ChoreoCraft: In-situ Crafting of Choreography in Virtual Reality through Creativity Support Tool

Hyunyoung Han*
hyhan@kaist.ac.kr
KAIST
Daejeon, Republic of Korea

Kyungeun Jung*
kyungeun.jung@kaist.ac.kr
KAIST
Daejeon, Republic of Korea

Sang Ho Yoon†
sangho@kaist.ac.kr
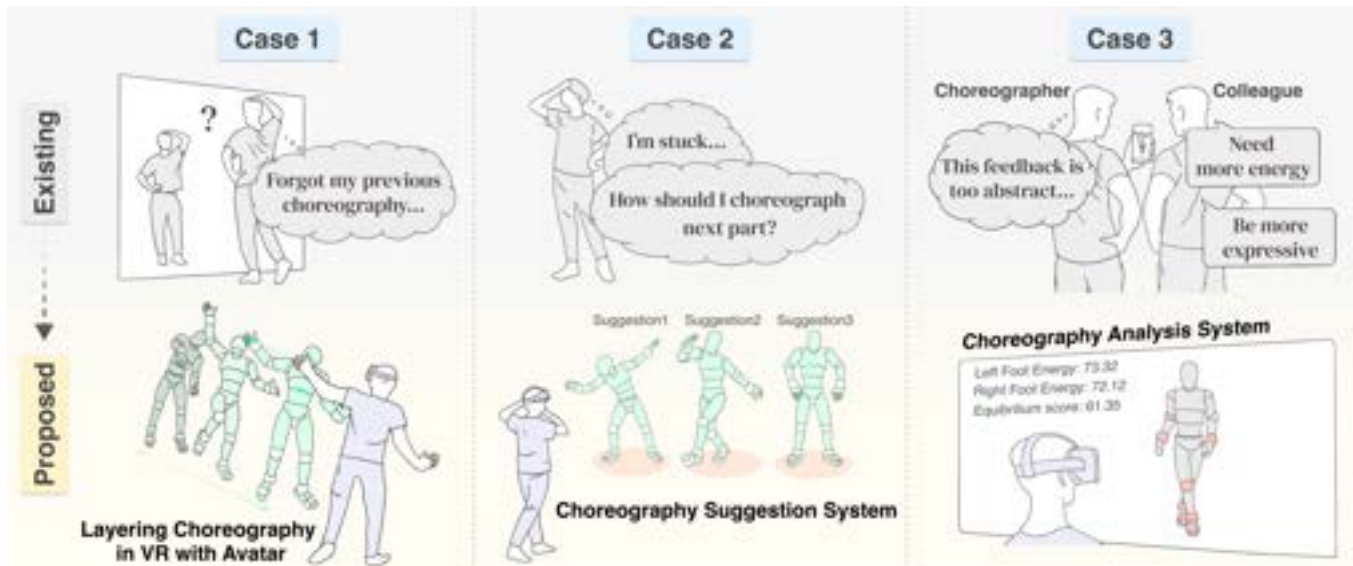KAIST
Daejeon, Republic of Korea

Figure 1: ChoreoCraft addresses three primary challenges in the traditional choreographic process. (Case 1) Choreographers often forget their compositions during the choreographic process. Our system establishes a VR choreography creation environment and provides a *snapshot* function to review and build upon created choreographies. (Case 2) Choreographers frequently encounter creative plateaus during the composition process. We propose a choreography suggestion system that suggests choreographies based on musical and motion similarity. (Case 3) Choreographers typically depend on abstract feedback on their creations, so we come up with a choreography analysis system to provide kinematic feature-based quantitative feedback.

ABSTRACT

Choreographers face increasing pressure to create content rapidly, driven by growing demand in social media, entertainment, and commercial sectors, often compromising creativity. This study introduces ChoreoCraft, a novel in-situ virtual reality (VR) choreographic system designed to enhance the creation process of choreography. Through contextual inquiries with professional choreographers, we identified key challenges such as memory dependency, creative plateaus, and abstract feedback to formulate design implications. Then, we propose a VR choreography creation system embedded with a context-aware choreography suggestion system and a choreography analysis system, all grounded in choreographers' creative processes and mental models. Our study results demonstrated that ChoreoCraft fosters creativity, reduces memory dependency, and improves efficiency in choreography creation. Participants reported high satisfaction with the system's ability to overcome creative plateaus and provide objective feedback. Our work advances creativity support tools by providing digital assistance in dance composition that values artistic autonomy while fostering innovation and efficiency.

CCS CONCEPTS

• **Human-centered computing → User interface toolkits**; **Virtual reality**; *User centered design*.

KEYWORDS

Creativity Support, Entertainment, Virtual/Augmented Reality

---

*Both authors contributed equally to this research.
†Corresponding author

1 INTRODUCTION

In recent years, the choreography industry has experienced growing demands for rapid content creation, particularly in the commercial sector [76] with a surge in dance challenge content on social media platforms [41] and rising demand for dance tutorials from dance learners [89]. These environments require professional choreographers to deliver original and creative movements in a shorter time period [82, 95]. This leads to limited exploration and refinement opportunities to craft choreography. Therefore, there is a need for interactive choreography tools to streamline choreography creation while maintaining artistic integrity. We propose a novel tool that enables choreography suggestion and analysis to enhance the existing choreographic creation process.

Several studies have enhanced the choreographic process with tools to facilitate ideation and prototyping [18, 46]. These studies focused on creative plateaus [46], which refer to the hardship of generating new movements or breaking free from repetitive patterns. These plateaus often come from mental fatigue to deal with complex spatial and temporal elements [17, 83]. Recent research proposed systems that created whole dance motion sequences based on audio input [78, 88]. In our work, we further enhance the choreographic process by introducing a motion suggestion system that takes care of both users' motion input and audio features. This allows choreographers to obtain more relative motion suggestions geared toward their artistic goals.

On the other hand, feedback is a critical element in refining and finalizing choreography. Although post-performance analysis or peer feedback has been commonly used, choreographers often rely on subjective intuition or abstract evaluations from colleagues [84]. This indicates there is a lack of systematic feedback for the current choreography creation process. To this end, previous works explored objective and data-driven approaches for analyzing dance motion through motion capture and computational techniques [26, 81]. Aligning with this trend, we employ a motion analysis tool that supports a systematic understanding of choreographers' movements which are used to provide objective feedback.

Among various interface types, the virtual reality (VR) interface has a high potential to provide seamless and uninterrupted workflow [27]. This leads users to easily carry out tasks in-situ with high task performance [7, 92]. Moreover, VR removes spatial and temporal limitations where choreographers do not need to work in a traditional studio to carry out the creative process. Thus, we implemented our tools in VR to better support the creative process. Moreover, VR allows users to view and refine movements from multiple perspectives in real-time to support spatial visualization features. These facilitate an intuitive understanding of their body balance and foster an immersive choreography creation process. This enables users to observe and refine their movements from multiple perspectives, breaking free from the spatial limitations of physical environments.

In this work, we propose ChoreoCraft, a VR choreography support tool that integrates avatar interactions, a choreography suggestion system, and a choreography analysis system. To the best of our knowledge, this is the first study to propose an interface that supports the choreographic creation process in-situ, without interrupting the creative flow, among the existing choreography support tools. Through an exploratory study, we found that choreographers gradually build their choreographies—a process we refer to as *layering*—and rely heavily on memory to retain their created choreographies. They often face creative plateaus and depend on abstract feedback. Our approach integrates VR technology into the choreographic workflow, providing an immersive environment for creative expression utilizing avatar interaction and a *snapshot* function to resolve memory dependency. We propose a choreography suggestion system, which reflects motion similarity and musical harmony to generate context-aware choreography motions. Within the system, we propose DanceDTW, a dynamic time warping-based algorithm specifically tailored to analyze motion similarity in the dance domain. DanceDTW ensures robustness against variations in dancers' physical characteristics and recording conditions. Furthermore, we introduce a choreography analysis system that quantitatively evaluates critical kinematic metrics, including motion stability, equability, and engagement.

Our key contributions are as follows:

- We developed a choreography crafting VR platform to resolve memory dependence and stimulate creativity for the choreographic creation process.
- We introduced a context-aware choreography suggestion system that incorporates multimodal inputs (music & prior motions) with DanceDTW to ensure smooth choreographic connections.
- We developed a kinematic-based analysis system to provide objective feedback for choreographic movements.
- We conducted user studies with expert choreographers to validate ChoreoCraft as a creative support tool.

2 RELATED WORK

2.1 Supporting the Choreographic Process

The choreography creation process is an intensively challenging and iterative one that requires creativity and a structured workflow. Felice et al. [17] conducted an ethnographic study to categorize this process into four stages: *Preparation*: This stage involves conceptualizing the choreography and using tools to prototype ideas such as sketching. *Studio*: Choreographers express their conceptual ideas through physical movement, utilizing various choreographic materials and engaging in interactions with fellow choreographers. *Performance*: This stage represents the live performance, where the prepared choreography is executed for an audience. *Reflection*: Choreographers identify areas for improvement based on observations from the *studio* or *performance* stages, leading to further development and refinement of the choreography.

Choreographers carry out *preparation*, *studio*, and *reflection* stages consecutively before the first *performance* (hereafter referred to as '*premiere*' according to [17]). In this study, we focused on enhancing these three consecutive stages. Previous research has explored

various methods to support these stages. For example, visualization techniques [28, 36, 52, 61] and interactive tools [8, 18] have been employed during the *preparation* stage to help choreographers ideate and prototype choreography. For the *studio* stage, interactive technologies have been utilized to assist in the physical expression of movement, with tools that facilitate collaboration with robots [31, 38], drones [24, 25], wall displays [77], and wearable suits [40, 43]. Additionally, there have been attempts to stimulate creativity through various visualization methods [25, 36, 61] and support the annotation process [12, 69]. Moreover, supportive systems for dance feedback have been developed to help choreographers analyze and refine their work for the *reflection* stage. Typically, this reflection is performed by observing self-recorded dance videos or discussing improvement directions with fellow choreographers. Previous research has primarily provided abstract feedback based on Laban Movement Analysis or complexity analysis [26, 57, 81]. Additionally, Singh et al. [69] and Carlson et al. [12] developed systems to support choreography creation by enabling dancers to self-reflect by viewing, evaluating, and improving their choreographic works within the system.

Our study builds on these efforts by addressing the need for a more integrated approach to supporting the *preparation*, *studio*, and *reflection* stages. We propose a system that enhances the choreographic process in VR environment, allowing choreographers to maintain a continuous and immersive creative flow. This system not only supports ideation and prototyping but also facilitates real-time feedback and iterative refinement, thereby improving the overall efficiency and creativity of the choreography creation process.

2.2 Dance Motion Creation and Suggestions for Choreography

The integration of algorithmic assistance in the choreographic process has been a topic of recent research, with a focus on generating dance motions through various input modalities such as music [2, 48, 78, 88, 90, 94], music & motion [30, 45, 71, 72, 74, 80], music & text [32], video [13], motion & text [46]. These algorithms can serve as catalysts for generating diverse motions during the ideation stage. Existing dance motion generation algorithms use evaluation metrics such as fidelity (comparison with ground truth, naturalness, physical plausibility), diversity (intra-motion, inter-motion), condition consistency (text-motion, audio-motion, scene-motion), and user study (subjective evaluation: preference, rating) [93]. Evaluations should be based on indicators from the decision-making process that dancers undergo when incorporating generated choreography into their actual creative process. This study proposes evaluation criteria that establish choreographers' mental models while creating new choreography and assist in the ideation process. Additionally, by conducting a study utilizing these criteria, we compared and evaluated the performance of several currently prominent algorithms. Liu & Sra [46] leveraged the dance motion generation algorithm and designed a system capable of dance motion generation and modification, as well as prototyping and documentation. However, it has limitations in synchronous usage with the choreography creation and supporting systems. Our system aims to build a choreography suggestion system that reflects

choreographers' mental models within the VR environment, enabling repetitive prototyping and 3D documentation. Moreover, by using this system during the choreography creation process, choreographers can receive "in-process" support.

2.3 VR-based Creativity Support Tool

VR has emerged as a promising medium for creativity support tools (CSTs) across various creative disciplines. These tools enhance user creativity and productivity by providing immersive, spatially unrestricted workspaces where users can interact with 3D objects and ideas in real-time [16, 44]. VR-based CSTs have been successfully applied in fields such as storytelling [86, 91], animation crafting [92], drawing [22, 75], fashion design [37], enabling users to explore new creative possibilities and refine their work with greater flexibility. In the domain of dance, VR-based tools have been primarily explored in educational and entertainment contexts. For example, VR has been used to teach dance by providing virtual instructors and detailed visualizations of dance moves, allowing students to practice in an immersive and risk-free environment [4, 42, 66]. Additionally, VR dance games [19, 33] have gained popularity for their ability to combine physical activity with interactive entertainment, offering users an engaging way to participate in dance routines [59, 62].

However, the potential of VR as a creativity support tool in the choreographic process still needs to be explored. While there have been some advancements in using VR for creating avatar animations and other creative tasks [92], the application of VR to support the entire choreographic process—encompassing ideation, prototyping, and refinement—has yet to be fully realized. The immersive nature of VR presents unique opportunities to enhance the choreographic process by enabling choreographers to visualize complex movements in real-time, experiment with new ideas in a 3D space, and receive immediate feedback on their creations. Our research seeks to explore and expand the potential of VR-based CST in the choreographic process.

3 EXPLORATORY STUDY: UNDERSTANDING CHALLENGES FOR CHOREOGRAPHERS' CREATIVE PROCESS

To clearly define the purpose and detailed design of a tool that assists in choreography creation, we conducted an exploratory study on the choreography creation process. We aimed to investigate choreographers' experiences including thoughts, actions, and emotions during the creation of solo dance pieces, and to identify the main challenges and requirements they face in the choreographic process. We conducted semi-structured interviews and contextual inquiries [6] to establish user-centered design implications for supporting the choreography creation in a VR environment.

3.1 Participants

We conducted the study with 8 choreographers (3 female) with varying levels of choreographic experience (2~14 years, $\mu = 6.25$, $\sigma = 4.71$). Specific demographic details are provided in the table 1. The choreographers participated in the study as consultants and were compensated with 50 USD. The study was approved under the IRB protocols.

Table 1: Demographic and Background Information of Participants in the Contextual Inquiry

Participants	Gender	Choreographic Experiences	Selected Song (Artists)	Length of Creation
P1	M	13 years	Poppin' (Chris Brown, 2005)	50 minutes (eight 8-counts)
P2	F	2 years	Say It. (Ebz the Artist, 2018)	50 minutes (eight 8-counts)
P3	M	3 years	Trip (Ella Mai, 2018)	30 minutes (eight 8-counts)
P4	F	2 years	SG (DJ Snake, 2022)	50 minutes (eight 8-counts)
P5	M	14 years	Savage Love (BTS, 2020)	23 minutes (eight 8-counts)
P6	F	5 years	MACARONI CHEESE (YOUNG POSSE, 2023)	33 minutes (eight 8-counts)
P7	M	5 years	Magnetic (R&B ver.) (Aaron Young, 2024)	63 minutes (four 8-counts)
P8	M	6 years	Come True (Summer Walker, 2019)	53 minutes (six 8-counts)

3.2 Procedure and Analysis

We conducted a 90-minute study in a dance studio. The study consisted of *Introduction*, *Pre-Interview*, *Contextual Inquiry (Observation & Interview)*, and *Post-Interview* phases. During the *Introduction*, we provided information about the purpose and procedure of the interview. The *Pre-Interview* aimed to explore the *Ideation* and *Prototyping* stages within the choreography creation process [17]. During the *Contextual Inquiry*, we observed the participants and asked questions while they created at least four 8-counts of the choreography. We refrained from giving additional instructions beyond the request to observe the participants' natural choreographic process. In the *Post-Interview*, we conducted a semi-structured interview regarding the queries raised during the *Contextual Inquiry* geared toward *preparation*, *studio*, and *reflection* stages. We conducted an analysis based on the interview transcriptions and recorded videos and performed affinity diagramming [34] to identify user needs for improvement in the choreography creation process.

3.3 Findings and Design Implication

3.3.1 Exploring Layering in Choreography and Addressing Memory Dependency. During the choreography creation process, we observed that all participants showed a pattern of gradually building a choreography sequence. Participants built choreography count by count to compose detailed sequences. Here, we define this process as *layering*. P4 and P8 noted that *layering* process allowed them to focus on the nuances of choreographic movements and continuity among them. On average, choreographers engaged in 12.53 (σ=7.33) repetitive *layering* to complete an 8-count sequence. Participants expressed frustration and anxiety when they forgot parts of their choreography (P3~P5), which necessitated starting from scratch. To avoid the risk of losing track of created choreography, all participants preferred to record their works frequently using smartphones. However, they found this process highly inconvenient and disruptive to their creative flow. While reviewing their recordings helped recall sequences, the frequent pauses to record and review were described as tedious and time-consuming.

To reduce the cognitive load for memorizing previously crafted sequences, we introduce a *snapshot* feature, virtual avatar-based record and replay feature for *layering*. This feature allows specified intervals of choreography to be recorded and supports replay for choreographers to seamlessly add newly composed sequences.

This approach mitigates the memorization issue and prevents losing any fleeting motions that occur spontaneously and are easily forgettable. We aim to support an 8-count-based recording system predicated on the scenario wherein choreographers create and organize choreography based on 8-count phrases.

3.3.2 Choreographers' Creative Plateaus and Designing Choreography Suggestion System. Choreographers invested significant time in the *ideation* phase. All participants expressed a desire to create new and original choreography but reported difficulties in the process. To inspire their creativity, participants engaged in improvisational movements (P2, P3, P5) or referred other choreographers' works on media platforms such as YouTube[1], Instagram[2], and Tik-Tok[3] (P1~P3, P5~P8). These references helped them understand how other choreographers think, feel, and express the given song. Participants referred to videos of other choreographies for the same song they were working on (P2, P3, P5~P8), different songs from similar genres (P1, P3, P6-P8), and even different genres (P1, P5, P6). They expressed concerns about directly copying the referenced choreography (P2, P3, P6~P8). As a result, participants focused more on emotional and visual impressions of movements or interaction with the music when reinterpreting the referenced choreographies into their own (P1, P2, P6, P8).

To help choreographers overcome these creative plateaus, previous dance motion generation algorithms [45, 71, 78] and system [46] have made valuable contributions. There still remains room for improvement in aligning generated movements with choreographers' creative processes and intentions. In this work, we propose a choreography suggestion system that takes into account the choreographers' scenario context. We aim to meet two key criteria for high-quality choreography: harmony with the current music and continuity between existing and generated motions. Inspired by choreographers' tendency to refer to dance videos on platforms like YouTube when facing creative plateaus, our system suggests a set of dance motions exerted from online dance videos in two steps. First, we identify the audio source that shows high similarity with the music the choreographers are working with and use it to extract potential dance motions. From these motions, we then evaluate their similarity to the choreographer's existing sequence to suggest the most compatible motions to support a smooth creation

---

[1]https://www.youtube.com/
[2]https://www.instagram.com/
[3]https://www.tiktok.com/

flow. In this way, we ensure both harmony with music and motion continuity while resolving creative plateaus in dance composition.

3.3.3 Towards a Systematic Choreography Analysis with Quantified Feedback. Participants mentioned that they carry out peer or self-evaluations on their choreography to enhance the completeness (P1, P3, P6, P8) which ensures the seamless alignment with the flow of the music and the previously crafted motions (P1, P2, P7). However, the current way of getting feedback from peers is often hard to interpret since the comments often include an abstract and subjective context. Moreover, the choreographers expressed a strong desire to receive quantitative feedback. To support these needs, we introduce a choreography analysis system based on kinematic features.

Participants wished to receive objective data to better understand body balance and correct asymmetries (P7, P8) for improving body control. To this end, we came up with *Motion Equability* metric that computes the distance between the center of gravity and foot placement. Participants also emphasized the importance of quantitative feedback on movement counts within the same beat. They valued data on which joints were active or inactive, enabling them to explore and utilize underused joints (P5~P8). They expressed a desire to visualize the activation levels within sequential dance movements to ensure harmony and coherence between the movements (P1, P2, P5). Here, we implemented *Motion Engagement* which highlights both salient and non-salient joints for visual feedback. Furthermore, participants wanted feedback on how they used spatial elements, particularly asymmetrical spaces, to improve spatial awareness (P3, P8). To support this, we propose *Motion Stability* metrics where one evaluates foot stability by tracking kinetic energy to detect unnecessary movements while another one tracks joint movement in spatial quadrants to assess space utilization.

4 CHOREOCRAFT

We propose ChoreoCraft (Figure 2) that enables an in-situ choreography craft in VR. Users can craft choreography with interactive virtual avatars along with creative support tools including suggestion and analysis interfaces. For our in-situ choreography craft interface, we apply *layering* concept as snapshot function. Overall ChoreoCraft consists of (1) in-situ VR record-and-replay, (2) choreography suggestion, and (3) choreography analysis. We built an entire system in the Unity[4] game engine.

4.1 Choreography Supportive System in VR

We aim to support choreographers in crafting choreography within a VR environment. VR was chosen as the interface for its ability to offer an immersive and focused environment, minimizing distractions and enhancing creative engagement. VR lets choreographers seamlessly transit between ideation and execution in-situ to form an uninterrupted workflow [27]. By removing the constraints of physical space and time, VR also enables the creative process to occur beyond the traditional studio setting, offering flexibility and accessibility. To achieve this, we implemented a robust motion capture configuration that allows dancers to record their movements and replay motion clips. Our system utilizes Azure Kinect [51],

which supports real-time human motion tracking of 32 joints' positional and rotational data. For the *snapshot* function, we capture the dancer's series of motions at each segmented frame interval (60 [sec]/BPM × 8) and generate a corresponding avatar in VR, which replays the recorded movements upon clicking. These features were designed to harness VR's potential to better support choreographers by creating an immersive, flexible, and distraction-free creative environment.

To compare input motion sequences from Azure Kinect with the SMPL-represented motion database, we pre-process motion data from different representations using the Iterative Closest Point algorithm [87]. The process involved computing optimal rotation matrices and translation vectors to align with two datasets, focusing on a common joint representation between the two systems. More details can be found in Appendix B.

4.2 Choreography Suggestion System

We propose a choreography suggestion system to alleviate choreographers' creative plateaus and assist in the creation of high-quality choreography (Figure 3). To consider the harmony between music and choreography, our system utilized the Spotify recommender API[5] that offers high-performance cognitive evaluations [5]. This API extracts lists of similar music tracks based on the input music and then collects choreography videos corresponding to identified music tracks (solo and group performances) from social media platforms. We converted these videos into 3D poses using a state-of-the-art (SOTA) 3D human pose estimation model called WHAM model [67]. When users encounter creative plateaus during choreography creation, they input ongoing motion sequences to receive suggestions for following choreography. The system employs dynamic time warping to measure the motion similarity between the user's input and database sequences. Then, the system returns the subsequent 8-count choreography sequence from the most similar motion. Overall, our system suggests choreography considering both the musical context and preceding movements.

4.2.1 Dynamic Time Warping for Dance Motion Similarity. Dynamic time warping (DTW) [53] is a widely adopted algorithm for analyzing time-series data across various fields, including kinesiology [64, 73] and musicology [11, 60]. DTW showed robust performance when comparing data with spatiotemporal distortions. To compare dance motion similarity, we propose **DanceDTW** (Figure 4), an extension of the multi-dimensional similarity comparison algorithm by Shokoohi-Yekta et al. [68] and incorporating the rotation information handling technique from [65]. We mainly used positional and rotational information of joints from two dance motions to assess the similarity. The lower final cost (C_{total}) indicates a higher degree of similarity among compared motions. Our approach mitigates the challenges associated with similarity comparison in sequential motion data, addressing issues arising from anthropometric variations and data capture differences across subjects, leading to enhanced performance. Please refer to Appendix A for a detailed DanceDTW algorithm.
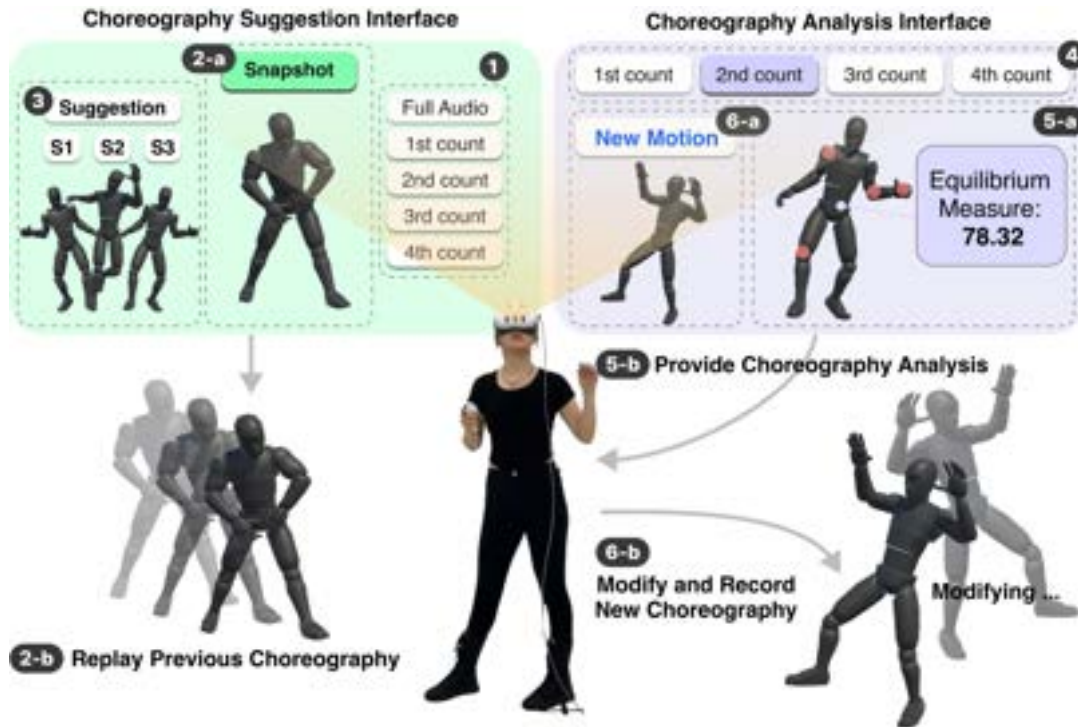

Figure 2: ChoreoCraft System Overview. ① Upon entering the VR interface, choreographers listen to the full audio and hover around segmented counts to start listening and recording the motion. ②-a The system displays the identical avatar of the choreographer. ②-b *Snapshot* function to instantly playback what has been recorded. ③ Upon activating the suggestion function, three different suggestion motions visualized with an avatar are shown. ④ Choreographers select an 8-count chunk recorded motion for review. ⑤-a The system displays a designated motion analysis value with corresponding visualization on an avatar. ⑤-b Then, choreographers interpret the analysis. ⑥-a After reviewing and making modifications, the choreographer records a new motion.⑥-b This example showcases where the motion analysis showed high motion dependency on the left wrist alerting the choreographer to focus on the right side.
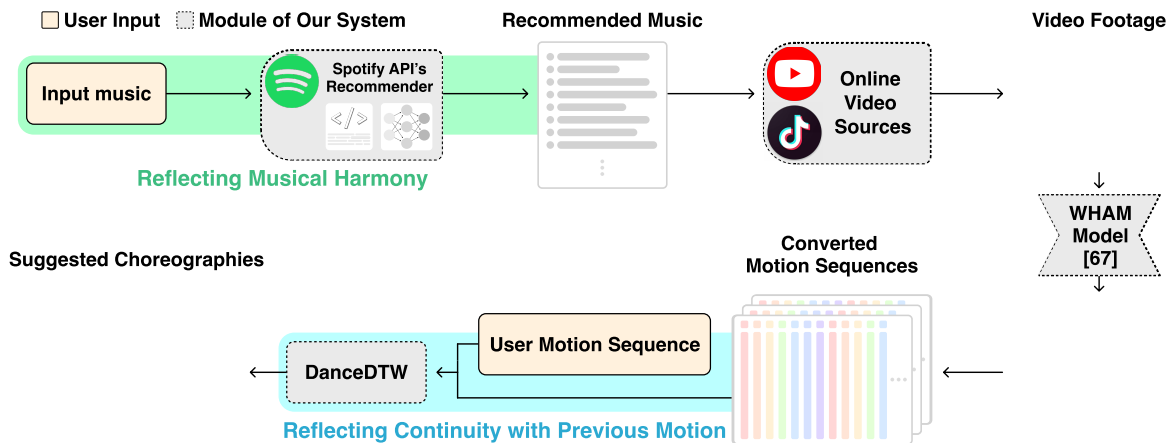
Figure 3: We suggest choreographies based on the user's input music and motion sequence. We extract similar songs and then extract choreography videos related to these songs from online sources. Subsequently, we undergo a 3D pose estimation to convert the videos into 3D motion sequences and employ DanceDTW to compare the similarity between these sequences and the user's input motion sequence. Going through these processes, we ultimately suggest choreographies that reflect both musical harmony and continuity with previous motions.

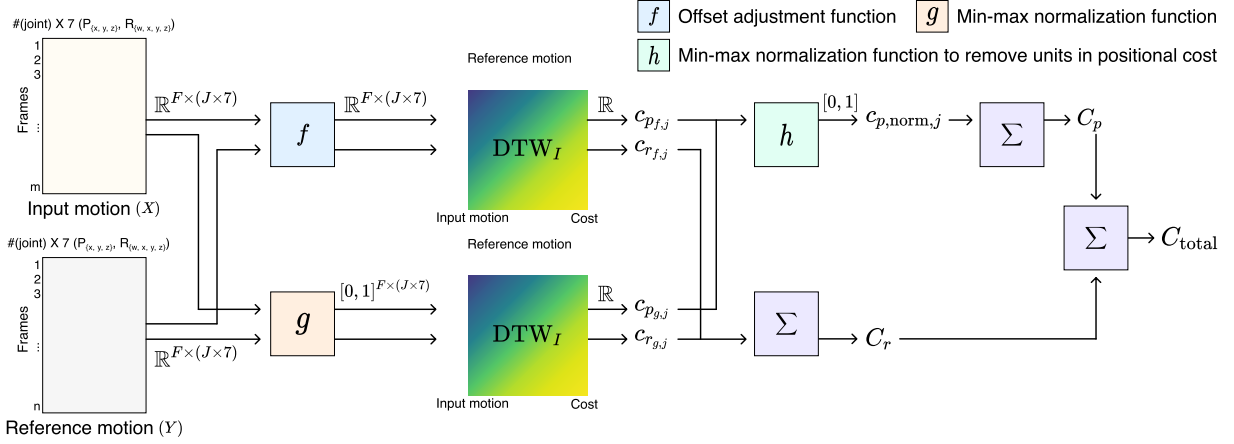**Figure 4: The structure of DanceDTW. Input motion (X) and reference motion (Y) undergo offset adjustment (f) and min-max normalization (g), respectively. Positional costs are normalized using function h and summed to obtain the total positional cost (C_p). Rotational costs are directly summed to obtain the total rotational cost (C_r). The final cost (C_{total}) represents the overall similarity between the motions.**

Let (X, Y) be a pair of input and reference motion data sequences, where:

$$X, Y \in \mathbb{R}^{F \times (J \times 7)}. \tag{1}$$

Here, F is the number of frames, J is the number of joints (with $J = 24$), and 7 represents positional (in mm) and rotational (quaternion) information: $(P_x, P_y, P_z, R_w, R_x, R_y, R_z)$ The articular structure and representation in this instance adhere to the methodology employed in the SMPL model [47].

We adjusted the offset between input datasets and applied min-max normalization as follows:

$$\overline{Z} = \frac{1}{2}(\overline{X} + \overline{Y}), \quad \overline{Z} \in \mathbb{R}^{(J \times 7)}, \tag{2}$$

$$f : \mathbb{R}^{F \times (J \times 7)} \rightarrow \mathbb{R}^{F \times (J \times 7)}, \quad f(x) = x - \overline{Z}, \tag{3}$$

$$g : \mathbb{R}^{F \times (J \times 7)} \rightarrow [0, 1]^{F \times (J \times 7)}, \quad g(x) = \frac{x - \min(x)}{\max(x) - \min(x)}. \tag{4}$$

The preprocessing methodology adjusts data ranges to compensate for variations stemming from anthropometric differences and disparate recording conditions. Using offset adjustment and min-max normalization (Equations 3 and 4), we compensate for physical differences, variations in a range of motion, and disparities in data sensing methodology. This compensation enables the robust recognition of similarities between two target motions performing the same action regardless of recording conditions. After adjusting the position and rotation data, we applied multi-dimensional DTW and multiplied each resulting cost by a coefficient to modulate each processing function.

We define the multi-dimensional DTW operation by employing the Independent DTW approach from [68], which allows us to handle the translational and rotational dimensions of each joint independently:

$$\text{DTW}_I(X, Y) : \mathbb{R}^{F \times (\text{dim.})} \times \mathbb{R}^{F \times (\text{dim.})} \rightarrow \mathbb{R}. \tag{5}$$

For each joint j, we calculate the costs for position and rotation as follows:

$$\begin{aligned}
c_{p_{f,j}} &= \lambda_f \cdot \text{DTW}_I(f(X_{p,j}), f(Y_{p,j})), \\
c_{p_{g,j}} &= \lambda_g \cdot \text{DTW}_I(g(X_{p,j}), g(Y_{p,j})), \\
c_{r_{f,j}} &= \lambda_f \cdot \text{DTW}_I(f(X_{r,j}), f(Y_{r,j})), \\
c_{r_{g,j}} &= \lambda_g \cdot \text{DTW}_I(g(X_{r,j}), g(Y_{r,j})).
\end{aligned} \tag{6}$$

Here, λ_f and λ_g adjust the range of the results from Equations 3 and 4, respectively, with the constraint that $\lambda_f + \lambda_g = 1$. This adjustment process compensates for the results obtained between offset adjustment and min-max normalization.

Next, we gather the position-related costs for all joints:

$$c_{p,\text{all}} = \{c_{p_{f,j}}, c_{p_{g,j}} \mid j = 1, \ldots, J, \ldots, J \times 3\}. \tag{7}$$

To incorporate both position and rotation costs into the total cost, we normalize each position-related cost to remove units. Through this process, the value of $c_{p,\text{norm},j}$ is transformed into an actual number within the range of $[0, 1]$.

$$c_{p,\text{norm},j} = \frac{c_{p,j} - \min(c_{p,\text{all}})}{\max(c_{p,\text{all}}) - \min(c_{p,\text{all}})} \quad \text{for each } c_{p,j} \in c_{p,\text{all}}. \tag{8}$$

Finally, we calculate the total position and rotation costs, C_p and C_r, as follows:

$$C_p = \sum_{j=1}^{2 \times J \times 3} (c_{p,\text{norm},j}) \quad \text{(unitless)}, \tag{9}$$

$$C_r = \sum_{j=1}^{J \times 4} (c_{r_{f,j}} + c_{r_{g,j}}) \quad \text{(as quaternions are unitless)}. \tag{10}$$

The final total cost is given by:

$$C_{\text{total}} = \mu_p C_p + \mu_r C_r, \quad \text{where } \mu_p + \mu_r = 1 \tag{11}$$

Here, μ_p and μ_r adjust for the range and scale differences between C_p and C_r, respectively. These parameters were fitted using Bayesian optimization, as discussed in the next section.

4.2.2 Evaluation of DanceDTW. In the context of similarity comparisons within the dance domain, we conducted fitting procedures for the modulating features λ and μ in the aforementioned Equations 6 and 11 based on actual choreography similarity comparison datasets. These fitting procedures were essential to optimize DanceDTW's performance and ensure its applicability to actual choreographic data. The dataset is primarily comprised of dance learning class videos featuring the instructor's reference choreography and students' performances. These {reference, student} video pairs capture the same choreography performed by different individuals or recorded from slightly different angles and timings. We evaluate the algorithm's efficacy to compensate for physical disparities between subjects and range of motion variations. Consequently, the dataset spans 9 genres and 11 dance videos, averaging 7.15 8-count chunks (σ=2.38) per video. Each dataset is structured as an {ref., stu.} pair, containing 24 joints with both position and rotation information.

We conducted coefficient fitting using 90 pairs of video data (1,300 video chunks) and evaluated 9 pairs (116 video chunks). This process involved fitting coefficients to classify identical 8-count chunks within the same genre and dance for paired dance motion data. Bayesian optimization [29] was employed for coefficient fitting with the objective function set to maximize the number of correctly classified pairs. We set the number of evaluations of the objective function as 50. The derived coefficients through data fitting are $\lambda_f =$ 0.9386, $\lambda_g = 0.0614$, $\mu_p = 0.0008$, $\mu_r = 0.9992$, yielding an accuracy of 86.31% on the training set and 89.66% on the evaluation set. Details about Bayesian optimization can be found in the appendix A.

Subsequently, we validated the classification performance using the UCR Time Series Classification Archive [14] and the choreographic similarity comparison dataset. In the comparison with standard DTW on the UCR dataset, our method demonstrated superior or equivalent performance in 50.6% (42/83) of cases, with 48.2% (40/83) showing superior performance and 2.4% (2/83) exhibiting equivalent results. When considering a 10% error tolerance, the performance improved to 91.57% (76/83), and with a 20% error tolerance, it further increased to 96.38% (80/83). When evaluated against a choreographic similarity comparison dataset, the standard DTW algorithm achieved an accuracy of 15.96%, while our proposed method attained an accuracy of 89.66%. This substantial improvement suggests that our approach not only performs the function of DTW but also demonstrates high applicability to actual choreographic similarity comparison tasks.

4.3 Choreography Analysis System

Drawing from the design implications, we develop a choreography analysis system, which provides analytical metrics for a sequence of crafted choreographies (Figure 5). For all the Equations, i indicates the number of frames, j indicates joints and t implies the time unit (1[sec]) used in Unity.

F_{eq} : **Measure of Equilibrium.** Any imbalance is considered a significant error that affects the assessment of dance choreography [9, 55]. To identify loss of balance, we track the distance from the body's center of mass to the rectangle Z formed by the performer's feet [55]. Rectangle Z is defined with values four vertical

components from feet ankles (Left Ankle: (x_l, y_l, z_l), Right Ankle: (x_r, y_r, z_r)), rectangle Z: (x_l, z_l), (x_l, z_r), (x_r, z_r), and (x_r, z_l)). Then measures the distance between the barycenter of the upper body $B = (x_b, y_b, z_b)$ [55]. To provide the objective score based on the distance, we tracked the maximum distance ($\max(F_{\text{eq}})$) peaked throughout their crafted choreography and calculated the overall score (Equation 13).

$$F_{\text{eq}} = \sqrt{\left(\frac{x_r + x_l}{2} - x_b\right)^2 + y_b^2 + \left(\frac{z_r + z_l}{2} - z_b\right)^2}, \quad (12)$$

$$\text{Overall Score for } F_{\text{eq}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{F_{\text{eq},i}}{\max(F_{\text{eq}})} \times 100\right). \quad (13)$$

F_{sb} : **Feet Position Stability.** Here, we measure the stability of the feet' positions during the performance. Feet need to remain steady and calm in between consecutive frames, and better choreography avoids any unnecessary support from movements, slipping, or leg vibrations [3, 10]. For each frame i, two values are calculated: (1) the kinetic energy of the feet, E_i^f, and (2) the filtered kinetic energy, Ef_i^f, using a low-pass filter (settled cutoff frequency: 10hz). $|E - Ef|$ represents the absolute difference between the energy and the filtered energy, while $\max(E, Ef)$ normalizes the stability measure [9, 55]. The score indicates that a higher overall score for F_{sb} corresponds to greater stability for both feet (Equation 15). Here, L refers to left, and R refers to right.

$$F_{\text{sb},l} = 1 - \frac{|E_l - Ef_l|}{\max(E_l, Ef_l)}, \ F_{\text{sb},r} = 1 - \frac{|E_r - Ef_r|}{\max(E_r, Ef_r)}, \quad (14)$$

$$\text{Overall Score for } F_{\text{sb}} = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{|E_{i,l} - Ef_{i,l}|}{\max(E_{i,l}, Ef_{i,l})}\right) \times 100. \quad (15)$$

F_{sj} and F_{nsj} : **Visualization of Salient/Non-Salient Joints.** For F_{sj} and F_{nsj}, we aim to identify the salient joints by utilizing both linear and angular acceleration for each input joint. Previous approach [39] focused solely on angular acceleration to decide salient joints. We further consider both linear (L_i) and angular acceleration (A_i) for determining salient joints since the choreography involves sophisticated motions [10, 56]. We compute the average of linear and angular velocities for all joints to compare each joint's velocity to the overall mean (Equation 16).

$$A_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\alpha_{ij}, \quad L_i = \frac{1}{n_i}\sum_{j=1}^{n_i}a_{ij}, \quad F_{\text{sj},i} = c_1 \cdot L_i + c_2 \cdot A_i. \quad (16)$$

We establish real-time adaptive coefficients (c_1,c_2) to interpret acceleration data and apply it to the complexity of salient joints. A pair of coefficient was designed for both linear and angular acceleration ensuring the sum of each pair equals 1. This pair remain inverted when comparing the calculated acceleration for each frame, suggesting that the coefficient pairs are assigned depending on the dominant form of acceleration [20, 63]. Since linear and angular accelerations have different units, we calculated their proportion across all joints. Among various coefficient combinations, we adopted $(0.9, 0.1)$ as our coefficients, since this configuration enables highly notable reflections in the calculated joint saliency values. This configuration enhances visualization by effectively
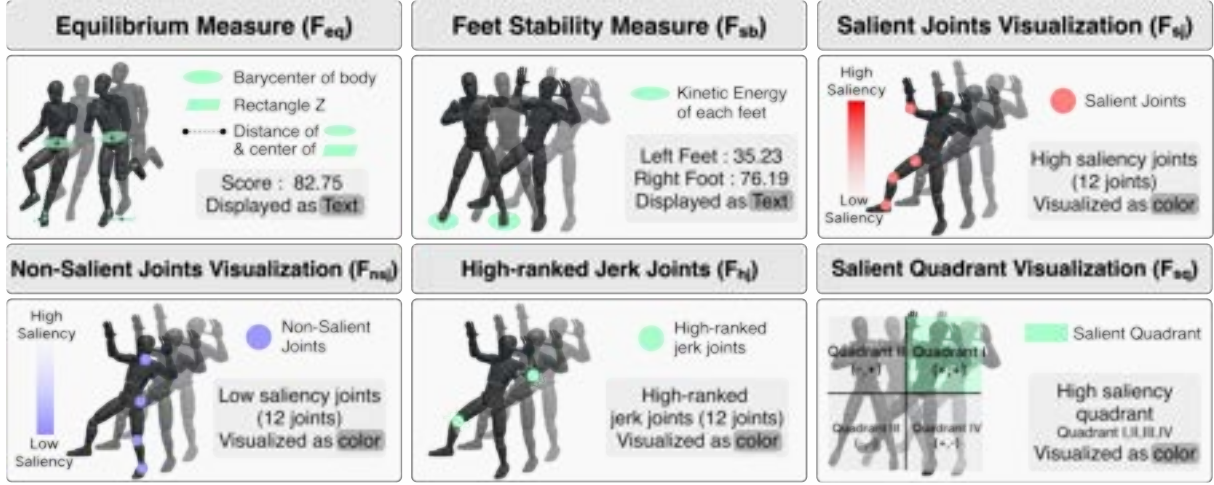
Figure 5: Kinematic factors used for choreography analysis derived from our exploratory study. We considered these factors to cover *Motion Equability*, *Motion Stability*, and *Motion Engagement* for our analysis.

highlighting differences in joint activation and reflecting the complexity of the choreographers' motion patterns, making the salient joints more visually distinguishable. We internally determined the number of joints to activate, if more than 50% of the input joints (32 joints) are visualized, it creates visual clutter, diminishing the clarity and effectiveness of information delivery. We therefore adopted 35% (12 joints) as the number of activated joints. As these values accumulate across all motion frames, the joints are ranked in descending order, and for F_{sj}, numbers of high-ranked joints are highlighted in Red (RGBA(255,0,0,0)) [1]. In contrast, we visualize F_{nsj} with numbers of low-ranked joints to show deactivated joints.

F_{hj} : **Visualization of high-ranked jerk joints.** For F_{hj}, we extend our analysis to consider both linear and angular jerks showing a dynamic view of how joint movements accelerate or decelerate over time. The linear jerk refers to changes in positional acceleration, whereas the angular jerk captures shifts in rotational acceleration (Equation 17). We average these across all frames and combine them using coefficients ($c = 0.5$).

$$J_{A_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{d\alpha_{ij}}{di}, J_{L_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{da_{ij}}{di}, \; F_{hj,i} = c \cdot J_{L_i} + (1-c) \cdot J_{A_i}.$$
$$(17)$$

F_{sq} : **Visualization of salient quadrant** Here, F_6 tracks the time spent in each quadrant to show the spatial dynamics of joint movement. By accumulating the F_6 indicator in every frame, we provide a clearer understanding of the spatial distribution of joint movements and highlight any asymmetries or biases in the choreography. We express this in a general Equation that accumulates the time a joint spends in each of these four quadrants. By denoting the position of the joint as $\vec{p}_{joint} = (x_{joint}, y_{joint})$, the barycenter position as $\vec{p}_{bary} = (x_{bary}, y_{bary})$, and the accumulated time in each quadrant as $T_{quadrant}$. The function $\delta_{quadrant}(\vec{p}_{joint}, \vec{p}_{bary})$ returns 1 if the joint is in the quadrant based on x and y comparisons, and 0 otherwise (Equation 18). As shown in the Figure 5, four quadrants ($x_{joint}, y_{joint}, x_{bary}, y_{bary}$)'s origin (0,0) is rooted in bary-center

of body, which simultaneously rotates and moves.

$$T_{quadrant} = \int_0^t \delta_{quadrant}(x_{joint}, y_{joint}, x_{bary}, y_{bary}) \, dt. \qquad (18)$$

5 USER EVALUATION

To validate the proposed system, we conducted an IRB-approved user evaluation that consisted of two sessions. First, we assessed how well our system suggests motions compared to recent dance motion generation algorithms. Then, we carried out another evaluation to figure out the most preferred choreography analysis factors and determine design parameters for visualizing salient joints. We recruited 18 choreographers (12 female) with diverse experience levels (2~13 years, $\mu = 5.75$, $\sigma = 3.91$) (See Table 2). Of the 18 participants in the user evaluation, six had previously participated in the exploratory study. This evaluation involved a total of four surveys (two surveys per session), each lasting for 30~60 minutes, and we compensated 75 USD.

5.1 User Evaluation 1: Validation of Choreography Suggestion System

Prior research [46] demonstrated the use of generative-AI (GenAI) to overcome creative plateaus in choreographic processes. Building on this, we evaluated how effectively GenAI met the criteria from our exploratory study. Specifically, we focused on its ability to support "high-quality choreography" and reduce creative plateaus. We conducted a user evaluation to compare the proposed choreography suggestion system against two baseline GenAI-based motion generation models. This evaluation aims to assess each model's ability to reflect the contextual nuances of choreographic creation. The baseline models include 1) a GPT-based motion generation from audio and starting pose inputs (Bailando [71]) and 2) A transformer-based diffusion model that generates dance motions in response to audio input (EDGE [78]). To ensure a fair comparison, we chose these as baseline models to represent SOTA performance

Table 2: Demographic and Background Information of Participants for User Evaluation

ID	Gender	Dance Experiences	Choreographic Experiences	Main Genre
P1	F	8 years	4 years	B-Boying, Breaking
P2	F	6 years	3 years	Choreography
P3	M	3 years	3 years	Choreography, Hip-hop
P4	F	3.6 years	3 years	Choreography
P5	F	12 years	10 years	Choreography
P6	F	14 years	10 years	Choreography, Hip-hop
P7	F	6 years	5 years	Choreography, Hip-hop
P8	F	15 years	13 years	Choreography, Hip-hop
P9	F	8 years	8 years	Choreography
P10	F	11 years	3 years	Choreography
P11	F	16 years	2 years	Choreography
P12	M	5 years	5 years	Choreography, Hip-hop
P13	F	2.5 years	2.5 years	Heel Choreography
P14	F	8 years	6 years	Modern Dance
P15	F	6 years	6 years	Heel Choreography
P16	M	2 years	2 years	Choreography, Hip-hop
P17	M	5 years	3 years	Choreography, Hip-hop
P18	M	16 years	15 years	Choreography, Hip-hop

in motion generation [35, 48, 70, 72] to minimize artifacts generated by GenAI. Additionally, extremely unnatural motions such as static poses or joint hierarchy distortions were excluded during the survey preparation process to avoid biases.

Our study compared a total of 45 dances (9 genres × 5 dances). The dance genres include breaking, hip-hop, house, jazz, k-pop, krump, locking, popping, and waacking. We designated one 8-count sequence for each method to carry out motion suggestion/generation and either use the music and motion data (our method & Bailando) or only music data (EDGE) as input. For music data, we provide the segment of the part that needs to be suggested/generated, whereas we feed the preceding 8-count motion data, which is required for our method and Bailando. We used the Bailando and EDGE models pre-trained on the AIST++ dataset [79], encompassing all 9 genres above, with the default model configuration.

We conducted the evaluation through an online survey via Tally Forms[6]. Participants first watched a reference choreography, and they subsequently viewed three dance clips generated by different methods for the reference choreography. Then, the participants responded to six questions regarding these observations. We constructed the question shown in Table 3 with criteria related to high-quality choreography (Q1~2) and utility (Q3~6), which we identified from the exploratory study. Participants reported priority ranking (from 1st to 3rd) for given methods for Rank questions and a 7-point Likert scale for Preference questions. We randomized the three generated motions across questions using a Latin square design. We divided the survey into two parts to avoid the fatigue effects. Each participant evaluated a total of 135 motions (45 reference motions × 3 methods).

Table 3: Types and Content of Questions Employed in User Evaluation 1

Types	Questions
Rank	Q1. Rank the continuity between the preceding motion and each generated motion.
Rank	Q2. Rank the congruence between the background music and derived dance motions.
Preference	Q3. Evaluate the continuity between the preceding and the three dance motions.
Preference	Q4. Assess the level of harmony between the music and dance motions.
Preference	Q5. Evaluate the extent to which each dance motion stimulates choreographic creativity.
Preference	Q6. Assess how the presented choreography could be incorporated into actual dance creation.

This study collected 4,860 responses (45 items × 6 questions × 18 participants). We used Friedman test [96] because each participant independently performed and evaluated multiple conditions (non-parametric method). We statistically validated the differences in evaluations across the conditions to ensure the reliability of our findings. Regarding the Rank evaluation, our method achieved a winning rate of 93.70% and 92.72% for Q1 and Q2 accordingly, securing the first position (Figure 6).

In Preference results, we observed significant differences among the three approaches in Q3 as indicated by the Friedman test statistics ($\chi^2(2) = 1197.5$, $p < 0.001$). The post-hoc Dunn's test further confirmed the superiority of our method with significant results against EDGE ($p < 0.001$) and Bailando ($p < 0.001$). Similarly, the proposed system demonstrated significant superiority over the baselines in the other three criteria (Q4~6) as well. In Q4, the test results
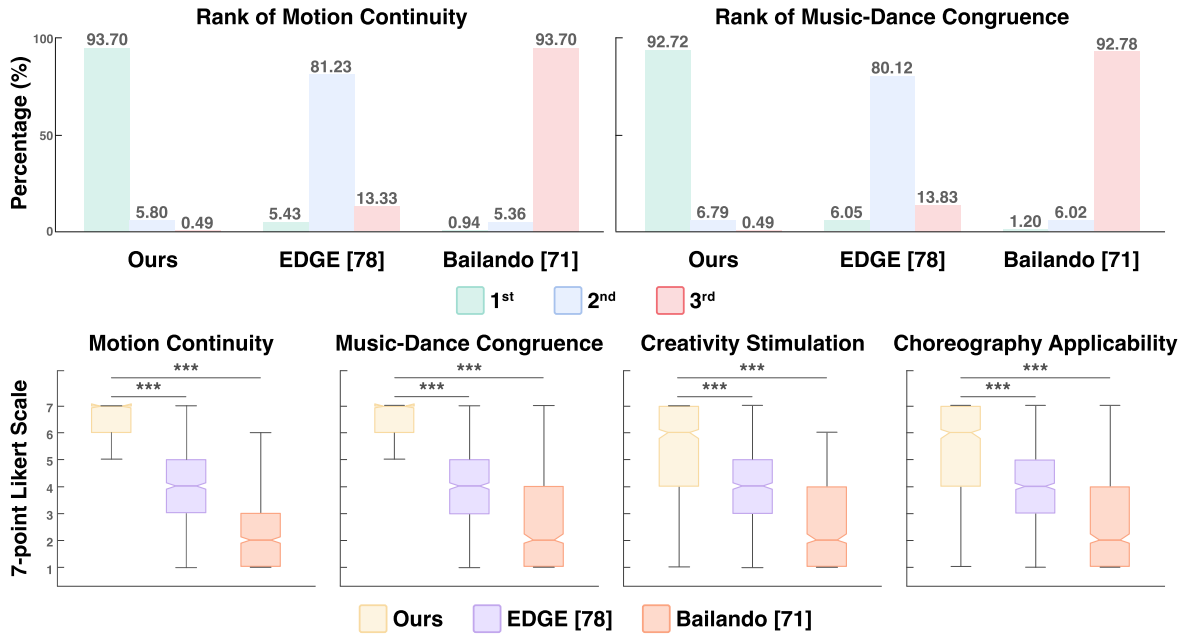
---

[6]https://tally.so/

Figure 6: Result of User Evaluation 1. The above two graphs represent the results of Q1 and Q2, and the below four graphs represent the results of Q3~6 from Table 3. The below four graphs aim to examine the significance between our method and the two GenAI models. The pairs indicated by asterisks showed statistically significant differences according to Dunn's post-hoc test ($* * * : p < 0.001$).

were $\chi^2(2) = 1185.4$, $p < 0.001$ with significant differences against EDGE ($p < 0.001$) and Bailando ($p < 0.001$). For Q5, $\chi^2(2) = 986.47$, $p < 0.001$ with significant differences against EDGE ($p < 0.001$) and Bailando ($p < 0.001$). Lastly, the results were $\chi^2(2) = 945.95$, $p < 0.001$ with significant differences against EDGE ($p < 0.001$) and Bailando ($p < 0.001$). Based on the results, we confirmed the favorable user inclination on our suggestion system over baseline models across all criteria.

5.2 User Evaluation 2: Validation of Choreography Analysis Factors

We have drawn six main choreography analysis factors from design implications in Section 4.3. In this session, we explored the user preference on six analysis factors (F_{eq}, F_{sb}, F_{sj}, F_{nsj}, F_{hj}, F_{sq}.) In order to inspect the users' preference, we prepared four dance motion clips of upper-body and lower-body activated motions using Mixamo[7]. Then, we showed six different analysis factors along with the motion clips. For evaluation, we conducted evaluations through surveys with Google Form[8].

We asked participants "Please rank factors from 1 (most important) to 6 (least important) based on what you need the most in your choreography process." Participants reported a preference for choreography analysis factors with a 7-point Likert scale, along with short responses related to their choices. Here, Friedman test revealed significant differences in participant preferences ($p <$
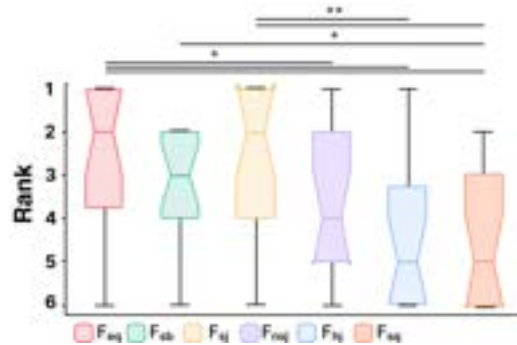

Figure 7: Result of User Evaluation 2. The Y-axis indicates Rank (1: most important, 6: least important). Pairs grouped by asterisks were significantly different by Durbin-Conover post-hoc tests ($* : 0.01 < p < 0.05$, $* * : p < 0.01$).

0.01). Shown in Figure 7, Post-hoc pairwise comparisons (Durbin-Conover) revealed significant differences between F_{hj} and F_{sq} compared to F_{eq}, F_{sb}, and F_{sj} ($p < 0.05$). F_{sb}, and F_{sj} were consistently ranked as the most important factors while there were no significant differences found between F_{sb} and F_{nsj}, F_{eq}. This strong preference for equilibrium (F_{eq}) and stability-related metrics (F_{sb} and F_{sj}) highlights their critical role in ensuring balance and stability during choreography creation. Participants likely prioritized these metrics because they directly contribute to the physical alignment and control needed to craft and refine movements (P3, P4, P15). The

---

[7]https://www.mixamo.com
[8]https://www.google.com/forms

overlap in rankings between F_{nsj}, F_{hj}, and F_{sq}, however, indicates that these factors were perceived as less crucial, suggesting a more specialized or supplementary role in the creative process. Participants preferred equilibrium and feet stability measurements (F_{eq} and F_{sb}) since these factors directly represent the physical balance of the choreography which is crucial component in creating new motion (P2, P7, P10). P7 mentioned "I think F_{sb}, F_{eq}, F_{sj} necessary in the process of creating techniques because it requires an understanding of skilled physical and balance." For visualization of salient joint (F_{sj}), participants (P3, P4, P7, P12) favored the easiness of capturing the activated joints information during the choreography. Participants showed moderate interest in F_{nsj}, F_{hj} and F_{sq} since they were involved in small movements or showed visualization that limited movement for creating choreography.

6 USER STUDY

From the previous evaluation (Section 5), we confirmed the validity of our approach, and preferred choreography motion analysis factors. Then, we moved on to a user study consisting of 2 sessions to investigate our system's main features, including choreography creation in VR, choreography suggestion, and choreography analysis. Each study aimed to validate our motion suggestion and analysis systems. The studies were approved under the IRB protocols.

6.1 Participants and Setup

Participants. We recruited 10 professional choreographers (5 female, mean age of 28.9, σ = 6.06) skilled in diverse dance genres (Table 4). We recruited choreographers who had not been involved in any prior studies to minimize potential biases from prior perceptions. To assess our system for people with a deeper knowledge of professional choreographic processes, we enlisted choreographers who have careers of an average of 8.6 years (σ = 5.4). The user study was conducted for up to 3 hours, and participants were compensated with 100 USD.

Apparatus. We configured our system (Figure 8) in real-time using a VR HMD (Meta Quest 3) [50] and Azure Kinect. Participants used the right-hand controller to control VR interface functions. For the baseline experiment, we set up a mirror (Width: 1m and Height: 1.7m) to mimic a dance studio environment. To avoid disturbance for participants for the baseline experiment, we installed a wall screen to avoid direct observation of their movements. For video recording, the screen was not used for ChoreoCraft.

6.2 Study Design

For the baseline experiment, we asked participants to craft choreography in the original routine, crafting choreography in front of the mirror. There was no time limit set for creating the choreography, and each choreographer was instructed to create the choreography in their usual manner. According to Figure 8, to validate ChoreoCraft, we had a training session with a random song and went through all functions with guidance. Then, we conducted the study starting with the session 1. We conducted the question consecutively and proceeded to session 2. After all sessions ended, we conducted a short qualitative interview.

6.2.1 Study Session 1: Observing the Impact of the Creative Process in VR Environments and Choreography Suggestion Systems on Crafting Choreography. The objectives of this experiment are as follows: (1) verifying the enhancement of creativity in choreographic creation within a VR environment, (2) examining the reduction in memory dependence during the choreographic process, (3) assessing the augmentation of creativity and the inspiration process facilitated by the choreography suggestion system, and (4) evaluating the increase in efficiency compared to the baseline.

The musical selections encompassed a range of BPMs capable of accommodating all 9 dance genres within our choreographic similarity comparison dataset and participants' dance genres. Two songs were chosen at 108 BPM: "One Time Comin'" (YG, 2016) and "In the Club" (Taufiq Akmal, 2022). To facilitate choreography creation for dance genres that are not typically associated with slower BPM genres like house [79], two additional songs at 134 BPM were selected: "Beggin'" (Måneskin, 2017) and "That's What I Like" (Bruno Mars, 2017). We selected songs with high popularity (YouTube view counts exceeding 1 million). We confirmed that participants had no prior experience creating choreography for these specific tracks. Within each BPM category, one song was designated for use with our system ("One Time Comin'" and "Beggin'"), while the other was allocated for the baseline creative approach ("In the Club" and "That's What I Like").

In the actual experiment process, users create choreography in 8-count segments while interacting with the VR environment and avatar using the *snapshot* function. To assess the utility of the choreography suggestion system, participants were instructed to use the suggestions function at least twice during the creation of four 8-count segments. While users composed choreography through the proposed system, the experimenter observed the creative process and recorded it using an additional camera.

Following the completion of Study 1, we administered a survey to the participants using a 7-point Likert scale, addressing the following questions written in Table 5. Q1~4 were designed to evaluate the efficacy of the choreographic creation process utilizing the VR environment and avatars, and Q5–9 were formulated to assess the effectiveness of the choreographic creation process using the choreography suggestion system.

6.2.2 Study Session 2: Assessment of Choreography Analysis System. Before conducting a motion analysis system study, we briefly explained the meanings of each factor that we are providing. Participants are required to playback the final recording from each 8-count they recorded that brought up with associated with choreography analysis factors. We let users freely playback the motions anytime before they are ready to modify the previous choreography. We asked for participants to record the new motion if they were willing to modify their previously recorded motion based on the feedback from the analysis system. If not, they completed the study with a new motion recording. Across two repeated sessions, evaluations were provided sequentially, with each 8-count segment receiving its corresponding analysis. In the final 8-count segment, a combination of F_{eq}, F_{sb}, and F_{sj} was presented. Following the completion of Study 2, we also carried out a short survey of the participants using a 7-point Likert scale with Q10~14 from Table 5.
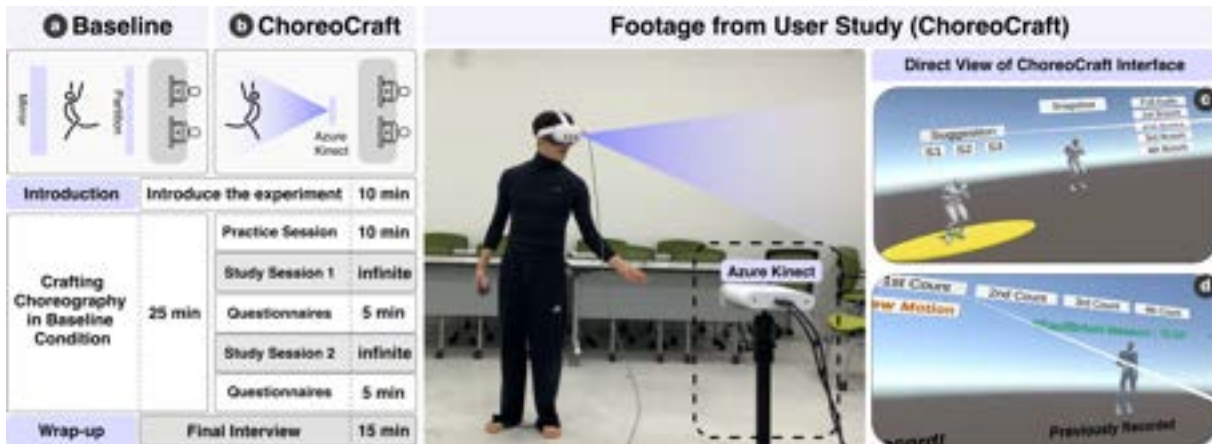
Figure 8: (a) refers to the environment of baseline condition, we settled the mirror and a partition. (b) refers to an environment using ChoreoCraft, wearing HMD, with Azure Kinect set-up. We conducted two sessions, before initiating ChoreCraft, we proceeded with a practice session. Baseline condition and ChoreoCraft conditions were randomly mixed. Here, (c) and (d) refer to the direct view of the ChoreoCraft interface. (c) is when the user is in suggestion mode, while (d) is getting motion analysis.

Table 4: Demographic and Background Information of Participants for User Study

ID	Gender	Dance Experiences	Choreographic Experiences	Main Genre
P1	F	5 years	2 years	Choreography
P2	M	10 years	7 years	Poppin, Bebop
P3	M	8 years	6 years	Hip-Hop, Choreography
P4	M	25 years	22 years	Modern Dance, Hip-Hop
P5	M	15 years	10 years	Locking
P6	F	8 years	6 years	Heel Choreography
P7	M	14 years	12 years	Hip-Hop, Modern Dance
P8	F	10 years	8 years	All genre
P9	F	8 years	6 years	Hip-Hop
P10	F	9 years	7 years	Waacking

6.3 Results

The results from the study showed (1) the impact of VR environment and avatar interaction, (2) the influence of the choreography suggestion system, and (3) the effect of the choreography feedback system on the choreography creation process.

6.3.1 The impact of VR environment and avatar interaction. Choreographers demonstrated positive responses to the process of creating choreography in a VR environment and interacting with avatars. To evaluate the system's effectiveness, participants rated stimulating creativity from creation in VR and interaction with avatars, resolving the memory dependency, improving the creative plateaus, and efficiency of the choreographic process on a 7-point Likert scale as shown in Table 5. Participants rated the extent to which VR choreography creation stimulated creativity (Q1) at an average of 4.85 (σ=1.53). Similarly, they rated the degree to which interaction with avatars in the VR environment stimulated creativity in choreography creation (Q2) at an average of 4.7 (σ=1.38). Participants highlighted several aspects of creation in VR and interaction with avatars. Choreographers expressed that in-situ VR supportive

features enhanced the choreographic process by providing an immersive experience. P2 stated *"VR choreography creation seems to induce high levels of immersion and concentration."* Also, they mentioned opening up new creative possibilities for choreographers. P4 commented *"I felt that the fusion of choreography and technology could present good possibilities for dancers."* In addition, choreography with avatar offers choreographers a refreshing perspective on their movements, inspiring creativity and unique choreographic expressions beyond traditional mirror-based practices. P6 noted, *"Seeing the avatar mimic my movements was novel and entertaining. This aspect could potentially spark creativity and lead to unique movements."* P10 found the experience of dancing with an avatar representation refreshing, saying, *"The experience of re-expressing myself through an avatar and dancing together was new, compared to always seeing myself in the mirror."* However, participants expressed discomfort with the system's inability to accurately track detailed expressions such as the avatar's delays, waves, facial expressions, and finger gestures. The *snapshot* function was favored by participants. All participants utilized *snapshot* feature to review their created choreography and continue the creative process. When

Table 5: Questions and Objectives Employed in User Study

Questions	Objectives
Q1. Please evaluate the extent to which choreographic creation in a VR environment stimulates creativity.	To examine the impact of the overall choreographic creation process in a VR environment on creativity enhancement.
Q2. Assess the degree to which interaction with avatars in the VR environment enhances creative processes in choreography creation.	To investigate the effect of avatar interaction in the VR environment on augmenting creativity in dance composition.
Q3. Evaluate whether the *snapshot* feature in our system reduces instances of choreographic forgetting compared to usual practices.	To assess the influence of the *snapshot* feature in our system on reducing choreographic forgetting.
Q4. Compare the sense of unfamiliarity in the VR and avatar-assisted choreographic process to your typical dance creation environment.	To explore the sense of unfamiliarity when using this system compared to the actual choreographic creation process.
Q5. Assess the convenience of utilizing the dance suggestion system compared to the traditional process of seeking inspiration for choreography (e.g., browsing dance videos on a mobile device during the creative process).	To evaluate the convenience offered by the choreography suggestion system in generating new motions compared to the traditional choreographic process.
Q6. Rate the extent to which the dance suggestion system aids in reducing deliberation time for subsequent choreography creation.	To determine whether the system alleviates creative plateaus in choreographic composition.
Q7. Assess the degree to which the dance suggestion system contributes to inspiring new choreographic ideas.	To investigate whether the choreography suggestion system provides inspiration for dance creation.
Q8. Evaluate the impact of the dance suggestion system on the time efficiency of the choreographic process.	To examine the impact of the choreography suggestion system on the temporal efficiency of dance creation.
Q9. Evaluate your willingness to utilize the choreography suggestion system in future choreographic processes.	To assess the willingness to use the choreography suggestion system in future dance compositions.
Q10. Evaluate how well you generally accept feedback from others on the choreography you crafted.	To assess the openness to external feedback on choreographic decisions in the creative process.
Q11. Evaluate the overall usefulness of the motion analysis system in assisting your choreography creation process.	To determine the utility of the motion analysis system in supporting and enhancing the choreography development process.
Q12. Evaluate the effectiveness of reviewing each evaluation factor individually during the choreography process.	To explore the impact of individually assessing evaluation factors on choreographic refinement and decision-making.
Q13. Evaluate the effectiveness of reviewing evaluation factors all at once during the choreography process.	To examine the benefits of reviewing multiple evaluation factors simultaneously.
Q14. Evaluate whether the numerical analysis effectively supported autonomous interpretation and revision of the choreography.	To investigate the effectiveness of numerical analysis in aiding independent interpretation and iterative refinement of choreography.

asked if the *snapshot* function could reduce instances of forgetting choreography compared to their usual process (Q3), participants gave an average rating of 5.95 (σ=0.86). P2, who struggles with memorizing choreography, found the feature helpful: *"The snapshot function seems capable of showing the movements through the avatar, which greatly aided in reducing the memory dependence."* P3 noted its utility in revising choreography: *"Often when thinking about changing the flow of earlier movements while working on later parts, I forget the choreography. The snapshot function seems to compensate for this."* P4 highlighted its potential for improvisation: *"The system can record spontaneous movements, which could be good for reconstructing dancers' natural expressions."* Other participants (P5, P8) appreciated how the *snapshot* function streamlined their usual process of recording and reviewing choreography with a smartphone. P10 emphasized its ability to capture fleeting moments of inspiration: *"There are many moments when I think, 'Oh, that was a great move, what was it?' But you can't record every moment. This function could help not to forget flashes of ideas."* Additionally, participants valued the system's ability to overcome spatial and temporal constraints. P3 noted, *"It feels like a significant advantage to be able to create choreography at any place and time I want, without needing a mirror."* P6 added, *"...it seems like it could increase the efficiency of choreography creation while using it comfortably at home."*

These findings suggest that the VR choreography system, with the avatar and its *snapshot* function, has the potential to enhance the creative process, and reduce memory dependence, capture spontaneous ideas, and provide flexibility in terms of time and location for choreography creation.

6.3.2 The influence of the choreography suggestion system. Our study revealed distinctive patterns in how choreographers utilized the suggestion system. Participants typically employed one of two approaches: either reconstructing a portion of a single suggested choreography into their composition (P1, P3, P5, P7~P10) or creatively combining elements from two or more suggested choreographies to form a new, unique composition (P2, P4, P6). These patterns demonstrate the system's flexibility in accommodating various creative processes. When assessing the system's convenience compared to traditional inspiration methods during creative plateaus (Q5), participants gave an average rating of 5.3 (σ = 1.56). The system's ability to alleviate creative plateaus (Q6) was rated even higher, with a mean of 5.85 (σ = 1.67). Participants also positively evaluated the system's capacity to provide choreographic inspiration (Q7), giving an average rating of 5 (σ = 1.47).

Qualitative feedback from participants provided deeper insights into the system's impact on their creative process. P1 noted the

system's ability to suggest innovative motions, stating, *"The suggested motions were innovative, offering body utilizations and basic movements I had yet to consider."* P2 and P4 also added, *"It aids in establishing the overall structure of the choreography. I can lay out the framework and then refine the details myself."* P5 found that it encouraged them to break free from genre-specific constraints, noting, *"The system reminded me of forgotten movements and inspired me to attempt new motions. It allows for breaking free from genre-specific choreography and exploring new creative motions."*

The system's impact on choreographic efficiency (Q8) was notably positive, with a high average rating of 5.6 (σ=1.76). P1 reported, *"It seems to maximize efficiency by eliminating wasted time in the choreographic process."* P5 relied more on the system as their process progressed, saying, *"I found myself increasingly relying on the suggestion system as the process progressed, thinking, 'Oh, there is this step? Let me adapt it to my style.'"* P6 highlighted its potential for quickly creating choreography under time constraints: *"The system could be particularly beneficial when needing to create choreography quickly and without interruptions, such as when preparing for student classes."* Participants expressed a strong willingness to use the system in the future (Q9) with a high average rating of 5.5 (σ=1.78).

In conclusion, these results indicate that the choreography suggestion system effectively supports the creative process, particularly in overcoming creative plateaus and enhancing choreographic efficiency. The system demonstrates promise for both experienced and novice choreographers, offering a valuable tool for inspiration and workflow optimization in dance composition. Its ability to suggest unexpected motions while allowing choreographers to maintain creative autonomy is key to its success.

6.3.3 The impact of the choreography analysis system. Through comparative data, we observed significant changes in key aspects of choreography: motion equability (F_{eq}), stability (F_{sb}), and engagement (F_{sj}). The extent of these modifications was conducted to determine the degree of influence the system had solely on the choreographers' decisions, providing concrete evidence of its impact on refining and improving dance sequences. Therefore, participants (P2, P7) skipped the modification process and mentioned *"The score of factor is very well reflected, but this is already high enough (F_{eq}), so I would not want to disrupt the score"* and reported they normally prefer not to accept external feedback (Q10) with averaging rate of 5.4 (σ = 1.07). However, we observed that 85% of participants considered modifying their crafted choreography after going through analysis in our system and rated the usefulness (Q11) with a high averaging rate of 5.6 (σ = 0.94). P4, P5, P8 utilized F_{eq} factor to edit motions, noting *"I initially felt that the energy level of my motion was slightly low, but the review of the precise numerical data made this more evident"*, *"This analysis enhanced my ability to perceive and analyze the details thoroughly, inspiring new ideas for motion creation"*. Participants also reported their autonomous interpretation and understanding of their choreography (Q14) in a high averaging score rate of 5.6 (σ = 1.26).

In terms of numerical changes observed during the modification process, F_{eq} score increased from an average of 85.46 (σ = 6.29) to 90.85 (σ = 3.77), with a maximum possible score of 100 (See Figure 9a). All the data net the normality assumption according to the Shapiro-Wilk test (W = 0.965, p = 0.860), and a paired samples t-test

was conducted to assess the significance of these changes, revealing a statistically significant improvement in motion scores (t(5) = -2.16, p < 0.05). Similarly, the Wilcoxon signed-rank test confirmed the significance of the difference (p < 0.05). Furthermore, the effect size, as measured by Cohen's d (d = -0.880), indicates a large practical impact of these modifications, demonstrating an improvement in the F_{eq} level and overall equability of choreography.

ChoreoCraft facilitated a quantitative analysis of choreography under two distinct conditions: participants were exposed to either a single factor representation or all three selected factors simultaneously. According to Table 5 (Q12) and (Q13), participants expressed a preference for receiving individual factor feedback (μ=6.1, σ=0.85) over-viewing multiple factors at once (μ=4.75, σ=1.29), noting *"It is hard to think what to focus on when there are too many factors on my movement."* (P6). This result suggests that presenting factors individually is more effective for users.
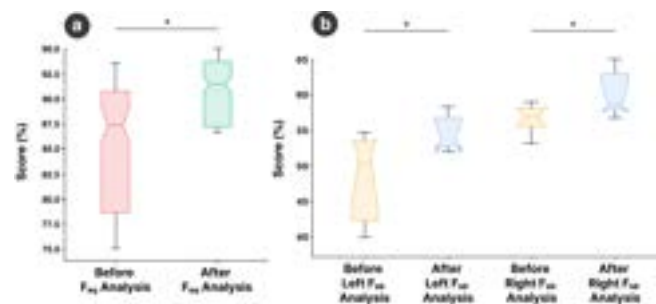
Figure 9: (a) shows the results of F_{eq} score before and after analysis, pairs showed significance in t-test, (b) describes the results of F_{sb} score before and after analysis of each left and right foot, pairs showed significance in t-test (* : p < 0.05).

Regarding F_{sb}, Participants (P1, P3, P8, and P10) commented *"The evaluated difference in energy between both feet was noticeable, which dragged me to refer and modify the motion to equalize them."* The quantitative rate showed that modified choreography did increase the F_{sb}. The average F_{sb} factors for the left and right sides were average of 47.96% and 52.27% for their initial trial and improved to 55.64% and 56.46% in the modified trial. The paired t-test results showed t(7) = -2.54, p < 0.05, indicating a meaningful improvement in F_{sb} factors after modification. The Wilcoxon signed-rank test yielded W = 1.00, p < 0.05, reinforcing this finding. The effect sizes, Cohen's d = -0.898 and a rank bi-serial correlation of -0.929, both suggest a substantial reduction in the energy imbalance between the feet, leading to more balanced and stable movements (See Figure9b).

F_{sj} was the most utilized analysis factor where 85% (8/10) of the participants referred to the analysis and modified the choreography. Noting *"Being left-handed, I tend to rely on the left side of my body. The system accurately detects this, which encourages me to engage other body segments"*. Among the participants who adjusted their movements based on the F_{sj} feedback, an average of 2 non-salient joints (σ = 0.86) were activated. This corresponds to improvements of about 10% for promoting joint activation of the 20 remaining joints that had not been previously activated.

6.3.4 Post Study on Peer-Reviewed Choreography Evaluation. We conducted a post-study on the peer-review process to evaluate choreographies created under three distinct conditions: baseline (using mirrors), utilizing the VR environment with choreography suggestion (VR), and choreography revised after using the choreography analysis system (Revised). Choreographer pairs with similar experience levels were assigned to assess each other's work, ensuring consistent feedback. For instance, P2 and P10, both with seven years of experience, were paired to evaluate the completeness and creativity of each other's choreographies. Participants evaluated each choreography based on three key questions represented in Table 6. We collected a total of 180 responses (10 participants × 6 choreography videos × 3 questions). Each participant rated four choreography creation videos on a 10-point scale, assessing how each choreography aligned with the posed questions. For tasks where a large number of responses could not be generated compared to the user evaluation, we adopted a 10-point scale to ensure more definitive differentiation. This approach allowed us to aggregate more conservative results, thereby enhancing reliability [23].

Table 6: Evaluation Questions for Post Study

Questions for Post Study
Q1. Please evaluate how well the choreography harmonizes with the music.
Q2. Please assess how organically the movements within the choreography are structured.
Q3. Please evaluate the creativity of the choreography.

The Friedman test was employed to determine significance across the three conditions for each of the three questions. Although significant results were not obtained for Q1 and Q3, Q2 showed some statistical significance. For Q2, Nemenyi post-hoc tests were conducted to explore differences between the videos, but no significant pairwise differences were identified. Despite the lack of statistical significance, there were observable trends in the mean scores: Q1 (Harmonization with music): Mirror ($\mu = 7.75, \sigma = 1.51$), VR ($\mu = 8.05, \sigma = 1.53$), Revised ($\mu = 8.15, \sigma = 1.42$); Q2 (Organic movement structure): Mirror ($\mu = 7.55, \sigma = 1.60$), VR ($\mu = 8, \sigma = 1.58$), Revised ($\mu = 8.05, \sigma = 1.66$); Q3 (Creativity): Mirror ($\mu = 7.7, \sigma = 1.24$), VR ($\mu = 7.8, \sigma = 1.36$), Revised ($\mu = 8.05, \sigma = 1.20$) (Figure 10). These results suggest that choreographies created using the VR environment, and especially those refined with the choreography analysis system, exhibited slightly higher completeness and creativity compared to those created using mirrors alone. While these findings were not statistically significant, the observed trends imply that dancers accustomed to mirror-based choreography may produce higher-quality results when utilizing our system. The absence of strong statistical significance in these results highlights the need for further research. A larger sample size and more choreography creations would likely provide a more robust assessment of the system's effectiveness. Such extended studies could help to establish stronger evidence for the positive impact of our system on the choreographic creation process.

When considering the previous results, this system offers an opportunity to create choreography with high efficiency while addressing the memory dependence of the traditional choreographic process. This suggests that compared to the method of creating with mirrors, it enables the production of choreography of comparable or higher quality more efficiently without being constrained by spatial or temporal limitations.

6.3.5 Assessing the Impact of Choreography Analysis Comparing to Self-Reflection From Baseline Condition. To confirm if the effect in Section 6.3.3 directly stems from the ChoreoCraft analysis system, we conducted an additional study focusing solely on the analysis component. We recruited 8 professional choreographers (6 females) with varied levels of choreographic experience (5~22 years, $\mu = 9.12, \sigma = 5.48$). In contrast to Section 6, the additional study focused exclusively on the analysis feature of the ChoreoCraft system.

Participants alternated between baseline and ChoreoCraft conditions using counterbalanced sequences across trials to mitigate the order effect. For a baseline, participants relied on traditional methods, crafting choreography in front of a mirror and conducting self-reflection using their own recordings. ChoreoCraft condition adapted the same initial choreography crafted in the baseline condition but allowed participants to revise it using quantitative feedback provided by the system. We conducted the Shapiro-Wilk Test and confirmed that both baseline and ChoreoCraft conditions satisfy the assumption of normality. For further analysis, we conducted a paired t-test, and the results confirmed that ChoreoCraft significantly outperformed the baseline condition across all metrics F_{eq}, F_{sb}. For F_{eq}, the percent improvement under ChoreoCraft was substantially greater ($\mu = 7.92\%$) compared to baseline ($\mu = 3.06\%$), with a $p < 0.01$. Similarly, left F_{sb} showed a mean improvement of 5.49% (ChoreoCraft) versus 2.35% (Baseline), with a $p < 0.05$. While the improvement in right F_{sb} was not statistically significant ($p = 0.054$), the trend still favored ChoreoCraft.

By employing a counterbalanced design where participants alternated between baseline and ChoreoCraft conditions, we aimed to mitigate potential sequence biases. This approach allowed us to assess whether ChoreoCraft consistently facilitated better choreographic refinement and ensured that the observed improvements were not merely the result of task repetition or familiarity. Through conducting t-test, analysis of relative delta ($\Delta_{relative} = \Delta_{ChoreoCraft} - \Delta_{baseline}$) for counterbalanced groups (Group A: baseline → ChoreoCraft; Group B: ChoreoCraft → baseline) revealed no significant difference ($p = 0.587$). This finding confirms that ChoreoCraft's effectivness is independent of sequence effects and is directly attributable to ChoreoCraft's choreography analysis feedback.

7 DISCUSSION

Addressing key considerations for improved VR experiences. A significant challenge involved the fidelity of the avatar's movements, with participants noting discrepancies between their physical movements and the avatar's rendering. This created a sense of disconnection, particularly for detailed facial expressions or delicate dance movements (body waves or isolation movements). Addressing these issues requires integrating advanced motion tracking
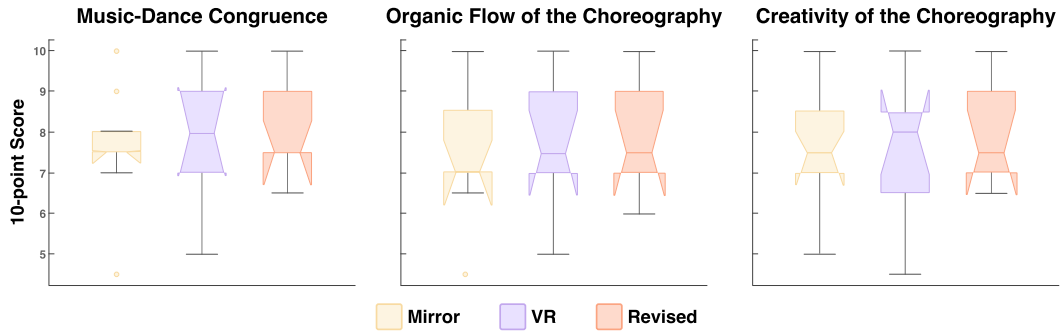
Figure 10: Result of Post Study. Three graphs represent the result of Q1~3 from Table 6.

Table 7: Demographic and Background Information of Participants for Additional User Study

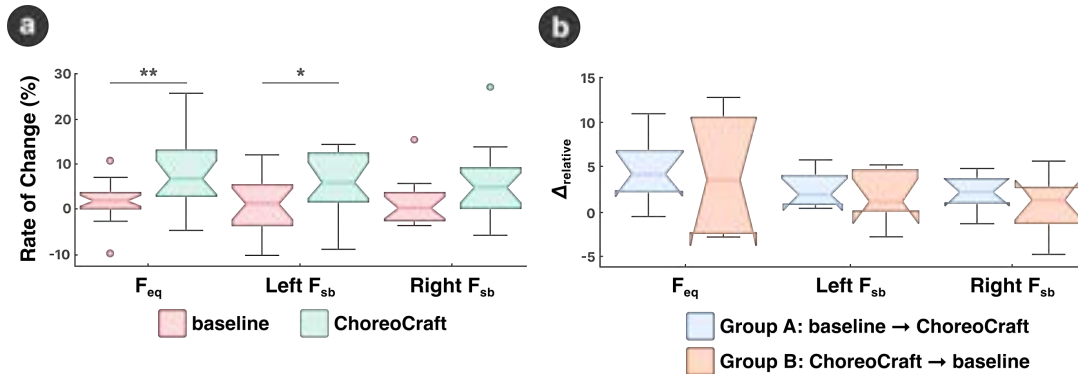| ID	Gender	Dance Experiences	Choreographic Experiences	Main Genre
P1	F	5.5 years	5.5 years	Girls Hip-Hop, Waacking
P2	F	25 years	10 years	Korean Modern Dance
P3	F	10 years	7 years	Waacking
P4	F	5 years	4 years	Hip-Hop, Waacking
P5	M	13 years	13 years	Poppin
P6	F	15 years	10 years	Jazz, Korean Modern Dance
P7	F	22 years	20 years	Modern Dance
P8	M	5 years	3.5 years	Choreography

Figure 11: Result of comparative study demonstrating ChoreoCraft's effectiveness relative to the baseline condition. (a) illustrates the percentage improvements across the conditions (baseline and ChoreoCraft) for metrics F_{eq}, left and right F_{sb}. (b) refers the delta comparison between Group A (baseline → ChoreoCraft) and Group B (ChoreoCraft → baseline).

systems [54, 85] to improve the synchronization between physical and virtual motions, thereby enhancing user immersion. Additionally, multiple participants mentioned the physical burden and constraints on dance motions such as moving hands close to the head or rolling on the floorcaused by weight of the HMD and wired connection. Recent advancements in the lightweight design of VR HMDs [15, 49] can help prevent physical fatigue caused by the weight of the device. While the current system utilizes wired connections to employ real-time motion tracking, adopting a wireless and real-time motion tracking method [21] enables more unrestricted dance movements.

Enhanced creative inspiration driven by diverse choreography analysis factors. The system exclusively provides an objective score or visualization without additional prescribing steps of what to proceed to improve the score. The result of Q10 from Table 5 demonstrates that encouraging choreographers to independently interpret the evaluation stimulated their ability to refine their movements with their own autonomous decisions. P8 from user study supports our hypothesis that *"Encountering analyzed score makes me reflect on my previous movements and encourages me to explore new approaches"*. However, some novice choreographers (<2 years of experience) often struggled to make progress after interpreting the

score. This suggests that ChoreoCraft further needs to consider developing an automated score system to provide adaptive guidance considering the user's expert knowledge.

Suggestion and future directions for representing choreography analysis factors. From the results, we found out that providing individual factor feedback one at a time is easier and more approachable for users rather than providing multiple factors simultaneously. Choreographers mentioned it was more clear and intuitive to investigate the movement when a sole factor was provided. Therefore, we suggest providing a single analysis factor rather than a combined presentation for future iterations. For future versions of ChoreoCraft, we plan to allow high flexibility so users can select their favored analysis factors to work with. Additionally, while the current use of Azure Kinect for motion capture has proven effective in quantifying macro-level dynamics and analyzing foot stability through raw and filtered data, it has limitations. Specifically, the system limits to capture subtle movements. To address these challenges, we aim to incorporate wearable sensors in future versions of ChoreoCraft. This enhancement would enable a more comprehensive analysis of muscle engagement and balance, providing choreographers with richer and more detailed feedback to support their creative process.

Towards better understanding of choreography analysis system. We found that additional practice sessions for understanding motion analysis scores were necessary. As the sessions progressed, participants gained a better understanding of the numerical feedback and were able to refine their choreography details with more creative and informed decisions. P5 from the additional user study specifically noted that they initially relied on the instructor's explanations to understand the scoring metric. Then, they got used to the score system and visual indicators to refine their movements over time. This highlights that structured guidance and some practice sessions would be required for the first time users to effectively leverage choreography analysis system.

8 CONCLUSION

In this paper, we proposed ChoreoCraft, the first VR choreography crafting supportive tool. Our results demonstrated that ChoreoCraft helped overcome the creative plateau with the choreography suggestion system and supported constructive feedback with choreography analysis factors. To bring these forward, we incorporated a novel dance similarity comparison method (DanceDTW) to suggest motion aligned with the context. By offering kinematic feature-based motion analysis, systematic feedback for the choreography becomes possible. Throughout multiple studies with professional choreographers, participants favored the aspect of an in situ inspirational system with objective feedback availability. We believe that all levels of choreographers, either novice or expert, would benefit from using ChoreoCraft in various stages of the choreography creation process. With continued advancement in motion generation and analysis technologies, we envision ChoreoCraft would become a central creative support tool to further support choreographers to nurture and shape the future of the choreography creation process.

REFERENCES

[1] Omid Alemi, Philippe Pasquier, and Chris Shaw. 2014. Mova: Interactive movement analytics platform. In *Proceedings of the 2014 International Workshop on Movement and Computing*. 37–42.

[2] Simon Alexanderson, Rajmund Nagy, Jonas Beskow, and Gustav Eje Henter. 2023. Listen, denoise, action! audio-driven motion synthesis with diffusion models. *ACM Transactions on Graphics (TOG)* 42, 4 (2023), 1–20.

[3] Dimitrios S Alexiadis, Philip Kelly, Petros Daras, Noel E O'Connor, Tamy Boubekeur, and Maher Ben Moussa. 2011. Evaluating a dancer's performance using kinect-based skeleton tracking. In *Proceedings of the 19th ACM international conference on Multimedia*. 659–662.

[4] Andreas Aristidou, Nefeli Andreou, Loukas Charalambous, Anastasios Yiannakidis, Yiorgos Chrysanthou, V Hulusic, and A Chalmers. 2021. Virtual Dance Museum: the Case of Greek/Cypriot Folk Dancing.. In *GCH*. 51–60.

[5] Ian Baracskay, Donald J Baracskay III, Mehtab Iqbal, and Bart Piet Knijnenburg. 2022. The diversity of music recommender systems. In *Companion Proceedings of the 27th International Conference on Intelligent User Interfaces*. 97–100.

[6] Hugh Beyer and Karen Holtzblatt. 1999. Contextual design. *interactions* 6, 1 (1999), 32–42.

[7] Klara Brandstätter and Anthony Steed. 2023. Dialogues For One: Single-User Content Creation Using Immersive Record and Replay. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*. 1–11.

[8] Thomas W. Calvert, Christopher Welman, Severin Gaudet, and Catherine Lee. 1989. Composition of multiple figure sequences for dance and animation. In *New Advances in Computer Graphics: Proceedings of CG International'89*. Springer, 245–255.

[9] Antonio Camurri, Corrado Canepa, Nicola Ferrari, Maurizio Mancini, Radoslaw Niewiadomski, Stefano Piana, Gualtiero Volpe, Jean-Marc Matos, Pablo Palacio, and Muriel Romero. 2016. A system to support the learning of movement qualities in dance: a case study on dynamic symmetry. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. 973–976.

[10] Antonio Camurri, Gualtiero Volpe, Stefano Piana, Maurizio Mancini, Radoslaw Niewiadomski, Nicola Ferrari, and Corrado Canepa. 2016. The dancer in the eye: towards a multi-layered computational framework of qualities in movement. In *Proceedings of the 3rd International Symposium on Movement and Computing*. 1–7.

[11] Julio José Carabias-Orti, Francisco J Rodriguez-Serrano, Pedro Vera-Candeas, Nicolás Ruiz-Reyes, and Francisco J Canadas-Quesada. 2015. An Audio to Score Alignment Framework Using Spectral Factorization and Dynamic Time Warping.. In *ISMIR*. 742–748.

[12] Kristin Carlson, Thecla Schiphorst, Karen Cochrane, Jordon Phillips, Herbert H Tsang, and Tom Calvert. 2015. Moment by moment: Creating movement sketches with camera stillframes. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*. 131–140.

[13] Caroline Chan, Shiry Ginosar, Tinghui Zhou, and Alexei A Efros. 2019. Everybody dance now. In *Proceedings of the IEEE/CVF international conference on computer vision*. 5933–5942.

[14] Yanping Chen, Eamonn Keogh, Bing Hu, Nurjahan Begum, Anthony Bagnall, Abdullah Mueen, and Gustavo Batista. 2015. The UCR Time Series Classification Archive. www.cs.ucr.edu/~eamonn/time_series_data/.

[15] Dewen Cheng, Qichao Hou, Yang Li, Tian Zhang, Danyang Li, Yilun Huang, Yue Liu, Qiwei Wang, Weihong Hou, Tong Yang, et al. 2022. Optical design and pupil swim analysis of a compact, large EPD and immersive VR head mounted display. *Optics express* 30, 5 (2022), 6584–6602.

[16] John Joon Young Chung, Shiqing He, and Eytan Adar. 2021. The intersection of users, roles, interactions, and technologies in creativity support tools. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference*. 1817–1833.

[17] Marianela Ciolfi Felice, Sarah Fdili Alaoui, and Wendy E Mackay. 2016. How do choreographers craft dance? Designing for a choreographer-technology partnership. In *Proceedings of the 3rd International Symposium on Movement and Computing*. 1–8.

[18] Marianela Ciolfi Felice, Sarah Fdili Alaoui, and Wendy E Mackay. 2018. Knotation: exploring and documenting choreographic processes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.

[19] Valve Corporation. 2024. *Dance Dash*. https://store.steampowered.com/app/2005050/Dance_Dash.

[20] Wojciech J Cynarski, Wiesław Wajs, and António VencesBrito. 2014. Improving the movements of basic karate techniques with the use of motion capture and mathematical modeling. The concept of a research project. *Ido Movement for Culture. Journal of Martial Arts Anthropology* 1, 14 (2014), 39–53.

[21] Peng Dai, Yang Zhang, Tao Liu, Zhen Fan, Tianyuan Du, Zhuo Su, Xiaozheng Zheng, and Zeming Li. 2024. HMD-Poser: On-Device Real-time Human Motion Tracking from Scalable Sparse Observations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 874–884.

[22] Nicholas Davis, Chih-PIn Hsiao, Kunwar Yashraj Singh, Lisa Li, Sanat Moningi, and Brian Magerko. 2015. Drawing apprentice: An enactive co-creative agent

for artistic collaboration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*. 185–186.

[23] John Dawes. 2008. Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International journal of market research* 50, 1 (2008), 61–104.

[24] Kaixu Dong, Zhiyuan Zhang, Xiaoyu Chang, Pakpong Chirarattananon, and RAY LC. 2024. Dances with Drones: Spatial Matching and Perceived Agency in Improvised Movements with Drone and Human Partners. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–16.

[25] Sara Eriksson, Åsa Unander-Scharin, Vincent Trichon, Carl Unander-Scharin, Hedvig Kjellström, and Kristina Höök. 2019. Dancing with drones: Crafting novel artistic expressions through intercorporeality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[26] Sarah Fdili Alaoui, Kristin Carlson, Shannon Cuykendall, Karen Bradley, Karen Studd, and Thecla Schiphorst. 2015. How do experts observe movement?. In *Proceedings of the 2nd International Workshop on Movement and Computing*. 84–91.

[27] Andrew S Forsberg, JJ LaViola, Lee Markosian, and Robert C Zeleznik. 1997. Seamless interaction in virtual reality. *IEEE Computer Graphics and Applications* 17, 6 (1997), 6–9.

[28] Jules Françoise, Sarah Fdili Alaoui, and Yves Candau. 2022. Co/da: Live-coding movement-sound interactions for dance improvisation. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–13.

[29] Peter I Frazier. 2018. A tutorial on Bayesian optimization. *arXiv preprint arXiv:1807.02811* (2018).

[30] Jibin Gao, Junfu Pu, Honglun Zhang, Ying Shan, and Wei-Shi Zheng. 2022. Pc-dance: Posture-controllable music-driven dance synthesis. In *Proceedings of the 30th ACM international conference on multimedia*. 1261–1269.

[31] Petra Gemeinboeck and Rob Saunders. 2017. Movement matters: How a robot becomes body. In *Proceedings of the 4th international conference on movement Computing*. 1–8.

[32] Kehong Gong, Dongze Lian, Heng Chang, Chuan Guo, Zihang Jiang, Xinxin Zuo, Michael Bi Mi, and Xinchao Wang. 2023. Tm2d: Bimodality driven 3d dance generation via music-text integration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 9942–9952.

[33] Inc. Harmonix Music Systems. 2024. *Dance Central*. https://www.dancecentral.com/

[34] Karen Holtzblatt and Hugh Beyer. 1997. *Contextual design: defining customer-centered systems*. Elsevier.

[35] Qiaochu Huang, Xu He, Boshi Tang, Haolin Zhuang, Liyang Chen, Shuochen Gao, Zhiyong Wu, Haozhi Huang, and Helen Meng. 2024. Enhancing Expressiveness in Dance Generation Via Integrating Frequency and Music Style Information. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 8185–8189.

[36] Leva Janauskaitė and George Palamas. 2019. Establishing dialogues between movement and atmospheric ambiances. In *Proceedings of the 6th International Conference on Movement and Computing*. 1–11.

[37] Youngseung Jeon, Seungwan Jin, Patrick C Shih, and Kyungsik Han. 2021. FashionQ: an ai-driven creativity support tool for facilitating ideation in fashion design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.

[38] Elizabeth Jochum and Jeroen Derks. 2019. Tonight we improvise! real-time tracking for human-robot improvisational dance. In *Proceedings of the 6th International Conference on Movement and Computing*. 1–11.

[39] Kyungeun Jung, Sangpil Kim, Seungjae Oh, and Sang Ho Yoon. 2024. HapMotion: motion-to-tactile framework with wearable haptic devices for immersive VR performance experience. *Virtual Reality* 28, 1 (2024), 13.

[40] Pavel Karpashevich, Eva Hornecker, Michaela Honauer, and Pedro Sanches. 2018. Reinterpreting Schlemmer's Triadic Ballet: interactive costume for unthinkable movements. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[41] Daniel Klug. 2020. " It took me almost 30 minutes to practice this". Performance and Production Practices in Dance Challenge Videos on TikTok. *arXiv preprint arXiv:2008.13040* (2020).

[42] Markus Laattala, Roosa Piitulainen, Nadia M Ady, Monica Tamariz, and Perttu Hämäläinen. 2024. WAVE: Anticipatory Movement Visualization for VR Dancing. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–9.

[43] Kate Ladenheim, Reika McNish, Wali Rizvi, and Amy LaViers. 2020. Live dance performance investigating the feminine cyborg metaphor with a motion-activated wearable robot. In *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*. 243–251.

[44] Jingyi Li, Eric Rawn, Jacob Ritchie, Jasper Tran O'Leary, and Sean Follmer. 2023. Beyond the Artifact: Power as a Lens for Creativity Support Tools. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–15.

[45] Ruilong Li, Shan Yang, David A Ross, and Angjoo Kanazawa. 2021. Ai choreographer: Music conditioned 3d dance generation with aist++. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 13401–13412.

[46] Yimeng Liu and Misha Sra. 2024. DanceGen: Supporting Choreography Ideation and Prototyping with Generative AI. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 920–938.

[47] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. 2015. SMPL: A Skinned Multi-Person Linear Model. *ACM Trans. Graphics (Proc. SIGGRAPH Asia)* 34, 6 (Oct. 2015), 248:1–248:16.

[48] Zhenye Luo, Min Ren, Xuecai Hu, Yongzhen Huang, and Li Yao. 2024. POPDG: Popular 3D Dance Generation with PopDanceSet. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 26984–26993.

[49] Andrew Maimone and Junren Wang. 2020. Holographic optics for thin and lightweight virtual reality. *ACM Transactions on Graphics (TOG)* 39, 4 (2020), 67–1.

[50] Meta. 2023. Meta Quest 3. https://www.meta.com/quest/quest-3. Accessed: 2024-09-09.

[51] Microsoft. 2020. Azure Kinect DK. https://learn.microsoft.com/ko-kr/azure/kinect-dk/.

[52] Luis Molina-Tanco, Carmen García-Berdonés, and Arcadio Reyes-Lecuona. 2017. The Delay mirror: A technological innovation specific to the dance studio. In *Proceedings of the 4th International Conference on Movement Computing*. 1–6.

[53] Meinard Müller. 2007. Dynamic time warping. *Information retrieval for music and motion* (2007), 69–84.

[54] NaturalPoint, Inc. 2024. OptiTrack Motion Capture Systems. https://optitrack.com. Accessed: 2024-12-11.

[55] Radoslaw Niewiadomski, Ksenia Kolykhalova, Stefano Piana, Paolo Alborno, Gualtiero Volpe, and Antonio Camurri. 2019. Analysis of movement quality in full-body physical activities. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 9, 1 (2019), 1–20.

[56] Andrea Orlandi, Emily S Cross, and Guido Orgs. 2020. Timing is everything: Dance aesthetics depend on the complexity of movement kinematics. *Cognition* 205 (2020), 104446.

[57] Antti Oulasvirta, Teemu Roos, Arttu Modig, and Laura Leppänen. 2013. Information capacity of full-body movements. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1289–1298.

[58] Chaeyong Park and Seungmoon Choi. 2024. Augmenting Perceived Length of Handheld Controllers: Effects of Object Handle Properties. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–11.

[59] Roosa Piitulainen, Perttu Hämäläinen, and Elisa D Mekler. 2022. Vibing together: Dance experiences in social virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.

[60] Aggelos Pikrakis, Sergios Theodoridis, and Dimitris Kamarotos. 2003. Recognition of isolated musical patterns using context dependent dynamic time warping. *IEEE Transactions on Speech and Audio Processing* 11, 3 (2003), 175–183.

[61] Katerina El Raheb, George Tsampounaris, Akrivi Katifori, and Yannis Ioannidis. 2018. Choreomorphy: A whole-body interaction experience for dance improvisation and visual experimentation. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*. 1–9.

[62] Bhuvaneswari Sarupuri, Richard Kulpa, Andreas Aristidou, and Franck Multon. 2024. Dancing in virtual reality as an inclusive platform for social and physical fitness activities: a survey. *The Visual Computer* 40, 6 (2024), 4055–4070.

[63] R Scattone-Silva, GC Lessi, DFM Lobato, and FV Serrão. 2012. Acceleration time, peak torque and time to peak torque in elite karate athletes. *Science & Sports* 27, 4 (2012), e31–e37.

[64] Rodrigo Schramm, Cláudio Rosito Jung, and Eduardo Reck Miranda. 2014. Dynamic time warping for music conducting gestures evaluation. *IEEE Transactions on Multimedia* 17, 2 (2014), 243–255.

[65] Samsu Sempena, Nur Ulfa Maulidevi, and Peb Ruswono Aryan. 2011. Human action recognition using dynamic time warping. In *Proceedings of the 2011 international conference on electrical engineering and informatics*. IEEE, 1–5.

[66] Simon Senecal, Niels A Nijdam, Andreas Aristidou, and Nadia Magnenat-Thalmann. 2020. Salsa dance learning evaluation and motion analysis in gamified virtual reality environment. *Multimedia Tools and Applications* 79, 33 (2020), 24621–24643.

[67] Soyong Shin, Juyong Kim, Eni Halilaj, and Michael J Black. 2024. Wham: Reconstructing world-grounded humans with accurate 3d motion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2070–2080.

[68] Mohammad Shokoohi-Yekta, Bing Hu, Hongxia Jin, Jun Wang, and Eamonn Keogh. 2017. Generalizing DTW to the multi-dimensional case requires an adaptive approach. *Data mining and knowledge discovery* 31 (2017), 1–31.

[69] Vikash Singh, Celine Latulipe, Erin Carroll, and Danielle Lottridge. 2011. The choreographer's notebook: a video annotation system for dancers and choreographers. In *Proceedings of the 8th ACM Conference on Creativity and Cognition*. 197–206.

[70] Li Siyao, Tianpei Gu, Zhitao Yang, Zhengyu Lin, Ziwei Liu, Henghui Ding, Lei Yang, and Chen Change Loy. 2024. Duolando: Follower gpt with off-policy reinforcement learning for dance accompaniment. *arXiv preprint arXiv:2403.18811*

(2024).

[71] Li Siyao, Weijiang Yu, Tianpei Gu, Chunze Lin, Quan Wang, Chen Qian, Chen Change Loy, and Ziwei Liu. 2022. Bailando: 3d dance generation by actor-critic gpt with choreographic memory. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 11050–11059.

[72] Li Siyao, Weijiang Yu, Tianpei Gu, Chunze Lin, Quan Wang, Chen Qian, Chen Change Loy, and Ziwei Liu. 2023. Bailando++: 3d dance gpt with choreographic memory. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023).

[73] Chuan-Jun Su. 2013. Personal rehabilitation exercise assistant with kinect and dynamic time warping. *International Journal of Information and Education Technology* 3, 4 (2013), 448–454.

[74] Jiangxin Sun, Chunyu Wang, Huang Hu, Hanjiang Lai, Zhi Jin, and Jian-Fang Hu. 2022. You never stop dancing: Non-freezing dance generation via bank-constrained manifold projection. *Advances in Neural Information Processing Systems* 35 (2022), 9995–10007.

[75] Tan Tang, Yanhong Wu, Peiquan Xia, Wange Wu, Xiaosong Wang, and Yingcai Wu. 2023. PColorizor: Re-coloring Ancient Chinese Paintings with Ideorealm-congruent Poems. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–15.

[76] Yayuan Tian. 2023. Modern Demands and Strategies of Dance Choreography. *International Journal of Education and Humanities* 11, 3 (2023), 487–492.

[77] Milka Trajkova, Duri Long, Manoj Deshpande, Andrea Knowlton, and Brian Magerko. 2024. Exploring Collaborative Movement Improvisation Towards the Design of LuminAI—a Co-Creative AI Dance Partner. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–22.

[78] Jonathan Tseng, Rodrigo Castellon, and Karen Liu. 2023. Edge: Editable dance generation from music. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 448–458.

[79] Shuhei Tsuchida, Satoru Fukayama, Masahiro Hamasaki, and Masataka Goto. 2019. AIST Dance Video Database: Multi-genre, Multi-dancer, and Multi-camera Database for Dance Information Processing. In *Proceedings of the 20th International Society for Music Information Retrieval Conference, ISMIR 2019*. Delft, Netherlands, 501–510.

[80] Guillermo Valle-Pérez, Gustav Eje Henter, Jonas Beskow, Andre Holzapfel, Pierre-Yves Oudeyer, and Simon Alexanderson. 2021. Transflower: probabilistic autoregressive dance generation with multimodal attention. *ACM Transactions on Graphics (TOG)* 40, 6 (2021), 1–14.

[81] Eduardo Velloso, Andreas Bulling, and Hans Gellersen. 2013. MotionMA: Motion modelling and analysis by demonstration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1309–1318.

[82] Jiahao Wang, Yunhong Wang, Nina Weng, Tianrui Chai, Annan Li, Faxi Zhang, and Sansi Yu. 2022. Will you ever become popular? learning to predict virality of dance clips. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 18, 2 (2022), 1–24.

[83] Edward C Warburton, Margaret Wilson, Molly Lynch, and Shannon Cuykendall. 2013. The cognitive benefits of movement reduction: evidence from dance marking. *Psychological Science* 24, 9 (2013), 1732–1739.

[84] Lauren Wingenroth. 2022. *How to Get Meaningful Feedback on Your Choreographic Work—and What to do With it*. https://www.dancemagazine.com/choreography-feedback/#gsc.tab=0

[85] Xsens Technologies B.V. 2024. Xsens Motion Capture Systems. https://www.xsens.com. Accessed: 2024-12-11.

[86] Zihan Yan, Chunxu Yang, Qihao Liang, and Xiang'Anthony' Chen. 2023. XCreation: A Graph-based Crossmodal Generative Creativity Support Tool. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–15.

[87] Li Yao, Xiaodong Peng, Yining Guo, Hongjie Ni, Yan Wan, and Cairong Yan. 2018. A data-driven approach for 3D human body pose reconstruction from a kinect sensor. In *Journal of Physics: Conference Series*, Vol. 1098. IOP Publishing, 012024.

[88] Zijie Ye, Haozhe Wu, Jia Jia, Yaohua Bu, Wei Chen, Fanbo Meng, and Yanfeng Wang. 2020. Choreonet: Towards music to dance synthesis with choreographic action unit. In *Proceedings of the 28th ACM International Conference on Multimedia*. 744–752.

[89] Eleni Zaggelidou, Evan gelos Tsamourtzis, Alexandros Malkogeorgos, and Georgios Zaggelidis. 2012. Dimensions of market demand associated with dance schools. *Sport Science Review* 21, 3-4 (2012), 101.

[90] Mingao Zhang, Changhong Liu, Yong Chen, Zhenchun Lei, and Mingwen Wang. 2022. Music-to-dance generation with multiple conformer. In *Proceedings of the 2022 International Conference on Multimedia Retrieval*. 34–38.

[91] Zheng Zhang, Ying Xu, Yanhao Wang, Bingsheng Yao, Daniel Ritchie, Tongshuang Wu, Mo Yu, Dakuo Wang, and Toby Jia-Jun Li. 2022. Storybuddy: A human-ai collaborative chatbot for parent-child interactive storytelling with flexible parental involvement. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–21.

[92] Qian Zhou, David Ledo, George Fitzmaurice, and Fraser Anderson. 2024. Time-Tunnel: Integrating Spatial and Temporal Motion Editing for Character Animation in Virtual Reality. In *Proceedings of the CHI Conference on Human Factors in*

Computing Systems. 1–17.

[93] Wentao Zhu, Xiaoxuan Ma, Dongwoo Ro, Hai Ci, Jinlu Zhang, Jiaxin Shi, Feng Gao, Qi Tian, and Yizhou Wang. 2023. Human motion generation: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023).

[94] Wenlin Zhuang, Congyi Wang, Jinxiang Chai, Yangang Wang, Ming Shao, and Siyu Xia. 2022. Music2dance: Dancenet for music-driven dance generation. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 18, 2 (2022), 1–21.

[95] Gabriella Zimányi and Anita Lanszki. 2020. The Influence of Social Media on Hip-Hop Dancers and their Classes. *Tánc és Nevelés* 1, 1 (2020), 97–112.

[96] Donald W Zimmerman and Bruno D Zumbo. 1993. Relative power of the Wilcoxon test, the Friedman test, and repeated-measures ANOVA on ranks. *The Journal of Experimental Education* 62, 1 (1993), 75–86.

A ALGORITHMS

Algorithm 1: DanceDTW Algorithm

Input: $X, Y \in \mathbb{R}^{F \times (J \times 7)}$: Motion sequences (e.g., X is input motion sequence and Y is reference motion sequence), where:

- F: Number of frames in the sequence (temporal dimension)
- J: Number of joints in the motion (spatial dimension, $J = 24$)
- 7: Positional (P_x, P_y, P_z) and rotational (R_w, R_x, R_y, R_z) data for each joint

Parameters: $\lambda_f, \lambda_g, \mu_p, \mu_r$

Output: C_{total}: Final similarity cost between X and Y (lower values indicate higher similarity)

Function DanceDTW($X, Y, \lambda_f, \lambda_g, \mu_p, \mu_r$)**:**

  Compute average motion data: $\overline{Z} = \frac{1}{2}(\overline{X} + \overline{Y})$

  Adjust offset: $f(X) \leftarrow X - \overline{Z}, f(Y) \leftarrow Y - \overline{Z}$

  Normalize data: $g(X) \leftarrow \frac{X - \min(X)}{\max(X) - \min(X)}$,

    $g(Y) \leftarrow \frac{Y - \min(Y)}{\max(Y) - \min(Y)}$

  for *each joint* $j = 1$ *to* J **do**

    Compute DTW costs for position:

    $c_{p_{f},j} = \lambda_f \cdot \text{DTW}_I(f(X_{p,j}), f(Y_{p,j}))$

    $c_{p_{g},j} = \lambda_g \cdot \text{DTW}_I(g(X_{p,j}), g(Y_{p,j}))$

    Compute DTW costs for rotation:

    $c_{r_{f},j} = \lambda_f \cdot \text{DTW}_I(f(X_{r,j}), f(Y_{r,j}))$

    $c_{r_{g},j} = \lambda_g \cdot \text{DTW}_I(g(X_{r,j}), g(Y_{r,j}))$

  Gather all position costs:

  $c_{p,\text{all}} = \{c_{p_{f},j}, c_{p_{g},j} \mid j = 1, \ldots, J, \ldots, J \times 3\}$

  for *each* $c_{p,j} \in c_{p,all}$ **do**

    Normalize position cost:

    $c_{p,\text{norm},j} = \frac{c_{p,j} - \min(c_{p,\text{all}})}{\max(c_{p,\text{all}}) - \min(c_{p,\text{all}})}$

  Compute total costs:

  $C_p = \sum_{j=1}^{2 \times J \times 3} c_{p,\text{norm},j}$

  $C_r = \sum_{j=1}^{J \times 4} (c_{r_{f},j} + c_{r_{g},j})$

  $C_{\text{total}} = \mu_p C_p + \mu_r C_r$

  return C_{total}

Algorithm 2: Parameter Fitting for DanceDTW using Bayesian Optimization

---

Input: TrainSet: 90 pairs of {reference, student} motion sequences

TestSet: 9 pairs of {reference, student} motion sequences

Output: Optimal parameters: $\lambda_f, \lambda_g, \mu_p, \mu_r$ and Test accuracy

Function `OptimizeParameters`(*TrainSet*, *TestSet*):

  Initialize Bayesian Optimization process

  for *each evaluation (50 iterations)* **do**

    Sample $\lambda_f, \lambda_g, \mu_p, \mu_r$ using Bayesian Optimization

    for *each pair (X_{ref}, X_{stu}) in TrainSet* **do**

      $C_{\text{total}} \leftarrow$ `DanceDTW` $(X_{\text{ref}}, X_{\text{stu}}, \lambda_f, \lambda_g, \mu_p, \mu_r)$

      Compute classification accuracy based on C_{total}

    Calculate average training accuracy

    Update Bayesian Optimization model

  Select best performing $\lambda_f, \lambda_g, \mu_p, \mu_r$

  for *each pair (X_{ref}, X_{stu}) in TestSet* **do**

    $C_{\text{total}} \leftarrow$ `DanceDTW` $(X_{\text{ref}}, X_{\text{stu}}, \lambda_f, \lambda_g, \mu_p, \mu_r)$

    Compute classification accuracy based on C_{total}

  Calculate final test accuracy

  return *Best $\lambda_f, \lambda_g, \mu_p, \mu_r$, Train accuracy, Test accuracy*

Run `OptimizeParameters`(*TrainSet*, *TestSet*) to find optimal parameters

return *Optimal $\lambda_f, \lambda_g, \mu_p, \mu_r$, Train accuracy, Test accuracy*

B   ITERATIVE CLOSEST POINT ALGORITHM FOR COMPARISON WITH SMPL AND KINECT MOTION DATA

The input consists of two sets of motion data, $\mathbf{P}_{\text{kinect}}$ and $\mathbf{P}_{\text{smpl}}$, each representing the 3D positions of common joints. First, the motion data is loaded from CSV files, where each joint's position and rotation are parsed into a structured dictionary format. The common joints between the two datasets are identified to ensure consistent alignment. For ICP, the first step involves calculating the centroids $\mathbf{c}_{\text{kinect}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{P}_{\text{kinect}}^i$ and $\mathbf{c}_{\text{smpl}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{P}_{\text{smpl}}^i$ of the corresponding joints. Next, the datasets are centered by subtracting the centroids from each point. The optimal rotation matrix $\mathbf{R}$ is computed using singular value decomposition (SVD) from the covariance matrix $\mathbf{H} = \sum_{i=1}^{n} (\mathbf{P}_{\text{kinect}}^i - \mathbf{c}_{\text{kinect}})(\mathbf{P}_{\text{smpl}}^i - \mathbf{c}_{\text{smpl}})^T$. Performing SVD on $\mathbf{H}$ gives $\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, and the optimal rotation matrix is then obtained as $\mathbf{R} = \mathbf{V}\mathbf{U}^T$. If the determinant of $\mathbf{R}$ is negative, the reflection is corrected by flipping the sign of the last column of $\mathbf{V}$. The translation vector is computed as $\mathbf{t} = \mathbf{c}_{\text{smpl}} - \mathbf{R}\mathbf{c}_{\text{kinect}}$. The output of the ICP process is the transformation consisting of the optimal rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$, which is applied to all frames of the Kinect dataset to align it with the SMPL dataset. Thus, the transformed Kinect data for each frame is computed as $\mathbf{P}_{\text{kinect}}^{\text{transformed}} = \mathbf{R}\mathbf{P}_{\text{kinect}} + \mathbf{t}$, ensuring that the Kinect and SMPL data are aligned for further analysis.

Algorithm 3: ICP Alignment for SMPL and Kinect Data

Input: $\mathbf{P}_{\text{kinect}}, \mathbf{P}_{\text{smpl}}$: Motion data from Kinect and SMPL systems

Each contains the 3D positions of common joints across frames.

Output: $\mathbf{R}, \mathbf{t}$: Optimal rotation matrix and translation vector for alignment

$\mathbf{P}_{\text{kinect}}^{\text{transformed}}$: Transformed Kinect data aligned with SMPL data

Function `ICPAlignment`($\mathbf{P}_{kinect}, \mathbf{P}_{smpl}$):

  Compute centroids of common joints:

  $\mathbf{c}_{\text{kinect}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{P}_{\text{kinect}}^i$,

  $\mathbf{c}_{\text{smpl}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{P}_{\text{smpl}}^i$

  Center the data by subtracting the centroids:

  $\mathbf{P}_{\text{kinect}}^{\text{centered}} = \mathbf{P}_{\text{kinect}} - \mathbf{c}_{\text{kinect}}$,

  $\mathbf{P}_{\text{smpl}}^{\text{centered}} = \mathbf{P}_{\text{smpl}} - \mathbf{c}_{\text{smpl}}$

  Compute covariance matrix $\mathbf{H}$:

  $\mathbf{H} = \sum_{i=1}^{n} \mathbf{P}_{\text{kinect}}^{\text{centered},i} \mathbf{P}_{\text{smpl}}^{\text{centered},i^T}$

  Perform SVD on $\mathbf{H}$:

  $\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^T$

  Compute optimal rotation matrix:

  $\mathbf{R} = \mathbf{V}\mathbf{U}^T$

  Check for reflection and adjust if necessary:

  if $det(\mathbf{R}) < 0$ **then**

    Flip sign of last column of $\mathbf{V}$: $\mathbf{V}_{-1} \leftarrow -\mathbf{V}_{-1}$

    Recompute $\mathbf{R} = \mathbf{V}\mathbf{U}^T$

  Compute translation vector:

  $\mathbf{t} = \mathbf{c}_{\text{smpl}} - \mathbf{R}\mathbf{c}_{\text{kinect}}$

  Apply transformation to all frames of Kinect data:

  $\mathbf{P}_{\text{kinect}}^{\text{transformed}} = \mathbf{R}\mathbf{P}_{\text{kinect}} + \mathbf{t}$

  return $\mathbf{R}, \mathbf{t}, \mathbf{P}_{kinect}^{transformed}$
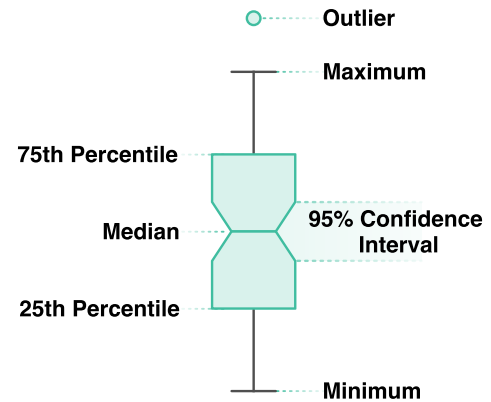

C   EXPLANATION OF NOTCHED BOXPLOT

Figure 12: Visual Explanation of Notched Boxplot. This figure was represented with reference to [58]