Computers & Graphics 138 (2026) 104643

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Special Section on SMI2026

Data-driven Inverse Kinematics using Laban Movement Analysis☆

Mehmet Akif Şahin[1], Sinan Sonlu[1], Uğur Güdükbay[*]

Department of Computer Engineering, Bilkent University, Ankara, Turkiye

ARTICLE INFO

Dataset link: https://github.com/Bilkent-ModVis/LMA-IK

Keywords:
Inverse kinematics
Laban Movement Analysis
Data-driven animation
Deep learning

ABSTRACT

Inverse Kinematics (IK) provides control over animation, facilitating the creation of full-body poses by utilizing target end-effector locations. Many approaches address the physical aspects of arranging limb configurations; however, systems that consider the psychological aspects of human motion are lacking. To this end, we introduce a quantitative translation of the qualitative concepts of Laban Movement Analysis (LMA) into computable, continuous style descriptors. Building upon this formulation, we also propose a data-driven Inverse Kinematics (IK) method that directly utilizes these LMA parameters to refine generated animations. Specifically, we refer to LMA Shape Qualities and the attitude towards the Kinesphere to control the orientation of the generated pose along the vertical and horizontal axes. Our *Interpolator* upsamples sparse end-effector keyframes into dense paths and modulates Time Effort at the trajectory level. Flow Effort is controlled by a pose-similarity objective that deliberately reduces pose similarity to the dataset examples. Through a perception user study, we show that the system can successfully apply LMA-based changes to the motion to express different personality traits. This data-driven system can ease the process of controlling the psychological aspect of generative animation.

1. Introduction

Simulating human body motion is an emerging research field with many applications in biomechanics, robotics, and animation. Studies often use motion capture techniques that require performance actors, expensive equipment, or complex procedural algorithms to generate plausible human motion. Applications also use Inverse Kinematics (IK) to determine the rotational configuration of body parts that satisfy certain contact requirements, enabling animations from a series of interpolated end-effector positions. IK is a well-studied area with different numerical and data-driven solutions [1]. At the same time, many IK approaches focus on the physical aspects of motion, and research on its psychological aspects is limited. To this end, this work aims to introduce a psychological layer to IK using Laban Movement Analysis (LMA).

LMA is a theoretical framework used originally for dance choreography. It is also used to analyze general human motion. Studies utilize LMA for various tasks including motion style synthesis [2], action recognition [3], emotion understanding [4], semantic segmentation [5], dance evaluation [6], personality expression [7] and personality recognition [8]. Although a popular approach to expressive animation, LMA lacks standard algorithms for calculating quantitative parameters, and studies often employ task-oriented calculations that follow verbal descriptions of LMA. LMA describes motion through four interrelated components — Body, Effort, Shape, and Space — which we detail in Section 2.

This work focuses on style-based generative animation using descriptors inspired by LMA Shape Qualities and Effort parameters, which strongly connect to expressive motion [9–11]. We use LMA Shape Qualities to describe posture along the vertical, horizontal, and frontal axes. Although Shape Qualities describe the change in motion, such as in Spreading motion, which goes from narrow to wide, we use them in relation to the Kinesphere to describe the spatial space around the body in relation to different axes. Time and Flow Efforts concern the continuous aspect of motion and are high-level parameters to our pose generation system, influencing the interpolation of the end effectors. Time Effort describes the quickness of motion; it influences the selective keyframe reduction. Flow describes the progression; it captures whether the motion is bound or free. We introduce Free Flow by a pose-similarity control during synthesis. Lower similarity to training examples yields freer, less constrained motion, whereas higher similarity yields more bound motion. Rather than using LMA parameters directly, we utilize logical combinations in our LMA-inspired style descriptors that control the style of motion synthesis.

☆ This article is part of a Special issue entitled: 'CAG_SS_SMI2026' published in Computers & Graphics.

* Corresponding author.
E-mail addresses: akif.sahin@ug.bilkent.edu.tr (M.A. Şahin), sinan.sonlu@bilkent.edu.tr (S. Sonlu), gudukbay@cs.bilkent.edu.tr (U. Güdükbay).
[1] Joint first authors.

https://doi.org/10.1016/j.cag.2026.104643
Received 10 May 2026; Accepted 31 May 2026
Available online 5 June 2026

We model the articulated human body as a tree structure; our approach also functions with arbitrary skeletons, but the role of LMA is specific to the human body. We train our data-driven IK system using samples from Bandai [12], Dance Motion Capture [6], and LMA Effort [13] datasets, using their respective armature configurations. We also use samples from these datasets for evaluation. The system processes motion segments as fixed-length windows of 50 frames. Using motion capture samples yields more realistic results than training on randomly generated configurations.

We evaluate the resulting system based on how well it can alter the personality and emotion perception of animation samples. We employ both quantitative metrics (Fréchet Inception Distance [14] between generated and real motion distributions) and qualitative assessment through user studies. To this end, we compare pairs of generated samples using opposing style descriptors to observe our system's influence on the perceived personality and emotion with a user study. The results mostly confirm the positive effect of utilizing LMA-based parameters for motion stylization.

Our system offers a generalizable solution for altering motion regarding LMA-inspired style descriptors; this approach can benefit full-body motion reconstruction from sparse data in Virtual Reality systems [15], improve personality expression in conversational agents [16], and aid in designing IK-based animations with controlled variation [17].

Our contributions are as follows:

- a quantitative translation of qualitative LMA Shape and Effort qualities into computable, continuous style descriptors, making them directly applicable to deep learning objective functions,
- a proof-of-concept data-driven IK system for generating animation utilizing sparse end-effector positions and the proposed LMA-based descriptors, and
- a qualitative analysis of the proposed approach for altering generated samples' emotional content and personality.

Our source code is accessible in our public repository for future studies[2]. The remainder of this article includes related work (Section 2), a detailed explanation of our system (Section 3), evaluation (Section 4), limitations (Section 5), and conclusion (Section 6).

2. Related work

Laban Movement Analysis (LMA), originally developed by Rudolf Laban for choreography and movement notation, provides a framework for describing human motion through four interrelated components [6]. The *Body* component examines which body parts are active and how they interrelate during motion. The *Effort* component captures the dynamic, inner attitude of movement and is decomposed into four factors, each on a bipolar continuum: Weight (Light ↔ Strong), Space (Indirect ↔ Direct), Time (Sustained ↔ Sudden), and Flow (Free ↔ Bound) [18]. The *Shape* component describes how the body changes form, with Shape Qualities expressing directional change along three dimensions: Rising/Sinking (vertical), Spreading/Enclosing (horizontal), and Advancing/Retreating (sagittal) [18]. The *Space* component analyzes the mover's relationship with the surroundings, including the *Kinesphere*—the personal reach-space defined by the limbs around the body [6]. These qualitative descriptors form the basis of expressive motion analysis but lack standardized quantitative formulations, which motivates task-specific computational interpretations such as ours. We mention related works on LMA-based motion adjustments for stylization and personalization, as well as IK-based studies on human animation.

[2] https://github.com/Bilkent-ModVis/LMA-IK

2.1. LMA-based motion adjustments

Studies use LMA-based high-level motion parameters to generate and classify expressive motion. LMA Effort and Shape parameters help generate more natural synthetic gestures through systematic adjustments [18]. Using LMA-based motion adjustments establishes observable personality cues in animation [7]. Such observable cues help express different personality factors more accurately with dialogue, voice, and facial expressions [16]. LMA-based features ease emotion [19, 20] and personality [8] recognition. Studies use handcrafted measurements to directly calculate LMA qualities from motion data [21]. Human motion analysis, which utilizes LMA parameters, reduces the dimensionality of motion capture [22], thereby easing semantic segmentation [5]. LMA offers learnable motion parameters that enable the application of user-specified styles to motion sequences [2]. LMA Effort parameters also apply to animals concerning various psychological patterns [23].

LMA parametrization can model interpersonal behavior and group dynamics [24] and can help express emotions in humanoid robots [25], even only using a robotic arm [26]. Humans can perceive psychological attributes in simple motion cues [27], which helps the expressive movement of non-humanoid robots to appear emotional [28]. The perceptual consistency of singular LMA components shows mixed results [29]; this is likely due to the correlations between LMA factors [13]. Even when the motion focuses on a single LMA Effort, observers perceive co-occurring qualities together. This phenomenon could enhance the effect of LMA-based motion changes; slight alterations in the body pose can suggest different emotions [4]. Studies also use proximity to convey different emotions [30], which can be analyzed using the LMA Space category. Analyzing limb trajectories can provide an overview of human actions, enabling automatic control of camera settings [31]. Gesture qualities such as arm swivel or stroke scale, together with language usage, have strong connections to perceived extraversion [32].

2.2. IK-based animation

Kinematic methods are used to generate motion. *Forward Kinematics* calculates the final state of the model using given parameters; *Inverse Kinematics* (IK) is the opposite; given the final state of the end effectors, IK solves the parameters to achieve a requested state. There are three main approaches to IK: The first uses *analytical* solutions, which involve solving parameters using geometric relations. These methods are efficient as they do not require iterative computations. However, their applicability is generally limited to a few degrees of freedom. Tolani et al. [33] propose an analytical algorithm to solve IK problems for a 7-degree-of-freedom chain structured like a human limb. Kallmann [34] proposed an analytical IK method with flexible body control for animating reaching tasks while preventing collisions. Another approach uses *numerical* methods, such as the Jacobian inverse technique, which are flexible and can handle complex structures with many degrees of freedom [35]. These approaches iteratively adjust the parameters to achieve the desired state. Kang et al. [36] introduce an accelerated IK solver that accurately reconstructs skinned 3D models by precomputing duplicate terms in the standard IK computations for future use. The third approach includes *data-driven* solutions, where a network model learns the required set of joint rotations to achieve the desired configuration [37].

Grochow et al. [38] presented an IK pose generation system that uses a Scaled Gaussian Process Latent Variable Model to represent the probability distribution of poses. The system aims to generate the most likely pose that satisfies the given constraints in real-time. Other notable deep-learning-based approaches to IK include [39], which attempts to solve the problem for a two-limb kinematic chain in two-dimensional space; Vats [40]; and Starke [41], which attempt to predict angles for a six-degree-of-freedom robot arm. Data-driven systems often

use the figure's base pose to produce plausible results, leveraging full-body correlations [42]. Additional IK controls to further adjust body shape, including spine curvature and shoulder angle, can help achieve more expressive stances [43]. Voss and Kopp [44] present a real-time IK solver that generates realistic human-like movement by addressing the error accumulation and complex joint limit problems using forward and inverse kinematics differentiation. Prior work establishes IK-based motion stylization using statistical models [38]. LMA-based motion parameters can control the IK end effectors to adjust dance animations to express emotion [45]. Individual body parts can incorporate arbitrary motion styles to convey diverse global styles [46]. Unlike existing work, our focus is on expressing personality and emotions with the same set of parameters rather than an arbitrary style.

Most IK solutions are skeleton-based; however, mesh-based IK systems, such as [47], also exist. HybrIK-X [48] is a hybrid inverse kinematics framework that integrates 3D keypoint estimation with body-mesh recovery in a unified manner, yielding accurate 3D joint information and a realistic human model structure. MANIKIN [49] is a neural inverse kinematics solver predicting accurate physical poses for full-body motion tracking by embedding anatomical constraints into the model. Moreno-Villamarin et al. [50] propose a generative model for synthesizing human motion using sparse training data by adopting multi-scale conditional generation and blending, producing a human mesh for each frame by exploiting the SMPL body model.

Data-driven animation research is not limited to IK-based solutions; studies utilize speech [51,52], music [53], or video [54] inputs to drive animations. Physically simulated characters can display diverse behaviors [55], and recent models can generate text-driven stylized motion [56]. Spatial–temporal graphs enable the application of a learned motion style to neutral animation samples [57]. Current generative models can leverage dense input features to provide fine-grained control over generated animations [58]; by contrast, we focus on motion synthesis using sparse end-effector positions. Studies also focus on adaptive techniques to store motion capture data more efficiently [59,60], where this reduced representation can enable stylistic variation.

3. Method

This study aims to generate continuous dense motion by mapping sparse end-effector locations onto IK goals. One primary expectation is the realism and continuity of the generated movement, ensuring that gaps between given sparse input are filled realistically. Another critical objective is the accurate positioning of end effectors at their intended locations, fulfilling the IK targets. Furthermore, the generated motion must align with the specified LMA-based style descriptors.

Fig. 1 demonstrates the workflow, organized into two layers, each addressing the above-mentioned expectations. The initial layer maps sparse input end-effector positions to dense end-effector positions while obeying LMA-based style descriptors and preserving realism, using a Conditional Variational Autoencoder (CVAE) decoder named *Interpolator*. The second layer uses a Long Short-term Memory (LSTM)-based recurrent neural network, *Synthesizer*, to map dense end-effector positions to human motion. This model focuses on the realism of the motion, IK goals, and compliance with given LMA-based style descriptors.

3.1. Architecture

The two stages serve distinct roles, motivating different architectural choices. The *Interpolator* maps sparse boundary keyframes and a target style descriptor to a dense trajectory. We adopt a Conditional Variational Autoencoder for this stage: the variational autoencoder provides a generative model rather than a deterministic mapping, which is appropriate because the same sparse boundary conditions can correspond to many plausible dense trajectories. The conditional formulation then lets us inject the boundary end-effector positions and

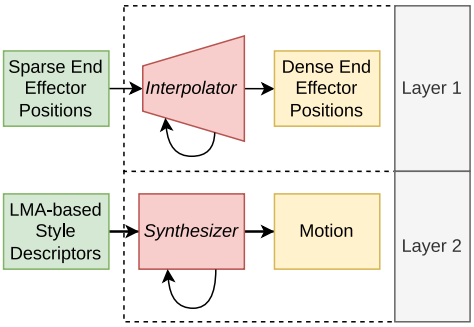

Fig. 1. The two-layer workflow of the proposed approach: first layer: filling the gaps between sparse input, second layer: mapping end-effector positions to human motion.

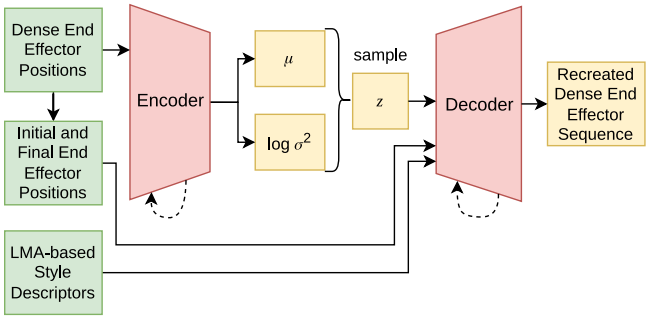

Fig. 2. The proposed CVAE architecture, designed to recreate dense end-effector sequences with conditions of initial and final end-effector positions with LMA-based style descriptors.

the LMA-based style descriptors as inputs that shape generation, so that the Interpolator does not merely produce *some* plausible trajectory but one consistent with the requested boundaries and style. The *Synthesizer*, by contrast, performs a sequence-to-sequence mapping along an already-specified trajectory; pose at any frame needs to be consistent with its near neighbors but not with frames at the opposite end of the window, since long-range temporal context is already encoded in the trajectory provided by the Interpolator. An LSTM captures this local temporal dependency efficiently, and attention-based architectures such as Transformers would not provide additional benefit given the local temporal scope at this stage.

The *Interpolator* model utilizes a CVAE and uses its decoder to generate a sequence of end-effector positions based on the conditions of the initial and final end-effector positions and LMA-based style descriptors. Fig. 2 depicts the proposed architecture. The following configuration is used in order to recreate 50 frames of a motion sequence. The encoder and decoder use a single-layer LSTM network with a hidden state dimension of 512, augmented by dedicated linear layers for input and output mapping, and a latent space dimension of 64. The exact input and output dimensions can vary based on the skeleton configuration.

The *Synthesizer* generates full-body joint angles for a motion sequence using given dense end-effector positions and additional LMA-based style descriptors. It utilizes recurrent networks to capture the temporal nature of motion. Fig. 3 illustrates the proposed architecture. Specifically, the model employs a single-layer LSTM network with a hidden state dimension of 128 and a dense network with three layers of 512 nodes each. The input is fed to the model by concatenating the dense end-effector positions with the specified LMA-based style descriptors. The output is in 6D rotation representation [61], and the output dimension varies depending on skeletal configurations.

A common concern with CVAE-based generation is posterior collapse, in which the decoder ignores the latent variable, leading to a

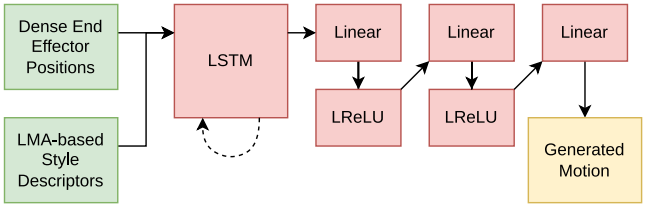

Fig. 3. The proposed Synthesizer architecture, designed to synthesize full-body motion using given dense end-effector positions and LMA-based style descriptors.

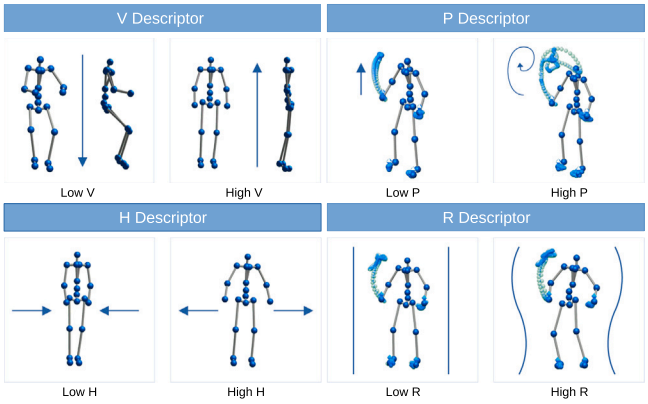

Fig. 4. Summary of the effect of defined style descriptors on final motion generation.

loss of output diversity. We mitigate this in both the Interpolator and the Synthesizer by employing cyclical β-annealing on the KL term [62], which prevents the KL loss from prematurely dominating the objective and forcing the latent towards an uninformative prior. The strong conditioning on boundary end-effector positions and style descriptors further reduces the risk that the decoder will ignore its inputs.

3.2. Style descriptors

Four style descriptors, $\mathscr{V}$ (vertical), $\mathscr{H}$(horizontal), $\mathscr{P}$(pace), and $\mathscr{R}$(regularity), were determined, inspired by LMA Shape and Effort parameters Weight, Space, Time, and Flow, as in Aristidou et al. [6]. Fig. 4 illustrates the effects of the style descriptors on motion generation. Each descriptor maps a motion sequence (in terms of joint positions or angles) to a real number. $\mathscr{V}$ represents the changes in the longitudinal axis. High $\mathscr{V}$ attracts the body upwards, resulting in a rising posture, while Low $\mathscr{V}$ results in a downward, sinking posture. $\mathscr{H}$ represents the changes in the frontal axis; High $\mathscr{H}$ results in a spreading posture while Low $\mathscr{H}$ represents an enclosing posture. $\mathscr{P}$ represents the changes in Time Effort; High $\mathscr{P}$ represents Sudden Time while Low $\mathscr{P}$ represents Sustained Time. $\mathscr{R}$ is inspired by the Flow Effort; High $\mathscr{R}$ corresponds to Free Flow and Low $\mathscr{R}$ represents Bound Flow.

General Notation: Let $P_j(t)$ denote the 3D position vector of joint j at time t. Let J be the set of all joints, and N_J be the total number of joints. Let T be the total number of time steps (frames) in the motion sequence.

3.2.1. $\mathscr{V}$ descriptor

This descriptor quantifies the average angular configuration across 11 key joint triplets. Let $A(P_{j_1}, P_{j_2}, P_{j_3})$ be the angle (in radians) formed by the triplet of joints j_1, j_2, j_3 with j_2 as the center joint. Specifically, for vectors $\mathbf{v}_1 = P_{j_3} - P_{j_2}$ and $\mathbf{v}_2 = P_{j_1} - P_{j_2}$, the angle is $\arccos\left(\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\|\|\mathbf{v}_2\|}\right)$.

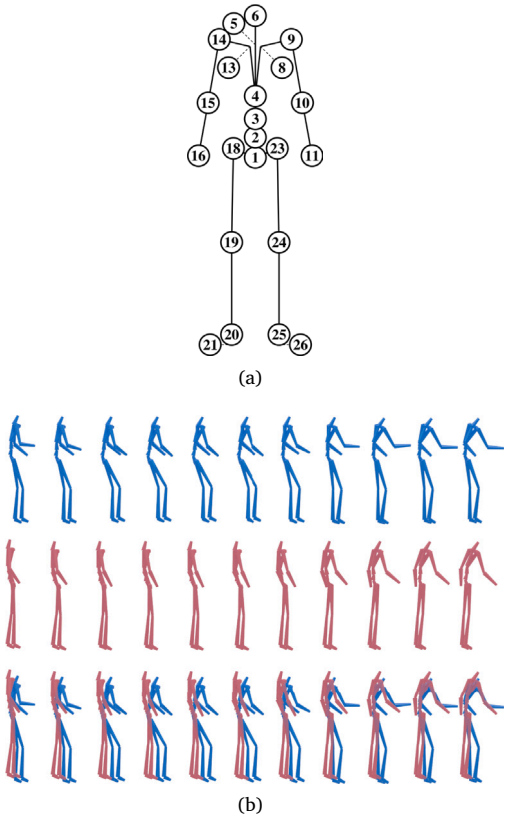

Fig. 5. (a) Visual breakdown of the skeletal structure and (b) comparison of the generated low (blue) and high (red) $\mathscr{V}$ style descriptor poses.

Let $K_{\mathscr{V}}$ be the set of the 11 specified joint triplets.

$$\mathscr{V} = \frac{1}{T \cdot |K_{\mathscr{V}}|} \sum_{t=1}^{T} \sum_{(j_1,j_2,j_3) \in K_{\mathscr{V}}} A(P_{j_1}(t), P_{j_2}(t), P_{j_3}(t)).$$

The triplets are listed below, with their corresponding joint numbers from Fig. 5(a) shown in parentheses.

1. Left hand (11), Left lower arm (10), Left upper arm (9),
2. Right hand (16), Right lower arm (15), Right upper arm (14),
3. Left foot (25), Left lower leg (24), Left upper leg (23),
4. Right foot (20), Right lower leg (19), Right upper leg (18),
5. Head (6), Neck (5), Spine (3),
6. Chest (4), Spine (3), Hips (1),
7. Left lower arm (10), Left upper arm (9), Left shoulder (8),
8. Right lower arm (15), Right upper arm (14), Right shoulder (13),
9. Left hand (11), Hips (1), Right hand (16),
10. Left upper arm (9), Hips (1), Right upper arm (14), and
11. Left foot (25), Hips (1), Right foot (20).

The effect of the $\mathscr{V}$ style descriptor is shown in Fig. 5(b). A high $\mathscr{V}$ value raises the body in the longitudinal axis, while a low value lowers it. Raised postures with elevated chest and lengthened torso have been linked to happiness [4] and extraversion [32], while shortened, dropped postures have been linked to sadness [4], motivating $\mathscr{V}$ as a candidate descriptor for personality and emotion control.

3.2.2. $\mathscr{H}$ descriptor

This descriptor quantifies the average spatial extent of the body by computing L2 distances between 13 specific joint pairs. While traditional LMA distinguishes between the lateral (side-to-side) and sagittal

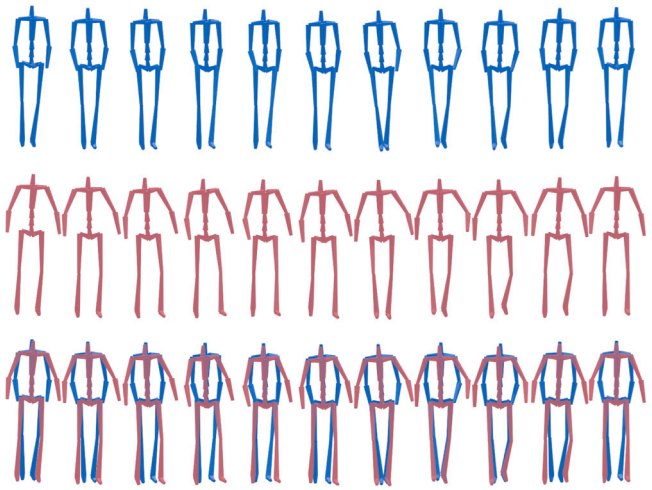

Fig. 6. Comparison of the generated low (blue) and high (red) $\mathscr{H}$ style descriptor poses.

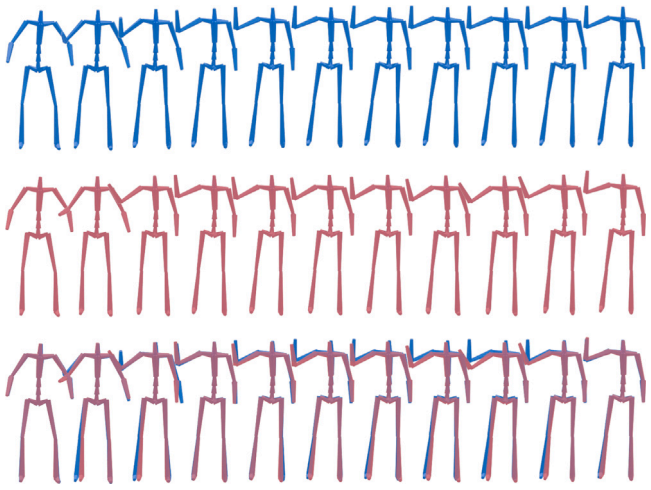

Fig. 7. Comparison of the generated low (blue) and high (red) $\mathscr{P}$ style descriptor poses.

(forward-and-back) dimensions of the horizontal plane, our formulation calculates the 3D Euclidean distance between opposing limbs. This effectively captures the overall volumetric expansion or contraction of the body's posture. Let $D(P_{j_1}, P_{j_2})$ denote the L2 (Euclidean) distance between joints j_1 and j_2, i.e., $\|P_{j_1} - P_{j_2}\|$. Let $K_{\mathscr{H}}$ be the set of the 13 specified joint pairs.

$$\mathscr{H} = \frac{1}{T \cdot |K_{\mathscr{H}}|} \sum_{t=1}^{T} \sum_{(j_1,j_2)\in K_{\mathscr{H}}} D(P_{j_1}(t), P_{j_2}(t))$$

The pairs are as follows, with their corresponding joint numbers from Fig. 5a shown in parentheses.

1. Left hand (11), Right hand (16),
2. Left lower arm (10), Right lower arm (15),
3. Left upper arm (9), Right upper arm (14),
4. Left foot (25), Right foot (20),
5. Left lower leg (24), Right lower leg (19),
6. Left upper leg (23), Right upper leg (18),
7. Head (6), Left hand (11),
8. Head (6), Right hand (16),
9. Hips (1), Head (6),
10. Hips (1), Left hand (11),
11. Hips (1), Right hand (16),
12. Left lower arm (10), Right hand (16), and
13. Right lower arm (15), Left hand (11).

Fig. 6 illustrates the effects of varying the $\mathscr{H}$ style descriptor. A higher $\mathscr{H}$ value produces a wider body layout along the frontal axis, whereas a lower value yields a more compact posture. Wider, spreading body layouts have been linked to happiness [4] and to extraversion through broad gestures and outward orientation [7,32], motivating $\mathscr{H}$ as a control parameter for both personality and emotion expression.

3.2.3. $\mathscr{P}$ descriptor

This descriptor quantifies the overall pace and tempo of the movement by computing the average speed for every individual joint. Let $V_j(t) = \|P_j(t) - P_j(t-1)\|$ be the speed (L2 norm of displacement) of joint j at time t. For $t = 1$, $P_j(0)$ is typically taken as $P_j(1)$ or the speed is set to zero for that initial frame.

$$\mathscr{P} = \frac{1}{T \cdot N_J} \sum_{t=1}^{T} \sum_{j\in J} V_j(t).$$

Fig. 7 illustrates how varying the $\mathscr{P}$ style descriptor primarily affects the speed of motion in the active limbs, while the rest of the body remains relatively stable. As the $\mathscr{P}$ value increases, the end-effectors move faster and cover a larger range of motion. Conversely, lower values result in slower and more restricted movements. Faster movement has been associated with higher extraversion and neuroticism in prior work [7,32], and is also among the most reliably perceived motion qualities [13], motivating $\mathscr{P}$ as a stylization parameter for personality control.

3.2.4. $\mathscr{R}$ descriptor

This descriptor is associated with the geodesic distance between the predicted and ground-truth joint rotations. It quantifies how accurately and consistently the predicted joint rotations match the true rotations.

Let $R_j^{pred}(t)$ be the predicted 3×3 rotation matrix for joint j at time t. Let $R_j^{gt}(t)$ be the ground truth 3×3 rotation matrix for joint j at time t. The geodesic distance $d_{geo}(M_1, M_2)$ between two rotation matrices M_1 and M_2 is defined as [63]:

$$d_{geo}(M_1, M_2) = \cos^{-1}\left(\frac{\mathrm{tr}(M_1 M_2^T) - 1}{2}\right).$$

To capture the standard deviation over the J dimension for each time step t, we first define the mean geodesic difference at time t as

$$\mu_J(t) = \frac{1}{N_J} \sum_{j\in J} d_{geo}(R_j^{pred}(t), R_j^{gt}(t)).$$

The standard deviation for joint differences at time t is then

$$\sigma_J(t) = \sqrt{\frac{1}{N_J} \sum_{j\in J} \left(d_{geo}(R_j^{pred}(t), R_j^{gt}(t)) - \mu_J(t)\right)^2}.$$

The updated $\mathscr{R}$ descriptor, incorporating this variability, is

$$\mathscr{R} = \frac{1}{T \cdot N_J} \sum_{t=1}^{T} \sum_{j\in J} d_{geo}(R_j^{pred}(t), R_j^{gt}(t)) + \lambda \frac{1}{T} \sum_{t=1}^{T} \sigma_J(t),$$

where λ is a hyperparameter balancing the contribution of the average difference and its joint-wise variability and is empirically determined as $\lambda = 1$. The comparison of motions generated with lower and higher values of the $\mathscr{R}$ descriptor is shown in Fig. 8. Higher $\mathscr{R}$ values lead to poses that deviate from those in the dataset, resulting in more fluid, less constrained movements. In contrast, lower $\mathscr{R}$ values yield poses that closely resemble those found in the dataset. Unlike the previous three descriptors, $\mathscr{R}$ does not have a clean precedent in prior body-expression literature; it functions as an exploratory data-driven control over how closely generated poses follow the training data, with its perceptual correlates established empirically in Section 4.

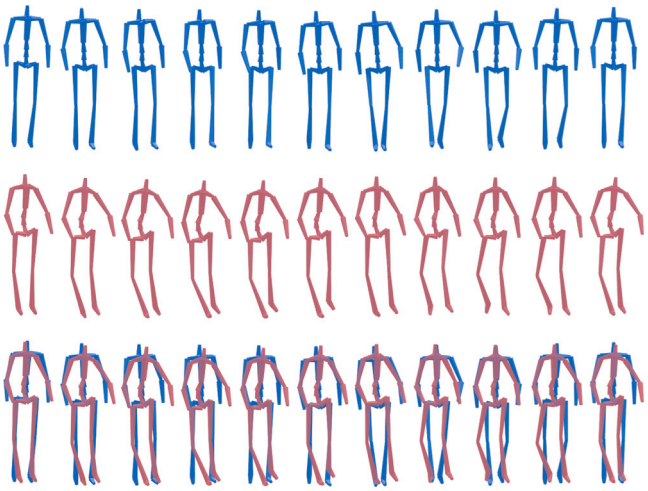

Fig. 8. Comparison of the generated low (blue) and high (red) $\mathscr{R}$ style descriptor poses.

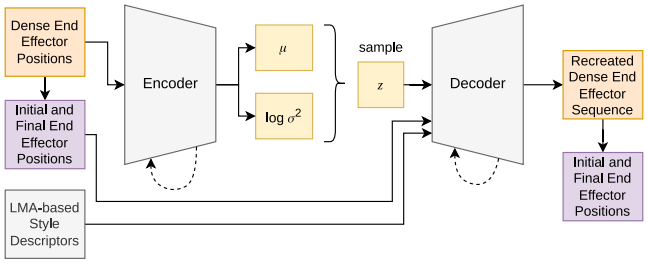

Fig. 9. The *Interpolator* training diagram: colors other than yellow (KL divergence loss) indicate a Mean Squared Error (MSE) loss between the same colored variables.

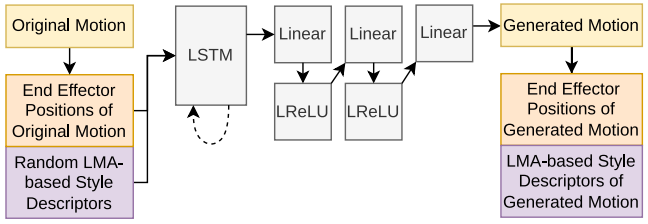

Fig. 10. The *Synthesizer* training diagram: yellow indicates geodesic distance loss, orange indicates MSE loss, and purple indicates L1 loss.

3.3. Loss functions

Our system employs two distinct neural architectures with specialized loss functions: *Interpolator* for trajectory interpolation and a *Synthesizer* for inverse kinematics-based pose generation. Both models are trained on motion databases consisting of Biovision Hierarchy (BVH) files, with each database producing a model trained on a normalized skeleton specific to that dataset. Figs. 9 and 10 show the Interpolator and Synthesizer training diagrams, respectively.

3.3.1. Interpolator loss

Interpolator operates on 3D end effector position trajectories and uses cyclical beta annealing for improved latent space regularization. The training process extracts data from motion sequences, including

- P_{seq}, the 3D site positions across the sequence, and
- C_{style}, the LMA-based style descriptors including $\mathscr{V}$, $\mathscr{H}$, and $\mathscr{P}$ parameters normalized to $[0, 1]$.

Table 1
Final *Synthesizer* training and validation losses.

Split	$\mathscr{L}_{total}$	$\mathscr{L}_{IK}$	$\mathscr{L}_{Sim}$	$\mathscr{L}_{\mathscr{V}}$	$\mathscr{L}_{\mathscr{H}}$	$\mathscr{L}_{\mathscr{P}}$	$\mathscr{L}_{\mathscr{R}}$
Train	0.9638	0.000585	0.0766	0.0007	0.0009	0.0270	0.0409
Valid	0.8447	0.000555	0.0734	0.0007	0.0010	0.0093	0.0165

The *Interpolator* generates reconstructed sequences P_{recon} and latent parameters μ and σ. The total loss is calculated as

$$\mathscr{L}_{CVAE} = \lambda_{recon}\mathscr{L}_{recon} + \beta(t)\mathscr{L}_{KLD},$$

where,

- *Reconstruction Loss*, $\mathscr{L}_{recon}$, uses MSE between reconstructed and target sequences, with additional emphasis on start and end frame accuracy: $\mathscr{L}_{recon}$+10$\mathscr{L}_{start}$+10$\mathscr{L}_{end}$,
- *KL Divergence Loss*, $\mathscr{L}_{KLD}$, enforces a smooth latent space distribution, and
- $\beta(t)$ implements cyclical annealing [62], varying between 0.0 and 0.01 over training cycles to balance reconstruction fidelity and latent space regularity.

After training for 800 epochs, the *Interpolator* converged with a final reconstruction loss of 0.0106 and a KL divergence loss of 1.9811 on the training set, while the corresponding validation losses were 0.0128 and 1.9827, respectively. These results indicate that the model achieved accurate trajectory reconstruction with well-regularized latent representations.

3.3.2. Synthesizer loss

The *Synthesizer* performs inverse kinematics by predicting joint rotations from 3D trajectory constraints and LMA-based style conditions. During training, the system extracts the following:

- P_{sites}, the mean-centered 3D positions of specific body sites,
- Θ_{target}, target joint rotations as 6D rotation representations [61], and
- $C_{Laban} = [\mathscr{V}, \mathscr{H}, \mathscr{P}, \mathscr{R}]$, LMA-based style descriptors.

The synthesizer generates Θ_{pred}, which, through forward kinematics, produces P_{recon} and derived LMA-based style descriptors. The comprehensive loss function is

$$\mathscr{L}_{synth} = \lambda_{IK}\mathscr{L}_{IK} + \lambda_{synth}\mathscr{L}_{synth} + \sum_{i\in\{\mathscr{V},\mathscr{H},\mathscr{P},\mathscr{R}\}} \lambda_i\mathscr{L}_i.$$

The individual loss components are as follows.

- *Inverse Kinematics Loss*, $\mathscr{L}_{IK}$, ensures generated end-effector positions match target trajectories using MSE between mean-centered site positions.
- *Angular Similarity Loss*, $\mathscr{L}_{synth}$, employs *geodesic distance* [63] to measure rotational differences between predicted and target joint angles.
- $\mathscr{V}$, $\mathscr{H}$, $\mathscr{P}$, $\mathscr{R}$ *Losses* $\mathscr{L}_{\mathscr{V}}$, $\mathscr{L}_{\mathscr{H}}$, $\mathscr{L}_{\mathscr{P}}$, and $\mathscr{L}_{\mathscr{R}}$ are L1 losses between computed and target style descriptor values.

The loss weights are empirically determined: $\lambda_{IK} = 40.0$, $\lambda_{synth} = 1.0$, $\lambda_W = 1.0$, $\lambda_S = 1.0$, $\lambda_T = 1.0$, and $\lambda_F = 2.0$. During training, noise is injected into $\mathscr{V}$ and $\mathscr{H}$ style descriptor values (± 0.2 uniform noise) while $\mathscr{R}$ style descriptor values are randomized to improve generalization. Both models employ *AdamW* [64] optimization with learning rate scheduling via *ReduceLROnPlateau* [65] for stable convergence. Table 1 summarizes the final training and validation losses for each component of the *Synthesizer*, demonstrating consistent convergence across inverse kinematics, angular similarity, and style descriptor terms.

3.4. Motion generation

The motion generation system operates through a two-stage pipeline that transforms user-specified 3D trajectory constraints and LMA-based style descriptors into full-body motion sequences. This system separates trajectory interpolation from pose synthesis, enabling more precise control over the spatial and stylistic aspects of the generated motion.

3.4.1. Trajectory interpolation

The first stage employs the trained *Interpolator* to generate smooth trajectory sequences between user-defined keyframes. Users specify sparse trajectory points for the four body sites (hands and feet), and the system generates complete trajectory sequences of 50 frames. The interpolation process operates recursively for longer sequences. Given start and end poses, the *Interpolator* generates intermediate trajectory segments, where each segment's end pose becomes the start pose for the subsequent segment. This chunked generation approach maintains temporal consistency across extended motion sequences while respecting the model's fixed sequence length constraint. During generation, LMA-based style descriptors are incorporated directly into the CVAE decoder.

3.4.2. Pose synthesis

The second stage transforms the interpolated 3D trajectories into full-body pose sequences using the trained *Synthesizer*. This inverse kinematics approach takes the generated site trajectories and produces joint rotations that satisfy both kinematic constraints and LMA-based style descriptors. The synthesis process operates frame-by-frame, simultaneously predicting 6D rotation representations for all joints. The input consists of mean-centered site positions (to achieve translation invariance) concatenated with the four LMA condition parameters. The LSTM-based architecture processes the entire sequence, ensuring temporal consistency in the generated poses.

Root translation is handled through a two-step process. First, the *Synthesizer* operates on normalized input data to remove global translation, primarily predicting relative joint rotations. Initially, joint positions are computed assuming the root joint remains at a fixed origin. Input end-effector positions are utilized to re-establish the global trajectory. The mean global position of input end-effector positions is compared to that of generated end-effector positions. The difference between these two mean paths yields the necessary absolute root translation correction for each frame, which is then applied to all generated joint positions, thus aligning the entire synthesized skeleton with the global path of the input.

The absolute root translation for each frame t, denoted as $T_{\text{root}}(t)$, is calculated as

$$T_{\text{root}}(t) = \left(\frac{1}{|S|} \sum_{s \in S} P_s^{\text{input}}(t) \right) - \left(\frac{1}{|S|} \sum_{s \in S} P_s^{\text{generated}}(t) \right),$$

where $P_s^{\text{input}}(t)$ is the absolute position of the end effector joint s from the input trajectory at time t, $P_s^{\text{generated}}(t)$ is the position of end effector joint s computed by forward kinematics (assuming a fixed root at the origin) at time t, and S is the set of end effectors. The final absolute world position $P_j^{\text{final}}(t)$ for any joint j at time t is then

$$P_j^{\text{final}}(t) = P_j^{\text{generated}}(t) + T_{\text{root}}(t).$$

3.5. Quantitative evaluation

We employ Fréchet Inception Distance (FID) [14] adapted for motion data to quantitatively assess the quality and diversity of our generated motion.

Table 2
FID scores comparing generated and real motion samples. Lower values indicate greater similarity to real motion, while higher values indicate greater deviation.

Category	FID score	Samples
Overall	29.67	32
$\mathscr{H}$	27.49	8
$\mathscr{P}$	28.47	8
$\mathscr{V}$	40.86	8
$\mathscr{R}$	70.78	8

3.5.1. Motion embedding

We train a Variational Autoencoder (VAE) to learn a compact, style-preserving embedding space for motion sequences. The encoder maps motion sequences to a 256-dimensional latent space, serving as our feature representation for calculating the FID.

The VAE operates on joint angles in 6D rotation representation [61]. The encoder uses an LSTM (512 hidden units) to output latent distribution parameters $\mu, \log \sigma^2$. The decoder reconstructs motions from latent distribution samples $z \sim \mathcal{N}(\mu, \sigma^2)$.

The training is also conducted on the LMA Effort Database [13] with loss $\mathscr{L} = \mathscr{L}_{\text{recon}} + \beta \cdot \mathscr{L}_{\text{KLD}}$, where $\mathscr{L}_{\text{recon}}$ is geodesic loss and $\mathscr{L}_{\text{KLD}}$ is KL divergence. We employ cyclical β-annealing [62] (10 cycles, β: $0.0 \rightarrow 0.01$) to prevent posterior collapse. After 200 epochs, the model achieves a validation loss of 19.28 radians ($\approx 1.1°$ per joint) with a stable KLD of 495.89.

3.5.2. Fréchet Inception Distance (FID)

For each motion, we extract 50-frame windows (with a stride of 25) and average their embeddings. Given real embeddings $\{\mathbf{r}_i\}$ and generated embeddings $\{\mathbf{g}_j\}$, we calculate FID as:

$$\text{FID} = \|\mu_r - \mu_g\|_2^2 + \text{Tr}\left(\Sigma_r + \Sigma_g - 2\left(\Sigma_r \Sigma_g\right)^{1/2} \right),$$

where μ and Σ are sample means and covariances.

Table 2 shows FID scores based on 141 real and 32 generated samples used in our study. Our model achieves an overall FID of 29.67, suggesting that the generated samples resemble real motions in the embedding space, with some variance. Lower FID scores for $\mathscr{H}$ (27.49) and $\mathscr{P}$ (28.47) indicate that generated motions in these categories exhibit both realistic kinematics and sufficient variability. In contrast, $\mathscr{V}$ (40.86) and especially $\mathscr{R}$ (70.78) suggest larger discrepancies from real data, likely due to exaggerated styling, which is related to the types of losses used for each descriptor. These results highlight potential directions for improvement; furthermore, given the relatively small sample size (32 generated sequences), the reported FIDs should be interpreted as indicative rather than absolute measures of distributional similarity.

3.6. Inference performance

We profiled the inference latency of the trained models on a passively cooled MacBook Air (Apple M4). Per-call latencies for the Interpolator, the Synthesizer, and the end-to-end pipeline are summarized in Table 3. The end-to-end pipeline averages 11.2 ms (p95: 13.0 ms), corresponding to approximately 90 Hz on this hardware, which is comfortably within typical motion-synthesis budgets for live VR scenarios where end-effector keyframes are streamed from positional trackers.

4. User study

4.1. Stimuli

We generate the stimuli for our user study using four motion-capture categories: walk, wave, sit, and put. The selected animations

Table 3
Inference latency of the proposed pipeline on a passively-cooled MacBook Air (Apple M4).

Stage	No. of parameters	Mean (ms)	p95 (ms)	Frame rate (fps)
Interpolator	3.30 million	5.9	7.7	169
Synthesizer	5.62 million	6.6	7.2	151
End-to-end pipeline	–	11.2	13.0	90

Table 4
Survey questions used in the evaluation.

No.	Dimension	Traits
Q1	Extraversion	Extraverted, enthusiastic
Q2	Agreeableness[a]	Critical, quarrelsome
Q3	Conscientiousness	Dependable, self-disciplined
Q4	Emotional stability[a]	Anxious, easily upset
Q5	Openness	Critical, quarrelsome
Q6	Extraversion[a]	Reserved, quiet
Q7	Agreeableness	Sympathetic, warm
Q8	Conscientiousness[a]	Disorganized, careless
Q9	Emotional stability	Calm, emotionally stable
Q10	Openness[a]	Conventional, uncreative
Q11	Happiness	Happy
Q12	Sadness	Sad
Q13	Fear	Afraid
Q14	Disgust	Disgusted
Q15	Anger	Angry
Q16	Surprise	Surprised
Q17	Human likeness	Human like, realistic

[a] The dimensions are reverse-scored.

involve active hand and leg movements with room for personality-related changes. Each sample produces eight alternatives, two polarities (low and high) for each of the style descriptors $\mathscr{V}$, $\mathscr{H}$, $\mathscr{P}$, and $\mathscr{R}$. We compare opposing polarities; four tasks per style descriptor yield 16 tasks in total. We focus on individual style changes to limit the resulting output combinations to manageable amounts. For example, we input the same base animation to synthesize versions with high and low $\mathscr{V}$, which are compared for personality and emotional expression. The generated animations are rendered using stick figures [66].

4.2. Design

We designed an online user study[3] to compare our system outputs regarding personality and emotion. We kept the tasks simple so that participants could focus better on the subject without fatigue. Each task displays a stimulus pair in a loop and includes 17 questions (see Table 4) that compare the two animations. The animations are rendered in real time using the same skeleton, and the participants can manage the orbiting camera around each figure. The left–right placement of the high and low samples is randomized for each task.

We use the Ten-Item Personality Inventory (TIPI) [67] to compare the high and low versions of personality expression. We also compare the output samples with respect to the six universal emotions [68] and human likeness. Participants can view and change their answers before submitting, but they cannot return to a previously submitted task. We randomized the order of appearance of the focused style descriptors. We did not inform the participants of the changes made to the animations.

A screenshot from the user study is included in Fig. 11. We have included the survey questions in a scrollable division, allowing participants to easily view the samples while answering the questions. Each participant answers the 17 questions in 16 tasks (4 motion styles × 4 style descriptor pairs compared) before being directed to the completion page. The average time to complete the experiment was 26 min.

[3] Bilkent University, Ethics Committee for Research Projects Involving Human Participants approved the user study with Decision Number 622 at the meeting İAEK_2024_12_06_01.

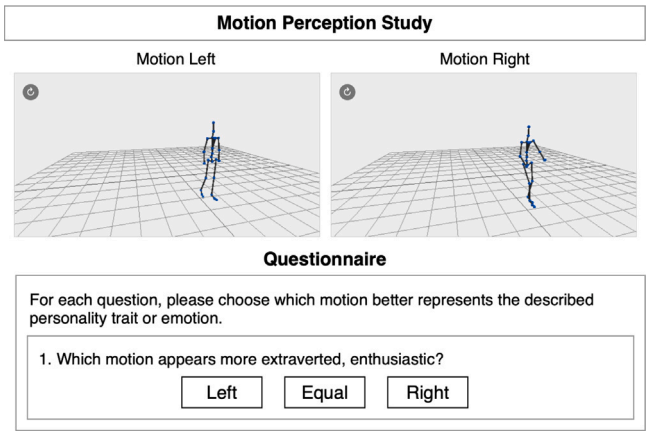

Fig. 11. The user study screenshot comparing altered versions of the same input animation: 17 questions measure the psychological motion attributes, and one additional question per sample measures the perceived human likeness.

We map the user answers for each question to a 3-point Likert scale where low (−1), equal (0), and high (1) choices are represented. If the changes due to a specific style descriptor influence a measured aspect, we expect a significant preference for that sample. For example, if the High $\mathscr{V}$ sample represents extraverted traits, most participants will choose that sample for the high extraversion question. TIPI includes two questions for each personality factor: one measures the trait directly, and the other measures it inversely. We expect the choices for these reversed questions to correspond to opposing polarities when a certain descriptor significantly alters the perception of the related trait.

4.3. Demographics

We recruited participants from Cloud Research's crowd-sourcing platform Connect; 31 unique participants (18 female, 13 male) contributed to our study. We did not have any age, education, or employment restrictions, as our study relied on the general perception of common traits among individuals. The only precondition was fluency in English, the language of the study questions. We did not require participants to have any prior experience with computer animation. The average participant age was 37.38 ± 9.18. The participants were from diverse countries, including 17 from North America, 10 from Europe, and 4 from Africa.

4.4. Analysis

4.4.1. Style descriptor effects

Fig. 12 depicts the number of participants' low, equal, and high choices per question for each style descriptor's effect on different motion types. We generally observe that the type of motion has a minimal effect on user response, with the primary effect arising from the style descriptor. However, in certain cases, such as $\mathscr{V} - Q3$, specific motion styles yield a different perception. For example, the effect of the style descriptor $\mathscr{V}$ for the sitting motion is reversed in this case. The figure also depicts the human-likeness scores on a 5-point Likert scale for the low and high variants. We generally observe a similar effect on realism except for $\mathscr{V}$.

We utilize Intraclass Correlation Coefficients (ICC) to measure the agreement of participant answers in Table 5. $ICC > .4$ is generally considered a fair inter-rater agreement, and $ICC > .75$ is excellent. Small, negative ICC values indicate no agreement between the user choices for that specific question and style descriptor combination.

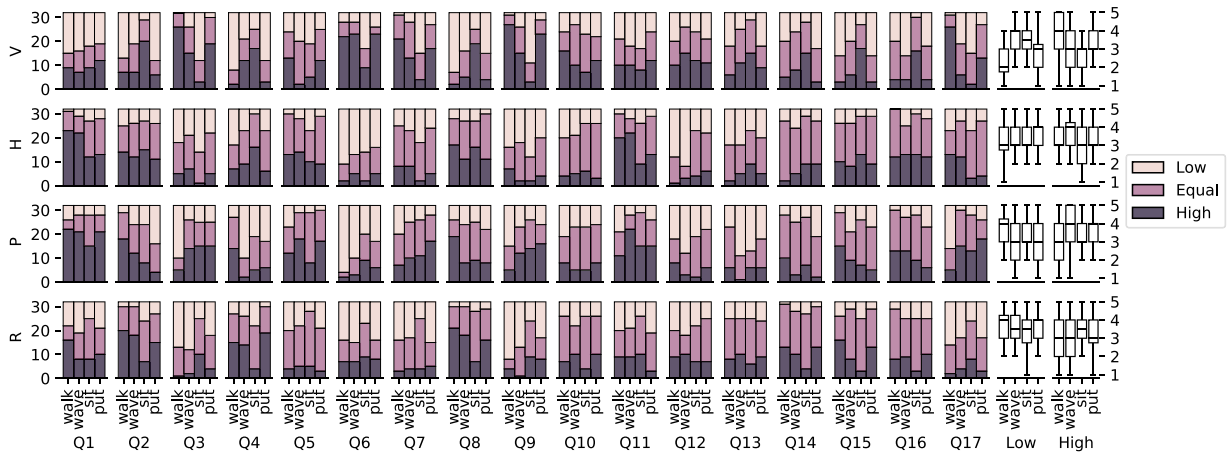

Fig. 12. Histogram plots of the user answers per question (Q1–17) and style descriptor combination for each motion style, including human-likeness measures for Low and High polarities on a 5-point Likert scale (5: Excellent, 1: Poor).

Table 5
Intraclass Correlation Coefficient (ICC) and the corresponding p-values for each question and style descriptor effect, measuring the agreement of the participant choices for different motion types. gray cells indicate $p < 0.05$ and dark-gray cells indicate $ICC > 0.75$.

No.	$\mathscr{V}_{ICC}$	$\mathscr{V}_{\rho}$	$\mathscr{H}_{ICC}$	$\mathscr{H}_{\rho}$	$\mathscr{P}_{ICC}$	$\mathscr{P}_{\rho}$	$\mathscr{R}_{ICC}$	$\mathscr{R}_{\rho}$
Q1	−.733	.631	.840	.001	−.251	.722	.025	.385
Q2	.906	.001	−.065	.806	.894	.001	.783	.005
Q3	.961	.001	.660	.037	.889	.001	.870	.001
Q4	.920	.001	.857	.001	.897	.001	.831	.001
Q5	.759	.008	.450	.150	.638	.046	.344	.213
Q6	.872	.001	.448	.150	.848	.001	.115	.341
Q7	.923	.001	.575	.077	.724	.016	.362	.203
Q8	.916	.001	−.347	.530	.655	.039	.774	.006
Q9	.962	.001	.331	.221	.764	.007	.853	.001
Q10	−.217	.485	−.042	.415	−.414	.830	−.404	.742
Q11	−.031	.411	.813	.002	.682	.029	.559	.086
Q12	.044	.376	.880	.001	.537	.098	−.530	.909
Q13	.791	.004	.577	.076	.712	.019	−.935	.894
Q14	.845	.001	.540	.097	.764	.007	.735	.013
Q15	.914	.001	.118	.340	.747	.011	.726	.015
Q16	.912	.001	−.295	.512	.645	.043	.126	.335
Q17	.956	.001	.658	.038	.886	.001	.671	.033

Table 6
T-test results per style descriptor and question comparing the answers to uniform randomness. We display the Benjamini–Hochberg adjusted p-values; dark-gray cells indicate $p < 0.05$. The gray cells were previously significant but were dropped after adjustment. The dashed cells are ignored because $ICC < .4$. Δ shows the mean of the distribution; a value close to 1 represents a choice towards the high variant, and vice versa.

No.	$\mathscr{V}_{\Delta}$	$\mathscr{V}_{\rho}$	$\mathscr{H}_{\Delta}$	$\mathscr{H}_{\rho}$	$\mathscr{P}_{\Delta}$	$\mathscr{P}_{\rho}$	$\mathscr{R}_{\Delta}$	$\mathscr{R}_{\rho}$
Q1	–	–	0.45	.001	–	–	–	–
Q2	−0.12	.453	–	–	0.05	.601	0.34	.003
Q3	0.27	.036	−0.27	.010	0.05	.601	−0.34	.003
Q4	−0.22	.104	0.02	.813	−0.22	.072	0.23	.045
Q5	−0.06	.581	0.22	.036	0.30	.007	–	–
Q6	0.38	.006	−0.48	.001	−0.45	.001	–	–
Q7	0.22	.100	−0.12	.263	0.12	.351	–	–
Q8	−0.27	.036	–	–	0.10	.435	0.40	.001
Q9	0.29	.036	–	–	0.05	.601	−0.34	.003
Q10	–	–	–	–	–	–	–	–
Q11	–	–	0.38	.001	0.30	.007	−0.09	.545
Q12	–	–	−0.38	.001	−0.30	.007	–	–
Q13	0.02	.817	−0.23	.030	−0.34	.003	–	–
Q14	−0.06	.581	0.03	.788	−0.05	.601	0.22	.039
Q15	−0.19	.153	–	–	0.05	.601	0.16	.136
Q16	−0.14	.293	–	–	0.16	.175	–	–
Q17	0.09	.526	0.03	.788	0.16	.196	−0.30	.004

In this case, either the style descriptor does not affect the measured aspect or results in different interpretations. Since we focus on the more universal effect, we consider only cases where $ICC > .4$.

We mainly observe that participants mostly agree on the effect of $\mathscr{V}$ and $\mathscr{P}$. $\mathscr{H}$ and $\mathscr{R}$ received more indecisive answers. The personality-related questions show greater agreement than the emotion-related ones. The reverse scored openness question (Q10) received the least overall agreement. On the other hand, the reverse-scored emotional stability question (Q4) shows the highest agreement.

We report the T-test results for each question and style descriptor in Table 6, grouping answers by motion style. For example, the user choices for the style descriptor $\mathscr{V}$ are similar across the different actions of walking, waving, sitting, and putting. Therefore, the overall effect is clearer when the answers are combined. For each question, we compare the answer means to the uniform random distribution, where each choice has a 33.3% chance of being selected. To correct for False Discovery Rate (FDR), we apply the Benjamini–Hochberg p-value adjustment.

The results suggest that $\mathscr{V}$ mostly influences personality. High $\mathscr{V}$ is perceived as high in conscientiousness, agreeableness, and emotional stability, and low in extraversion. $\mathscr{H}$ influences both personality and emotion perception. High $\mathscr{H}$ is perceived as high in extraversion and openness, and low in conscientiousness. It is also perceived as high in happiness and low in sadness and fear. $\mathscr{P}$ has a similar effect on the measurements. High $\mathscr{P}$ is associated with high emotional stability, openness, and extraversion, as well as high happiness and low sadness and fear. High $\mathscr{R}$ is associated with low agreeableness, conscientiousness, and emotional stability. It is also perceived as high disgust. The only style descriptor that causes a significant difference for human likeness is $\mathscr{R}$. A high $\mathscr{R}$ is perceived as having low human likeness.

We report ANOVA statistics that compare the effect due to different motion styles for each question and style descriptor in Table 7, for cases where $p < 0.05$. For each case, we also report the individual means for each motion style. We label these cases based on the highlighted motion style, indicating they differ significantly from a random distribution. We label the cases where different motion styles agree as *same*. If disagreement exists for at least one motion style, we use the label *inverse*, which means at least two significant means have the inverse sign. We exclude cases where there is only one significant motion style.

The cases labeled with *same* do not negatively impact the overall effect. In such cases, the difference in means is likely due to differences in range of motion across motion categories. For example, in the case of *Q1-*$\mathscr{H}$, sitting and putting motions give less freedom to the model, so these two motion styles result in less difference. In contrast, the walking

Table 7
ANOVA statistics of the cases with $p < 0.05$. We report the mean distributions for each motion style. The highlighted motion styles significantly differ from random choice ($T - test_p < 0.05$). We label the cases as *same* when the means agree, *inverse* when they disagree.

Case	F	ρ	Δ_{Walk}	Δ_{Wave}	Δ_{Sit}	Δ_{Put}	Label
Q1-$\mathscr{H}$	4.002	.009	0.69	0.59	0.22	0.28	Same
Q2-$\mathscr{V}$	10.594	.001	−0.38	−0.19	0.53	−0.44	Inverse
Q2-$\mathscr{R}$	5.071	.002	0.56	0.50	−0.03	0.31	Same
Q3-$\mathscr{V}$	26.964	.001	0.81	0.28	−0.53	0.53	Inverse
Q3-$\mathscr{R}$	7.137	.001	−0.56	−0.56	0.09	−0.31	Same
Q4-$\mathscr{V}$	12.750	.001	−0.69	0.03	0.31	−0.53	Same
Q5-$\mathscr{P}$	2.793	.043	0.09	0.47	0.16	0.47	Same
Q6-$\mathscr{V}$	7.531	.001	0.56	0.59	−0.19	0.53	Same
Q6-$\mathscr{P}$	6.581	.001	−0.81	−0.59	−0.09	−0.28	Same
Q7-$\mathscr{V}$	13.507	.001	0.62	0.28	−0.41	0.38	Inverse
Q8-$\mathscr{V}$	13.070	.001	−0.72	−0.34	0.38	−0.41	Same
Q8-$\mathscr{R}$	3.883	.011	0.59	0.50	0.09	0.41	Same
Q9-$\mathscr{V}$	27.901	.001	0.81	0.28	−0.56	0.62	Inverse
Q9-$\mathscr{R}$	5.870	.001	−0.62	−0.56	0.03	−0.22	Same
Q11-$\mathscr{H}$	3.658	.014	0.56	0.56	0.09	0.31	Same
Q12-$\mathscr{H}$	6.190	.001	−0.59	−0.66	−0.16	−0.12	Same
Q15-$\mathscr{V}$	10.074	.001	−0.47	−0.19	0.38	−0.47	Same
Q16-$\mathscr{V}$	10.835	.001	−0.25	−0.44	0.44	−0.31	Inverse
Q17-$\mathscr{V}$	22.772	.001	0.78	−0.22	−0.47	0.25	Inverse
Q17-$\mathscr{P}$	9.223	.001	−0.41	0.41	0.28	0.38	Inverse

Table 8
Measurements affected by style descriptors; the non-empty cells show the polarity of the parameter to achieve high values of that measurement. The cells indicating 'High' show direct proportion, and those indicating 'Low' show inverse proportion. The highlighted measurements show the strong effects.

Measurement	$\mathscr{V}$	$\mathscr{H}$	$\mathscr{P}$	$\mathscr{R}$
Openness	–	High	High	–
Conscientiousness	High[a]	Low	–	Low
Extraversion	Low	High	High	–
Agreeableness	–	–	–	Low
Emotional stability	High[a]	–	–	Low
Happiness	–	High	High	–
Sadness	–	Low	Low	–
Fear	–	Low	Low	–
Disgust	–	–	–	High
Human likeness	–	–	–	Low

[a] Cells have an inverse effect on certain motion styles.

motion, in which legs and arms play an active role, enables the model to produce a readily perceptible effect.

On the other hand, the cases labeled with *inverse* require a closer inspection, as the corresponding style descriptor causes opposing effects. For example, in the case of *Q2-*$\mathscr{V}$, the vertical adjustments of $\mathscr{V}$ cause a high perception for the sitting action while the general trend is in the negative direction. This behavior diminishes the overall perception; we do not observe a significant effect for *Q2-*$\mathscr{V}$ in Table 6. However, in certain cases, such as *Q9-*$\mathscr{V}$, the remaining motion styles dominate the inverse effect on the sitting action, and a significant effect is still observed in the grouped pairwise comparison. While deciding on the overall effect of each style descriptor, we mark the significant effects with *inverse* label for future inspection.

4.4.2. Overall effect

Table 8 presents the overall effect of each style descriptor. For each measurement, we indicate whether the high or low alternative represents the high values of the corresponding trait.

The results suggest that our style descriptors influence different perceptual measurements. $\mathscr{V}$ directly correlates with conscientiousness and emotional stability and is inversely related to extraversion. This style descriptor's effect on conscientiousness and emotional stability is opposite to its effect on sitting motion. This behavior is likely as $\mathscr{V}$ adjusts the general posture in the vertical axis, which could alter the semantic meaning of the sitting action. $\mathscr{H}$ directly correlates with openness, extraversion, and happiness, and is inversely related to conscientiousness, sadness, and fear. The effect of $\mathscr{P}$ is very similar to $\mathscr{H}$, except that $\mathscr{P}$ does not significantly impact conscientiousness, and its effect on extraversion is weaker. The impact of $\mathscr{R}$ is significantly different from that of the remaining style descriptors. $\mathscr{R}$ has a direct proportion with disgust, and an indirect relation with conscientiousness, agreeableness, emotional stability, and human likeness. The effect on human likeness is expected, as this style descriptor aims to create an animation that deviates from the dataset's rules.

4.4.3. Discussion

Our results suggest that the proposed model parameters successfully distinguish between high and low variants across different affect measures. $\mathscr{V}$ influences conscientiousness, extraversion, and emotional stability. This style descriptor can be used to alter perceptions of personality without influencing emotion. However, since this parameter primarily adjusts the pose along the vertical axis, certain motion styles, such as sitting, yield unexpected results.

The influences of $\mathscr{H}$ and $\mathscr{P}$ are very similar. Both influence openness, extraversion, happiness, sadness, and fear perception. $\mathscr{H}$ also affects conscientiousness. $\mathscr{R}$ affects conscientiousness, agreeableness, emotional stability, and disgust. This style descriptor also influences human likeness, which can be a side effect of not following regular motion. However, it is also possible that creating a strong effect on emotional stability requires certain differences in human likeness.

Compared to the literature-based predictions made when the descriptors were introduced, our findings only partially confirm them. The associations of $\mathscr{H}$ with extraversion and happiness match prior work [4,7,32]. For $\mathscr{V}$, the predicted link between raised posture and extraversion [32] inverts: a raised body reads instead as conscientious, agreeable, and emotionally stable—likely because our uniform raising of the body differs from the prior cue's outward chest-forward projection. For $\mathscr{P}$, the extraversion direction matches [7] but the predicted link to neuroticism inverts, with high $\mathscr{P}$ read as emotionally stable. These divergences suggest that the geometric operationalization of a qualitative motion property shapes its perceived effects, with prior literature offering a useful starting point that empirical evaluation refines.

5. Limitations

Our user study was constrained by the number of base animation samples available. Future studies could expand the tested animation types and the compared variations to test the generalizability of the proposed approach. For example, the base animation could be compared separately to the low and high variants. We also did not test for different style descriptor combinations to keep the sample count manageable. However, the system can generate motion using a mixture of different style descriptors. In this case, the effects of each style descriptor are additive, as they use distinct losses. Future work could test the influence of different combinations where leveraging data-driven methods for human animation assessment can improve the quality of the evaluation [69].

Certain users may focus on specific body parts, and their answers could be affected by those parts. We deliberately used stick-figure rendering in the user study to isolate the perceptual contribution of motion itself, avoiding confounding factors such as character body type, clothing, and facial features that could otherwise influence personality and emotion judgments alongside the motion. Future work can test the animations using skinned meshes and use anatomically correct deformations to support realism [70]. The body pose can drive finger motion to achieve more versatile characters [71]. The current system focuses solely on altering the motion style, and since we do not use additional constraints for foot contacts or rotation limits, the resulting

animations may contain artifacts in these areas. Such constraints are orthogonal to motion style and can be composed with our pipeline. Future work can elaborate on introducing additional motion constraints to better control the output. For example, using contact loss could reduce foot sliding.

Our data-driven approach offers high-level parametric control via LMA-based descriptors but lacks the explicit, per-joint editability of procedural or analytical IK methods. For animators who require fine-grained pose adjustment, the generated motion can be combined with a downstream procedural editing stage.

The descriptor formulation and pipeline are dataset-agnostic: $\mathscr{V}$, $\mathscr{H}$, and $\mathscr{P}$ are computed directly from joint positions and velocities, and our system can, in principle, be retrained on any motion-capture dataset. The *learned mapping* between descriptor values and perceived personality and emotion, however, is shaped by the stylistic tendencies of the training data. Generalization to motion types under-represented in our datasets — particularly highly stylized cultural forms (e.g., specific dance traditions) or extreme dynamic motions such as athletics and combat — is therefore not guaranteed and would require either retraining on suitable data or further evaluation.

6. Conclusion

We propose a data-driven solution for IK-based human animation that uses LMA-inspired parameters as mediating features to control apparent personality and emotions. Our evaluation suggests the resulting system could introduce differences across all personality factors and the emotions of happiness, sadness, fear, and disgust. One major drawback of earlier procedural approaches was that handcrafted algorithms make assumptions about input motion, whereas our data-driven IK-based approach can utilize animations across different action categories, albeit with slightly different effectiveness for certain traits. Another advantage of an IK-based approach is that the system can operate without animation input; end effectors can be controlled in real time using positional tracking. Future work can expand our LMA-inspired parameters to introduce more refined control over the generated animations.

CRediT authorship contribution statement

Mehmet Akif Şahin: Data curation, Formal analysis, Investigation, Software, Validation, Writing – original draft, Writing – review & editing. **Sinan Sonlu:** Data curation, Formal analysis, Investigation, Software, Validation, Writing – original draft, Writing – review & editing. **Uğur Güdükbay:** Conceptualization, Formal analysis, Supervision, Validation, Writing – review & editing.

Declaration of Generative AI and AI-assisted technologies

During the preparation of this work, the authors used Grammarly and its Generative AI tool in order to improve the clarity and style of the writing. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Please add the following Acknowledgment text that includes funding information and thanks to the native speaker who proofread the manuscript.

Acknowledgments

This research is supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under Grant No. 122E123. We are grateful to Lori Russell Dağ for proofreading the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cag.2026.104643.

Data availability

Most of the datasets used in the experiments are from public resources. The source code is available in the GitHub repository https://github.com/Bilkent-ModVis/LMA-IK.

References

[1] Aristidou A, Lasenby J, Chrysanthou Y, Shamir A. Inverse kinematics techniques in computer graphics: A survey. Comput Graph Forum 2018;37(6):35–58.
[2] Torresani L, Hackney P, Bregler C. Learning motion style synthesis from perceptual observations. In: Proceedings of the 20th international conference on neural information processing systems. NIPS '06, Cambridge, MA, USA: MIT Press; 2006, p. 1393–400.
[3] Ramezanpanah Z, Mallem M, Davesne F. Human action recognition using Laban movement analysis and dynamic time warping. Procedia Comput Sci 2020;176:390–9.
[4] Wu C, Davaasuren D, Shafir T, Tsachor R, Wang JZ. Bodily expressed emotion understanding through integrating Laban movement analysis. Patterns 2023;4(10):100816, 14 pages.
[5] Bouchard D, Badler N. Semantic segmentation of motion capture using Laban movement analysis. In: Proceedings of the 7th international conference on intelligent virtual agents. IVA '07, Berlin, Heidelberg, Germany: Springer-Verlag; 2007, p. 37–44.
[6] Aristidou A, Stavrakis E, Charalambous P, Chrysanthou Y, Himona SL. Folk dance evaluation using Laban movement analysis. J Comput Cult Herit 2015;8(4). Article no. 20, 19 pages.
[7] Durupinar F, Kapadia M, Deutsch S, Neff M, Badler NI. PERFORM: Perceptual approach for adding OCEAN personality to human motion using Laban movement analysis. ACM Trans Graph 2016;36(1). Article no. 6, 16 pages.
[8] Erkoç Z, Demirci S, Sonlu S, Güdükbay U. Skeleton-based personality recognition using Laban movement analysis. In: Understanding social behavior in dyadic and small group interactions, proceedings of machine learning research. vol. 173, Birmingham, UK: PMLR; 2022, p. 74–87.
[9] Samadani A-A, Burton S, Gorbet R, Kulic D. Laban effort and shape analysis of affective hand and arm movements. In: Proceedings of the humaine association conference on affective computing and intelligent interaction. Piscataway, NJ, USA: IEEE; 2013, p. 343–8.
[10] Knight H, Simmons R. Expressive motion with x, y and theta: Laban effort features for mobile robots. In: Proceedings of the 23rd IEEE international symposium on robot and human interactive communication. ROMAN '14, Piscataway, NJ, USA: IEEE; 2014, p. 267–73.
[11] Broughton MC, Davidson JW. An expressive bodily movement repertoire for marimba performance, revealed through observers' Laban effort-shape analyses, and allied musical features: Two case studies. Front Psychol 2016;7. Article no. 1211, 20 pages.
[12] Kobayashi M, Liao C-C, Inoue K, Yojima S, Takahashi M. Motion capture dataset for practical use of AI-based motion editing and stylization. 2023, arXiv:2306.08861.
[13] Kim HJ, Neff M, Lee S-H. The perceptual consistency and association of the LMA effort elements. ACM Trans Appl Percept 2022;19(1). http://dx.doi.org/10.1145/3473041, Article no. 1 17 pages.
[14] Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. 2018, arXiv:1706.08500 URL https://arxiv.org/abs/1706.08500.
[15] Ponton JL, Yun H, Aristidou A, Andujar C, Pelechano N. SparsePoser: Real-time full-body motion reconstruction from sparse data. ACM Trans Graph 2023;43(1). Article no. 5, 14 pages.
[16] Sonlu S, Güdükbay U, Durupinar F. A conversational agent framework with multi-modal personality expression. ACM Trans Graph 2021;40(1). Article no. 7, 16 pages.
[17] Ong E-J, Hilton A. Learnt inverse kinematics for animation synthesis. Graph Model 2006;68(5–6):472–83.
[18] Chi D, Costa M, Zhao L, Badler N. The EMOTE model for effort and shape. In: Proceedings of the 27th annual conference on computer graphics and interactive techniques. SIGGRAPH '00, New York, NY, USA: ACM; 2000, p. 173–82.
[19] Lourens T, Van Berkel R, Barakova E. Communicating emotions and mental states to robots in a real time parallel framework using Laban movement analysis. Robot Auton Syst 2010;58(12):1256–65.
[20] Dewan S, Agarwal S, Singh N. Laban movement analysis to classify emotions from motion. In: Proceedings of the tenth international conference on machine vision. ICMV '17, 10696, Bellingham, WA, USA: SPIE; 2018, p. 717–24.

[21] Bernstein R, Shafir T, Tsachor R, Studd K, Schuster A. Laban movement analysis using Kinect. Int J Comput Inf Eng 2015;9(6):1567–71.

[22] Chen J-F, Lin W-C, Tsai K-H, Dai S-Y, et al. Analysis and evaluation of human movement based on Laban movement analysis. J Appl Sci Eng 2011;14(3):255–64.

[23] Foroud A, Pellis SM. The development of "roughness" in the play fighting of rats: A Laban movement analysis perspective. Dev Psychobiol: J Int Soc Dev Psychobiol 2003;42(1):35–43.

[24] Roudposhti KK, Santos L, Aliakbarpour H, Dias J. Parameterizing interpersonal behaviour with Laban movement analysis—A Bayesian approach. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition workshops. CVPRW '12, Piscataway, NJ, USA: IEEE; 2012, p. 7–13.

[25] Masuda M, Kato S. Motion rendering system for emotion expression of human form robots based on Laban movement analysis. In: Proceedings of the 19th international symposium in robot and human interactive communication. RO-mAN '10, Piscataway, NJ, USA: IEEE; 2010, p. 324–9.

[26] La Viola C, Fiorini L, Mancioppi G, Kim J, Cavallo F. Humans and robotic arm: Laban movement theory to create emotional connection. In: Proceedings of the 31st IEEE international conference on robot and human interactive communication. RO-mAN '22, Piscataway, NJ, USA: IEEE; 2022, p. 566–71.

[27] Koppensteiner M. Perceiving personality in simple motion cues. J Res Pers 2011;45(4):358–63.

[28] Lakatos G, Gácsi M, Konok V, Brúder I, Bereczky B, Korondi P, Miklósi Á. Emotion attribution to a non-humanoid robot in different social situations. PLOS One 2014;9(12):e114207, 32 pages.

[29] Bernardet U, Fdili Alaoui S, Studd K, Bradley K, Pasquier P, Schiphorst T. Assessing the reliability of the Laban movement analysis system. PLOS One 2019;14(6):e0218179, 6 pages.

[30] Perrinet J, Olivier A-H, Pettré J. Walk with me: Interactions in emotional walking situations, a pilot study. In: Proceedings of the ACM symposium on applied perception. 2013, p. 59–66.

[31] Assa J, Cohen-Or D, Yeh I-C, Lee T-Y. Motion overview of human actions. ACM Trans Graph 2008;27(5). Article no. 115, 11 pages.

[32] Neff M, Wang Y, Abbott R, Walker M. Evaluating the effect of gesture and language on personality perception in conversational agents. In: Allbeck J, Badler N, Bickmore T, Pelachaud C, Safonova A, editors. Intelligent virtual agents. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010, p. 222–35.

[33] Tolani D, Goswami A, Badler NI. Real-time inverse kinematics techniques for anthropomorphic limbs. Graph Model 2000;62(5):353–88. http://dx.doi.org/10.1006/gmod.2000.0528.

[34] Kallmann M. Analytical inverse kinematics with body posture control. Comput Animat Virtual Worlds 2008;19(2):79–91.

[35] Girard M, Maciejewski AA. Computational modeling for the computer animation of legged figures. ACM Comput Graph (Proc. SIGGRAPH '85) 1985;19(3):263–70. http://dx.doi.org/10.1145/325165.325244.

[36] Kang D, Park H, Kwon T. Efficient inverse-kinematics solver for precise pose reconstruction of skinned 3D models. Comput Graph 2024.

[37] Ho ESL, Shum HPH, Cheung Y-m, Yuen PC. Topology aware data-driven inverse kinematics. Comput Graph Forum 2013;32(7):61–70. http://dx.doi.org/10.1111/cgf.12212.

[38] Grochow K, Martin SL, Hertzmann A, Popović Z. Style-based inverse kinematics. ACM Trans Graph 2004;23(3):522–31. http://dx.doi.org/10.1145/1015706.1015755.

[39] Itani O. Deep-neural-network-for-solving-the-inverse-kinematics-problem. 2023, [Accessed 15 March 2026] https://github.com/OmarJItani/Deep-Neural-Network-for-Solving-the-Inverse-Kinematics-Problem.

[40] Vats U. Inverse Kinematics Using Neural Networks. 2021, [Accessed: 15 March 2026] https://github.com/Utkarsh-Vats-2000/Inverse-Kinematics-using-Neural-Networks.

[41] Starke S. DeepIK. 2017, [Accessed 15 March 2026] https://github.com/sebastianstarke/DeepIK.

[42] Agrawal D, Guay M, Buhmann J, Borer D, W. Sumner R. Pose and skeleton-aware neural IK for pose and motion editing. In: SIGGRAPH Asia 2023 conference papers. New York, NY, USA: ACM; 2023, Article no. 49, 10 pages.

[43] Neff M, Fiume E. Methods for exploring expressive stance. In: Proceedings of the 2004 ACM SIGGRAPH/eurographics symposium on computer animation. 2004, p. 49–58.

[44] Voss H, Kopp S. Real-time inverse kinematics for generating multi-constrained movements of virtual human characters. In: Proceedings of the 25th ACM international conference on intelligent virtual agents. IVA '25, New York, NY, USA: ACM; 2025, http://dx.doi.org/10.1145/3717511.3747066, Article no. 8, 10 pages.

[45] Aristidou A, Zeng Q, Stavrakis E, Yin K, Cohen-Or D, Chrysanthou Y, Chen B. Emotion control of unstructured dance movements. In: Proceedings of the ACM SIGGRAPH / eurographics symposium on computer animation. SCA '17, New York, NY, USA: Association for Computing Machinery; 2017, http://dx.doi.org/10.1145/3099564.3099566, Article no. 9, 10 pages.

[46] Jang D-K, Park S, Lee S-H. Motion puzzle: Arbitrary motion style transfer by body part. ACM Trans Graph 2022;41(3). http://dx.doi.org/10.1145/3516429, Article no. 33, 16 pages.

[47] Sumner RW, Zwicker M, Gotsman C, Popović J. Mesh-based inverse kinematics. ACM Trans Graph 2005;24(3):488–95. http://dx.doi.org/10.1145/1073204.1073218.

[48] Li J, Bian S, Xu C, Chen Z, Yang L, Lu C. HybrIK-X: Hybrid analytical-neural inverse kinematics for whole-body mesh recovery. IEEE Trans Pattern Anal Mach Intell 2025;47(4):2754–69. http://dx.doi.org/10.1109/TPAMI.2025.3528979.

[49] Jiang J, Streli P, Luo X, Gebhardt C, Holz C. MANIKIN: Biomechanically accurate neural inverse kinematics for human motion estimation. In: Leonardis A, Ricci E, Roth S, Russakovsky O, Sattler T, Varol G, editors. Computer vision – ECCV 2024. Cham: Springer Nature Switzerland; 2025, p. 128–46.

[50] Moreno-Villamarin DE, Hilsmann A, Eisert P. Multi-resolution generative modeling of human motion from limited data. In: Proceedings of 21st ACM SIGGRAPH conference on visual media production. CVMP '24, New York, NY, USA: Association for Computing Machinery; 2024, http://dx.doi.org/10.1145/3697294.3697309, Article no. 7, 10 pages.

[51] Ji X, Zhou H, Wang K, Wu W, Loy CC, Cao X, Xu F. Audio-driven emotional video portraits. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021, p. 14080–9.

[52] Liu J, Hui B, Li K, Liu Y, Lai Y-K, Zhang Y, Liu Y, Yang J. Geometry-guided dense perspective network for speech-driven facial animation. IEEE Trans Vis Comput Graphics 2021;28(12):4873–86.

[53] Feng B, Ao T, Liu Z, Ju W, Liu L, Zhang M. Robust Dancer: Long-term 3D dance synthesis using unpaired data. 2023, arXiv:2303.16856 [Cs.CV] Available at https://arxiv.org/abs/2303.16856, Online; [Accessed 15 March 2026].

[54] Qiu L, Yu C, Li Y, Wang Z, Huang H, Ma C, Zhang D, Wan P, Han X. ViMo: Generating motions from casual videos. 2024, arXiv:2408.06614 [Cs.CV] Available at https://arxiv.org/abs/2408.06614, [Accessed 15 March 2026].

[55] Won J, Gopinath D, Hodgins J. A scalable approach to control diverse behaviors for physically simulated characters. ACM Trans Graph 2020;39(4). Article no. 33, 12 pages.

[56] Zhong L, Xie Y, Jampani V, Sun D, Jiang H. SMooDi: Stylized motion diffusion model. In: Computer vision – ECCV 2024. Cham, Switzerland: Springer Nature; 2025, p. 405–21.

[57] Park S, Jang D-K, Lee S-H. Diverse motion stylization for multiple style domains via spatial-temporal graph-based generative model. Proc ACM Comput Graph Interact Tech 2021;4(3):1–17.

[58] Tang X, Wu L, Wang H, Wu Y, Hu B, Li S, Gong X, Liao Y, Kou Q, Jin X. Decoupling contact for fine-grained motion style transfer. In: Proceedings of SIGGRAPH Asia 2024 conference papers. SA '24, New York, NY, USA: Association for Computing Machinery; 2024, http://dx.doi.org/10.1145/3680528.3687609, Article no. 54, 11 pages.

[59] Lin I-C, Peng J-Y, Lin C-C, Tsai M-H. Adaptive motion data representation with repeated motion analysis. IEEE Trans Vis Comput Graphics 2010;17(4):527–38.

[60] Hou J, Chau L-P, Magnenat-Thalmann N, He Y. Human motion capture data tailored transform coding. IEEE Trans Vis Comput Graphics 2015;21(7):848–59.

[61] Zhou Y, Barnes C, Lu J, Yang J, Li H. On the continuity of rotation representations in neural networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. CVPR '19, 2019, p. 5738–46. http://dx.doi.org/10.1109/CVPR.2019.00589.

[62] Fu H, Li C, Liu X, Gao J, Celikyilmaz A, Carin L. Cyclical annealing schedule: A simple approach to mitigating KL vanishing. In: Burstein J, Doran C, Solorio T, editors. Proceedings of the conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers). Minneapolis, MN, USA: Association for Computational Linguistics; 2019, p. 240–50. http://dx.doi.org/10.18653/v1/N19-1021.

[63] Hempel T, Abdelrahman AA, Al-Hamadi A. 6D rotation representation for unconstrained head pose estimation. In: Proceedings of the IEEE international conference on image processing. ICIP '22, San Francisco, CA, USA: IEEE; 2022, p. 2496–500. http://dx.doi.org/10.1109/ICIP46576.2022.9897219.

[64] PyTorch. AdamW. 2019, [Accessed 15 March 2026] https://docs.pytorch.org/docs/stable/generated/torch.optim.AdamW.html.

[65] PyTorch. ReduceLROnPlateau - PyTorch 2.10 documentation. 2019, [Accessed 15 March 2026] https://pytorch.org/docs/stable/generated/torch.optim.lr_scheduler.ReduceLROnPlateau.html.

[66] Aberman K, Li P, Olga S-H, Lischinski D, Cohen-Or D, Chen B. Skeleton-aware networks for deep motion retargeting. ACM Trans Graph 2020;39(4). Article no. 62, 14 pages.

[67] Gosling SD, Rentfrow PJ, Swann Jr WB. A very brief measure of the Big-Five personality domains. J Res Pers 2003;37(6):504–28.

[68] Ekman P, Sorenson ER, Friesen WV. Pan-cultural elements in facial displays of emotion. Science 1969;164(3875):86–8.

[69] Rekik R, Wuhrer S, Hoyet L, Zibrek K, Olivier A-H. Quality assessment of 3D human animation: Subjective and objective evaluation. 2025, arXiv:2505.23301 [Cs.GR] Available at https://arxiv.org/abs/2505.23301, Online; [Accessed 15 March 2026].

[70] Nedel LP, Thalmann D. Modeling and deformation of the human body using an anatomically-based approach. In: Proceedings of the computer animation. CA '98', IEEE; 1998, p. 34–40.

[71] Jörg S, Hodgins J, Safonova A. Data-driven finger motion synthesis for gesturing characters. ACM Trans Graph 2012;31(6). Article no. 189, 7 pages.