*, 2005.

[21] N. Fligge, J. McIntyre, and P. van der Smagt, "Minimum jerk for human catching movements in 3d," in *IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2012, pp. 581–586.

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

[23] P. G. Morasso, "Spatial control of arm movements," *Experimental Brain Research*, vol. 42, no. 2, pp. 223–227, Apr. 1981.

[24] W. Abend, E. Bizzi, and P. Morasso, "Human arm trajectory formation," *Brain : a journal of neurology*, vol. 105, pp. 331–48, Jul. 1982.

[25] J. F. Soechting and M. Flanders, "Sensorimotor representations for pointing to targets in three-dimensional space," *Journal of neurophysiology*, vol. 62, no. 2, pp. 582–594, 1989.

[26] T. Kang, Y. Shimansky, and J. He, "Angle of elbow elevation depends on the reach target coordinates," in *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] [Engineering in Medicine and Biology*, vol. 3, 2002, pp. 2571–2572 vol.3.

[27] S. Kim, C. Kim, and J. H. Park, "Human-like arm motion generation for humanoid robots using motion capture database," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, 2006, pp. 3486–3491.

[28] T. Asfour and R. Dillmann, "Human-like motion of a humanoid robot arm based on closed-form solution of the inverse kinematics problem." in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2003, pp. 407–1412.

[29] M. V. Liarokapis, P. K. Artemiadis, and K. J. Kyriakopoulos, "Functional anthropomorphism for human to robot motion mapping," in *IEEE Intl. Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2012, pp. 31–36.

[30] H. Klein, N. Jaquier, A. Meixner, and T. Asfour, "A riemannian take on human motion analysis and retargeting," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2022, pp. 5210–5217.

[31] M. J. Gielniak and A. L. Thomaz, "Spatiotemporal correspondence as a metric for human-like robot motion," in *International Conference on Human-Robot Interaction (HRI)*. IEEE, 2011, pp. 77–84.

[32] M. J. Gielniak, C. K. Liu, and A. L. Thomaz, "Generating human-like motion for robots," *Intl. Journal of Robotics Research*, vol. 32, no. 11, pp. 1275–1301, 2013.

[33] A. Dragan and S. Srinivasa, "Integrating human observer inferences into robot motion planning," *Autonomous Robots*, vol. 37, no. 4, pp. 351–368, 2014.

[34] C. Lamperti, A. M. Zanchettin, and P. Rocco, "A redundancy resolution method for an anthropomorphic dual-arm manipulator based on a musculoskeletal criterion," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2015, pp. 1846–1851.

[35] M. Santello, M. Flanders, and J. Soechting, "Postural hand synergies for tool use," *Journal of neuroscience*, vol. 18, pp. 10 105–15, Dec. 1998.

[36] P. Chen, H. Zhao, X. Zhao, D. Ge, and H. Ding, "Dimensionality reduction for motion planning of dual-arm robots," in *IEEE International Conference on Mechatronics and Automation (ICMA)*, 2018, pp. 718–723.

[37] N. García, J. Rosell, and R. Suárez, "Motion Planning by Demonstration With Human-Likeness Evaluation for Dual-Arm Robots," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 11, pp. 2298–2307, Nov. 2019.

[38] J. Lim, H. J. Chang, and J. Y. Choi, "Pmnet: Learning of disentangled pose and movement for unsupervised motion retargeting," in *British Machine Vision Conference*, 2019.

[39] K. Aberman, P. Li, D. Lischinski, O. Sorkine-Hornung, D. Cohen-Or, and B. Chen, "Skeleton-aware networks for deep motion retargeting," *ACM Transactions on Graphics*, vol. 39, no. 4, Jul. 2020.

[40] 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)*, Jun. 2018.

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

[42] T. Asfour, K. Regenstein, P. Azad, J. Schroder, A. Bierbaum, N. Vahrenkamp, and R. Dillmann, "ARMAR-III: An integrated humanoid platform for sensory-motor control," in *IEEE/RAS Intl. Conf. on Humanoid Robots (Humanoids)*, 2006, pp. 169–175.

[43] T. Asfour, M. Wächter, L. Kaul, S. Rader, P. Weiner, S. Ottenhaus, R. Grimm, Y. Zhou, M. Grotz, and F. Paus, "ARMAR-6: A High-Performance Humanoid for Human-Robot Collaboration in Real World Scenarios," *IEEE Robotics and Automation Magazine*, 2019.

[44] C. Lenz and S. Behnke, "Bimanual telemanipulation with force and haptic feedback and predictive limit avoidance," *European Conference on Mobile Robots (ECMR)*, pp. 1–7, 2021.
[45] K. Luttgens and N. Hamilton, *Kinesiology: Scientific Basis of Human Motion*. Brown & Benchmark, 1997.
[46] 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)*, 2023.
[47] N. Jaquier, L. Rozo, D. G. Caldwell, and S. Calinon, "Geometry-aware manipulability learning, tracking, and transfer," *Intl. Journal of Robotics Research*, vol. 40, no. 2-3, pp. 624–650, 2021.
[48] T. F. Chan and R. Dubey, "A weighted least-norm solution based scheme for avoiding joint limits for redundant joint manipulators," *IEEE Transactions on Robotics and Automation*, vol. 11, no. 2, pp. 286–292, 1995.
[49] S. G. Johnson, "The NLopt nonlinear-optimization package," https://github.com/stevengj/nlopt, 2007.
[50] N. Vahrenkamp, T. Asfour, and R. Dillmann, "Robot placement based on reachability inversion," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2013, pp. 1970–1975.
[51] J. de Winter, S. Gosling, and J. Potter, "Comparing the pearson and spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data," *Psychological Methods*, vol. 21, pp. 273–290, Sep. 2016.
[52] J. H. W. Jr., "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963.